



# Deusto

Universidad de Deusto  
Deustuko Unibertsitatea

Facultad de Ingeniería

**TESIS DOCTORAL**

**NUEVO MÉTODO BASADO EN EL *HRV*  
PARA LA EVALUACIÓN DE *HMI*s Y  
SISTEMAS *ITS* PARA TRANSPORTE  
POR CARRETERA INTEGRANDO  
FACTORES PERSONALES,  
TEMPORALES Y AMBIENTALES**

**Programa: CIENCIA DE LA COMPUTACIÓN**

**Presentada por:**

**Jesús Murgoitio Larrauri**

**Dirigida por:**

**Dr. D. José Luis Gutiérrez Temiño**

Bilbao, octubre de 2015





Universidad de Deusto  
Unibertsitatea

Facultad de Ingeniería

## TESIS DOCTORAL

**NUEVO MÉTODO BASADO EN EL *HRV*  
PARA LA EVALUACIÓN DE *HMI*s Y  
SISTEMAS *ITS* PARA TRANSPORTE  
POR CARRETERA INTEGRANDO  
FACTORES PERSONALES,  
TEMPORALES Y AMBIENTALES**

**Programa: CIENCIA DE LA COMPUTACIÓN**

**Presentada por: Jesús Murgoitio Larrauri**

**Dirigida por: Dr. D. José Luis Gutiérrez Temiño**

Bilbao, octubre de 2015



A mi familia



## **AGRADECIMIENTOS**

Quiero expresar mi agradecimiento, en primer lugar, al Dr. D. José Luis Gutiérrez Temiño de ESIDE por su acertada dirección, paciencia y colaboración, y por haberme apoyado con su espíritu universitario. En segundo lugar, a la Dra. María José Gil Larrea, también de ESIDE, por la preocupación, ánimo y paciencia puestos en este trabajo, así como por sus acertadas ideas.

Hago extensible mi agradecimiento a los compañeros de la Fundación Tecnalía Research & Innovation, especialmente a los del grupo *ITS* (Sistemas Inteligentes de Transporte), así como al departamento de cardiología del hospital de Usánsolo, por su diligencia y disponibilidad.

Quiero agradecer también a todos y cada uno de los participantes que tomaron parte en los experimentos realizados para esta monografía, así como a los desarrolladores de algunas de las herramientas SW que se emplearon para la obtención de datos relativos al *HRV*, es decir, D. Miguel Antúnez Salgado y D. Roberto Montero Miguel.

Por último, quiero agradecer a mis amistades, entre otras muchísimas cosas, por su ayuda en la corrección de esta memoria y, especialmente a mi familia por su paciencia y comprensión.



## RESUMEN

Esta tesis doctoral profundiza en el análisis de la variabilidad de la frecuencia cardíaca o *HRV (Heart Rate Variability)* calculada a partir de la señal eléctrica cardíaca o *ECG (Electro Cardio Grama)* y de los cambios que se producen en sus propiedades relacionados con el nivel de atención de los sujetos que conducen vehículos terrestres. El objetivo es su utilización en aplicaciones donde es necesaria la estimación de aquel nivel y enfocarla principalmente a medir la influencia que tienen en la seguridad vial los nuevos *HMI (Human Machine Interface)* que se experimentan en los vehículos de última generación.

Más específicamente, esta tesis se centra en la identificación de los parámetros mejor correlacionados con la variabilidad debida a características personales (edad), temporales (ciclos circadianos), y del entorno (tipo de carretera) para su posterior integración en un sistema completo. Este carácter integrador es muy escaso en la literatura, casi inexistente, y persigue conseguir criterios de clasificación lo más universales posibles, condición indispensable para su utilización en aplicaciones como la conducción de vehículos, la cual demanda un estado de alerta mantenido. Como consecuencia, se ha desarrollado una aplicación que permite generar patrones según un sistema de cuatro variables normalizadas inferidas a partir del *HRV*, lo cual permite evaluar y comparar diferentes maneras de interactuar con vehículos de transporte por carretera.

Por último, resaltar que todo el análisis contenido en esta tesis es potencialmente extensible a otros ámbitos donde la medida relativa (versus patrones) del consumo de la atención, entendida según el modelo de Kahneman, sea útil o necesaria (e.g. entornos o actividades de riesgo).

## **ABSTRACT**

This doctoral thesis makes an in-depth analysis of HRV (Heart Rate Variability) as calculated from an ECG (Electro Cardio Gram) and how it varies related to the attention span shown by individuals when driving land vehicles. The aim is to use this knowledge in applications where this attention span needs to be estimated, and to focus primarily on measuring the influence that the new HMIs (Human Machine Interfaces) being tried out in the latest generation of vehicles may have on road safety.

More specifically, the thesis concentrates on the identification of those parameters which best relate to variability due to personal characteristics such as age, temporary features such as circadian cycles, or the surrounding environment as in road category, in order to integrate all of these features into a complete system. It is this integrating feature which is most lacking in the literature, indeed it is almost non-existent, so that the present work seeks to achieve the most universal classification criteria. This is essential for its use in applications such as vehicle driving, which require a constant state of alert. To this end, an application has been developed which allows patterns to be generated according to a system of four standard variables derived from the HRV, allowing different ways of interacting with on-road vehicles to be compared and assessed.

Finally, a further point of note is that the analysis contained in this thesis may be extended to other fields (e.g. risky activities or environments) where it is useful or indeed necessary to take a relative measurement of attention span rather than follow a simple patterns, as understood by the Kahneman model.

# ÍNDICE

	pág.
<b>ÍNDICE</b> .....	<b>i</b>
<b>ACRÓNIMOS</b> .....	<b>v</b>
<b>LISTA DE FIGURAS</b> .....	<b>vii</b>
<b>LISTA DE TABLAS</b> .....	<b>xi</b>
<b>1. INTRODUCCIÓN</b> .....	<b>1</b>
1.1. Antecedentes.....	1
1.2. Justificación del estudio .....	13
1.3. Hipótesis y objetivos.....	20
<b>2. ESTADO DEL ARTE</b> .....	<b>23</b>
2.1. Naturaleza de la atención.....	23
2.1.1. Definiciones.....	23
2.1.2. Modelos teóricos .....	25
2.1.3. Características .....	27
2.1.4. Factores determinantes .....	28
2.1.5. Técnicas de medida.....	32
2.2. Nivel de alerta o “ <i>Arousal</i> ” .....	38
2.2.1. Introducción .....	39
2.2.2. Conceptos asociados a bajos niveles de alerta .....	40
2.2.3. Factores que inciden sobre el estado de alerta .....	42
2.2.4. Métodos para la evaluación del nivel de alerta .....	44
2.3. Fundamentos del HRV .....	54
2.3.1. Naturaleza del ECG .....	54
2.3.2. Causas de la variabilidad del ritmo cardíaco .....	56
2.3.3. Métodos para la obtención de la serie RR .....	59
2.3.4. Aplicaciones del análisis HRV .....	61
2.4. Métodos del análisis HRV .....	63
2.4.1. Métodos estadísticos .....	64

2.4.2.	Métodos espectrales .....	78
2.4.3.	Métodos espectro-temporales .....	87
<b>3.</b>	<b>METODOLOGÍA .....</b>	<b>99</b>
3.1.	Etapas de la metodología.....	99
3.1.1.	Adquisición.....	100
3.1.2.	Pre-procesado.....	101
3.1.3.	Edición (Recuperación manual).....	101
3.1.4.	Post-Procesado 1: Herramienta de análisis HRV .....	102
3.1.5.	Post-Procesado 2: Compilación .....	104
3.1.6.	Conversión.....	104
3.1.7.	Minería de datos .....	105
3.2.	Descripción del recorrido .....	110
3.3.	Sujetos participantes .....	111
3.4.	Diseño: Variables, Modelo de datos y Casos de estudio .....	113
3.4.1.	Variables y Modelo de datos .....	113
3.4.2.	Casos de estudio .....	117
<b>4.</b>	<b>ANÁLISIS Y RESULTADOS .....</b>	<b>121</b>
4.1.	Caso de estudio 1: Factores personales (Edad) .....	126
4.1.1.	Análisis en el dominio del tiempo (DT).....	127
4.1.2.	Análisis en el dominio de la frecuencia (DF).....	131
4.1.3.	Análisis combinado (DC) .....	134
4.1.4.	Conclusiones.....	137
4.2.	Caso de estudio 2: Factores temporales (Ciclos Circadianos).....	138
4.2.1.	Análisis en el dominio del tiempo.....	139
4.2.2.	Análisis en el dominio de la frecuencia .....	143
4.2.3.	Análisis en el dominio combinado .....	146
4.2.4.	Conclusiones.....	149
4.3.	Caso de estudio 3: Factores ambientales (Tipo de carretera) .....	151
4.3.1.	Análisis en el dominio del tiempo.....	153
4.3.2.	Análisis en el dominio de la frecuencia .....	158
4.3.3.	Análisis en el dominio combinado .....	161
4.3.4.	Conclusiones.....	164
4.4.	Caso de estudio 4: “Fatiga” (Tipo de tramo) .....	166
4.4.1.	Análisis en el dominio del tiempo.....	168

4.4.2.	Análisis en el dominio de la frecuencia .....	173
4.4.3.	Análisis en el dominio combinado .....	178
4.4.4.	Conclusiones.....	183
<b>5.</b>	<b>CONCLUSIONES FINALES .....</b>	<b>187</b>
5.1.	Conclusiones.....	187
5.2.	Líneas futuras.....	190
<b>6.</b>	<b>APLICACIÓN .....</b>	<b>193</b>
6.1.	Fase 1: “Patrón” .....	193
6.2.	Fase 2: “Proyecto” .....	195
6.3.	Fase 3: “Evaluación” .....	198
<b>7.</b>	<b>BIBLIOGRAFÍA.....</b>	<b>201</b>
<b>8.</b>	<b>ANEXOS .....</b>	<b>225</b>
8.1.	Anexo 1: Participantes y Autorizaciones .....	225
8.1.1.	Participantes.....	225
8.1.2.	Autorizaciones .....	227
8.2.	Anexo 2: Cuadro resumen de resultados.....	267
8.3.	Anexo 3: Equipamiento – Especificaciones técnicas.....	269
8.3.1.	Netbook: DELL – Inspiron 1110: Especificaciones.....	269
8.3.2.	Sistema de adquisición: NI – DAQ 6009 .....	272
8.3.3.	Sensor ECG: EKG-Flex/Pro: SA9306M.....	278
8.3.4.	Otros equipos utilizados .....	281
8.3.4.1.	Fuente de alimentación para sensor ECG.....	281
8.3.4.2.	Electrodos.....	282
8.4.	Anexo 4: Manual de usuario de la aplicación. ....	283



## ACRÓNIMOS

<b>Acrónimo</b>	<b>Descripción</b>
ABS	<i>Antilock Brake System</i>
AED	Actividad Electro-Dérmica
AR	Auto-Regresivo
ARFF	<i>Attribute-Relation File Format</i>
ASDA	<i>American Sleep Disorders Association</i>
AuNS	<i>Autonomous Nervous System</i>
CNS	<i>Central Nervous System</i>
CSA	<i>Compressed Spectral Array</i> o Matriz Espectral Comprimida
CWT	<i>Continuous Wavelet Transform</i>
DC	Dominio Combinado
DF	Dominio de la Frecuencia
DGT	Dirección General de Tráfico
DRR	Diferencias de la serie RR
DT	Dominio del Tiempo
DWT	<i>Discrete Wavelet Transform</i>
ECG	Electro-Cardio-Grama
EEG	Electro-Encéfalo-Grama
EMG	Electro-Mio-Grama
EOG	Electro-Oculo-Grama
ESIDE	Escuela Superior de Ingeniería DEusto
FFT	<i>Fast Fourier Transform</i>
FP	<i>False Positives</i>
FWT	<i>Fast Wavelet Transform</i>
HF	<i>High Frequency</i>
HMI	<i>Human Machine Interface</i>
HRV	<i>Heart Rate Variability</i>
IA	Inteligencia Artificial
INTRAS	Instituto Universitario de Tráfico y Seguridad Vial
IPFM	<i>Integral Pulse Frequency Modulation</i>
ISTAG	<i>Information Society Technologies Advisory Group</i>
ITS	<i>Intelligent Transport Systems</i>
LF	<i>Low Frequency</i>
LFC	<i>Low Frequency-Component</i>
LFP	<i>Low Frequency-Power</i>
LIN	<i>Local Interconnect Network</i>
MSPWVD	<i>Modified and Smoothed Pseudo-Wigner-Ville Distribution</i>
MSTFT	<i>Modified Short-Time Fourier Transform</i>
NHTSA	<i>National Highway Traffic Safety Administration</i>
NREM	<i>Non Rapid Eye Movement</i>
PaNS	<i>Parasympathetic Nervous System</i>
PET	<i>Positron Emission Tomography</i>
PIB	Producto Interior Bruto
PNB	Producto Nacional Bruto

<i>PNS</i>	<i>Peripheral Nervous System</i>
<i>PSD</i>	<i>Power Spectral Density</i>
<i>RACC</i>	Real Automóvil Club de Cataluña
<i>RACE</i>	Real Automóvil Club de España
<i>RED</i>	Respuesta ElectroDérmica
<i>REM</i>	<i>Rapid Eye Movement</i>
<i>ROC</i>	<i>Receiver Operating Characteristic</i> o Curva ROC
<i>RSA</i>	<i>Respiratory Sinus Arrhythmia</i>
<i>SLAM</i>	<i>SeLective Attention Model</i>
<i>SoNS</i>	<i>Somatic Nervous System</i>
<i>SPWVD</i>	<i>Smoothed Pseudo-Wigner-Ville Distribution</i>
<i>SyNS</i>	<i>Sympathetic Nervous System</i>
<i>SW</i>	<i>SoftWare</i>
<i>TP</i>	<i>True Positives</i>
<i>UE</i>	Unión Europea
<i>VLf</i>	<i>Very Low Frequency</i>
<i>VNC</i>	Variación Negativa Contingente

## LISTA DE FIGURAS

<b>Figura</b>	<b>Pág.</b>
<i>Figura 1: Posicionamiento tecnológico del proyecto “Caring Cars”</i> .....	4
<i>Figura 2: Diseño con sensores dedicados</i> .....	6
<i>Figura 3: Diseño con red de sensores</i> .....	7
<i>Figura 4: Tasa de variación de víctimas mortales en carretera en España (2001-2014)</i> .....	13
<i>Figura 5: Víctimas mortales en carretera en España (1960-2014).</i> .....	13
<i>Figura 6: Equipo EEG</i> .....	17
<i>Figura 7: Equipo ECG</i> .....	18
<i>Figura 8: Fisiología del corazón humano”</i> .....	55
<i>Figura 9: Complejo “PQRST”</i> .....	56
<i>Figura 10: Estructura del sistema nervioso</i> .....	57
<i>Figura 11: Artefactos en la serie temporal RR (ejemplos)</i> .....	61
<i>Figura 12: Anchura de la base del histograma (registro de 2000 latidos)</i> .....	68
<i>Figura 13: Detalle de secuencia RR en registro de corta duración</i> .....	73
<i>Figura 14: Plot de Poincaré: Significado gráfico de SD1 y SD2</i> .....	76
<i>Figura 15: Esquema de procesamiento de la señal RR para el análisis espectral</i> .....	87
<i>Figura 16: Representación de secuencia RR (Izqda.: CSA, Dcha.: mapa de nivel)</i> .....	89
<i>Figura 17: Etapas de la metodología (Adquisición – Análisis)</i> .....	99
<i>Figura 18: Sistema de adquisición utilizado</i> .....	100
<i>Figura 19: HMI de aplicación de adquisición</i> .....	101
<i>Figura 20: HMI de aplicación “Editor”</i> .....	102
<i>Figura 21: HMI de aplicación de análisis previo</i> .....	103
<i>Figura 22: División de la señal en intervalos de 5 minutos (ventanas)</i> .....	103
<i>Figura 23: Ejemplo de formato “arff”</i> .....	105
<i>Figura 24: Ayuda WEKA sobre “CfsSubsetEval”</i> .....	106
<i>Figura 25: Ayuda WEKA sobre “ExhaustiveSearch”</i> .....	107
<i>Figura 26: Ayuda WEKA sobre “RandomTree”</i> .....	108
<i>Figura 27: Ayuda WEKA sobre “KStar”</i> .....	109
<i>Figura 28: Descripción del recorrido</i> .....	110
<i>Figura 29: Modelo de datos</i> .....	113
<i>Figura 30: Gráfico del PSD de una ventana de 5 minutos</i> .....	116
<i>Figura 31: Ejemplo de resultado en la selección de atributos y ranking de los mismos</i> .....	122
<i>Figura 32: Ejemplo de resultado del clasificador 1 “RandomTree”</i> .....	123
<i>Figura 33: Ejemplo de resultado del clasificador 2 “Lazy KStar”</i> .....	124

<i>Figura 34: Ranking de atributos en el DT (Edad)</i> .....	128
<i>Figura 35: Resultado del clasificador “RandomTree” en el DT (Edad)</i> .....	129
<i>Figura 36: Resultado del clasificador “Lazy KStar” en el DT (Edad)</i> .....	130
<i>Figura 37: Ranking de atributos en el DF (Edad)</i> .....	131
<i>Figura 38: Resultado del clasificador “RandomTree” en el DF (Edad)</i> .....	132
<i>Figura 39: Resultado del clasificador “Lazy KStar” en el DF (Edad)</i> .....	133
<i>Figura 40: Resultado del clasificador “RandomTree” en el DC (Edad)</i> .....	135
<i>Figura 41: Resultado del clasificador “Lazy KStar” en el DC (Edad)</i> .....	136
<i>Figura 42: Ranking de atributos en el DT (Ciclos Circadianos)</i> .....	140
<i>Figura 43: Clasificador “RandomTree” en el DT (Ciclos Circadianos)</i> .....	141
<i>Figura 44: Clasificador “Lazy KStar” en el DT (Ciclos Circadianos)</i> .....	142
<i>Figura 45: Ranking de atributos en el DF (Ciclos Circadianos)</i> .....	144
<i>Figura 46: Clasificador “RandomTree” en el DF (Ciclos Circadianos)</i> .....	144
<i>Figura 47: Clasificador “Lazy KStar” en el DF (Ciclos Circadianos)</i> .....	145
<i>Figura 48: Clasificador “RandomTree” en el DC (Ciclos Circadianos)</i> .....	147
<i>Figura 49: Resultado del clasificador “Lazy KStar” en el DC (Ciclos Circadianos)</i> .....	148
<i>Figura 50: Ranking de atributos en el DT (Tipo de Carretera)</i> .....	154
<i>Figura 51: Clasificador “RandomTree” en el DT (Tipo de Carretera)</i> .....	155
<i>Figura 52: Histograma de la media normalizada (Tipo de Carretera)</i> .....	156
<i>Figura 53: Clasificador “Lazy KStar” en el DT (Tipo de Carretera)</i> .....	157
<i>Figura 54: Clasificador “Lazy KStar” en el DT (Tipo de Carretera – 3 variables)</i> .....	157
<i>Figura 55: Ranking de atributos en el DF (Tipo de Carretera)</i> .....	159
<i>Figura 56: Clasificador “RandomTree” en el DF (Tipo de Carretera)</i> .....	159
<i>Figura 57: Clasificador “Lazy KStar” en el DF (Tipo de Carretera)</i> .....	160
<i>Figura 58: Clasificador “RandomTree” en el DC (Tipo de Carretera)</i> .....	162
<i>Figura 59: Clasificador “Lazy KStar” en el DC (Tipo de Carretera)</i> .....	163
<i>Figura 60: Clasificador “Lazy KStar” en el DC (Tipo de Carretera – 3 variables)</i> .....	164
<i>Figura 61: Clasificador “RandomTree” en el DT (Tipo de Tramo)</i> .....	169
<i>Figura 62: Tendencia de errores para “RandomTree” en el DT (Tipo de Tramo)</i> .....	170
<i>Figura 63: Histograma de la media normalizada (Tipo de Tramo)</i> .....	170
<i>Figura 64: Clasificador “Lazy KStar” en el DT (Tipo de Tramo)</i> .....	171
<i>Figura 65: Tendencia de errores para “Lazy KStar” en el DT (Tipo de Tramo)</i> .....	172
<i>Figura 66: Clasificador “Lazy KStar” en el DT (Tipo de Tramo – 4 variables)</i> .....	172
<i>Figura 67: Tendencia de errores para “Lazy KStar” en el DT (Tipo de Tramo – 4 variables)</i> .....	173
<i>Figura 68: Clasificador “RandomTree” en el DF (Tipo de Tramo)</i> .....	174
<i>Figura 69: Tendencia de errores para “RandomTree” en el DF (Tipo de Tramo)</i> .....	175

<i>Figura 70: Clasificador “Lazy KStar” en el DF (Tipo de Tramo) .....</i>	<i>176</i>
<i>Figura 71: Tendencia de errores para “Lazy KStar” en el DF (Tipo de Tramo) .....</i>	<i>176</i>
<i>Figura 72: Clasificador “Lazy KStar” en el DF (Tipo de Tramo – 4 variables) .....</i>	<i>177</i>
<i>Figura 73: Tendencia de errores para “Lazy KStar” en el DF (Tipo de Tramo – 4 variables).....</i>	<i>178</i>
<i>Figura 74: Resultado del clasificador “RandomTree” en el DC (Tipo de Tramo) .....</i>	<i>179</i>
<i>Figura 75: Tendencia de errores para “RandomTree” en el DF (Tipo de Tramo).....</i>	<i>180</i>
<i>Figura 76: Clasificador “Lazy KStar” en el DC (Tipo de Tramo).....</i>	<i>181</i>
<i>Figura 77: Tendencia de errores para “Lazy KStar” en el DC (Tipo de Tramo) .....</i>	<i>181</i>
<i>Figura 78: Clasificador “Lazy KStar” en el DC (Tipo de Tramo – 4 variables).....</i>	<i>182</i>
<i>Figura 79: Tendencia de errores para “Lazy KStar” en el DC (Tipo de Tramo – 4 variables).....</i>	<i>183</i>
<i>Figura 80: Obtención del “Patrón” .....</i>	<i>194</i>
<i>Figura 81: Gráfico del “Patrón”.....</i>	<i>195</i>
<i>Figura 82: Selección del “Proyecto” .....</i>	<i>196</i>
<i>Figura 83: Gráfico del “Proyecto” .....</i>	<i>197</i>
<i>Figura 84: Menú “Evaluación” (Patrón=“OK”, Proyecto=“OK”).....</i>	<i>199</i>
<i>Figura 85: Menú “Evaluación” (Patrón=“OK”, Proyecto= “OK” &amp; Cambio NO significativo) ...</i>	<i>200</i>
<i>Figura 86: Recorrido del experimento.....</i>	<i>228</i>



## LISTA DE TABLAS

<b>Tabla</b>	<b>Pág.</b>
<i>Tabla 1: Modelos teóricos sobre la atención</i> .....	27
<i>Tabla 2: Definiciones sobre el concepto “arousal”</i> .....	39
<i>Tabla 3: Correlación entre diversos índices empleados para caracterizar la HRV en medidas de 24 horas [Kleiger et al., 1993]</i> .....	70
<i>Tabla 4: Resumen de índices estadísticos convencionales</i> .....	71
<i>Tabla 5: Recorrido y Tipos de carretera</i> .....	111
<i>Tabla 6: Listado resumen de participantes</i> .....	112
<i>Tabla 7: Listado de variables personales</i> .....	114
<i>Tabla 8: Listado de variables propias del experimento</i> .....	115
<i>Tabla 9: Listado de variables propias de cada ventana temporal de cinco minutos</i> .....	115
<i>Tabla 10: Conjunto de variables en el dominio del tiempo para cada cinco minutos</i> .....	116
<i>Tabla 11: Conjunto de variables en el dominio de la frecuencia para cada cinco minutos</i> .....	117
<i>Tabla 12: Tipos de carretera</i> .....	119
<i>Tabla 13: Información resultante en la selección de atributos</i> .....	122
<i>Tabla 14: Información resultante con clasificador 1 “RandomTree”</i> .....	123
<i>Tabla 15: Información resultante con clasificador 2 “Lazy KStar”</i> .....	124
<i>Tabla 16: Participantes seleccionados para el estudio de la edad</i> .....	126
<i>Tabla 17: Tabla resumen de los rendimientos de los clasificadores (Edad)</i> .....	137
<i>Tabla 18: Participantes seleccionados para el estudio de los ciclos circadianos</i> .....	138
<i>Tabla 19: Tabla resumen de los rendimientos de los clasificadores (Ciclos Circadianos)</i> .....	150
<i>Tabla 20: Participantes seleccionados para el estudio del tipo de carretera</i> .....	152
<i>Tabla 21: Tabla resumen de los rendimientos de los clasificadores (Tipo de Carretera)</i> .....	165
<i>Tabla 22: Participantes seleccionados para el estudio del tipo de tramo</i> .....	167
<i>Tabla 23: Resumen de variables significativas en el DT (casos de estudio 1 a 3)</i> .....	168
<i>Tabla 24: Resumen de variables significativas en el DF para casos de estudio 1 a 3</i> .....	174
<i>Tabla 25: Resumen de variables significativas en el DT y DF para casos de estudio 1 a 3</i> .....	178
<i>Tabla 26: Tabla resumen de los rendimientos de los clasificadores (Tipo de Tramo)</i> .....	185
<i>Tabla 27: Identificadores de los participantes</i> .....	226
<i>Tabla 28: Tabla resumen de los rendimientos de los clasificadores (Casos de estudio 1 y 2)</i> .....	267
<i>Tabla 29: Tabla resumen de los rendimientos de los clasificadores (Casos de estudio 3 y 4)</i> .....	268



# 1. INTRODUCCIÓN

## 1.1. Antecedentes

La presente tesis doctoral se enmarca dentro de las actividades que en los últimos tiempos se vienen realizando en relación con la monitorización del estado del conductor de vehículos motorizados, principalmente conductores de automóviles o camiones para tramos largos, y los métodos objetivos para medir el nivel de atención, alerta y fatiga en base a la importante influencia que el **factor humano** tiene sobre el número de accidentes de tráfico.

Este trabajo forma parte de algunos proyectos con financiación local y/o nacional donde se profundiza en el concepto de “inteligencia ambiental” y en la necesidad de medir objetivamente la influencia que pueden tener en la seguridad vial los nuevos *HMI (Human Machine Interface)* que se están experimentando en los vehículos de última generación integrables en los nuevos ITS (*Intelligent Transport Systems*), e.g. joysticks *versus* volantes, etc. En definitiva, la tesis ahonda en la necesidad de desarrollar herramientas y/o aplicaciones que permitan valorar de forma objetiva unas formas de interactuar con los vehículos frente a otras (nuevas *versus* existentes).

### **Inteligencia Ambiental**

Una visión de futuro podría incluir entornos inteligentes equipados de tal forma que tanto las señales derivadas del comportamiento humano como los parámetros mecánicos pudieran ser monitorizados con el propósito de evaluar dicho entorno y detectar con anticipación determinadas situaciones. Desde este punto de vista se proyecta una convergencia hacia lo que vienen llamándose los ambientes inteligentes o “Inteligencia Ambiental”, según el término elegido por el *ISTAG (Information Society Technologies Advisory Group)* de la Comisión Europea para hacer referencia a la capacidad del entorno para facilitar la interacción humana y alcanzar sus objetivos y necesidades.

Este paradigma de la sociedad de la información nos promete que pronto estaremos rodeados de dispositivos inteligentes e intuitivos que responderán a nuestras necesidades, pero de manera que no nos demanden hacer complejas o tediosas operaciones, es decir, que su presencia no se perciba, que sean completamente transparentes.

Estos ambientes inteligentes, estén o no lejos, sean o no como los imaginamos, sí nos ofrecen un marco estable del que tomar conceptos y en el cual intercambiar tecnologías, beneficiándonos de su globalidad intrínseca y proporcionando un caldo enriquecido para la creatividad. Este marco se construye sobre tres pilares tecnológicos:

- Computación omnipresente e invisible, pervasiva, materializada en redes inteligentes de dispositivos de captura de datos y proceso de la información relativa a los objetos que nos rodean,
- Comunicaciones ubicuas que soportan el diálogo permanente entre aquellos dispositivos, por medio de redes inalámbricas,
- E interfaces de usuario inteligentes que permiten a los sujetos interactuar con su entorno de manera personalizada y transparente.

La “*inteligencia ambiental*” es un nuevo concepto de interacción inteligente con la tecnología que permitirá mejorar la calidad de vida de los usuarios y que en el futuro modificará el entorno doméstico, industrial, médico, laboral e incluso el ocio personal. Esta nueva idea abrirá una nueva perspectiva a nuestro entorno físico más próximo; un conjunto de dispositivos electrónicos mimetizados en el ambiente crearán una atmósfera que responderá y se adaptará a nuestros deseos y necesidades. Específicamente, en relación a las actividades de investigación tanto básica como aplicada ligadas a esta tesis, se aplicarán al entorno de la conducción y sus problemáticas asociadas, principalmente la evaluación del estado del conductor (factor humano).

### **Inteligencia Ambiental aplicada al automóvil**

El automóvil es uno de los espacios donde la *computación ambiental* o ubicua está avanzando a pasos agigantados. Un entorno inteligente revestirá al vehículo para aumentar la seguridad y comodidad del conductor y los acompañantes, así aportará información para modelos del contexto cada vez mejores, e.g. [Urquiza et al., 2011].

La mayoría de los dispositivos necesarios para crear estos vehículos inteligentes ya existen en el mercado, bastantes son de serie en automóviles de gama alta, y tan solo falta un pequeño impulso para que el funcionamiento de cada dispositivo se coordine con el resto. Resulta en este caso más fácil conseguir su implantación que en el caso de la domótica,

pues en el ensamblaje de las diferentes partes de un vehículo se parte de cero y se depende sólo de la estrategia del fabricante, por lo que resulta más sencillo conseguir que los diferentes componentes se comuniquen armónicamente entre sí, aunque los fabriquen diferentes proveedores.

El automóvil podrá comunicarse con otro área de computación ambiental como puede ser el edificio de oficinas donde trabaja el conductor, para activar tareas como arrancar su ordenador personal, y cargar de aplicaciones a la espera que el usuario las utilice a su llegada; o el propio domicilio, en donde el propio vehículo podrá iniciar tareas domésticas como encender la calefacción, cualquier electrodoméstico, informar de la programación de TV y una larga lista de actividades a la medida del sujeto.

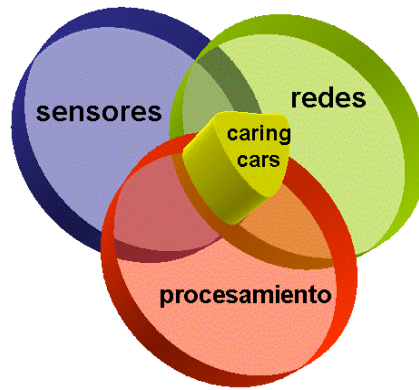
La inteligencia ambiental aplicada a la seguridad vial es otro aspecto de gran impacto potencial: El sistema inteligente puede controlar el grado de alcoholismo o cansancio del conductor [Vicente, 2011], y en caso extremo, impedir el uso del vehículo a su propio dueño. De igual forma puede mantener la velocidad del vehículo dentro de los límites legales en cada tramo de la carretera o controlar las luces de acuerdo con la luminosidad y la proximidad de otros vehículos, así como mantener la distancia prudencial con el vehículo precedente en función de la velocidad y condiciones ambientales.

El proyecto “*Caring Cars*”, en el que el autor de este trabajo tomó parte, se enmarca dentro del concepto de inteligencia ambiental y su aplicación al desarrollo de electrónica para el sector del transporte por carretera.

### **Proyecto “Caring Cars”**

El objetivo principal del proyecto “*Caring Cars*” consistía en mejorar el bienestar del conductor mediante una plataforma de sensores, procesamiento y comunicaciones, que evaluara la aptitud física y estado del conductor y le asistiera en la conducción, garantizando en todo momento la seguridad y el confort al volante.

El proyecto perseguía este objetivo buscando la sinergia entre tres áreas tecnológicas: sensores, plataforma de procesamiento y comunicaciones. La Figura 1 ilustra el posicionamiento tecnológico de este proyecto:



**Figura 1: Posicionamiento tecnológico del proyecto “Caring Cars”**

En concreto, “Caring Cars” desarrolló una nueva red de sensores y un *gateway* embarcado para su uso en automóviles, centrándose en el conductor del vehículo. En este sentido, “Caring Cars” se diferencia de la mayoría de los proyectos sobre vehículos inteligentes, que suelen desarrollar sólo sensores para monitorizar el entorno del vehículo.

El proyecto aplica tecnología de sensores existentes de una forma más innovadora [Reina et al., 2012; Peltokangas et al., 2012; Murgoitio et al., 2007, 2008, 2009, 2010] para conseguir una reducción de costes y una mayor riqueza y flexibilidad del sistema de información en el automóvil, así como que todos los coches, no sólo los de lujo, se puedan dotar de sistemas avanzados de asistencia a la conducción y reducción de accidentes.

La propuesta de “Caring Cars” es fruto de la colaboración con un consorcio internacional integrado por participantes de varios países europeos (España, Países Bajos y Turquía), y fue entregado al programa EUREKA en Marzo 2006 para la obtención de la etiqueta MEDEA+. El Comité de MEDEA+ dio una evaluación favorable a la propuesta preliminar (*Project Outline*) de “Caring Cars”, y la propuesta completa (*Full Project Proposal*) obtuvo la etiqueta MEDEA+ (2A403–CARING CARS) el 2 de junio del 2006.

La salud del conductor y la seguridad del tránsito tienen una relación estrecha. Aunque la obtención del carné de conducir en España está sujeta a un examen médico, esto no garantiza la aptitud del conductor en todos los momentos. La mayoría de los accidentes se producen por error humano y, en particular, por la incapacidad física del conductor (embriaguez, sueño, distracción, actitud) para reaccionar adecuadamente a la situación.

Los sensores y los dispositivos de comunicación en el coche jugarán un papel decisivo en la reducción de accidentes y en el apoyo a los servicios de emergencia en caso de accidente. El proyecto “*Caring Cars*” trata de desarrollar una plataforma de sensores y comunicación en el automóvil, para asistir al conductor, mejorar su bienestar, y ayudarle en caso de accidente [Reina et al., 2012; Murgoitio et al., 2007, 2008, 2009, 2010].

Los proyectos internacionales actuales que trabajan sobre sensores en el automóvil muestran las siguientes carencias:

- La mayoría de los proyectos se centran en sensores para el vehículo, y no para su conductor. Sin embargo, la mayoría de los accidentes se deben al error humano. Además, en caso de accidente los datos sobre el estado físico de los ocupantes del vehículo pueden ser de importancia vital para los servicios de asistencia.
- Los proyectos que desarrollan sensores para monitorizar el conductor se centran sólo en su estado físico, y no en su estado mental. Sin embargo, el estado mental (distracción, aburrimiento, agitación) es causa de muchos accidentes.
- Los proyectos existentes suelen plantear una integración rígida y vertical de sensores en el automóvil, lo cual limita la extensibilidad y escalabilidad.

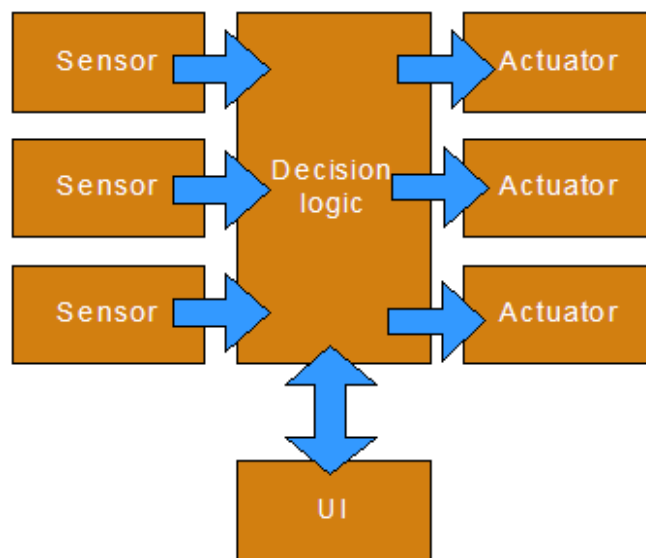
El proyecto “*Caring Cars*” tiene un enfoque claro dirigido hacia el bienestar del conductor y, además propone una plataforma de integración dinámica de sensores que permite comunicaciones con sistemas externos de información [Reina et al., 2012; Murgoitio et al., 2007, 2008, 2009, 2010, 2011].

### **Electrónica para el automóvil**

El número de sistemas electrónicos y sensores en los automóviles se ha incrementado sensiblemente en las últimas décadas. Hoy en día el coste de la electrónica en los vehículos de gama alta puede llegar a una cuarta parte del coste total de fabricación del vehículo. Toda esta electrónica se conecta por cables. En vehículos de gama alta la cantidad de cable puede llegar a ser de unos 4 kilómetros (frente a los 45 metros para automóviles fabricados en 1955).

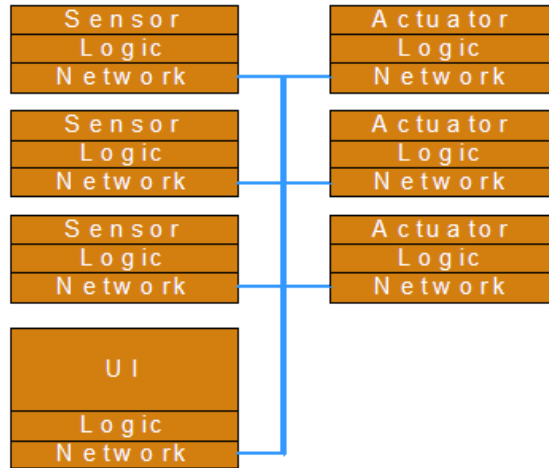
En la actualidad, los automóviles están equipados con una gran cantidad de sensores, entre 50 y 100. Son utilizados en todo tipo de mediciones, desde procesos del motor hasta la temperatura externa. En los vehículos de gama alta esta información se utiliza para transmitir información al fabricante de forma que se puedan tomar medidas antes de que se produzca un problema. Sin embargo, estas aplicaciones son sólo válidas para una cierta aplicación de un fabricante, y los sensores sólo son utilizados para una única funcionalidad. Las redes de sensores propuestas en “*Caring Cars*” buscan disminuir la cantidad de cables en los vehículos modernos así como aumentar la flexibilidad de las funcionalidades para las que se utilizan los sensores.

En la Figura 2 se puede ver la disposición actual de los sensores en el automóvil. Se conectan directamente a una unidad de proceso, y los actuadores son controlados de la misma forma. Aunque se trata de un diseño muy sencillo no permite flexibilidad y requiere cables dedicados por cada sensor.



**Figura 2: Diseño con sensores dedicados**

En cambio, en la Figura 3 se muestra el diseño propuesto para este proyecto, donde los sensores se conectan a una lógica y a una interfaz de red:



**Figura 3: Diseño con red de sensores**

De esta manera, la conexión con la lógica dedicada, que puede estar o no presente, se realiza a través de una red, de forma que se puede compartir el cableado, pero aún más importante, la información de un sensor estará disponible también para otras aplicaciones.

Los principales competidores en esta área de aplicación son los fabricantes Visteon, Delphi y Freescale:

- Visteon anunció TACNET, un sistema de comunicaciones para servicios de emergencia. Este sistema aporta una solución de control integrada para las fuerzas de seguridad pública. Sin embargo, en lugar del sistema distribuido propuesto en el presente proyecto, Visteon ha diseñado un sistema que tiene una sola unidad de control central que puede controlar los sistemas del vehículo.
- Delphi trabaja en un sistema de seguridad integrado en el vehículo que se centra en tecnologías que necesitan un nuevo tipo de interfaz de usuario. Esta solución incluye un sistema de control de la carga de trabajo, y usa varios sensores biométricos para monitorizar el nivel de alerta del conductor, como el seguimiento ocular, cadencia de respiración y pulso que mide las variaciones de presión sanguínea mediante los pequeños cambios de potencial eléctrico en las manos del conductor cuando reposan en el volante. Este sistema tiene algunos aspectos en común con *Caring Cars*.
- Freescale Semiconductor ha mostrado su vehículo *Local Interconnect Network (LIN)*, que incluye una red de instrumentos y una aplicación de control. *LIN* es una

arquitectura de red especialmente desarrollada para redes de sensores y actuadores para la automoción. Estrictamente hablando no se trata de competencia directa con el proyecto, ya que en el proyecto “*Caring Cars*” se utilizan tecnologías de redes pero yendo mucho más allá.

### **Otros proyectos relacionados**

Otros proyectos de investigación en automoción desarrollan sensores capaces de detectar si el conductor se está quedando dormido, y se enmarcan dentro de la iniciativa i2010 de la Unión Europea cuyo proyecto estrella es el “Coche Inteligente”. La Comisión Europea está patrocinando varias líneas en el área de la seguridad electrónica, algunas de las cuales están relacionadas con esta propuesta:

- ADASE (*Advanced Driver Assistance Systems in Europe*, Sistemas Avanzados de Ayuda al Conductor en Europa) es un conjunto de 30 proyectos de menor alcance que trabajan en diversos aspectos de la asistencia al usuario, incluyendo conducción autónoma o cooperativa, prevención de colisiones, interfaces hombre máquina y mapas multimedia.
- AIDE (*Adaptive Integrated Driver-vehicle Interface*, Interfaz Adaptativo Integrado Conductor-Vehículo) es un proyecto integrado FP6 que atiende al modelado de comportamiento, concentrándose principalmente en la interfaz hombre máquina para sistemas avanzados de asistencia al conductor.
- APROSYS (*Advanced Protection Systems*, Sistemas de Protección Avanzados) es un proyecto integrado en mecanismos pasivos de seguridad. Las tecnologías estudiadas en APROSYS están más relacionadas con un sistema inteligente de disminución de daños, mientras que *Caring Cars* está más centrado en la prevención de accidentes.
- EASIS (*Electronic Architecture and System Engineering for Integrated Safety Systems*, Arquitectura Electrónica e Ingeniería de Sistemas para Sistemas de Seguridad Integrados) es un proyecto orientado a mejorar la integración de sistemas de sensores y control en automóviles. EASIS busca una integración ajustada de los sistemas embarcados, contempla que su sistema forme parte de los sistemas electrónicos del automóvil. *Caring Cars* apunta hacia un mayor nivel de

integración de los sensores y aplicaciones embarcadas, de forma inteligente y autoconfigurable.

- EFCD (*Enhanced Floating Car Data*, Datos Flotantes del Automóvil Mejorados) es un subproyecto del proyecto integrado *Global System for Telematics* (Sistema Global Telemático). Atañe específicamente a la recolección de información de los sensores y la comunicación de la misma a centros de servicios externos. La tecnología desarrollada por este proyecto es relevante, aunque *Caring Cars* también se centra en el flujo de información en sentido contrario, llevando información de servidores externos al vehículo.
- HUMANIST (*HUMAN centred design for Information Society Technologies*, diseño centrado en personas para Tecnologías de la Sociedad de la Información) es una Red de Excelencia que estudia el factor humano en la seguridad electrónica. Sus actividades incluyen el modelado cognitivo del conductor, análisis de usabilidad, el diseño centrado en el usuario, y la educación y entrenamiento de los conductores. Aunque algunos de estos aspectos son relevantes para *Caring Cars*, HUMANIST no cuenta con el enfoque hardware de *Caring Cars*.
- PReVENT (*Preventive and Active Safety Applications*, Aplicaciones Preventivas y Activas de Seguridad) es un proyecto integrado muy amplio, que trabaja en varios sistemas para detectar riesgos, monitorear a los conductores y evitar accidentes. *Caring Cars* complementa este proyecto, poniendo más énfasis en sensores innovadores y su integración flexible.
- MyHeart (Prevención de enfermedades cardiovasculares mediante un estilo de vida preventivo y el diagnóstico temprano) es otro proyecto integrado enfocado en las mejoras sanitarias para los ciudadanos. Este proyecto se basa en la integración de sistemas en la propia ropa mediante sensores textiles integrados.

Resulta obvio que algunas de las áreas abordadas por este proyecto están siendo investigadas en otros trabajos. Sin embargo, una aproximación integrada como la que se propone en *Caring Cars*, en el que los sensores están conectados a través de redes que permiten compartir información de forma abierta y flexible, es lo que diferencia los

resultados esperados de este proyecto respecto al actual estado del arte [Reina et al., 2012; Murgoitio et al., 2007, 2008, 2009, 2010].

### **Accidentes de tráfico y factor humano**

La salud y el tráfico son dos aspectos de la vida europea que representan costes muy elevados para la sociedad. Según informes del *European Environment Agency* [*European Environment Agency*, 2014], los costes de tráfico en carretera son muy altos, y se pueden estimar en un 8% del PIB europeo. Los accidentes y la contaminación son los responsables de gran parte de estos costes.

### **Accidentes de tráfico**

De todos los modos de transporte, el traslado por carretera es con creces el más peligroso y el que se cobra más vidas humanas. Los accidentes de tráfico representan en la actualidad uno de los principales problemas de salud pública y se sitúan en el quinto lugar entre las principales causas de muerte, tras las enfermedades del aparato circulatorio, los tumores, las infecciones respiratorias y los trastornos digestivos.

Según informes del Banco Mundial [Banco Mundial, 2014], cada año mueren en todo el mundo 1,17 millones de personas en accidentes de tráfico, y más de 10 millones resultan heridas. Según el citado organismo, el coste de este tipo de accidentes se sitúa entre el 1 y el 3% del Producto Nacional Bruto, dependiendo de la zona.

La propia Comisión Europea introdujo esta dimensión económica en su política de seguridad vial, y ha evaluado el coste directo o indirecto en 160 mil millones de euros, lo que representa un 2% del PNB de la Unión Europea (UE). Algunos grupos de población o categorías de usuarios son especialmente vulnerables: los jóvenes de 15 a 24 años (10.000 muertos al año), los peatones (7.000 muertos) o los ciclistas (1.800 muertos).

En España, el coste anual de los accidentes de tráfico equivale a cerca del 2% del producto interior bruto, muriendo cada año más de 5.000 personas, y resultando heridas de mayor o menor gravedad unas 130.000, según las cifras de la Dirección General de Tráfico [DGT, 2014].

Es estremecedor pensar que entre los 15 y 30 años uno de cada dos muertos lo son por causa del tráfico, un problema que además cuesta al año a los españoles unos 15.400 millones de euros.

Por esta razón, la última estrategia de seguridad Vial 2011-2020 [DGT, 2014] persigue un marco de actuación que impulse, facilite y coordine las iniciativas de seguridad vial de todos los agentes: políticos, económicos y sociales a nivel nacional y, a pesar de que los accidentes de tráfico se han venido reduciendo, este descenso en la siniestralidad, cercano a un 3%, continúa lejos de las exigencias de la Unión Europea, que pretende que los accidentes mortales desciendan en un 50%.

A pesar del esfuerzo de numerosos organismos para luchar contra la alta siniestralidad de nuestro país, los objetivos marcados por el órgano comunitario se hallan muy por encima de los conseguidos en nuestro país, aunque esto es extensible a la mayoría de los países de la zona euro. Según la Dirección General de Tráfico (DGT), “lo conseguido no es como para estar excesivamente orgulloso, pero sí representa un pequeño descenso”.

Si pensamos que en España cada año tenemos un millón más de automóviles en circulación, y la movilidad de los vehículos ahora es muy grande, estamos en una buena línea. No obstante, desde la DGT subrayan que les gustaría que ese 3% se acercara a un valor entre un 10% o un 15% de reducción en la siniestralidad.

### **Factor humano**

Aproximadamente, en el 80% de los siniestros está implicado el factor humano, ya sea de forma directa o indirecta (distracciones, cansancio, somnolencia, alcohol, drogas, medicamentos, teléfonos móviles, nuevas tecnologías que aturden al automovilista).

Según informes de la NHTSA –órgano rector del tráfico en EEUU– [NHTSA, 2014] los fallos humanos que con mayor regularidad se repiten, son los relacionados con una deficiente atención en la conducción y la fatiga. Un estudio de este organismo señala que “los despistes a los mandos del volante son la causa del 25% de los seis millones de accidentes que se producen en el mundo al año”.

En efecto, según los datos de la DGT [DGT, 2014], una cuarta parte de los accidentes de tráfico son causados por la fatiga o la somnolencia del conductor, ya que ambas circunstancias dificultan la concentración, aumentan nuestro tiempo de reacción ante un imprevisto (tardamos más en pisar el freno) y consiguen que nos equivoquemos más fácilmente en la estimación de velocidades y distancias.

Las distracciones al volante son, por su coste en vidas humanas, una epidemia comparable con la más temida de las enfermedades. Para luchar contra ellas, las líneas de actuación pasan, en primer lugar, por incrementar y mejorar la seguridad activa (e.g. el ABS o los controles de estabilidad y tracción que corrigen el fallo humano) y pasiva de los automóviles (e.g. airbags, cinturones de seguridad, etcétera que minimizan las consecuencias de los accidentes mediante) y, en segundo, desarrollar infraestructuras más seguras.

Otro frente abierto en esta lucha está ligado directamente al conductor. Al suceder algo que no es previsible se lo cataloga de accidente, pero todos los informes elaborados por distintos organismos, como el Instituto Universitario de Tráfico y Seguridad Vial (INTRAS) o el Real Automóvil Club de España [RACE, 2014], subrayan que muy pocos “sinistros de circulación” son “accidentales”. Por ejemplo, no podemos evitar que se produzca una tormenta, pero el conductor puede adoptar una serie de medidas para reducir el peligro de sufrir un choque, como aminorar la velocidad o aumentar la distancia de seguridad.

Los factores que componen el fallo humano, responsable del 80% de los accidentes, son múltiples. El INTRAS destaca en primer lugar los errores cometidos por el conductor. Entre éstos, los más comunes son: el incorrecto reconocimiento o identificación de la vía, los errores en la ejecución de una maniobra y la falta de precisión en la toma de decisiones. En segundo lugar, existen una serie de agentes que dificultan la conducción. Entre ellos destacan: los problemas físicos del conductor que alteran la conducción, como la fatiga; los problemas psíquicos como la falta de atención y la agresividad al volante; el consumo de sustancias que alteran nuestras aptitudes (alcohol, drogas y fármacos); y las enfermedades como la depresión y el estrés. Por último, cabe mencionar los estados inhibidores de la

prudencia. Los más comunes son la sobrestimación de la propia habilidad o una adaptación inadecuada de la velocidad a las características de la vía.

Para circular con seguridad, todo conductor que se precie debe mantener un alto grado de atención. Partiendo de esta premisa, debemos saber que nuestra capacidad para gestionar la información es limitada, por lo que tenemos que elegir qué datos procesar y cuáles no. Todo ello nos acerca a los conceptos de “**atención**” y/o “**nivel de alerta**”.

## 1.2. Justificación del estudio

Las estadísticas sobre la siniestralidad en las carreteras españolas en 2014 plantea dudas sobre la capacidad de mejora del sistema de seguridad vial. Según datos de la Dirección General de Tráfico [DGT, 2014] y del Real Automóvil Club de Cataluña [RACC, 2014], la siniestralidad en carretera en España sólo se redujo un 0,3% en 2014, la más baja desde 2003 (Figura 4, Figura 5):

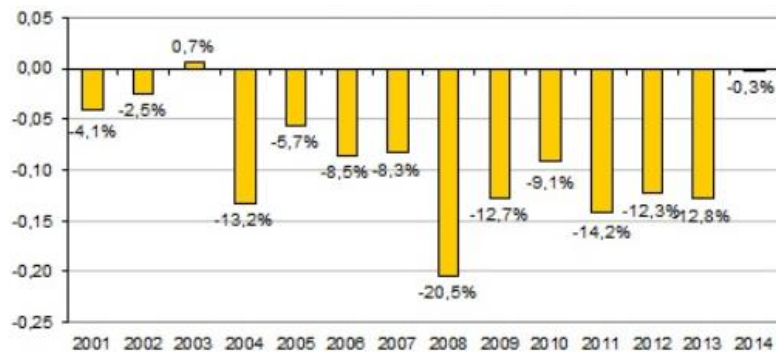


Figura 4: Tasa de variación de víctimas mortales en carretera en España (2001-2014)

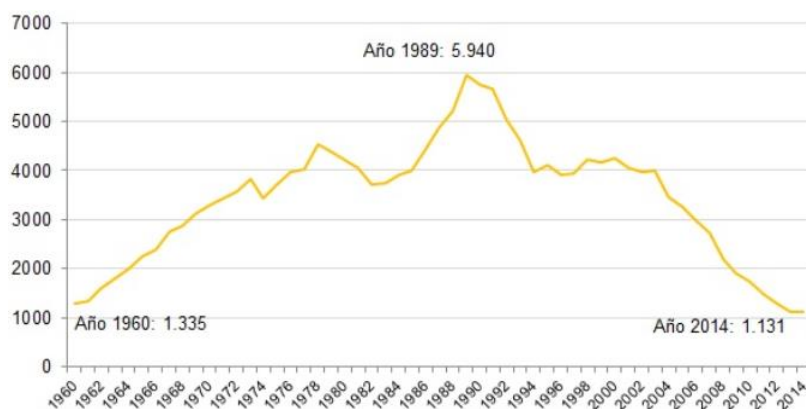


Figura 5: Víctimas mortales en carretera en España (1960-2014).

Estas estadísticas de siniestralidad vial dependen de circunstancias muy diferentes: volumen de tráfico, tipo de accidente, tipo de vehículo y su antigüedad, ocupación del mismo, climatología, tipo de carretera, cumplimiento de las normas, en especial límites de velocidad o conducir bajo el influjo de alcohol o medicamentos que alteren la capacidad de conducir. Además, la actitud del conductor, puede verse afectada por circunstancias del entorno que incidan en su capacidad de atención.

Analizar la “no reducción” significativa de la siniestralidad en 2014 significa intentar responder a estas cuestiones, pero en cualquier caso uno de los principales aspectos a mejorar estaría relacionado con el factor humano, i.e. la actitud de los conductores al cumplir las normas, una mayor prudencia al volante, etc.

Relacionado con lo anterior y el concepto de “Inteligencia Ambiental” (ver capítulo de “Antecedentes”) aplicada a la seguridad vial, se vienen evaluando las posibilidades de medir mediante sensores en los vehículos y/o conductores de manera que se identifique la información más objetiva posible y que cuantifique la fatiga y la correspondiente disminución del nivel de alerta en el conductor.

Desde este punto de vista, la sensorización mencionada en el párrafo anterior plantea principalmente dos cuestiones a resolver:

- 1- Obtención de la información básica de forma no intrusiva y con suficiente calidad para ser posteriormente analizada. Básicamente esta cuestión suele ser un problema de eliminación de ruidos, nada sencillo de resolver en muchos casos.
- 2- Análisis de la información obtenida con posibilidades de reportar medidas relacionadas con el nivel de alerta y la fatiga suficientemente objetivas, una vez obtenida la información básica con suficiente calidad.

La investigación que se recoge aquí **se centra en la segunda de estas cuestiones**. Más específicamente en el estudio de la denominada “Variabilidad de la frecuencia cardíaca” (HRV), inferida a partir de la señal fisiológica del ECG, su comportamiento en la conducción de vehículos durante recorridos largos de más de dos horas [RACE, 2011], y el análisis de la influencia de factores personales como la edad, así como de los ciclos

circadianos (diferentes momentos del día) sobre diferentes factores derivados de este tipo de medidas objetivas relacionadas con el nivel de alerta y la fatiga.

Además, esta monografía orienta el resultado de todo lo anterior a su posterior utilización en aplicaciones o herramientas de evaluación, que basadas en las últimas técnicas de inteligencia artificial y minería de datos, sean capaces de producir estimadores cuantificables sobre la idoneidad de los nuevos HMI (*Human Machine Interface*) que se están integrando en los vehículos de última generación para su manejo por carretera; todo ello con objeto de poder ofrecer información más completa y objetiva al sector de automoción (Wolkswagen, Ficosa, Seat, CIE, ...) sobre sus nuevos desarrollos, e.g. manejo de vehículos mediante joysticks en lugar de volantes, pedales hápticos, comandos por voz, vehículos autónomos, etc... [Murgoitio et al., 2007, 2013; Murgoitio, 2013].

### **Factor humano**

La principal razón por la que se justifica este interés en la profundización de la evaluación del estado del conductor de vehículos motorizados, principalmente conductores de automóviles o camiones para tramos largos, y la utilización de métodos objetivos para medir el nivel de alerta y la fatiga, en muchos casos deterioradas por los significativos cambios en los nuevos HMI (mayor número de dispositivos, más información, cambios en las formas de interactuar con las máquina, etc.) de los vehículos, es por tanto la importante influencia que el factor humano tiene sobre el número de accidentes de tráfico.

Según un informe sobre fatiga elaborado por el RACE [RACE, 2014], durante estos años de trabajo, se ha investigado la incidencia de la fatiga, sus causas para mejorar el conocimiento de los usuarios en los siguientes temas:

- Identificación de los síntomas
- Análisis del riesgo
- Incorporación de hábitos saludables al volante
- Información sobre la parada de descanso
- Descansos y tipos de vehículos
- La prevención desde el entorno laboral

En definitiva, los últimos datos revelan un cambio de tendencia, tanto en el número de fallecidos como de siniestros mortales relacionados con la fatiga y la somnolencia en las carreteras de España. Y es que los datos referidos a fatiga y somnolencia también rompen con el resto de factores principales de la accidentalidad, como velocidad, distracción o infracción, que sí siguen la reducción en la siniestralidad.

### **Métodos de evaluación basados en medidas fisiológicas**

Asumido el interés que existe en encontrar métodos objetivos para la valoración del nivel de alerta y fatiga en conductores de vehículos, los métodos que se pueden utilizar se agrupan en las categorías siguientes:

- Efectos del nivel de alerta en los parámetros mecánicos del vehículo (velocidad, movimientos del volante, trayectorias del vehículo, etc.).
- Efectos del nivel de alerta en el comportamiento físico del sujeto (posición y movimientos corporales, expresión facial, etc.).
- Efectos del nivel de alerta en el comportamiento de variables fisiológicas (EEG, ECG, EOG, etc.).
- Sistemas mixtos.

En cualquier caso, en base a los últimos avances en sensórica, microelectrónica y técnicas de procesamiento de señal, se puede resumir que las dos aproximaciones más relevantes para determinar el nivel de alerta de los conductores son:

- El comportamiento de los parpadeos por medio de cámaras de video.
- La señales fisiológicas.

La primera tiene la ventaja de ser no intrusiva, atrayendo el interés de diferentes investigadores en todo el mundo, aunque tiene dificultades técnicas y de elevados costes que no son fácilmente superables, al menos en el corto plazo. Además, las diferentes fuentes bibliográficas ya anotadas señalan dificultades mucho mayores como las debidas a la gran variabilidad interpersonal y a las conclusiones contradictorias en cuanto a la tasa de parpadeos.

Respecto a la segunda posibilidad, son patentes algunas lagunas para su aplicación práctica: la gran **variabilidad interpersonal** de estos métodos que, junto con el problema de intrusividad al requerir de electrodos puede ser la causa de que este tipo de sistemas sean más difíciles de aplicar. En cualquier caso, creemos que existen suficientes estudios que soportan su idoneidad, por lo que en el presente trabajo de investigación se optó por valorar la utilización de métodos basados en señales fisiológicas, y más concretamente en las señales EEG y ECG, para finalmente profundizar en esta última.

En cuanto al EEG se llevaron a cabo varios contactos con el servicio de salud OSAKIDETZA del gobierno autónomo de País Vasco, concretamente con la unidad de NeuroFisiología del hospital de Cruces en Baracaldo (Vizcaya), el cual cedió uno de sus equipos con el objetivo de llevar a cabo determinadas pruebas. Así mismo, fueron contactados varios suministradores de este tipo de equipos, en este caso más ligeros y portátiles, con los que las casas comerciales realizaron demostraciones, tanto de sus equipos como de sus programas de análisis de las señales EEG recogidas (Figura 6).



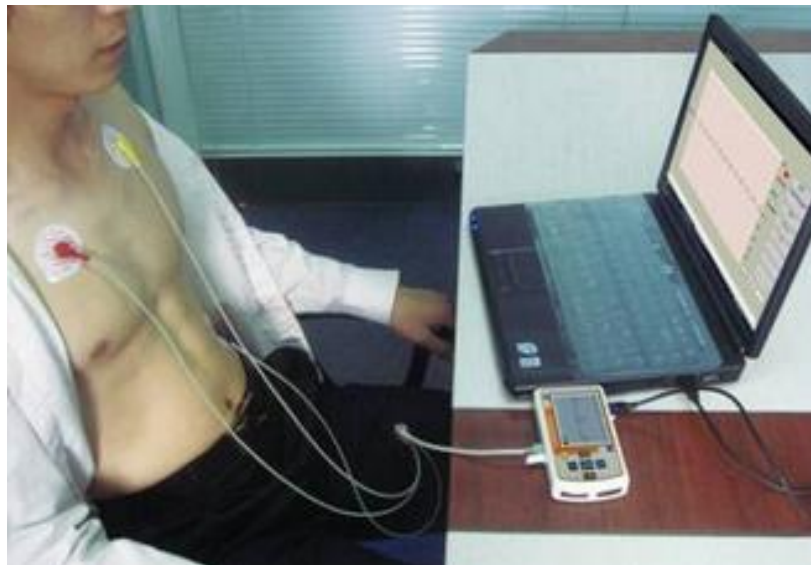
**Figura 6: Equipo EEG**

### **Métodos de evaluación basados en el ECG**

Finalmente, la opción de profundizar en la utilización de esta técnica fue desechada por la complejidad de tratamiento de esta señal (numerosas señales del orden de micro-voltios), la necesidad de utilizar equipamiento más difícil de usar (gorros especiales con gel específicos) y mucho más caros, así como su mayor dificultad para su posible adquisición

de forma lo menos intrusiva posible para el entorno de conducción. Hay que tener en cuenta que el proyecto “*Caring Cars*”, origen de las actividades de este trabajo de investigación (véase capítulo de “Antecedentes”), contemplaba la posibilidad de acercar al mercado productos que, como resultado de los desarrollos del mismo, permitan aumentar la seguridad durante la conducción de automóviles y por ende la disminución de accidentes de tráfico. En base a este criterio, los desarrollos basados sobre el ECG ofrecen más posibilidades que aquellas basadas en el EEG.

Según esto, se optó por probar con equipos que pudieran obtener información de la señal ECG, y para ello se contó con la colaboración inestimable de la unidad de cardiología del hospital de Galdakao perteneciente al servicio vasco de salud OSAKIDETZA (Figura 7). De esta manera, la mencionada unidad de cardiología cedió uno de sus equipos para realizar varias pruebas preliminares de adquisición de la señal ECG en conducción real durante trayectos largos (más de dos horas), así como la aplicación de análisis de los datos correspondiente.



**Figura 7: Equipo ECG**

Las conclusiones fueron que, confirmada la mayor facilidad de uso de este tipo de señal con respecto al EEG y su análisis posterior, el tratamiento del ECG ofrecía más posibilidades a una futura integración como sensor especial en el entorno de la conducción, minimizando el nivel intrusivo de los dispositivos necesarios. Hay que tener en cuenta que, aunque existen equipos de adquisición ECG con diez electrodos y mayor calidad de la

señal recogida, también es posible utilizar únicamente tres electrodos si el objetivo es simplemente detectar la denominada onda “R” del complejo PQRST típico con el que se caracteriza al ECG (ver apartado “Naturaleza del ECG”), más fácil de detectar y diferenciar.

Así, a pesar de ser considerado menos fiable por los expertos que la señal EEG, se ha optado por profundizar en la utilización de la señal ECG y el análisis de parámetros derivados de la misma como el HRV (*Heart Rate Variability*), cuyo estudio se basa en la detección de la mencionada onda “R”, más fácil de ser detectada y tratada. Además, desde el punto de vista eléctrico, mientras en el tratamiento de la señal EEG se trata con señales originales de micro-voltios, en la adquisición del ECG hablamos de milivoltios.

### **Ritmos circadianos**

De entre los numerosos parámetros que condicionan el factor humano el denominado ritmo circadiano es uno de los más importantes. Los humanos vivimos sujetos a un ciclo vigilia-sueño llamado ritmo circadiano donde dentro de un ciclo circadiano de 24 horas hay dos períodos donde el nivel de somnolencia disminuye. El primero ocurre durante la noche y en las primeras horas de la mañana y el segundo en la tarde. Durante estos períodos de sueño, el rendimiento y estado de ánimo subjetivo se degradan [Rosekind, 1999].

La influencia del ritmo circadiano sobre los accidentes de tráfico fue demostrada por Pack et al. [1995] quienes analizaron los datos de los accidentes ocurridos en Carolina del Norte y encontraron un máximo principal en su ocurrencia durante la noche y otro a media tarde. Así mismo, Wylie [1996] examinó la pérdida de alerta y la degradación del rendimiento de 80 conductores de camiones de USA y Canadá, encontrando la existencia de picos de alerta reducida tanto en la noche como al amanecer. También llegó a la conclusión de que la hora del día es **el factor más consistente que influye en la fatiga de los conductores**. También pueden verse otros estudios relacionados en: [Cinaz et al., 2013; Gil et al., 2013; Peltokangas et al. 2012; Yoshizaki et al. 2013].

La consecuencia práctica que se deriva de estas evidencias es la necesidad de tener en cuenta la hora del día a la que se conduce para controlar el efecto de la fase circadiana y evaluar el efecto de la misma.

## Períodos de descanso

Según estudios de la DGT y el RACE [RACE, 2011, 2014], mediante pruebas médicas y psicotécnicas, se ha comprobado que no descansar cada dos horas, puede aumentar el error de percepción de las distancias en un 63% o reducir a un 60% la capacidad de percibir y procesar información, como a la hora de ver señales o cualquier otro estímulo relevante para la circulación. Además, se incrementa por encima del 85% el tiempo de reacción al volante.

En dichos estudios, también se recuerda la importancia de prevenir el cansancio parando cada 150 kilómetros, tras 1,5 h de conducción o realizar unos estiramientos, lo que ayuda a que la capacidad de atención no disminuya, y de esta manera el tiempo de respuesta ante un obstáculo es un 26% más rápido.

Como consecuencia, esta monografía se justifica en la necesidad de profundizar en el análisis de la influencia de estas variables con el objetivo de ser consideradas en algoritmos mejorados para la evaluación del nivel de alerta y la fatiga en la conducción de vehículos motorizados durante trayectos largos y basados en la adquisición de la señal ECG. En particular el estudio se fundamenta en el análisis del HRV, asentado en la detección de la onda “R” del complejo “PQRST” típico del ECG, y su aplicación a herramientas *SoftWare* (SW) que permitan medir y validar las nuevas maneras de interactuar con los vehículos de última generación, al menos de forma relativa, integrando la información que mejor se correlaciona con **factores personales** (edad), **temporales** (ciclos circadianos) y **ambientales** (tipo de carretera),.

### 1.3. Hipótesis y objetivos

Teniendo en cuenta todo lo anterior la hipótesis de esta monografía se enuncia de la siguiente manera:

**Hipótesis:** La información contenida en la variabilidad de la frecuencia cardíaca o HRV forma un sistema completo y suficiente para evaluar de forma comparativa diferentes sistemas inteligentes de transporte o ITS aplicables en los interfaces hombre-máquina o HMI de los vehículos por carretera.

Con el objetivo de encarar esta hipótesis, este trabajo analiza la influencia de varios factores de diferente naturaleza, i.e. personales (edad), temporales (ciclos circadianos) y ambientales (tipo de carretera), con el **objetivo general** de integrarlos y optimizar el uso de la señal ECG en estudios de evaluación y validación para nuevas soluciones que se propongan en la mejora de la seguridad del transporte por carretera, i.e. estudios de investigación y desarrollo que requieren medidas objetivas para la validación de nuevos productos relacionados con sistemas inteligentes de transporte o ITS y que puedan implicar modificaciones en los niveles de seguridad vial.

Al servicio del objetivo general mencionado anteriormente, el presente trabajo de investigación se centra en los siguientes **objetivos específicos**:

- 1- Analizar la influencia de factores **personales** como la edad sobre el HRV durante la conducción prolongada de vehículos motorizados, y seleccionar los parámetros más relacionados.
- 2- Analizar la influencia de factores **temporales** como los ciclos circadianos sobre el HRV durante la conducción prolongada de vehículos motorizados (mañana/tarde), y detectar los parámetros más relacionados.
- 3- Analizar la influencia de factores **ambientales** como el tipo de carretera sobre el HRV durante la conducción prolongada de vehículos motorizados, y seleccionar los parámetros más relacionados.
- 4- Analizar la influencia de todos los factores (personales, temporales y ambientales) analizados con anterioridad e integrarlos para la evaluación de la **fatiga** durante la conducción prolongada de vehículos motorizados.
- 5- Desarrollar una nueva **aplicación**, que basada en los análisis anteriores, permita evaluar y comparar diferentes maneras de interactuar con vehículos de transporte por carretera.

Los objetivos específicos anteriores se sustentan en la mejora de los algoritmos de medida de la fatiga basados en la señal biológica del ECG obtenida durante la conducción de

vehículos terrestres motorizados y su análisis HRV, y su utilización como indicador objetivo del estado del conductor para la validación de los nuevos interfaz Hombre-Máquina (HMI). Estos algoritmos, basados en técnicas de IA y minería de datos, se utilizarán en nuevas aplicaciones informáticas al servicio de proyectos relacionados con los nuevos sistemas inteligentes de transporte (ITS) que vienen aplicándose a las últimas gamas de vehículos y que sean potenciales modificadores de la seguridad vial.

Para ello, según toda buena tesis doctoral, esta monografía se **estructura** de la siguiente manera: El capítulo 1 contiene de manera introductoria los antecedentes, la justificación del estudio, así como la hipótesis de trabajo y los objetivos. El capítulo 2 analiza el estado del arte sobre la naturaleza de la atención, el nivel de alerta, los fundamentos del HRV, y sus métodos de análisis. El capítulo 3 expone la metodología utilizada, donde se explican las etapas en las que se ha estructurado, la descripción del recorrido por carretera diseñado para la conducción del vehículo, los sujetos participantes, el modelos de datos y variables tenidas en cuenta, así como los casos de estudio en los que se dividido toda la investigación. El capítulo 4 recoge el análisis y resultados de los cuatro casos de estudio en los que se ha basado este trabajo, con sus conclusiones parciales por cada uno de ellos. En el capítulo 5 se describen las conclusiones finales resumidas, y se enumeran los potenciales trabajos futuros que extiendan lo logrado en esta tesis doctoral. El capítulo 6 hace una descripción de la aplicación desarrollada en base a dichas conclusiones. Finalmente, se incluyen la bibliografía utilizada (capítulo 7), y los anexos (capítulo 8) con la lista de participantes y autorizaciones, cuadro resumen de resultados, equipamiento utilizado, y el manual de usuario de la aplicación obtenida.

## **2. ESTADO DEL ARTE**

### **2.1. Naturaleza de la atención**

La atención, considerada dentro del entorno de la conducción, puede considerarse un proceso selectivo en el que no siempre interviene la voluntad del conductor.

Un conductor que desvía su atención de la carretera cuando su vehículo circula a 120 kilómetros por hora, recorre casi 100 metros sin control, y cada año se producen mayor número de siniestros por este tipo de circunstancias (falta de atención).

Entre el amplio abanico de factores que reducen la atención del conductor, unos proceden del propio individuo -internos-, otros tienen su origen en el ambiente que envuelve al conductor y, por último, otros son las distracciones que ocasionan todos los elementos telemáticos que incorporan los coches.

Los expertos han concretado los principales efectos de las distracciones en la conducción: la desviación de la trayectoria respecto a la posición del carril, circular con una velocidad inadecuada con relación al resto del tráfico -normalmente se circula más despacio-, mantener una distancia de seguridad inapropiada -se tiende a circular muy cerca del vehículo que nos precede- y un mayor tiempo de reacción ante una situación imprevista.

A continuación vamos a analizar algunas definiciones, características y modelos descritos por los expertos, para que nos acerquen a la naturaleza de la atención y el nivel de alerta en general.

#### **2.1.1. Definiciones**

Desde distintas opciones teóricas y empíricas, siguen siendo discutidas muchas cuestiones relacionadas con la atención, su definición y su naturaleza. Hay autores, como Dixon [1981], que se preguntan si la atención es un proceso, un mecanismo, una característica de la voluntad o un estado concreto de la mente. El tema se complica cuando se analizan las funciones y características de la atención y su relación con los procesos psicológicos, como queda patente en el análisis que realizan distintos autores desde diferentes

posicionamientos teóricos [Posner, 1975, 1978, 1987; Dember y Warm, 1979; Ruiz-Vargas y Botella, 1981; Marrero y Torres, 1986; Bermejo, 1987; Tudela, 1992, Moñivas, 1993; Rosselló i Mir, 1996; García Sevilla, 1997].

En lo que parecen coincidir la mayoría de los autores es en considerar que el concepto de la atención hace referencia a un conjunto de fenómenos diversos. Así, por ejemplo, Posner y Boies [1971] destacaron la atención como un mecanismo de selección de información, como un mecanismo de capacidad limitada y como un mecanismo de alerta. Los tres fenómenos, según M. de Vega [1984] están estrechamente relacionados, por lo que se puede entender la atención como un sistema de capacidad limitada que realiza operaciones de selección de información y cuya disponibilidad o estado de alerta varía notablemente. Por otro lado, Luria [1975] concibe la atención en términos de selección de la información y de su control. Por su parte, Tudela [1985] destaca como fenómenos atencionales el reflejo de orientación, la alerta, la selectividad, la capacidad y la conciencia. Rosselló i Mir [1996] sostiene que a partir del análisis factorial de diferentes tareas atencionales se pueden distinguir, al menos, tres factores diferenciados relacionados con la atención: la selectividad, la resistencia a la distracción y la habilidad para pasar de un foco de atención a otro. En un análisis similar, García Sevilla [1997] destaca como procesos implicados en el mecanismo de la atención a los procesos selectivos, a los procesos de distribución y, por último, a los procesos de mantenimiento o sostenimiento de la atención. Por otro lado, Taylor [1995] destaca como principales dimensiones del fenómeno atencional, la atención intensiva, la atención sostenida, la atención selectiva y la atención controlada.

Entre las definiciones que consideramos más completas destacamos la propuesta por Posner, debido a su amplitud teórica y apoyo empírico derivado de las aportaciones más recientes de la Psicología cognitiva y de las denominadas neurociencias, especialmente de la Neuropsicología. Así, Posner [1978, 1987] concibe la atención como un mecanismo múltiple de naturaleza modular, que consta de diferentes submecanismos u operaciones elementales interrelacionadas y coordinadas entre sí. En una línea similar, Tudela define la atención como un "mecanismo central de capacidad limitada cuya función primordial es controlar y orientar la actividad consciente del organismo de acuerdo con un objetivo determinado" [Tudela, 1992]. Se trata, según Tudela, de un mecanismo de control de estructura modular cuya función básica es la regulación y dirección de los mecanismos

responsables del procesamiento. Desde una perspectiva más descriptiva, también cabe mencionar la definición inicial que nos proporciona García Sevilla [1997]: "mecanismo implicado directamente en la activación y el funcionamiento de los procesos y/u operaciones de selección, distribución y mantenimiento de la actividad psicológica".

Las teorías actuales se inclinan mayoritariamente hacia modelos explicativos de la atención, al concebirla como un agente responsable del control del procesamiento de la información y de la actividad del sujeto. La atención no se concibe como un proceso, en el sentido de una serie de cambios o transformaciones durante los distintos momentos del procesamiento, sino que es el mecanismo encargado de regular dicha actividad [García Sevilla, 1997]; es decir, el mediador funcional básico de todos los procesos cognitivos [Ruiz-Vargas y Botella, 1987]. La atención actuaría como un "mecanismo vertical", que activaría o inhibiría los diferentes procesos (percepción, memoria, razonamiento, toma de decisiones, etc.) según las necesidades de cada momento y situación, orientando de forma "horizontal" dichos procesos, siendo así como un mecanismo de control responsable de la organización jerárquica de los procesos psicológicos que elaboran la información [Rosselló i Mir, 1996]. En este mismo sentido, Posner y Friedrich [1986] conciben la atención como un mecanismo de control o regulador interno de la actividad cognitiva. El control de la ejecución de los procesos mentales también está reflejado en el modelo de Logan [1981], así como en el de Simon [1986], que añade además el control de la conducta.

### 2.1.2. Modelos teóricos

La Tabla 1 presenta una síntesis ordenada cronológicamente de los distintos modelos teóricos que sobre la atención se han venido teniendo en cuenta a lo largo de los últimos años:

MODELOS	AUTORES	SEGUIDORES	SÍNTESIS
DE FILTRO	Broadbent [1958]	Cherry [1953]	Existe una estructura central que no permite procesar más de un mensaje a la vez. Dicha estructura actúa a manera de un cuello de botella o filtro que regula la entrada de la información.
DE FILTRO	Treisman [1960]	Broadbent [1958]	El sistema central de capacidad limitada distribuye su capacidad entre todos los mensajes y no solo y exclusivamente entre mensajes atendidos; sin embargo, el mensaje relevante recibe un tratamiento especial en el filtro; el resto de los mensajes no atendidos son atenuados.

MODELOS	AUTORES	SEGUIDORES	SÍNTESIS
DE FILTRO	Deutsch y Deutsch [1963]	Broadbent [1954]	Señalaron, que mientras su teoría afirmaba que todas las informaciones entrantes eran analizadas, aseguraba asimismo que sólo las importantes originaban una respuesta.
DE FILTRO	Norman [1968]	Deutsch y Deutsch [1963]	Siguiendo el esquema básico de Deutsch y Deutsch, elabora un modelo atencional más articulado, dedicando el máximo esfuerzo al mecanismo analizador y haciendo intervenir los mismos de memoria.
DE RECURSOS	Kahneman [1973]	Norman y Bobrow [1975] Navon y Gopher [1979] Wickens [1980]	El modelo atencional de Kahneman partía de que la capacidad que posee el sujeto, siempre es la misma, con independencia del tipo de tarea que se tenga que desarrollar. El sujeto posee una cantidad de recursos indiferenciados que están disponibles para ser consumidos en cualquier momento.
DE RECURSOS	Baddeley y Hitch [1974]	Atkinson y Shiffrin [1968] Watkins y otros [1973] Reitman [1974]	Propusieron la sustitución del concepto de "almacenamiento a corto plazo" por el de "memoria operativa"; una de las diferencias principales consiste en que la noción de memoria operativa es más amplia que la de almacenamiento a corto plazo.
DE RECURSOS	Norman y Bobrow [1975]	Kahneman [1973]	Ampliaron el modelo de Kahneman y desarrollaron un marco conceptual para tratar los problemas de administración de capacidad o recursos, La idea central fue construir una teoría general del reparto de capacidad cuando ésta es variable pero limitada y cuando son más de uno los procesos.
DE RECURSOS	Johnston y Heinz [1978]	Shiffrin y Schneider [1977]	La atención selectiva consume capacidad de procesamiento y la cantidad consumida aumenta al atravesar las diferentes etapas o modalidades atencionales.
DE RECURSOS	Navon y Gopher [1979]		Este modelo es el más complejo y articulado, aún cuando puede considerarse una variante de los anteriores.
DE AUTOMA- TICIDAD	Posner y Snyder [1975] Schneider y Shiffrin [1977] Hasher y Zacks [1979]	Jonides [1985] Schneider y Shiffrin [1985] Naveh-Benjamin y Jonides Zbrodoff y Logan [1986]	Estos modelos son una continuación de los modelos de recursos y una crítica a los modelos atencionales de filtro.  Resaltamos, que no solo han estudiado este proceso sino otros, como la memoria, la activación semántica, etc.  Las teorías clásicas de la automaticidad han postulado la existencia de dos formas de procesamientos: los procesos automáticos y los controlados.  Mientras un proceso controlado consume una gran cantidad de recursos de procesamiento, un proceso automático no consume apenas capacidad atencional.

MODELOS	AUTORES	SEGUIDORES	SÍNTESIS
DE CONTROL ATENCIONAL	Logan [1981]	Norman y Shallice [1986] Baars [1987-88] Shallice [1988] Ruiz Vargas [1993]	A partir de la década de los 80, la atención se conceptualiza como un mecanismo que es capaz de controlar la ejecución de los procesos mentales. Este sistema actúa en situaciones diversas: cuando hay que planificar o tomar decisiones, cuando no existen soluciones familiares para un problema, o hay que inhibir una respuesta habitual.
APORTACIONES DEL CONEXIONISMO	Rumelhart y McClelland [1986]	Phaf, Van der Heijden y Hudson [1990]	Los modelos conexionistas se han interesado por simular el funcionamiento de los procesos de aprendizaje. Este modelo de la atención, es conocido con el nombre de modelo SLAM (Selective Attention Model). y simula el funcionamiento de los mecanismos de control de la atención.

**Tabla 1: Modelos teóricos sobre la atención**

Destacamos entre todos ellos el **modelo de Kahneman** [1973] por su influencia en todos los modelos posteriores, y sobre todo aquellos que están más relacionados con nuestra investigación referidos a las neurociencias y las aportaciones del conexionismo.

### 2.1.3. Características

Aunque depende de la orientación teórica de la que se parta, en la línea de las últimas investigaciones, podemos especificar como principales características: la amplitud, la intensidad, el oscilamiento y el control [García Sevilla, 1997].

- **Amplitud:** La amplitud o "ámbito de la atención" hace referencia a la cantidad de información a la que podemos atender a la vez, así como al número de tareas que el ser humano es capaz de ejecutar de forma simultánea.
- **Intensidad:** La intensidad o "tono atencional" se puede definir como la cantidad de atención que presta un sujeto a un objeto, actividad o tarea, caracterizándose por estar relacionada de forma directa con el nivel de vigilia y alerta del sujeto. Un aspecto fundamental del tono atencional es su variabilidad de unas situaciones a

otras, dentro del mismo día y en un mismo sujeto, dando lugar a “fluctuaciones de la atención”.

- **Oscilamiento:** El oscilamiento o “desplazamiento de la atención” (“*shifting*”) se refiere a la característica de la atención según la cual ésta cambia continuamente debido a la necesidad de procesar dos o más fuentes de información o bien porque tenemos que ejecutar más de una tarea de forma simultánea, dirigiéndose a una y otra de forma alterna.
- **Control atencional:** Esta característica se refiere a la activación de la atención y al despliegue de sus mecanismos de funcionamiento de una forma eficiente, respondiendo a las demandas y exigencias de cada situación. Frente a la denominada “atención libre” o no controlada, el control atencional supone un determinado nivel de esfuerzo cognitivo. Cada vez son más los autores que destacan la función de control de la atención dentro del funcionamiento del sistema cognitivo [Keele y Neill, 1978; Logan, 1978, 1980; Posner, 1978, 1987; Reason, 1979, 1984; Norman y Shallice, 1986; Baars, 1988; Shallice, 1988, 1989; Tudela, 1992; Ruiz-Vargas, 1993; Rosselló i Mir, 1996; García Sevilla, 1997].

#### **2.1.4. Factores determinantes**

Con el término de “factores determinantes de la atención” hacemos referencia a aquellas variables o situaciones que influyen de forma directa sobre el funcionamiento del mecanismo atencional [García Sevilla, 1997]. Estos factores se suelen agrupar en “**extrínsecos**” e “**intrínsecos**”. Ambos tipos de factores suelen actuar de forma conjunta e interdependiente, influyéndose y modulándose mutuamente.

##### **Factores extrínsecos**

Los factores extrínsecos son aquellos que, basados en las características físicas de los objetos, mejor captan y mantienen la atención, tales como:

- El tamaño: En general, los objetos de mayor tamaño atraen más la atención.

- La posición: Las investigaciones sobre este aspecto han puesto de manifiesto que la mitad superior izquierda de nuestro campo visual es la zona que primero capta la atención.
- El color: Normalmente, los estímulos en color atraen más la atención del sujeto que los estímulos en blanco y/o negro.
- El movimiento: Los estímulos en movimiento captan antes y mejor la atención del sujeto que los inmóviles.
- La novedad: El cambio de uno o varios atributos que forman el estímulo definen la novedad del mismo. La novedad es uno de los factores más importantes en la captación de la atención, constatado en diversas investigaciones [Berlyne, 1960, 1961; Berlyne y McDonnel, 1965],
- La intensidad: En general, los estímulos más intensos atraen más la atención.
- La complejidad: Viene definida por el grado de información que un estímulo transmite al sujeto. Normalmente, los complejos atraen más la atención que los simples.
- La relevancia: Un estímulo se define como relevante o significativo cuando es capaz de provocar cambios estimulares importantes para el sujeto, siendo un factor claro de captación de la atención.
- La repetición: Hasta un determinado momento, los estímulos que se repiten suelen captar la atención mejor que los estímulos que únicamente aparecen una vez y con escaso tiempo de permanencia.

### **Factores intrínsecos**

Los factores intrínsecos, denominados también factores determinantes internos o endógenos, se definen como el conjunto de factores que se relacionan de forma directa con el estado del organismo. Destacaremos los factores biológicos y fisiológicos, los factores evolutivos, los factores cognitivos y los factores dependientes de los estados transitorios del organismo.

**Biológicos y fisiológicos:** La Psicofisiología, la Neuroanatomía y la Neuropsicología han puesto de manifiesto el papel que juegan los sistemas cerebrales del organismo en los mecanismos de funcionamiento de la atención, constituyendo en la actualidad un punto de referencia fundamental. Las concepciones clásicas en el estudio de la Psicobiología de la atención, se refieren por un lado, a la concepción de la atención como condición energética del organismo, como activación, despertar o alerta, facilitando el procesamiento de la información y la ejecución de determinadas tareas; y, por otro lado, a la concepción de la atención como ser consciente de los cambios ambientales, tal como ocurre en la respuesta de orientación. Las investigaciones más recientes también han desarrollado nuevos campos de estudio, tales como el análisis de la atención visual, los correlatos psicofisiológicos que acompañan a la respuesta de orientación y a las tareas que exigen esfuerzo mental.

En este contexto hay que destacar que el **nivel de activación o “arousal”** constituye el principal determinante de la atención a nivel psicofisiológico, estando definido por el nivel de receptividad y capacidad de respuesta que el sistema nervioso posee en un determinado momento ante los estímulos ambientales, derivando en una serie de correlatos fisiológicos y conductuales. En general, a un mayor nivel de activación suele corresponder un mayor nivel de atención y de concentración, mayor capacidad de mantener mejor la atención durante períodos de tiempo relativamente amplios.

**Evolutivos:** Las últimas investigaciones han puesto de manifiesto que la atención también desempeña un papel fundamental en el desarrollo de la cognición, por lo que es necesario tener en cuenta los determinantes evolutivos del mecanismo atencional. A nivel selectivo, los oscilamientos de la atención son más lentos, en general, durante la infancia que en la edad adulta, lentitud que va disminuyendo progresivamente según avanza la edad para volver a aumentar a medida que vamos acercándonos a la que podríamos denominar “tercera edad”. El control del nivel de distracción también aumenta con la edad para volver a disminuir a medida que vamos envejeciendo [Faust et al., 2013; Voss et al., 2012].

**Cognitivos:** Entre los determinantes cognitivos cabe destacar los intereses y expectativas. En este contexto, conviene recordar, tal como hace el profesor Tudela [1992] que la Psicología de la motivación es previa a la Psicología de la atención, especialmente cuando hacemos referencia a la atención voluntaria, la más importante en el funcionamiento

cognitivo del ser humano. En general, parece constatarse que los estímulos que se encuentran dentro del ámbito de interés de un sujeto se perciben antes y mejor que aquellos otros que, en igualdad de condiciones, son neutros para él.

**Estados transitorios:** Por último, otro factor determinante de la atención está representado por los denominados "estados transitorios", los cuales son definidos como situaciones que se producen en un momento más o menos amplio de la vida de un sujeto y que influyen en su actividad mental y conductual [García Sevilla, 1997]. Respecto a los problemas relacionados con la atención destacan la fatiga, el estrés, el efecto de ciertas drogas y psicofármacos y el sueño:

- La fatiga es un estado psicofísico caracterizado por la disminución de la capacidad energética del sujeto, pudiéndose ser física y/o mental. Es un factor importante relacionado con la disminución de la capacidad para mantener la atención a medida que la fatiga aumenta.
- El estrés es un factor que aumenta el nivel de activación del sujeto, siendo un factor positivo respecto del funcionamiento de la atención hasta llegar a un nivel óptimo. Una vez que se supera ese nivel, el estrés puede ser considerado excesivo, y entonces perjudica el funcionamiento eficiente del mecanismo atencional, cuya consecuencia más evidente es el estrechamiento del foco atencional.
- Entre las drogas y psicofármacos cabe distinguir dos grupos con efectos totalmente distintos: los tranquilizantes y los estimulantes. En general, puede afirmarse que los tranquilizantes producen enlentecimiento de la atención [Coull, Robbins, Middleton y Sahakian, 1995; Fafrowicz, Unrug, Marek y Van-Luijtelaar, 1995], así como en los niveles de alerta, que disminuyen [Unrug, Van-Luijtelaar y Coenen, 1992]. Los estimulantes ayudan a incrementar el nivel de activación del sujeto, por lo que también afectan en este mismo sentido sobre la atención.
- Respecto al sueño, se ha comprobado que la privación del mismo produce un descenso considerable en el nivel de activación, lo cual conlleva una reducción o disminución en la capacidad para enfocar la atención sobre los estímulos relevantes o críticos, además de producir el aumento de la susceptibilidad a los

efectos de los estímulos distractores y disminución de la intensidad de la atención [Eysenck, 1982].

### **2.1.5. Técnicas de medida**

La constatación de las distintas manifestaciones de la atención, puestas de manifiesto mediante la investigación científica, va a suponer una diversidad de técnicas para medir el mecanismo atencional, caracterizadas todas ellas por utilizar un conjunto de estrategias estandarizadas. Dichas estrategias reciben el nombre de “paradigmas experimentales”, siendo características básicas de cada una de ellas: plantear un procedimiento específico de presentación de los estímulos y/o información al sujeto, por un lado; y por utilizar unas tareas concretas, por otro.

La evaluación o medida de las distintas manifestaciones de la atención se lleva a cabo en la actualidad, básicamente, por medio o a través de la utilización de diversas técnicas objetivas que evalúan la actividad psicofisiológica, la motora y la cognitiva.

#### **Actividad psicofisiológica**

Las medidas de la actividad psicofisiológica se basan en la serie de correlatos que presenta el mecanismo atencional a nivel psicofisiológico. Cada correlato psicofisiológico puede ser registrado mediante unas técnicas de registro concretas, entre las que destacan las siguientes:

- **Electroencefalografía:** las neuronas de la corteza cerebral emiten constantemente una actividad eléctrica con un determinado potencial, constituyendo la “actividad electroencefalográfica (EEG)”. Según el estado de activación, de sueño / vigilia, el potencial eléctrico varía, pudiéndose analizar mediante aparatos de registro que describen ondas con diversa frecuencia y voltaje. Este análisis consiste en el registro y estudio de la actividad eléctrica de las neuronas corticales. Dicho potencial se recoge en un electroencefalógrafo, aparato en el que la actividad EEG aparece como una onda sinusoidal en la que se puede apreciar un conjunto variado de frecuencias (alfa, beta, theta, delta, ...), en el que suele predominar una de ellas. Cada tipo de onda se diferencia en la banda de frecuencia, que se mide en ciclos

por segundo, en la amplitud y en las variaciones de frecuencia, que se relaciona con la persistencia del ritmo de la onda. Así, las ondas alfa se registran cuando el sujeto está despierto y relajado, presentando unas ondas que oscilan entre 8 y 10 Hz. Cuando el sujeto está muy concentrado, con alta activación, aparecen las ondas beta, con ciclos más rápidos, entre 18 y 30 Hz. También se pueden registrar ondas con menor frecuencia, especialmente en estados de sueño y en niños pequeños (ondas delta y theta). La principal aportación del registro de la actividad electroencefalográfica (EEG) es la información sobre el tono atencional general del sujeto en un momento determinado.

- **Registro de potenciales evocados:** las neuronas corticales también emiten otro tipo de actividad eléctrica que se superpone a la EEG espontánea de fondo. Aparecen en situaciones concretas, como cuando se producen cambios notables en la estimulación externa, y producen la respuesta conocida como potenciales evocados, de los cuales se pueden registrar su latencia, amplitud y polaridad. El procedimiento de registro de esta técnica se basa en la presentación de un estímulo en una determinada modalidad sensorial (visual, auditivo,...), que produce una respuesta eléctrica en la correspondiente región de la corteza cerebral. Al ser menor la amplitud de un potencial evocado que la de la EEG, una serie de electrodos se encargan del registro de la actividad que sigue al estímulo, eliminando las ondas de mayor amplitud propias de la EEG. El registro mediante ordenador permite la presentación repetida del estímulo y la obtención del promedio de las respuestas para cada electrodo. El potencial evocado se presenta como una onda senoide que registra, además, estas propiedades: la latencia, la polaridad, la frecuencia y la amplitud. Actualmente, se ha podido relacionar una serie de estos potenciales evocados con el funcionamiento de la atención, destacando los siguientes: el P100, registro de activación de las áreas corticales relacionadas con el procesamiento visual; el P300, en situaciones de espera activa o incertidumbre y que implique la concentración del sujeto; el N1, relacionado con la estimulación auditiva y con diversas variables personales como el nivel de activación; la VNC o Variación Negativa Contingente, relacionada con la espera atenta en situaciones a la presencia de un estímulo y el posterior movimiento cuando aparece.

- **Registro de la actividad electrodérmica:** se basa en el conjunto de cambios que sufren las propiedades eléctricas de la piel, denominándose también “respuesta electrodérmica (RED)” o "resistencia eléctrica de la piel”. Los cambios de la actividad eléctrica de la piel pueden ser tónicos o fásicos. La actividad tónica indica cambios que se producen en el nivel basal durante un período de tiempo determinado; mientras que la actividad fásica supone cambios breves y pasajeros que se producen en el nivel tónico de la AED (Actividad ElectroDérmica). Esta actividad puede tener lugar en ausencia de un estímulo ("respuestas no específicas") o bien ante la presencia de estímulos específicos ("respuestas específicas"). Consiste en el registro de la actividad eléctrica de la piel. Uno de los electrodos, colocados en áreas activas de la piel, produce una pequeña corriente eléctrica que atraviesa la piel y se dirige a un segundo electrodo, encargado de la captación del cambio que se produce en su resistencia, enviándola para su amplificación y registro a través de un polígrafo. Las medidas de respuestas utilizadas en la actividad eléctrica de la piel son las siguientes: el tiempo de latencia, la intensidad de la respuesta, el tiempo de recuperación media y el tiempo de elevación.
- **Electrocardiografía (ECG):** se basa en el número de latidos del corazón durante un período de tiempo determinado. Es un índice del nivel de activación del sujeto y de la atención que se presta a un determinado estímulo. Además, hay que tener en cuenta que las tareas que exigen atención y esfuerzo pueden provocar notables fluctuaciones de la frecuencia cardiaca. El método consiste en la medición de la frecuencia cardiaca mediante un aparato denominado electrocardiógrafo. Se colocan electrodos en la superficie y el electrocardiógrafo registra la señal ECG emitida por el organismo. Dicha actividad está formada por un conjunto de diversas ondas. La distancia entre las denominadas ondas R relacionadas con la contracción y la despolarización ventricular, permite calcular el tiempo que transcurre entre latido y latido, obteniéndose la correspondiente tasa cardíaca.
- **Electromiografía (EMG):** hace referencia a la actividad eléctrica desencadenada por los músculos cuando éstos se hallan en situaciones de actividad y reposo. Relacionado con el mecanismo atencional, la EMG está considerada como un índice de la actividad cognitiva, en el sentido de que aumenta de forma

significativa con las tareas que suponen un mayor esfuerzo atencional. También sirve como índice del nivel de activación o vigilia del sujeto. Se colocan unos electrodos sobre un músculo superficial, técnica predominante en psicofisiología ("electromiografía de superficie"), y se registran el número de despolarizaciones que se producen cuando dicho músculo es activado. La electromiografía se presenta como un conjunto irregular de ondas, cuyas medidas más utilizadas son la frecuencia y la amplitud de la onda registrada.

- **Pupilometría:** Es el registro y análisis de los cambios en el tamaño de la pupila. La pupila del ojo puede aumentar o disminuir como resultado de la contracción y relajación de la musculatura del iris y según determinadas circunstancias, como es el caso de la activación de la atención, ya que una de las manifestaciones más frecuentes de la atención es la dilatación pupilar. Este tipo de respuesta aparece cuando se producen estímulos nuevos o interesantes para el sujeto, así como cuando realiza tareas que suponen atención y esfuerzo psicológico. Se suelen utilizar dos técnicas diferentes. Por un lado, la denominada "fotometría refleja infrarroja", basada en la proyección sobre un ojo de un haz de rayos infrarrojos. El ojo refleja una luz que incide sobre una célula fotoeléctrica situada a unos tres centímetros de distancia del ojo, siendo registrada su intensidad por un osciloscopio. La otra técnica, conocida como "procesamiento de imágenes por video" consiste en la utilización de una cámara de vídeo que es sensible a los rayos infrarrojos y un dispositivo que analiza las imágenes de forma digital. Posteriormente, diversas y complejas técnicas de rastreo por ordenador permiten analizar el tamaño pupilar con una gran precisión. Las medidas más utilizadas mediante estas técnicas son la latencia y la velocidad de la constricción y dilatación pupilar y los diámetros máximos y mínimos de la pupila.
- **Técnicas de neuroimagen:** Este tipo de técnicas se emplea actualmente en una gran cantidad de ciencias, incluidas las denominadas "neurociencias", ya que ofrecen una información amplia y variada de la actividad cerebral. Se basan en la computación de imágenes cerebrales, es decir, en el registro de imágenes cerebrales por ordenador. Una de las técnicas más conocidas y posiblemente la más utilizada en el estudio de la atención es la "tomografía de emisión de positrones (PET)", que consiste en analizar el flujo sanguíneo en áreas localizadas

del cerebro. En el campo de la atención se registra el flujo cerebral de aquellas áreas que supuestamente están implicadas en las actividades atencionales. Previamente se plantean al sujeto una serie de tareas cognitivas y se registra el flujo cerebral que se produce durante la fase de realización de la tarea.

### **Actividad motora**

Las medidas de la conducta motora se realizan, en su mayor parte, a partir del registro y la cuantificación en las denominadas "hojas de registro", siguiendo el siguiente procedimiento: se enumera todo un conjunto de posibles respuestas que el sujeto puede manifestar en una situación determinada (como enderezar la cabeza; enderezar cabeza y tronco; orientar la cabeza en busca de información; señalar hacia la fuente de estimulación; desplazarse hacia la fuente de información; mirar hacia la fuente de estimulación, etc.) y a continuación se evalúan una serie de parámetros conductuales anteriormente prefijados a partir de un criterio establecido.

Los parámetros o unidades de medida más utilizados son la tasa de respuesta, la latencia y la duración de la respuesta. En determinadas ocasiones también es posible el uso de técnicas de registro más sofisticadas, como ocurre, por ejemplo, en el registro de los movimientos oculares a partir del uso de la técnica de la pupilografía, la cual está a medio camino entre las técnicas psicofisiológicas y las técnicas motoras. Dicha técnica permite los siguientes tipos de medida: por un lado, el recorrido, la duración y la longitud de los movimientos; y, por otro lado, las pausas y los puntos de fijación. Se distinguen entre "movimientos sacádicos" o movimientos bruscos del globo ocular cuya función principal es la de centrar la fovea sobre un estímulo, y "movimientos lentos", que se producen cuando el sujeto sigue a un objeto, así como cuando mantiene la línea de visión sobre un objeto estático. Ambos tipos de movimientos están considerados como índices destacados del mecanismo atencional. Los movimientos sacádicos se relacionan casi siempre con los procesos selectivos de exploración y búsqueda, asociados también con posibles oscilaciones de la atención, salvo en algunos casos ya que la atención puede variar sin que se produzcan movimientos oculares. Por otra parte, los movimientos lentos se suelen registrar para el estudio de la atención selectiva a determinados objetos móviles.

## **Actividad cognitiva**

Respecto a las medidas de la actividad cognitiva, la forma más utilizada es a través del análisis del rendimiento obtenido en la ejecución de una tarea determinada, siguiendo este procedimiento: el sujeto realiza una tarea o situación problemática, siendo ésta evaluada y/o medida en función de unas dimensiones de respuesta, entre las que cabe destacar las siguientes:

- Tiempo de reacción o latencia de respuesta: Es el tiempo que el sujeto tarda en comenzar una respuesta desde el momento en que aparece el estímulo que la suscita.
- Precisión de la respuesta: Número de aciertos y errores cometidos a lo largo de la tarea.
- Duración de la respuesta: Medición del intervalo temporal que transcurre desde que se inicia una respuesta hasta que ésta acaba.
- Tasa de respuesta o frecuencia de respuesta: Número de veces en que aparece la respuesta esperada, bajo determinadas condiciones.

Entre las tareas más utilizadas en la investigación básica destacan la detección, la discriminación, la identificación, el recuerdo, el reconocimiento y la búsqueda. La medida y evaluación de estas tareas se suele realizar mediante aparatos eléctricos y electrónicos, destacando el taquistoscopio y el ordenador en los estudios relacionados con la atención, aunque también se utilizan las pruebas psicométricas o test psicológicos, especialmente en el ámbito de la Psicología aplicada.

Aunque todos estos parámetros (medidas de la actividad psicofisiológica, medidas de la conducta motora y medidas de la actividad cognitiva) se utilizan en las investigaciones sobre la atención [Saroj et al., 2001; Qiang Ji y Xiaojie Yang, 2002; Zhao, 2012], lo que realmente les hace relevantes, dependiendo de los objetivos de cada investigación, es su participación en los diferentes paradigmas experimentales.

## 2.2. Nivel de alerta o “Arousal”

En el capítulo previo hablábamos de los “factores determinantes” de la atención, entendidos como variables o situaciones que influyen de forma directa sobre el funcionamiento del mecanismo atencional [García Sevilla, 1997], agrupábamos a dichos factores en “extrínsecos” e “intrínsecos”.

Entre los factores intrínsecos, (también denominados factores determinantes internos o endógenos, y definidos como el conjunto de factores que se relacionan de forma directa con el estado del organismo) se enumeran los factores biológicos y fisiológicos, destacando el nivel de activación o “*arousal*” (nivel de alerta) como el principal determinante de la atención a nivel psicofisiológico, y se define por el nivel de receptividad y capacidad de respuesta que el sistema nervioso posee en un determinado momento ante los estímulos ambientales, derivando en una serie de correlatos fisiológicos y conductuales.

También se evidenciaba la relación directa existente entre el nivel de activación y los tres mecanismos básicos de la atención: procesos selectivos, procesos de distribución y procesos de mantenimiento o sostenimiento de la atención [García Sevilla, 1997], poniendo de manifiesto que normalmente a un mayor nivel de activación corresponde lo siguiente: un mayor nivel de atención y de concentración, o una mayor capacidad de dar respuestas o de realizar tareas como de mantener mejor la atención durante períodos de tiempo relativamente amplios.

Sin embargo, también hay que tener en cuenta que las situaciones de gran estrés y “*arousal*” o nivel de alerta pueden provocar un deficiente funcionamiento del mecanismo atencional.

En este capítulo primeramente se hace una introducción al concepto y definiciones del “*arousal*” que se han dado, para posteriormente llevar a cabo una revisión de los términos y conceptos que se emplean en el trabajo relacionado con la alerta o “*arousal*”, como la fatiga y somnolencia. A continuación haremos un análisis de las referencias bibliográficas que tratan de los factores que inciden sobre ese estado de alerta, como la falta de sueño, fases circadianas, carga de trabajo, etc. Además, aunque ya en los capítulos precedentes sobre la atención ya se trataban, analizaremos el estado del arte de los métodos publicados

para la evaluación del nivel de alerta, pero esta vez ligándolos más a la conducción de vehículos y analizando sus características.

### 2.2.1. Introducción

El término *arousal* es sobre todo de amplio uso en medicina del sueño, aunque no siempre se emplea con el mismo sentido; proviene del inglés “*to arouse*” que significa tanto despertar como activar. Ambos significados concuerdan con el concepto de *arousal*: despertar y activación (cerebral y no cerebral: *arousal* vegetativo). Se consideran sinónimos los términos *microarousal*, fase de activación transitoria, activación cerebral y, sobre todo, despertar y/o **nivel de alerta**.

A pesar de su uso continuado, pocos autores han dado una definición precisa de este concepto. Las definiciones más ampliamente difundidas (Tabla 2) son la de Retschchaffen y Kales [1968], la del Glosario de Grass [1981], la de la ASDA [1992] y más recientemente la de Terzano et al. [1997].

Retschchaffen y Kales [1968]	Cualquier incremento EMG, en cualquier canal, que se acompañe de un cambio en el patrón de cualquier otro canal.
Grass, Glossary [1981]	Un cambio brusco durante el sueño de una fase profunda de sueño NREM a otra más ligera, o desde fase REM a vigilia con la posibilidad de un despertar total. El <i>arousal</i> puede acompañarse por un incremento de la actividad EMG tónica y de la frecuencia cardíaca y de los movimientos corporales.
ASDA [1992]	Un cambio brusco de la frecuencia EEG, que puede incluir theta, alfa y/o frecuencias mayores de 16 Hz pero no <i>spindles</i> , sujeto a una serie de reglas y condiciones.
Terzano et al. [1997]	Desincronización brusca del EEG (intrusión de ritmo alfa y/o beta durante fases de NREM); intrusión de frecuencias rápidas especialmente ritmos beta durante sueño REM. El cambio EEG puede asociarse con cambios EMG (ej. refuerzo del tono muscular).

**Tabla 2: Definiciones sobre el concepto “*arousal*”**

Todas estas definiciones se refieren preferentemente al *arousal*, o alertamiento electroencefalográfico, que se produce de manera transitoria durante el sueño, lo cual constituye el concepto más extendido de *arousal*. Sin embargo, el término *arousal* se entiende también como despertar comportamental y como situación de vigilancia.

Electrofisiológicamente el alertamiento o activación cortical se presenta como una desincronización de la actividad electroencefalográfica y un incremento generalizado de la actividad eléctrica y excitabilidad de los sistemas tanto motores como sensoriales. Esta desincronización cortical se relaciona con cambios eléctricos talámicos. Por otra parte, se ha descubierto en el tálamo [Mariotti M, Gritti I, Mancina M, 1998] una actividad rítmica muy lenta que se correlaciona estrechamente con osciladores del sistema simpático tanto en el animal como en el hombre. Estas oscilaciones lentas se expresan unas en el EEG por la variabilidad del intervalo R-R del electro cardiograma (ECG) y otras con la respiración. Las primeras serían reflejo de la actividad simpática y las segundas, de la actividad vagal. El origen de estas oscilaciones se situaría en el bulbo en las áreas reguladoras vasomotoras.

### **2.2.2. Conceptos asociados a bajos niveles de alerta**

En todos los estudios relacionados con accidentes producidos por un deterioro en el nivel de alerta se menciona repetidamente conceptos como fatiga psicofisiológica (diferente a la fatiga muscular), hipo-vigilancia y somnolencia, entre otros muchos. Conviene hacer una breve referencia a esos términos.

Desafortunadamente, todos son conceptos sumamente subjetivos y vagos por lo que difícilmente se puede hablar de definiciones universalmente aceptadas ni de conceptos de fácil aplicación práctica a la hora de diseñar experimentos y elaborar sus conclusiones. Esta indefinición se convierte en una barrera importante siempre que intentemos cuantificar el nivel de alerta de un sujeto de manera objetiva o reproducible. En la mayor parte de las definiciones se emplean, además, otros conceptos, igualmente subjetivos, como cansancio, letargo, etc., que configuran en las pretendidas definiciones, círculos viciosos con difícil salida. Conviene, sin embargo, tener a mano algunas definiciones de dichas voces y de sus limitaciones. Incluimos por sencillez y proximidad las que ofrece la Enciclopedia Médica en Español: (<http://www.nlm.nih.gov/medlineplus/spanish/encyclopedia.html>):

## **Fatiga**

La fatiga es una sensación de falta de energía, de agotamiento o de cansancio. Otros nombres alternativos son: Cansancio, abatimiento, letargo.

Otras definiciones de fatiga pueden extraerse de la literatura, poco coherentes entre sí. Grandjean [1979, 1988] definió la fatiga como un estado marcado por una reducida eficiencia y falta general de disposición para el trabajo. En un intento de aproximarse a definiciones más objetivas, se proponen referencias a indicadores psicofisiológicos, electrofisiológicos, de eficiencia y bioquímicos [Williamson et al., 2001]. La evaluación de esos conceptos es el siguiente gran problema; tanto si se basa en la propia percepción subjetiva de los sujetos de experimentación o la valoración por observadores, donde difícilmente se puede disponer de “*gold standards*”, como si se emplean métodos indirectos para la caracterización de aquellas variables mediante magnitudes físicas relacionadas.

## **Somnolencia**

La somnolencia es un deseo irresistible de quedarse dormido en un momento inapropiado. Otras definiciones de somnolencia, al igual que en el caso de la fatiga, pueden verse en la literatura, donde resalta el escaso consenso entre ellas. La somnolencia en conductores está caracterizada por una experiencia subjetiva de cansancio, alerta reducida y por el decremento en la capacidad de mantener atención y responder rápidamente [Charlton y Baas, 2001; Lisper et al., 1986; Nilsson et al., 1997; Williamson, et al., 1996].

Fatiga no es lo mismo que somnolencia. Por lo general, la somnolencia es la sensación de una necesidad de dormir, mientras que la fatiga involucra la falta de energía y de motivación. La fatiga puede ser una respuesta normal e importante al esfuerzo físico, al estrés emocional, al aburrimiento, a un descanso previo insuficiente o a la falta de estímulos. Sin embargo, también puede ser un signo o una señal no específica de un trastorno psicológico o fisiológico. La fatiga se ha descrito asociada o conectada a múltiples patologías.

Basados en los argumentos anteriores, hemos considerado conveniente que los términos subjetivos de cansancio, decaimiento, letargo, deterioro de la eficiencia, atención reducida, fatiga, etc., se refieran o estén asociados a bajos niveles de alerta, mientras que somnolencia se entienda como un estado extremo en el que el sujeto tiene la necesidad inevitable de cerrar sus ojos y descansar.

### **2.2.3. Factores que inciden sobre el estado de alerta**

Los factores que inciden en el estado de alerta de un individuo son innumerables. En los apartados que siguen no se pretende analizar ampliamente esos factores, cosa que no aportaría ningún conocimiento útil al presente estudio, sino identificar algunos de esos factores con unas relaciones causa-efecto claras con el nivel de alerta, para ser utilizados en el control del mismo nivel de alerta en los experimentos que se han diseñado. Incidimos en tres grupos principales: falta de sueño, factores circadianos, y carga de trabajo.

#### **Falta de sueño**

La privación acumulada de sueño provoca bajos niveles de alerta y somnolencia. Una pérdida regular de dos horas de sueño por noche puede conducir a una somnolencia crónica [Horne J., 1995].

El grado de eficiencia humano parece ser dependiente del efecto acumulativo de la fatiga [Féry, Ferry, vom Hofe y Rieu, 1997; O'Neil et al., 1999]. En nuestra investigación, la privación de sueño será utilizada para provocar situaciones reproducibles de baja alerta.

#### **Ritmos circadianos**

Los humanos vivimos sujetos a un ciclo vigilia-sueño llamado ritmo circadiano [Halberg, 1974; Cinaz et al., 2013; Gil et al., 2013; Peltokangas et al. 2012]. Varias investigaciones [Ferrara y Gennaro, 2001; Horne y Reyner, 1995; Maycock, 1997] han demostrado que dentro de un ciclo circadiano de 24 horas hay dos períodos donde el nivel de somnolencia disminuye.

El primero ocurre durante la noche y en las primeras horas de la mañana y el segundo en la tarde. Durante estos períodos de sueño, el rendimiento y estado de ánimo subjetivo se degradan [Rosekind, 1999]. La influencia del ritmo circadiano sobre los accidentes de tráfico fue demostrada por Pack et al. [1995] quienes analizaron los datos de los accidentes ocurridos en Carolina del Norte y encontraron un máximo principal en su ocurrencia durante la noche y otro a media tarde. Wylie [1996] examinó la pérdida de alerta y la degradación del rendimiento de 80 conductores de camiones de USA y Canadá. El método de estudio incluyó un registro continuado utilizando una cámara de video que monitorizaba la cara de los conductores para observar el comportamiento ocular y la expresión facial. Con este método se encontró que hay picos de alerta reducida tanto en la noche como al amanecer. Este estudio también encontró que la hora del día es el factor más consistente que influye en la fatiga de los conductores. La consecuencia práctica que se deriva de estas evidencias es la necesidad de controlar, en los diseños experimentales, la hora del día a la que se realicen las pruebas, tanto para controlar el efecto de la fase circadiana como para evaluar el efecto de la misma.

En este tema ha de hacerse necesariamente referencia a la organización del trabajo por turnos, tema esencial para minimizar los accidentes laborales debidos a niveles de alerta insuficientes de los operadores. Perfiles de alerta, accidentes y hora del día deberían ser elementos principales en cualquier estrategia de trabajo de turnos. Se ha comprobado por medio de encuestas que los grupos sometidos a turnos de trabajo sufren ataques de sueño durante el trabajo [Akerstedt T., 1994]. Un estudio similar demuestra que más de un 30% de los conductores han reconocido haber dado alguna cabezada al volante alguna vez y más de un tercio de estos trabajadores desarrollan su jornada laboral en turnos de trabajo [Akerstedt T., 1998]. El criterio, generalmente asumido en las organizaciones para la planificación de los turnos de trabajo, de que todas las personas se adaptan de manera similar a estas condiciones es claramente erróneo. En realidad, los horarios de los trabajadores se deberían diseñar teniendo en cuenta las peculiaridades de su comportamiento rítmico propio. Otros estudios relacionados pueden verse también en: [Cinaz et al., 2013; Gil et al., 2013; Peltokangas et al. 2012; Yoshizaki et al. 2013].

## **La carga de trabajo**

Según la definición dada por la Enciclopedia Médica en español, la carga de trabajo es la cantidad de horas de trabajo requeridas para realizar una tarea. Es de sentido común suponer que, al igual que la actividad física sin descanso conduce a una fatiga muscular, una prolongada carga mental sin descanso debe conducir a una fatiga mental que implica una reducción del nivel de alerta. Esto se verificó en un estudio en el que se comprobó que la fatiga se incrementa con esta carga de trabajo y el tiempo transcurrido realizando tareas cotidianas durante el día, e.g. tareas en oficina [Cinaz et al., 2013], o más concretamente tareas de conducción [Dinges y Kribbs, 1991; Hancock y Verwey, 1997; Matthews y Desmond, 2002]. Estos estudios demuestran una correlación entre el número de horas de conducción y diversos indicadores de somnolencia (salida del vehículo de la carretera, inexistencia de marcas de frenado, el conductor no intentó evitar el accidente).

Al realizar estudios de accidentes ocurridos tanto en autopistas como en el interior de la ciudad relacionados con bajos niveles de alerta [Harrison y Horne, 2000; Wilkinson, 1992], por medio de encuestas subjetivas se comprobó que los primeros tienen mayor probabilidad de ocurrencia, debido posiblemente a la menor cantidad de estímulos en la conducción por autopistas, sugiriéndose que este factor actúa como modulador del nivel de alerta.

### **2.2.4. Métodos para la evaluación del nivel de alerta**

La valoración del nivel de alerta en conductores de vehículos ha sido abordada en la literatura de varias maneras, y aunque los métodos utilizados se derivan de los ya mencionados en capítulos anteriores, para su análisis crítico comparativo los hemos agrupado en las categorías siguientes:

- Efectos del nivel de alerta en los parámetros mecánicos del vehículo (velocidad, movimientos del volante, trayectorias del vehículo, etc.).
- Efectos del nivel de alerta en el comportamiento físico del sujeto (posición y movimientos corporales, expresión facial, etc.).

- Efectos del nivel de alerta en el comportamiento de variables fisiológicas (EEG, ECG, EOG, etc.).
- Sistemas mixtos.

### **Efectos del nivel de alerta en los parámetros mecánicos del vehículo**

Numerosos estudios han intentado identificar y cuantificar la correlación entre el nivel de alerta de los conductores y los parámetros del comportamiento mecánico del vehículo. Se ha detectado, durante las fases de somnolencia, una disminución en las variaciones de la velocidad del vehículo [Ranney A.T. et al., 1999; Cnossen F. et al., 2000; Arnedt, J.T. et al., 2001] y un aumento tanto en número como en amplitud en los movimientos del volante [Verwey W.B., y Zaidel D.M., 2000]. La exactitud en el direccionamiento y en el seguimiento del carril se deterioran cuando el conductor está con niveles bajos de alerta [Stein, 1995]. Asimismo, encuestas realizadas a conductores de camiones demuestran que disminuyen los cambios de marchas y se ralentizan los tiempos de reacción como primeros síntomas de somnolencia [Williamson et. al., 1992; Williamson et. al., 2001]. Además, otros estudios o trabajos donde se evalúan parámetros del vehículo aparecen e.g. en [Urquiza et al. 2011].

### **Efectos del nivel de alerta en el comportamiento corporal del sujeto**

Muchos de estos métodos, basados en variables como la posición lateral del conductor [Godthelp H. and Milgram P., 1984; Ranney A.T. et al., 1999; Verwey W.B., y Zaidel D.M., 2000], medida por el ángulo de inclinación del tronco del conductor respecto de la vertical, concluyen que hay una dependencia entre los valores de ese ángulo y los estados de somnolencia. Sin embargo, este tipo de información no permite evaluar el nivel de alerta del sujeto con un tiempo suficientemente holgado como para prevenir accidentes y por tanto puede no ser útil para aplicaciones de conductores de vehículos.

La expresión facial también ha sido utilizada para determinar el estado de alerta por medio del tratamiento de imágenes recogidas por una cámara de vídeo. Este método lo detallaremos en la sección de EOG ya que las dos fuentes de información son procesadas en paralelo para determinar el grado de alerta.

## **Efectos del nivel de alerta en variables fisiológicas**

Las medidas fisiológicas se basan en señales de tipo somático que guardan una estrecha relación con variables de tipo psicológico, de ahí que el estudio de señales tales como la actividad electro-encefalográfica, la actividad electro-oculográfica o la tasa cardiaca permitan inferir ciertas variables de orden psicológico como la alerta atencional, el esfuerzo mental, etc. Una de las ventajas que presenta este tipo de medidas es su objetividad, en el sentido de que el sujeto no sabe en qué aspecto de su comportamiento se le está evaluando, por lo que es difícil que pueda modificarlas a voluntad, a no ser que esté especialmente entrenado para ello. Prácticamente todas las señales psico-fisiológicas producen cambios eléctricos que pueden captarse desde la superficie de la piel mediante la utilización de aparatos que recogen y amplifican estas señales para que puedan ser analizadas posteriormente.

Dentro de las señales fisiológicas, se han revisado las fuentes bibliográficas en las que se utilizan las señales de Electroencefalografía (EEG), Electro-oculografía (EOG), y Electrocardiografía (ECG) para la detección del nivel de alerta.

**EEG:** La relación entre el espectro de potencia del EEG y la alerta humana ha sido estudiada desde 1967 [Matousek M., 1967; Kallaway P. y Maulsby R., 1967; Zhao et al., 2012]. Estas y otras muchas publicaciones han reportado cambios en el espectro de potencias relacionados con la alerta, bajo una amplia variedad de condiciones experimentales. Varios estudios usan, para describir el espectro de la señal de EEG, las bandas de frecuencia tradicionales (delta, theta, alfa y beta) y principalmente la presencia de estallidos de ondas alfa en sujetos sometidos a tareas monótonas [Idogawa, 1989], aunque no se realizó ningún intento por cuantificar el nivel de alerta. En esa investigación se consideró la siguiente distribución de frecuencias del EEG: delta (menores a 3,4 Hz); theta (3,5 a 7,4 Hz); alfa lentas (7,5 a 9,4 Hz); alfa rápidas (9,5 a 11,4 Hz); beta 1(13,5 a 17,4 Hz); beta2 (17,5 a 21,4 Hz) y beta3 (21,5 a 30,4 Hz) y gamma (30,5 a 50 Hz). El tiempo transcurrido desde el inicio de la experimentación hasta el momento de la aparición de estos estallidos de ondas alfa cambió de un sujeto a otro; por ejemplo, en uno de los sujetos apareció a los 13 minutos y en otro a los 60 minutos. Curiosamente se concluye que a pesar de someter a los sujetos a largos tiempos de duración de la tarea monótona, el

tiempo de reacción medido desde la presentación de un dígito en una pantalla y su respuesta, el pulsado de ese número en un teclado, no cambió a lo largo de la experimentación.

También se ha reportado una significativa correlación entre el número de errores cometidos al realizar tareas específicas y el incremento en la potencia espectral en 3,7 Hz y en 14,7 Hz [Tzyy-Ping, et. Al., 1997]; sin embargo la ausencia de una clara descripción de las condiciones experimentales, en donde los sujetos no fueron privados de sueño, hace difícil una asociación fiable entre tasas de error y nivel de alerta.

Se ha sugerido la utilización de la función discriminante aplicada al espectro del EEG a la banda entre 1 y 24 Hz para clasificar patrones de EEG en estados de alerta y somnolencia (US Pat. Appl. 2002/0183644) y en (US Pat. 1998/5813993) se describe la posibilidad de detección de somnolencia con frecuencias superiores a los 30 Hz, en la banda de 80 a 420 Hz.

En 2003, Saroj y Craig sometieron a un grupo de 35 sujetos a pruebas de conducción, a una velocidad de 80 Km/h, y 2 horas de duración. Los sujetos no fueron privados de sueño previamente y se reportó que si bien las variaciones en la banda alfa están asociadas a somnolencia, se encontraron diferencias más grandes en las bandas delta y theta, mucho más fáciles de detectar. Todo ello mediante el promediado de 19 canales de registro; una opción lejana a ser implementada en dispositivos portátiles que requieren un reducido número de canales. En este estudio se analizaron tanto los valores medios como las desviaciones estándar de cada una de las bandas: delta (0 a 4 Hz), theta (4 a 8 Hz), alfa (8 a 13 Hz) y beta (13 a 20 Hz) y se los correlacionó con diferentes situaciones de alerta, desde normal hasta somnolencia. Los autores concluyen anotando que su atención estuvo más dirigida a conseguir un correcto funcionamiento del prototipo, contrastado mediante EOG y una cámara de vídeo, y que se debe realizar una experimentación adicional para validarlo. En este estudio no se menciona ningún análisis de la variabilidad interpersonal y no está muy claro cómo se consiguió que, con solo dos horas de conducción y sin privación de sueño, los sujetos llegarán a tener signos de fatiga o incluso somnolencia; tampoco se contrastaron los resultados con la auto evaluación subjetiva de los sujetos, concluyendo, sin embargo, que este es un asunto muy importante que requiere una investigación adicional.

Un incremento en las bandas alfa y theta, como consecuencia de la privación de sueño fue reportada por Bittner et al., [2002], pero no hacen una clasificación del nivel de alerta en función de este comportamiento, mencionando, además, que la complejidad del EEG y la variabilidad interindividual son los principales problemas en su utilización para la construcción de sistemas de detección de alerta.

**EOG y TV:** El comportamiento ocular y la expresión facial han sido relacionadas con el nivel de alerta de los individuos en conductores de vehículos [Qiang Ji y Xiaojie Yang, 2002], caracterizados y cuantificados con diferentes variables: frecuencia de parpadeos, duración del cierre de párpados o su velocidad y el porcentaje de ojos cerrados en el tiempo (denominado en el trabajo PERCLOSE). En este estudio se llegó a determinar que cuando el PERCLOSE es mayor al 30% o cuando la duración del cierre de párpados es mayor a 0,5 segundos, el sujeto ha alcanzado un estado de somnolencia pero, desafortunadamente, el mayor esfuerzo se dedica a vencer las dificultades tecnológicas y utiliza sujetos bajo condiciones simuladas de somnolencia concluyendo que el sistema debe ser probado en sujetos bajo condiciones de somnolencia real.

Este enfoque tiene la gran ventaja de ser completamente no invasivo ya que ningún dispositivo debe ser colocado en los individuos como es el caso de los sistemas que requieren electrodos y gel conductor sobre la piel. Sin embargo, esta es la razón de sus principales desventajas: es difícil conseguir algoritmos fiables si se considera que el sistema tiene que trabajar mientras el conductor mueve su cabeza y además se requieren cámaras de TV con una alta tasa de cuadros por segundo para hacer frente a la alta velocidad de los parpadeos. Estas dificultades se incrementan cuando el conductor utiliza gafas mientras conduce. Por otro lado, si bien las dificultades tecnológicas se van resolviendo y los costes se van reduciendo, el principal problema al que se tienen que enfrentar este tipo de sistemas es la gran variabilidad interindividual en los parámetros anteriores por lo cual se observan contradicciones en las conclusiones de las diferentes investigaciones publicadas [Grace R. et al., 1998; Hayami et al., 2002; Ito T. et al., 2002; Hamada et. al., 2003] y se hace evidente la utilización de complejas rutinas de personalización. En especial esos dos últimos trabajos son muy reveladores por los resultados completamente diferentes que se obtienen para dos sujetos registrados en fases de alerta normal y de somnolencia reales, por lo que no es posible determinar umbrales

universales que discriminen el estado de alerta y por tanto, es necesaria una personalización del sistema para cada uno de los sujetos. Bittner et al., [2002] después de someter a varios sujetos (no especifica el número) a pruebas de conducción realizadas en autopistas reales tanto en estado de alerta normal como después de 24 y 48 horas de privación de sueño, concluyen que los resultados obtenidos al comparar la frecuencia de parpadeos son especialmente desalentadores y conducen a conclusiones contradictorias. Como mencionamos anteriormente, en la actualidad se puede considerar que las dificultades técnicas y económicas están siendo superadas. Sin embargo, una posible explicación de las razones por las que este tipo de sistemas no ha prosperado podría deberse a dos factores: a) la variabilidad interpersonal y b) conclusiones contradictorias en cuanto a la frecuencia de parpadeos, factor que influye directamente incluso en el PERCLOSE, uno de los más utilizados y considerados más prometedores para determinar el nivel de alerta de los conductores.

**ECG:** Se ha encontrado un decremento en la tasa de latidos por minuto durante conducción nocturna prolongada [Riemersma et al., 1977] y en conductores fatigados [Lal y Craig, 2000]. Sin embargo, en estos estudios se manifiesta que este área de trabajo requiere de una investigación adicional de los cambios producidos en el sistema autónomo antes de emitir cualquier conclusión firme sobre su validez como detector de fatiga, siendo una de las razones que explican la escasa literatura sobre ECG, asociada con este tema, a pesar de ser una señal más fácil de registrar que el EEG. De todas maneras, en 2014 Durantin et al. continuaban utilizando el análisis de frecuencias cercanas al infrarrojo y el HRV para la detección de la sobrecarga mental.

**Sistemas Mixtos:** Muchos estudios publicados consideran simultáneamente varias variables que suministran información acerca del nivel de alerta del individuo:

Así, Zhao et. al., en 2012 utilizaban el EEG y el ECG para el estudio y evaluación de la fatiga en simuladores de conducción. Además, un modelo simple del nivel de alerta en conductores, basado en la monitorización por medio de una cámara de video enfocada al rostro del conductor, fue propuesto por Hattori, et al., 1987, quienes demuestran tres estados en su rendimiento en conducción prolongada: El conductor comienza en un estado 1, caracterizado por un estado de alerta normal en el que no existen signos ni síntomas de

fatiga. En un estado 2, los conductores tienden a estar somnolientos y a disminuir su atención mirando vagamente a un punto específico. En el estado 3, el nivel de consciencia del conductor baja y los parpadeos fueron extremadamente reducidos. Durante esta fase, hubo movimientos de zigzag y a veces se cruzó la línea de carril parcialmente o a veces por completo. En este estado, los niveles de alerta son bajos y es difícil mantener el control del vehículo. La velocidad del vehículo se mantuvo más o menos constante pero hubo a menudo retardos en el cambio de velocidad en respuesta a cambios en el gradiente de la carretera. Este tipo de conducción ha sido referida como hipnosis de la carretera [Williams, 1963].

Bittner et al., [2002] realizaron un estudio del comportamiento del nivel de alerta en conductores con y sin privación de sueño. Abordaron varios métodos:

- a) Utilizando un polígrafo, se registraron las señales fisiológicas de EEG, EOG y ECG pero no se dan detalles de los sitios ni del número de canales utilizados,
- b) Se utilizó un vehículo (Octavia Skoda proporcionado por Volkswagen AG) equipado con varios sensores para medir temperatura, nivel de luz, posición del volante, aceleración,
- c) Dos cámaras de video que registraron el comportamiento facial del conductor y los acontecimientos que sucedieron en frente del vehículo.

La idea fue correlacionar el comportamiento fisiológico del conductor, durante las diferentes fases de alerta, con el comportamiento de la información de video para determinar el nivel de alerta a partir del sistema de TV no invasivo. Si bien en la metodología se menciona la utilización de todos estos métodos, al final no se ofrecen resultados en relación al nivel de alerta obtenido mediante el EEG ni tampoco mediante el ECG, y en cuanto al EOG se manifiesta que los resultados son contradictorios. En cuanto a otros métodos, se presentan algunas gráficas de su variación en el tiempo pero en ninguna de ellas se especifica su correlación con el grado de alerta.

### **Efectos del nivel de alerta en otros escenarios**

Una referencia obligada de este análisis de estado del arte son los clásicos estudios polisomnográficos relacionados con el análisis de las diferentes fases de sueño [Davis et

al., 1938; Santamaría y Chiapa, 1987; Makeig y Inlow, 1993; Vuckovic and Popovic, 2002], para caracterizar desórdenes específicos de sueño y para investigar sus causas. Si bien los estados de alerta y somnolencia están bien documentados, su relación con nuestros objetivos es relativa ya que las condiciones de registro y su propósito de definir los perfiles de las fases de sueño, están lejos de la detección de estados de alerta reducida y condiciones de somnolencia en sujetos que están con sus ojos abiertos y realizando tareas regulares como la conducción de vehículos.

La valoración de la profundidad de anestesia y sedación mediante el EEG es otro campo relacionado que ofrece resultados útiles en términos de las herramientas de procesamiento del EEG para eliminación automática de artefactos y de la valoración del nivel de consciencia, relacionada a nuestros objetivos de valoración de los estados de alerta. La US Pat. Appl. 2004/0193068 ofrece un amplio conjunto de herramientas de procesamiento de EEG que reclaman el derecho de ser válidas para la valoración tanto del nivel de consciencia como para estados de alerta reducida en estados de vigilia de los sujetos.

Otros escenarios donde se han aplicado estas técnicas pueden encontrarse en: [Zhao et al., 2012; Yoshizaki et al., 2013; Shields et al. 2013; Schnell et al. 2013; Gao et al., 2013].

### **Comparación de los métodos anteriores**

Algunas fuentes hacen una comparación entre los métodos anteriormente mencionados que no llegan a un consenso para determinar cuál de estos es el mejor. En 1994 Ueno H. et al. del *Vehicle Research Laboratory* de la *Nissan Research Center*, evalúa comparativamente los métodos de detección de alerta agrupados en las tres mismas categorías que hemos usado nosotros, los procedimientos basados en:

- las señales fisiológicas (EEG, ECG y EOG),
- el comportamiento del sujeto (postura corporal, grado de inclinación de la cabeza, comportamiento de los parpadeos),
- y el comportamiento mecánico del vehículo (velocidad, cambios de marcha, frenado, movimientos del volante).

En cuanto a la exactitud de detección del nivel de alerta, se asigna a los dos primeros la máxima puntuación. En cuanto a su cualidad de constituir sistemas de detección prácticos, coloca al último método como el mejor y al primero como el peor. Horne y Reyner [1995] concluyen que el índice más válido para la determinación de la alerta en un conductor es el que se obtiene a partir del EEG, lo que también ha sido ratificado por los trabajos de otros autores como Lal y Craig [2000].

Varios proyectos financiados por la Unión Europea han tratado estos temas de seguridad de conductores de vehículos (proyectos consecutivos: SAVE, AWAKE, SENSATION y HARKEN). Estos proyectos han integrado una variedad de tecnologías con el objetivo de incrementar la seguridad del conductor en general y han promovido una intensa investigación sobre la valoración del nivel de somnolencia basada en señales fisiológicas así como la búsqueda de criterios bien definidos sobre la transición entre las fases de vigilia y somnolencia.

El proyecto **SAVE** (*“System for Effective Assessment of the Driver State and Vehicle Control in Emergency Situations”*) cuyo objetivo inicial era el desarrollo de un sistema integrado, concluye al final que los sistemas identificados tienen demasiadas limitaciones para su uso real, e.g. la plataforma desarrollada es válida solo para caminos rectos. De todas maneras, no se tiene información sobre los algoritmos de clasificación implementados ni de los resultados experimentales obtenidos.

El proyecto **AWAKE** (*“System for effective Assessment of driver vigilance and Warning According to traffic risk Estimation”*) es la continuación del proyecto anterior SAVE y su objetivo fue desarrollar un sistema no intrusivo y fiable para monitorizar al conductor por medio de una cámara de video, con frecuencia de muestreo de 50 cuadros por segundo, y detectar situaciones de baja alerta en tiempo real. El sistema considera varios parámetros: monitorización del conductor (parpadeos, agarre y posición del volante) y monitorización del medio (seguimiento del carril, sistema de posicionamiento, radar anticollisiones y sensor de punto de mirada). Sus principales limitaciones son, por un lado que solo vale en autopistas, y por otro lado que está muy infuido por factores medioambientales tales como el uso de gafas y en general por condiciones de luz adversas (amanecer y anochecer) debido al tipo de sensores utilizados.

Las conclusiones son las siguientes: **a)** los sensores tienen serias limitaciones (la cámara de video utilizada necesita tener una tasa de muestreo mayor, y para evitar el problema del uso de gafas los sensores deben estar físicamente más cercanos al conductor); **b)** Es imprescindible registrar señales fisiológicas para contrastar los resultados obtenidos por medio de la cámara de video; **c)** la experimentación en este tema es difícil al no disponer de criterios claros para la caracterización de la somnolencia y la falta de *gold standards* para poder comparar la fiabilidad de los nuevos métodos y sistemas.

El proyecto **SENSATION** (*“Advanced Sensor Development For Attention, Stress, Vigilance & Sleep/Wakefulness Monitoring”*) dedica una gran parte de los recursos a la creación de un sistema de monitorización del nivel de alerta basado en tecnología de nuevos micro y nano-sensores de bajo coste y alta eficiencia para monitorizar el estado fisiológico, especialmente la actividad cerebral, incluyendo tanto las fases de vigilia y sueño como su transición.

El proyecto **HARKEN** (*“Heart and respiration in-car embedded nonintrusive sensors”*) ha desarrollado un dispositivo integrado en textiles que mide la actividad cardíaca y la respiración del conductor con el objetivo de avisarle en caso de fatiga y riesgo de somnolencia. Este proyecto es un trabajo conjunto entre diversas empresas, universidades y centros tecnológicos europeos.

A partir de este análisis del estado del arte y considerando los avances recientes en el área de sensores, la microelectrónica y las tecnologías de procesamiento de señales, podemos concluir que las dos aproximaciones principales para determinar el nivel de alerta de los conductores son dos:

- La señales fisiológicas.
- El comportamiento de los parpadeos monitorizados por medio de cámaras de video.

En cuanto a la primera posibilidad, aunque creemos que existen suficientes estudios que soportan su idoneidad, son patentes las lagunas que impiden su aplicación práctica: la gran variabilidad interpersonal de estos métodos lo que, junto con el problema de intrusividad al requerir de electrodos, puede ser la causa de que este tipo de sistemas no hayan prosperado.

En cuanto a la segunda posibilidad, su ventaja de ser totalmente no intrusivo ha atraído la atención de varios grupos de investigación en todo el mundo pero, como ya se anotó en esta misma sección, presenta varias dificultades técnicas y elevados costes que no parecen fácilmente superables a corto plazo. Además, las diferentes fuentes bibliográficas ya anotadas mencionan problemas mucho más graves como los debidos a la gran variabilidad interpersonal y a las conclusiones contradictorias en cuanto a la tasa de parpadeos.

Durante el presente trabajo de investigación se ha profundizado en la utilización de métodos basados en señales fisiológicas, y más concretamente en las señales EEG y ECG. De esta manera se hicieron pruebas con equipos EEG comerciales resultando ser equipos considerablemente más caros y difíciles de utilizar (i.e. necesidad de utilizar geles para un buen contacto, gorros que suponen mayores niveles de intrusión y tiempos de preparación por experimento, así como manejo de señales con niveles eléctricos del orden de microvoltios, mucho más bajos que los necesarios en otras técnicas, e.g. ECG).

Finalmente, tras varias pruebas preliminares llevadas a cabo con equipos ECG comerciales prestados por el departamento de cardiología del hospital de Usánsolo (Vizcaya), a pesar de ser menos preciso que el análisis sobre la señal EEG, se ha optado por la profundización en la utilización de la señal ECG y el análisis de parámetros derivados de la misma basados principalmente en el **HRV (*Heart Rate Variability*)**, cuyo estudio se basa en la detección de la onda “R”, más fácil de ser detectada y tratada (los niveles eléctricos manejados en el ECG son del orden de milivoltios, y solo requiere de tres puntos de contacto en el cuerpo para su registro; además para una buena calidad de la señal solo es necesario utilizar unos electrodos comerciales muy baratos según se explica más adelante con más detalle).

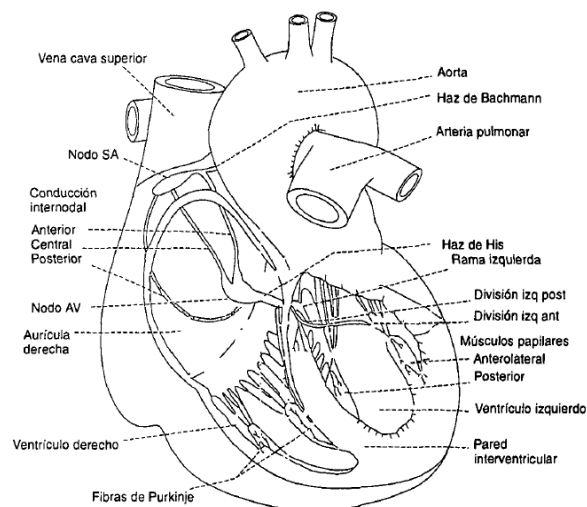
## **2.3. Fundamentos del HRV**

### **2.3.1. Naturaleza del ECG**

El corazón es uno de los órganos necesarios para la sustentación de la vida. Su misión es el bombeo de sangre a todas las partes del cuerpo gracias a la acción combinada de sus cuatro cámaras: aurículas derecha e izquierda y ventrículos derecho e izquierdo. La contracción y

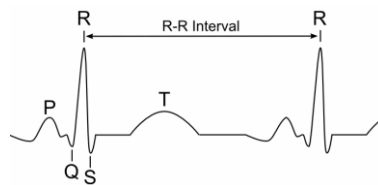
distensión de dichas cámaras se realiza de manera síncrona de forma que se optimiza el gasto cardíaco [Geddes, 1984].

El corazón es un músculo compuesto por células excitables [Webster, 1992] que se contraen al ser estimuladas por sus vecinas. El origen de la excitación está ubicado (en un latido normal) en el nodo sinusal (nodo SA de la Figura 8) que está compuesto por células auto-excitables (o marcapasos) dotadas de una cierta cadencia cuya frecuencia está afectada por diversos sistemas. Una vez iniciada la excitación en el nodo sinusal (cuando se excede un determinado potencial de umbral), las células vecinas se contraen propagando asimismo el llamado potencial de acción. La conducción se realiza a través de todo el músculo y a través de los llamados caminos preferenciales que transmiten el impulso de excitación de forma rápida. De esta forma el estímulo se transmite a las dos aurículas que se contraen y llega hasta el nodo aurículo-ventricular (nodo AV de la figura). Este conjunto de células se caracterizan por tener una conducción muy lenta. Por otro lado, la separación entre aurículas y ventrículos no es conductora salvo por un punto. Por lo tanto, la excitación queda "retenida" en el nodo aurículo-ventricular durante un cierto tiempo (el suficiente para que las aurículas hayan bombeado la sangre a los ventrículos). Pasado este tiempo, la excitación sigue su camino hacia los ventrículos a través del haz de *His* de forma rápida. A partir de aquí la velocidad de la excitación se acelera hasta alcanzar las fibras de Purkinje donde se transmite a todos los puntos de ambos ventrículos causando la contracción de éstos y, por consiguiente, la expulsión de la sangre venosa hacia los pulmones y de la sangre arterial hacia el resto del cuerpo (Figura 8).



**Figura 8: Fisiología del corazón humano”**

El ciclo anteriormente descrito es el que controla la contracción síncrona de las aurículas y ventrículos, el cual puede ser observado desde el electrocardiograma superficial (ECG). En la Figura 9 se muestra un electrocardiograma convencional obtenido mediante una derivación estándar II [Webster, 1992]. Aparecen claramente las ondas constitutivas de todo ECG, a saber: onda P, complejo QRS y onda T. La onda P corresponde a la despolarización (contracción) de las aurículas, el complejo QRS corresponde a la despolarización (contracción) de los ventrículos y la onda T corresponde a la repolarización (distensión) de los ventrículos. La repolarización de las aurículas queda enmascarada por el complejo QRS:



**Figura 9: Complejo “PQRST”**

Observando cuidadosamente los sucesivos latidos en un electrocardiograma superficial se puede asegurar que:

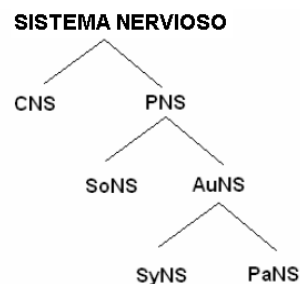
- La morfología del ECG varía latido a latido.
- La separación temporal entre latidos varía.

La segunda observación es la que sirve de base para el estudio de la variabilidad del ritmo cardíaco. Es conocido desde hace mucho [Hales, 1733] que el corazón no se comporta como un oscilador periódico sino que su ritmo está modulado. Por ejemplo, al realizar un ejercicio físico el ritmo cardíaco se acelera y durante períodos de reposo el ritmo se reduce. No tan conocido es el hecho de que el ritmo cardíaco varía latido a latido. Las causas de esta variación son las que se discuten en el apartado siguiente.

### **2.3.2. Causas de la variabilidad del ritmo cardíaco**

El corazón, como todo músculo, está conectado al sistema nervioso. Es éste el que se encarga de modular la respuesta cardiovascular en función de las necesidades del sujeto.

Dicho sistema nervioso se divide en dos estructuras anatómicas diferentes: el sistema nervioso central (CNS), que incluye el cerebro y médula espinal, y el sistema nervioso periférico (PNS) (Figura 10). Este último se suele considerar a su vez el formado por el sistema nervioso somático (SoNS), relacionado con la activación voluntaria de los músculos, y el sistema nervioso autónomo (AuNS) que controla órganos internos y su autonomía. Así mismo el sistema nervioso autónomo está formado por el sistema nervioso simpático (SyNS), encargado de activar el cuerpo ante situaciones de estrés o emergencia, y el parasimpático (PaNS), responsable de mantener las funciones vitales:



**Figura 10: Estructura del sistema nervioso**

Muchos órganos están enervados tanto por el sistema simpático como por el parasimpático. El caso del corazón es uno de ellos y su actividad cardíaca está modulada por las dos ramas del sistema nervioso autónomo: el sistema simpático y el sistema parasimpático. Debido a que la rama parasimpática actúa sobre el corazón a través del nervio vago, en el estudio de la variabilidad del ritmo cardíaco se suelen emplear como sinónimos sistema parasimpático y sistema vago, e.g. [Park et al., 2013].

Una regla generalmente aceptada en el estudio del ritmo cardíaco es que la excitación proveniente del sistema simpático acelera el ritmo cardíaco mientras que la excitación proveniente del sistema parasimpático decelera el ritmo cardíaco. Como ambos sistemas actúan simultáneamente se producen oscilaciones alrededor del ritmo cardíaco medio. Por lo tanto, el origen de la variabilidad del ritmo cardíaco es la interacción entre los sistemas simpático y parasimpático.

El sistema nervioso autónomo recibe información de muchos otros sistemas y actúa sobre diferentes órganos (corazón, sistema digestivo, sistema renal, sistema respiratorio, etc.). No es de extrañar pues que estos sistemas modulen el ritmo cardíaco.

Los sistemas que afectan de forma más patente la variabilidad del ritmo cardíaco son:

- **Sistema respiratorio:** Es bien conocido que la serie RR presenta oscilaciones a la frecuencia respiratoria. Durante la inspiración el ritmo cardíaco aumenta y durante la espiración disminuye. Este fenómeno es conocido como la arritmia sino-respiratoria [Hirsch et al., 1981]. Se conoce que está mediado por activaciones y desactivaciones del sistema vagal. Además, la función de transferencia entre la respiración y la variabilidad del ritmo cardíaco tiene una característica paso-bajo.
- **Sistema vasomotor:** Es otro fenómeno bien conocido que la presión sanguínea y el ritmo cardíaco interactúan entre sí. Los baro-receptores miden la presión sanguínea y esta información es integrada por ambas ramas del sistema nervioso autónomo modulando el ritmo cardíaco. La variación en el ritmo cardíaco afecta a su vez a la presión sanguínea. Por lo tanto, la presión sanguínea y el ritmo cardíaco forman un oscilador cuyo período se sitúa alrededor de los 10s formando, en el caso de presión, las llamadas ondas de Mayer [Ten Harkel et al., 1990].
- **Sistema termorregulador:** Cambios en la diferencia de temperatura entre el sujeto y el ambiente causan variaciones de muy baja frecuencia en la variabilidad del ritmo cardíaco [Kitney, 1975].
- **Sistema renina-angiotensina:** El aumento, descenso o desplazamiento de fluidos corporales así como cambios en la cantidad de sal pueden modificar la variabilidad del ritmo cardíaco. El sistema renina-angiotensina está modulado por estas causas y modula a su vez la rama simpática. El efecto sobre el ritmo cardíaco se presenta como oscilaciones de muy baja frecuencia [Akselrod et al., 1981].
- **Sistema nervioso central:** Es muy conocido que el estado emocional del sujeto modifica el ritmo cardíaco. Hay estudios sobre cómo afecta el estrés [Cerutti et al., 1991; Park et al., 2013] y sobre las diversas etapas de sueño [Somers et al., 1993].

Enumerados ya los diferentes sistemas que modifican la variabilidad del ritmo cardíaco queda patente la utilidad de su estudio: la variabilidad del ritmo cardíaco no sólo aporta información sobre el estado del sistema cardiovascular sino que informa sobre el estado de diferentes sistemas fisiológicos cuya medida directa es forzosamente invasiva.

### 2.3.3. Métodos para la obtención de la serie RR

Se pueden distinguir dos procedimientos diferentes para la obtención de la serie RR en función de la instrumentación empleada:

- Obtención de la serie RR a partir de registros electrocardiográficos ya obtenidos y almacenados sobre cualquier soporte de memoria.
- Obtención directa de la serie RR en tiempo real.

Sea cual sea el procedimiento, para obtener la serie temporal RR se deben realizar los siguientes pasos:

- Amplificación del ECG.
- Detección del complejo QRS.
- Obtención de RR midiendo el tiempo respecto al latido anterior.

Supóngase amplificado ya el ECG (complejos PQRST). Dicha amplificación la habrá realizado un Holter o un electrocardiógrafo en el caso del primer procedimiento o un amplificador realizado *ad hoc* en el caso del segundo procedimiento. El detector de QRS es el punto crítico en la medida de la serie RR y puede ser de dos tipos: detector software y detector hardware.

El detector software es el que se empleará siempre cuando el registro electrocardiográfico esté ya almacenado en memoria y en algunos casos de obtención en tiempo real de la serie RR cuando el electrocardiograma es digitalizado tras ser amplificado. El detector hardware se empleará cuando la única información que se desea guardar en memoria es, precisamente, la serie RR. Ambos tipos de detectores se basan en la misma descripción del complejo QRS para detectarlo: el complejo QRS es una onda de gran amplitud y con transiciones bruscas.

Se han propuesto un gran número de detectores QRS. Entre los que se basan en métodos software se pueden enumerar desde los más sencillos que emplean un umbral fijo que cuando se sobrepasa indica la presencia de un QRS [Restivo et al., 1985] hasta los más sofisticados que emplean una plantilla o modelo y buscan el punto en que se maximiza la

correlación del ECG con la plantilla [Engel et al., 1990], pasando por aquellos que buscan un punto de alta amplitud y derivada [Craelius et al., 1986]. En [Friesen et al., 1990] se comparan diferentes tipos de detectores QRS software y se comentan los diferentes problemas que se encuentran en la detección, a saber:

- Interferencia de red.
- Ruido del contacto de electrodos.
- Artefactos de movimiento.
- Electromiograma.
- Deriva de línea base.
- Modulación de la amplitud del ECG debida a la respiración.
- Ruido asociado a la electrónica de medida.
- Ruido de aparatos de electro-cirugía.

Cuando un detector no identifica un QRS correctamente se acusa en la serie RR la presencia de un artefacto. Si el detector no ha detectado la presencia del QRS en su verdadera posición se dice que se ha producido un falso negativo mientras que si ha detectado una onda (o ruido) que no corresponde a un QRS se dice que se ha producido un falso positivo. A partir de aquí puede darse cualquier combinación de falsos positivos y negativos que pueden falsear completamente la información que se desea medir.

En la Figura 11 se muestran diferentes combinaciones de artefactos obtenidas de un detector de QRS hardware. Se representa en el eje de abscisas el número de latido y en el eje de ordenadas la serie RR expresada en segundos. Un falso positivo se localiza como un valor inusualmente bajo de la serie RR mientras que un falso negativo tiene un valor muy elevado respecto a la media de la serie RR. Un proceso previo al análisis de la serie RR empleado por la mayoría de los investigadores es la corrección de artefactos.

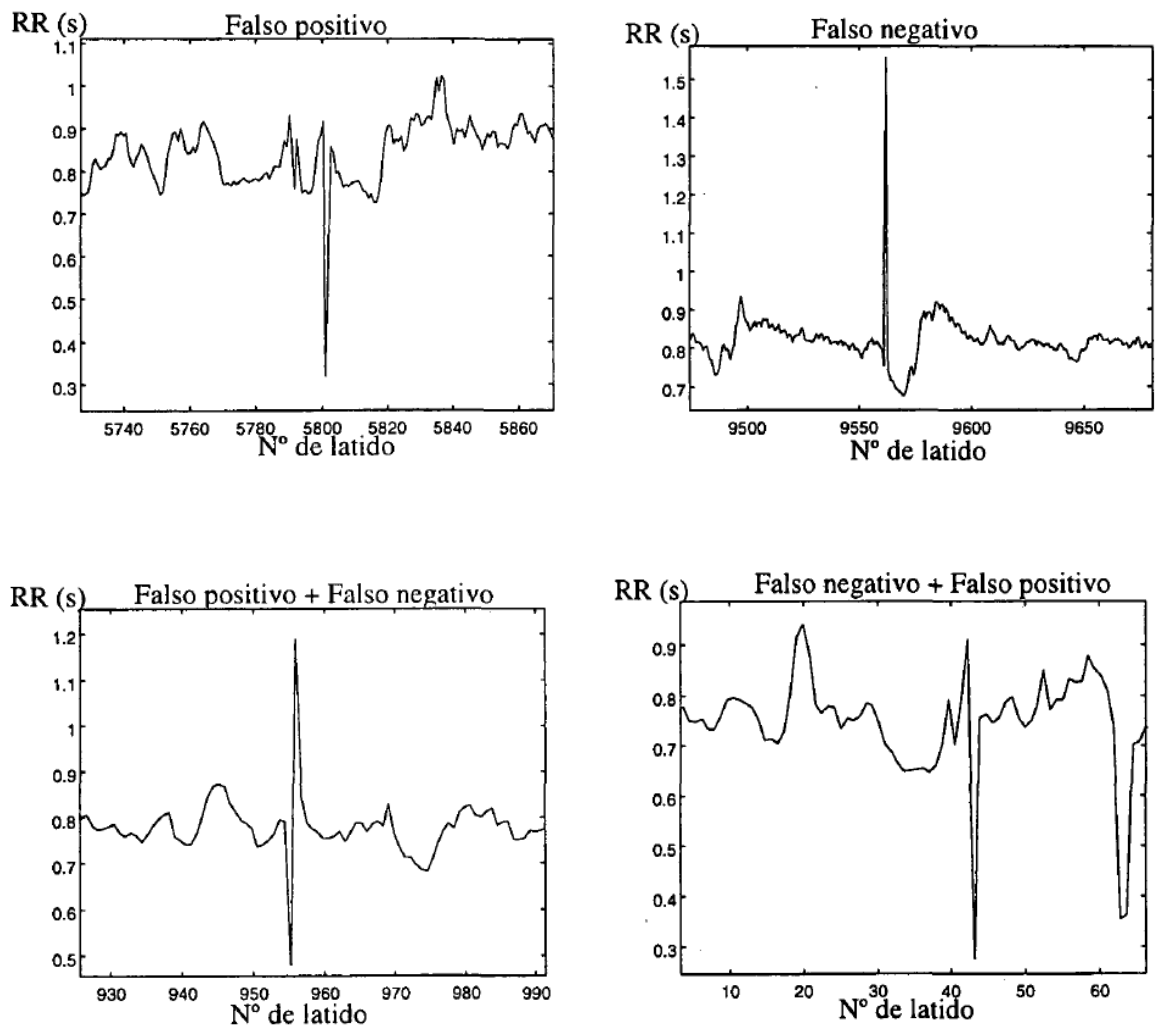


Figura 11: Artefactos en la serie temporal RR (ejemplos)

### 2.3.4. Aplicaciones del análisis HRV

#### Antecedentes médicos

Dado que la variabilidad del ritmo cardíaco refleja la modulación del sistema nervioso autónomo sobre el sistema cardiovascular, no es de extrañar que las principales aplicaciones estén asociadas a testear el estado de este sistema. La variabilidad del ritmo cardíaco se ha propuesto para la predicción de arritmias malignas tras infarto [Kleiger et al., 1987] así como para la predicción de muerte cardíaca súbita [Malik et al., 1989]. Tras un infarto el grado de inervación del corazón disminuye con lo cual se pierde un cierto control sobre este órgano. La variabilidad del ritmo cardíaco refleja esta pérdida de control de forma que se pueden separar grupos de riesgo de muerte cardíaca súbita. Por otro lado,

la propia degeneración del sistema nervioso autónomo se puede medir con un análisis de variabilidad del ritmo cardíaco.

Por lo tanto, se ha aplicado también a la diagnosis de neuropatía, especialmente en diabéticos [Ewing et al., 1981]. La relación entre presión sanguínea y ritmo cardíaco es la explicación que también se pueda diagnosticar la hipertensión [Furlan et al., 1990].

Asimismo, como los cambios posturales afectan a la regulación de la presión sanguínea, un estudio de variabilidad del ritmo cardíaco ayuda al estudio de colapsos y de insuficiencia ortostática aplicado sobre todo en astronautas [Diedrich et al., 1994].

Centrándonos en los efectos del sistema nervioso central sobre el ritmo cardíaco, otros campos de estudio son: estudios de polisomnografía [Somers et al., 1993], estudios de estrés mental [Cerutti et al., 1991; Hoover et al., 2012; Takahiro Y., 2010; Shaun T., 2010; Saus et al., 2012; Murgoitio et al., 2013; Park et al., 2013] y pruebas de esfuerzo [Shin et al., 1993].

### **HRV y esfuerzo mental**

Numerosos estudios han utilizado la frecuencia cardíaca (o su inverso, el intervalo entre latidos) como un indicador de esfuerzo mental y, en última instancia, de atención en diversas tareas. Por ejemplo, Svebak, Dalen y Storfiell [1981] hallaron que se producía un incremento en la tasa cardiaca conforme aumentaba la dificultad de la tarea y McDonald y Cameron [1974] hallaron un descenso en la tasa cardiaca conforme disminuía la dificultad de la tarea. También se ha constatado un descenso en la tasa cardiaca durante la conducción prolongada implicando altos niveles de fatiga [Mackie y O'Hanlon, 1977; Fagerstrom y Lisper, 1977; O'Hanlon y Kelley, 1977]. Otros estudios también han mostrado una relación inversa entre la tasa cardiaca y las horas de conducción [Burns, Baker, Simonson y Kelper, 1966; Lisper, Laurell y Stenig, 1971]. Por tanto, todo parece indicar que la tasa cardiaca aumenta conforme aumenta el esfuerzo o carga mental en la ejecución de una tarea.

Sin embargo, la interpretación de la dirección de los cambios en la tasa cardiaca como una función de las demandas de la tarea y del esfuerzo mental puede resultar complicado, debido a que esta medida se ve influida globalmente por la interacción entre las ramas

simpáticas y parasimpáticas del sistema nervioso autónomo o vegetativo [Sayers, 1973], lo que hace que resulte excesivamente sensible a las influencias autonómicas que continuamente se producen en el organismo [Mascord y Heath, 1992]. Por ello, los estudios suelen derivar a partir de la tasa cardíaca otra medida: la variabilidad de la tasa cardíaca (HRV: *heart rate variability*) o la variabilidad del intervalo entre latidos cardíacos, puesto que es menos sensible a las influencias autonómicas. Diversos estudios han mostrado que la variabilidad del ritmo cardíaco disminuye si se incrementa la carga de trabajo mental [Kalsbeek y Etterna, 1963; Blitz, Hoogstraten y Mulder, 1970; Etterna y Zielhuis, 1971; Luczak y Laurig, 1973; Mulder y Mulder-Hajonides, 1973; Park et al., 2013]. Estas medidas cardíacas se representan en el dominio del tiempo.

Naturalmente, pueden ir apareciendo nuevas aplicaciones sobre el estudio de la variabilidad del ritmo cardíaco debido a la gran cantidad de sistemas por los que es afectada. No obstante, el problema de la aplicación ya está implícito: son tantos los sistemas que la afectan que la serie RR tiene un aspecto pseudo-aleatorio. Por ello, los métodos para la cuantificación de la serie RR son cada vez más sofisticados: [Tarvainen et al., 2014; Murgoitio et al., 2011, 2013; Reyes del Paso et al. 2013; Gil et al., 2013; Huo et al., 2012; Hsu et al., 2012; Xhyheri et al., 2012]

## **2.4. Métodos del análisis HRV**

Existen diversos métodos para el análisis y cuantificación de la serie RR. Cada uno de ellos tiene su aplicación a diferentes disciplinas y tiene una mayor o menor complejidad.

Los más sencillos (y a la vez, los más utilizados) son los métodos **estadísticos**. Éstos se basan en la cuantificación de la serie RR a partir de medidas estadísticas como es la media o la desviación estándar. Su aplicación práctica parte de la observación empírica: a mayor cantidad de variabilidad del ritmo cardíaco se asocia un estado más sano del sistema cardiovascular. Es por ello que los métodos estadísticos se emplean mucho en la prognosis de muerte cardíaca súbita y en la detección de neuropatía ya que el grado de inervación del corazón está relacionado con la cantidad de variabilidad del ritmo cardíaco.

Mientras que los métodos estadísticos son incapaces de separar la cantidad de variabilidad del ritmo cardíaco debida a un cierto sistema, los métodos **espectrales** se aprovechan de la virtud que tienen dichos sistemas de afectar en una banda determinada del espectro de la serie RR. Realizando una separación de bandas se puede obtener información de la modulación de un sistema concreto sobre la variabilidad del ritmo cardíaco.

No obstante, los métodos espectrales sólo son aplicables cuando la señal es estacionaria. Cuando se desean analizar registros largos donde la señal ya es fuertemente no estacionaria se requieren aplicar métodos **espectro-temporales** dentro de los cuales se incluyen las distribuciones tiempo-frecuencia, los métodos espectrales variantes con el tiempo y las ondículas ("wavelets"). Dichos métodos son capaces, además, de seguir la variación de la modulación de los diferentes sistemas.

Mención aparte merecen los métodos que realizan la **caracterización de la serie RR** desde el campo de la dinámica no lineal. Últimamente ha cobrado gran interés el estudio de la complejidad de la señal (caracterización a partir del comportamiento caótico de la señal). Además, si consideramos que la interacción entre sistemas es no lineal, merece la pena realizar un análisis con la ayuda del biespectro para investigar las posibles no linealidades asociadas a la génesis de la variabilidad del ritmo cardíaco.

Por último, a la hora de caracterizar la variabilidad del ritmo cardíaco debe considerarse la información aportada por cada uno de estos métodos. Esta información será redundante en algunos casos pero en otros proporcionará claves para la mejora de la caracterización.

#### **2.4.1. Métodos estadísticos**

Los métodos estadísticos o temporales fueron los que se utilizaron primero en el análisis de la variabilidad del ritmo cardíaco debido, fundamentalmente, al mayor conocimiento de las técnicas estadísticas por parte de los médicos frente a otros métodos (e.g. los espectrales). Las primeras aplicaciones de los métodos estadísticos se realizaron para cuantificar la respuesta a diversas pruebas o maniobras. Se trataba, por lo tanto, de caracterizar la variabilidad del ritmo cardíaco a corto plazo y en situaciones muy concretas. Más adelante se empezó a caracterizar la variabilidad del ritmo cardíaco a partir de electrocardiogramas

de 24 horas de duración (adquiridos mediante sistemas Holter). Así pues, la caracterización de la variabilidad del ritmo cardíaco mediante métodos estadísticos se puede dividir en dos clases: medidas de corta duración y medidas de larga duración.

### **Medidas de corta duración**

Las medidas de corta duración reflejan generalmente la actividad vagal [Kleiger et al., 1993]. Se utilizan muy a menudo, por ejemplo, para la diagnosis de neuropatía en diabéticos y en pruebas de esfuerzo. Por ejemplo, durante una batería de pruebas para la determinación de neuropatía se suele medir la variabilidad del ritmo cardíaco durante respiración profunda, durante la maniobra de Valsalva o la incorporación del paciente.

Una caracterización generalizada de la secuencia RR obtenida durante las pruebas es la correspondiente a la **media** y la **desviación estándar** de los intervalos, índices que también son utilizados, como se verá, en medidas a largo plazo. No obstante, hay otros índices que también han sido muy empleados.

En medidas con respiración profunda se conoce como índice de arritmia sino-respiratoria (**RSA<sub>index</sub>**) definido como [Lopes et al., 1992]:

$$\mathbf{RSA}_{\text{index}} = (\mathbf{RR}_{\text{max}} - \mathbf{RR}_{\text{min}}) / \mathbf{RR}$$

donde **RR<sub>max</sub>** y **RR<sub>min</sub>** son los intervalos RR máximo y mínimo durante una respiración profunda. Obsérvese que dicho índice está normalizado por el valor medio de los intervalos RR. Suele expresarse más comúnmente como porcentaje que como tanto por uno. Otro índice utilizado es la llamada diferencia en respiración profunda (**DBD**) definida como [Smith et al., 1981]:

$$\mathbf{DBD} = \mathbf{RR}_{\text{max}} - \mathbf{RR}_{\text{min}}$$

Este índice no es más que el **RSA<sub>index</sub>** sin normalizar. Un tercer índice que se ha considerado es la relación entre espiración e inspiración (E:I) [Sundkvist et al., 1979]:

$$\mathbf{E:I} = \mathbf{RR}_{\text{max}} / \mathbf{RR}_{\text{min}}$$

Durante la realización de la maniobra de Valsalva aparece una taquicardia, seguida, al finalizar, por una bradicardia. Un índice que se ha propuesto es la relación de Valsalva (Valsalva ratio) [Smith et al., 1981]:

$$\mathbf{Valsalva_{ratio}} = RR_{\max}(\text{bradicardia}) / RR_{\min}(\text{taquicardia})$$

Si se desea hallar el nivel de taquicardia conseguido, se puede utilizar la llamada relación de taquicardia (Tachycardia ratio) [Baldwa et al., 1977]:

$$\mathbf{Tachycardia_{ratio}} = RR_{\min}(\text{taquicardia}) / RR_{\max}(\text{control})$$

donde por control se entiende el conjunto de intervalos RR (entre 10 y 60) medidos antes de la maniobra.

Para cuantificar la secuencia RR durante el proceso de incorporación del paciente se suele utilizar la **Relación 30:15 (30:15 ratio)** [Ewing et al., 1981]:

$$\mathbf{30:15_{ratio}} = RR(30) / RR(15)$$

donde RR(30) es el intervalo RR número 30 tras la incorporación del paciente y RR(15) el número 15.

### **Medidas de larga duración**

En las medidas de larga duración aparece el problema de los artefactos debido a que la obtención de la secuencia RR suele partir de un registro ambulatorio. Este problema queda parcialmente solventado si no se emplean índices donde intervengan el máximo o el mínimo de la secuencia RR que son altamente sensibles a los artefactos. Por ello, los índices utilizados en la caracterización estadística de la variabilidad del ritmo cardíaco de registros de larga duración suelen ser diferentes a los empleados en medidas de corta duración. Veamos algunos de los índices propuestos.

Se conoce como **diferencia noche/día** [Kleiger et al., 1993] a la diferencia entre la media de los intervalos RR normales durante la noche y la que se obtiene durante el día:

$$\text{Diferencia noche/día} = \text{Media}(\text{RR}_{\text{noche}}) - \text{Media}(\text{RR}_{\text{día}})$$

Dicho valor es positivo debido a que la descarga simpática es más acusada en el período diurno.

Un índice muy utilizado es el SDNN, CLV o SDRR [Kleiger et al., 1993] que corresponde, por definición, a la desviación estándar de los RR normales en un registro de 24 horas:

$$SDNN = \sqrt{\frac{\sum_{n=1}^N (\text{RR}(n) - \overline{\text{RR}})^2}{(N-1)}}$$

donde N es el número de intervalos RR normales del registro (se han eliminado los artefactos y latidos ectópicos).

El índice **SDNN<sub>index</sub>** [Kleiger et al., 1993] se define como la media de las desviaciones estándar de intervalos de 5 minutos de duración obtenidos al fragmentar el registro de 24 horas (**SD<sub>5min</sub>**),

$$SDNN_{index} = 1/L \sum_{n=1}^L SD(n)_{5\text{min}}$$

donde L es el número de intervalos de 5 minutos presentes en el registro.

El índice SDANN [Scherer et al., 1994] se define como la desviación estándar de la media de los intervalos de 5 minutos (**RRm<sub>5min</sub>**):

$$SDANN = \sqrt{\frac{\sum_{n=1}^L (\text{RRm}_{5\text{min}}(n) - \overline{\text{RRm}_{5\text{min}}})^2}{(L-1)}}$$

Un problema importante de los índices asociados a la desviación estándar es su gran sensibilidad a la presencia de artefactos. No es de extrañar, pues, que diferentes investigadores se interesaran por índices más robustos a artefactos. Por ello se introdujo el **HRV<sub>index</sub>** [Cripps et al., 1991] como una alternativa al SDNN: es el cociente entre el número total N de intervalos RR en un registro de 24 horas y la frecuencia modal:

$$\text{HRV}_{index} = N / \text{Frecuencia modal}$$

Con el mismo objetivo se puede utilizar la anchura de la base del histograma, que se puede obtener por interpolación triangular del histograma de los RR [Odemuyiwa et al., 1991] o bien utilizando un parámetro que será función del grado de contaminación de la secuencia [Malik et al., 1989]. En la Figura 12 se muestra la anchura (línea roja) de la base del histograma de un registro que contiene 20.000 latidos donde se ha decidido que la anchura viene dada por los RR cuya frecuencia es superior a 25. El valor resultante es de 400 ms.

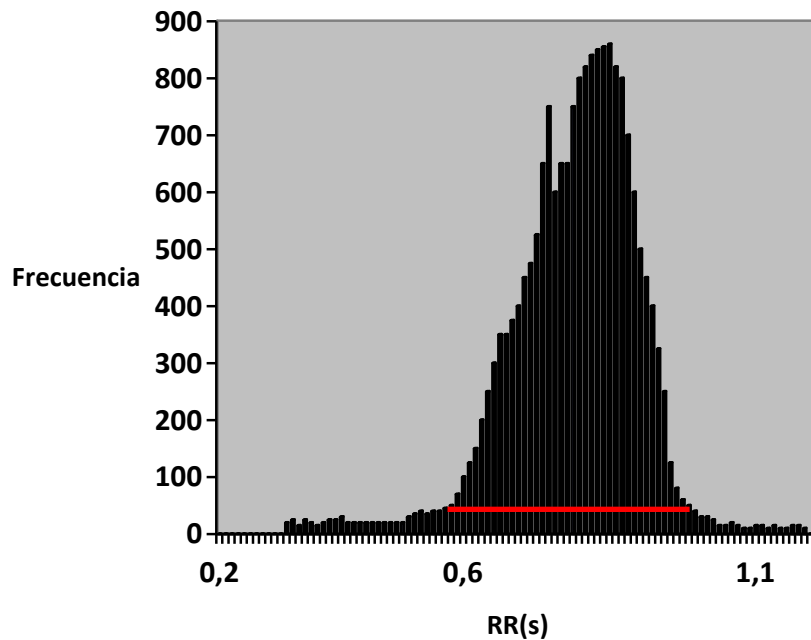


Figura 12: Anchura de la base del histograma (registro de 2000 latidos)

Tanto el  $HRV_{index}$  como la anchura de la base del histograma están relacionados con el SDNN en el sentido de que los tres índices cuantifican la anchura del histograma.

Otra forma de cuantificar la variabilidad del ritmo cardíaco a largo plazo es mediante el llamado **índice de carga** [Tkacz et al., 1993]. Si se define  $M_0$  como la moda de la secuencia RR,  $f(M_0)$  la frecuencia modal,  $N$  el número de intervalos RR de la secuencia y  $\Delta RR = (RR_{max} - RR_{min})$ , entonces el índice de carga (*load index*) es:

$$Load_{index} = f(M_0) / (2 \times M_0 \times \Delta RR \times N)$$

Otros índices propuestos a partir de la secuencia RR han sido las frecuencias relativas de máximos y mínimos relativos [LeBlanc, 1986], donde se caracteriza la variabilidad del ritmo cardíaco de una forma muy diferente a los índices anteriormente citados. La utilidad de estos índices reside en la detección de arritmias y bloqueos. Dado que en el análisis de la variabilidad del ritmo cardíaco se descartan aquellos intervalos que son anormales, estos índices han caído en desuso.

Un grupo de índices que se han utilizado últimamente con cierta asiduidad son los que cuantifican las diferencias absolutas sucesivas de los intervalos RR. Dichos índices están relacionados con la actividad vagal. Uno de los propuestos es el **pNN50** [Kleiger et al., 1993] definido como el cociente entre el número de veces que la secuencia de diferencias absolutas sucesivas de los intervalos RR de un registro de 24 horas supera el umbral de 50 ms y el número total de intervalos RR. Formalmente se puede expresar como:

$$pNN50 = \sum_{n=2}^{n=N} V(n) / N \quad \text{donde} \quad \begin{cases} V(n) = 1 \Leftrightarrow |RR(n) - RR(n-1)| \geq 50ms \\ V(n) = 0 \Leftrightarrow |RR(n) - RR(n-1)| < 50ms \end{cases}$$

En la misma línea está el **r-MSSD** o **SDSD** [Kleiger et al., 1993] que es la desviación estándar de las diferencias absolutas sucesivas de los intervalos RR:

$$r - MSSD = \sqrt{\frac{\sum_{n=1}^{N-1} (DARR(n) - \overline{DARR})^2}{(N-2)}}$$

donde DARR indica  $|RR(n+1)-RR(n)|$  y N es el número total de intervalos RR normales.

Ante la abundancia de índices para caracterizar la variabilidad del ritmo cardíaco, cabe preguntarse si tal cantidad es realmente necesaria. Se observa que para pruebas de corta y larga duración muchos de los índices aportan esencialmente la misma información. Hay pocos estudios de correlación entre índices. Respecto a medidas de corta duración no hemos encontrado ningún estudio, mientras que en [Kleiger et al., 1993] se recoge información sobre la relación de diversos índices de caracterización de registros de 24 horas. Sus resultados más destacables se recogen en la Tabla 3:

ÍNDICE 1	ÍNDICE2	CORRELACIÓN
r-MSSD	pNN50	0,96
SDNN <sub>index</sub>	pNN50	0,94
SDNN	Diferencia noche-día	0,78
SDNN	Media RR (noche)	0,92
SDNN	Media RR (día)	0,66

**Tabla 3: Correlación entre diversos índices empleados para caracterizar la HRV en medidas de 24 horas [Kleiger et al., 1993]**

Obsérvese la alta correlación entre indicadores de actividad vagal y la gran relación entre SDNN y la media del ritmo cardíaco en períodos nocturnos. Todo esto hace suponer que con dos índices se tiene gran parte de la información que aporta el conjunto total de índices. Uno de estos índices está relacionado con la anchura del histograma de los RR (SDNN, HRV<sub>index</sub>, etc.) y otro con las diferencias absolutas de los RR (pNN50, r-MSSD, etc.). El primero es una medida absoluta de la HRV y está influenciado por el sistema simpático y el parasimpático mientras que el segundo refleja el nivel de actividad vagal.

La elección entre uno u otro índice de entre los propuestos debe basarse en diferentes criterios como son:

- **Robustez del índice a la presencia de artefactos.** La robustez frente a artefactos es necesaria siempre que además se desee prescindir del complicado proceso de corrección de artefactos. Es un requisito imprescindible si se quiere realizar una monitorización directa de índices.
- **Nivel de información aportado en casos anormales.** Un índice que sólo aporte información a partir de un valor dicotómico pierde gran parte de su utilidad. Debe ser capaz de indicar si el sujeto bajo test presenta la anomalía y en qué grado de desarrollo está.
- **Independencia del sistema de medida.** Un índice que dependa del sistema de medida es poco general y no aporta información si no se acompaña con la documentación necesaria. La frecuencia de muestreo del ECG es un factor del que pueden depender diversos índices.
- **Rapidez computacional.** Éste suele ser un factor determinante en la elección de un índice u otro. Las limitaciones de tiempo suelen ser más importantes en el

análisis de registros largos donde el volumen de información es mucho más elevado.

La Tabla 4 es un resumen de los índices estadísticos considerados anteriormente:

ÍNDICE	Corta duración (maniobras)			Larga duración
	Maniobra Valsalva	Respiración profunda	Incorporación	
Media(RR)	$1/N \sum_{n=1}^N RR(n)$	$1/N \sum_{n=1}^N RR(n)$	$1/N \sum_{n=1}^N RR(n)$	$1/N \sum_{n=1}^N RR(n)$
SD	$\sqrt{\frac{\sum_{n=1}^N (RR(n) - \overline{RR})^2}{(N-1)}}$	$\sqrt{\frac{\sum_{n=1}^N (RR(n) - \overline{RR})^2}{(N-1)}}$	$\sqrt{\frac{\sum_{n=1}^N (RR(n) - \overline{RR})^2}{(N-1)}}$	$\sqrt{\frac{\sum_{n=1}^N (RR(n) - \overline{RR})^2}{(N-1)}}$
RSA <sub>index</sub>		$\frac{RR_{\max} - RR_{\min}}{RR}$		
DBD		$RR_{\max} - RR_{\min}$		
E:I		$RR_{\max} / RR_{\min}$		
Valsalva <sub>ratio</sub>	$\frac{RR_{\max} (bradicardia)}{RR_{\min} (taquicardia)}$			
Tachycardia <sub>ratio</sub>	$\frac{RR_{\min} (taquicardia)}{RR_{\max} (control)}$			
30:10 <sub>ratio</sub>			RR(30) / RR(10)	
Diferencia noche/día				$\overline{RR}_{noche} - \overline{RR}_{dia}$
SDNN				$\sqrt{\frac{\sum_{n=1}^N (RR(n) - \overline{RR})^2}{(N-1)}}$
SDNN <sub>index</sub>				$1/L \sum_{n=1}^L SD(n)_{5\min}$
SDANN				$\sqrt{\frac{\sum_{n=1}^L (RRm_{5\min}(n) - \overline{RRm_{5\min}})^2}{L-1}}$
HRV <sub>index</sub>				N/Frecuenc. modal
Load <sub>index</sub>				$\frac{f(M_o)}{2 \times M_o \times \Delta RR \times N}$
pNN50				$\sum_{n=2}^{n=N} V(n)$
r-MSSD				$\sqrt{\sum_{n=1}^{N-1} (DARR(n) - \overline{DARR})^2}$

Tabla 4: Resumen de índices estadísticos convencionales

Los índices estadísticos descritos hasta ahora (convencionales) presentan algunas limitaciones. Por ejemplo, la desviación estándar tiene alta correlación con el resto de índices propuestos para cuantificar la variabilidad del ritmo cardíaco durante la realización de maniobras. Por tanto, se sugiere la utilización de SD para caracterizar estadísticamente la secuencia RR.

Por otro lado, la dependencia del  $HRV_{index}$  respecto a la frecuencia de muestreo. Una solución posible es la normalización de esta frecuencia o lo que es lo mismo: para calcular el  $HRV_{index}$  se recomienda realizar el histograma a partir del redondeo de los valores de la serie RR (expresada en segundos) al segundo decimal. Esto equivale a una frecuencia de muestreo de 100 Hz (resolución temporal de 10 ms) que se ha considerado como un límite inferior en la adquisición de electrocardiogramas para la detección de ondas R.

Otros problemas importantes son:

- Sensibilidad de la desviación estándar (SD) frente a artefactos.
- Falta de información de la cantidad de variabilidad latido a latido en pNN50.

Para mejorar estos problemas proponemos la sustitución de estos índices convencionales por otros nuevos que resuelvan las limitaciones sin aumentar en exceso la complejidad de cálculo.

### **Índices mejorados: MIRR**

La desviación estándar (SD) es un índice estadístico que cuantifica la anchura del histograma. La presencia de artefactos hace que la distribución de la serie RR se ensanche formando colas largas y aumentando la asimetría del histograma. Por ello, la estimación de la desviación estándar en presencia de artefactos es muy sensible a éstos. Una solución posible consiste en ignorar la información contenida en las colas. Cuanto más nos limitemos a las proximidades de la media del histograma, mayor robustez tendremos frente a artefactos; pero, a la vez, a mayor proximidad con la media, mayor es la información perdida, disminuyendo así la correlación con la SD. Debemos optar por una solución de compromiso.

Definimos el índice MIRR como el margen intercuartil de la secuencia RR. Si  $Q_1$  es el cuartil de menor valor de los RR y  $Q_3$  es el de mayor valor ( $Q_2$  es la mediana), podemos expresar el índice MIRR como:

$$\text{MIRR} = Q_3 - Q_1$$

*A priori*, el MIRR debe correlacionar bien con la SD. Si se supone que la distribución de los RR es gaussiana se cumple [Croxtton, 1959] que:

$$\text{MIRR} = 1,35 * \text{SD}$$

Esto implica una relación lineal entre ambos. Obsérvese que el MIRR es otra medida de la anchura del histograma.

Para analizar la robustez de este índice frente a artefactos, llamamos secuencia SRR a la ordenación de menor a mayor de los N intervalos RR que deseamos cuantificar. Así pues,  $\text{SRR}(1) = \text{mínimo} \{ \text{RR}(n) \mid n \in (1, N) \}$  y  $\text{SRR}(N) = \text{máximo} \{ \text{RR}(n) \mid n \in (1, N) \}$ . La señal SRR es escalonada y creciente. En la Figura 13 se observa un detalle de la señal SRR para un registro de corta duración:

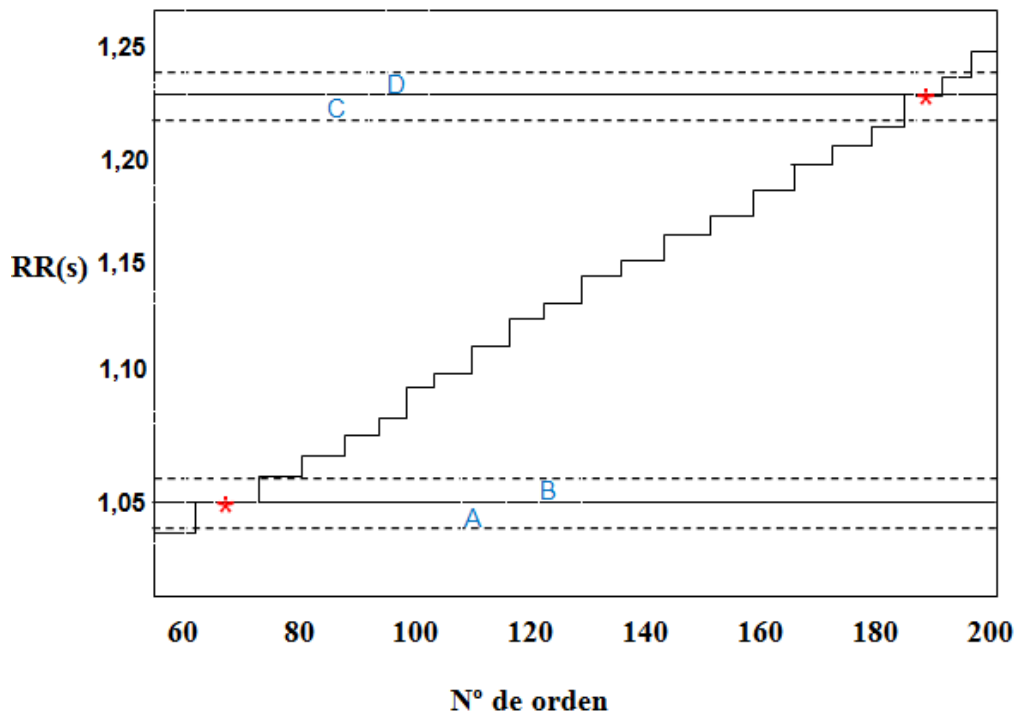


Figura 13: Detalle de secuencia RR en registro de corta duración

Con un asterisco se indica la posición en el eje de ordenadas de  $Q_1$  y  $Q_3$ . Las letras A, B, C y D indican los posibles cambios en la estimación de los cuartiles ante la presencia de un falso positivo o negativo. Estos cambios son iguales (en la mayoría de los casos) a la resolución temporal de la secuencia RR.

El error en el peor caso viene dado por:  $E_{\text{peor-caso}} = \text{Máximo}\{A+D, B+C\}$

que, a efectos prácticos, corresponde a  $2/f_s$ , siendo  $f_s$  la frecuencia de muestreo. Se debe remarcar que en la mayoría de casos este error será nulo y, aún en el peor caso, el MIRR es mucho más robusto que la SD frente a artefactos.

El cálculo del MIRR requiere los siguientes pasos:

- a) Calcular la secuencia SRR.
- b) Hallar  $Q_1 = \text{SRR}(N/4)$  y  $Q_3 = \text{SRR}(3N/4)$ .
- c) Realizar la diferencia  $\text{MIRR} = Q_3 - Q_1$ .

Obsérvese que la rapidez computacional para el cálculo de este índice es elevada. El paso a) implica una ordenación fácilmente realizable con rutinas clásicas como el *Quick-Sort* o el método de la burbuja. El paso b) es la identificación de dos posiciones en el vector donde está almacenada la secuencia SRR. El paso c) se reduce a una simple diferencia.

### Índices mejorados: MDARR

Se ha visto que el pNN50 es un índice muy robusto a artefactos pero que no informa de forma directa de la variabilidad latido a latido. La elección del umbral de 50 ms hace que si no hay en ningún momento un cambio superior a éste el índice valga 0. Por otro lado, si la variabilidad latido a latido es muy elevada, el pNN50 tiende a valer 1. A este efecto lo llamaremos saturación del índice. Por lo tanto, se debe buscar un índice que no presente este problema. Una posible solución podría ser el índice r-MSSD pero es muy sensible a artefactos. Sea DARR la secuencia obtenida de la siguiente forma:

$$\mathbf{DARR(n)} = |\text{RR}(n+1) - \text{RR}(n)| \quad \forall n \in \{1, N-1\}$$

La presencia de un artefacto en la secuencia RR provoca dos artefactos en la secuencia DARR. Es por ello que el empleo del r-MSSR es desaconsejable.

Definimos el índice MDARR como la mediana de las diferencias absolutas de los RR, es decir, como la mediana de la secuencia DARR. Igual que el índice MIRR, está basado en la robustez de los cuartiles frente a artefactos. De hecho, por definición, es el valor del segundo cuartil de la secuencia DARR. Por un razonamiento similar al realizado con el MIRR, el error en el peor caso será  $1/f_s$  aunque la mayoría de veces dicho error será nulo.

Es de prever una alta correlación entre el MDARR y el pNN50 ya que el MDARR es el valor para el cual la mitad de variabilidades latido a latido del registro están por encima de él y la otra mitad por debajo. Si MDARR es igual a 50 ms, entonces se tendrá que  $pNN50 = 0,5$ . La ventaja que presenta es que mientras pNN50 vale cero, el MDARR da un valor pequeño pero diferente de cero (siempre y cuando la variabilidad latido a latido sea superior a la frecuencia de muestreo) lo cual permite discriminar más entre sujetos con anormalidades y aporta más información. El extremo contrario también es cierto. Cuando el pNN50 vale 1, el MDARR dará un valor elevado pero no presentará el efecto de saturación que presenta el pNN50 para variabilidades altas y bajas.

El cálculo de MDARR requiere los siguientes pasos:

- a) Obtener la secuencia DARR definida según la fórmula anterior.
- b) Transformar dicha secuencia en SDARR realizando una ordenación de menor a mayor.
- c) Calcular  $MDARR = SDARR((N-1)/2)$ . Éste es el segundo cuartil de DARR.

Nuevamente, este índice se puede calcular de forma sencilla y rápida. El paso a) implica un bucle en el cual se restan dos números y se prescinde del signo. El paso b) es idéntico al paso a) del índice MIRR. El paso c) es la identificación del valor situado en una posición determinada del vector donde se almacena DARR.

## Índices mejorados: SD1 y SD2 (*Plot de Poincaré*)

El *plot* de Poincaré es un mapa de puntos en coordenadas cartesianas [Huo et al., 2012; Hsu et al., 2012; Xhyheri et al., 2012], construido a partir de los valores de los intervalos RR obtenidos, donde cada punto es representado en el eje X (horizontal/abscisa) por el intervalo RR normal precedente y, en el eje Y (vertical/ordenada), por el intervalo RR siguiente. Para el análisis cuantitativo del *plot* los puntos del gráfico son ajustados a una elipse con centro determinado por la media de los intervalos RR, y son calculados los índices SD1, que mide el desvío-estándar de las distancias de los puntos a la diagonal  $y = x$ , y SD2 que mide el desvío-estándar de las distancias de los puntos a la recta  $y = -x + RR_m$ , donde  $RR_m$  es la media de los intervalos RR. El SD1 es un índice de registro instantáneo de la variabilidad latido-a-latido y representa la actividad parasimpática, mientras que el índice SD2 representa el HRV, en registros de larga duración, y refleja la variabilidad global. La relación de ambos (SD1/SD2) muestra la razón entre las variaciones corta y larga de los intervalos RR. Estas dos medidas y su relación son parámetros que condensan toda la información de la actividad simpático-vagal cardíaca. Además de eso, el análisis del *plot* de Poincaré es considerado por algunos autores como **contenedor de la dinámica no lineal** (Figura 14).

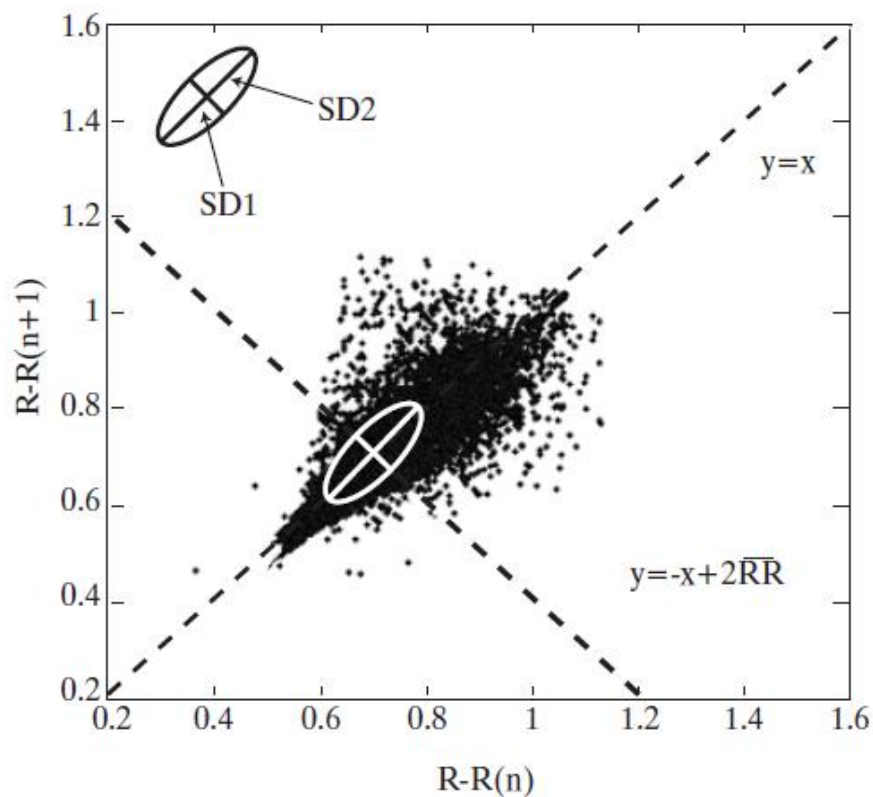


Figura 14: Plot de Poincaré: Significado gráfico de SD1 y SD2

## Conclusiones

En este capítulo se ha presentado en primer lugar el conjunto de índices convencionales más comúnmente empleados para caracterizar la variabilidad del ritmo cardíaco. Se ha observado la gran disparidad de definiciones que es consecuencia de una falta de estandarización. A grandes trazos, los índices estadísticos se pueden clasificar en dos categorías en función de la variable que se desea cuantificar, a saber:

- índices que estiman la anchura del histograma de la secuencia RR.
- índices que estiman la anchura del histograma de las diferencias absolutas de la secuencia RR.

Mientras los primeros reflejan la variabilidad del ritmo cardíaco debida a todas las causas que lo modulan, los segundos son más selectivos y dan información del nivel de excitación vagal producido por reflejos rápidos (por ejemplo, mediados por los baroreceptores). No obstante, estos últimos índices no evalúan completamente el nivel de excitación vagal.

Otra posible clasificación de los índices es en función de la duración del registro. Podemos hablar de índices para maniobras o de índices para registros de larga duración. Mientras que en los primeros hay un claro predominio del empleo de los máximos y mínimos de la secuencia, los segundos emplean herramientas más robustas a artefactos aunque son susceptibles de ser mejorados.

Se ha observado que la desviación estándar tiene buena correlación con la mayoría de índices utilizados para la cuantificación de la variabilidad del ritmo cardíaco en maniobras. Se podría proponer, por lo tanto, como índice estándar para la estimación de la anchura del histograma de la secuencia RR ya que es un índice de uso muy extendido en muchas y diversas disciplinas. No obstante, la desviación estándar es un índice muy sensible a la presencia de artefactos. Se han propuesto en diversos trabajos de la bibliografía diversos índices robustos a artefactos para estimar la anchura del histograma de la secuencia RR pero el grado de correlación con la desviación estándar en ausencia de artefactos es reducido y su nivel predictivo o de diagnóstico es inferior. Por lo tanto, interesa tener un índice bien correlado con la desviación estándar en ausencia de artefactos y, por otro lado, que sea robusto frente a artefactos. Se propone para ello el índice **MIRR** (margen

intercuartil de la secuencia RR). El MIRR es un índice de cálculo rápido, muy robusto frente a artefactos y con alta correlación con la desviación estándar.

Otro índice muy utilizado para caracterizar la anchura del histograma de la secuencia RR es el **HRV<sub>index</sub>**. Se definió para intentar eliminar el efecto de la presencia de artefactos en el registro. No obstante, dicho índice es función de la frecuencia de muestreo del ECG y no tiene una correlación alta con la desviación estándar. Como solución al primer problema se sugiere normalizar la secuencia RR a una resolución temporal de 10 ms.

Uno de los índices más utilizados para estimar la anchura del histograma de las diferencias absolutas de la secuencia RR es el índice **pNN50**. Dicho índice es muy robusto frente a artefactos. No obstante, presenta problemas de saturación para bajas y altas variabilidades latido a latido del ritmo cardíaco. Para solucionar este problema se propone el índice **MDARR** (mediana de las diferencias absolutas del índice RR). El MDARR es un índice de cálculo rápido, muy robusto frente a artefactos, con alta correlación con el pNN50 y que no presenta limitaciones de saturación.

Por lo tanto, si deseamos cuantificar la variabilidad del ritmo cardíaco debido a todas las causas que lo modulan utilizaremos el índice MIRR. Si deseamos estimar el nivel de excitación vagal utilizaremos el índice MDARR.

Finalmente, el análisis del *plot* de Poincaré [Huo et al., 2012; Hsu et al., 2012; Xhyheri et al., 2012] es considerado por algunos autores como un contenedor muy completo de la dinámica no lineal (información compleja y cualitativa) de la actividad cardíaca, donde SD1 es un índice de registro instantáneo de la variabilidad latido-a-latido y representa la actividad parasimpática, mientras que el índice SD2 representa el HRV, en registros de larga duración, y refleja la variabilidad global. Estas dos medidas y su relación son parámetros que condensan toda la información de la actividad simpático-vagal cardíaca.

#### **2.4.2. Métodos espectrales**

Los métodos espectrales han comenzado a ser utilizados últimamente debido a su capacidad de separar la variabilidad del ritmo cardíaco debida a la excitación vagal de la producida por la excitación simpática. En los métodos estadísticos sólo había una somera

aproximación a este objetivo cuando se cuantificaba la variabilidad del ritmo cardíaco latido a latido (con el índice pNN50 o el MDARR, por ejemplo). La utilidad de separar las dos fuentes que provocan la variabilidad del ritmo cardíaco es mejorar su capacidad de análisis y diagnóstico. No obstante, los métodos espectrales presentan varias limitaciones que no siempre son tenidas en cuenta por el investigador. Si no se realiza un correcto preprocesado de la señal y los métodos de estimación espectral no son los correctos, se pueden obtener resultados aberrantes que pueden conducir a diagnósticos incorrectos.

El análisis espectral tiene como objetivo caracterizar el contenido frecuencial de una señal a partir de métodos de procesado de señal. Es útil para la detección de variaciones periódicas y se ha empleado exhaustivamente en el procesado de series temporales. Se sabe que en la variabilidad del ritmo cardíaco aparecen variaciones casi periódicas a distintas frecuencias y que son debidas a diferentes sistemas fisiológicos. El origen de estas oscilaciones está influido por los sistemas simpático y parasimpático. Se sabe que las oscilaciones de alta frecuencia sólo se generan debido a la excitación vagal, mientras que las oscilaciones de baja frecuencia están causadas por las dos ramas del sistema nervioso autónomo.

En [DeBoer et al., 1985] y en otras muchas referencias bibliográficas se comenta que aquellas oscilaciones periódicas que aparecen a frecuencias entre 0,2 Hz y 0,35 Hz son las asociadas a la arritmia sino-respiratoria. En las frecuencias próximas a 0,1 Hz aparecen las oscilaciones conocidas como ondas de Mayer que están relacionadas con la regulación del sistema vasomotor. Estas oscilaciones son debidas a la interacción entre la variabilidad del ritmo cardíaco y la variabilidad de la presión sanguínea. A frecuencias menores de 0,05 Hz pueden aparecer otras oscilaciones debidas a termorregulación y al sistema renina-angiotensina.

Para cuantificar el efecto de cada una de las fuentes se debe realizar, *a priori*, una separación del dominio frecuencial en bandas. No hay un criterio formal sobre cuáles deben ser los límites de cada banda ya que éstos deben ser necesariamente flexibles dependiendo de la aplicación. Por ejemplo, normalmente el límite entre la banda de baja y alta frecuencia se sitúa en 0,15 Hz, pero si se trata de analizar un registro en el cual el

sujeto ha estado respirando a una frecuencia próxima a 0,15 Hz, dicho límite se deberá bajar a 0,12 Hz.

Para registros de más de 5 minutos se suele separar el dominio frecuencial en tres bandas. Una elección bastante común puede ser la siguiente:

- Banda de muy baja frecuencia (VLF) para frecuencias inferiores a 0,04 Hz.
- Banda de baja frecuencia (LF) para frecuencias entre 0,04 Hz y 0,15 Hz.
- Banda de alta frecuencia (HF) para frecuencias entre 0,15 Hz y 0,4 Hz.

La actividad en la banda HF es debida a excitación vagal mientras que la asociada a las bandas de VLF y LF es debida a la interacción entre los sistemas simpático y parasimpático. La separación del espectro en bandas se puede emplear para cuantificar la potencia presente en cada una de ellas o para dar la relación de potencia entre dos bandas. Este hecho explica que el análisis espectral sea tan atractivo. Si un sujeto que respira de forma normal tiene una potencia en la banda HF excesivamente pequeña se puede sospechar un inicio de neuropatía autónoma. De hecho, en las etapas iniciales de neuropatía autónoma el sistema parasimpático es el primero en verse afectado. Si la potencia en la banda LF es elevada, se puede suponer una excesiva preponderancia del sistema simpático lo cual es un indicador de riesgo de taquicardias y arritmias.

### **Muestreo no uniforme**

La secuencia RR se define como el tiempo entre complejos QRS sucesivos. La forma más usual de expresarla es en forma de tacograma. Si definimos como  $R(n)$  el instante en el cual aparece el complejo QRS  $n$ -ésimo, entonces el tacograma queda definido como una secuencia discreta en la cual  $RR(n)=R(n+1)-R(n)$ . Esta es la misma secuencia RR que hemos utilizado en el capítulo anterior. No obstante, si consideramos la secuencia RR como una señal que refleja la actividad del sistema nervioso autónomo, se debe tener en cuenta que el muestreo de dicha actividad no es uniforme.

La gran mayoría de métodos espectrales parten de señales muestreadas uniformemente. Cabe preguntarse si el empleo del tacograma provocará falsas estimaciones espectrales o errores aberrantes. Es decir:

- ¿Se incurre en un gran error si se hace una estimación espectral del tacograma? y
- ¿Cuál es el preprocesado necesario de la secuencia RR para realizar una correcta estimación espectral en caso que la respuesta a la pregunta anterior sea afirmativa?

Las soluciones propuestas en la bibliografía son varias. Una respuesta satisfactoria a la primera pregunta se recoge en [Lisenby et al, 1977]. En dicha referencia se utiliza la FFT para realizar la estimación espectral del tacograma tras ser enventanado. La gran aportación de este artículo es hablar del dominio del latido (*beatquency domain*) en lugar del dominio frecuencial. Es decir, la frecuencia no se mide en ciclos por segundo sino en ciclos por latido. Por lo tanto, el tacograma no necesita ser preprocesado puesto que el reloj es el propio corazón y en este dominio el tacograma es una señal muestreada uniformemente. No obstante, muchas de las causas que provocan la variabilidad del ritmo cardíaco no dependen de éste. A veces es necesario cambiar de un dominio a otro. Un ejemplo claro es el cálculo de la frecuencia asociada a la arritmia sino-respiratoria (RSA) en casos de respiración periódica. En este caso, la frecuencia de la RSA deberá ser expresada en hertzios para comparar con la frecuencia a la cual ha respirado el sujeto. En [Baselli et al., 1986] se explica cómo realizar el cambio entre dominios. La idea es que siempre y cuando la variabilidad del ritmo cardíaco no sea excesivamente elevada, el cambio de la frecuencia en el dominio del latido ( $f_{c/l}$ ) a frecuencia en el dominio frecuencial ( $f_{c/s}$ ) se realiza mediante:

$$f_{c/s} = f_{c/l} / \overline{RR}$$

Es decir, se normaliza la frecuencia por la media de la secuencia RR.

Otros investigadores se decantan por aplicar un preprocesado a la señal RR para realizar la estimación espectral independientemente del nivel de variabilidad del ritmo cardíaco. [Berger et al., 1986] comparan diversos métodos de preprocesado de la señal RR. Comentan que el espectro del tacograma (o espectro de intervalos) presenta armónicos debido al muestreo no uniforme. Lo mismo ocurre con el espectro del inverso del tacograma (espectro del ritmo cardíaco instantáneo). Otro método que se describe en la misma referencia es el espectro de las cuentas. Dicha estimación se realiza a partir de una señal que ya está muestreada uniformemente. La señal consiste en un tren de deltas situadas

en los instantes  $R(n)$ . Dicho espectro presenta espurios. La comparación entre los diversos espectros se realiza mediante simulación con el modelo IPFM (*Integral Pulse Frequency Modulation*). No obstante, [DeBoer et al., 1984] observaron que la diferencia entre el espectro del tacograma y el espectro de las cuentas es pequeña.

Una técnica muy utilizada para re-muestrear la señal es el re-muestreo por interpolación polinómica. Hay trabajos publicados donde se emplea interpolación de orden 0 [Shin et al., 1989], de orden 1 [Grönlund et al., 1989] o de orden 3 [Birkett et al., 1991]. Los problemas de espurios son mínimos en el caso de interpolación de orden 3 con lo cual, si el re-muestreo es necesario, optaremos aquí por la interpolación cúbica.

A pesar de este abanico de posibilidades para realizar la estimación espectral de la variabilidad del ritmo cardíaco, no hay en la bibliografía ninguna norma sobre cuál de estos métodos escoger. [DeBoer et al., 1984] no citan qué fuente de variabilidad deseaban identificar.

### **FFT y métodos autorregresivos (AR)**

Sea cual sea la señal a procesar, en la caracterización espectral de la variabilidad del ritmo cardíaco se emplean principalmente dos métodos: la FFT y la estimación espectral del modelo autorregresivo (AR) que mejor se ajusta a la señal a procesar (método AR). En este apartado se comparan sus ventajas y limitaciones.

La FFT es un algoritmo de computación rápido que permite realizar la estimación espectral de series temporales siempre y cuando la longitud de éstas sea potencia de 2. En caso contrario se puede extender la serie a la potencia superior de 2 mediante la introducción de ceros ("*zero padding*"). Este proceso produce una interpolación del espectro pero no aumenta la resolución frecuencial, que está limitada por el número de muestras y la frecuencia de muestreo. El principal problema asociado con el análisis frecuencial mediante FFT surge de la propia naturaleza de la señal a analizar que tiene un alto grado de aleatoriedad [Baselli et al, 1987]. Esto provoca la aparición en el espectro de picos espurios asociados a componentes no periódicos de la señal. Por otro lado, debido a que la señal analizada tiene un número finito de muestras, se hace necesario el empleo de ventanas, y *a priori* no se sabe cuál es la ventana óptima.

Parte de estos problemas quedan solventados con el empleo de la estimación espectral a partir de métodos autorregresivos. En un modelo AR se supone que la entrada al sistema es ruido blanco gaussiano. De manera formal se puede representar como:

$$x(n) = - \sum_{k=1}^p a(k) * x(n-k) + u(n)$$

La salida en un instante  $n$ ,  $x(n)$ , dependerá de la entrada en ese instante,  $u(n)$ , y de las salidas en los  $p$  instantes anteriores, donde  $p$  es el orden del modelo. Para hallar los coeficientes  $a(k)$  se debe establecer un criterio de error que debe ser minimizado. Dichos criterios determinan el algoritmo a emplear. El problema a resolver consiste en hallar una secuencia  $x(n)$  que se parezca a la señal a analizar, designada como  $y(n)$ . De esta forma, una vez ajustado el modelo,  $x(n)$  poseerá la parte determinista de la señal  $y(n)$ . Esto implica que se da por supuesto que  $y(n)$  tiene una cierta componente aleatoria que se superpone a una componente de origen determinista o periódico. Conociendo los coeficientes  $a(k)$  y la estimación de la potencia de ruido a la entrada se puede realizar la estimación espectral de la señal. Este método es una estimación de máxima entropía que tiene mayor resolución espectral que la FFT [Cerutti et al., 1989]. Además, realiza una estimación del espectro más fiable y permite utilizar técnicas de descomposición espectral [Baselli et al., 1986]. Una ventaja adicional del método es que no hace falta inventanar la señal [Pagani et al., 1986].

Una seria limitación de los métodos AR es la determinación del orden correcto para la estimación espectral ( $p$ ). Varios autores utilizan un orden fijo [Sapoznikov et al., 1990] pero en este caso la estimación puede ser completamente incorrecta. Supongamos que existe un orden óptimo para la estimación espectral (sea  $p_{opt}$ ). Si el orden escogido es inferior a  $p_{opt}$  el espectro obtenido estará suavizado respecto al real, es decir, los picos espectrales serán más anchos y podemos perder información sobre la presencia de algún oscilador. Si el orden escogido es superior a  $p_{opt}$  aparecen picos espurios que pueden ser mal interpretados. Hay diversos criterios para la determinación de  $p_{opt}$  pero el más extendido, y también el utilizado aquí, es el criterio de la información de Akaike. En [Baselli et al., 1985] se comenta que el orden del modelo escogido por el criterio de la información de Akaike (que fija el número y la posición de los polos del modelo) indica la complejidad del mecanismo de generación de la serie RR. Por otro lado, la presencia de

latidos anormales (artefactos, latidos ectópicos, etc.) decrece el orden y aumenta el ancho de banda del espectro.

Dado que la FFT tiene más limitaciones que los métodos AR, a partir de ahora, cuando se hable de estimación espectral se supondrá que se realiza mediante el ajuste de la señal a un modelo autorregresivo empleando el criterio de la información de Akaike para la determinación del orden óptimo.

### **Índices para la cuantificación**

La cuantificación de la variabilidad del ritmo cardíaco a partir de métodos espectrales se realiza mediante un número reducido de índices, el cual es función del número de bandas frecuenciales que se estudie o analice. Es decir, una vez establecidas las bandas de interés, se procede a cuantificar de alguna forma la potencia asociada a cada una de estas bandas mencionadas.

Hay tres formas diferentes de cuantificar la potencia asociada a cada banda:

- Cálculo del área bajo la curva de la estimación espectral para la banda de interés.
- Cálculo de la potencia asociada a los polos del modelo AR que contribuyen a la potencia de la banda de interés a partir del teorema de los residuos.
- Cálculo de la altura del pico frecuencial más elevado de la banda de interés.

Los dos primeros métodos son los más utilizados. El tercer método es útil para cuantificar la arritmia sino-respiratoria. En este caso la banda de interés es la banda HF y el índice resultante recibe el nombre de potencia vagal [Reisman et al., 1991].

Para clarificar la nomenclatura, el índice calculado en la banda x por el primer método se denominará xP (sufijo P de potencia) mientras que si lo calculamos mediante el segundo método se denominará xC (sufijo C de componente). Así, para la banda LF tendremos los índices LFP y LFC. Por otro lado, a menudo no se da el valor absoluto de la potencia en cada banda sino que se expresa como porcentaje de la potencia total del espectro [Desai et al., 1991]. En ese caso, añadiremos al sufijo correspondiente el símbolo %. La información aportada por las diferentes formas de cuantificación es esencialmente la misma siempre y

cuando las frecuencias de los osciladores que modulan la variabilidad del ritmo cardíaco estén suficientemente lejos de los límites de las bandas.

Aparte de la cuantificación por separado de cada banda, se suele relacionar la potencia entre bandas. El índice más utilizado es el LF/HF (con los sufijos correspondientes) que es un indicador de la interacción entre el sistema simpático y parasimpático [Pagani et al., 1986].

Algunas observaciones que han probado la utilidad de los métodos espectrales son las siguientes:

- [Pomeranz et al., 1985] comentan que el índice HFP es proporcional a la profundidad de la respiración. Además, disminuye cuando el individuo pasa de posición supina a posición erguida. El índice LFP está relacionado con la presencia de ondas de Mayer. Depende del sistema parasimpático en posición supina y del sistema simpático en posición erguida.
- [Pagani et al., 1986] comentan que el índice LFC da información sobre la regulación del sistema vasomotor (principalmente del reflejo baro-receptor) y aumenta al pasar de supino a erguido pues aumenta la actividad simpática. El índice HFC prácticamente se anula con la inyección de atropina que es un inhibidor vagal... Por otro lado, el índice LFC/HFC indica la interacción entre las dos ramas del sistema nervioso autónomo. Aumenta al pasar de supino a erguido y disminuye con respiración sincrona durante la cual aumenta la actividad vagal.
- [Bianchi et al., 1991] comentan que el índice VLFC aumenta en estados REM (durante el sueño) así como durante actividad física. Todo esto se asocia a activaciones del sistema simpático inusitadamente elevadas.

Todos estos trabajos avalan el interés del estudio de la variabilidad del ritmo cardíaco mediante métodos espectrales, y por lo tanto los perfeccionamientos en el cálculo de índices que permitan mejorar el análisis.

## Conclusiones

Se han mostrado las ventajas y limitaciones de los métodos espectrales con la intención de plantear una metodología para el procesado de series RR.

Se ha comentado que los métodos espectrales son una buena herramienta para investigar con mayor detalle las causas que originan la variabilidad del ritmo cardíaco. Tras revisar diversos métodos para la caracterización espectral de la variabilidad del ritmo cardíaco se propone el empleo de los métodos autorregresivos debido a la naturaleza casi aleatoria de la señal RR. Se ha escogido como orden del modelo aquel dictado por el criterio de la información de Akaike.

Se debe tener en cuenta la influencia del re-muestreo de la señal RR sobre la estimación espectral y se ha observado que sólo es necesario cuando se desea analizar la influencia de un oscilador exógeno y se produce un cambio en el ritmo medio que es superior a la amplitud del oscilador. El re-muestreo es contraproducente cuando se desea analizar un oscilador endógeno de gran amplitud.

Para realizar una estimación espectral fiable, se requiere que la serie RR sea estacionaria. Se ha comprobado que los test de estacionariedad son poco fiables y se recomienda una inspección visual de la señal.

La presencia de artefactos en la serie RR produce grandes errores en la estimación de los índices espectrales y se propone la señal DRR (diferencias en la serie RR) como alternativa, donde  $DRR(n) = RR(n+1) - RR(n)$ . El análisis espectral de la señal DRR produce unos índices espectrales de alta correlación con los que se obtienen en la serie RR en ausencia de artefactos. Si hay artefactos, el análisis mediante la señal DRR es mucho más robusto.

Por otro lado, la variabilidad de la frecuencia de los osciladores que modulan la variabilidad del ritmo cardíaco puede ser una causa de sobrestimación de la potencia en las bandas de baja frecuencia y de muy baja frecuencia. Si no se posee una medida sobre el oscilador, se debe recurrir a métodos espectro-temporales para cuantificar este efecto.

Recuérdese que una sobrestimación de la banda de baja frecuencia produce índices LF/HF elevados que pueden llevar a una diagnosis incorrecta.

En la Figura 15 se presenta el esquema del procesado para la señal RR. Si la señal es estacionaria se realiza el análisis espectral. En caso contrario se realiza el análisis espectro-temporal. Sólo si se está interesado en una fuente exógena y la variabilidad del ritmo medio es elevada, se realiza un re-muestreo. Si hay riesgo de artefactos se analiza la señal DRR. Tras realizar la estimación espectral mediante modelos autorregresivos se realiza la cuantificación de las bandas según los límites escogidos por el usuario. Recuérdese que el análisis espectro-temporal puede ser útil en la detección de la modulación de los osciladores.

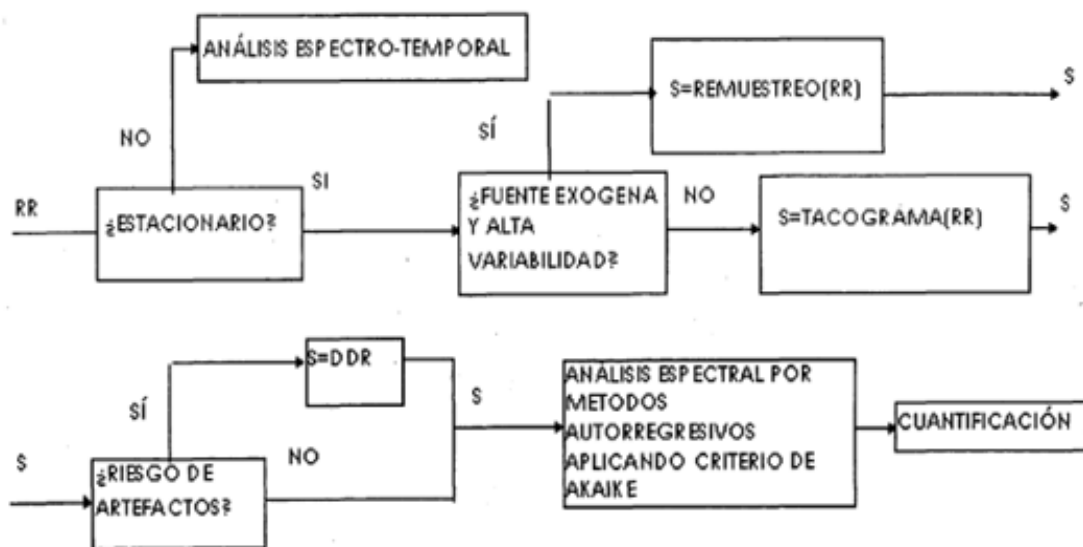


Figura 15: Esquema de procesado de la señal RR para el análisis espectral

### 2.4.3. Métodos espectro-temporales

En esta sección se habla de los diferentes métodos espectro-temporales distinguiendo para qué tipo de aplicaciones son más idóneos cada uno de los tipos de análisis posibles. Se proponen como única solución posible para el análisis espectral de registros largos y no estacionarios, y se sugiere el empleo de las ondículas para el reconocimiento de ondas características dentro de la serie RR.

El objetivo es presentar y aplicar al estudio de la variabilidad del ritmo cardíaco diversos métodos de análisis espectral que aún no se utilizan comúnmente pero que pueden ser útiles cuando la señal RR no es estacionaria. Los métodos que se estudian son: los métodos espectrales variantes con el tiempo, los métodos espectro-temporales y los métodos de escala-tiempo.

Los **métodos espectrales variantes con el tiempo** son una simple extensión de los métodos espectrales en la que se realizan diferentes estimaciones espectrales utilizando una ventana temporal que se desliza a lo largo del registro. De esta forma, se consigue analizar registros no estacionarios y observar cómo varía el contenido espectral de la señal con el tiempo.

Los **métodos espectro-temporales** son una sofisticación de los anteriores. Se emplean diferentes métodos para la transformación en frecuencia al mismo tiempo que se utilizan ventanas de suavizado en tiempo y en frecuencia. De esta forma se consigue paliar el efecto negativo que provoca la aparición de términos cruzados debido a que se utilizan funciones cuadráticas en el cálculo de la distribución espectro-temporal.

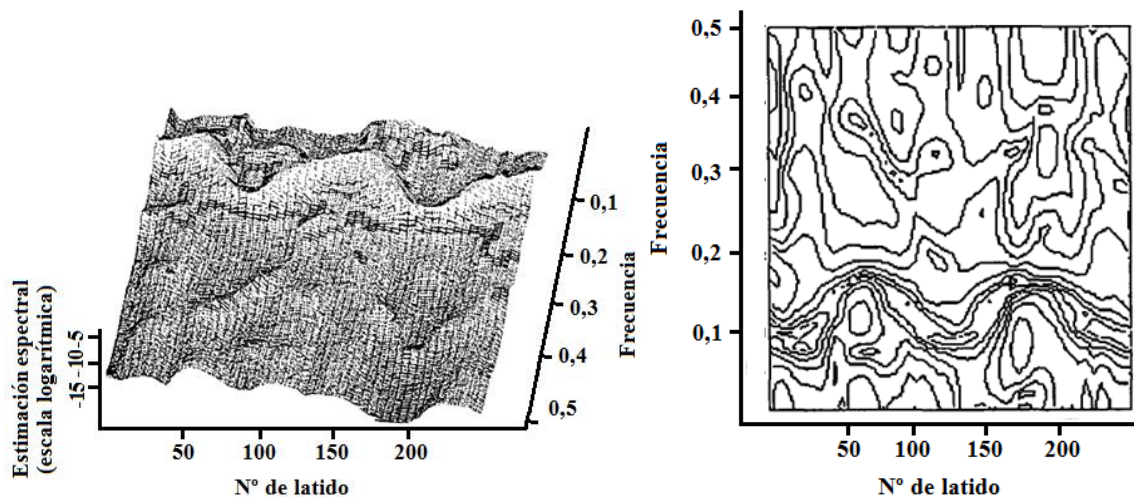
Una visión radicalmente distinta es la que ofrecen los **métodos de escala-tiempo**. Mientras los métodos espectrales y espectro-temporales parten de la descomposición de la señal en senoides, los métodos de escala-tiempo la realizan mediante otras funciones arbitrarias (es decir, a escoger por el usuario) que deben cumplir unos requisitos mínimos. La utilidad de estos métodos es doble: por un lado, si hacemos la abstracción de asignar a una escala una frecuencia determinada, tenemos un nuevo método espectral con ciertas ventajas añadidas; y por otro, se puede realizar una descomposición de la señal en componentes de variación rápida y lenta.

### **Métodos espectrales variantes con el tiempo**

Anteriormente se ha comentado que los métodos espectrales clásicos fallan a la hora de aportar información en series temporales cuyas componentes espectrales y energía cambian con el tiempo. Por esta razón, varios autores han desarrollado diversos métodos para la estimación espectral de señales no estacionarias.

La aproximación al problema que proponen los métodos espectrales variantes con el tiempo es realizar sucesivas estimaciones espectrales de segmentos de la señal de interés que estén solapados [Bianchi et al, 1991]. Evidentemente, cuanto menor sea la duración del segmento mayor definición temporal se tendrá aunque se perderá resolución frecuencial. Los métodos más comúnmente empleados para realizar este análisis se basan en modelos autorregresivos debido a su alta resolución frecuencial. De esta forma, se calcula un espectro para una ventana temporal determinada; deslizando dicha ventana a lo largo del registro se obtiene la variación de las diferentes componentes espectrales a lo largo del tiempo.

La forma más común de representación de la estimación espectral variante con el tiempo es la matriz espectral comprimida (CSA). Este esquema tridimensional tiene por eje de abscisas el tiempo y por eje de ordenadas la frecuencia, mientras que la altura es proporcional a la potencia espectral para una frecuencia determinada en un instante determinado. Otra forma de visualización es mediante un mapa de nivel. Esta representación bidimensional representa en el eje de abscisas el tiempo y en el eje de ordenadas la frecuencia. Se dibujan en este plano las curvas de igual nivel de potencia.



**Figura 16: Representación de secuencia RR (Izqda.: CSA, Dcha.: mapa de nivel)**

En la Figura 16 se ha representado la estimación espectral variante con el tiempo de una secuencia RR donde un sujeto estaba respirando con una frecuencia variante con el tiempo. El orden escogido por el modelo es fijo e igual a 10. La longitud de la ventana es de 50 muestras y se desliza en pasos de 5 muestras. Así mismo, se pone de manifiesto las

diferentes elecciones que se deben realizar cuando se desea realizar un análisis espectral variante con el tiempo:

- Orden del modelo para cada ventana de observación.
- Longitud de la ventana.
- Desplazamiento en muestras de la ventana.

Un orden igual a 10 evita tener que calcular el orden por el criterio de la información de Akaike y agiliza los cálculos puesto que el orden es pequeño, aunque lo suficientemente elevado como para poder representar la información en una ventana pequeña. Más crítica es la elección de la longitud de la ventana y del desplazamiento de ésta.

Los métodos espectrales variantes con el tiempo han demostrado ser muy útiles en la determinación de variaciones de frecuencia y en la caracterización de registros no estacionarios. Sin embargo, aunque pueden detectarse variaciones en la potencia de la señal, estos métodos no son muy aconsejables para la cuantificación de potencia de las señales ya que por cada simulación aparecen variaciones que pueden causar errores de cuantificación. A pesar de ello, estos métodos pueden ser aplicados en una primera aproximación ya que su complejidad de cálculo es baja y dan excelentes resultados en la detección de picos espectrales. Si el interés reside en una determinación exacta de la variación de la potencia de la señal con el tiempo se requerirán métodos más sofisticados como las distribuciones tiempo-frecuencia.

### **Representaciones tiempo-frecuencia**

El principal problema de la estimación espectral variante con el tiempo mediante métodos autorregresivos es su deficiente estimación de la potencia de la señal. En [Pola et al., 1996] se cita que estos métodos ofrecen malos resultados en señales simuladas tal y como hemos comprobado en la sección anterior. No obstante, sirven para discernir la variación en frecuencia de alguno de los osciladores. Si se desea una cuantificación mejor de la potencia instantánea de la señal y de sus componentes espectrales respecto al tiempo se requiere el uso de métodos espectro-temporales más complejos. Estos métodos utilizan transformaciones matemáticas, similares a la transformada de Fourier, que suelen conocerse como representaciones tiempo-frecuencia.

La necesidad de aplicar representaciones tiempo-frecuencia surgió históricamente con el afán de analizar espectralmente registros no estacionarios. Recuérdese que el análisis espectral variante con el tiempo es una aproximación burda al problema: escogiendo una ventana corta se supone que la señal dentro de ella es estacionaria. En cambio, las representaciones tiempo-frecuencia permiten analizar más fácilmente todo tipo de señales no estacionarias.

Las transformadas tiempo-frecuencia muestran cómo el contenido espectral de la señal a analizar cambia con el tiempo. Son extensiones de la transformada de Fourier que representan una superficie sobre el llamado plano tiempo-frecuencia. Las representaciones tiempo-frecuencia más utilizadas son las cuadráticas. Las propiedades deseables de toda representación tiempo-frecuencia son:

- La integral en todo el plano tiempo-frecuencia de la representación debe ser igual a la energía total de la señal.
- La integral en el eje temporal debe dar el espectro de potencia de la señal mientras que la integral en el eje frecuencial debe dar la potencia instantánea de la señal. Ambas integrales son conocidas como los marginales de la representación tiempo-frecuencia.
- La representación debe ser real y, a ser posible, definida positiva.
- Sea  $P(t, f)$  la representación tiempo-frecuencia de la señal  $s(t)$ . Entonces, la representación de  $s(t-t_0)$  debe ser  $P(t-t_0, f)$  y la representación de  $s(t)e^{j2\pi f_0 t}$  debe ser  $P(t, f-f_0)$ .

Otras propiedades que interesa que posean las representaciones tiempo-frecuencia son:

- El valor medio de la frecuencia en un instante determinado debe coincidir con la frecuencia instantánea de la señal. El valor medio del tiempo debe coincidir con el retardo de grupo de la señal.
- Si no hay señal fuera del intervalo temporal  $[t_1, t_2]$ , entonces  $P(t, f)$  debe ser nula fuera de dicho intervalo. Similarmente, si no hay componentes espectrales fuera del intervalo frecuencial  $[f_1, f_2]$ , entonces  $P(t, f)$  debe ser nula fuera de dicho intervalo.
- La representación de la señal  $\sqrt{a}s(at)$  debe ser  $P(at, f/a)$ .

- Si  $y(t)$  es la convolución de  $x(t)$  y  $h(t)$ , su representación debe ser el producto de las representaciones de  $x(t)$  y  $h(t)$ . Si  $y(t)$  es el producto de  $x(t)$  y  $h(t)$  entonces su representación es la convolución frecuencial de las representaciones de  $x(t)$  y  $h(t)$ .
- Si  $x_1(t)=s(t)$  tiene por representación  $P_1(t,f)$  y  $x_2(t)=S(t)$ , donde  $S(f)$  es la transformada de Fourier de  $s(t)$ , se debe cumplir que  $P_2(t,f)=P_1(-f,t)$ .
- El producto escalar de dos representaciones debe ser igual al producto de las energías de las dos señales.

Uno de los problemas principales que poseen las representaciones cuadráticas es la aparición de términos cruzados debido a batidos de frecuencia. Dichos términos deberían ser nulos en una representación tiempo-frecuencia ideal. Para intentar eliminarlos se opta por modificar las representaciones empleando ventanas espectro-temporales adecuadas.

Hay un gran número de representaciones tiempo-frecuencia. Algunas de las más utilizadas son:

- Espectrograma
- Choi-Williams
- Margenau-Hill
- Page
- SPWVD (*Smoothed Pseudo-Wigner-Ville Distribution*)
- Wigner-Ville
- MSTFT (*Modified Short-Time Fourier Transform*)
- MSPWVD (*Modified and Smoothed Pseudo-Wigner-Ville Distribution*)

Una mayor información sobre cada una de estas (y otras) representaciones y sus propiedades matemáticas y formulaciones puede encontrarse en [Boualem B. et al, 2003].

### **Métodos de escala-tiempo (ondículas)**

Se han presentado hasta aquí dos métodos para el estudio de la variación del contenido espectral de una señal con el tiempo. Ambos métodos tienen en común que asumen que la

señal es fruto de la superposición de un cierto número de señales senoidales. Es decir, son modificaciones o ampliaciones de un análisis de Fourier.

Recientemente [Akay, 1995] se ha introducido otra metodología en el análisis de la señal. Son los métodos conocidos como tiempo-escala y se basan en las transformadas por ondículas. Estas transformadas realizan la descomposición de la señal sobre un conjunto de funciones obtenidas a partir de otra llamada madre mediante dilataciones o contracciones y desplazamientos en el tiempo de ésta. En este caso, no se hablará de frecuencia sino de escala (aunque la frecuencia puede ser relacionada con la escala para las ondículas utilizadas habitualmente).

Se pueden distinguir, básicamente, dos tipos de transformadas por ondículas: la continua (CWT: *Continuous Wavelet Transform*) donde los parámetros de escala y tiempo varían de forma continua y la discreta (DWT: *Discrete Wavelet Transform*) donde se reducen las parejas posibles de parámetros a una red diádica.

La transformada por ondículas es un método aplicable a cualquier tipo de señal y constituye una transformación invertible siempre y cuando la ondícula cumpla una serie de condiciones. Las diferencias fundamentales entre las ondículas y los métodos espectro-temporales son:

- La ondícula tiene característica pasobanda. En cambio, las ventanas asociadas a los métodos espectro-temporales tenían características pasobajo.
- La transformación o descomposición en ondículas se realiza mediante desplazamientos temporales y con una operación de escalado. En los métodos espectro-temporales se emplea el desplazamiento temporal pero en lugar de la operación de escalado se utiliza una modulación (desplazamiento frecuencial).
- La resolución en tiempo y en frecuencia de los métodos espectro-temporales es constante en todo el plano (t,f). En cambio, en la transformada por ondículas se tiene una buena resolución frecuencial a bajas frecuencias y una buena resolución temporal a altas frecuencias. Eso sí, el producto duración por ancho de banda es constante.

En la CWT, las variables que controlan la escala y el desplazamiento temporal pueden tener cualquier valor real excepto cero. La ondícula con la cual se realiza el análisis debe cumplir la condición de admisibilidad si se desea que la transformada sea reversible [Daubechies, 1991]. Si se desea descomponer la señal en una red mínima de variables de escala y desplazamiento temporal, se debe emplear la DWT. Para poder reconstruir *a posteriori* la señal, se requiere que, en este caso, las ondículas cumplan un conjunto mayor de propiedades. Un método rápido para la computación de la DWT es la FWT (*Fast Wavelet Transform*), método también conocido como el algoritmo de Mallat [Mallat, 1989]. Además, con modificaciones de este algoritmo se puede calcular la CWT.

En los últimos años, se han desarrollado conjuntos de ondículas con buenas propiedades para el análisis tiempo-frecuencia. Dichas ondículas constituyen bases ortonormales que permiten una descomposición de la señal a diferentes escalas. Algunas de las cualidades de estas bases de funciones son:

- Deben ser regulares y simétricas: De esta forma se evitan los problemas asociados a la ausencia de algunos valores de escala y desplazamiento (debido a la resolución del procesador o a la misma cuantificación de la señal).
- Gran número de momentos nulos: Cuanto mayor sea este número mejor será la localización de la energía de la señal para una cierta escala e instante.
- Decaimiento hacia el infinito: Dependiendo de este decaimiento tendremos una mejor resolución temporal o frecuencial.

La idea que subyace en el análisis multirresolución es expresar una señal como el límite de una sucesión de aproximaciones cada vez mejores de la señal. Cada una de las aproximaciones tiene una resolución creciente.

Las bases ortonormales más empleadas son:

- Las ondículas de Daubechies [Daubechies, 1988].
- Las *coiflets* [Daubechies, 1993].
- Las *splines* [Unser et al, 1992] para el cálculo de la CWT.
- Las ondículas de Morlet [Morlet et al, 1993].

Se debe tener en cuenta que la resolución de estos métodos no es constante en el plano tiempo-frecuencia por lo cual un cambio rápido en la señal (por ejemplo, un artefacto) estará muy bien localizado en el tiempo. Por otro lado, las derivas de baja frecuencia serán estimadas con una alta exactitud en frecuencia aunque con baja resolución en tiempo.

La transformada por ondículas no puede ser utilizada para un análisis espectro-temporal debido a que la resolución frecuencial no es muy elevada para las frecuencias de interés comúnmente analizadas en variabilidad del ritmo cardíaco. No obstante, tiene otros alcances como, por ejemplo, la detección de transitorios rápidos. Según esto se podría llegar a la falsa conclusión de que el empleo de ondículas no aporta ninguna mejora sobre los métodos espectro-temporales.

Los métodos tiempo-escala empleando ondículas se caracterizan por tener muy buena resolución frecuencial cuando la frecuencia a analizar es baja. En el caso de los métodos espectro-temporales, la resolución es constante sea cual sea el punto de interés del plano tiempo-frecuencia. Se ha observado, no obstante, que los métodos espectro-temporales no tienen la resolución frecuencial deseada a baja frecuencia (a no ser que se escoja una ventana lo suficientemente grande con lo cual queda comprometida la velocidad de procesado).

Se pueden utilizar ondículas para realizar un análisis multirresolución. Este análisis descompone la señal en diferentes escalas de manera que al final del proceso se tiene una serie de señales que son la salida de un banco de filtros. Dichas señales pueden ser utilizadas (tras un procesado no lineal) para la detección de diversas formas de onda dentro de la serie RR.

Por tanto, sobre la utilización de los métodos basados en ondículas se puede concluir que:

- El análisis mediante ondículas ha presentado un doble interés: el empleo de las transformadas por ondículas para el análisis espectro-temporal de la señal y el empleo del análisis multirresolución para la detección de ondas características.
- El empleo de ondículas en el estudio de las bandas más utilizadas en el análisis de la variabilidad del ritmo cardíaco (LF y HF) no aporta ninguna mejora y podría ser comparado al empleo de métodos espectrales variantes con el tiempo. Eso sí, su

alta resolución temporal en alta frecuencia puede ser utilizada para la detección de transitorios bruscos en la señal. La gran ventaja del empleo de ondículas reside en el estudio de oscilaciones de muy baja frecuencia.

- Por otro lado, el análisis multirresolución unido a otras técnicas que emplean transformaciones no lineales ha demostrado ser útil para la detección de ondas características dentro de la serie RR. Así se ha podido detectar con suma facilidad la posición temporal de las llamadas ondas de corte.

## **Conclusiones**

La necesidad del empleo del análisis espectro-temporal en la caracterización de la variabilidad del ritmo cardíaco había aparecido en los estudios sobre los métodos espectrales, donde se ha visto que la no estacionariedad de la serie RR o la variación de la frecuencia de los osciladores con el tiempo podía provocar un falso diagnóstico mediante índices espectrales. El análisis espectro-temporal puede ser empleado también como marcador de eventos: un cambio súbito en la frecuencia de uno de los osciladores indica el inicio de un nuevo estado de regulación cardiovascular.

Si se desea realizar únicamente una comprobación de la estabilidad de la frecuencia de los osciladores respecto al tiempo, un análisis espectral variante con el tiempo empleando métodos autorregresivos puede ser suficiente. No obstante, la determinación de la potencia de estos osciladores con estos métodos es muy burda con lo cual resultan inútiles para una cuantificación.

Las representaciones tiempo-frecuencia permiten realizar las mismas comprobaciones que los métodos espectrales variantes con el tiempo pero, además, aportan una determinación de la potencia de los osciladores muy exacta. En cambio, el tiempo de computación con estos métodos es mayor. La representación SPWVD y el espectrograma han demostrado ser muy eficaces en el análisis de la variabilidad del ritmo cardíaco.

El empleo de ondículas es útil para el seguimiento de la variación de frecuencia de los osciladores pero, además, permite determinar la presencia y variación de frecuencias muy lentas, lo cual es una clara ventaja respecto a las representaciones tiempo-frecuencia.

Además, pueden ser empleadas para la detección de ondas características en la serie RR.  
No obstante, la estimación de la potencia espectral deja mucho que desear.



### 3. METODOLOGÍA

Con objeto de afrontar los objetivos mencionados en el apartado anterior, se han definido cuatro casos de estudio (uno por objetivo), para un mismo recorrido por carretera que asegure la conducción cercana a las tres horas, seleccionándose hasta **38 personas diferentes** que pudieran satisfacer las condiciones óptimas del estudio, sobre todo desde el punto de vista de la salud (personas sin problemas de salud que afectaran la señal ECG) y contemplando los distintos grupos de edad (factor personal analizado).

Según esto se realizaron en total **64 experimentos** (conducción por participante de más de 2 horas y 45 minutos) donde, con objeto de analizar la influencia de los ciclos circadianos en el HRV, más o menos el 50% de los experimentos son en el turno de día (06:00 a 12:59) y el otro 50% en el turno de tarde (13:00 a 19:59). Además, se diseñó un recorrido con diferentes entornos (tipo de carretera) que demandaran distintos niveles de alerta y así analizar la influencia del entorno (ambiental) en el HRV (caso de estudio correspondiente).

Se describen a continuación las etapas del método seguido, la descripción del recorrido seleccionado, la lista de participantes, las herramientas hardware y software utilizadas, así como los casos de estudio en los que se estructura este trabajo.

#### 3.1. Etapas de la metodología

Las etapas seguidas para los 64 experimentos se describen según la Figura 17:

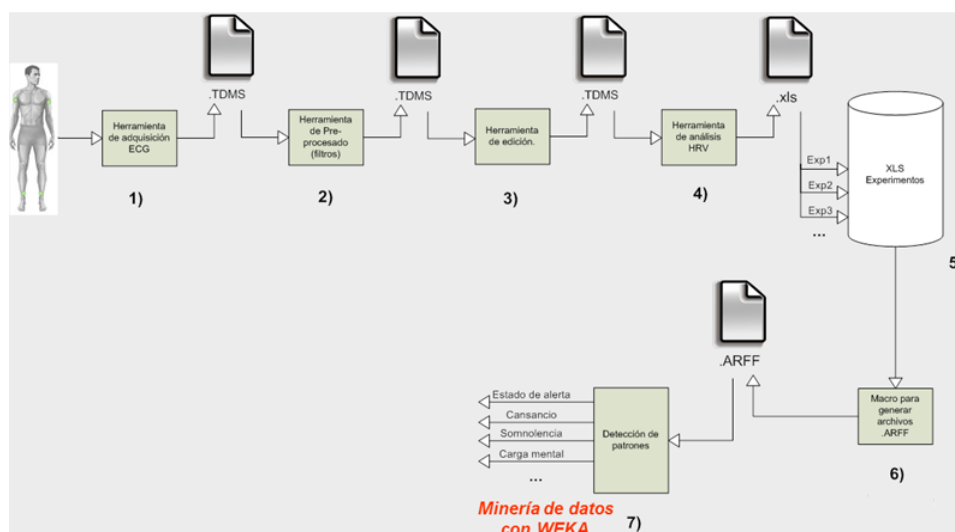
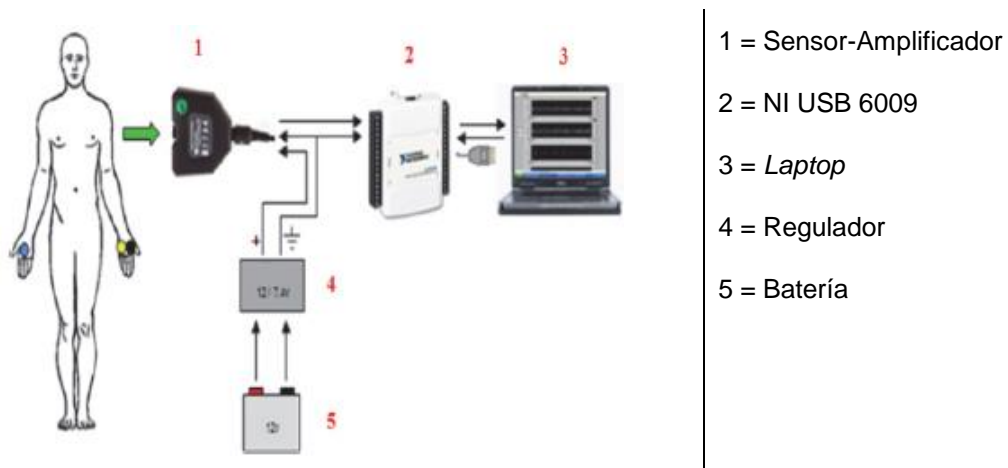


Figura 17: Etapas de la metodología (Adquisición – Análisis)

A continuación se describe cada una de las etapas:

### 3.1.1. Adquisición

En la etapa de adquisición se instruye a los participantes sobre la colocación de tres electrodos conectados a un sistema de adquisición según la Figura 18 mediante el cual se registra la señal ECG durante más de 2 horas y 45 minutos de conducción del vehículo por el recorrido que se define más adelante:

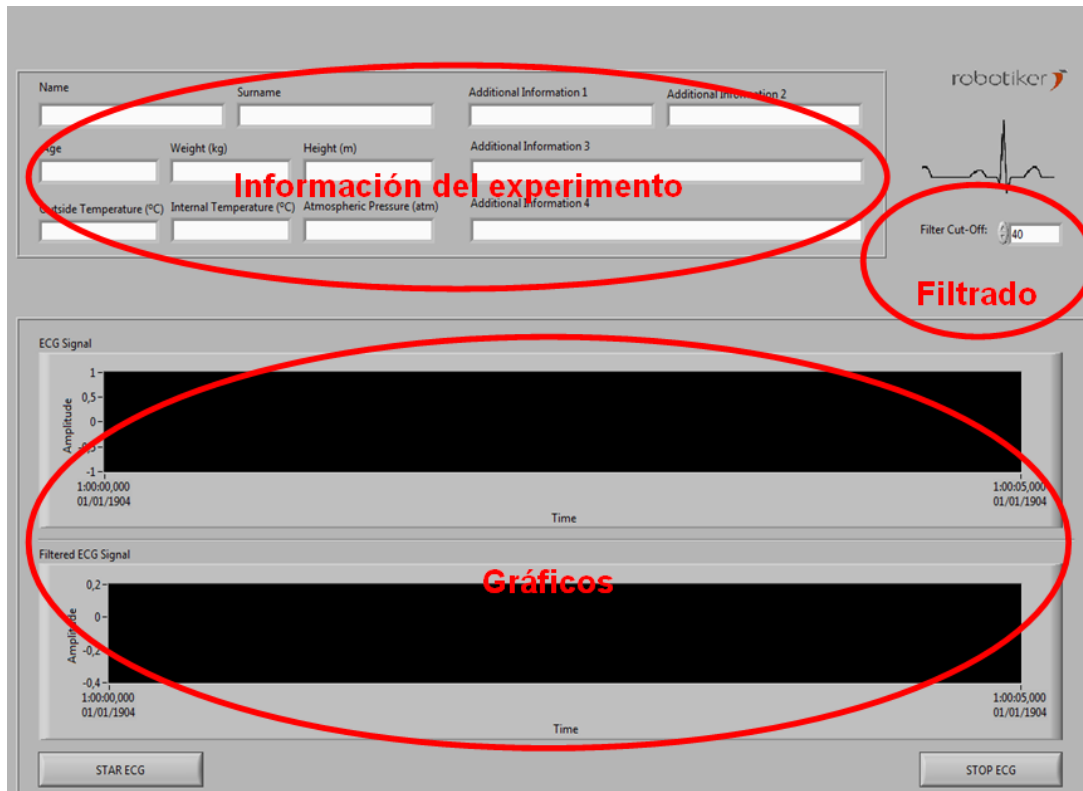


**Figura 18: Sistema de adquisición utilizado**

Era importante que la alimentación del amplificador del ECG se encontrara entre 7V y 7,3V del sensor-amplificador de lo contrario la señal adquirida no era lo más limpia posible (sin ruidos eléctricos ni artefactos).

Para esta etapa, en conjunto con el sistema de adquisición de la Figura 18, se ha desarrollado una aplicación en el entorno gráfico LabVIEW que almacena la señal ECG en el “Laptop” en ficheros con extensión TDMS (uno por cada experimento). Esta herramienta consta de 3 partes (Figura 19):

- Información del experimento: En este apartado se incluye la información relacionada con el experimento: nombre y apellido de la persona, edad, peso, estatura, etc.
- Filtrado: Se selecciona la frecuencia de corte del filtro pasa-bajos.
- Gráfico: En esta parte se monitoriza constantemente la señal ECG adquirida y la filtrada.



**Figura 19: HMI de aplicación de adquisición**

Para más detalle sobre el equipamiento utilizado ver en anexos datos técnicos de los equipos (también se puede consultar el manual de usuario correspondiente a la mencionada aplicación de adquisición).

### 3.1.2. Pre-procesado

En este paso se aplican los filtros necesarios para la eliminación del ruido de frecuencias bajas en el ECG registrado anteriormente, principalmente para estabilizar todo el registro. La señal resultante se guarda con el mismo formato que el archivo original (TDMS).

### 3.1.3. Edición (Recuperación manual)

En esta etapa se obtienen los intervalos RR de la señal ECG pre-procesada mediante la detección de picos de una aplicación-editor desarrollada, la cual permite ir al detalle para corregir los intervalos RR que se han seleccionado erróneamente (artefactos, falsos positivos y falsos negativos). La Figura 20 muestra el HMI de esta aplicación:

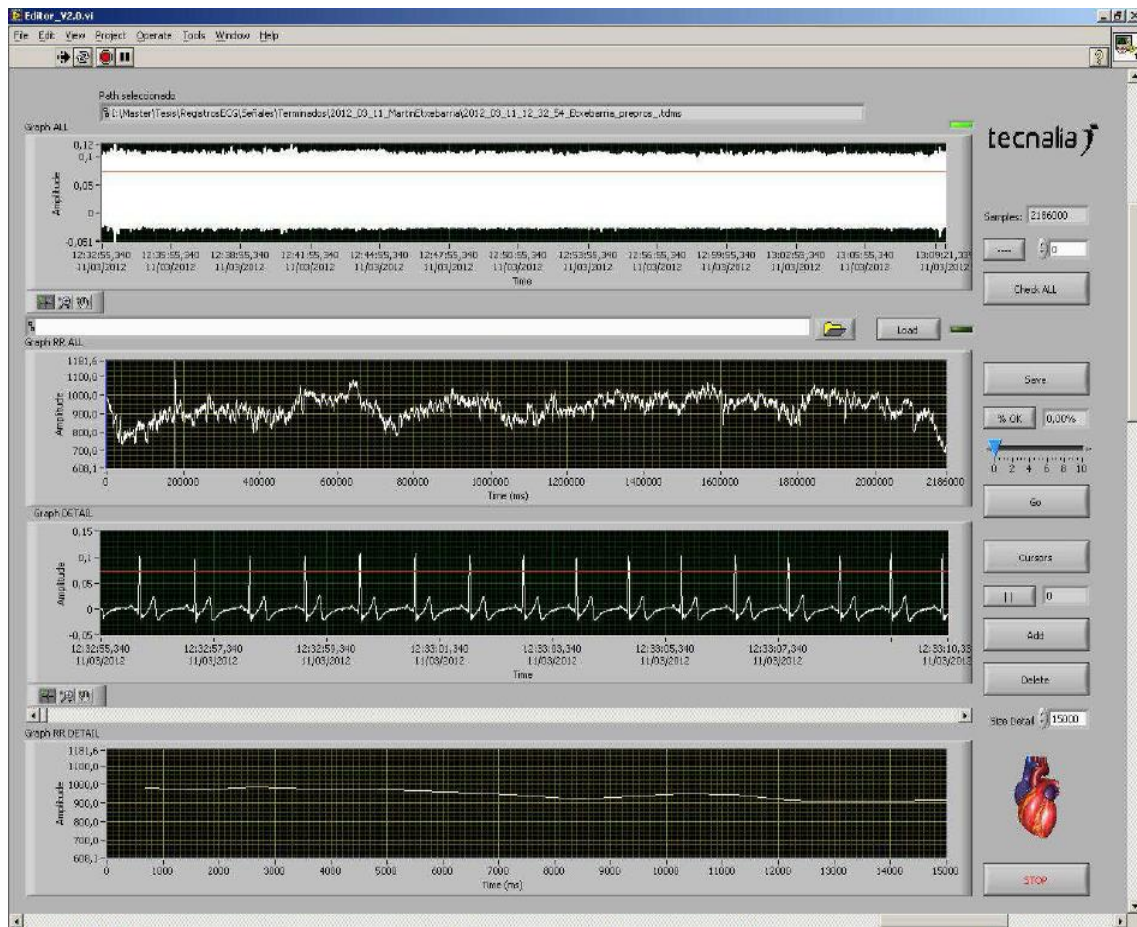


Figura 20: HMI de aplicación “Editor”

Para más detalle sobre esta aplicación ver manual de usuario correspondiente.

### 3.1.4. Post-Procesado 1: Herramienta de análisis HRV

En esta etapa se parte de los ficheros TDMS corregidos por el editor anterior y mediante una nueva aplicación propia desarrollada en Labview se analizan los intervalos RR en el espacio temporal y de frecuencias. La Figura 21 muestra el interfaz de usuario de dicha herramienta (para más detalle ver manual de usuario correspondiente):

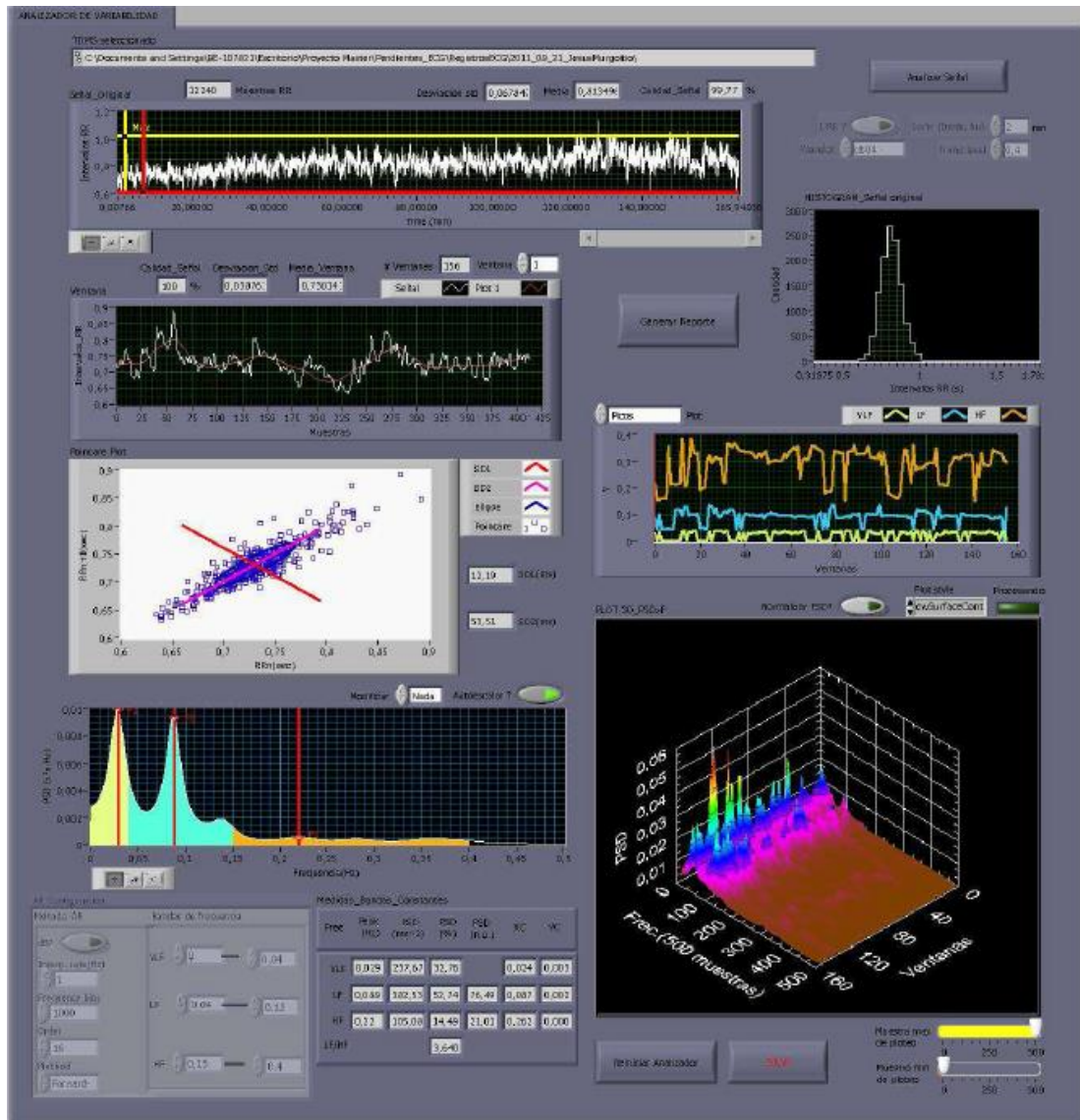


Figura 21: HMI de aplicación de análisis previo

Como resultado de esta etapa se obtiene un fichero “.xls” de excel para cada experimento donde, con el fin de ver su evolución en el tiempo de los intervalos RR en las aproximadamente 3 horas de conducción ininterrumpida, se divide en ventanas de 5 minutos, estando desfasada cada una de las ventanas respecto de la anterior en 2,5 minutos (*overlapping* del 50%):

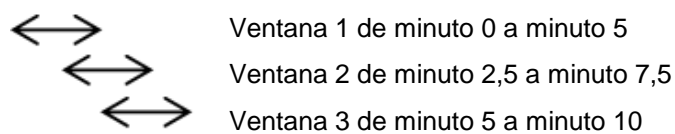


Figura 22: División de las señal en intervalos de 5 minutos (ventanas)

Los datos obtenidos para cada ventana de cinco minutos son entre otros los siguientes:

- Parámetros estadísticos de toda la señal como la media y desviación estándar.
- Parámetros estadísticos (media, desviación estándar, SD1 y SD2) de cada ventana de 5 min.
- Porcentaje de calidad de toda la señal y por cada ventana de 5 min.
- PSD de las frecuencias VLF, LF y HF por cada ventana de 5 min.
- Relación LF/HF por cada ventana de 5 min.
- Componentes “X” e “Y” de los picos de VLF, LF y HF por cada ventana de 5 min.
- Componentes “X” y “Y” de los centros de masa de cada frecuencia VLF, LF y HF por cada ventana de 5 min.

Para más detalle sobre toda la información recogida ver capítulo sobre modelo de datos y variables.

### **3.1.5. Post-Procesado 2: Compilación**

En esta segunda etapa de post-procesado, en base al modelo de datos diseñado (ver capítulo de diseño de variables) se compila toda la información referente a los participantes (datos personales), experimentos y ventanas temporales para, a partir de todos los ficheros excel de experimentos (uno por cada experimento) generar un único fichero “.xls” con todos los parámetros para cada ventana temporal de cinco minutos de todos los experimentos realizados. De esta manera los datos quedan preparados para su conversión a archivos ARFF de WEKA según la etapa que se describe a continuación.

### **3.1.6. Conversión**

La etapa de conversión no es más que el cambio de formato mediante una macro de excel de los datos obtenidos en el Post-Procesado 2 (xls) al formato que entiende WEKA denominado “arff”, acrónimo de *Attribute-Relation File Format* (Figura 23). Este formato está compuesto por una estructura claramente diferenciada en tres partes: cabecera, declaración de atributos (@attribute) y sección de datos (@data):

```

% Archivo de prueba para Weka.
@relation prueba

@attribute nombre STRING
@attribute ojo_izquierdo {Bien,Mal}
@attribute dimension NUMERIC
@attribute fecha_analisis DATE "dd-MM-yyyy HH:mm"

@data
Antonio,Bien,38.43,"12-04-2003 12:23"
'Maria Jose',?,34.53,"14-05-2003 13:45"
Juan,Bien,43,"01-01-2004 08:04"
Maria,?,?, "03-04-2003 11:03"

```

**Figura 23: Ejemplo de formato "arff"**

### 3.1.7. Minería de datos

Por último, en base a las ventanas de información de cinco minutos de conducción generadas según las etapas anteriores, se ha llevado a cabo el análisis final para cada caso de estudio utilizando técnicas de minería de datos en base a herramientas para selección de atributos (Evaluador de atributos + Método de búsqueda) y clasificadores, todas ellas pertenecientes a una extensa colección de algoritmos de Máquinas de conocimiento desarrollados por la universidad de Waikato (Nueva Zelanda) implementados en Java; útiles para ser aplicados sobre datos mediante los interfaces que ofrece o para embeberlos dentro de cualquier aplicación, la cual se conoce como WEKA [Ian H. Witten, Eibe Frank, Len Trigg, Mark Hall Geoffrey Holmes, and Sally Jo Cunningham, 1999] y [Ian H. Witten and Eibe Frank, 2000]. La versión de Weka utilizada ha sido la 3.6.10 y contiene las herramientas necesarias para realizar transformaciones sobre los datos, tareas de clasificación, regresión, *clustering*, asociación y visualización, estando diseñada como una herramienta orientada a la extensibilidad por lo que añadir nuevas funcionalidades es una tarea sencilla.

Así, para cada caso de estudio se han definido **tres estudios** según el tipo de variables independientes analizadas (variables en el dominio del tiempo, la frecuencia y combinada), donde para cada estudio se ha realizado un primer proceso de **selección de los atributos** mejor correlacionados con la variable dependiente a analizar, y por último un análisis

basado en dos **clasificadores** con estrategias bien diferentes cuyo rendimiento ha demostrado su eficacia en una fase preliminar.

### Selección de atributos:

El proceso de selección de atributos se ha llevado a cabo una vez definidas las variables independientes para cada uno de los tres estudios de cada caso de estudio, y para ello se ha utilizado como **evaluador de atributos** a “weka.attributeSelection.CfsSubsetEval” (Figura 24):

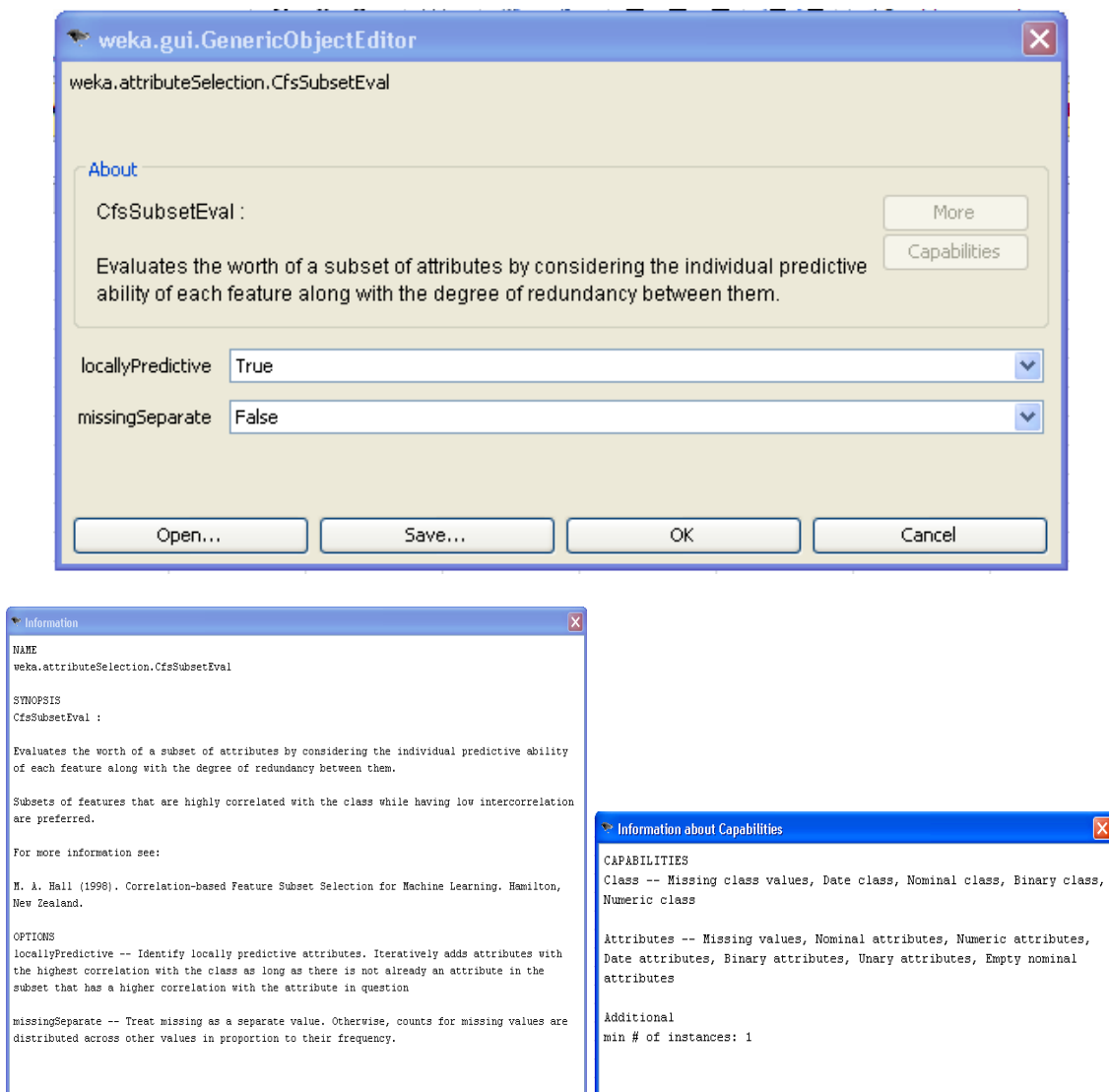
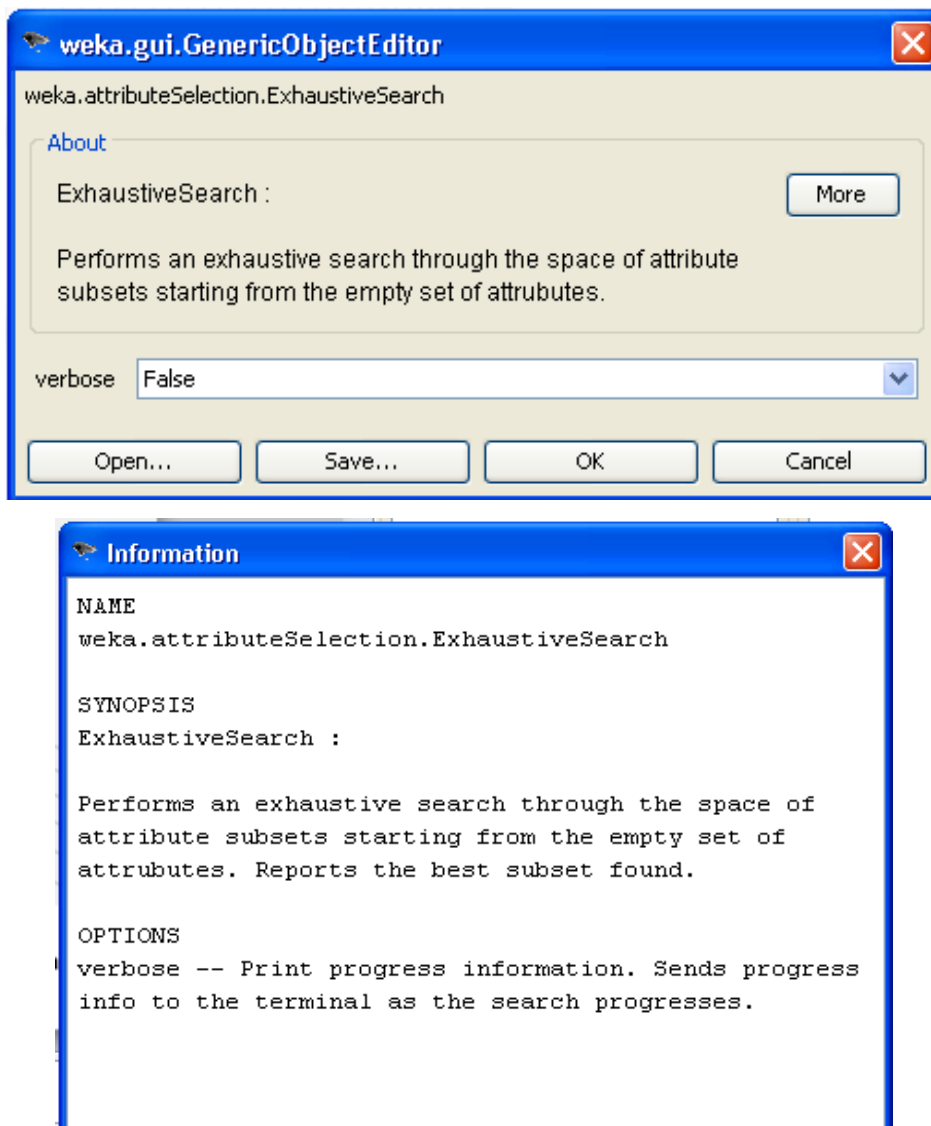


Figura 24: Ayuda WEKA sobre “CfsSubsetEval”

y como **método de búsqueda** a “weka.attributeSelection.ExhaustiveSearch” (Figura 25):



**Figura 25: Ayuda WEKA sobre “ExhaustiveSearch”**

En base a los niveles de correlación obtenidos con estos dos métodos se establece un **ranking de atributos** según el criterio de agrupamiento que se describe más adelante en el capítulo “Análisis y resultados”, y sobre el que se ha llevado a cabo el análisis posterior con clasificadores (ver manual de WEKA para las dos herramientas anteriores).

### **Clasificación:**

Una vez obtenido el ranking de atributos para cada estudio, se utilizan dos clasificadores con buenos rendimientos para las variables dependientes según dos estrategias diferentes.

## Clasificador 1 (CL1) (Figura 26): “weka.classifiers.trees.RandomTree”

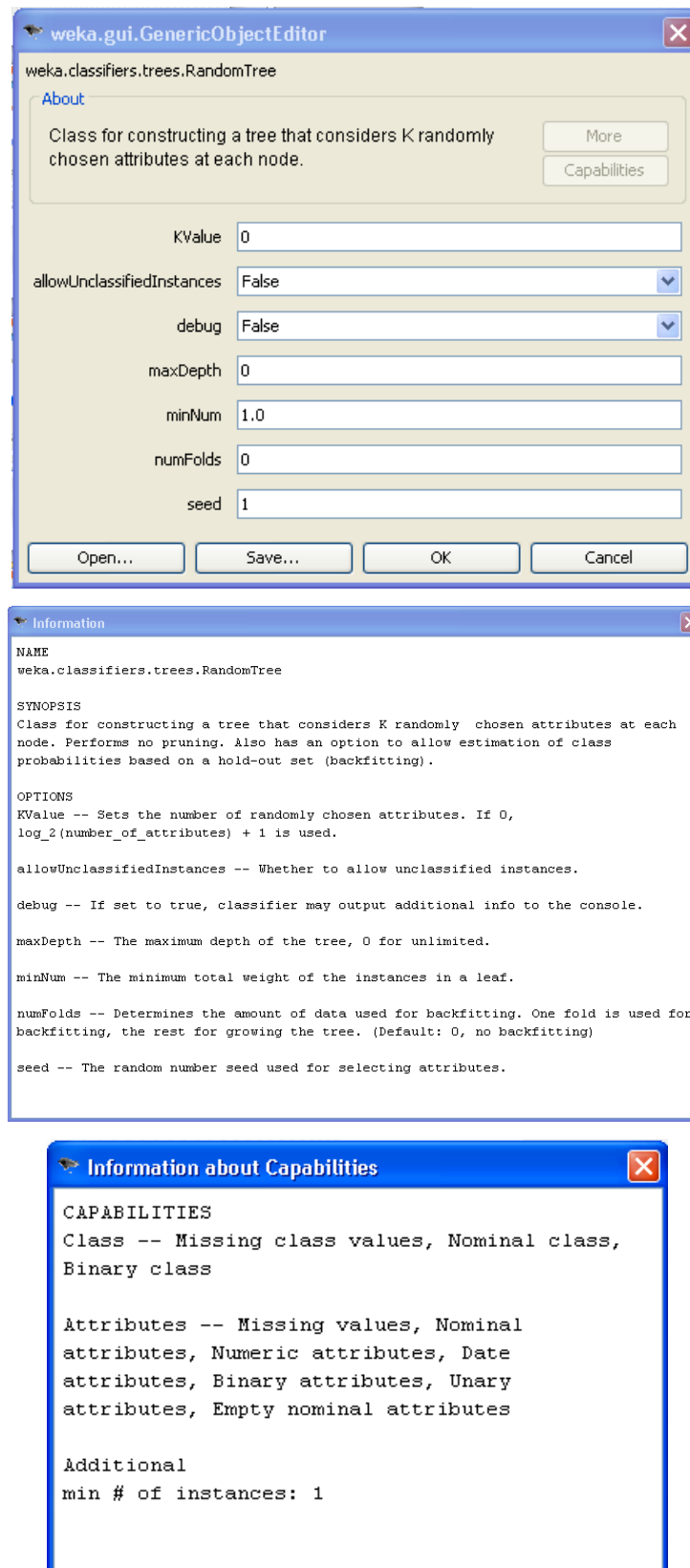
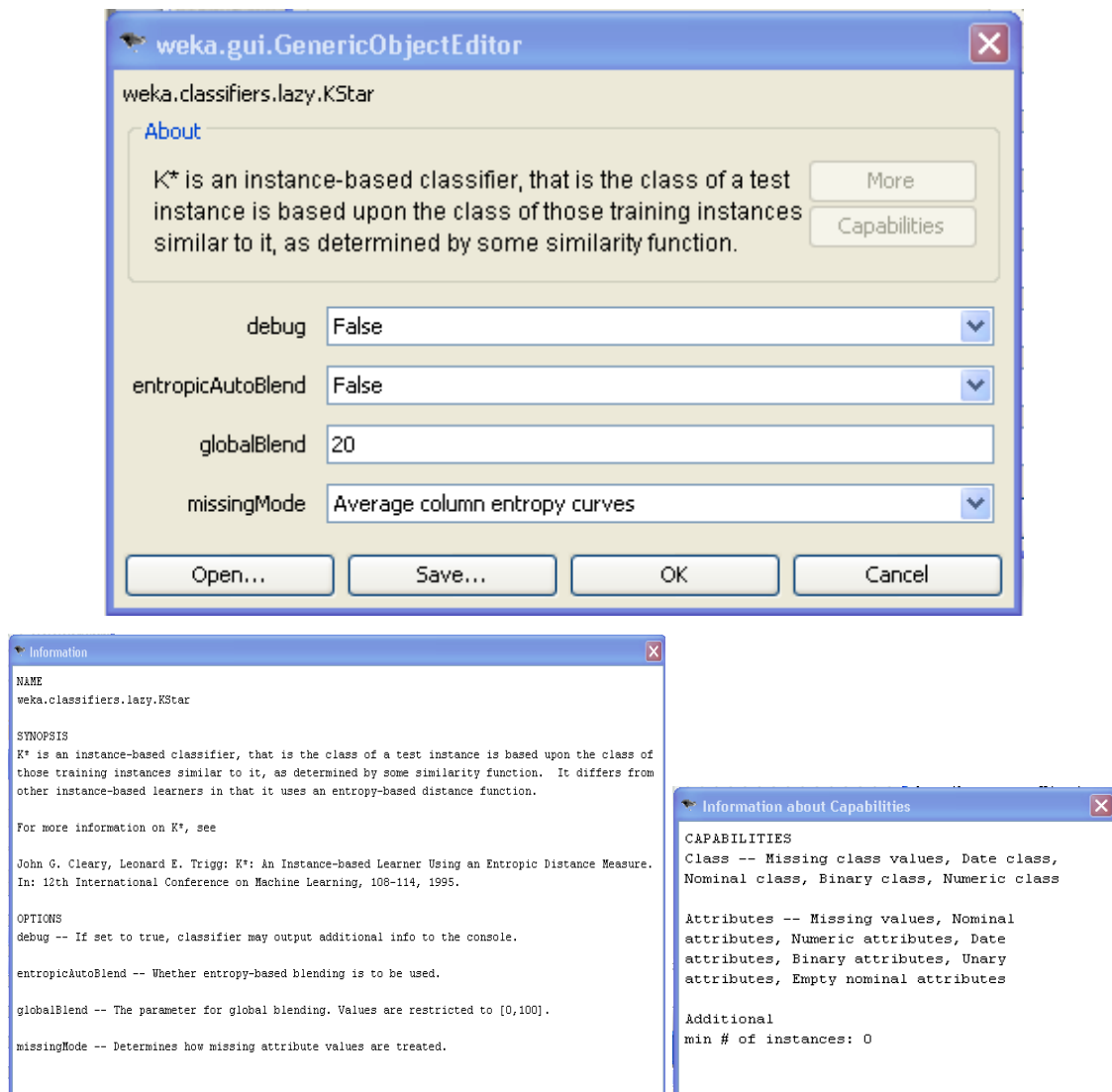


Figura 26: Ayuda WEKA sobre “RandomTree”

## Clasificador2 (CL2) (Figura 27): “weka.classifiers.lazy.KStar”



**Figura 27: Ayuda WEKA sobre “KStar”**

Para más detalle ver manual o ayuda de WEKA para ambas herramientas. La parametrización de los clasificadores se detalla en los resultados del análisis de cada uno de los tres estudios, en base a los cuales se hace una evaluación que atiende principalmente al número de instancias (ventanas de información relativas a cinco minutos de conducción) correctamente clasificadas, la ROC área [Concejero Pedro, 2004] y el estadístico Kappa [Everitt, B., 1992; López et al., 1999; Altman DG, 1991], este último como medida de la fiabilidad desde el aspecto de la concordancia entre observaciones al tratarse de datos de tipo categórico (clases).



En dicho recorrido se definen cinco tipos de carretera, según su exigencia en cuanto a nivel de alerta reclamada, i.e. más nivel de alerta necesaria para C5 que para C1 (Tabla 5):

	<b>RECORRIDO</b>		
	<b>KM</b>	<b>Tipo Carretera</b>	<b>Descripción</b>
<b>Inicio</b>	0,0		
	0,5	C5	Urbana
	3,0	C1	Autovía
	44,0	C2	Nacional 1 carril
	83,5	C1	Autovía
	97,1	C2	Nacional 1 carril
	115,3	C3	Comarcal
	138,1	C4	Comarcal - Montaña
	151,5	C3	Comarcal
	158,0	C5	Urbana
	169,4	C1	Autovía
	210,0	C2	Nacional 1 carril
	212,8	C1	Autovía
<b>Fin</b>	213,3	C5	Urbana

**Tabla 5: Recorrido y Tipos de carretera**

### **3.3. Sujetos participantes**

Los participantes se han seleccionado según dos criterios básicos: Por un lado que sean sanos y no tengan ninguna enfermedad relacionada con trastornos del corazón, y por otro lado que pertenezcan a grupos de edad diferente. Resaltar que ha sido especialmente complicado seleccionar participantes con edad superior a los 55 años, sobre todo teniendo en cuenta el primer criterio y que sean conductores en activo.

Según esto, se analiza un grupo de 38 personas diferentes, todas ellas de sexo “Varón”, las cuales han sido informadas del objetivo del experimento, autorizando la utilización de la información recogida según autorizaciones firmadas por los mismos. La Tabla 6 muestra su identificador (Id) dentro de este estudio:

<b>Id</b>	<b>Nombre:</b>	<b>Apellido1</b>	<b>Apellido2</b>
1	Jesus	Murgoitio	Larrauri
2	Manuel	Marín	Andrés
3	Martin	Etxebarria	Ezpeleta
4	Txomin	Rodriguez	Villa
5	Juan	Nevado	Durán
6	Eneko	Nevado	Murgoitio
7	Alberto Demetrio	DeDiego	Ibáñez
8	Jose Ignacio	Beitia	Sertucha
9	Alberto	DeDiego	Campos
10	Jose Ignacio	Rivada	González
11	Juan Antonio	Mentxaka	Alberdi
12	Javier	Markaida	Markaida
13	Jon Koldo	Barcenilla	Sagardui
14	Joseba	Arruza	Elguezabal
15	Juan Antonio	Diaz	Nevado
16	Manu	Jauregui	Larringan
17	Jose Luis	Zabaleta	LasHeras
18	Javier	Olazar	Gangoiti
19	Pedro	Artetxe	Atucha
20	Jose Maria	Arieta-araunabeña	Bustinza
21	Josu	Barcenilla	Sagardui
22	Aitor	Arin	Bilbao
23	Josu	Jaka	Aspe
24	Alfredo	Madina	Jugo
25	Asier	Arambarri	Beldarrain
26	Ekain	Aguirre	García
27	Xabier	Razkin	Euba
28	Jon Iker	Arambarri	Beldarrain
29	Alain	Bilbao	Abaunza
30	Juan Jose	Amarika	Sagarminaga
31	Victor	Sagardui	Larrondo
32	Pedro	Intxauralde	Etxebarria
33	Jose Antonio	Larringan	Uriarte
34	Benito	Nevado	Durán
35	Francisco Javier	Saez	Martínez
36	Gaspar	Nevado	Durán
37	Jon Andoni	Murgoitio	Larrauri
38	Santos	Goti	Eguiguren

**Tabla 6: Listado resumen de participantes**

Para más detalle sobre los sujetos participantes ver “Anexo 1” incluyendo autorizaciones de los mismos.

### 3.4. Diseño: Variables, Modelo de datos y Casos de estudio

En esta sección se describen las variables recogidas, tratadas y analizadas, así como los distintos casos de estudio en los se ha diseñado y estructurado todo el análisis de esta monografía.

#### 3.4.1. Variables y Modelo de datos

En cuanto a las variables recogidas y analizadas, se ajustan según el modelo de datos de la Figura 29:

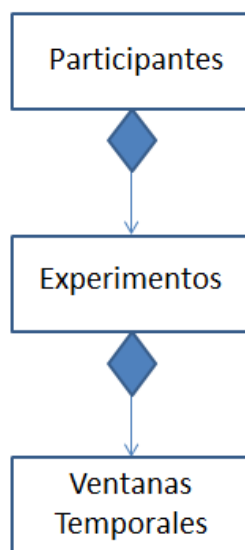


Figura 29: Modelo de datos

De esta manera cada “Participante” se relaciona con los “Experimentos” en los que hubiera intervenido, y cada “Experimento” con las “Ventanas Temporales” (ventanas de cinco minutos de conducción) que contienen las variables, sobre todo relacionadas con información del HRV, tanto en el dominio de la frecuencia como en el dominio del tiempo. Aclarar que dos ventanas consecutivas tienen un “*overlapping*” del 50%. Todas las variables según este modelo se describen a continuación.

- **Variables personales (Participantes)** (Tabla 7): Id asignado al participante, nombre, apellido1, apellido2, fecha de nacimiento, sexo, teléfono, móvil, dirección, población, provincia y DNI.

Variable	Descripción
Id	Id asignado al participante
Nombre	Nombre del participante
Apellido1	Primer apellido
Apellido2	Segundo apellido
Fecha nacimiento	Fecha de nacimiento
Sexo	V=Varón. M=Mujer (Solo se contemplaron varones)
Teléfono	Número de teléfono de contacto
Móvil	Número de móvil de contacto
Dirección	Dirección del participante
Población	Población correspondiente a "Dirección"
Provincia	Provincia correspondiente a "Dirección + Población"
DNI	Documento Nacional de identidad del participante

**Tabla 7: Listado de variables personales**

- **Variables propias del experimento (Experimentos)** (Tabla 8): "Id asignado al participante" según "Variables personales", fecha del experimento, directorio del experimento donde se guarda los archivos del registro, archivo TDMS, archivo excel, ciclo circadiano, grupo de edad, peso, altura, hora de inicio, duración, distancia recorrida, frecuencia de muestreo, *AR order*, banda de frecuencia para VLF, LF y HF, Media y desviación del tiempo entre picos RR, % de señal OK, tiempo de corte inicial, forma física, horas de sueño, tiempo desde la última ingesta.

Variable	Descripción
<b>Id_Participante</b>	Id asignado al participante
<b>Fecha_Experimento</b>	Fecha de cuando se realizó el experimento (AAAAMMDD)
<b>Ciclo_Circadiano</b>	M = Mañana (experimentos realizados entre 06:00 y 12:59). T = Tarde (experimentos realizados entre 13:00 y 19:59). N = Noche (experimentos realizados entre 20:00 y 05:59). (Solo se contemplaron experimentos de noche "M" y "T")
Directorio	Directorio repositorio del experimento.
Archivo_TDMS	Archivo TDMS donde se recoge la señal ECG del experimento.
Archivo_Excel	Archivo Excel donde se recoge la información del experimento.
Edad	Edad (años)
Grupo_Edad	Cuatro grupos de edad: A= hasta 25 años, B= desde 26 hasta 40 años, C= desde 41 hasta 55 años, D: mayores de 55 años.
Peso	Peso en Kg del participante.
Altura	Altura en cm del participante.
Hora_Inicio	Hora de inicio del recorrido.
Duración	Duración del recorrido en minutos.
Distancia_Recorrida	Distancia recorrida = +- 215 Km.

Variable	Descripción
Frecuencia_Muestreo	Frecuencia de muestreo del ECG = 1 KHz.
AR_Order	Orden del filtro AR aplicado = 16.
Banda_VLF	Frecuencia frontera entre VLF y LF = 0,04 Hz.
Banda_LF	Frecuencia frontera entre LF y HF = 0,15 Hz.
Banda_HF	Frecuencia máxima de HF = 0,40 Hz.
Media	Media del intervalo RR en todo el experimento
Desviación_Típica	Desviación típica del intervalo RR en todo el experimento
%_Señal_OK	% del tiempo en que la señal ECG ha sido de buena calidad.
Tiempo_Corte_Inicial	Minutos de registro ECG iniciales no contemplados = 2 min.
Horas_Sueño	Horas de sueño del participante la noche anterior
Tiempo_Ingesta	Tiempo en horas desde la última ingesta.

**Tabla 8: Listado de variables propias del experimento**

- **Variables de cada ventana de cinco minutos (Ventanas Temporales)** (Tabla 9): “Id asignado al participante” + “fecha del experimento” + “ciclo circadiano” según variables propias del experimento”, número de ventana, km de inicio, km de fin, distancia recorrida, tiempo de inicio, velocidad, tipo de tramo, porcentaje de señal OK, tipo de velocidad asignada, tipo de carretera asignada, conjunto de variables en el dominio del tiempo, conjunto de variables en el dominio de la frecuencia.

Variable	Descripción
<b>Id_Participante</b>	Id asignado al participante
<b>Fecha_Experimento</b>	Fecha del experimento
<b>Ciclo_Circadiano</b>	Ciclo circadiano
<b>Ventana</b>	Nº correlativo de ventana de 5 min. desde inicio del recorrido.
<b>Km_Inicio</b>	Km inicio de la ventana.
<b>Km_Fin</b>	Km fin de ventana.
<b>Distancia</b>	Distancia recorrida (Km).
<b>Tiempo_Inicio</b>	Minutos transcurridos desde inicio del recorrido.
<b>Velocidad</b>	Velocidad media en Km/h asignada a la ventana.
<b>Tipo_Tramo</b>	Seis tipos de tramos asignados cada 30 minutos de conducción: A=0' a 30', B=31' a 60', C=61' a 90', D=91' a 120', E=121' a 150', F= desde 151'
<b>%_Señal_OK</b>	Porcentaje del tiempo en que la señal ECG es de buena calidad.
<b>Tipo_Carretera</b>	C1=Autovía, C2=Nacional 1 carril, C3=Comarcal, C4=Comarcal Montaña, C5=Urbana (ver “Descripción del recorrido”).
<b>Conjunto DT</b>	Variables en el dominio del tiempo (ver más adelante).
<b>Conjunto DF</b>	Variables en el dominio de la frecuencia (ver más adelante).

**Tabla 9: Listado de variables propias de cada ventana temporal de cinco minutos**

A continuación se detallan los dos últimos conjuntos (DT y DF).

### Variables en el dominio del tiempo (Conjunto DT):

Dentro del conjunto de variables en el dominio del tiempo, para cada una de las ventanas se calcularon y asignaron los siguientes parámetros o variables de la Tabla 10.

Parámetro	Descripción
Media	Media del tiempo entre ondas R del HRV (ms).
Desviación	Desviación del tiempo entre ondas R del HRV (ms).
SD1	Índice mejorado: SD1 (Ver “Métodos estadísticos”).
SD2	Índice mejorado: SD2 (Ver “Métodos estadísticos”).
SD1/SD2	Ratio de los índices SD1 y SD2.

Tabla 10: Conjunto de variables en el dominio del tiempo para cada cinco minutos

### Variables en el dominio de la frecuencia

Adicionalmente, a cada una de las ventanas se le aplicó el método auto regresivo para conocer el PSD (*Power Spectral Density* o Densidad Espectral de potencia) y poder extraer los parámetros espectrales del HRV. Esta información se puede apreciar en la gráfica de la Figura 30.

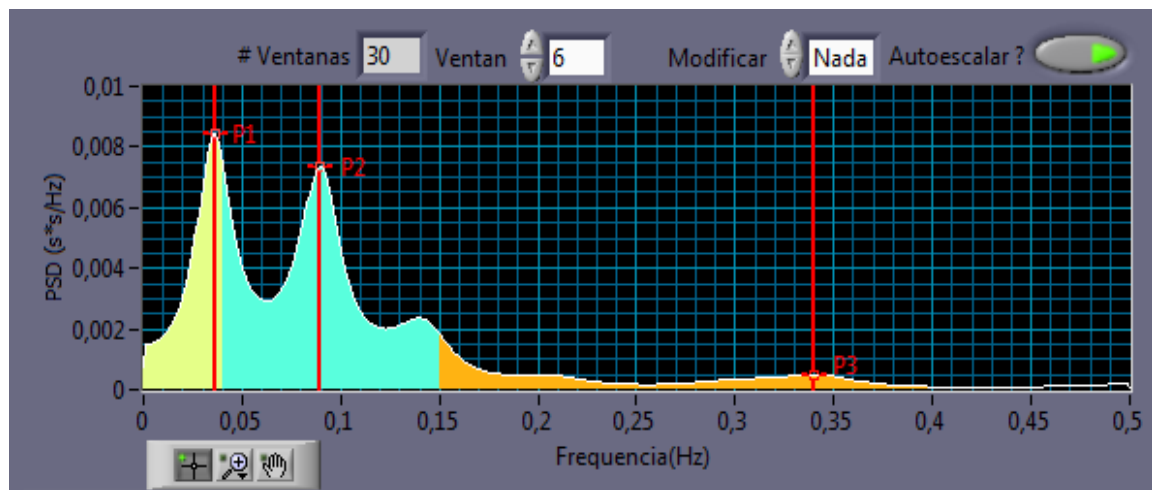


Figura 30: Gráfico del PSD de una ventana de 5 minutos

Los parámetros que definen las bandas de corte entre las frecuencias VLF, LF y HF se definieron según las variables propias del experimento “Banda\_VLF”, “Banda\_LF” y

“Banda\_HF”. Según esto, los parámetros obtenidos en el dominio de la frecuencia, y que completan la base para el análisis del HRV, son los que se muestran en la Tabla 11:

<b>Parámetro</b>	<b>Descripción</b> (VLF=Very Low Frequency, LF=Low Frequency, HF=High Frequency)
<b>VentVLF Peak_X</b>	La abscisa del pico de la frecuencia VLF
<b>VentVLF Peak_Y</b>	La ordenada del pico de la frecuencia VLF
<b>VentLF Peak_X</b>	La abscisa del pico de la frecuencia LF
<b>VentLF Peak_Y</b>	La ordenada del pico de la frecuencia LF
<b>VentHF Peak_X</b>	La abscisa del pico de la frecuencia HF
<b>VentHF Peak_Y</b>	La ordenada del pico de la frecuencia HF
<b>VentVLF Power</b>	El PSD de la frecuencia VLF
<b>VentLF Power</b>	El PSD de la frecuencia LF
<b>VentHF Power</b>	El PSD de la frecuencia HF
<b>VentPorcVLF Power</b>	El PSD en % de la frecuencia VLF
<b>VentPorcLF Power</b>	El PSD en % de la frecuencia LF
<b>VentPorcHF Power</b>	El PSD en % de la frecuencia HF
<b>VentLF/HF</b>	Relación entre el PSD de las frecuencia LF y HF
<b>VentVLF_XC</b>	La abscisa centro de masa del PSD de la frecuencia VLF
<b>VentVLF_YC</b>	La ordenada centro de masa del PSD de la frecuencia VLF
<b>VentLF_XC</b>	La abscisa centro de masa del PSD de la frecuencia LF
<b>VentLF_YC</b>	La ordenada centro de masa del PSD de la frecuencia LF
<b>VentHF_XC</b>	La abscisa centro de masa del PSD de la frecuencia HF
<b>VentHF_YC</b>	La ordenada centro de masa del PSD de la frecuencia HF

**Tabla 11: Conjunto de variables en el dominio de la frecuencia para cada cinco minutos**

### 3.4.2. Casos de estudio

Enmarcado en el objetivo general de analizar la influencia de varios factores de diferente naturaleza para integrarlos y de esta manera optimizar el uso de la señal ECG y la parametrización de algoritmos de monitorización del estado del conductor (nivel de alerta y fatiga) basados en el análisis HRV durante la conducción de vehículos a motor en recorridos largos, el presente trabajo de investigación se estructura en cuatro casos de estudio, cada uno de los cuales encara uno de los cuatro primeros objetivos específicos en los que se ha descompuesto esta monografía. Así, cada denominado caso de estudio define una variable dependiente (objetivo) sobre la que se centra el análisis.

### **Caso de estudio 1: “Edad”**

Con objeto de analizar la influencia de factores **personales (objetivo 1)** sobre la medida del HRV, y seleccionar las variables independientes más relacionadas con este tipo de factores, así como estudiar su comportamiento, se ha escogido la variable “Grupo\_Edad” como variable dependiente para este caso de estudio o estudio.

De esta manera se definen cuatro grupos de edad entre los participantes (A: hasta 25 años inclusive, B: desde 26 años hasta 40 años ambos inclusive, C: desde 41 hasta 55 años ambos inclusive, D: mayores de 55 años) que finalmente se agrupan en dos: hasta 55 años (grupos A, B y C) y mayores de 55 años (grupo D).

### **Caso de estudio 2: “Ciclos circadianos”**

Así mismo, con el objetivo de estudiar la influencia de factores **temporales** sobre el HRV (**objetivo 2**), para detectar las variables independientes más relacionadas con este tipo de factores, así como analizar el comportamiento de las mismas, se ha escogido como variable dependiente para este caso de estudio el parámetro “Ciclo\_Circadiano”.

Según esto, se definen tres valores posibles para esta variable: M (Mañana) para los experimentos realizados entre las 06:00 y las 12:59, T (Tarde) para los que se han llevado a cabo entre las 13:00 y las 19:59, y N (Noche) para el resto de horas de una jornada de 24 horas (desde 20:00 a 05:59). El valor de la variable (M, T o N) se asigna a cada experimento según donde ha transcurrido más del 75 % del tiempo de conducción. Esto se hace porque algunas veces, debido a la necesidad de adaptarse a la disponibilidad de los participantes, algunas de las últimas ventanas del recorrido pertenecían al ciclo circadiano siguiente en el que comenzaron (principalmente experimentos de tarde (T) que terminaban algo más tarde de 19:59. El caso de estudio se ha centrado en los ciclos M y T.

### **Caso de estudio 3: “Tipo de carretera”**

Con objeto de estudiar el comportamiento del HRV debido a la influencia de factores **ambientales (objetivo 3)**, filtrando previamente las variables independientes más

relacionadas con este tipo de factores, y analizándolas posteriormente, se ha escogido la variable “Tipo\_Carretera” como variable dependiente para este caso de estudio.

De esta manera se definen cinco tipos de carretera (Tabla 12), según su exigencia en cuanto a nivel de alerta reclamada, i.e. más nivel de alerta necesaria para C5 que para C1 (para conocer el tipo de carretera asignado a cada tramo del recorrido ver “Descripción del recorrido”):

<b>Tipo Carretera</b>	<b>Descripción</b>	<b>Nivel de alerta (*)</b>
C1	Autovía	- -
C2	Nacional 1 carril	-
C3	Comarcal	+
C4	Comarcal - Montaña	++
C5	Urbana	++

(\*) Nivel de alerta requerido por el tipo de vía: (- - => mínimo), (++ => máximo).

**Tabla 12: Tipos de carretera**

Según esta primera asignación del tipo de carretera en función de la distancia recorrida desde el inicio, debido a que cinco minutos de conducción podían transcurrir por diferentes tipos de carretera, para cada ventana de cinco minutos de conducción se calcula el porcentaje de tiempo que se conduce por cada tipo de carretera y se le asigna el tipo de carretera dominante, es decir, aquella que supera el 75 % del tiempo.

Finalmente, con el objetivo de resaltar el nivel de alerta inferido a los diferentes parámetros recogidos, a partir de los cinco tipos definidos anteriormente se seleccionan solo aquellos que requieran niveles de alerta muy diferentes entre sí, es decir, por una parte todas las ventanas cuyo “Tipo de carretera” sea C1 (mínimo nivel de alerta necesario), y la agrupación de las ventanas con C4 o C5 como tipo asignado (máximos niveles de alerta necesario), prescindiendo de aquellas cuyo tipo asignado es C2 o C3.

#### **Caso de estudio 4: “Fatiga”**

Finalmente, orientado a optimizar el uso del HRV como marcador de fatiga durante la conducción prolongada de vehículos motorizados, integrando todos los factores (personales, temporales y ambientales) que en los casos de estudio anteriores han

demostrado ser significativos (**objetivo 4**), se ha escogido la variable “Tipo\_Tramo” como variable dependiente para este caso de estudio.

Según esto, se divide el recorrido en seis tramos (uno cada 30’) en función del tiempo transcurrido desde el comienzo del experimento y se asigna a la variable “Tipo\_Tramo” de cada ventana temporal un valor (A-F): A = de 0’ a 30’, B = de 31’ a 60’, C = de 61’ a 90’, D = de 91’ a 120’, E = de 121’ a 150’, F = desde 151’. Este parámetro “Tipo\_Tramo” es el que se ha escogido como variable dependiente para este caso de estudio. Para más detalles sobre el análisis y resultados recogidos para este caso de estudio ver capítulo correspondiente en “Análisis y Resultados”.

## 4. ANÁLISIS Y RESULTADOS

A continuación se detallan el análisis y resultados recogidos para cada uno de los casos de estudio en los que se estructura esta monografía según el siguiente esquema común para todos ellos:

- 1- Análisis en el dominio del tiempo (DT):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 2- Análisis en el dominio de la frecuencia (DF):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 3- Análisis combinado (DC):
  - a. Selección de atributos: según selecciones en DT y DF.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 4- Conclusiones.

**Selección de atributos - Descripción:** En la selección de atributos, tal y como se indica en el capítulo “Minería de datos” de la metodología, se utiliza como evaluador a “weka.attributeSelection.CfsSubsetEval” y como método de búsqueda a “weka.attributeSelection.ExhaustiveSearch” (ver manual o ayuda de WEKA para más detalle sobre estas dos herramientas). En base a ello, una vez definidas las variables independientes y la correspondiente variable dependiente para cada uno de los tres estudios (DT, DF y DC) en cada caso de estudio, se obtiene un cuadro resumen de resultados como el de la Figura 31, en el que se establece un **ranking de atributos** según el criterio de agrupamiento que se adjunta. El análisis posterior con clasificadores se hace según este criterio de agrupación de basado en el nivel de correlación (0-100 %).

```

=== Run information ===
Evaluator: weka.attributeSelection.CfsSubsetEval
Search: weka.attributeSelection.ExhaustiveSearch
Relation: ExpGrupoEdadPer-weka.filters.unsupervised.attribute.StringToNominal-R2-
Instances: 2689
Attributes: 9
VentMedialNorm
VentDesviaciónS
VentMedia
VentSD1(ms)
VentSD2(ms)
VentSD1SD2_Rati
Vent_DS1_Norm
Vent_DS2_Norm
ExpGrupoEdadPer
Evaluation mode: 10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===
number of folds (%) attribute
0( 0 %) 1 VentMedialNorm
4( 40 %) 2 VentDesviaciónS
1( 10 %) 3 VentMedia
10(100 %) 4 VentSD1(ms)
7( 70 %) 5 VentSD2(ms)
0( 0 %) 6 VentSD1SD2_Rati
2( 20 %) 7 Vent_DS1_Norm
0( 0 %) 8 Vent_DS2_Norm

```

**Grupos de correlación (máx X=100%)**

**Grupo 1: X = 100%** (e.g. VentSD1)

**Grupo 2: >= 50 X <100%** (e.g. VentSD2)

**Grupo 3: = > 0 X <50%** (e.g. VentMedia)

**Grupo 4: X = 0%** (e.g. VentSD1SD2\_Rati)

**Figura 31: Ejemplo de resultado en la selección de atributos y ranking de los mismos**

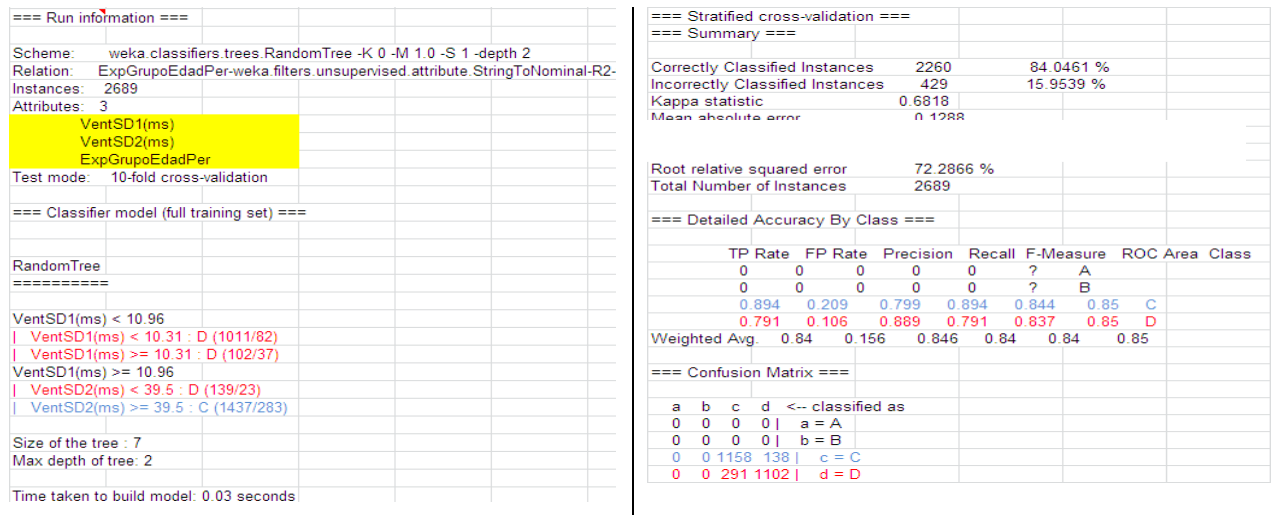
La Tabla 13 describe el significado de los datos que se reporta en esta etapa como resultado en cada uno de los tres análisis (DT, DF, DC):

Información	Descripción (para más detalle ver manual de Weka 3.6.10)
<b>Evaluator</b>	Evalgador utilizado.
<b>Search</b>	Método de búsqueda y selección de parámetros más correlacionados.
<b>Relation</b>	Atributo seleccionado como objetivo (variable dependiente).
<b>Instances</b>	Número de instancias evaluadas = Nº de ventanas de 5' de conducción, i.e. 2689.
<b>Attributes</b>	Variabes independientes evaluadas: Números de las mismas, i.e. 6, y listado de las mismas con fondo amarillo.
<b>Evaluation mode</b>	Método de evaluación = Validación cruzada en base a 10 grupos.
<b>Attribute selection</b>	Ranking de atributos según el cual se ha llevado a cabo el análisis posterior con clasificadores. Solo se analizaron los atributos más correlacionados con la variable dependiente, i.e. aquellos que aparecen con % más altos y que se indican con color rojo según la Figura 31. En el "paso1" se indican los atributos en rojo como los más correlacionados (100%). En el paso siguiente (paso 2) se vuelve a repetir la evaluación los atributos, esta vez sin los parámetros seleccionados en el paso anterior (paso1).

**Tabla 13: Información resultante en la selección de atributos**

**Clasificación - Descripción:** En la etapa de clasificación, se utilizan dos clasificadores con estrategias bien diferentes: RandomTree y Lazy KStar. A continuación se describen los datos resumen de salida de cada uno de estas dos técnicas de inferencia del conocimiento en el campo de la minería de datos, y sobre los que se lleva a cabo el análisis posterior para cada uno de los casos de estudio identificados por su variable dependiente o clase objetivo.

La Figura 32 muestra los datos de salida que se generan como resultado del clasificador 1 (**RandomTree**) en cada uno de los tres análisis (DT, DF, DC):



**Figura 32: Ejemplo de resultado del clasificador 1 “RandomTree”**

La Tabla 14 describe el significado de los datos que se reporta en esta etapa como resultado del clasificador 1 (**RandomTree**) en cada uno de los tres análisis (DT, DF, DC):

Información	Descripción (para más detalle ver manual de Weka 3.6.10)
<b>Scheme</b>	Comando Weka parametrizado.
<b>Relation</b>	Atributo seleccionado como objetivo o variable dependiente, i.e. “ExpGrupoEdad”= Grupo_Edad.
<b>Instances</b>	Número de instancias evaluadas = Nº de ventanas de 5’ de conducción.
<b>Attributes</b>	Variables independientes evaluadas: Números de las mismas, i.e. 3, y listado de las mismas con fondo amarillo.
<b>Test mode</b>	Método de evaluación = Validación cruzada en base a 10 grupos.
<b>RandomTree</b>	Arbol “RandomTree”: en azul Grupo_Edad = C, en rojo Grupo_Edad = D.
<b>Summary</b>	Resumen de parámetros del rendimiento de la clasificación.
<b>Detailed accuracy by class</b>	Detalle por clase según “Relation” del rendimiento de la clasificación.
<b>Confusion Matrix</b>	Matriz de confusión.

**Tabla 14: Información resultante con clasificador 1 “RandomTree”**



**Método de análisis:** Según las descripciones anteriores, durante la etapa de **selección de atributos** se establece un ranking de atributos según figura y descripciones en Figura 31.

Tras la etapa de selección de atributos, y en base al ranking de variables independientes obtenido anteriormente, se lleva a cabo el análisis de las mismas en base al **clasificador “RandomTree”** generando el correspondiente árbol de decisión, y siguiendo principalmente el criterio de máxima “*ROC Area*” [Concejero Pedro, 2004] con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), según figura 32 y tabla 14.

De igual manera, en base al mismo ranking de variables independientes, se realiza el análisis de las mismas en base al **clasificador “Lazy KStar”** generando el correspondiente cuadro resumen según figura 33 y tabla 15, y siguiendo el criterio de máxima “*ROC Area*” como principal método de evaluación [Fawcett T., 2004].

#### 4.1. Caso de estudio 1: Factores personales (Edad)

Este caso de estudio se caracteriza por tener como variable dependiente la edad, para el que se han seleccionado los recorridos de 12 participantes mayores de 55 años (grupo D) y 12 participantes de menos de 55 años (Grupos A-C) a los que se les ha informado del experimento a realizar, tal y como se describe en los sub-apartados del capítulo de “Metodología”.

Los participantes seleccionados para este análisis se muestran en la Tabla 16:

<b>Id</b>	<b>Nombre</b>	<b>Apellido 1</b>	<b>Apellido 2</b>	<b>Grupo Edad</b>
1	Jesus	Murgoitio	Larrauri	C
2	Manuel	Marin	Andrés	C
8	Jose Ignacio	Beitia	Sertucha	C
9	Alberto	DeDiego	Campos	D
10	Jose Ignacio	Rivada	González	C
14	Joseba	Arruza	Elguezabal	C
16	Manu	Jauregui	Larringan	D
19	Pedro	Artetxe	Atucha	D
20	Jose María	Arieta-araunabeña	Bustinza	D
21	Josu	Barcenilla	Sagardui	C
22	Aitor	Arin	Bilbao	C
23	Josu	Jaka	Aspe	C
25	Asier	Aranbarri	Beldarrain	C
26	Ekain	Aguirre	García	C
27	Xabier	Razkin	Euba	C
29	Alain	Bilbao	Abaunza	C
30	Juan Jose	Amarika	Sagarminaga	D
31	Víctor	Sagarduy	Larrondo	D
32	Pedro	Intxauralde	Etxebarria	D
33	Jose Antonio	Larringan	Uriarte	D
34	Benito	Nevado	Durán	D
35	Francisco Javier	Saez	Martínez	D
36	Gaspar	Nevado	Durán	D
38	Santos	Goti	Eguiguren	D

**Tabla 16: Participantes seleccionados para el estudio de la edad**

Como resultado del experimento se obtiene un registro ECG por participante correspondiente a su trayecto de conducción prolongada durante alrededor de tres horas, que tras las etapas uno a seis de la metodología han generado 2.689 ventanas de información correspondiendo cada una de ellas a cinco minutos de conducción de un participante, y suponiendo en total más de 112 horas de conducción.

A continuación se detallan el análisis y resultados recogidos para el estudio de la edad, caracterizado por tener la variable “Grupo\_Edad” como variable dependiente, según el siguiente esquema:

- 1- Análisis en el dominio del tiempo (DT):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 2- Análisis en el dominio de la frecuencia (DF):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 3- Análisis combinado (DC):
  - a. Selección de atributos: según selecciones en DT y DF.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 4- Conclusiones.

#### **4.1.1. Análisis en el dominio del tiempo (DT)**

El análisis en el dominio del tiempo para el caso de estudio 1 se caracteriza por definir como variable dependiente a la variable “**Grupo\_Edad**” y como variables independientes a las variables en el dominio del tiempo denominadas como “**Conjunto DT**” en el apartado “Variables y Modelo de datos”, i.e. Media, Desviación, SD1 y SD2. Derivado de ellos también se consideran el ratio SD1/SD2, así como las medidas normalizadas de la Media, SD1 y SD2 (8 variables independientes).

## Selección de atributos:

Según lo anteriormente expuesto, en la etapa de selección de atributos se obtienen los resultados que se reflejan en la Figura 34:



**Figura 34: Ranking de atributos en el DT (Edad)**

De esta manera, según la Figura 34, en la etapa de selección de atributos se ha valorado la variable SD1 (VentSD1=100%) como claramente la más correlacionada con “Grupo\_Edad” (véase en color rojo en la Figura 34), seguida de SD2 (VentSD2=70%), la desviación (VentDesviaciónS=40%), SD1 normalizada (Vent\_SD1\_Norm=20%) y Media (VentMedia=10%).

## Resultados del clasificador 1: “RandomTree”

Basado en el ranking de variables independientes obtenido anteriormente se ha llevado a cabo el análisis de las mismas en base al clasificador “RandomTree” generando el correspondiente árbol de decisión, y siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), seleccionándose como mejor resultado el de la Figura 35:

=== Run information ===		=== Stratified cross-validation ===					
=== Summary ===							
Scheme:	weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1 -depth 2	Correctly Classified Instances	2260 84.0461 %				
Relation:	ExpGrupoEdadPer-weka.filters.unsupervised.attribute.StringToNominal-R2	Incorrectly Classified Instances	429 15.9539 %				
Instances:	2689	Kappa statistic	0.6818				
Attributes:	3	Mean absolute error	0.1288				
	VentSD1(ms)						
	VentSD2(ms)						
	ExpGrupoEdadPer						
Test mode:	10-fold cross-validation						
=== Classifier model (full training set) ===							
RandomTree							
=====							
VentSD1(ms) < 10.96							
VentSD1(ms) < 10.31 : D (1011/82)							
VentSD1(ms) >= 10.31 : D (102/37)							
VentSD1(ms) >= 10.96							
VentSD2(ms) < 39.5 : D (139/23)							
VentSD2(ms) >= 39.5 : C (1437/283)							
Size of the tree :	7						
Max depth of tree :	2						
Time taken to build model :	0.03 seconds						
		Total Number of Instances 2689					
		=== Detailed Accuracy By Class ===					
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0	0	0	0	?	?	A
	0	0	0	0	?	?	B
	0.894	0.209	0.799	0.894	0.844	0.85	C
	0.791	0.106	0.889	0.791	0.837	0.85	D
Weighted Avg.	0.84	0.156	0.846	0.84	0.84	0.85	
		=== Confusion Matrix ===					
	a	b	c	d	<-- classified as		
	0	0	0	0	a = A		
	0	0	0	0	b = B		
	0	0	1158	138	c = C		
	0	0	291	1102	d = D		

**Figura 35: Resultado del clasificador “RandomTree” en el DT (Edad)**

Tal y como se puede apreciar, únicamente con información de las variables SD1 y SD2 se consigue clasificar correctamente al 84,05% (2.260) de las instancias (ventanas temporales de 5’) de un total de 2.689, alcanzando una **ROC área del 85%** tanto para las instancias del grupo de edad C (hasta 55 años) como las del grupo D (mayores de 55 años).

Así mismo, la simpleza del árbol de decisión obtenido (*Size of the tree = 7, Max depth of tree = 2*) indican claramente un comportamiento muy diferenciado de estas variables para los dos grupos de edad: La variabilidad detectada según **SD1** en el grupo de mayor edad (D) es sensiblemente inferior a la detectada para el grupo de edad hasta los 55 años (C).

Además, según el árbol de decisión generado, allí donde podría haber dudas entre el comportamiento del HRV entre los grupos de diferente edad, la variable **SD2** termina por optimizar la diferenciación entre los grupos. El grupo de instancias del grupo C conserva mayores valores de SD2 que los del grupo D.

En resumen, según la información obtenida a partir del análisis de la clasificación obtenida con el clasificador “RandomTree”, la simple información del SD1 y del SD2 como variables independientes en el dominio del tiempo conforman un sistema simple y valioso como marcadores estrechamente ligados a una variable dependiente como la edad y representativa de un factor personal.

## Resultados del clasificador 2: “Entropía”

De forma similar, y en base al mismo ranking de variables, se ha realizado el análisis del comportamiento de las mismas en base al segundo clasificador “Lazy KStar”, cuya inferencia se sustenta en el concepto de la entropía (Cleary, 1995), generando como mejor resultado el de la Figura 36, siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*):

=== Run information ===		=== Stratified cross-validation ===	
=== Summary ===			
Scheme:	weka.classifiers.lazy.KStar -B 20 -M a	Correctly Classified Instances	2321 86.3146 %
Relation:	ExpGrupoEdadPer-weka.filters.unsupervised.attribute.StringToNominal-R2	Incorrectly Classified Instances	368 13.6854 %
Instances:	2689	Kappa statistic	0.7268
Attributes:	3	Mean absolute error	0.1176
	VentSD1(ms)		
	VentSD2(ms)		
	ExpGrupoEdadPer		
Test mode:	10-fold cross-validation	Total Number of Instances	2689
=== Classifier model (full training set) ===		=== Detailed Accuracy By Class ===	
KStar Beta Verion (0.1b)		TP Rate	FP Rate
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz)		Precision	Recall
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz)		F-Measure	ROC Area
KStar options : -B 20 -M a			Class
Time taken to build model: 0 seconds		0	0
		0	0
		0.904	0.174
		0.826	0.096
		0.828	0.904
		0.864	0.862
		0.932	0.932
		0.863	0.134
		0.866	0.866
		0.863	0.863
		0.863	0.932
=== Confusion Matrix ===			
a	b	c	d
0	0	0	0
0	0	0	0
0	0	1171	125
0	0	243	1150

Figura 36: Resultado del clasificador “Lazy KStar” en el DT (Edad)

Una vez más, únicamente con información de las variables SD1 y SD2 se consigue clasificar correctamente el 86,31% (2.321) de las ventanas temporales de 5’ del total de 2.689, alcanzando una **ROC área del 93,2%** para ambos tipos de instancias: las del grupo C (hasta 55 años) y las del grupo D (mayores de 55 años).

Así mismo, la **matriz de confusión** muestra una clara diagonal principal y señala claramente un comportamiento muy diferenciado de estas dos variables según el concepto de entropía manejado por el clasificador para los dos grupos de edad, con un rendimiento ligeramente superior para el ratio TP o “True-Positives” (90,4%) del grupo C que para el grupo D (82,6%).

En definitiva, según un criterio de clasificación completamente diferente al anterior, una vez más la simple información del SD1 y del SD2 conforman un sistema muy significativo como marcadores ligados a la variable dependiente de la edad (representativa de factores personales para esta monografía), todo ello en base a la información generada con el

clasificador “*Lazy KStar*”, esta vez incluso con rendimientos ligeramente superiores: ROC área del 85,0% con el “*RandomTree*” frente al 93,2% con el “*Lazy KStar*”.

#### 4.1.2. Análisis en el dominio de la frecuencia (DF)

El análisis en el dominio de la frecuencia para el caso de estudio 1 se caracteriza por definir como variable dependiente a la variable “**Grupo\_Edad**” y como variables independientes a aquellas en el dominio de la frecuencia denominadas como “**Conjunto DF**” en el apartado “Variables y Modelo de datos”. Derivadas de ellas también se consideran algunas medidas normalizadas de las mismas (hasta 19 variables independientes).

#### Selección de atributos:

Así, en la etapa de selección de atributos se obtienen los resultados de la Figura 37:

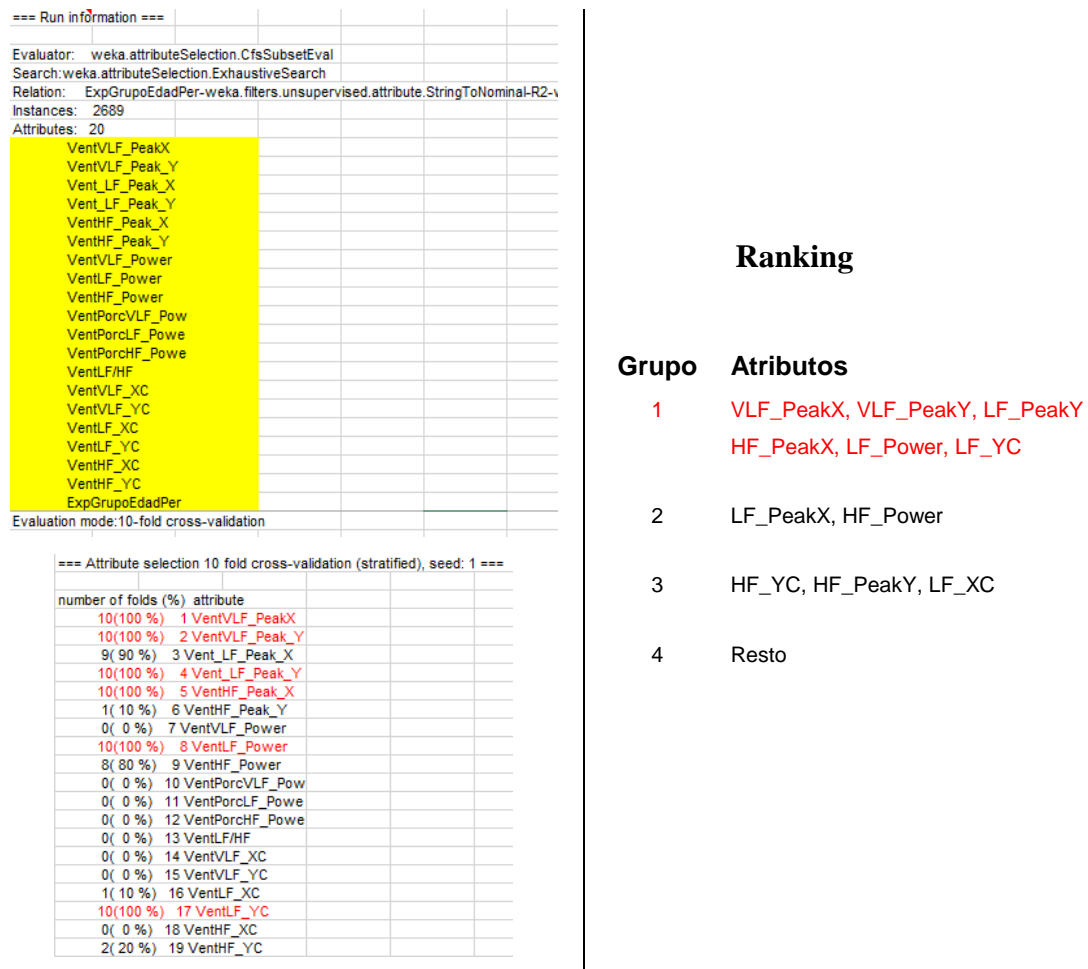


Figura 37: Ranking de atributos en el DF (Edad)

En base a estos resultados, y como se verá en el análisis con clasificadores que sigue, sería suficiente la información del PSD para LF para conseguir buenos rendimientos.

### Resultados del clasificador 1: “RandomTree”

Basado en las variables anteriores se analizan las PSD con el clasificador “RandomTree”, se genera el correspondiente árbol de decisión, y siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), se selecciona como mejor resultado el de la Figura 38:

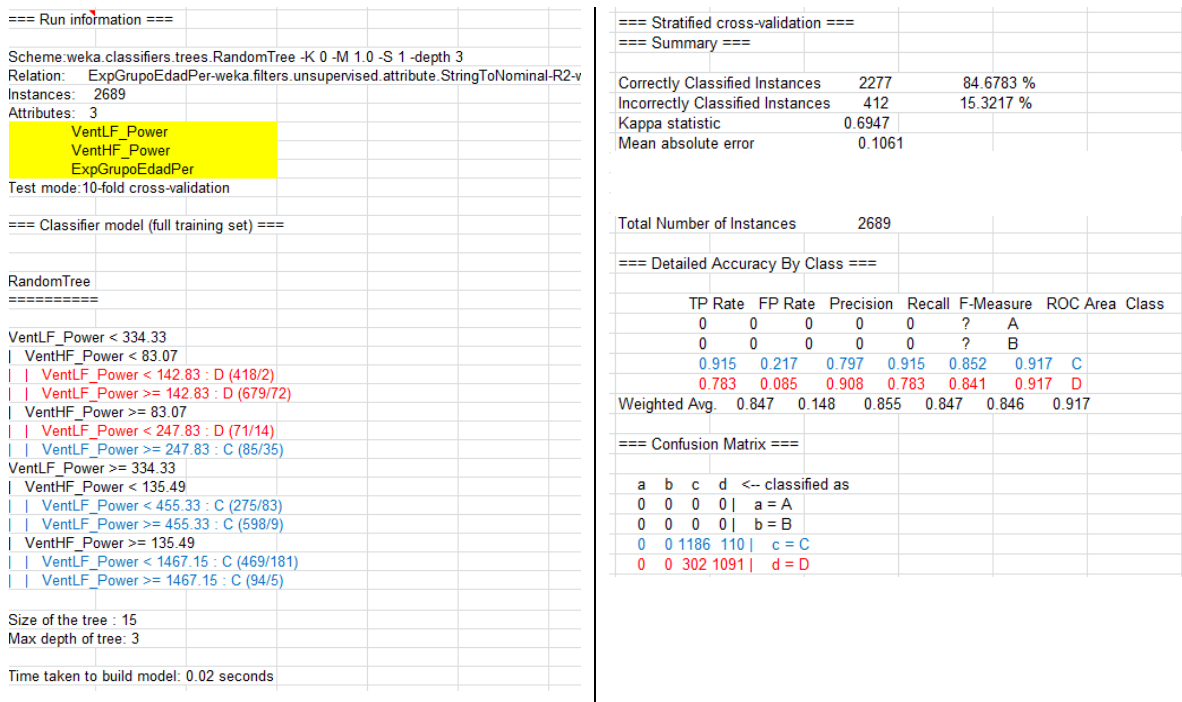


Figura 38: Resultado del clasificador “RandomTree” en el DF (Edad)

Según este resultado, solo es necesario tener información del “Power” en las frecuencias LF y HF, para clasificar correctamente al 84,68% (2.277) de las instancias (ventanas temporales de 5’) de un total de 2.689, alcanzando en su evaluación una **ROC área del 91,7%** para las instancias del grupo de edad C (hasta 55 años y de color azul en el *RandomTree* generado) y las del grupo D (mayores de 55 años y de color rojo en el *RandomTree* generado).

Así mismo, la simpleza del árbol de decisión obtenido (*Size of the tree* = 15, *Max depth of tree* = 3) indican claramente un comportamiento muy diferenciado de estas variables para los dos grupos de edad: La LF\_Power detectada para el grupo de mayor edad (D) es

sensiblemente inferior a la detectada para el grupo de edad hasta los 55 años (C). Además, según el árbol de decisión generado, podría evitarse el uso de HF (Power), sin gran pérdida del rendimiento del clasificador.

En resumen, según la información obtenida a partir del análisis de la clasificación obtenida con el clasificador “*RandomTree*”, la simple información del PSD (LF+HF) conforma un sistema simple y valioso estrechamente ligado a la edad, variable representativa de un factor personal. Este sistema podría incluso limitarse a la LF (Power) sin gran pérdida de rendimiento (ROC área = 83’8 %) y clasificando correctamente al 83,00% (2.232) de las instancias del total de 2.689 (ventanas temporales de 5’). Esta pérdida de rendimiento es achacable a la frontera definida entre HF y LF, que como puede apreciarse en la Figura 38, podría ser otra y alcanzarse rendimientos de clasificación mejores solo con información del LF.

### Resultados del clasificador 2: “Entropía”

Igualmente, y en base al mismo conjunto de variables, se realiza el análisis del comportamiento de las mismas en base al segundo clasificador “*Lazy KStar*”, generando la Figura 39 como mejor resultado (máxima “*ROC Area*” y mínimo número de atributos):

=== Run information ===		=== Stratified cross-validation ===		
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===		
Relation: ExpGrupoEdadPer-weka.filters.unsupervised.attribute.StringToNominal-R2		Correctly Classified Instances	2342 87.0956 %	
Instances: 2689		Incorrectly Classified Instances	347 12.9044 %	
Attributes: 3		Kappa statistic	0.742	
VentLF_Power		Mean absolute error	0.0987	
VentHF_Power				
ExpGrupoEdadPer				
Test mode: 10-fold cross-validation				
=== Classifier model (full training set) ===		Total Number of Instances	2689	
KStar Beta Verion (0.1b).		=== Detailed Accuracy By Class ===		
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).		TP Rate	FP Rate	
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).		Precision	Recall	
KStar options : -B 20 -M a		F-Measure	ROC Area	
Time taken to build model: 0 seconds		Class		
		0	0	
		0	0	
		0.89	0.147	
		0.853	0.11	
		0.849	0.89	
		0.893	0.853	
		0.869	0.873	
		0.944	0.944	
		Weighted Avg.	0.871	
			0.128	
			0.872	
			0.871	
			0.871	
			0.944	
=== Confusion Matrix ===				
a	b	c	d	<- classified as
0	0	0	0	a = A
0	0	0	0	b = B
0	1154	142		c = C
0	0	205	1188	d = D

Figura 39: Resultado del clasificador “*Lazy KStar*” en el DF (Edad)

Tal y como se puede apreciar en la Figura 39, únicamente con información del PSD para LF + HF se logra clasificar correctamente el 87,10% (2.342) de las ventanas temporales de 5' del total de 2.689, alcanzando una **ROC área del 94,4%**.

Así mismo, la **matriz de confusión** muestra una clara diagonal principal y señala claramente un comportamiento muy diferenciado de estas dos variables según el concepto de entropía manejado por el clasificador para los dos grupos de edad, con un rendimiento algo superior para el ratio TP o “*True-Positives*” (89,0%) del grupo D que para el grupo D (85,3%).

Como consecuencia, en base a un criterio de clasificación completamente diferente al anterior, una vez más la información basada en la PSD en las frecuencias LF + HF forman un sistema suficientemente significativo o correlacionado a la edad (representativa de factores personales para esta monografía), incluso con rendimientos algo superiores a los obtenidos con el clasificador anterior: ROC área del 91,7% con el “*RandomTree*” frente al 94,4% con el “*Lazy KStar*”.

#### **4.1.3. Análisis combinado (DC)**

El análisis al combinar variables en el dominio del tiempo y la frecuencia para el caso de estudio 1 se caracteriza por tener como variable dependiente a la variable “**Grupo\_Edad**” y como variables independientes aquellas que según los estudios DT y DF anteriores han mostrado mayor correlación con “Grupo\_Edad”, i.e. SD1 y SD2 para DT junto a LF y HF (Power) para DF, las cuales son la base para el análisis combinado de este caso de estudio con clasificadores.

#### **Resultados del clasificador 1: “RandomTree”**

Basado en la información de variables anterior se lleva a cabo el análisis con el clasificador “*RandomTree*” generándose el correspondiente árbol de decisión, y siguiendo el criterio de máxima “*ROC Area*” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), seleccionándose como mejor resultado el de la Figura 40:

=== Run information ===		=== Stratified cross-validation ===					
Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1 -depth 2		=== Summary ===					
Relation: ExpGrupoEdadPer-weka.filters.unsupervised.attribute.StringToNominal-R2-v		Correctly Classified Instances	2298 85.4593 %				
Instances: 2689		Incorrectly Classified Instances	391 14.5407 %				
Attributes: 3		Kappa statistic	0.7093				
VentLF_Power		Mean absolute error	0.1196				
VentSD1(ms)							
ExpGrupoEdadPer							
Test mode: 10-fold cross-validation							
=== Classifier model (full training set) ===							
RandomTree							
=====							
VentLF_Power < 334.33							
VentLF_Power < 142.83 : D (427/2)							
VentLF_Power >= 142.83 : D (826/136)							
VentLF_Power >= 334.33							
VentSD1(ms) < 10.22 : D (78/25)							
VentSD1(ms) >= 10.22 : C (1358/225)							
Size of the tree: 7							
Max depth of tree: 2							
Time taken to build model: 0.01 seconds							
		Total Number of Instances	2689				
		=== Detailed Accuracy By Class ===					
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0	0	0	0	?	A	
	0	0	0	0	?	B	
	0.873	0.163	0.833	0.873	0.853	0.867	C
	0.837	0.127	0.877	0.837	0.856	0.867	D
	Weighted Avg.	0.855	0.144	0.856	0.855	0.855	0.867
		=== Confusion Matrix ===					
	a	b	c	d	← classified as		
	0	0	0	0	a = A		
	0	0	0	0	b = B		
	0	0	1132	164	c = C		
	0	0	227	1166	d = D		

**Figura 40: Resultado del clasificador “RandomTree” en el DC (Edad)**

Según este resultado, solo haría falta tener información del “Power” en la frecuencia LF, además de SD1, para clasificar correctamente al 85,46% (2.298) de las instancias (ventanas temporales de 5’) de un total de 2.689, alcanzando en su evaluación una **ROC área del 86,7%** para las instancias del grupo de edad C (hasta 55 años y de color azul en el *RandomTree* generado) y las del grupo D (mayores de 55 años y de color rojo en el *RandomTree* generado).

Así mismo, se obtiene un árbol de decisión muy simple (*Size of the tree = 7, Max depth of tree = 2*) indicando un comportamiento muy diferenciado de estas variables para los dos grupos de edad.

En resumen, según la información obtenida a partir del análisis de la clasificación obtenida con el clasificador “*RandomTree*”, la información del PSD de LF junto a SD1 conforman un sistema simple como marcadores ligados a la edad y representativa de un factor personal.

## Resultados del clasificador 2: “Entropía”

De forma similar, y en base a la misma base de variables, se realiza el análisis del comportamiento de las mismas en base al segundo clasificador “*Lazy KStar*” según la

entropía (Cleary, 1995), generando como mejor resultado el de la Figura 41 con la máxima “ROC Area” obtenida con el mínimo número de atributos (*Attributes*):

=== Run information ===		Correctly Classified Instances		2353		87.5046 %		
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		Incorrectly Classified Instances		336		12.4954 %		
Relation: ExpGrupoEdadPer-weka.filters.unsupervised.attribute.StringToNominal-R2-		Kappa statistic		0.7503				
Instances: 2689		Mean absolute error		0.1026				
Attributes: 3								
VentLF_Power								
VentSD1(ms)								
ExpGrupoEdadPer								
Test mode: 10-fold cross-validation								
=== Classifier model (full training set) ===								
KStar Beta Verion (0.1b).								
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).								
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).								
KStar options : -B 20 -M a								
Time taken to build model: 0 seconds								
		Total Number of Instances		2689				
		=== Detailed Accuracy By Class ===						
		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
		0	0	0	0	?		A
		0	0	0	0	?		B
		0.898	0.146	0.851	0.898	0.874	0.939	C
		0.854	0.102	0.9	0.854	0.876	0.939	D
		Weighted Avg.	0.875	0.123	0.876	0.875	0.875	0.939
		=== Confusion Matrix ===						
		a	b	c	d	<-- classified as		
		0	0	0	0	a = A		
		0	0	0	0	b = B		
		0	0	1164	132	c = C		
		0	0	204	1189	d = D		

**Figura 41: Resultado del clasificador “Lazy KStar” en el DC (Edad)**

Una vez más, únicamente con información de las variables del PSD de LF y SD1 se consigue clasificar correctamente el 87,50% (2.353) de las ventanas temporales de 5’ del total de 2.689, alcanzando una **ROC área del 93,9%** para ambos tipos de instancias: las del grupo C (hasta 55 años) y las del grupo D (mayores de 55 años).

La **matriz de confusión** muestra una clara diagonal principal y señala claramente un comportamiento diferenciado de estas dos variables según el concepto de entropía manejado por el clasificador para los dos grupos de edad, con un rendimiento ligeramente superior para el ratio TP o “*True-Positives*” (89,8%) del grupo C que para el grupo D (85,4%).

En definitiva, según este segundo criterio de clasificación, muy diferente al anterior, con la única información de estas dos variables (PSD de LF y SD1) se conforma un sistema muy significativo ligado a la variable dependiente de la edad (representativa de factores personales para esta monografía). Además, según la información generada con el clasificador “*Lazy KStar*”, con rendimientos ligeramente superiores: ROC área del 86,7% con el “*RandomTree*” frente al 93,9% con el “*Lazy KStar*”.

#### 4.1.4. Conclusiones

Como resumen (Tabla 17), en base a los tres estudios anteriores (DT + DF + DC) llevados a cabo entorno a la edad (variable dependiente) como representativa de factores personales, se puede concluir que:

##### **Dominio del Tiempo (DT):**

- 1- Edad (DT): SD1 y SD2 son las variables independientes más significativas en el dominio del tiempo.
- 2- Edad (DT): SD1 tiene valores más altos para el grupo C (participantes más jóvenes).
- 3- Edad (DT): SD2 optimiza la discriminación entre C y D allí donde SD1 no puede.
- 4- Edad (DT): El rendimiento del clasificador 2 basado en la entropía (ROC área = 93,2%) es ligeramente superior al del clasificador 1 (ROC área = 85,0%).

##### **Dominio de la Frecuencia (DF):**

- 5- Edad (DF): El PSD de LF (Power) es la variable independiente más significativa en el dominio de la frecuencia.
- 6- Edad (DF): El PSD de LF tiene valores más altos para el grupo C (participantes más jóvenes).
- 7- Edad (DF): El PSD de HF optimiza la clasificación entre C y D allí donde LF ya no puede.
- 8- Edad (DF): El rendimiento obtenido con el clasificador 2 basado en la entropía (ROC área = 94,4%) es algo superior al del clasificador 1 (ROC área = 91,7%).

##### **Dominio Combinado (DC):**

- 9- Edad (DC): LF, y SD1 son las variables independientes más significativas en el dominio combinado.
- 10- Edad (DC): El rendimiento del clasificador 2 basado en la entropía (ROC área = 93,9%) es un poco superior al del clasificador 1 (ROC área = 86,7%).
- 11- Edad (DC): El SD1 optimiza la clasificación entre C y D allí donde LF ya no puede.
- 12- Edad (DC): La combinación de parámetros en ambos dominios no mejora significativamente la mejor clasificación obtenida en el DF.

##### **Cuadro resumen (Caso de estudio 1):**

	DT		DF		DC	
	CL1	CL2	CL1	CL2	CL1	CL2
Nº Variables:	2	2	2	2	2	2
Variables:	SD1 SD2	SD1 SD2	LF_Power HF_Power	LF_Power HF_Power	SD1 LF_Power	SD1 LF_Power
Nº Clases (n):	2	2	2	2	2	2
Instancias (n):	2.689	2.689	2.689	2.689	2.689	2.689
ROC área (%):	85,0	<b>93,2</b>	91,7	<b>94,4</b>	86,7	93,9
Correctly Classified (%):	84,0	86,3	84,7	87,1	85,5	87,5
Kappa statistic:	0,6818	0,7268	0,6947	0,7420	0,7093	0,7503
Mean absolute error:	0,1288	0,1176	0,1061	0,0987	0,1196	0,1026

**Tabla 17: Tabla resumen de los rendimientos de los clasificadores (Edad)**

## 4.2. Caso de estudio 2: Factores temporales (Ciclos Circadianos)

Este caso de estudio se caracteriza por tener como variable dependiente los ciclos circadianos (factor temporal), concretado en esta monografía en la variable “Ciclo\_Circadiano” (ver apartado “Variables y Modelo de datos”), para el que se han seleccionado los recorridos de 12 participantes en ciclo de mañana (Ciclo\_Circadiano=M) y 12 participantes en ciclo de tarde (Ciclo\_Circadiano=T) a los cuales se les ha informado del experimento a realizar, tal y como se describe en los sub-apartados del capítulo de “Metodología”. No se han registrado ciclos nocturnos (Ciclo\_Circadiano=N) y, aunque se considera una línea abierta de interés para las recomendaciones de tiempos de descanso durante la conducción prolongada, no ha sido objeto de este estudio.

Los participantes seleccionados para este análisis se muestran en la Tabla 18:

<b>Id</b>	<b>Nombre</b>	<b>Apellido 1</b>	<b>Apellido 2</b>	<b>Ciclo Circadiano</b>
1	Jesus	Murgoitio	Larrauri	M
2	Manuel	Marin	Andrés	M
3	Martin	Etxebarria	Ezpeleta	M
4	Txomin	Rodriguez	Villa	T
5	Juan	Nevado	Durán	M
6	Eneko	Nevado	Murgoitio	T
7	Alberto Demetrio	DeDiego	Ibáñez	T
8	Jose Ignacio	Beitia	Sertucha	M
9	Alberto	DeDiego	Campos	M
11	Juan Antonio	Mentxaka	Alberdi	M
12	Javier	Markaida	Markaida	M
14	Joseba	Arruza	Elguezabal	T
16	Manu	Jauregui	Larringan	T
22	Aitor	Arin	Bilbao	T
23	Josu	Jaka	Aspe	M
27	Xabier	Razkin	Euba	T
28	Jon Iker	Aranbarri	Beldarrain	T
29	Alain	Bilbao	Abaunza	T
30	Juan Jose	Amarika	Sagarminaga	T
32	Pedro	Intxauralde	Etxebarria	T
33	Jose Antonio	Larringan	Uriarte	T
34	Benito	Nevado	Durán	M
35	Francisco Javier	Saez	Martínez	M
36	Gaspar	Nevado	Durán	M

**Tabla 18: Participantes seleccionados para el estudio de los ciclos circadianos**

Al igual que para el estudio anterior centrado en los grupos de edad, como resultado del experimento se obtiene un registro ECG por participante correspondiente a su trayecto de conducción prolongada durante alrededor de tres horas, que tras las etapas uno a seis de la metodología han generado 2.705 ventanas de información correspondiendo cada una de ellas a cinco minutos de conducción de un participante, y suponiendo en total más de 112 horas de conducción.

A continuación se detallan el análisis y resultados recogidos para el estudio de los ciclos circadianos, centrado en la variable “Ciclo\_Circadiano” como variable dependiente, según el siguiente esquema:

- 1- Análisis en el dominio del tiempo (DT):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 2- Análisis en el dominio de la frecuencia (DF):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 3- Análisis combinado (DC):
  - a. Selección de atributos: según selecciones en DT y DF.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
- 4- Conclusiones.

#### **4.2.1. Análisis en el dominio del tiempo**

El análisis en el dominio del tiempo para el caso de estudio 2 se caracteriza por definir como variable dependiente a la variable “**Ciclo\_Circadiano**” y como variables independientes a las variables en el dominio del tiempo denominadas como “**Conjunto DT**” en el apartado “Variables y Modelo de datos”, i.e. Media, Desviación, SD1 y SD2. Derivado de ellas también se consideran el ratio SD1/SD2, así como las medidas normalizadas de la Media, SD1 y SD2 (8 variables independientes).

## Selección de atributos:

Así, en la etapa de selección de atributos se obtienen los resultados que se reflejan en la Figura 42 (para más detalle de los métodos de selección ver “Minería de datos” de la metodología):

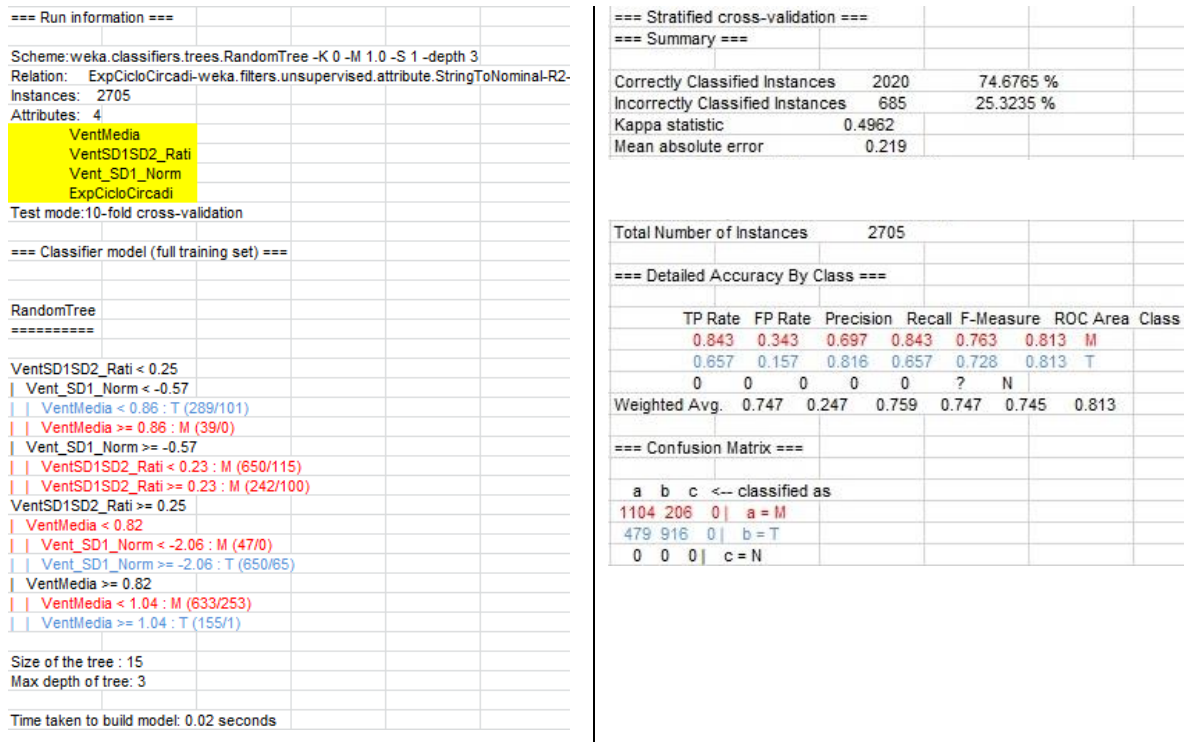


**Figura 42: Ranking de atributos en el DT (Ciclos Circadianos)**

De esta manera, según la Figura 42, en la etapa de selección de atributos se valoran las variables Media (VentMedia=100%), SD1 (VentSD1=100%), ratio SD1/SD2 (VentSD1SD2\_Rati=100%) y SD1 normalizada (Vent\_SD1\_Norm=100%) como claramente las más correlacionadas con “Ciclo\_Circadiano” (véase en color rojo en la Figura 42).

## Resultados del clasificador 1: “RandomTree”

Basado en el ranking de variables independientes obtenido anteriormente se ha llevado a cabo el análisis de las mismas en base al clasificador “RandomTree” generando el correspondiente árbol de decisión, y siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), seleccionándose como mejor resultado el de la Figura 43:



**Figura 43: Clasificador "RandomTree" en el DT (Ciclos Circadianos)**

Según el árbol obtenido, solamente con información de la media, el ratio SD1/SD2 y el SD1 normalizado se consigue clasificar correctamente al 74,68% (2.020) de las instancias (ventanas temporales de 5') de un total de 2.705, alcanzando una **ROC área del 81,3%** para las instancias de la mañana (M) y las de la tarde (T).

De esta manera, según el árbol de decisión anterior obtenido (*Size of the tree* = 15, *Max depth of tree* = 3) se aprecia un comportamiento más diferenciado en variables como el ratio SD1/SD2, siendo mayor para gran parte de los ciclos de tarde, y allí donde hay confusión son necesarias las otras dos variables (media y SD1 normalizado) para conseguir que el ROC área supere el 80 %.

En resumen, el ratio SD1/SD2 es la variable independiente más estrechamente ligada a los ciclos circadianos (SD1/SD2 menores para los ciclos de mañana) y representativa de un factor temporal (ella sola consigue una ROC área del 69,4% para árboles con profundidad o *Max depth of tree* = 2). Aún y todo, necesita complementarse con información de la media y el SD1 normalizado para optimizar la inferencia de conocimiento relacionada con los ciclos Mañana/Tarde y superar rendimientos del 80%.

## Resultados del clasificador 2: “Entropía”

De forma similar, y en base al mismo ranking de variables, se ha realizado el análisis del comportamiento de las mismas en base al segundo clasificador “Lazy KStar” (Cleary, 1995), generando como mejor resultado el de la Figura 44, siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*):

=== Run information ===		=== Stratified cross-validation ===	
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===	
Relation: ExpCicloCircadi-weka.filters.unsupervised.attribute.StringToNominal-R2-v		Correctly Classified Instances	2455 90.7579 %
Instances: 2705		Incorrectly Classified Instances	250 9.2421 %
Attributes: 4		Kappa statistic	0.8151
VentMedia		Mean absolute error	0.1207
VentSD1(ms)			
VentSD1SD2_Rati			
ExpCicloCircadi			
Test mode: 10-fold cross-validation			
=== Classifier model (full training set) ===			
KStar Beta Verion (0.1b).		Total Number of Instances	2705
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).		=== Detailed Accuracy By Class ===	
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).		TP Rate	FP Rate
KStar options : -B 20 -M a		Precision	Recall
Time taken to build model: 0 seconds		F-Measure	ROC Area
		Class	
		0.92	0.104
		0.893	0.92
		0.906	0.971
		0.896	0.08
		0.923	0.896
		0.909	0.971
		0	0
		0	0
		0	0
		0.908	0.092
		0.908	0.908
		0.908	0.971
		=== Confusion Matrix ===	
		a	b
		c	<-- classified as
		1205	105
		0	a = M
		145	1250
		0	b = T
		0	0
		0	c = N

Figura 44: Clasificador “Lazy KStar” en el DT (Ciclos Circadianos)

Según esto, con información similar a la utilizada en el clasificador 1 (esta vez se ha utilizado SD1 en lugar de su normalizada) se consigue clasificar correctamente el 90,76% (2.455) de las ventanas temporales de 5’ del total de 2.705, alcanzando una **ROC área del 97,1%** para ambos tipos de instancias: las del grupo M (mañana) y las del grupo T (tarde).

Conforme a ello, la **matriz de confusión** muestra una marcada diagonal principal y señala claramente un comportamiento muy diferenciado de este grupo de variables según el concepto de entropía manejado por el clasificador para los dos grupos, con un rendimiento ligeramente superior para el ratio TP o “*True-Positives*” (92,0%) del grupo M que para el grupo T (89,6%).

Una vez más, según este criterio de clasificación completamente diferente al anterior, el conjunto de la información formado por la media, el ratio SD1/SD2 y el SD1 (sin normalizar) significan un sistema de marcadores muy ligados a la variable dependiente “Ciclo\_Circadiano” (representativa de factores temporales en esta monografía), todo ello

en base a la información generada por el clasificador “*Lazy KStar*”, con rendimientos ligeramente superiores: ROC área del 81,3% con el “*RandomTree*” frente al 97,1% con el “*Lazy KStar*”.

#### 4.2.2. Análisis en el dominio de la frecuencia

El análisis en el dominio de la frecuencia para el caso de estudio 2 se caracteriza por tener de variable dependiente a la variable “**Ciclo\_Circadiano**” y como variables independientes a aquellas en el dominio de la frecuencia denominadas como “**Conjunto DF**” en el apartado “Variables y Modelo de datos”. Derivadas de ellas también se consideraron algunas medidas normalizadas de las mismas (hasta 19 variables independientes).

##### Selección de atributos:

Así, en la etapa de selección de atributos se obtuvieron los resultados que se reflejan en la Figura 45 (ver métodos de selección en “Minería de datos” de la metodología):

=== Run information ===	
Evaluator:	weka.attributeSelection.CfsSubsetEval
Search:	weka.attributeSelection.ExhaustiveSearch
Relation:	ExpCicloCircadi-weka.filters.unsupervised.attribute.StringToNomina
Instances:	2705
Attributes:	20
	VentVLF_PeakX
	VentVLF_Peak_Y
	Vent_LF_Peak_X
	Vent_LF_Peak_Y
	VentHF_Peak_X
	VentHF_Peak_Y
	VentVLF_Power
	VentLF_Power
	VentHF_Power
	VentPorcVLF_Pow
	VentPorcLF_Powe
	VentPorcHF_Powe
	VentLF/HF
	VentVLF_XC
	VentVLF_YC
	VentLF_XC
	VentLF_YC
	VentHF_XC
	VentHF_YC
	ExpCicloCircadi
Evaluation mode:10-fold cross-validation	

#### Ranking

Grupo	Atributos
1	LF_Power, HF_Power, PorcHF_Powe HF_XC_Power, HF_YC
2	VLF_PeakX, VLF_Peak_Y, HF_Peak_X
3	HF_PeakY, VLF_Power, PorcVLF_Pow PorcLF_Powe, VLF_XC
4	Resto

```

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===
number of folds (%) attribute
9( 90 %) 1 VentVLF_PeakX
8( 80 %) 2 VentVLF_Peak_Y
0( 0 %) 3 Vent_LF_Peak_X
0( 0 %) 4 Vent_LF_Peak_Y
6( 60 %) 5 VentHF_Peak_X
3( 30 %) 6 VentHF_Peak_Y
1( 10 %) 7 VentVLF_Power
10(100 %) 8 VentLF_Power
10(100 %) 9 VentHF_Power
3( 30 %) 10 VentPorcVLF_Pow
1( 10 %) 11 VentPorcLF_Powe
10(100 %) 12 VentPorcHF_Powe
0( 0 %) 13 VentLF/HF
1( 10 %) 14 VentVLF_XC
0( 0 %) 15 VentVLF_YC
0( 0 %) 16 VentLF_XC
0( 0 %) 17 VentLF_YC
10(100 %) 18 VentHF_XC
10(100 %) 19 VentHF_YC

```

Figura 45: Ranking de atributos en el DF (Ciclos Circadianos)

Siguiendo el **ranking de variables** establecido según la Figura 45 se ha llevado a cabo el análisis final de este caso de estudio en el DF en base al uso de los clasificadores.

### Resultados del clasificador 1: “RandomTree”

Basado en el ranking de variables anterior se ha analizado con el clasificador “RandomTree” generándose el correspondiente árbol de decisión, y siguiendo el criterio de máxima “ROC Area” con mínimo número de atributos (*Attributes*) y mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), siendo el mejor resultado el de la Figura 46:

<pre> === Run information === Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1 -depth 3 Relation: ExpCicloCircadi-weka.filters.unsupervised.attribute.StringToNominal-R2- Instances: 2705 Attributes: 4   VentVLF_Peak_Y   VentHF_Power   VentHF_XC   ExpCicloCircadi Test mode: 10-fold cross-validation </pre>		<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances 2062 76.2292 % Incorrectly Classified Instances 643 23.7708 % Kappa statistic 0.5239 Mean absolute error 0.2191 </pre>	
<pre> === Classifier model (full training set) ===  RandomTree ===== VentHF_Power &lt; 134.71   VentHF_XC &lt; 0.25     VentVLF_Peak_Y &lt; 0 : M (177/83)     VentVLF_Peak_Y &gt;= 0 : M (932/204)   VentHF_XC &gt;= 0.25     VentVLF_Peak_Y &lt; 0.02 : T (826/249)     VentVLF_Peak_Y &gt;= 0.02 : M (136/15) VentHF_Power &gt;= 134.71   VentHF_XC &lt; 0.27     VentHF_Power &lt; 184.98 : T (122/21)     VentHF_Power &gt;= 184.98 : T (387/7)   VentHF_XC &gt;= 0.27     VentHF_Power &lt; 260.66 : T (54/22)     VentHF_Power &gt;= 260.66 : M (71/3) </pre>		<pre> Total Number of Instances 2705  === Detailed Accuracy By Class === TP Rate FP Rate Precision Recall F-Measure ROC Area Class 0.749 0.225 0.758 0.749 0.753 0.813 M 0.775 0.251 0.767 0.775 0.771 0.813 T 0 0 0 0 0 ? N Weighted Avg. 0.762 0.239 0.762 0.762 0.762 0.813 </pre>	
<pre> Size of the tree : 15 Max depth of tree: 3 Time taken to build model: 0 seconds </pre>		<pre> === Confusion Matrix === a b c &lt;-- classified as 981 329 0   a = M 314 1081 0   b = T 0 0 0   c = N </pre>	

Figura 46: Clasificador “RandomTree” en el DF (Ciclos Circadianos)

Según este resultado, solo ha sido necesario tener información en las frecuencias altas (HF) y muy bajas (VLF), para clasificar correctamente al 76,23% (2.062) de las instancias (ventanas de 5') de un total de 2.705, alcanzando una **ROC área del 81,3%** para las instancias del grupo T (ciclos de tarde, color azul en el *RandomTree* generado) y las del grupo M (ciclos de mañana, color rojo en el *RandomTree* generado).

Así mismo, la simpleza del árbol de decisión obtenido (*Size of the tree* = 15, *Max depth of tree* = 3) indican un comportamiento muy diferenciado de estas variables para los dos grupos (M y T): La HF\_Power detectada para el grupo de mañana (M) es sensiblemente inferior a la detectada para el grupo de la tarde (T).

En resumen, según la información obtenida a partir del análisis con el clasificador “*RandomTree*”, solo con información de las frecuencias altas (Power + Peak + C) se conforma un sistema simple y valioso estrechamente ligado a los ciclos circadianos, variable representativa de un factor temporal. Cabe añadir que este sistema, incluso con información de HF (Power + XC) consigue buen rendimiento (ROC área = 80’8 %) clasificando correctamente al 72,57% (1.963) de todas las instancias (2.705).

## Resultados del clasificador 2: “Entropía”

De forma similar, y en base al mismo ranking de variables, se ha realizado el análisis del comportamiento de las mismas en base al segundo clasificador “*Lazy KStar*” (Cleary, 1995), generando como mejor resultado el de la Figura 47, siguiendo el criterio de máxima “*ROC Area*” con el mínimo número de atributos (*Attributes*):

=== Run information ===	=== Stratified cross-validation ===
Scheme: weka.classifiers.lazy.KStar -B 20 -M a	=== Summary ===
Relation: ExpCicloCircadi-weka.filters.unsupervised.attribute.StringToNominal-R2	Correctly Classified Instances 2101 77.671 %
Instances: 2705	Incorrectly Classified Instances 604 22.329 %
Attributes: 3	Kappa statistic 0.5544
VentHF_Power	Mean absolute error 0.2284
VentHF_XC	
ExpCicloCircadi	
Test mode: 10-fold cross-validation	
=== Classifier model (full training set) ===	Total Number of Instances 2705
KStar Beta Verion (0.1b).	=== Detailed Accuracy By Class ===
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).	
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).	TP Rate FP Rate Precision Recall F-Measure ROC Area Class
KStar options : -B 20 -M a	0.824 0.267 0.743 0.824 0.781 0.855 M
Time taken to build model: 0 seconds	0.733 0.176 0.816 0.733 0.772 0.855 T
	0 0 0 0 0 ? N
	Weighted Avg. 0.777 0.22 0.781 0.777 0.776 0.855
	=== Confusion Matrix ===
	a b c <- classified as
	1079 231 0   a = M
	373 1022 0   b = T
	0 0 0   c = N

Figura 47: Clasificador “*Lazy KStar*” en el DF (Ciclos Circadianos)

Según esto, con parecida información (un parámetro menos) a la utilizada en el clasificador 1 se consigue clasificar correctamente el 77,67% (2.101) de las ventanas temporales de 5' del total de 2.705, alcanzando una **ROC área del 85,5%** para ambos tipos de instancias: las del grupo M (mañana) y las del grupo T (tarde).

Conforme a ello, la **matriz de confusión** muestra una clara diagonal principal y señala claramente un comportamiento muy diferenciado de este grupo de variables según el concepto de entropía manejado por el clasificador para los dos grupos, con un rendimiento ligeramente superior para el ratio TP o “*True-Positives*” (82,4%) del grupo M que para el grupo T (73,3%).

Una vez más, según este criterio de clasificación completamente diferente al anterior, el conjunto de la información formado de HF (Power + XC) significan un sistema de marcadores muy ligados a la variable dependiente “Ciclo\_Circadiano” (representativa de factores temporales dentro de esta monografía), todo ello en base a la información generada con el clasificador “*Lazy KStar*”, con rendimientos ligeramente superiores: ROC área del 81,3% con el “*RandomTree*” frente al 85,5% con el “*Lazy KStar*”, aunque requiriendo menos información: (HF\_Power + HF\_XC + VLF\_PeakY) para el “*RandomTree*” frente a (HF\_Power + HF\_XC) para el “*Lazy KStar*”.

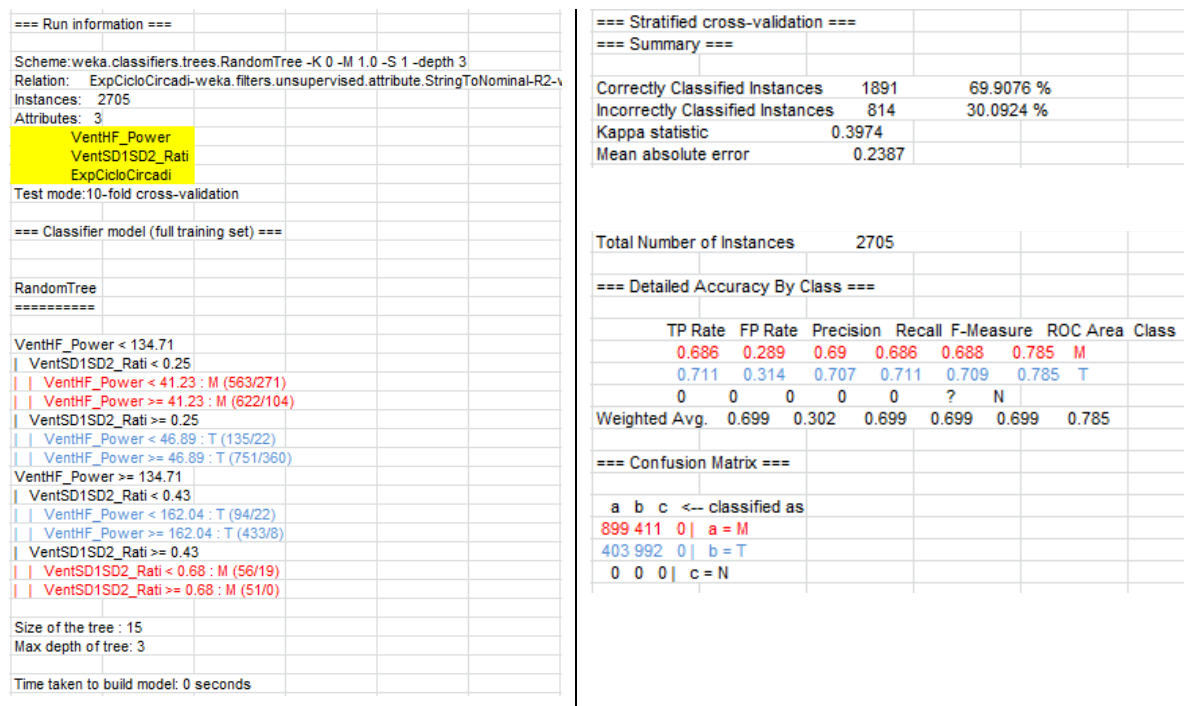
### **4.2.3. Análisis en el dominio combinado**

El análisis al combinar variables en el dominio del tiempo y la frecuencia para el caso de estudio 2, i.e. DC o dominio combinado, se caracteriza por tener como variable dependiente a la variable “**Ciclo\_Circadiano**” y como variables independientes aquellas que basadas en los estudios DT y DF anteriores han mostrado mayor correlación con “**Ciclo\_Circadiano**”, i.e. Media, SD1 y ratio SD1/SD2 para DT y HF (Power + C) para DF, y que serán la base para el análisis final de este caso de estudio con clasificadores.

#### **Resultados del clasificador 1: “RandomTree”**

Basado en la información de variables anterior se ha hecho el análisis con el clasificador “*RandomTree*” generándose el correspondiente árbol de decisión. De esta manera, siguiendo el criterio de máxima “*ROC Area*” con el mínimo número de atributos

(Attributes) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), se ha seleccionado como mejor resultado el de la Figura 48:



**Figura 48: Clasificador “RandomTree” en el DC (Ciclos Circadianos)**

Según este resultado, solo haría falta tener información del “Power” en la frecuencia HF, además del ratio SD1/SD2, para clasificar correctamente a prácticamente el 70 % de las instancias (ventanas temporales de 5’) de un total de 2.705, alcanzando en su evaluación una **ROC área del 78,5%** para las instancias del grupo M (ciclos de mañana y de color rojo en el *RandomTree* generado) y las del grupo T (ciclos de tarde y de color azul en el *RandomTree* generado). Añadir solamente que la utilización de un parámetro más del DF como es HF\_XC podría mejorar algo el rendimiento del clasificador (ROC área = 80,9%).

De esta manera, se mantiene la simpleza del árbol de decisión obtenido (*Size of the tree* = 15, *Max depth of tree* = 3) indicando un comportamiento diferenciado de estas variables para los dos grupos (M y T), donde las instancias del grupo M se caracterizan por valores más bajos del PSD para las HF y ratio SD1/SD2 menor.

En resumen, según la información obtenida a partir del análisis de la clasificación obtenida con el clasificador “*RandomTree*”, la combinación de información en el DT y el DF

formada por el PSD de HF junto al ratio SD1/SD2 conforman un sistema simple como marcadores ligados a los ciclos circadianos, variable representativa de un factor temporal.

## Resultados del clasificador 2: “Entropía”

En base al mismo grupo de variables, se ha realizado el análisis del comportamiento de las mismas en base al segundo clasificador “Lazy KStar” según la entropía (Cleary, 1995), generando como mejor resultado el de la Figura 49 con la máxima “ROC Area” obtenida y el mínimo número de atributos (*Attributes*):

=== Run information ===		=== Stratified cross-validation ===	
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===	
Relation: ExpCicloCircadi-weka.filters.unsupervised.attribute.StringToNominal-R2-		Correctly Classified Instances	1996 73.7893 %
Instances: 2705		Incorrectly Classified Instances	709 26.2107 %
Attributes: 3		Kappa statistic	0.478
Test mode: 10-fold cross-validation		Mean absolute error	0.2425
=== Classifier model (full training set) ===		Total Number of Instances	2705
KStar Beta Verion (0.1b).		=== Detailed Accuracy By Class ===	
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).		TP Rate	FP Rate
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).		0.816	0.335
KStar options : -B 20 -M a		Precision	Recall
Time taken to build model: 0 seconds		0.665	0.184
		F-Measure	ROC Area
		0.794	0.665
		?	0.826
		Weighted Avg.	0.738
			0.257
			0.746
			0.738
			0.737
			0.826
		=== Confusion Matrix ===	
		a	b
		1069	241
		468	927
		0	0
		<-- classified as	
		a = M	
		b = T	
		c = N	

Figura 49: Resultado del clasificador “Lazy KStar” en el DC (Ciclos Circadianos)

Una vez más, únicamente con información de las variables del PSD de HF y el ratio SD1/SD2 se consigue clasificar correctamente el 73,79% (1.996) de las ventanas temporales de 5’ del total de 2.705, alcanzando una **ROC área del 82,6%** para ambos tipos de instancias: las del grupo M (ciclo de mañana) y las del grupo T (ciclo de tarde).

La **matriz de confusión** muestra una marcada diagonal principal y señala claramente un comportamiento diferenciado de estas dos variables según el concepto de entropía manejado por el clasificador para los dos grupos, con un rendimiento ligeramente superior para el ratio TP o “*True-Positives*” (81,6%) del grupo M.

En definitiva, según este segundo criterio de clasificación, la simple información de estas dos variables (PSD de HF y ratio SD1/SD2) conforman un sistema bastante significativo

ligado a la variable dependiente de los ciclos circadianos (representativa de factores temporales para esta monografía), y según la información generada con el clasificador “*Lazy KStar*”, con rendimientos ligeramente superiores: ROC área del 78,5% con el “*RandomTree*” frente al 82,6% con el “*Lazy KStar*”. Además hay que tener en cuenta que si se tuviera en cuenta el parámetro HF\_XC del DF, el rendimiento de la clasificación con el “*Lazy KStar*” mejoraría hasta una ROC área del 91,0% frente al ROC área de 80,9% conseguido con el “*RandomTree*” para el mismo conjunto de parámetros.

#### 4.2.4. Conclusiones

Como resumen (Tabla 19), en base a los tres estudios anteriores (DT + DF + DC) llevados a cabo entorno a los ciclos circadianos (variable dependiente) como representativa de factores temporales, se puede concluir que:

##### **Dominio del Tiempo (DT):**

- 1- Ciclos circadianos (DT): El ratio SD1/SD2 es la variable independiente más significativa en el dominio del tiempo.
- 2- Ciclos circadianos (DT): Los ciclos de mañana (M) tienen en general ratios SD1/SD2 más bajos y el SD1 normalizado optimiza la clasificación donde el ratio SD1/SD2 no puede. Por el contrario, los ciclos de tarde (T) tienen en general ratios SD1/SD2 más altos y la Media optimiza la clasificación donde el ratio SD1/SD2 no puede.
- 3- Ciclos circadianos (DT): El SD1 normalizado mejora el rendimiento del clasificador 1, y el SD1 sin normalizar el del clasificador 2.
- 4- Ciclos circadianos (DT): El rendimiento del clasificador 2 basado en la entropía (ROC área = 97,1%) es superior al del clasificador 1 (ROC área = 81,3%).

##### **Dominio de la Frecuencia (DF):**

- 5- Ciclos circadianos (DF): La información relativa a HF (Power + C) es la más significativa en el dominio de la frecuencia.
- 6- Ciclos circadianos (DF): El PSD de HF tiende a valores más altos para el grupo T (ciclos de tarde).
- 7- Ciclos circadianos (DF): El Peak\_X de VLF optimizan la discriminación entre M y T allí donde HF no puede.
- 8- Ciclos circadianos (DF): El rendimiento obtenido con el clasificador 2 basado en la entropía (ROC área = 85,5%) es algo superior al del clasificador 1 (ROC área = 81,3%) y con menos información: Para los rendimientos anteriores el clasificador 1 necesita un parámetro más.

##### **Dominio Combinado (DC):**

- 9- Ciclos circadianos (DC): PSD de HF y SD1/SD2 son las variables independientes significativas en el dominio combinado. ROC área del clasificador 1 = 78,5%. ROC área del clasificador 2 = 82,6%.
- 10- Ciclos circadianos (DC): El rendimiento del clasificador 2 basado en la entropía (ROC área = 82,6%) es un poco superior al del clasificador 1 (ROC área = 78,6%).

11- Ciclos circadianos (DC): La combinación de parámetros en ambos dominios no mejora significativamente la mejor clasificación obtenida en el DT con el clasificador 2 aunque con un parámetro más.

**Cuadro resumen (Caso de estudio 2):**

	DT		DF		DC	
	CL1	CL2	CL1	CL2	CL1	CL2
Nº Variables:	3	3	3	2	2	2
Variables:	Media SD1_Nor SD1/SD2	Media SD1 SD1/SD2	VLF_PeakY HF_Power HF_XC	HF_Power HF_XC	SD1/SD2 HF_Power	SD1/SD2 HF_Power
Nº Clases (n):	2	2	2	2	2	2
Instancias (n):	2.705	2.705	2.705	2.705	2.705	2.705
ROC área (%):	81,3	<b>97,1</b>	81,3	<b>85,5</b>	78,5	82,6
Correctly Classified (%):	74,7	90,8	76,2	77,7	69,9	73,8
Kappa statistic:	0,4962	0,8151	0,5239	0,5544	0,3974	0,4780
Mean absolute error:	0,2190	0,1207	0,2191	0,2284	0,2387	0,2425

**Tabla 19: Tabla resumen de los rendimientos de los clasificadores (Ciclos Circadianos)**

### 4.3. Caso de estudio 3: Factores ambientales (Tipo de carretera)

Este caso de estudio se caracteriza por tener como variable dependiente el tipo de carretera (factor ambiental), concretado en esta monografía en la variable “**Tipo\_Carretera**” (ver apartado “Variables y Modelo de datos”), y para el que se han analizado los recorridos de 36 participantes a los cuales previamente se les ha informado del experimento a realizar, tal y como se describe en los sub-apartados del capítulo de “Metodología”.

Para este estudio, según se explica con más detalle en “**Descripción del recorrido**”, se han definido cinco tipos de carretera en función de la exigencia en cuanto a nivel de alerta reclamada, i.e. más nivel de alerta necesaria para C5 (vía urbana) o C4 (vía comarcal o de montaña) que para C1 (Autovía). Posteriormente, con el objetivo de resaltar el nivel de alerta inferido en los parámetros recogidos, a partir estos cinco tipos se seleccionaron solo aquellos que requieren niveles de alerta muy diferentes entre sí: por una parte todas las ventanas cuyo “Tipo de carretera” es C1 (mínimo nivel de alerta requerido), y la agrupación de las ventanas que tienen C4 o C5 como tipo asignado (máximos niveles de alerta requeridos), prescindiendo de aquellas cuyo tipo asignado es C2 o C3 (para más detalle ver en “Casos de estudio” lo referido al **caso de estudio 3: “Tipo de carretera**”).

Los participantes seleccionados para este análisis se muestran en la Tabla 20:

<b>Id</b>	<b>Nombre</b>	<b>Apellido1</b>	<b>Apellido 2</b>
1	Jesus	Murgoitio	Larrauri
2	Manuel	Marin	Andrés
3	Martin	Etxebarria	Ezpeleta
4	Txomin	Rodriguez	Villa
5	Juan	Nevado	Durán
6	Eneko	Nevado	Murgoitio
7	Alberto Demetrio	DeDiego	Ibáñez
8	Jose Ignacio	Beitia	Sertucha
10	Jose Ignacio	Rivada	González
11	Juan Antonio	Mentxaka	Alberdi
12	Javier	Markaida	Markaida
13	Koldo	Barcenilla	Sagardui

<b>Id</b>	<b>Nombre</b>	<b>Apellido1</b>	<b>Apellido 2</b>
14	Joseba	Arruza	Elguezabal
15	Juan Antonio	Ruiz	Nevado
16	Manu	Jauregui	Larringan
17	Jose Luis	Zabaleta	LasHeras
18	Javier	Olazar	Gangoiti
19	Pedro	Artetxe	Atucha
20	Jose Maria	Arieta-araunabeña	Bustinza
21	Josu	Barcenilla	Sagardui
22	Aitor	Arin	Bilbao
23	Josu	Jaka	Aspe
24	Alfredo	Madina	Jugo
25	Asier	Aranbarri	Beldarrain
26	Ekain	Aguirre	García
27	Xabier	Razkin	Euba
28	Jon Iker	Aranbarri	Beldarrain
29	Alain	Bilbao	Abaunza
30	Juan Jose	Amarika	Sagarminaga
31	Victor	Sagarduy	Larrondo
32	Pedro	Intxauralde	Etxebarria
33	Jose Antonio	Larringan	Uriarte
34	Benito	Nevado	Durán
35	Francisco Javier	Saez	Martínez
36	Gaspar	Nevado	Durán
37	Jon Andoni	Murgoitio	Larrauri

**Tabla 20: Participantes seleccionados para el estudio del tipo de carretera**

Al igual que para los dos estudios anteriores, como resultado del experimento se obtiene un registro ECG por participante correspondiente a su trayecto de conducción prolongada durante alrededor de tres horas. Así, tras las etapas uno a seis de la metodología han generado 1.854 ventanas de información, correspondiendo cada una de ellas a cinco minutos de conducción de un participante, y suponen en total 77,25 horas de conducción en tramos de carretera con muy diferente exigencia en cuanto al nivel de alerta necesaria.

A continuación se detallan el análisis y resultados recogidos para el estudio del tipo de carretera, centrado en la variable “Tipo\_Carretera” como variable dependiente, según el siguiente esquema:

- 1- Análisis en el dominio del tiempo (DT):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
  
- 2- Análisis en el dominio de la frecuencia (DF):
  - a. Selección de atributos.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
  
- 3- Análisis combinado (DC):
  - a. Selección de atributos: según selecciones en DT y DF.
  - b. Resultados según clasificador 1 (RandomTree).
  - c. Resultados según clasificador 2 (Lazy KStar).
  - d. Conclusión
  
- 4- Conclusiones.

#### **4.3.1. Análisis en el dominio del tiempo**

El análisis en el dominio del tiempo para el caso de estudio 3 se caracteriza por tener como variable dependiente a la variable “**Tipo\_Carretera**” y como variables independientes a las variables en el dominio del tiempo o el denominado “**Conjunto DT**” en el apartado “Variables y Modelo de datos” (i.e. Media, Desviación, SD1 y SD2). Derivado de ellos también se ha considerado el ratio SD1/SD2, así como las medidas normalizadas de la Media, SD1 y SD2 (8 variables independientes).

##### **Selección de atributos:**

Según lo anteriormente expuesto, en la etapa de selección de atributos se obtuvieron los resultados que se reflejan en la Figura 50:

```

=== Run information ===
Evaluator: weka.attributeSelection.CfsSubsetEval
Search:weka.attributeSelection.ExhaustiveSearch
Relation: VentTipCarReduc-weka.filters.unsupervised.attribute.StringToNominal-R2
Instances: 1854
Attributes: 9
VentMediaNorm
VentDesviaciónS
VentMedia
VentSD1(ms)
VentSD2(ms)
VentSD1SD2_Rati
Vent_SD1_Norm
Vent_SD2_Norm
VentTipCarReduc
Evaluation mode:10-fold cross-validation

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===

number of folds (%) attribute
10(100 %) 1 VentMediaNorm
1( 10 %) 2 VentDesviaciónS
10(100 %) 3 VentMedia
7( 70 %) 4 VentSD1(ms)
0( 0 %) 5 VentSD2(ms)
5( 50 %) 6 VentSD1SD2_Rati
10(100 %) 7 Vent_SD1_Norm
0( 0 %) 8 Vent_SD2_Norm

```

Ranking	
Grupo	Atributos
1	MediaNorm, Media, SD1_Norm
2	SD1, SD1/SD2
3	DesviaciónStandard
4	Resto

**Figura 50: Ranking de atributos en el DT (Tipo de Carretera)**

De esta manera, según la Figura 50, en la etapa de selección de atributos son la Media y el SD1, así como sus normalizadas, las más correlacionadas con “Tipo\_Carretera”. No se han detectado correlaciones de interés en el resto de variables, excepto una muy baja en la “desviación típica”.

### Resultados del clasificador 1: “RandomTree”

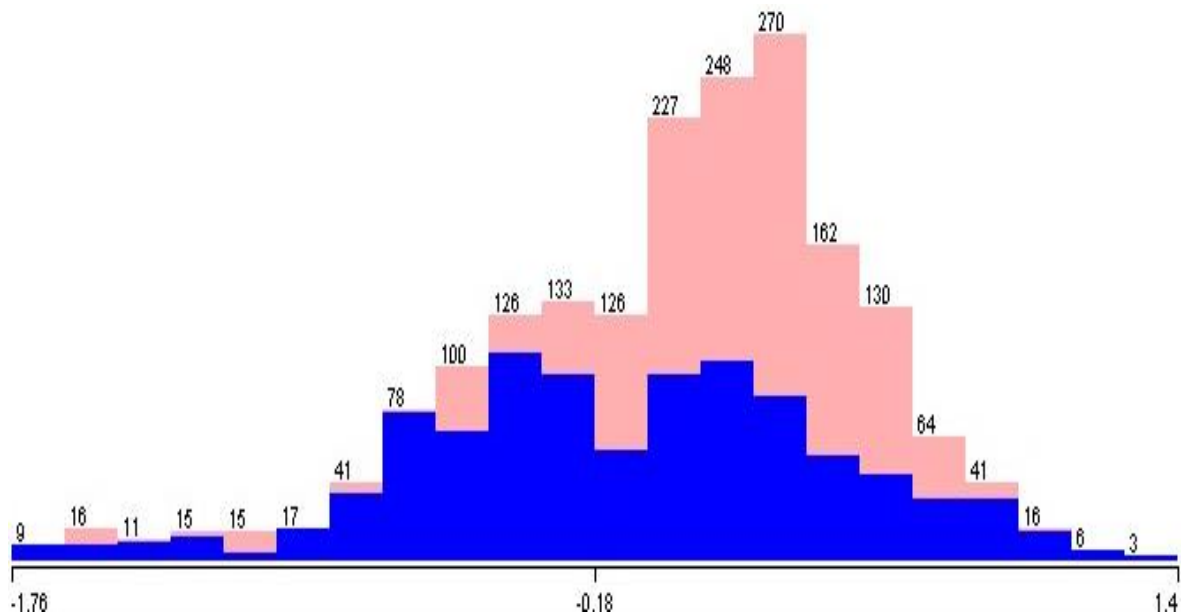
Basado en el ranking de variables anterior se ha hecho el análisis sobre las mismas en base al clasificador “RandomTree” generando el correspondiente árbol de decisión, y siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*). Se ha seleccionado como mejor resultado el de la Figura 51:

=== Run information ===		=== Stratified cross-validation ===																																	
Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1 -depth 4		=== Summary ===																																	
Relation: VentTipCarReduc-weka.filters.unsupervised.attribute.StringToNominal-R2		Correctly Classified Instances	1261 68.0151 %																																
Instances: 1854		Incorrectly Classified Instances	593 31.9849 %																																
Attributes: 3		Kappa statistic	0.3619																																
VentMediaNorm Vent_SD1_Norm VentTipCarReduc		Mean absolute error	0.1288																																
Test mode: 10-fold cross-validation		Total Number of Instances 1854																																	
=== Classifier model (full training set) ===		=== Detailed Accuracy By Class ===																																	
RandomTree		<table border="1"> <thead> <tr> <th></th> <th>TP Rate</th> <th>FP Rate</th> <th>Precision</th> <th>Recall</th> <th>F-Measure</th> <th>ROC Area</th> <th>Class</th> </tr> </thead> <tbody> <tr> <td></td> <td>0.645</td> <td>0.281</td> <td>0.713</td> <td>0.645</td> <td>0.677</td> <td>0.744</td> <td>C1</td> </tr> <tr> <td></td> <td>0.719</td> <td>0.355</td> <td>0.651</td> <td>0.719</td> <td>0.683</td> <td>0.744</td> <td>C5</td> </tr> <tr> <td>Weighted Avg.</td> <td>0.68</td> <td>0.317</td> <td>0.683</td> <td>0.68</td> <td>0.68</td> <td>0.744</td> <td></td> </tr> </tbody> </table>			TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class		0.645	0.281	0.713	0.645	0.677	0.744	C1		0.719	0.355	0.651	0.719	0.683	0.744	C5	Weighted Avg.	0.68	0.317	0.683	0.68	0.68	0.744	
	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class																												
	0.645	0.281	0.713	0.645	0.677	0.744	C1																												
	0.719	0.355	0.651	0.719	0.683	0.744	C5																												
Weighted Avg.	0.68	0.317	0.683	0.68	0.68	0.744																													
<pre> VentMediaNorm &lt; -0.22   Vent_SD1_Norm &lt; 0.83     Vent_SD1_Norm &lt; -0.91       VentMediaNorm &lt; -0.54 : C1 (114/0)       VentMediaNorm &gt;= -0.54 : C1 (39/8)       Vent_SD1_Norm &gt;= -0.91         VentMediaNorm &lt; -1.49 : C5 (8/0)         VentMediaNorm &gt;= -1.49 : C1 (346/71)       Vent_SD1_Norm &gt;= 0.83       VentMediaNorm &lt; -1.21         VentMediaNorm &lt; -1.25 : C5 (2/0)         VentMediaNorm &gt;= -1.25 : C1 (4/0)       VentMediaNorm &gt;= -1.21 : C5 (18/0) VentMediaNorm &gt;= -0.22   VentMediaNorm &lt; 0.89     Vent_SD1_Norm &lt; 0.04       VentMediaNorm &lt; 0.22 : C1 (376/174)       VentMediaNorm &gt;= 0.22 : C5 (263/109)       Vent_SD1_Norm &gt;= 0.04         Vent_SD1_Norm &lt; 1.11 : C5 (584/190)         Vent_SD1_Norm &gt;= 1.11 : C5 (61/2)       VentMediaNorm &gt;= 0.89       Vent_SD1_Norm &lt; -0.14         VentMediaNorm &lt; 0.94 : C1 (1/0)         VentMediaNorm &gt;= 0.94 : C5 (1/0)         Vent_SD1_Norm &gt;= -0.14 : C1 (37/0)           </pre>		=== Confusion Matrix === <table border="1"> <thead> <tr> <th></th> <th>a</th> <th>b</th> <th>c</th> <th>d</th> <th>e</th> <th>f</th> <th>&lt;-- classified as</th> </tr> </thead> <tbody> <tr> <td>a = C1</td> <td>622</td> <td>0</td> <td>0</td> <td>0</td> <td>343</td> <td>0</td> <td></td> </tr> <tr> <td>e = C5</td> <td>250</td> <td>0</td> <td>0</td> <td>0</td> <td>639</td> <td>0</td> <td></td> </tr> </tbody> </table>			a	b	c	d	e	f	<-- classified as	a = C1	622	0	0	0	343	0		e = C5	250	0	0	0	639	0									
	a	b	c	d	e	f	<-- classified as																												
a = C1	622	0	0	0	343	0																													
e = C5	250	0	0	0	639	0																													
Size of the tree : 27																																			
Max depth of tree: 4																																			
Time taken to build model: 0.01 seconds																																			

**Figura 51: Clasificador “RandomTree” en el DT (Tipo de Carretera)**

Según el árbol obtenido, la mejor clasificación requiere información de la media y el SD1 normalizados, clasificando correctamente al 68,02% (1.261) de las instancias (ventanas temporales de 5’) de un total de 1.854, y alcanzando una **ROC área del 74,4%**.

En esta ocasión ha hecho falta generar un árbol de decisión con mayor profundidad y mayor tamaño (*Size of the tree = 27, Max depth of tree = 4*) para llegar a rendimientos significativos y conseguir que el ROC área superara el 70 %. En todo caso, tal y como se aprecia en el histograma de la Figura 52 para la media normalizada (en azul “MediaNorm” = C1, en rojo “MediaNorm” = C4+C5), este parámetro muestra un comportamiento diferente para los tramos de C1 (histograma más amplio) que para los tramos de C4 + C5 (histograma más concentrado):



**Figura 52: Histograma de la media normalizada (Tipo de Carretera)**

En resumen, la media normalizada es la variable independiente más directamente relacionada con “Tipo de carretera” (medias normalizadas más altas y concentradas para los tramos de mayor necesidad de atención, i.e. C4+C5) y representativa de un factor ambiental. Aún y todo, necesita complementarse con información de SD1 para optimizar la clasificación y alcanzar rendimientos cercanos al 75%. En todo caso, hay que tener en cuenta que el nivel de alerta está inferido en este caso de estudio del tramo de carretera y no siempre un tramo de autovía (C1) requiere necesariamente un nivel de alerta bajo, e.g. al hacer un adelantamiento, etc..., razón por la que se explicaría que el árbol de decisión, aunque recoge la tendencia del histograma anterior, se hace menos evidente.

### **Resultados del clasificador 2: “Entropía”**

Una vez más, en base al mismo ranking de variables, se ha realizado el análisis del comportamiento de las mismas en base al segundo clasificador “Lazy KStar” (Cleary, 1995), generando como mejor resultado el de la Figura 53, siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*):

=== Run information ===		=== Stratified cross-validation ===	
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===	
Relation: VentTipCarReduc-weka.filters.unsupervised.attribute.StringToNominal-R2	Instances: 1854	Correctly Classified Instances	1332 71.8447 %
Attributes: 3		Incorrectly Classified Instances	522 28.1553 %
VentMediaNorm		Kappa statistic	0.4398
Vent_SD1_Norm		Mean absolute error	0.1314
VentTipCarReduc			
Test mode: 10-fold cross-validation			
=== Classifier model (full training set) ===		Total Number of Instances	1854
KStar Beta Verion (0.1b).		=== Detailed Accuracy By Class ===	
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).		TP Rate	FP Rate
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).		0.65	0.207
KStar options : -B 20 -M a		Precision	Recall
		0.773	0.65
		F-Measure	ROC Area
		0.793	0.35
		0.676	0.793
		0.73	0.79
		0.79	0.79
		Weighted Avg.	0.718 0.276 0.727 0.718 0.717 0.79
		=== Confusion Matrix ===	
		a b c d e f <-- classified as	
		627 0 0 0 338 0   a = C1	
		184 0 0 0 705 0   e = C5	

Figura 53: Clasificador “Lazy KStar” en el DT (Tipo de Carretera)

Así, con la misma información a la utilizada en el clasificador 1 se consigue clasificar correctamente el 71,84% (1.332) de las ventanas temporales de 5’ del total de 1.854, alcanzando una **ROC área del 79,0%**.

La **matriz de confusión** muestra una clara diagonal principal y marca claramente un comportamiento diferenciado de este grupo de variables según el concepto de entropía manejado por el clasificador para los dos grupos (Figura 54). Además, el rendimiento es ligeramente superior para el ratio TP o “*True-Positives*” (79,3%) del grupo C4+C5 que para el grupo de C1 (65,0%), ratios que mejoran considerablemente si se tuviera en cuenta una más de las variables significativas (+ ratio SD1/SD2 para ROC área del 88,3%):

=== Run information ===		=== Stratified cross-validation ===	
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===	
Relation: VentTipCarReduc-weka.filters.unsupervised.attribute.StringToNominal-R2	Instances: 1854	Correctly Classified Instances	1478 79.7195 %
Attributes: 4		Incorrectly Classified Instances	376 20.2805 %
VentMediaNorm		Kappa statistic	0.5952
VentSD1SD2_Rati		Mean absolute error	0.1062
Vent_SD1_Norm			
VentTipCarReduc			
Test mode: 10-fold cross-validation		Total Number of Instances	1854
=== Classifier model (full training set) ===		=== Detailed Accuracy By Class ===	
KStar Beta Verion (0.1b).		TP Rate	FP Rate
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).		0.763	0.165
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).		Precision	Recall
KStar options : -B 20 -M a		0.834	0.763
		F-Measure	ROC Area
		0.835	0.237
		0.764	0.835
		0.798	0.883
		0.883	0.883
		Weighted Avg.	0.797 0.2 0.8 0.797 0.797 0.883
		=== Confusion Matrix ===	
		a b c d e f <-- classified as	
		736 0 0 0 229 0   a = C1	
		147 0 0 0 742 0   e = C5	

Figura 54: Clasificador “Lazy KStar” en el DT (Tipo de Carretera – 3 variables)

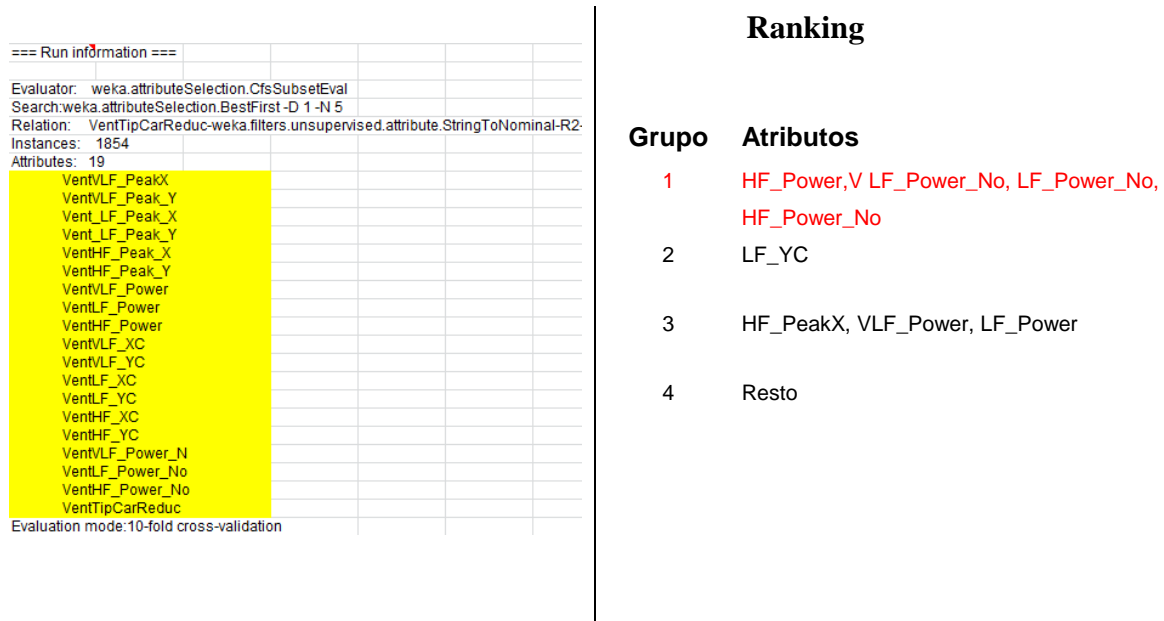
Una vez más, según este criterio de clasificación completamente diferente al anterior, el conjunto de la información formado por la media y el SD1 normalizados significan un sistema de marcadores ligados a la variable dependiente “Tipo\_Carretera” (representativa de factores temporales dentro de esta monografía), todo ello en base a la información generada con el clasificador “*Lazy KStar*”. Además, los rendimientos con este clasificador son ligeramente superiores para el mismo número de parámetros (2): ROC área del 74,4% con el “*RandomTree*” frente al 79,0% con el “*Lazy KStar*”.

### 4.3.2. Análisis en el dominio de la frecuencia

En este caso de estudio el análisis en el dominio de la frecuencia se caracteriza por definir como variable dependiente a la variable “**Tipo\_Carretera**” y como variables independientes a aquellas en el dominio de la frecuencia denominadas como “**Conjunto DF**” en el apartado “Variables y Modelo de datos”. Derivadas de ellas también se consideran algunas medidas normalizadas de las mismas (hasta 19 variables independientes).

#### Selección de atributos:

Según lo anteriormente expuesto, en la etapa de selección de atributos se obtienen los resultados que se reflejan en la Figura 55:



```

=== Attribute selection 10 fold cross-validation (stratified), seed: 1 ===
number of folds (%) attribute
0( 0 %) 1 VentVLF_PeakX
0( 0 %) 2 VentVLF_Peak_Y
0( 0 %) 3 Vent_LF_Peak_X
0( 0 %) 4 Vent_LF_Peak_Y
2( 20 %) 5 VentHF_Peak_X
0( 0 %) 6 VentHF_Peak_Y
1( 10 %) 7 VentVLF_Power
1( 10 %) 8 VentLF_Power
10(100 %) 9 VentHF_Power
0( 0 %) 10 VentVLF_XC
0( 0 %) 11 VentVLF_YC
0( 0 %) 12 VentLF_XC
9( 90 %) 13 VentLF_YC
0( 0 %) 14 VentHF_XC
0( 0 %) 15 VentHF_YC
10(100 %) 16 VentVLF_Power_N
10(100 %) 17 VentLF_Power_No
10(100 %) 18 VentHF_Power_No

```

Figura 55: Ranking de atributos en el DF (Tipo de Carretera)

Siguiendo el **ranking de variables** establecido según la Figura 55 se realiza el análisis final de este caso de estudio en el DF en base al uso de los clasificadores.

### Resultados del clasificador 1: “RandomTree”

Basado en el ranking de variables anterior se lleva a cabo el análisis con el clasificador “RandomTree” y se genera el correspondiente árbol de decisión, siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*). El mejor resultado es el de la Figura 56:

<pre> === Run information === Scheme:weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1 -depth 3 Relation: VentTipCarReduc-weka.filters.unsupervised.attribute.StringToNominal-R2 Instances: 1854 Attributes: 4   VentVLF_Power_N   VentLF_Power_No   VentHF_Power_No   VentTipCarReduc Test mode:10-fold cross-validation  === Classifier model (full training set) ===  RandomTree ===== VentHF_Power_No &lt; -0.58   VentLF_Power_No &lt; -0.65     VentVLF_Power_N &lt; -0.49 : C1 (186/17)     VentVLF_Power_N &gt;= -0.49 : C1 (115/34)     VentLF_Power_No &gt;= -0.65       VentHF_Power_No &lt; -1.07 : C1 (97/24)       VentHF_Power_No &gt;= -1.07 : C1 (161/78) VentHF_Power_No &gt;= -0.58   VentLF_Power_No &lt; -1.18     VentHF_Power_No &lt; -0.46 : C1 (22/6)     VentHF_Power_No &gt;= -0.46 : C1 (29/2)   VentLF_Power_No &gt;= -1.18     VentLF_Power_No &lt; -0.15 : C5 (515/256)     VentLF_Power_No &gt;= -0.15 : C5 (729/260)  Size of the tree : 15 Max depth of tree: 3  Time taken to build model: 0 seconds </pre>		<pre> === Stratified cross-validation === === Summary === Correctly Classified Instances 1139 61.4347 % Incorrectly Classified Instances 715 38.5653 % Kappa statistic 0.2315 Mean absolute error 0.1461  Total Number of Instances 1854  === Detailed Accuracy By Class === TP Rate FP Rate Precision Recall F-Measure ROC Area Class 0.567 0.334 0.648 0.567 0.605 0.67 C1 0.666 0.433 0.586 0.666 0.623 0.67 C5 Weighted Avg. 0.614 0.382 0.618 0.614 0.614 0.67  === Confusion Matrix ===  a b c d e f &lt;-- classified as 547 0 0 0 418 0   a = C1 297 0 0 0 592 0   e = C5 </pre>	
---	--	--	--

Figura 56: Clasificador “RandomTree” en el DF (Tipo de Carretera)



alcanzando una **ROC área del 80,4%** para ambos tipos de instancias: las del grupo C1 (mínimo nivel de alerta requerido) y las del grupo C4+C5 (máximo nivel de alerta requerido).

Conforme a ello, la **matriz de confusión** muestra una marcada diagonal principal y señala claramente un comportamiento muy diferenciado de este grupo de variables según el concepto de entropía manejado por el clasificador para los dos grupos. Además, el rendimiento es ligeramente superior para el ratio TP o “*True-Positives*” (75,7%) del grupo C4+C5 que para el grupo C1 (68,9%).

Una vez más, según un criterio de clasificación completamente diferente al anterior, el conjunto de la información formado por el PSD normalizado de las tres bandas de frecuencia (VLF, LF y HF) significan un sistema de variables ligados a la variable dependiente “*Tipo\_Carretera*” (representativa de factores ambientales dentro de esta monografía), todo ello en base a la información generada con el clasificador “*Lazy KStar*”, con rendimientos superiores: ROC área del 67,0% con el “*RandomTree*” frente al 80,4% con el “*Lazy KStar*”.

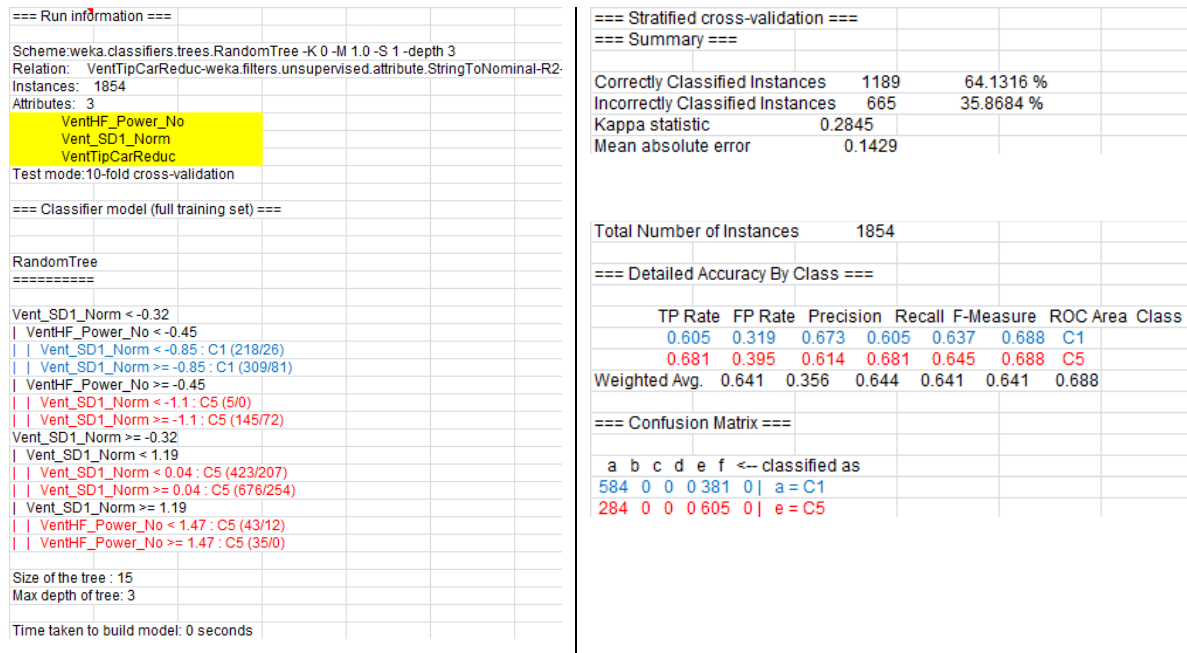
### **4.3.3. Análisis en el dominio combinado**

El análisis al combinar variables en el dominio del tiempo y la frecuencia para este caso de estudio, i.e. DC o dominio combinado, se caracteriza por tener como variable dependiente a la variable “**Tipo\_Carretera**” y como variables independientes a aquellas que basadas en los estudios DT y DF anteriores han mostrado mayor correlación con “**Tipo\_Carretera**”, i.e. Media y SD1 normalizados para DT, y PSD normalizado de todas las frecuencias (VLF, LF y HF) para DF. Estas variables serán el punto de partida del análisis final de este caso de estudio con clasificadores.

#### **Resultados del clasificador 1: “RandomTree”**

Basado en la información de variables anterior se lleva a cabo el análisis con el clasificador “*RandomTree*” generándose el correspondiente árbol de decisión, y siguiendo el criterio de máxima “*ROC Area*” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño

de árbol (*Size of tree & Max.depth of tree*). De esta manera, se ha seleccionado como mejor resultado el de la Figura 58:



**Figura 58: Clasificador “RandomTree” en el DC (Tipo de Carretera)**

Según este resultado, con la información del PSD en la frecuencia HF y del SD1 normalizado se logra clasificar correctamente al 64,13 % de las instancias (ventanas temporales de 5’) de un total de 1.854, alcanzando en su evaluación una **ROC área del 68,8%**, con un rendimiento ligeramente superior para el ratio TP o “*True-Positives*” (68,1%) del grupo C4+C5 que para el grupo C1 (60,5%).

Aunque los resultados no son muy buenos, según el árbol de decisión obtenido (*Size of the tree = 15, Max depth of tree = 3*) se aprecia cierta tendencia de un comportamiento diferenciado de estas variables para los dos grupos (C1 y C4+C5), donde las instancias del grupo C4+C5 se caracterizan por valores más altos del SD1 normalizado. Además, necesita complementarse con el PSD de las HF donde la normalización del SD1 no puede.

## Resultados del clasificador 2: “Entropía”

Soportado en la misma base de variables, se ha realizado el análisis del comportamiento de las mismas según el clasificador “Lazy KStar”, generando como mejor resultado el de la Figura 59 con la máxima “*ROC Area*” obtenida con el mínimo número de atributos (*Attributes*):

=== Run information ===		=== Stratified cross-validation ===	
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===	
Relation: VentTipCarReduc-weka.filters.unsupervised.attribute.StringToNominal-R2		Correctly Classified Instances	1244 67.0982 %
Instances: 1854		Incorrectly Classified Instances	610 32.9018 %
Attributes: 3		Kappa statistic	0.3444
VentHF_Power_No		Mean absolute error	0.1429
Vent_SD1_Norm			
VentTipCarReduc.			
Test mode: 10-fold cross-validation			
=== Classifier model (full training set) ===			
KStar Beta Verion (0.1b).		Total Number of Instances	1854
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).		=== Detailed Accuracy By Class ===	
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).		TP Rate	FP Rate
KStar options : -B 20 -M a		Precision	Recall
Time taken to build model: 0 seconds		F-Measure	ROC Area
		Class	
		0.621	0.274
		0.711	0.621
		0.663	0.731
		0.726	0.379
		0.638	0.726
		0.679	0.731
		0.671	0.67
		0.671	0.731
		=== Confusion Matrix ===	
		a	b
		c	d
		e	f
		←- classified as	
		599	0
		0	366
		0	0
		244	0
		0	645
		0	0

**Figura 59: Clasificador “Lazy KStar” en el DC (Tipo de Carretera)**

Según el resultado anterior, con información de las variables SD1 normalizado y el PSD de HF se consigue clasificar correctamente el 67,10% (1.244) de las ventanas temporales de 5’ del total de 1.854, alcanzando una **ROC área del 73,1%**.

La **matriz de confusión** muestra una clara diagonal principal y señala un comportamiento diferenciado de estas dos variables según el concepto de entropía manejado por el clasificador para los dos grupos, con un rendimiento ligeramente superior para el ratio TP o “*True-Positives*” (72,6%) del grupo C4+C5 que para el grupo C1 (62,1%).

En resumen, según este segundo criterio de clasificación, la simple información de estas dos variables (SD1 normalizado y PSD de HF) conforman un sistema de marcadores ligados a la variable dependiente del tipo de carretera (representativa de factores ambientales para esta monografía). Además, la información generada con el clasificador “*Lazy KStar*”, muestra rendimientos ligeramente superiores: ROC área del 68,8% con el “*RandomTree*” frente al 73,1% con el “*Lazy KStar*”. También hay que tener en cuenta que si se considera el parámetro de la Media normalizada, el rendimiento de la clasificación con el “*Lazy KStar*” subiría hasta una ROC área del 85,5%, ratios que no se alcanzan con el “*RandomTree*” al considerar más parámetros (Figura 60):

=== Run information ===		=== Stratified cross-validation ===	
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===	
Relation: VentTipCarReduc-weka.filters.unsupervised.attribute.StringToNominal-R2		Correctly Classified Instances	1444 77.8857 %
Instances: 1854		Incorrectly Classified Instances	410 22.1143 %
Attributes: 4		Kappa statistic	0.5592
VentMediaNorm		Mean absolute error	0.1147
VentHF_Power_No			
Vent_SD1_Norm			
VentTipCarReduc			
Test mode: 10-fold cross-validation		Total Number of Instances	1854
=== Classifier model (full training set) ===		=== Detailed Accuracy By Class ===	
KStar Beta Verion (0.1b)		TP Rate	FP Rate
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz)		Precision	Recall
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz)		F-Measure	ROC Area
KStar options : -B 20 -M a		Class	
Time taken to build model: 0 seconds		0.73	0.168
		0.825	0.73
		0.774	0.855
		0.832	0.27
		0.739	0.832
		0.783	0.855
		0.779	0.217
		0.784	0.779
		0.779	0.855
		=== Confusion Matrix ===	
		a	b
		c	d
		e	f
		<-- classified as	
		704	0
		0	0
		0	261
		0	0
		0	740
		0	0
		a = C1	
		e = C5	

**Figura 60: Clasificador “Lazy KStar” en el DC (Tipo de Carretera – 3 variables)**

### 4.3.4. Conclusiones

Como resumen (Tabla 21), en base a los tres estudios anteriores (DT + DF + DC) llevados a cabo entorno al tipo de carretera (variable dependiente) como representativa de factores ambientales, se puede concluir que:

#### **Dominio del Tiempo (DT):**

- 1- Tipo de carretera (DT): La media normalizada es la variable independiente más significativa en el dominio del tiempo.
- 2- Tipo de carretera (DT): El grupo C1 tiende a valores más distribuidos de la media normalizada mientras el grupo de C4+C5 lo hace a valores más concentrados sobre una media normalizada más alta.
- 3- Tipo de carretera (DT): La media normalizada optimiza mejor la clasificación del grupo C4+C5 mientras el SD1 la mejora para el grupo de C1, complementándose.
- 4- Tipo de carretera (DT): El rendimiento del clasificador 2 basado en la entropía (ROC área = 79,0%) es algo superior al del clasificador 1 (ROC área = 74,4%), para el mismo conjunto de variables formado por la media y el SD1 normalizados.

#### **Dominio de la Frecuencia (DF):**

- 5- Tipo de carretera (DF): La información relativa al PSD de LF y HF es la más significativa en el dominio de la frecuencia.
- 6- Tipo de carretera (DF): El PSD de LF y HF tiende a valores más bajos para el grupo C1 (mínimo nivel de alerta requerido).
- 7- Tipo de carretera (DF): El PSD de LF y VLF necesitan normalizarse para obtener información suficientemente correlacionada con la variable dependiente.
- 8- Tipo de carretera (DF): El rendimiento obtenido con el clasificador 2 basado en la entropía (ROC área = 80,4%) es significativamente superior al del clasificador 1 (ROC área = 67,0%).

#### **Dominio Combinado (DC):**

- 9- Tipo de carretera (DC): SD1 y PSD de HF normalizados son las variables independientes que mejor se complementan y mejores rendimientos de clasificación

consiguen en el dominio combinado. ROC área del clasificador 1 = 68,3%. ROC área del clasificador 2 = 73,1%.

- 10- Tipo de carretera (DC): El rendimiento del clasificador 2 basado en la entropía (ROC área = 73,1%) consigue rendimientos superiores al del clasificador 1 (ROC área = 68,3%) para la misma cantidad de información (variables).
- 11- Tipo de carretera (DC): El rendimiento del clasificador 2 basado en la entropía mejora considerablemente al considerar también la Media normalizada (ROC área = 85,5%), cosa que no se logra con el clasificador 1.
- 12- Tipo de carretera (DC): Los rendimientos más bajos de los clasificadores no son despreciables y marcan tendencias, principalmente teniendo en cuenta que la variable “Tipo\_Carretera” no necesariamente implica siempre un nivel de alerta superior para todos los tramos C5 (vía urbana) o C4 (vía comarcal o de montaña) que para C1 (Autovía), e.g. una tramo C1 podría haber requerido un nivel de alerta más alto al hacer un adelantamiento.

**Cuadro resumen (Caso de estudio 3):**

	DT		DF		DC	
	CL1	CL2	CL1	CL2	CL1	CL2
Nº Variables:	2	2	3	3	2	3
Variables:	Media_Norm SD1_Norm	Media_Norm SD1_Norm	VLF_Power_No LF_Power_No HF_Power_No	VLF_Power_No LF_Power_No HF_Power_No	SD1_Norm HF_Power_No	Media_Norm SD1_Norm HF_Power_No
Nº Clases (n):	2	2	2	2	2	2
Instancias (n):	1.854	1.854	1.854	1.854	1.854	1.854
ROC área (%):	74,4	<b>79,0</b>	67,0	<b>80,4</b>	68,8	85,5
Correctly Classified (%):	68,0	71,8	61,4	72,2	64,1	77,9
Kappa statistic:	0,3619	0,4398	0,2315	0,4445	0,2845	0,5592
Mean absolute error:	0,1288	0,1314	0,1461	0,1272	0,1429	0,1147

**Tabla 21: Tabla resumen de los rendimientos de los clasificadores (Tipo de Carretera)**

#### 4.4. Caso de estudio 4: “Fatiga” (Tipo de tramo)

Partiendo de la integración de todos los resultados anteriores (integración de las variables que mejor se correlacionan con factores “personales + temporales + ambientales”), este caso de estudio se caracteriza por tener como variable dependiente el tipo de tramo, concretado en esta monografía en la variable “**Tipo\_Tramo**” (ver apartado “Variables y Modelo de datos”). En este estudio se analizan los recorridos de 30 participantes a los cuales previamente se les ha informado del experimento a realizar, tal y como se describe en los sub-apartados del capítulo de “Metodología”.

Según se explica en “**Casos de estudio**”, se ha dividido el recorrido total en seis tramos, uno por cada 30 minutos transcurridos desde el comienzo del experimento, y se ha asignado a la variable “Tipo\_Tramo” de cada ventana temporal un valor (A-F): A = de 0’ a 30’, B = de 31’ a 60’, C = de 61’ a 90’, D = de 91’ a 120’, E = de 121’ a 150’, F = desde 151’. En base a ello, el estudio analiza los cambios en las variables que han demostrado buenos rendimientos en los tres casos de estudio anteriores. Como consecuencia se persigue inferir un sistema más integrado que establezca los patrones que mejor definan cada una de las fases de la conducción de un vehículo hasta el comienzo de las primeras señales de fatiga. El objetivo final es que sirva de referencia en los estudios de validación de nuevos equipamientos y/o servicios que se quieran introducir en los coches de última generación, pero basándonos en información obtenida exclusivamente a partir del ECG (e.g. si un nuevo *HMI* para el manejo de un vehículo de carretera provoca que estas fases o patrones se sucedan con más rapidez respecto de otro *HMI*, se concluye que sobrecarga más al conductor) y consume más recursos de la atención.

Los participantes seleccionados para este análisis se muestran en la Tabla 22:

<b>Id</b>	<b>Nombre</b>	<b>Apellido1</b>	<b>Apellido 2</b>
1	Jesus	Murgoitio	Larrauri
2	Manuel	Marin	Andrés
3	Martin	Etxebarria	Ezpeleta
4	Txomin	Rodriguez	Villa
5	Juan	Nevado	Durán
6	Eneko	Nevado	Murgoitio
7	Alberto	DeDiego	Ibáñez
8	Jose Ignacio	Beitia	Sertucha

<b>Id</b>	<b>Nombre</b>	<b>Apellido1</b>	<b>Apellido 2</b>
10	Jose Ignacio	Rivada	González
11	Juan Antonio	Mentxaka	Alberdi
12	Javier	Markaida	Markaida
13	Koldo	Barcenilla	Sagardui
14	Joseba	Arruza	Elguezabal
15	Juan Antonio	Ruiz	Nevado
16	Manu	Jauregui	Larringan
17	Jose Luis	Zabaleta	LasHeras
18	Javier	Olazar	Gangoiti
21	Josu	Barcenilla	Bustinza
22	Aitor	Arin	Bilbao
23	Josu	Jaka	Aspe
24	Alfredo	Madina	Jugo
25	Asier	Aranbarri	Beldarrain
30	Juan Jose	Amarika	Sagarminaga
31	Victor	Sagarduy	Larrondo
32	Pedro	Intxauralde	Etxebarria
33	Jose Antonio	Larringan	Uriarte
34	Benito	Nevado	Durán
35	Francisco Javier	Saez	Martínez
36	Gaspar	Nevado	Durán
37	Jon Andoni	Murgoitio	Larrauri

**Tabla 22: Participantes seleccionados para el estudio del tipo de tramo**

Al igual que para los dos estudios anteriores, como resultado de los experimentos se ha obtenido un registro ECG por participante correspondiente a su trayecto de conducción prolongada durante alrededor de tres horas, que tras las etapas uno a seis de la metodología han generado 5.887 ventanas de información correspondiendo cada una de ellas a cinco minutos de conducción de un participante.

A continuación se detallan el análisis y resultados recogidos para el estudio del tipo de tramo, centrado en la variable “Tipo\_Tramo” como variable dependiente, según el siguiente esquema (para una solución más integrada, la selección de atributos se obtiene de las conclusiones de los tres casos de estudio previos):

- 1- Análisis en el dominio del tiempo (DT):
  - a. Resultados según clasificador 1 (RandomTree).
  - b. Resultados según clasificador 2 (Lazy KStar).
  - c. Conclusión
- 2- Análisis en el dominio de la frecuencia (DF):
  - a. Resultados según clasificador 1 (RandomTree).

- b. Resultados según clasificador 2 (Lazy KStar).
  - c. Conclusión
- 3- Análisis combinado (DC):
- a. Resultados según clasificador 1 (RandomTree).
  - b. Resultados según clasificador 2 (Lazy KStar).
  - c. Conclusión
- 4- Conclusiones.

#### 4.4.1. Análisis en el dominio del tiempo

El análisis en el dominio del tiempo para el caso de estudio 4 se caracteriza por definir como variable dependiente a la variable “**Tipo\_Tramo**” y como variables independientes a las variables en el dominio del tiempo o el denominado “**Conjunto DT**” en el apartado “Variables y Modelo de datos” que han demostrado tener mejores rendimientos en los tres casos de estudio anteriores (Tabla 23): Media + SD1, y sus correspondientes normalizadas, así como SD2 y el ratio SD1/SD2:

	C.Estudio 1		C.Estudio 2		C.Estudio 3	
	CL1	CL2	CL1	CL2	CL1	CL2
<b>Variables:</b>	SD1	<b>SD1</b>	Media SD1_Nor	Media SD1	Media_Norm	<b>Media_Norm</b>
	SD2	SD2	SD1/SD2	<b>SD1/SD2</b>	SD1_Norm	SD1_Norm

**Tabla 23: Resumen de variables significativas en el DT (casos de estudio 1 a 3)**

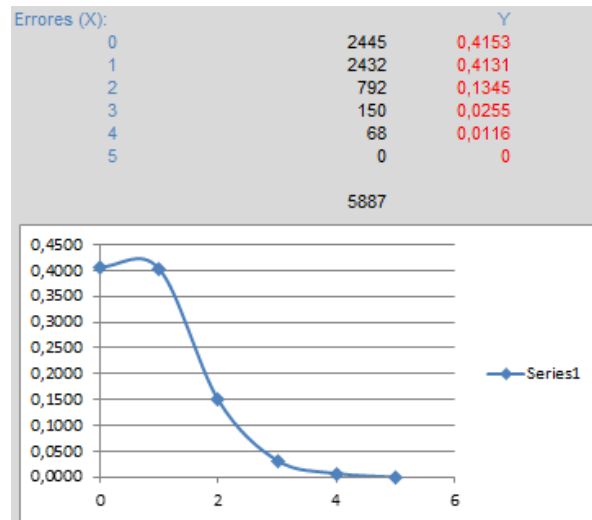
#### Resultados del clasificador 1: “RandomTree”

Basado en las variables anteriores se ha realizado el análisis de las mismas con el clasificador “RandomTree” generando el correspondiente árbol de decisión. Siguiendo el criterio de mejor rendimiento con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), se ha seleccionado como mejor resultado el de la Figura 61:

=== Run information ===		=== Stratified cross-validation ===	
Scheme: weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1 -depth 4		=== Summary ===	
Relation: VentanaTipoTram-weka.filters.unsupervised.attribute.String		Correctly Classified Instances	2445 41.5322 %
Instances: 5887		Incorrectly Classified Instances	3442 58.4678 %
Attributes: 3		Kappa statistic	0.2878
VentMediaNorm Vent_SD1_Norm VentanaTipoTram		Mean absolute error	0.2285
Test mode: 10-fold cross-validation		Total Number of Instances 5887	
=== Classifier model (full training set) ===		=== Detailed Accuracy By Class ===	
RandomTree		TP Rate	FP Rate
=====		Precision	Recall
VentMediaNorm < 0.25		F-Measure	ROC Area
VentMediaNorm < -0.59		Class	
Vent_SD1_Norm < 0.05		0.503	0.076
Vent_SD1_Norm < -0.85 : A (419/214)		0.578	0.503
Vent_SD1_Norm >= -0.85 : A (397/110)		0.339	0.132
Vent_SD1_Norm >= 0.05		0.366	0.339
Vent_SD1_Norm < 2.53 : A (105/64)		0.533	0.246
Vent_SD1_Norm >= 2.53 : C (19/1)		0.328	0.533
VentMediaNorm >= -0.59		0.406	0.406
VentMediaNorm < -0.24		0.341	0.114
Vent_SD1_Norm < -1.23 : C (39/7)		0.4	0.341
Vent_SD1_Norm >= -1.23 : B (846/505)		0.369	0.369
VentMediaNorm >= -0.24		0.535	0.14
VentMediaNorm < 0.16 : C (1439/974)		0.491	0.808
VentMediaNorm >= 0.16 : C (387/268)		0.114	0.005
VentMediaNorm >= 0.25		0.701	0.114
VentMediaNorm < 0.67		0.196	0.865
VentMediaNorm < 0.4		0.415	0.129
Vent_SD1_Norm < -0.31 : D (145/108)		0.451	0.415
Vent_SD1_Norm >= -0.31 : D (479/257)		0.415	0.406
VentMediaNorm >= 0.4		0.781	
Vent_SD1_Norm < 0.21 : D (316/191)			
Vent_SD1_Norm >= 0.21 : E (461/295)			
VentMediaNorm >= 0.67			
VentMediaNorm < 0.91			
Vent_SD1_Norm < 0.11 : E (94/49)			
Vent_SD1_Norm >= 0.11 : E (340/157)			
VentMediaNorm >= 0.91			
VentMediaNorm < 1.3 : E (306/142)			
VentMediaNorm >= 1.3 : F (95/22)			
Size of the tree : 31			
Max depth of tree: 4			
Time taken to build model: 0.03 seconds			
		=== Confusion Matrix ===	
		a	b
		c	d
		e	f
		-- classified as	
		507	228
		11	34
		0	0
		a = A	
		240	366
		359	90
		25	0
		b = B	
		79	274
		576	119
		32	0
		c = C	
		24	96
		363	366
		223	0
		d = D	
		27	29
		171	232
		562	29
		e = E	
		0	7
		61	96
		365	68
		f = F	

Figura 61: Clasificador “RandomTree” en el DT (Tipo de Tramo)

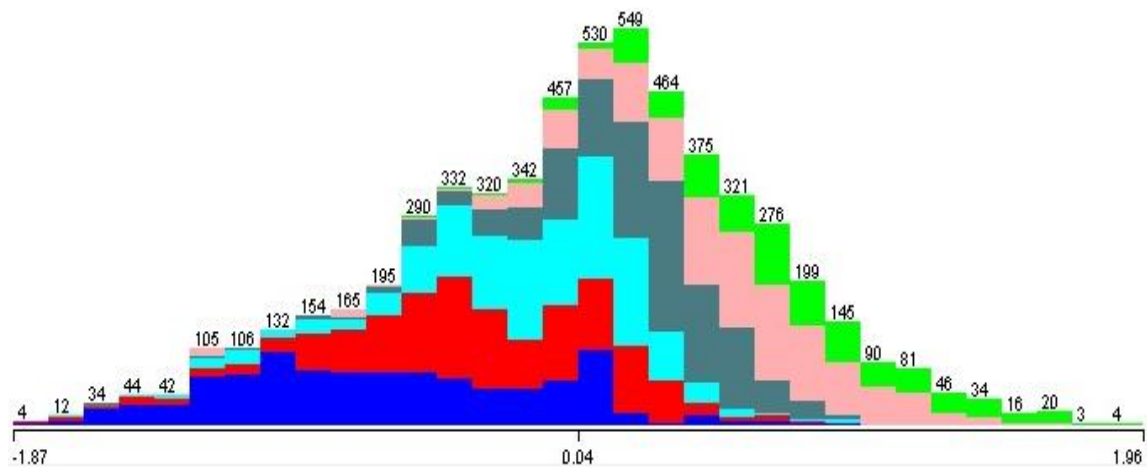
Según los resultados obtenidos, la mejor clasificación ha requerido de información normalizada de la media y el SD1, y señala una tendencia alineada con el tipo de tramo según la matriz de confusión obtenida. En la gráfica de la Figura 62 se puede apreciar el comportamiento decreciente de los diferentes niveles de error (x=0 corresponde a los correctamente clasificados o aciertos, x=1 significa errores de nivel 1, i.e. elemento de la matriz de confusión cuya columna se aleja en uno de la diagonal principal o confusión entre dos tramos seguidos, e.g. entre tramos A y B):



**Figura 62: Tendencia de errores para “RandomTree” en el DT (Tipo de Tramo)**

Tal y como se aprecia en la gráfica anterior, existe una buena tendencia de los errores (el error es decreciente a medida que dos tramos están más alejados en el tiempo) pero no de forma óptima: Aunque los dos primeros niveles cubren el 82,84% de todas las instancias, los errores de nivel 1 (41,31%) aún son tan importantes como los aciertos (41,53%).

En resumen, la media normalizada (Figura 63) es la variable independiente más directamente relacionada con “Tipo de tramo” (la media normalizada va creciendo a medida que se avanza en los tramos), y allí donde dicha media no puede, la normalización del SD1 ayuda a mejorar el rendimiento del clasificador:



**Figura 63: Histograma de la media normalizada (Tipo de Tramo)**

En cualquier caso, aunque la tendencia inferida por el clasificador “*RandomTree*” es de interés ya que el error es decreciente y la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) cubren el 82,84%, el rendimiento debería mejorar para que los aciertos (41,53%) fueran claramente superiores a los errores de nivel 1 (41,31%) y los “*TP Rate*” por clase mejoraran.

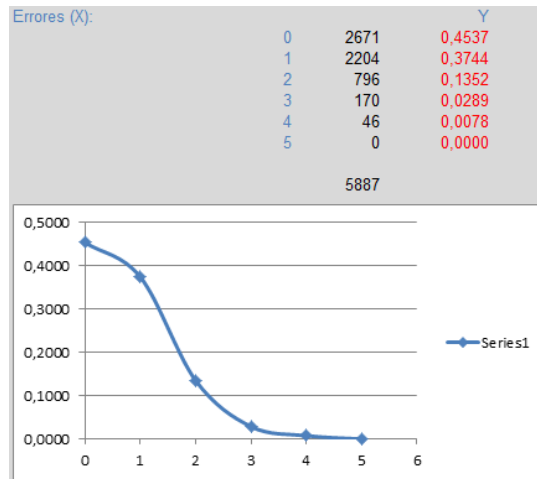
## Resultados del clasificador 2: “Entropía”

De forma similar, y en base al conjunto de variables, se ha realizado el análisis del comportamiento de las mismas en base al segundo clasificador “*Lazy KStar*”, generando como mejor resultado el de la Figura 64, siguiendo el criterio de mejor rendimiento con el mínimo número de atributos (*Attributes*):

=== Run information ===		=== Stratified cross-validation ===						
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===						
Relation: VentanaTipoTram-weka.filters.unsupervised.attribute.StringToNominal-R2		Correctly Classified Instances	2671 45.3712 %					
Instances: 5887		Incorrectly Classified Instances	3216 54.6288 %					
Attributes: 3		Kappa statistic	0.3342					
VentMediaNorm		Mean absolute error	0.2316					
Vent_SD1_Norm								
VentanaTipoTram								
Test mode: 10-fold cross-validation								
=== Classifier model (full training set) ===								
KStar Beta Verion (0.1b).								
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).								
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).								
KStar options : -B 20 -M a								
Time taken to build model: 0 seconds								
		Total Number of Instances	5887					
		=== Detailed Accuracy By Class ===						
		TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
		0.536	0.064	0.634	0.536	0.581	0.87	A
		0.412	0.117	0.441	0.412	0.426	0.781	B
		0.492	0.194	0.363	0.492	0.417	0.77	C
		0.507	0.164	0.407	0.507	0.452	0.779	D
		0.537	0.124	0.485	0.537	0.51	0.834	E
		0.079	0.003	0.734	0.079	0.142	0.88	F
		Weighted Avg.	0.454	0.12	0.491	0.454	0.442	0.813
		=== Confusion Matrix ===						
		a	b	c	d	e	f	<- classified as
		540	198	196	51	23	0	a = A
		183	445	302	132	18	0	b = B
		95	233	531	202	19	0	c = C
		20	105	234	544	169	0	d = D
		14	19	139	297	564	17	e = E
		0	9	62	110	369	47	f = F

Figura 64: Clasificador “*Lazy KStar*” en el DT (Tipo de Tramo)

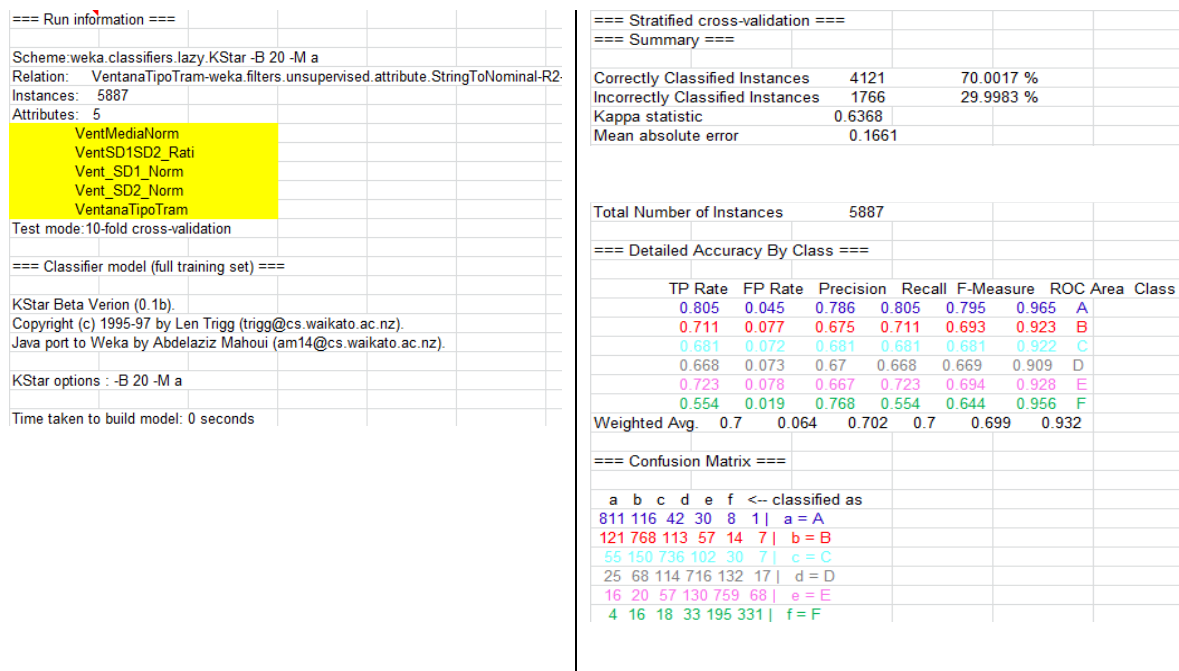
Según los resultados obtenidos, la clasificación con información normalizada de la media y el SD1 indica una tendencia más acusada que la inferida con el clasificador 1, y más alineada con el tipo de tramo según la matriz de confusión obtenida. En la gráfica de la Figura 65 se puede apreciar el comportamiento decreciente de los diferentes niveles de error:



**Figura 65: Tendencia de errores para “Lazy KStar” en el DT (Tipo de Tramo)**

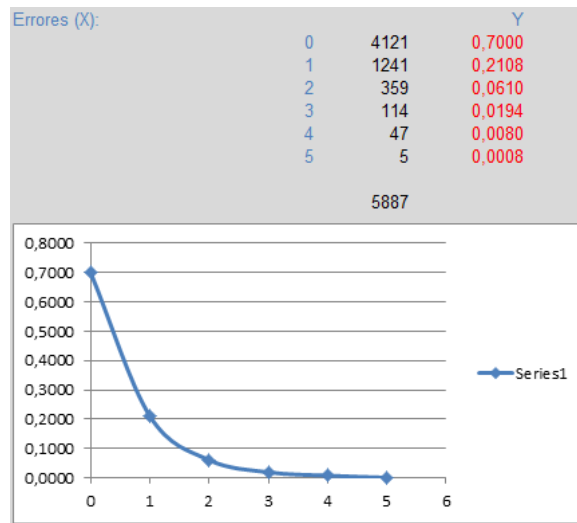
Tal y como se aprecia en la gráfica anterior, existe una tendencia muy marcada de los errores (el error es claramente decreciente a medida que dos tramos están más alejados en el tiempo): los dos primeros niveles cubren el 82,81% de todas las instancias y además los errores de nivel 1 (37,44%) son claramente inferiores a los aciertos (45,37%).

La **matriz de confusión** muestra una marcada diagonal principal, y señala claramente un comportamiento diferenciado de este grupo de variables según el concepto de entropía. Adicionalmente, los ratios mejoran considerablemente si se tuvieran en cuenta dos más de las variables significativas (SD2 normalizado y ratio SD1/SD2) (Figura 66):



**Figura 66: Clasificador “Lazy KStar” en el DT (Tipo de Tramo – 4 variables)**

La Figura 67 se muestra la significativa mejora en el comportamiento decreciente de los diferentes niveles de error:



**Figura 67: Tendencia de errores para “Lazy KStar” en el DT (Tipo de Tramo – 4 variables)**

En resumen, según este criterio de clasificación completamente diferente al anterior, el conjunto de la información formado por la media, el SD1 y el SD2 normalizados significan un sistema muy ligado a la variable dependiente “Tipo\_Tramo”, integrando los resultados obtenidos en los tres estudios anteriores. Todo ello en base a la información generada con el clasificador “*Lazy KStar*”, con rendimientos significativamente superiores para el mismo número de parámetros (2). Además, los resultados mejoran de forma importante al tener en cuenta también el SD2 normalizado y el ratio SD1/SD2: la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) llegan al 91,08%, el rendimiento de los aciertos puros (70%) son muy superiores a los errores de nivel 1 (21,08%) y los “*TP Rate*” por clase mejoran considerablemente.

#### 4.4.2. Análisis en el dominio de la frecuencia

El análisis en el dominio de la frecuencia (Tabla 24) para el caso de estudio 4 se caracteriza por definir como variable dependiente a la variable “**Tipo\_Tramo**” y como variables independientes a aquellas en el dominio de la frecuencia denominadas como “**Conjunto DF**” en el apartado “Variables y Modelo de datos” que han demostrado tener los mejores rendimientos en los tres casos de estudio anteriores:

	C.Estudio 1		C.Estudio 2		C.Estudio 3	
	CL1	CL2	CL1	CL2	CL1	CL2
<b>Variables:</b>	LF_Power HF_Power	LF_Power HF_Power	VLF_PeakY HF_Power HF_XC	HF_Power HF_XC	VLF_Power_No LF_Power_No HF_Power_No	VLF_Power_No LF_Power_No HF_Power_No

Tabla 24: Resumen de variables significativas en el DF para casos de estudio 1 a 3

### Resultados del clasificador 1: “RandomTree”

Basado en las variables anteriores se ha llevado a cabo el análisis de las mismas en base al clasificador “RandomTree” generando el correspondiente árbol de decisión, y siguiendo el criterio de mejor rendimiento con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*), se ha seleccionado como mejor resultado el de la Figura 68:

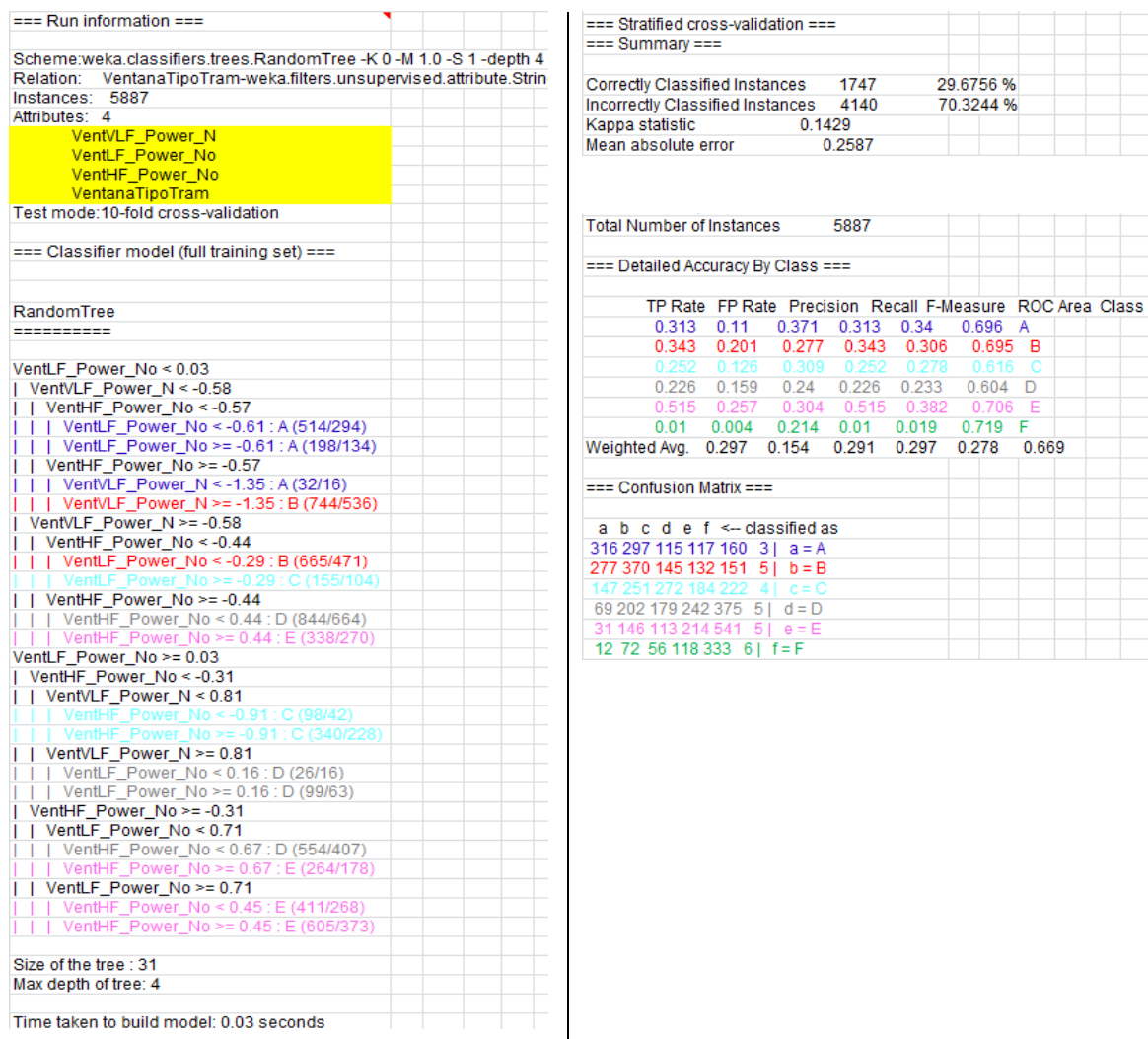
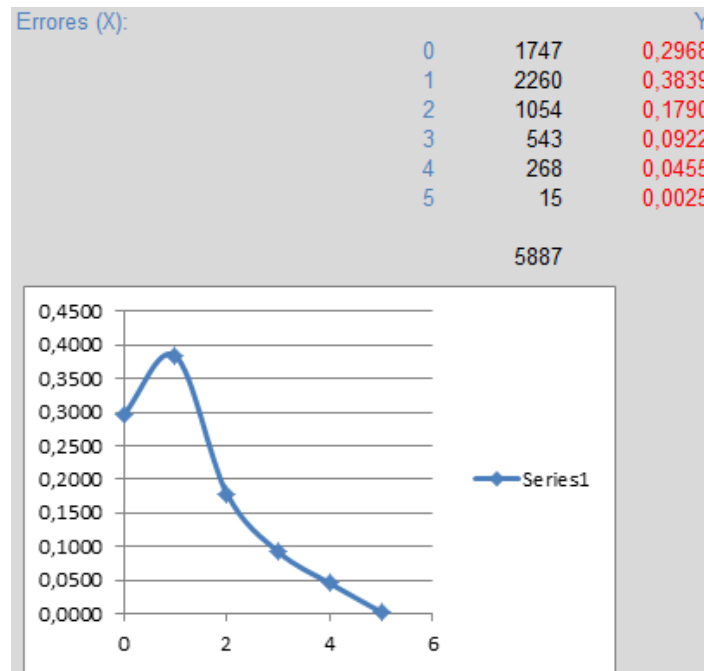


Figura 68: Clasificador “RandomTree” en el DF (Tipo de Tramo)

La mejor clasificación requiere de información normalizada del PSD en todas las frecuencias (VLF, LF y HF), señalando una **débil tendencia** alineada con el tipo de tramo según la matriz de confusión obtenida. Aunque en la gráfica de la Figura 69 se puede apreciar el comportamiento decreciente de los diferentes niveles de error el número de errores leves son superiores al de aciertos y los “*TP Rate*” por clase son bajos ( $x=0 \Rightarrow$  aciertos,  $x=1 \Rightarrow$  errores de nivel 1 o confusión entre dos tramos seguidos):



**Figura 69: Tendencia de errores para “RandomTree” en el DF (Tipo de Tramo)**

En cualquier caso, aunque la tendencia inferida por el clasificador “*RandomTree*” es de interés, ya que el error es decreciente y la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) cubren casi el 70%, el rendimiento debería mejorar para que los aciertos (29,68%) fueran superiores a los errores de nivel 1 (38,39%), y los “*TP Rate*” por clase aún tendrían que mejorar significativamente.

## **Resultados del clasificador 2: “Entropía”**

De igual manera, basándonos en el mismo conjunto de variables, se ha realizado el análisis del comportamiento de las mismas en base al segundo clasificador “*Lazy KStar*”, generando como mejor resultado el de la Figura 70. Se sigue el criterio de mejor rendimiento con el mínimo número de atributos (*Attributes*):

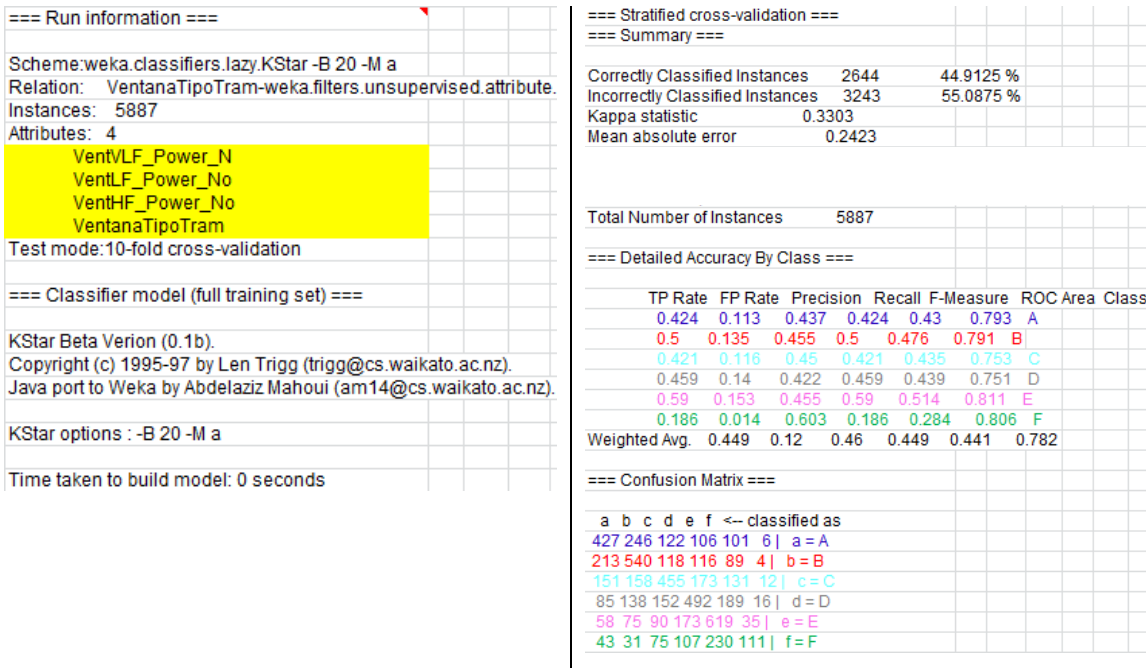


Figura 70: Clasificador “Lazy KStar” en el DF (Tipo de Tramo)

Según los resultados obtenidos, la clasificación con información normalizada de las tres bandas de frecuencia, señala una tendencia más clara que la inferida con el clasificador anterior. En consecuencia, según la matriz de confusión obtenida, los resultados son mejores. En la gráfica de la Figura 71 se puede apreciar el comportamiento decreciente de los diferentes niveles de error:

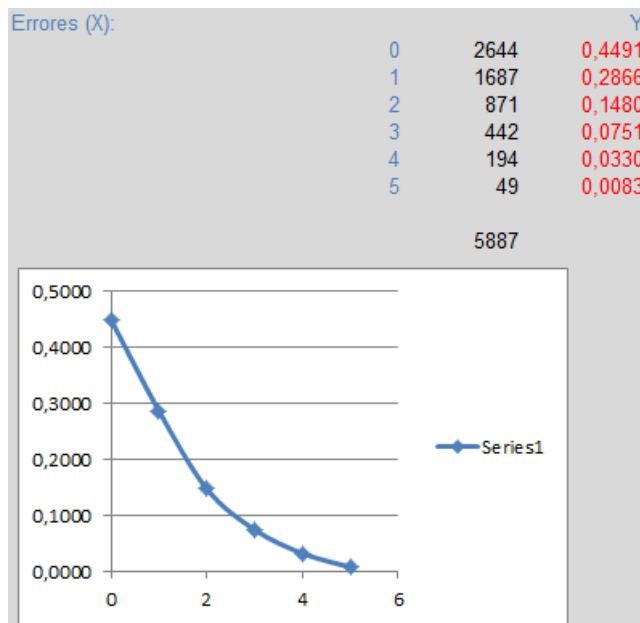
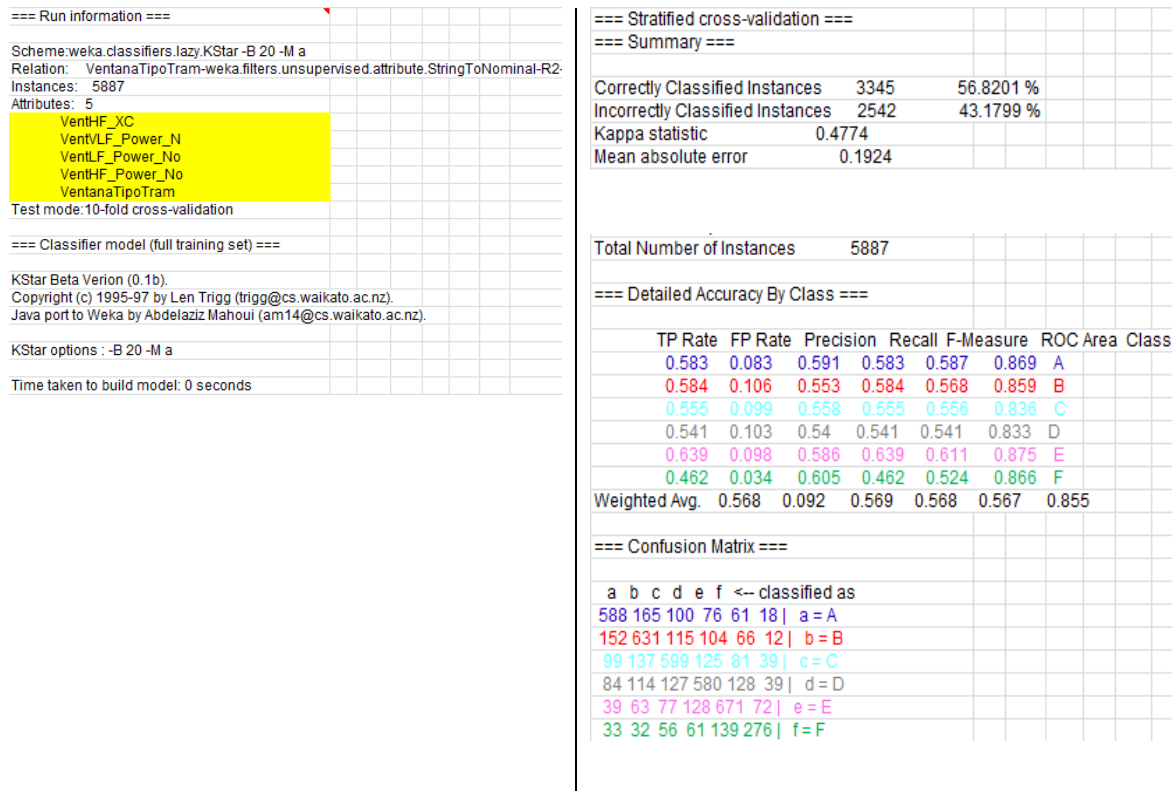


Figura 71: Tendencia de errores para “Lazy KStar” en el DF (Tipo de Tramo)

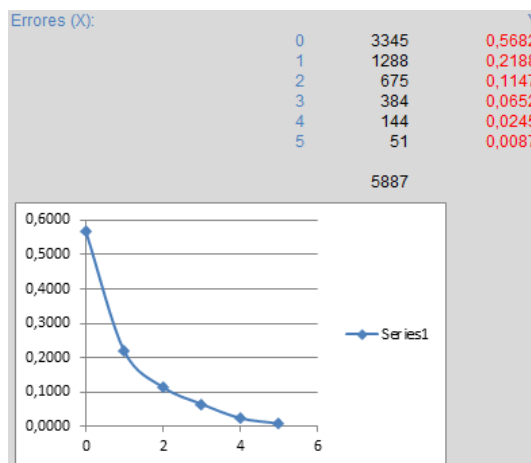
Tal y como se aprecia en la gráfica anterior, existe una tendencia muy acusada de los errores (el error es claramente decreciente a medida que dos tramos están más alejados en el tiempo): los dos primeros niveles cubren el 73,57% de todas las instancias y además los errores de nivel 1 (28,66%) son claramente inferiores a los aciertos puros (44,91%).

La **matriz de confusión** muestra una acusada diagonal principal y señala claramente un comportamiento diferenciado de este grupo de variables según el concepto de entropía. Adicionalmente (Figura 72) los ratios, mejoran significativamente si se tuviera en cuenta también una más de las variables significativas (+ HF\_XC):



**Figura 72: Clasificador “Lazy KStar” en el DF (Tipo de Tramo – 4 variables)**

La Figura 73 muestra la significativa mejora en el comportamiento decreciente de los diferentes niveles de error:



**Figura 73: Tendencia de errores para “Lazy KStar” en el DF (Tipo de Tramo – 4 variables)**

En resumen, según este criterio de clasificación completamente diferente al anterior, el conjunto de información formado por el PSD de VLF, LF y HF normalizados significan un sistema ligado a la variable dependiente “Tipo\_Tramo”, integradora de los resultados obtenidos en los tres estudios anteriores. Todo ello en base a la información generada con el clasificador “*Lazy KStar*”, con rendimientos significativamente superiores para el mismo número de parámetros (3). Además, los resultados mejoran al tener en cuenta la variable HF\_XC: la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) llegan el 78,70%, el rendimiento de los aciertos puros (56,82%) son muy superiores a los errores de nivel 1 (21,88%), y los “*TP Rate*” por clase mejoran considerablemente.

#### 4.4.3. Análisis en el dominio combinado

El análisis al combinar variables en el dominio del tiempo y la frecuencia para este caso de estudio se caracteriza por definir como variable dependiente a la variable “**Tipo\_Tramo**”, y como variables independientes aquellas que basadas en los tres casos de estudio previos han demostrado tener mejores rendimientos (Tabla 25):

	C.Estudio 1		C.Estudio 2		C.Estudio 3	
	CL1	CL2	CL1	CL2	CL1	CL2
<b>Variables DT:</b>	SD1 SD2	SD1 SD2	Media SD1_Nor SD1/SD2	Media SD1 SD1/SD2	Media_Norm SD1_Norm	Media_Norm SD1_Norm
<b>Variables DF:</b>	LF_Power HF_Power	LF_Power HF_Power	VLF_PeakY HF_Power HF_XC	HF_Power HF_XC	VLF_Power_No LF_Power_No HF_Power_No	VLF_Power_No LF_Power_No HF_Power_No

**Tabla 25: Resumen de variables significativas en el DT y DF para casos de estudio 1 a 3**

## Resultados del clasificador 1: “RandomTree”

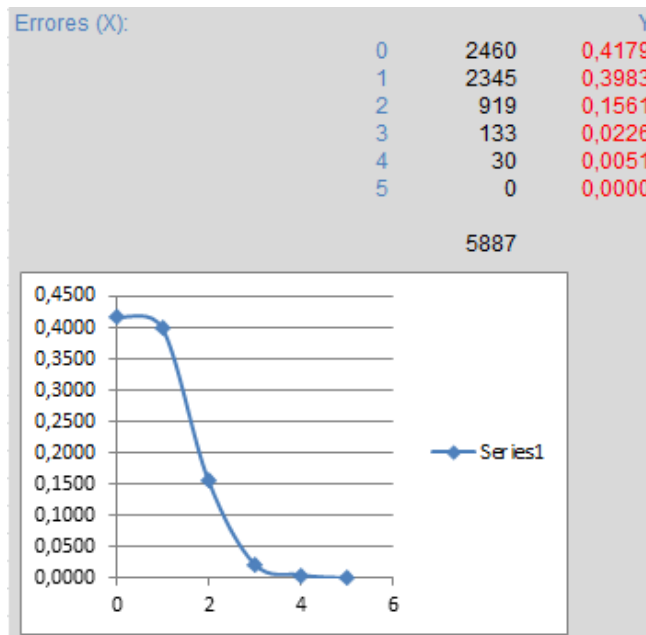
Según el cuadro resumen anterior se ha llevado a cabo el análisis con el clasificador “RandomTree” generándose el correspondiente árbol de decisión, y siguiendo el criterio de máxima “ROC Area” con el mínimo número de atributos (*Attributes*) y el mínimo tamaño de árbol (*Size of tree & Max.depth of tree*). Finalmente, se ha seleccionado como mejor resultado el de la Figura 74:

<pre> Scheme:weka.classifiers.trees.RandomTree -K 0 -M 1.0 -S 1 -depth 4 Relation: VentanaTipoTram-weka.filters.unsupervised.attribute.String Instances: 5887 Attributes: 3   VentMediaNorm   VentLF_Power_No   VentanaTipoTram Test mode:10-fold cross-validation  === Classifier model (full training set) ===  RandomTree =====  VentMediaNorm &lt; 0.25   VentMediaNorm &lt; -0.59     VentLF_Power_No &lt; 0.17       VentMediaNorm &lt; -0.92 : A (384/83)       VentMediaNorm &gt;= -0.92 : A (369/198)       VentLF_Power_No &gt;= 0.17         VentMediaNorm &lt; -1.52 : D (6/0)         VentMediaNorm &gt;= -1.52 : C (181/116)     VentMediaNorm &gt;= -0.59       VentMediaNorm &lt; -0.24         VentLF_Power_No &lt; -0.28 : B (539/284)         VentLF_Power_No &gt;= -0.28 : C (346/223)       VentMediaNorm &gt;= -0.24         VentLF_Power_No &lt; 0.13 : C (1251/870)         VentLF_Power_No &gt;= 0.13 : C (575/372)   VentMediaNorm &gt;= 0.25     VentMediaNorm &lt; 0.67       VentMediaNorm &lt; 0.4         VentLF_Power_No &lt; -0.08 : D (298/198)         VentLF_Power_No &gt;= -0.08 : D (326/167)         VentMediaNorm &gt;= 0.4           VentLF_Power_No &lt; 0.19 : D (360/232)           VentLF_Power_No &gt;= 0.19 : E (417/254)       VentMediaNorm &gt;= 0.67         VentMediaNorm &lt; 0.91           VentLF_Power_No &lt; -0.81 : E (51/28)           VentLF_Power_No &gt;= -0.81 : E (383/178)         VentMediaNorm &gt;= 0.91           VentMediaNorm &lt; 1.3 : E (306/142)           VentMediaNorm &gt;= 1.3 : F (95/22)  Size of the tree : 31 Max depth of tree: 4  Time taken to build model: 0.02 seconds </pre>	<pre> Time taken to build model: 0.02 seconds  === Stratified cross-validation === === Summary ===  Correctly Classified Instances   2460      41.787 % Incorrectly Classified Instances 3427      58.213 % Kappa statistic                  0.2908 Mean absolute error              0.2267  Total Number of Instances       5887  === Detailed Accuracy By Class ===        TP Rate  FP Rate  Precision  Recall  F-Measure  ROC Area  Class       0.494    0.073    0.585     0.494    0.535     0.834    A       0.207    0.057    0.449     0.207    0.284     0.765    B       0.686    0.325    0.322     0.686    0.438     0.74     C       0.367    0.13     0.386     0.367    0.376     0.746    D       0.51     0.12     0.479     0.51     0.494     0.809    E       0.114    0.005    0.701     0.114    0.196     0.861    F Weighted Avg. 0.418    0.128    0.468     0.418    0.401     0.786  === Confusion Matrix ===    a  b  c  d  e  f  &lt;-- classified as 498 116 353 31 10 0   a = A 233 224 512 101 10 0   b = B 80 110 741 127 22 0   c = C 21 46 423 393 189 0   d = D 20 3 207 255 536 29   e = E 0 0 68 110 351 68   f = F </pre>
--	---

Figura 74: Resultado del clasificador “RandomTree” en el DC (Tipo de Tramo)

La mejor clasificación requiere de información normalizada de la media y el PSD en la frecuencia LF, que indican cierta **tendencia** alineada con el tipo de tramo según la matriz de confusión obtenida. En la gráfica de la Figura 75 se puede apreciar el comportamiento decreciente de los diferentes niveles de error, aunque el número de errores leves (39,83%)

son similares a los de aciertos (41,79%) y los “*TP Rate*” por clase son bajos ( $x=0 \Rightarrow$  aciertos,  $x=1 \Rightarrow$  errores de nivel 1 o confusión entre dos tramos seguidos):



**Figura 75: Tendencia de errores para “RandomTree” en el DF (Tipo de Tramo)**

En cualquier caso, aunque la tendencia inferida por el clasificador “*RandomTree*” es de interés ya que el error es decreciente y la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) cubren casi el 81,62%, el rendimiento debería mejorar para que los aciertos (41,79%) fueran más claramente superiores a los errores de nivel 1 (39,83%) y los “*TP Rate*” por clase mejoraran significativamente.

## **Resultados del clasificador 2: “Entropía”**

Partiendo del mismo conjunto de variables, se ha realizado el análisis del comportamiento de las mismas según el clasificador “*Lazy KStar*”, generándose como mejor resultado el de la Figura 76 (mejor rendimiento de clasificación con el mínimo número de atributos):

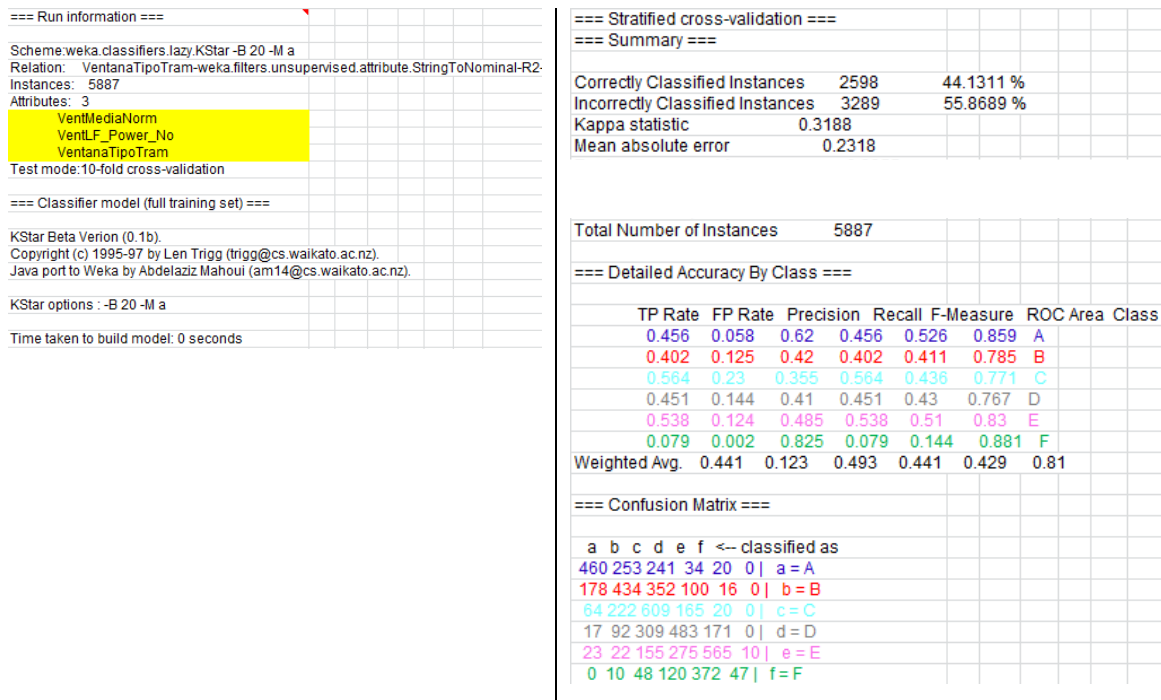


Figura 76: Clasificador “Lazy KStar” en el DC (Tipo de Tramo)

Según los resultados obtenidos, la clasificación con información normalizada de la media y el PSD de LF nuevamente señala una tendencia más clara que la inferida con el clasificador anterior. En consecuencia, todo apunta a unos resultados más alineados con el tipo de tramo según la matriz de confusión obtenida. En la gráfica de la Figura 77 se puede apreciar el comportamiento decreciente de los diferentes niveles de error:

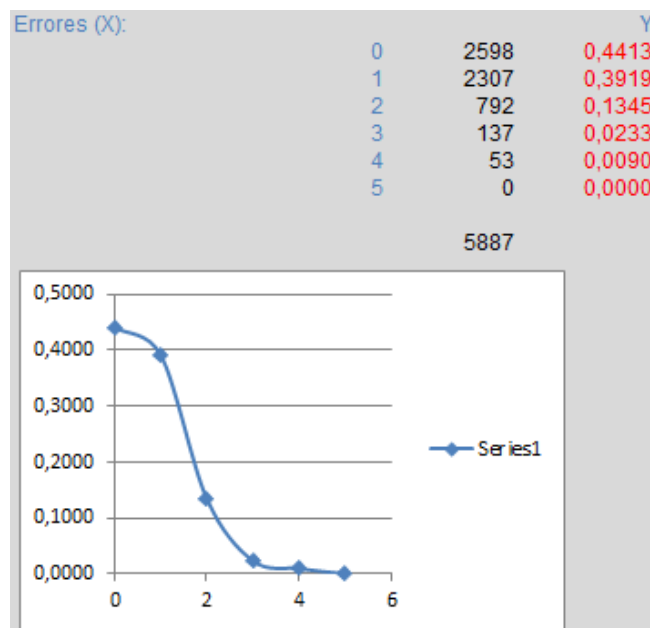


Figura 77: Tendencia de errores para “Lazy KStar” en el DC (Tipo de Tramo)

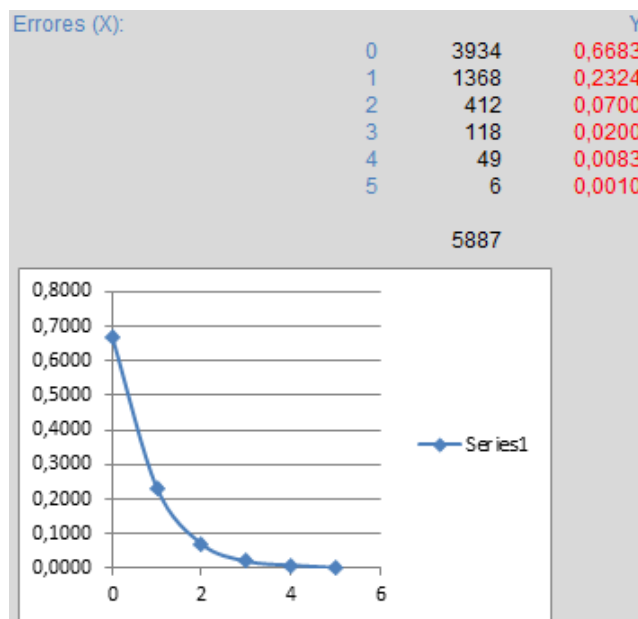
Tal y como se aprecia en la gráfica anterior, existe una tendencia muy acusada de los errores (el error es claramente decreciente a medida que dos tramos están más alejados en el tiempo): los dos primeros niveles cubren el 83,32% de todas las instancias y los errores de nivel 1 (39,19%) se mantienen inferiores a los aciertos (44,13%).

La **matriz de confusión** muestra una clara diagonal principal y señala un comportamiento diferenciado de estas dos variables según el concepto de entropía. Además, los ratios mejoran significativamente si se tienen en cuenta también dos más de las variables significativas (+ SD1 y PSD de HF normalizados) (Figura 78):

=== Run information ===		=== Stratified cross-validation ===				
Scheme: weka.classifiers.lazy.KStar -B 20 -M a		=== Summary ===				
Relation: VentanaTipoTram-weka.filters.unsupervised.attribute.StringToNominal-R2-		Correctly Classified Instances	3934 66.8252 %			
Instances: 5887		Incorrectly Classified Instances	1953 33.1748 %			
Attributes: 5		Kappa statistic	0.5981			
VentMediaNorm VentLF_Power_No VentHF_Power_No Vent_SD1_Norm VentanaTipoTram		Mean absolute error	0.1804			
Test mode: 10-fold cross-validation		Total Number of Instances 5887				
=== Classifier model (full training set) ===		=== Detailed Accuracy By Class ===				
KStar Beta Verion (0.1b).		TP Rate	FP Rate			
Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).		Precision	Recall			
Java port to Weka by Abdelaziz Mahoui (am14@cs.waikato.ac.nz).		F-Measure	ROC Area			
KStar options : -B 20 -M a		Class				
Time taken to build model: 0 seconds		0.768	0.054			
		0.746	0.768			
		0.757	0.951			
		0.951	A			
		0.656	0.081			
		0.646	0.656			
		0.651	0.903			
		0.903	B			
		0.641	0.081			
		0.641	0.641			
		0.641	0.898			
		0.898	C			
		0.649	0.09			
		0.616	0.649			
		0.632	0.892			
		0.892	D			
		0.719	0.083			
		0.653	0.719			
		0.684	0.915			
		0.915	E			
		0.516	0.015			
		0.796	0.516			
		0.626	0.936			
		0.936	F			
		Weighted Avg.	0.668 0.072 0.673 0.668 0.667 0.914			
=== Confusion Matrix ===						
a	b	c	d	e	f	<-- classified as
774	130	57	25	19	3	a = A
164	709	115	57	27	8	b = B
57	160	692	137	27	7	c = C
24	72	138	696	125	17	d = D
16	20	63	152	755	44	e = E
3	6	15	62	203	308	f = F

**Figura 78: Clasificador “Lazy KStar” en el DC (Tipo de Tramo – 4 variables)**

La Figura 79 muestra la significativa mejora en el comportamiento decreciente de los diferentes niveles de error:



**Figura 79: Tendencia de errores para “Lazy KStar” en el DC (Tipo de Tramo – 4 variables)**

En resumen, según este criterio de clasificación completamente diferente al anterior, el conjunto de la información formado por la media, SD1 y el PSD de LF y HF normalizados significan un sistema muy ligado a la variable dependiente “Tipo\_Tramo”, integradora de los resultados obtenidos en los tres estudios anteriores. Todo ello en base a la información generada con el clasificador “*Lazy KStar*”, con rendimientos algo superiores para el mismo número de parámetros (Media + PSD de LF normalizados), y que mejoran considerablemente al tener en cuenta dos más de las variables significativas (SD1 y el PSD de HF normalizados): la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) llegan el 90,07%, el rendimiento de los aciertos puros (66,83%) son muy superiores a los errores de nivel 1 (23,24%) y los “*TP Rate*” por clase mejoran considerablemente.

#### 4.4.4. Conclusiones

Como resumen (Tabla 26), en base a los tres estudios anteriores (DT + DF + DC) llevados a cabo entorno al tipo de tramo (variable dependiente), se puede concluir que:

##### **Dominio del Tiempo (DT):**

- 1- Tipo de tramo (DT): La normalización de algunas variables optimiza los rendimientos de clasificación ligados a la variable “Tipo\_Tramo”, donde la media normalizada es la variable independiente más significativa en el dominio del tiempo y se complementa con información del SD1 normalizado para optimizar rendimientos de clasificación.

- 2- Tipo de tramo (DT): La clasificación con el “*RandomTree*” obtiene ratios significativos que infieren una tendencia que liga la variable “Tipo\_Tramo” con el grupo de variables formado por la media y el SD1 normalizados, pero no de forma concluyente ya que aunque el error es decreciente y la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) cubren el 82,84%, los aciertos (41,53%) aún son un similares a los errores de nivel 1 (41,31%) y los “*TP Rate*” por clase deberían mejorar.
- 3- Tipo de tramo (DT): La clasificación con el “*Lazy KStar*” obtiene ratios muy superiores con una tendencia muy acusada de los errores (el error es claramente decreciente a medida que dos tramos están más alejados en el tiempo): los dos primeros niveles cubren el 82,81% de todas las instancias, los errores de nivel 1 (37,44%) son inferiores a los aciertos (45,37%), y los “*TP Rate*” por clase mejoran considerablemente.
- 4- Tipo de tramo (DT): El rendimiento del clasificador 2 basado en la entropía mejora significativamente al considerar también al SD2 normalizado y el ratio SD1/SD2 en el sistema de variables independientes: la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) llegan el 91,08%, el rendimiento de los aciertos puros es del 70%, muy superior a los errores de nivel 1 (21,08%), y los “*TP Rate*” por clase mejoraron considerablemente.

**Dominio de la Frecuencia (DF):**

- 5- Tipo de tramo (DF): La normalización de algunas variables optimiza los rendimientos de clasificación ligados a la variable “Tipo\_Tramo”, donde El PSD normalizado para las tres bandas de frecuencia consideradas forman el conjunto de variables independientes más significativas en el dominio de la frecuencia.
- 6- Tipo de tramo (DF): La clasificación con el “*RandomTree*” obtiene ratios bajos no concluyentes, y aunque indican una tendencia que liga la variable “Tipo\_Tramo” con el grupo de variables formado por el PSD de VLF, LF y HF normalizados (error decreciente y la suma de aciertos más errores leves suponen casi el 70%), los aciertos (29,68%) son inferiores a los errores de nivel 1 (38,39%) y los “*TP Rate*” por clase deberían ser mejores.
- 7- Tipo de tramo (DF): La clasificación con el “*Lazy KStar*” obtiene ratios muy superiores con una tendencia muy acusada de los errores (el error es claramente decreciente a medida que dos tramos están más alejados en el tiempo): los dos primeros niveles cubren el 73,57% de todas las instancias, los errores de nivel 1 (28,66%) son inferiores a los aciertos (44,91%), y los “*TP Rate*” por clase mejoran considerablemente.
- 8- Tipo de tramo (DF): El rendimiento del clasificador 2 basado en la entropía mejora si se considera la variable significativa HF\_XC en el sistema de variables independientes: la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) llegan el 78,70%, el rendimiento de los aciertos puros es del 56,82%, muy superior a los errores de nivel 1 (21,88%), y los “*TP Rate*” por clase mejoraron significativamente.

**Dominio Combinado (DC):**

- 9- Tipo de tramo (DC): La normalización de variables optimiza los rendimientos de clasificación ligados a la variable “Tipo\_Tramo”, donde la media y el PSD de LF normalizados forman el conjunto de variables independiente más significativas en el dominio combinado.
- 10- Tipo de tramo (DC): La clasificación con el “*RandomTree*” obtiene ratios bajos no concluyentes, y aunque indican una tendencia que liga la variable “Tipo\_Tramo”

con variables como la media y el PSD de LF normalizados (error decreciente y la suma de aciertos más errores leves suponen el 83,32%), los aciertos (41,79%) son similares a los errores de nivel 1 (39,83%) y los “*TP Rate*” por clase deberían mejorar.

11- Tipo de tramo (DC): La clasificación con el “*Lazy KStar*” para el mismo número de variables obtiene ratios superiores con una tendencia muy acusada de los errores (el error es claramente decreciente a medida que dos tramos están más alejados en el tiempo): los dos primeros niveles cubren el 83,32% de todas las instancias, los errores de nivel 1 (39,19%) son inferiores a los aciertos (44,13%), y los “*TP Rate*” por clase mejoran.

12- Tipo de tramo (DC): El rendimiento del clasificador 2 basado en la entropía mejora considerablemente al integrar las variables significativas SD1 y PSD de HF normalizadas en el sistema de variables independientes: la suma de aciertos más errores leves o de nivel 1 ( $x=1$ ) llegan el 90,07%, el rendimiento de los aciertos puros es del 66,83%, muy superior a los errores de nivel 1 (23,24%), y los “*TP Rate*” por clase mejoraron significativamente.

#### Cuadro resumen (Caso de estudio 4):

	DT		DF		DC	
	CL1	CL2	CL1	CL2	CL1	CL2
Nº Variables:	3	3	3	3	2	4
Variables:	Media_Norm SD1_Norm SD2_Norm	Media_Norm SD1_Norm SD2_Norm	VLF_Power_No LF_Power_No HF_Power_No	VLF_Power_N LF_Power_No HF_Power_No	SD1_Norm LF_Power_No	Media_Norm SD1_Norm LF_Power_No HF_Power_No
Nº Clases (n):	6	6	6	6	6	6
Instancias (n):	5.887	5.887	5.887	5.887	5.887	5.887
ROC área (%):	77,7	87,5	66,9	78,2	78,6	91,4
Correctly Classified (%):	40,3	55,9	29,7	44,9	41,8	66,8
Kappa statistic:	0,2753	0,4632	0,1429	0,3303	0,2908	0,5981
Mean absolute error:	0,2325	0,2102	0,2587	0,2423	0,2267	0,1804

Tabla 26: Tabla resumen de los rendimientos de los clasificadores (Tipo de Tramo)



## 5. CONCLUSIONES FINALES

### 5.1. Conclusiones

Este trabajo concluye en un sistema que se orienta a un nuevo método de evaluación y validación de interfaces de usuario para la conducción de vehículos de transporte por carretera, basado en la señal ECG y su derivada *HRV*. Además, según los patrones construidos con este sistema de variables y el clasificador basado en instancias (Lazy K-Star) y el concepto de entropía, se consigue descubrir la mejor correlación ligada al avance del tiempo y el consumo de los recursos de la atención.

De esta manera, como conclusión de todos los resultados obtenidos anteriormente, se define un sistema de cuatro de variables: **Media, SD1, PSD de LF y PSD de HF**, todas ellas **normalizadas**, como el más integrador de factores personales (e.g. Edad), temporales (e.g. Ciclos circadianos) y ambientales (e.g. Tipo de carretera).

La necesidad de normalizar las cuatro variables para alcanzar los mejores ratios de clasificación refuerza la calidad del método al considerar factores personales, (en el caso de estudio 1 ya se identificó al SD1 como la variable que mejor se correlacionaba con la edad como factor personal, y con cualquiera de las estrategias de clasificación utilizadas). Esto se debe a que la normalización aplicada no es más un proceso en el que teniendo en cuenta información exclusiva de cada participante (media y desviación típica de cada variable a lo largo del experimento) traslada cada variable a un espacio común más comparable. Las consecuencias de esto son, por una parte el mejor rendimiento del clasificador, pero por otra parte la restricción de poder utilizar el método solo *a posteriori* (las medias y desviaciones solo se conocen al final de cada experimento). En consecuencia, el método es especialmente interesante para las actividades de validación de sistemas *ITS* a la hora de evaluar si un proceso, dispositivo, actividad, etc... nueva es mejor o peor que otro (mayor o menor consumidor de recursos de la atención).

Basado en los resultados de esta investigación (sistema de cuatro variables normalizadas con información del *HRV*, tanto en el dominio del tiempo como en el de la frecuencia, analizado según un método de inteligencia artificial basado en el concepto de “desorden” o

“entropía” y tratado por técnicas de minería de datos), se describe a continuación la aplicación-objetivo que de forma novedosa se está comenzando a aplicar a la hora de comparar diferentes maneras de interactuar con los vehículos de transporte.

En consecuencia, la nueva aplicación permite **evaluar y comparar diferentes maneras de interactuar con vehículos de transporte por carretera, y es potencialmente extensible a otras aplicaciones de interés.**

Basado en todas las conclusiones anteriores, se puede afirmar que el conjunto de los resultados obtenidos confirman la hipótesis de partida:

**Hipótesis:** La información contenida en la variabilidad de la frecuencia cardíaca o HRV forma un sistema completo y suficiente para evaluar de forma comparativa diferentes sistemas inteligentes de transporte o ITS aplicables en los interfaces hombre-máquina o HMI de los vehículos por carretera.

Finalmente, en relación a las actividades relacionadas con la temática de esta tesis, como resultado de trabajos preliminares y como producto de la labor de diseminación que se ha hecho durante la misma, se pueden destacar el siguiente listado de publicaciones en las que el doctorando ha contribuido, y defendido en la mayoría de los casos:

- Murgoitio, J.; Ferros, M.; Goti, A.; Larburu, M.; Rodríguez, T. (2007): Contributions to the Ambient Intelligence: Networking Sensors & Actuators for a New Active Headrest, 11th International Forum on Advanced Microsystems for Automotive Applications, Berlin, Alemania; 09 – 10 May. 2007.
- Murgoitio, J.; Fernández, J.I. (2008): Car driver monitoring by networking vital data, AMAA, International Forum on Advanced Microsystems for Automotive Applications, 2008.
- Murgoitio, J.; Arejita, B., Antúnez, M. (2009): Wellness monitoring in cars with CAN, CAN Newsletter. December 2009.

- Murgoitio, J.; Arejita, B.; Larburu, M.; Sánchez, J. (2010): A Bluetooth Sensor Network Base on the IEEE1451 Standard. In Proceedings of the International Conference on Wireless Information Network and Systems WINSYS 2010. Athens. 2010.
- Murgoitio, J.; Antúnez, M. (2011): Herramienta software para el estudio de la variabilidad de la frecuencia cardíaca (HRV): Aplicación a la medida del nivel de alerta en la conducción de vehículos, (vol. 86, núm. 1, p. 28), Dyna (Febrero/2011).
- Murgoitio, J.; Arejita, B.; Larburu, M.; Sánchez, J. (2011): A Bluetooth Sensor Network Based on the IEEE 1451 Standard. A Sensor Network Solution to Evaluate the Wellness of the Passenger and Improve the Safety in Cars, Computer Science Research and Technology. (vol. 2, p. 107) (2011). ISBN: 978-1-61122-724-6.
- Murgoitio, J.; Sánchez, J.; Cases, J.I.; Izaguirre, M. (2013): Iniciativas ITS entorno al concepto PRT en el parque tecnológico de Miramón (San Sebastián). XIII Congreso Español sobre Sistemas Inteligentes de Transporte. Proceedings, San Sebastián (Junio/2013).
- Murgoitio, J.; Aranbarri, A.; Gutiérrez, J.L.; Gil, M.J. (2013): Heart Rate Variability. Knowing More about HRV Analysis and Fatigue in Transport Studies. Cardiotecnix 2013: International Congress on Cardiovascular Technologies (Vilamoura). Proceedings, 107-114.
- Murgoitio, J. (2013): Local, National and European projects on Autonomous Vehicles, ITS Ireland Conference in association with RITS-Net. Proceedings, Dublin (Nov/2013).
- Reina, N.A.; Martínez, N.; Seepold, R.; Murgoitio, J. (2012): *A UPnP Service to Control and Manage IEEE 1451 Transducers in Control Networks*. IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, vol. 61 Issue: 3, p. 791-800. DOI: 10.1109/TIM.2011.2170501. Mar 2012.
- Urquiza, A.; Murgoitio, J.; Larburu, M. (2011): A Context evaluator model for a multimodal user interface in driving activity. ICAART 2011 - 3rd International Conference on Agents and Artificial Intelligence, Rome, Italy; Jan. 2011.

## 5.2. Líneas futuras

En base a los análisis realizados con técnicas de minería de datos (Weka) y los métodos de inteligencia artificial utilizados, además de los cuatro estudios llevados a cabo se han identificado las siguientes otras líneas de investigación que se estiman serían de interés y que son susceptibles de avanzar en ellas según algunos de los resultados obtenidos y las tendencias detectadas:

- 1- Firma cardíaca: Definición de una firma cardíaca que en base a variables extraídas del análisis HRV permita identificar a una persona en una muestra de tamaño X.
- 2- SD1 y SD2: Análisis del comportamiento del SD1 y del SD2 como marcadores de salud.
- 3- LF: Estudio del comportamiento de las frecuencias (en especial LF) para determinar la edad fisiológica del corazón en relación a la edad cronológica (definición de un coeficiente) a lo largo de nuestra vida (revisiones periódicas).
- 4- Estudio de la frontera HF y LF (ver caso de estudio EDAD) para diagnósticos relacionados con determinados problemas de salud.
- 5- Estudio para el ajuste de los tiempos de descanso requeridos en función de la edad y los ciclos circadianos (M/T/N).
- 6- Extensión de este estudio HRV a los ciclos de conducción nocturna (en este trabajo no se han realizado registros nocturnos).
- 7- Extensión al análisis de la influencia del clima en el HRV.
- 8- Diseño y construcción de una base de datos (más registros de diferentes edades y turnos de mañana/tarde/noche) como referencia-patrón y en base a las variables significativas de esta monografía y el comportamiento del clasificador de "Entropía". El disponer de mayor cantidad de registros se podría orientar a realizar evaluaciones absolutas (proyecto versus patrón absoluto) más que relativas

(proyecto versus patrón particular), así como a pasar a resoluciones más precisas en el tiempo, e.g. con ventanas de información agrupadas en tramos de conducción de 15 minutos o menos en lugar de cada 30 minutos, tal y como se ha realizado en esta monografía.

- 9- Análisis y mejora en el diseño del concepto de vecindad o “*blending*” utilizado en el clasificador “Lazy K-Star” como representativo del “*jitter*” en el HRV.
- 10- Estudio del comportamiento del clasificador “Lazy K-Star” al variar el “*Blend*” y su correlación con la fatiga o problemas de salud.
- 11- Análisis comparativo de otras técnicas o métodos de minería de datos que optimicen los rendimientos de clasificación.
- 12- Extensión de la aplicación de los resultados de esta monografía a otros ámbitos de interés, i.e. “evaluación/comparación de otros HMIs en otros ámbitos y/o sectores.



## 6. APLICACIÓN

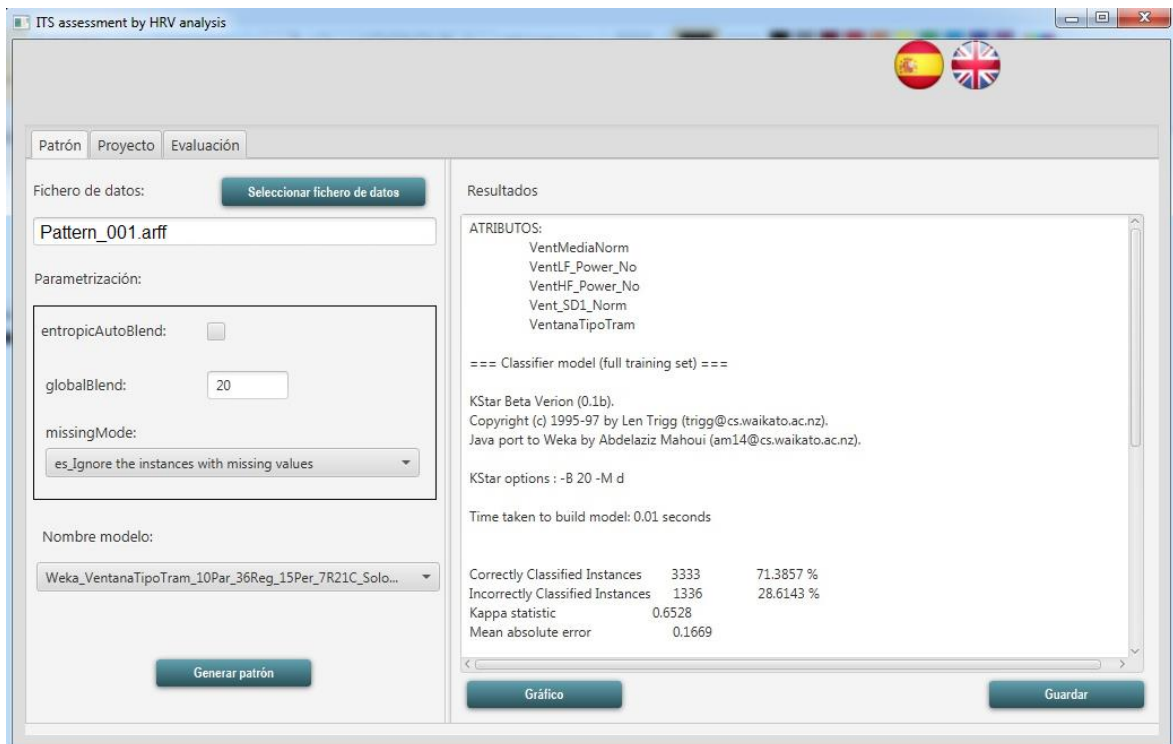
La aplicación que aquí se describe está desarrollada en Java y, como resultado de todo el análisis anterior, basa su motor de inferencia en los métodos de IA (Inteligencia Artificial) y las técnicas de minería de datos que mejores resultados han proporcionado para extraer el más amplio conocimiento (i.e. integración de las variables más ligadas a factores personales, temporales y ambientales) relacionado con el nivel de alerta y la fatiga, el cual se obtiene a partir de un sistema de información basado en cuatro parámetros normalizados extraídos exclusivamente a partir del *HRV*: Media, SD1, PSD de LF y PSD de HF.

Con objeto de mostrar la utilidad de esta aplicación, se explica a continuación el uso de la misma mediante un caso práctico basado en las tres fases (Patrón, Proyecto y Evaluación) de las que consta cualquier evaluación que se quiera realizar con esta herramienta (para más detalle del uso de la aplicación ver manual de usuario correspondiente en Anexos).

### 6.1. Fase 1: “Patrón”

El primer paso para poder hacer una evaluación es generar un patrón a partir de la información contenida en un archivo “.arff” donde se encuentre la información con las variables “Media, SD1, PSD de LF y PSD de HF” normalizadas obtenidas a partir de los registros ECG cuyas condiciones sirvan de referencia para la generación de un patrón sobre el que comparar más tarde en la fase 2 un proyecto.

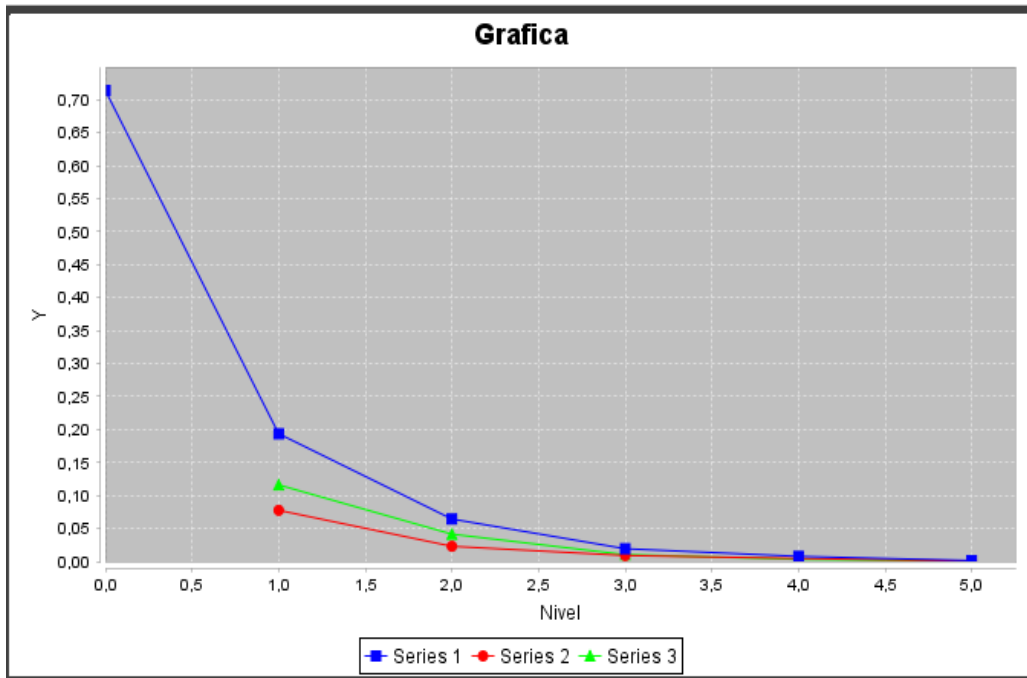
En la Figura 80 se muestra el resultado que se obtiene al definir con la ayuda del botón “**Seleccionar fichero de datos**” el archivo “.arff” donde está la información con las cuatro variables mencionadas anteriormente y, tras pulsar el botón “**Generar patrón**”, generar el resultado que aparece en “**Resultados**”:



**Figura 80: Obtención del “Patrón”**

Si el operador considera suficientemente bueno el resultado en función de los datos obtenidos podrá guardar como modelo de referencia al pulsar “**Guardar**”, lo que generará un archivo “\*.model” donde almacenar el patrón correspondiente basado en el clasificador “*Lazy KStar*”. La bondad del patrón generado es algo arbitrario aunque en general criterios como los siguientes suelen considerarse bastante buenos [López et al., 1999; Altman DG, 1991]: *Correctly Classified Instances* > 70%; *Kappa statistic* > 60% [Everitt, B., 1992; López et al., 1999; Altman DG, 1991]; y  $\sum Y$  para niveles 0 y 1 > 80% del gráfico de la Figura 81. El gráfico se obtiene al pulsar “**Grafico**” en la aplicación según la Figura 80 y se basa en la matriz de confusión y el concepto de “Tendencia de los errores” manejado en el análisis del caso de estudio 4 (línea azul).

Las líneas verde y roja distinguen el comportamiento de la matriz de confusión por encima (línea roja) y por debajo (línea verde) de la diagonal principal. La suma de las ordenadas de las líneas verde y roja debe ser el valor de la correspondiente ordenada en la línea azul. Además, para que el patrón sea bueno debería cumplirse que la tendencia de los errores por encima de la diagonal principal de la matriz de confusión sean de similar comportamiento a la que se dan por debajo de la misma, tal y como ocurre en la Figura 81:



**Figura 81: Gráfico del “Patrón”**

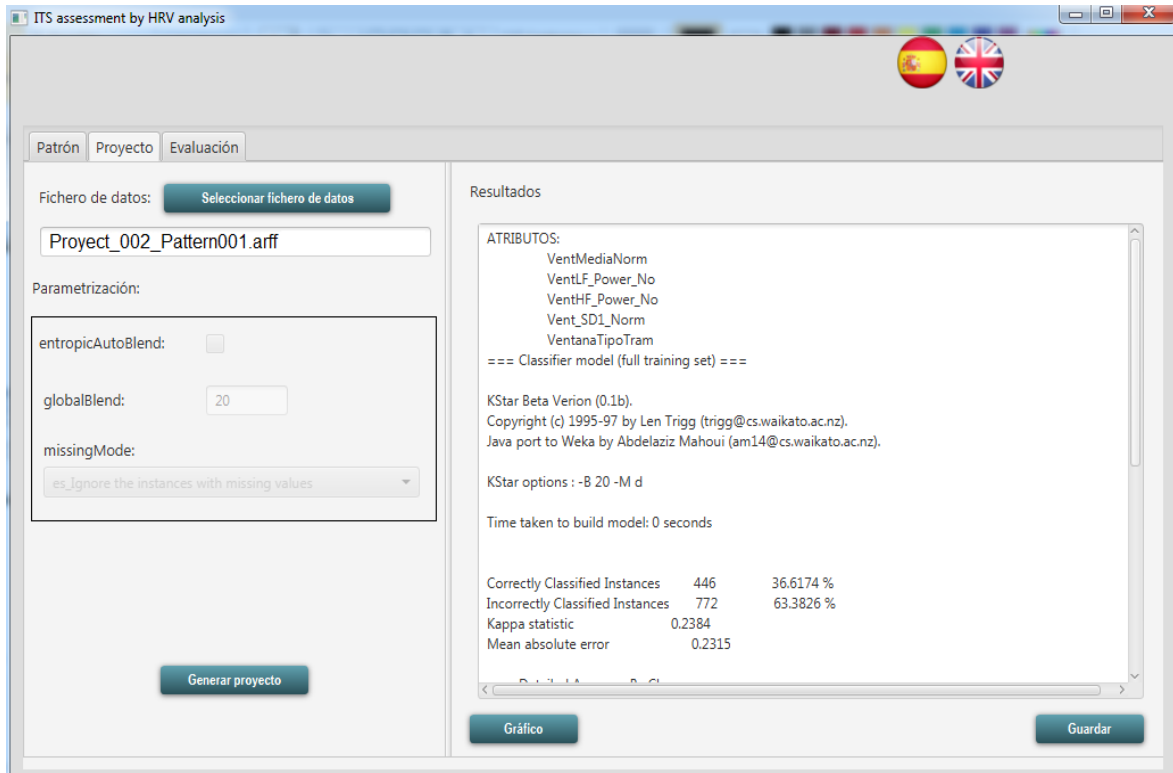
Este patrón (Pattern\_001.arff) está basado en los registros que sirvieron para el análisis del caso de estudio 4 y cuya conducción fue considerada más deportiva, menos segura y por tanto más consumidora de los recursos de la atención según el modelo de Kahneman [1973], que tal y como indicábamos en nuestro estado del arte es modelo de referencia en todos los modelos actuales (ver “Modelos teóricos” sobre la atención), y sobre todo aquellos más relacionados con esta investigación ligada a las neurociencias y el conexionismo.

En otro estudio, e.g. conducción de vehículo con volante tradicional versus conducción de vehículo con joystick, el patrón sería el basado en los registros obtenidos a partir de varios conductores **conduciendo con volante tradicional**.

## 6.2. Fase 2: “Proyecto”

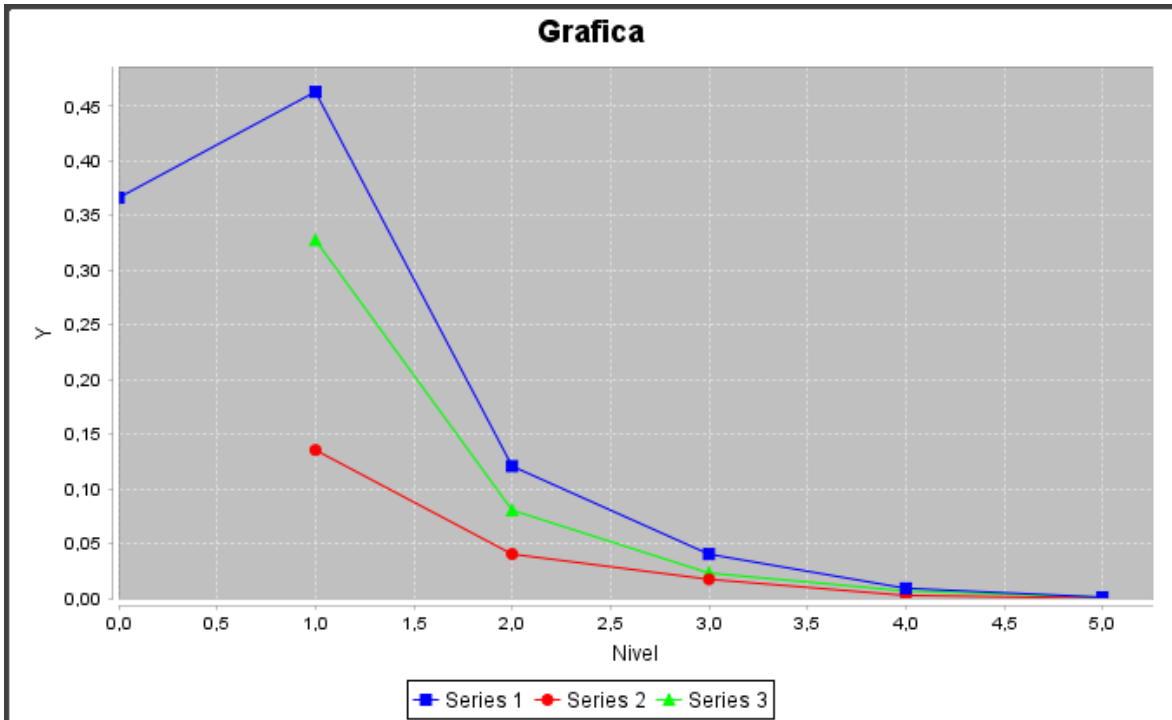
Una vez obtenido un patrón con un modelo consistente se puede pasar a la fase 2, la cual se completa por medio del menú “**Proyecto**” al pulsar el botón “**Seleccionar fichero de datos**” y elegir un archivo “.arff”, el cual también contiene las variables “Media, SD1, PSD de LF y PSD de HF” normalizadas, y cuya información se compara contra el modelo

construido según el patrón anterior al pulsar en “**Generar proyecto**”. La Figura 82 muestra el resultado obtenido al enfrentar el proyecto “Project002\_Pattern\_001.arff” contra el modelo o patrón anterior construido según “Pattern\_001.arff”:



**Figura 82: Selección del “Proyecto”**

Si el operador considera interesante el resultado en función de los datos obtenidos podrá guardar la comparación contra el modelo al pulsar “**Guardar**”, lo que generará un archivo donde se almacena esta valoración. El gráfico correspondiente se obtiene al pulsar “**Gráfico**” en la aplicación según la Figura 82 y se basa en la matriz de confusión y el concepto de “Tendencia de los errores” manejado en el análisis del caso de estudio 4 (línea azul). Al igual que lo explicado para el patrón, las líneas verde y roja distinguen el comportamiento de la matriz de confusión por encima (línea roja) y por debajo (línea verde) de la diagonal principal. La suma de las ordenadas de las líneas verde y roja es el valor de la ordenada en la línea azul (Figura 83):



**Figura 83: Gráfico del “Proyecto”**

Este proyecto (Project002\_Pattern\_001.arff) está basado en registros que sirvieron para el análisis del caso de estudio 4 y cuya conducción fue considerada menos deportiva y más segura, y por tanto menos consumidora de los recursos de la atención. Así, el resultado al compararlo con el patrón muestra diferencias sobre todo en los niveles 0 y 1 (comparar los dos gráficos anteriores): Los aciertos de la diagonal principal en el gráfico del patrón se desplazan sobre todo hacia el nivel 1. Como se verá en la fase 3, para la evaluación final es importante analizar para qué parte de la matriz de confusión se desplazan las diferencias entre ambos gráficos.

En otro estudio, e.g. conducción de vehículo con volante tradicional versus conducción de vehículo con joystick, el proyecto sería el basado en los registros obtenidos a partir de varios conductores **conduciendo con joystick**.

### 6.3. Fase 3: “Evaluación”

En la última fase, se recoge de forma conjunta la información obtenida en las dos fases anteriores y se valoran los resultados del proyecto de forma relativa al patrón generado en base a unos umbrales que determinarán:

- 1- Si el patrón es lo suficientemente consistente, es decir, si cumple las cuatro condiciones que se muestran en la zona de la Figura 84 etiquetada como “**Zona Evaluación Patrón**”. En caso afirmativo se pone el semáforo correspondiente en verde, dando sentido a valorar los datos del área etiquetada “**Zona Evaluación Proyecto**”.
- 2- Si los resultados del proyecto tienen información alineada con el patrón, así como coherente, para llevar a cabo una evaluación concluyente, i.e. la mayor y más relevante parte de la información se concentra en los niveles importantes (e.g. 0 y 1), y la diferencia de comportamiento entre los datos por encima y por debajo de la diagonal principal de la matriz de confusión es significativa (diferencia entre el comportamiento de las líneas verde y rojo del gráfico). En caso afirmativo se pone el semáforo correspondiente en verde y se activa la zona etiquetada “**Zona Evaluación Final**”.
- 3- Por último, si se han cumplido las condiciones anteriores según los umbrales establecidos (los umbrales son configurables), se nos permite indicar una banda (ver en “**Zona Evaluación Final**”) según la cual se determinará si los resultados del proyecto concluyen que los datos obtenidos en “Proyecto” con respecto al patrón (la valoración es relativa, no absoluta) son de alguna manera significativos de un mayor (semáforo rojo) o menor (semáforo verde) consumo de nuestros recursos de la atención, y por lo tanto, más o menos conveniente, seguro, fatigoso, etc.

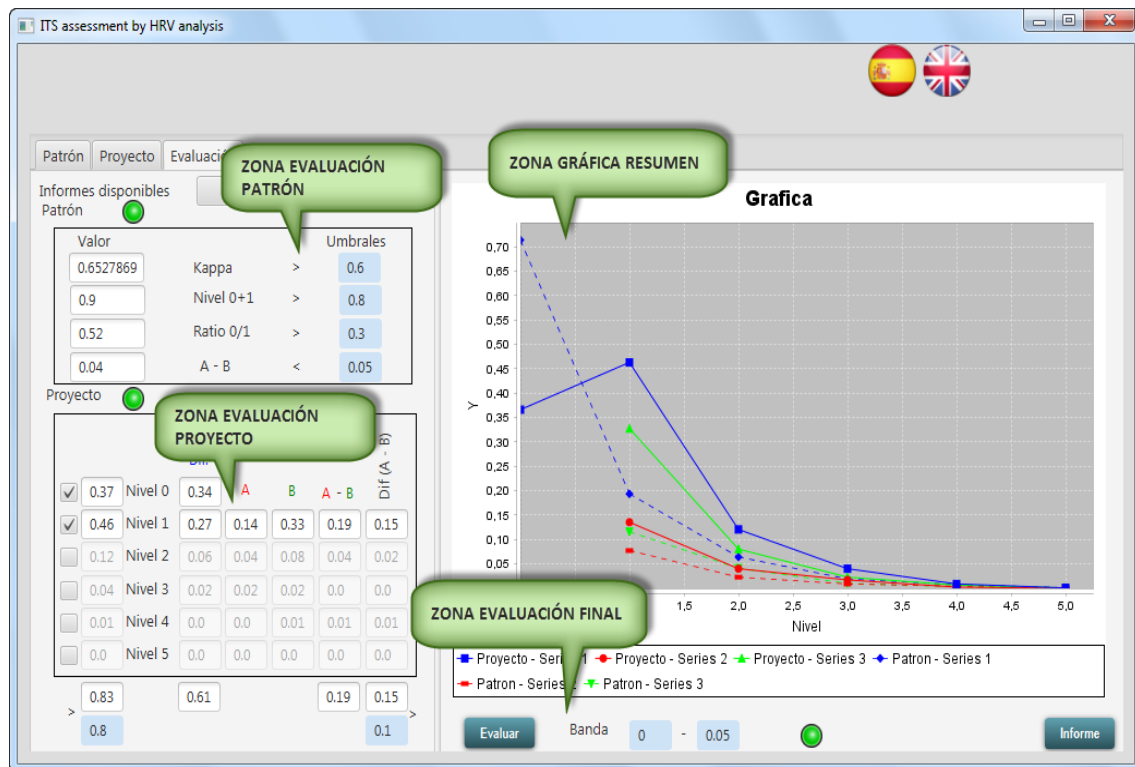


Figura 84: Menú “Evaluación” (Patrón=“OK”, Proyecto=“OK”)

Según la Figura 84 y los umbrales fijados se puede concluir que, en base al sistema de variables extraído del HRV y fijado por este trabajo de investigación, el motor de inferencia diseñado determina un menor consumo (semáforo de “Zona Evaluación Final” en verde) de los recursos de nuestra atención que los consumidos bajo las condiciones en las que se registraron los mismos datos para el patrón de referencia.

El informe correspondiente de todo esta evaluación se genera con el botón “Informe” y queda registrado según un archivo pdf que guarda para su trazabilidad toda la información relevante.

Ejemplo de pantalla cuando el patrón es consistente (Figura 85), los datos obtenidos para un proyecto superan las condiciones (umbrales) que permitan llevar a cabo una valoración con garantías, y la evaluación determina que las diferencias con el patrón NO son significativas (semáforo amarillo):

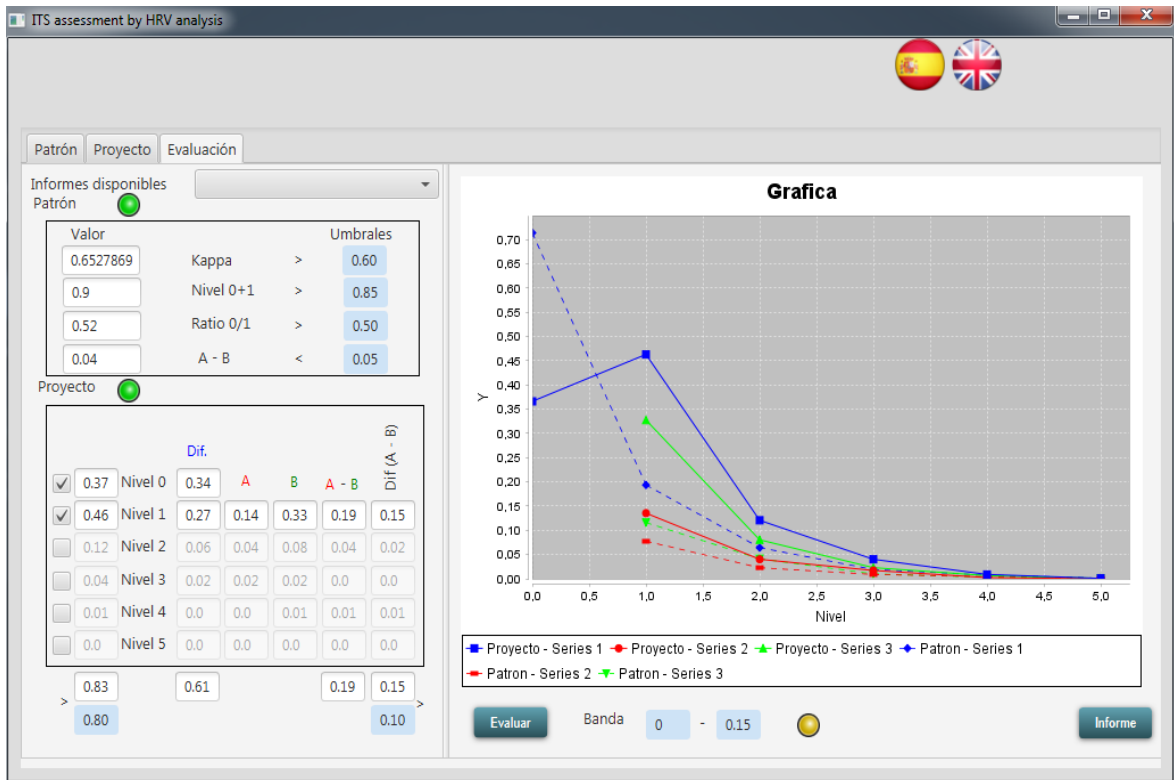


Figura 85: Menú “Evaluación” (Patrón=”OK”, Proyecto= ”OK” & Cambio NO significativo)

## 7. BIBLIOGRAFÍA

- Akay, G. (1995): Flow-induced phase inversion in powder structuring by polymers. In, *Polymer Powder Technology*, Narkis M. and Rosenweig N. (Eds.), Chapter 20, Wiley p. 542-587.
- Akerstedt, T. (1994): Accidents and sleepiness: a consensus statement from the International Conference on work hours, sleepiness and accidents, Stockholm. *J. Sleep Res.*, 3:195, 1994.
- Akerstedt, T. (1998): Is there an optimal sleep-wake pattern in shift work? Invited paper 1998; 24 suppl 3:18-27.
- Akselrod, S.; Gordon, D.; Ubel, F.A.; Shannon, D.C.; Barger A.C.; Cohen R.J. (1981): *Power spectrum analysis of heart rate fluctuation: a qualitative probe of beat to beat cardiovascular control*, *Science*, 213:220222, 1981.
- Altman DG. (1991): *Practical statistics for medical research*. New York: Chapman and Hall; 1991.
- Arnedt, J.T.; Wilde, G.J.; Munt, P.W.; MacLean, A.W. (2001): How do prolonged wakefulness and alcohol compare in the decrements they produce on a simulated driving task? *Accid Anal Prev* 33: 337–344.
- ASDA (1992): Sleep Disorders Atlas Task Force. EEG arousals: scoring rules and examples. *Sleep* 1992;15:173-84.
- Atkinson, R.C.; Shiffrin, R.M. (1968): Human memory: A proposed system and its control process. *The Psychology of learning and motivation*, 2, 2, 89-195.
- Baars, B.J. (1987): What is conscious in the control of action? A modern ideomotor theory of voluntary control. Learning. *The Ebbinghaus Centennial Conference*. Hillsdale, N.J.: LEA.
- Baars, B.J. (1988): *A Cognitive Theory of Consciousness*. Cambridge, MA: Cambridge University Press.

- Baddeley, A.; Hitch, G. (1974): Working memory. En G.H. Bower (ed.), The psychology of learning and motivation, vol. 8, p. 47-89. Nueva York: Academic.
- Baker, C.A.; Burns, N.M.; Simonson, E.; Keiper, C. (1966): Electrocardiogram changes in prolonged automobile driving. *Percept Mot Skills*. 1966 Aug; 23 (1): 210.
- Baldwa, R.A. et al. (1977): *The HP ECG analysis program, Trends in Computer-Processed Electrocardiograms*, J.H. vanBcmnel & J.L. Willerns, Eds. North Holland, p. 197-205, 1977.
- Banco Mundial (2014): <http://www.bancomundial.org/pubsdocs/>
- Baselli, G.; Bolis, D.; Cerutti, S.; Freschi, C. (1985): Autoregressive modelling and power spectral estimate of R-R interval time series in arrhythmic patients, *Computers and Biomedical Research*, vol. 18, p. 510-530.
- Baselli, G.; Cerutti, S.; Civardi, S.; Liberati, D.; Lombardi, F.; Malliani A.; Pagani, M. (1986): Spectral and Cross-Spectral Analysis of Heart Rate and Arterial Blood Pressure Variability Signals", *Computers and Biomedical Research*, vol. 19, p. 520-534.
- Baselli, G.; Cerutti, S.; Civardi, S.; Lombardi, F.; Malliani, A.; Merri, M.; Pagani, M.; Rizzo, G. (1987): Heart rate variability signal processing: a quantitative approach as an aid to diagnosis in cardiovascular pathologies, *International Journal of Bio-Medical Computing*, vol. 20, p. 51-70.
- Berger, R.D.; Akselrod, S.; Gordon, D.; Cohen, R.J. (1986): An efficient algorithm for spectral analysis of heart rate variability, *IEEE Transactions on Biomedical Engineering*, vol. 33, p. 900-904.
- Berlyne, D.E. (1960): Conflict, arousal, and curiosity. McGraw-Hill series in psychology. New York, NY, US: McGraw-Hill Book Company. (1960). xii 350 p. <http://dx.doi.org/10.1037/11164-000>.
- Berlyne, D.E. (1961): Conflict and the orientation reaction. *J. exp. Psychol.* 1961; 62:476-483

- Berlyne, D.E.; McDonnell, P. (1965): Effects of stimulus complexity and incongruity on duration of eeg desynchronization. *Electroencephalogr Clin Neurophysiol.* 1965 Feb; 18: 156-61.
- Bermejo, V. (1987): Procesos atencionales y aplicaciones educativas. *Psiquis, VIII*, 46-53.
- Bianchi, A.; Scholz, U.; Kubicki, S.; Cerutti, S. (1991): Heart rate variability and sleep states, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 13, p. 658-659.
- Birkett, C.L.; Kienzle, M.G.; Myers, G.A. (1991): Interpolation Over Ectopic Beats Increases Low Frequency Power in Heart Rate Variability Spectra, *Proceedings of Computers in Cardiology*, p. 257-259.
- Bittner, T.; Stell, J.G. (2002): Approximate qualitative spatial reasoning. *Spatial Cognition and Computation.* Vol. 2, Nr. 4, 435-466.
- Bittner, T.; Stell, J.G. (2002): Vagueness and Rough Location. *geoinformatica*, Vol. 6, 99-121.
- Blitz, P.S.; Hoogstraten, J.; Mulder, G. (1970): Mental load, heart rate and heart rate variability. *Psychologische Forschung*, 33, 277-288.
- Boualem Boashash et al. (2003): *Time Frequency Signal Analysis and Processing.* A Comprehensive reference. Elsevier Ltd.
- Broadbent, D.E. (1954): The role of auditory localization and attention in memory span. *Journal of experimental Psychology*, 47, 191-196.
- Broadbent, D.E. (1958): *Perception and communication.* London: Pergamon Press. (Trad. Cast.; *Percepción y comunicación.* Madrid: Debate).
- Charlton, S.G.; Baas, P.H. (2001): Fatigue, work-rest cycles, and psychomotor performance of New Zealand truck drivers, *New Zealand Journal of Psychology*, 30A, 32-39, 2001.

- Cherry, E.C. (1953): Some experiments on recognition of speech with one and two ears. *Journal of the acoustical Society of America*, 25, 975-979.
- Cerutti, S.; Civardi, S.; Bianchi, A.; Signorini, M.G.; Ferrazzi, E.; Pardi, G. (1989): Spectral analysis of antepartum heart rate variability, *Clinical Physics and Physiological Measurements*, vol. 10, supl. B, p. 27-31.
- Cerutti, C.; Gustin, M.-P.; Paultre, C.-Z.; Lo, M.; Julien, C.; Vincent, M.; Sassard, J. (1991): Autonomic nervous system and cardiovascular variability in rats: a spectral analysis approach. *American Journal of Physiology*, 261, H1292-1299.
- Cerutti, S.; Bianchi, A.; Signorini, M.G.; Cavezzale, M. Cobelli, R.; Lucini, D.; Pagani, M. (1991): Sympatho-vagal balance in humans under stressing conditions. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 13, n° 2, p. 662-663.
- Cinaz, B.; Arnrich, B.; La Marca, R.; Troster, G. (2013): *Monitoring of mental workload levels during an everyday life office-work scenario*. *Personal and ubiquitous computing*. 17, 229-239. 1617-4909; 1617-4917.
- Cleary, J.G.; Trigg, L.E. (1995): “*K\**: An Instance-Based Learner Using an Entropic Distance Measure”. In: 12th International Conference on Machine Learning, p. 108-114. Morgan Kaufmann, San Francisco (1995).
- Cnossen, F.; Rothengatter, T.; Meijman, T. (2000): Strategic changes in task performance in simulated car driving as an adaptive response to task demands. *Transportation Research Part F: Traffic Psychology and Behaviour* (3), p. 123- 140.
- Concejero, P. (2004): Comparación de modelos de curvas ROC para la evaluación de procedimientos estadísticos de predicción en investigación de mercados. Tesis Doctoral. Universidad Complutense de Madrid.
- Coull, J.T.; Middleton, H.C.; Robbins, T.W.; Sahakian, B.J. (1995): Clonidine and diazepam have differential effects on tests of attention and learning.” *Psychopharmacology (Berl)* 120(3):322-32

- Coull, J.T.; Middleton, H.C.; Robbins, T.W.; Sahakian, B.J. (1995): Contrasting effects of clonidine and diazepam on tests of working memory and planning. *Psychopharmacology (Berl)* 120 (3):311-21
- Craelius, W.; Restivo, M.; Assadi, M.A.; El-Sherif, N. (1986): Criteria for Optimal Averaging of Cardiac Signals, *IEEE Transactions on Biomedical Engineering*, vol. BME-33, n° 10, p. 957-966.
- Croxton, F.E. (1959): *Elementary statistics with application in medicine and the biological sciences*. New York: Dover Publications Inc.
- Daubechies, I. (1988): Orthonormal bases of compactly supported wavelets, *Communications on Pure and Applied Mathematics*, vol. 41, p. 909-996.
- Daubechies, I. (1991): A method for time-frequency localization in *Advances in spectrum stimation and array processing*, vol. 11. Englewood Cliffs, New Jersey: Prentice-Hall, 1991.
- Daubechies, I. (1993): Orthonormal bases of compactly supported wavelets II: Variations on a theme, *SIAM Journal of Mathematical Analysis*, vol. 24, n° 2, p. 449-519.
- Davis, H.; Davis, P.; Loomis, A. (1938): Human brain potential during the onset of sleep, *J Neurophysiol*, vol. 1, p. 24-38, 1938.
- De Vega, M. (1984): *Introducción a la Psicología Cognitiva*. Alianza Editorial, Madrid, 1984.
- DeBoer, R.W.; Karemaker, J.M.; Strackee, J. (1984): Comparing Spectra of a Series of Point Events Particularly for Heart Rate Variability Data, *IEEE Transactions on Biomedical Engineering*, vol. BME-31, n° 4, p. 384-387.
- DeBoer, R.W.; Karemaker, J.M.; Strackee, J. (1985): *Hemodynamic fluctuations and baroreflex sensitivity in humans: a beat-to-beat model*, *Am J Physiol*, 253: H680-H689.
- Dember, W.N.; Warm, J.S. (1979): *The psychology of perception* (2nd ed.). New York: Wiley.

- Deutsch, J.A.; Deutsch, D. (1963): Attention some theoretical considerations. *Physiological Review*, 70 (1), 80-90.
- DGT (2014): <http://www.dgt.es/es/seguridad-vial/>
- Diedrich, A.; Drescher, J.; Nalishitij, V.; Kirchner, F. (1994): Accute Effects of Simulated Microgravity on Heart Rate Variability, *Journal of Gravitational Physiology*, vol. 1, p. 35-36.
- Dinges, D.F.; Kribbs, N.B. (1991): Performing while sleepy: Effects of experimentally-induced sleepiness. In T. H. Monk (Ed.), *Sleep, sleepiness and performance*, 1991.
- Dixon, J.W. (1981): Weakness in engineer's professionalism. *Production Engineer* Vol. 60, Issue: 9.
- Durantin, G.; Gagnon, J.F.; Tremblay, S.; Dehais, F. (2014): *Using near infrared spectroscopy and heart rate variability to detect mental overload*. *Behavioural brain research*. 259,16-23. 0166-4328; 1872-7549.
- Engel, D.; Silny, J.; Rau, G.; Rubart M.; Hanrath P. (1990): Diagnosis of ischemia by continuous ECG-monitoring, *Proceedings of Computers in Cardiology*, p. 275-278.
- Ettema, J.H.; Zielhuis, R.L. (1971): Physiological parameters of mental load. *Ergonomics*, 14, 137-144.
- *European Environment Agency* (2014): <http://www.eea.europa.eu/themes/transport>
- Everitt, B. (1992): *The Analysis of Contingency Tables*. Chapman Hall, London.
- Ewing, D.J.; Campbell, I.W.; Clarke, B.F. (1981): Heart rate changes in diabetes mellitus", *The Lancet*, vol. 24, p. 183-185.
- Eysenck, M.W. (1982): *Attention and Arousal. Cognition and Performance*. Springer-Verlag 1, 8.

- Fafrowicz, M.; Unrug, A.; Marek, T.; van Luijtelaar, G.; Noworol, C.; Coenen, A. (1995): Effects of diazepam and buspirone on reaction time of saccadic eye movements. *Neuropsychobiology*. 1995; 32 (3): 156-60.
- Fagerstrom, S.; Lisper, H.O. (1977): Effect of listening to car radio, experience and personality of the driver on subsidiary reaction time and heart-rate in a long-term driving task. En R. R. Mackie (Ed.) *Vigilance*. Plenum Press, New York and London.
- Faust, O.; Yi, L.M.; Hua, L.M. (2013): *Heart Rate Variability Analysis for Different Age and Gender*. *Journal of medical imaging and health informatics*. 3, 395-400. 2156-7018.
- Fawcett, T. (2004): *ROC Graphs: Notes and Practical Considerations for Researchers*. Technical report. Palo Alto (USA): HP Laboratories, 2004.
- Féry, Y.A.; Ferry, A.; Vom Hofe, A.; Rieu, M. (1997): Effect of physical exhaustion on cognitive functioning. *Percept Mot Skills*. 1997 Feb; 84 (1): 291-8.
- Ferrara, M.; Gennaro, L.D. (2001): How much sleep do we need?, *Sleep Medicine Reviews*, 5:2, 155-179, 2001.
- Friesen, G.M.; Jannett, T.C.; Jadallah, M.A.; Yates, S.L.; Quint, S.R.; Nagle, H.T. (1990): *A Comparison of the Noise Sensitivity of Nine QRS Detection Algorithms*, *IEEE Transactions on Biomedical Engineering*, vol. 37, núm. 1, 1990.
- Furlan, R.; Guzzetti, S.; Crivellaro, W.; Dassi, S.; Tinelli, M.; Baselli, G.; Cerutti, S.; Lombardi, F.; Pagani, M.; Malliani, A. (1990): Continuous 24-hour assessment of the neural regulation of systemic arterial pressure and RR variabilities in ambulant subjects, *Circulation*, vol. 81, p. 537-547.
- Gao, Q.; Wang, Y.; Song, F.; Li, ZZ.; Dong, XL. (2013): *Mental workload measurement for emergency operating procedures in digital nuclear power plants*. *Ergonomics*. 56, 1070-1085.0014-0139.
- García Sevilla, J. (1997): *Psicología de la Atención*. Madrid. Editorial Síntesis S.A.P

- Geddes, L.A. (1984): Cardiovascular devices & their applications. New York: John Wiley & Sons.
- Gil, E.A.; Aubert, X.L.; Most, E.I.S.; Beersma, D.G.M.; (2013): *Human Circadian Phase Estimation from Signals Collected in Ambulatory Conditions Using an Autoregressive Model*. Journal of biological rhythms. 28, 152-163. 0748-7304.
- Godthelp, H.; Milgram, P.; Blaauw, G.J. (1984): The development of a time-related measure to describe driving strategy. *Human Factors*, 26, 257-268.
- Grace, R.; Byrne, V.E.; Bierman, D.M.; Legrand, J.M.; Gricourt, D.; Davis, B.K.; Staszewski, J.J.; Carnahan, B.A. (1998): Drowsy driver detection system for heavy vehicles, in Proc. Digital Avionics Systems Conference, 17th DASC. The AIAA/IEEE/SAE, vol. 2, p. 136/1-136/8, 1998.
- Grandjean, E. (1979): Physiologische Arbeitsgestaltung: Leitfaden der Ergonomie
- Grandjean, E. (1988): Fitting the Task to the Man: A Textbook of Occupational Ergonomics
- Grass (1981): <http://www.grasstechnologies.com/>
- Grönlund, J.U.; Antila, K.J.; Siimes, A.S.; Metsälä, T.; Oja, R.; Tuominen, J.; Välimäki, I.A. (1989): Beta-adrenergic control and inter-relationships between heart rate and blood pressure in neonatal lambs, *Medical & Biological Engineering & Computing*, vol. 27, p. 163-170.
- Halberg, F. (1974): Protection by timing treatment according to bodily rhythms - an analogy to protection by scrubbing before surgery. *Suppl 1:27-72*.
- Hales, S. (1733): *Statistical Essays: Containing Haemastaticks*. London, UK: Innys, Manby and Woodward, 1733.
- Hamada, T.; Ito, T.; Adachi, K.; Nakano, T.; Yamamoto, S. (2003): Detecting method for driver's drowsiness applicable to individual features, en Proc. Intelligent Transportation Systems, IEEE, vol. 2, p. 1405-1410, Oct. 2003.

- Hancock, P.A.; Verwey, W.B. (1997): Fatigue, workload and adaptive driver systems. *Accident Analysis and Prevention*, 29 (4), 495-506.
- Harrison, Y.; Horne, J.A. (2000): The impact of sleep deprivation on decision making: a review, *Journal of Experimental Psychology: Applied*, 6:3, 236-249, 2000.
- Hasher, L.; Zacks, R. T. (1979): Automatic and effortful processes in memory. *Journal of Experimental Psychology: General*, 108, 356-388.
- Hattori, S.; Sato, H.; Oh-hira, M.; Ishida, N.; Imamura, Y.; Kawakita, M. (1987): Molecular cloning and nucleotide sequence of P, M and F genes of Newcastle disease virus avirulent strain D26. *Virus Res.* 1987 May; 7 (3): 241-55.
- Hayami, T.; Matsunag, K.; Shidoji, K.; Matsuki, Y. (2002): Detecting drowsiness while driving by measuring eye movement - a pilot study, in *Proc. Intelligent Transportation Systems, The IEEE 5th International Conference on*, 2002, p. 156-161.
- Hirsch, J.A.; Bishop, B. (1981): Respiratory sinus arrhythmia in humans: how breathing pattern modulates the heart rate, *Am. J. Physiol.*, 241:11620-11629, 1981.
- Hoover, A.; Singh, A.; Fishel-Brown, S.; Muth, E. (2012): *Real-time detection of workload changes using heart rate variability*. *Biomedical signal processing and control.* 7, 333-341. 1746-8094.
- Horne, J.A.; Reyner, L.A. (1995): Sleep related vehicle accidents, *British Medical Journal*, vol. 310, núm. 6979, p. 565-567.
- Horne, J.A.; Reyner, L.A. (1995): Counteracting driver sleepiness: effects of napping, caffeine, and placebo, *Psychophysiology*, vol. 33, p. 306-309.
- Huo, C.Y.; Zhuang, J.J.; Huang, X.L.; Hou, F.Z.; Ning, X.B. (2012): *Heart rate variability analysis based on modified Poincare plot*. *Acta physica sinica.* 61, 190506. 1000-3290.

- Hsu, C.H.; Tsai, M.Y.; Huang, G.S.; Lin, T.C.; Chen, K.P.; Ho, S.T.; Shyu, L.Y.; Li, C.Y. (2012): *Poincare plot indexes of heart rate variability detect dynamic autonomic modulation during general anesthesia induction*. *Acta anaesthesiologica Taiwanica: official journal of the Taiwan Society of Anesthesiologists*. 50, 12-8. 1875-452X.
  
- Ian H. Witten; Eibe Frank; Len Trigg; Mark Hall Geoffrey Holmes, Sally Jo Cunningham (1999): *Weka: Practical machine learning tools and techniques with java implementations*. Department of Computer Science. University of Waikato. New Zealand. <http://www.cs.waikato.ac.nz/~ml/publications/1999/99IHW-EF-LT-MH-GH-SJC%-Tools-Java.pdf>.
  
- Idogawa, K. (1989): *A time variation of Professional Driver's EEC in Monotonous Work*, IEEE Engineering in Medicine & Biology Society, 11th annual international conference, p. 719-720, 1989.
  
- Ito, T.; Mita, S.; Kozuka, K.; Nakano, T.; Yamamoto, S. (2002): *Driver blink measurement by the motion picture processing and its application to drowsiness detection*, in *Proc. Intelligent Transportation Systems, The IEEE 5th International Conference on*, 2002, p. 168-173.
  
- Jonides, J.; Yantis, S. (1985): *Review of Visual Information Processing*. *American Journal of Psychology*, 1984, 97, 134-137
  
- Johnston, W.A.; Heinz, S.P. (1978): *Flexibility and capacity demands of attention*. *Journal of experimental Psychology: General*, 107 (4), 420-435.
  
- Kahneman, D. (1973): *Attention and Effort*. Prentice Hall, New Jersey.
  
- Kallaway, P.; Maulsby, R. (1967): *The Normative electroencephalographic data reference library, Final Report*, NASA Contractor Report NAS9-1200, National Aeronautics and Space Administration, Washington, D.C., 1967.
  
- Kalsbeek, J.W.H.; Ettema, J.H. (1963): *Scored regularity of the heart rate pattern and the measurement of perceptual or mental load*. *Ergonomics*, 6, 306-307.

- Keele, S.W.; Neill, W.T. (1978): Mechanisms of attention. En E. C. Carterette y M. P. Freedman (Eds.), *Handbook of perception. vol IX (p. 3-46). New York: Academic Press.*
- Kitney, R.I. (1975): Entrainment of human RR interval by thermal stimuli, *Journal of Physiology-London*, vol: 252, p. 37-38, Issn: 0022-3751-
- Kitney, R.I. (1975): Analysis of nonlinear behavior of human thermal vasomotor control-system, *Journal of theoretical biology*, vol: 52, pages: 231-248, issn: 0022-5193
- Kitney, R.I. (1975): The assessment of physiological control systems utilising biological signals. (Review lecture), Proc. of the 15th International Conference on Biomed. Eng, p. 33-47.
- Kleiger, R.E.; Miller, J.P.; Bigger, J.T.; Moss, A.J. (1987): Decreased heart rate variability and its association with increased mortality after myocardial infarction, *AM. J. Cardiol.*, 59 p. 256-262.
- Kleiger, R.E.; Bosner, M.S.; Rottman, J.N.; Stein, P.K. (1993): Time-domain measurements of heart rate variability. *Journal of Ambulatory Monitoring*, vol. 6, p. 1-18.
- Lal, S.K.; Craig, A. (2000): Psychophysiological effects associated with drowsiness: driver fatigue and electroencephalography. *International Journal of Psychophysiology*, p. 35, 39.
- Lisper, H.O.; Laurell, H.; Stenig, G. (1971): *Effects of prolonged driving and the experience of the driver on heart-rate, respiration-rate and a subsidiary auditory reaction time.* Department of Psychology. Report 111, University of Uppsala.
- Lisper, H.O.; Laurel, H.; Van Loon, J. (1986): Relation between time to falling asleep behind the wheel on a closed track and changes in subsidiary reaction time during prolonged driving on a motorway, *Ergonomics*, 29:3, 445-453, 1986.
- Logan, G.D. (1978): Attention in character-classification tasks: Evidence for the automaticity of component stages. *Journal of experimental psychology: General*, 107 (1), 32-63.

- Logan, G.D. (1980): Attention and automaticity in Stroop and priming task: Theory and data. *Cognitive psychology*, 12, 523-553.
- Logan, G.D. (1981): Attention, automaticity and the ability to stop a speeded choice response. En J. Long y A. Baddeley (Eds.), *Attention and Performance IX* (p. 205-222). Hillsdale, N. J.: Erlbaum.
- Lopes, P.; Mitchell, R.H.; White, J.A. (1992): The relationships between respiratory sinus arrhythmia and coronary heart rate risk factors, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 14, p. 769-770.
- López de Ullibarri Galparsoro I; Pita Fernández, S. (1999): Medidas de concordancia: el índice de Kappa. Unidad de Epidemiología Clínica y Bioestadística. Complejo Hospitalario-Universitario Juan Canalejo. A Coruña (España). *Cad Aten Primaria* 1999; 6: 169-171.
- Luczak, H.; Laurig, W. (1973): An analysis of heart rate variability. *Ergonomics*, 16, 85-94.
- Luria, A.R. (1975): Atención y memoria, Barcelona: Fontanella.
- Mackie, R.R.; O'Hanlon, J. (1977): A study of the combined effects of extended driving and heat stress on driver arousal and performance. En R. R. Mackie (Ed.) *Vigilance*. Plenum Press, New York and London.
- Makeig, S.; Inlow, M. (1993): Lapses in alert coherence of fluctuation in performance and EEC spectrum, *Electroencephalogr. Clin. Neurophysiol*, vol. 86, p. 23-35, 1993.
- Malik, N., Kallestad, J.C., Gunderson, N.L., Austin, S.D., Neubauer, M.G., Ochs, V., Marquardt, H., Zarling, J.M., Shoyab, M., Wei, C.M., (1989): Molecular cloning, sequence analysis, and functional expression of a novel growth regulator, oncostatin M.
- Mallat, S.G. (1989): A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Transactions on Pattern Recognition and Machine Intelligence* 11 (7): 674-693.

- Mariotti, M., Gritti, I., Mancina, M. (1998): GABAergic and cholinergic basal forebrain and preoptic-anterior hypothalamic projections to the mediodorsal nucleus of the thalamus in the cat. *Neuroscience*. 1998 Jul;85 (1): 149-78.
- Marrero, H.; Torres, E. (1986): Panorama general de los estudios atencionales. *Revista de Psicología General y Aplicada*, 41 (2), 241-263.
- Mascord, D.J.; Heath, R.A. (1992): Behavioral and physiological indices of fatigue in a visual tracking task. *Journal of Safety Research*, 23, 19-25.
- Matousek, M. (1967): Automatic evaluation of EEG background activity. *Electroencephalography and clinical neurophysiology*.
- Matthews, G.; Desmond, P.A. (2002): Task-induced fatigue states and simulated driving performance, *The Quarterly Journal of Experimental Psychology*, vol. 55A, núm. 2, p. 659-686.
- Maycock, G. (1997): Sleepiness and driving: the experience of U.K. car drivers. *Accident Analysis and Prevention*, 29 (4), 453-462.
- McDonald, W.; Cameron, C. (1974): The use of behavioral methods to assess traffic hazards. *Australian Road Research*, 14, 1-27.
- Moñivas, A. (1993): Procesos, teorías y modelos de la atención. En A. Puente (coord.). *Psicología Básica. Introducción al estudio de la conducta humana*, p. 116-150. Madrid: Eudema.
- Morlet, T., Collet, L., Salle, B., Morgon, A. (1993): Functional maturation of cochlear active mechanisms and of the medial olivocochlear system in humans. *Acta Otolaryngol (Stockh)* 1993; 113: 271-7.
- Mulder, G.; Mulder-Hajonides, W. R. (1973): Mental load and the measurement of heart rate variability. *Ergonomics*, 16, 69-84.
- Murgoitio, J.; Ferros, M.; Goti, A.; Larburu, M.; Rodríguez, T. (2007): Contributions to the Ambient Intelligence: Networking Sensors & Actuators for a New Active Headrest, 11th

International Forum on Advanced Microsystems for Automotive Applications, Berlin, Alemania; 09 – 10 May. 2007.

- Murgoitio, J.; Fernández, J.I. (2008): Car driver monitoring by networking vital data, AMAA, International Forum on Advanced Microsystems for Automotive Applications, 2008.
- Murgoitio, J.; Arejita, B., Antúnez, M. (2009): Wellness monitoring in cars with CAN, CAN Newsletter. December 2009.
- Murgoitio, J.; Arejita, B.; Larburu, M.; Sánchez, J. (2010): A Bluetooth Sensor Network Base on the IEEE1451 Standard. In Proceedings of the International Conference on Wireless Information Network and Systems WINSYS 2010. Athens. 2010.
- Murgoitio, J.; Antúnez, M. (2011): Herramienta software para el estudio de la variabilidad de la frecuencia cardíaca (HRV): Aplicación a la medida del nivel de alerta en la conducción de vehículos, (vol. 86, núm. 1, p. 28), Dyna (Febrero/2011).
- Murgoitio, J.; Arejita, B.; Larburu, M.; Sánchez, J. (2011): A Bluetooth Sensor Network Based on the IEEE 1451 Standard. A Sensor Network Solution to Evaluate the Wellness of the Passenger and Improve the Safety in Cars, Computer Science Research and Technology. (vol. 2, p. 107) (2011). ISBN: 978-1-61122-724-6.
- Murgoitio, J.; Sánchez, J.; Cases, J.I.; Izaguirre, M. (2013): Iniciativas ITS entorno al concepto PRT en el parque tecnológico de Miramón (San Sebastián). XIII Congreso Español sobre Sistemas Inteligentes de Transporte. Proceedings, San Sebastián (Junio/2013).
- Murgoitio, J.; Aranbarri, A.; Gutiérrez, J.L.; Gil, M.J. (2013): Heart Rate Variability. Knowing More about HRV Analysis and Fatigue in Transport Studies. Cardiotecnix 2013: International Congress on Cardiovascular Technologies (Vilamoura). Proceedings, 107-114.
- Murgoitio, J. (2013): Local, National and European projects on Autonomous Vehicles, ITS Ireland Conference in association with RITS-Net. Proceedings, Dublin (Nov/2013).

- Naveh-Benjamin, M. y Jonides, J. (1986): On the automaticity of frequency coding: Effects of competing task load, encoding strategy and intention. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 12, 378-386
- Navon, D.; Gopher, D. (1979): On the economy of the human processing system. *Psychological Review*, 86 (3), 214-255.
- NHTSA: *National Highway Traffic Safety Agency* (2014): <http://www.nhtsa.gov/>
- Nilsson, T.; Nelson, T. M.; Carlson, D. (1997): Development of fatigue symptoms during simulated driving. *Accident Analysis and Prevention*, 29 (4), 479-488.
- Norman, D.A. (1968): Toward a theory of memory and attention. *Physiological Review*, 75, 522-536.
- Norman, D.A.; Bobrow, D.G. (1975): On data-limited and resource-limited processes. *Cognitive Psychology*, 7, 44-64.
- Norman, D.A.; Shallice, T. (1986): Attention to action. Willed and automatic control of behavior. En R.J. Davison, G. E. Schwartz y D. D Shapiro (Eds.), *Consciousness and self-regulation*, vol. 4. New York: Plenum Press.
- O'Hanlon, J.; Kelley, G. (1977): *Comparison of performance and physiological changes between drivers who perform well and poorly during prolonged vehicular operation*. En R. R. Mackie (Ed.) *Vigilance*. Plenum Press, New York and London.
- O'Neil, J.M. & Nadeau, R.A. (1999): *Men's gender role conflict, defense mechanism, and self-protective defensive strategies: Explaining men's violence against women from a gender role socialization perspective*. In M. Harway & J. O'Neil (Eds.) *What causes men's violence against women?* Thousand Oaks, CA: Sage Publications.
- Odemuyiwa, O.; Malik, M.; Farrell, T., Bashir, Y.; Poloniecki, J.; Camm, J. (1991): *Comparison of the predictive characteristics of heart rate variability index and left ventricular ejection fraction for all-cause mortality, arrhythmic events and sudden death after acute myocardial infarction*, *Am J Cardiol*. 1991; 68:434-439

- Pack, A.L.; Pack, A.M.; Rodgman, E.; Cucchiara, A.; Dinges, D.F.; Schwab, C.W. (1995): *Characteristics of crashes attributed to the driver having fallen asleep*. *Accident Analysis and Prevention*, 27 (6), 769-775.
- Pagani, M.; Lombardi, F.; Guzzetti, S.; Rimoldi, O.; Furlan, R.; Pizzinelli, P.; Sandrone, G.; Malfatto, G.; Dell'Orto, S.; Piccaluga, E.; Turiel, M.; Baselli, G.; Cerutti, S.; Malliani, A. (1986): *Power Spectral Analysis of Heart Rate and Arterial Pressure Variabilities as a Marker of Sympatho-Vagal Interaction in Man and Conscious Dog*. *Circulation Research*, vol. 59, p. 178-193.
- Park, G.; Vasey, M.W.; Van Bavel, J.J.; Thayer, J.F. (2013): *Cardiac vagal tone is correlated with selective attention to neutral distractors under load*. *Psychophysiology*. 50, 398-406.0048-5772.
- Peltokangas, M.; Verho, J.; Vehkaoja, A. (2012): *Night-Time EKG and HRV Monitoring With Bed Sheet Integrated Textile Electrodes*. *IEEE transactions on information technology in biomedicine*. 16, 935-942. 1089-7771.
- Phaf, R.H.; Van der Heijden, A.H.C.; Hudson, P.T. (1990): *SLAM: A connectionist model for attention in visual selection tasks*. *Cognitive Psychology*, 22, 273-341.
- Pomeranz, B.; Macaulay, R.J.; Caudill, M.A.; Kutz, L.; Adam, D.; Gordon, D.; Kilborn, K.M.; Barger, A.C.; Shannon, D.C.; Cohen, R.J.; Benson, H. (1985): *Assessment of autonomic function in humans by heart rate spectral analysis*. *American Journal of Physiology*, 248, 15 1-153.
- Posner, M.I.; Boies, S. (1971): *Components of attention*. *Psychological Review*, 78 (5), 391-408.
- Posner, M.I. (1975): *Psychobiology of attention*. En M.S. Gazzaniga y C. Blakemore (Eds.), *Handbook of Psychobiology* (p. 441-480). New York: Academic Press.
- Posner, M.I.; Snyder, C.R. (1975): *Attention and cognitive control*. En R. Solso (Ed.), *Information processing and cognition. The Loyola symposium* (p.55-85). Hillsdale, N. J.: Erlbaum.

- Posner, M.I.; Snyder, C.R. (1975): *Facilitation and inhibition in the processing of signals*. En P. M. A. Rabbit y S. Dornic (Eds.), *Attention and performance V* (p.669-682). New York: Academic Press.
- Posner, M.I. (1978): *Chronometric explorations of mind*. Hillsdale: Erlbaum.
- Posner, M.I.; Friedrich, F.J. (1986): *Attention and control of cognition*. In S.L. Friedman, K. A. Klivington, & R.W. Peterson (Eds.), *Brain, Cognition and Education*. Orlando, FL: Academic Press.
- Posner, M.I. (1987): *Selective attention and cognitive control*. Trends in Neuroscience, January, 13-17.
- Qiang, J.; Xiaojie, Y. (2002): *Real-Time Eye, Gaze, and Face Pose Tracking for Monitoring Driver Vigilance*. Real-Time Imaging, 8, p. 357-377.
- RACC (2014): Real Automóvil Club de Cataluña: Claroscuros en el balance de accidentalidad vial (<http://tribunas.racc.es/es/analisis-accidentalidad-2014>). Tribunales de a movilidad.
- RACE (2011): Real Automóvil Club de España. Informe 2011: La fatiga en la conducción.
- RACE (2014): [www.fatigayconduccion.com](http://www.fatigayconduccion.com)
- Ranney, T.A.; Simmons, L.A.; Masalonis, A.J. (1999). Prolonged exposure to glare and driving time: effects on performance in a driving simulator. *Accid. Anal. Prev.* 31, 601-610.
- Reason, J. (1979): Actions not as planned: The price of automatization. En G. Underwood y R Stevens (Eds.), *Aspects of Consciousness*. Vol. 1. Londres: Academic Press.
- Reason, J. (1984): Lapses of attention in everyday life. En R. Parasuraman y D.R. Davies (Eds.), *Varieties of Attention*. New York: Academic Press.
- Reina, N.A.; Martínez, N.; Seepold, R.; Murgoitio, J. (2012): *A UPnP Service to Control and Manage IEEE 1451 Transducers in Control Networks*. IEEE TRANSACTIONS ON

INSTRUMENTATION AND MEASUREMENT, vol. 61 Issue: 3, p. 791-800. DOI: 10.1109/TIM.2011.2170501. Mar 2012.

- Reisman, S.; Tapp, W.; Findley, T.; Zhang, P. (1991): *Fatigue in multiple sclerosis using heart rate variability*, Annual International Conference of the IEEE Engineering in Medicine and Biology Society, vol. 13, n° 2, p. 664-665, 1991.
- Reitman, J.S. (1974): Without surreptitious rehearsal, information in short-term memory decays. *Journal of Verbal Learning and Verbal Behaviour*, 13, 365-377.
- Restivo, M.; El-Sherif, N.; Kelen, G.J.; Henkin, R.; Craelius, W.; Gough, W.B. (1985): Correlation of late potentials on the body surface and ventricular activation maps of reentrant circuits in the post-infarction dog heart, *Circulation*, vol. 72, supl. m, p. 11.
- Retschaffan, A.; Kales, A. (1968): A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects. National Institutes of Health, Publication no. 204, Neurological Information Network (1968).
- Reyes del Paso, G.A; Langewitz, W.; Mulder, L.JM; van Roon, A.; Duschek, S. (2013): *The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: a review with emphasis on a reanalysis of previous studies..* *Psychophysiology*. 50, 477-87. 1540-5958.
- Riemersma J.B.J., Sanders A.R. and Hildervanek C. (1977): *Performance decrement during prolonged night driving*. In: Mackie R. (ed.). *Vigilance: Theory, operational performance and physiological correlates*. FlenumPress, New York, 41-58.
- Rosekind, M.R. (1999): *Fatigue in Transportation: Physiological, Performance, and Safety Issues*. Alertness Solutions.
- Rosselló i Mir, J. (1996): *Psicología de la Atención. Introducción al estudio del mecanismo atencional*. Madrid: Piramide.
- Ruiz-Vargas, J.M.; Botella, J. (1981): Limitaciones de procesamiento y selectividad atencional. *Estudios de psicología*, 7, 30-41.

- Ruiz-Vargas, J.M.; Botella, J. (1987): Atención. En J.M. Ruiz-Vargas (Dir), *Esquizofrenia: Un enfoque cognitivo*. Madrid: Alianza.
- Ruiz-Vargas, J.M. (1993): Atención y control: Modelos y problemas para una integración teórica. *Revista de Psicología General y Aplicada*, 46 (2), 125-137.
- Rumelhart, D.E. y McClelland, J.L. (1986): Parallel distributed processing. Explorations in the Microstructure of Cognition: Foundations. Volume 1: Foundations.
- Santamaría, J.; Chiapa, K. (1987): The EEG of drowsiness in normal adults. *J. Clin. Neurophysiol.*, vol. 4, núm. 4, p. 327-382, 1987.
- Saroj, K.L.; Ashley C. (2001): *A critical review of psychophysiology of driver fatigue*. *Biological Psychology* 55, p. 173-194, 2001.
- Saus, E.R.; Johnsen, B.H.; Eid, J.; Thayer, J.F. (2012): *Who benefits from simulator training: Personality and heart rate variability in relation to situation awareness during navigation training*. *Computers in human behavior*. 28, 1262-1268. 0747-5632.
- Sayers, B. (1973): Analysis of heart-rate variability. *Ergonomics*, 16, 17-32.
- Schneider, W.; Shiffrin, R.M. (1977): Controlled and automatic human information processing I: Detection, search and attention. *Psychological Review*, 84, 1-66.
- Schneider, W.; Shiffrin, R.M. (1985): Categorization (restructuring) and automatization: Two separable factors. *Psychological Review*, 92, 424-428.
- Schnell, I.; Potchter, O.; Epstein, Y.; Yaakov, Y.; Hermesh, H.; Brenner, S.; Tirosh, E. (2013): *The effects of exposure to environmental factors on Heart Rate Variability: An ecological perspective*. *Environmental pollution*. 183, 7-13. 0269-7491; 1873-6424.
- Shallice, T. (1988): *From neuropsychology to mental structure*. Cambridge CUP.

- Shallice, T. (1988): Specialization within the semantic system. *Cognitive neuropsych* 5 (1), 133-142.
- Shallice, T.; Burgess, P.W.; Schon, F.; Baxter, D.M. (1989): The origins of utilization behavior. *Brain* 112, 1587-1598.
- Shaun, T. (2010): Performance based specification for heavy vehicle fatigue and speed monitoring. *17<sup>th</sup> ITS World Congress, Busan-KOREA*.
- Shields, K.N.; Cavallari, J.M.; Hunt, M.J.O.; Lazo, M.; Molina, M.; Molina, L.; Holguin, F. (2013): *Traffic-related air pollution exposures and changes in heart rate variability in Mexico City: A panel study*. *Environmental health*. 12, 7. 1476-06X.
- Shiffrin, R.M.; Schneider, W. (1977): Controlled and automatic human information processing II: Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84, 127-190.
- Shin, S.; Tapp, W.N.; Reisman, S.S.; Tapp, W.N.; Natelson, B.H. (1989): Assessment of autonomic regulation of heart rate variability by the method of complex demodulation, *IEEE Transactions on Biomedical Engineering*, vol. BME-36, p. 274-283.
- Shin, K.S.; Minamitani, H.; Onishi, S.; Yamazaki, H.; Lee, M.H. (1993): The power spectral analysis of heart rate variability during exercise, *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, vol. 15, p. 329 - 332.
- Simon, H.A. (1986): The role of attention in cognition. En S. H. Friedman y cols. (Eds), *The brain, cognition and education*. New York: Academic press.
- Smith, S.E.; Smith, S.A. (1981): Heart rate variability in healthy subjects measured with a bedside computer-based technique", *Clinical Science*, vol. 61, p. 379-383.
- Somers, V.; Dyken, M.E.; Mark, A.L.; Abboud, F.M. (1993): Sympathetic-nerve activity during sleep in normal subjects, *N Engl J Med*, vol. 328, n° 5, p. 303-307.

- Stein, E.W. (1995): Organization memory: Review of concepts and recommendations for management. *International Journal of Information Management*. Volume 15, Issue 1, February 1995, P. 17–32.
- Svebak, S.; Dalen, K.; Storfjell, O. (1981): The psychological significance of task-induced tonic changes in somatic and autonomic activity. *Psychophysiology*, 18, 403-409.
- Takahiro, Y. (2010): Estimating probability of automobile accident from driver's reaction time under different arousal states. *17<sup>th</sup> ITS World Congress, Busan-KOREA*.
- Tarvainen, M.P.; Niskanen, J.P.; Lipponen, J.A.; Ranta-aho, P.O.; Karjalainen, P.A.; (2014): Kubios HRV - *Heart rate variability analysis software*. *Computer methods and programs in biomedicine*. 113, 210-220. 0169-2607; 1872-7565.
- Taylor, E. (1995): Dysfunctions of attention. En C. Cicchetti y D. J. Cohen (Eds.), *Developmental Psychology*, vol. 2: Risk, disorder and adaptation (p. 243-273).
- Ten Harkel, A.D.; Van Lieshout, J.J.; Van Lieshout, E.J.; Wieling, W. (1990): Assessment of cardiovascular reflexes: influence of posture and period of preceding rest. *J Appl Physiol* (1985). 1990 Jan; 68(1):147-53.
- Terzano, M.G.; Parrino, L.; Boselli, M.; Spaggiari, M.C.; Di Giovanni, G.; Smerieri, A. (1997): Sensitivity of cyclic alternating pattern to prolonged pharmacotherapy: a 5-week study evaluating zolpidem in insomniac patients. *Clin Neuropharmacol*. 1997 Oct; 20 (5): 447-54.
- Treisman, A.M. (1960): Contextual cues in selective listening. *Quarterly Journal of experimental Psychology*, 12, 242-248).
- Tudela, P. (1985): *Psicología Experimental*. Vol. 1. Madrid: UNED.
- Tudela, P. (1992): Atención. En J. L. Fernández-Trespalacios y P. Tudela (Coords.), *Atención y percepción* (p. 119-163). En J. Mayor y J. L. Pinillos (Eds), *Tratado de Psicología General*, vol. 3. Madrid: Alhambra.

- Tzyy-Ping, J.; Scott, M.; Magnus, S. (1997): Estimating alertness from the EEG power spectrum, *IEEE Transactions on Biomedical Engineering*, vol. 45, p. 60-69, 1997.
- Ueno, H.; Kaneda, M.; Tsukino, M. (1994): Development of drowsiness detection system, in *Proc. Vehicle Navigation and Information Systems Conference*, 1994. p. 15-20.
- Unrug-Neervoort, A.; van Luijtelaar, G.; Coenen, A. (1992): Cognition and vigilance: differential effects of diazepam and buspirone on memory and psychomotor performance. *Neuropsychobiology*. 1992; 26 (3): 146-50.
- Unser, M.; Aldroubi, A.; Eden, M. (1992): Cardinal spline filters: Stability and convergence to the ideal sinc interpolator. *Signal Processing*, Volume 28, Issue 2, August 1992, p. 127-138
- Unser, M.; Aldroubi, A.; Eden, M. (1992): On the asymptotic convergence of B-spline wavelets to Gabor functions, *IEEE IT*, vol. 38, n° 2, p. 864-871, 1992.
- Urquiza, A.; Murgoitio, J.; Larburu, M. (2011): A Context evaluator model for a multimodal user interface in driving activity. *ICAART 2011 - 3rd International Conference on Agents and Artificial Intelligence*, Rome, Italy; Jan. 2011.
- Verwey, W.B.; Zaidel, D.M. (2000): Predicting drowsiness accidents from personal attributes, eye blinks and ongoing driving behaviour, *Personality and Individual Differences*, vol. 28, p 123-142.
- Vicente, J.; Laguna, P.; Bartra, A.; Bailón, R. (2011): Análisis HRV en la detección de somnolencias en conductores. *CASEIB 2011-XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica*. p 193.
- Voss, A.; Heitmann, A.; Schroeder, R.; Peters, A.; Perz, S. (2012): *Short-term heart rate variability-age dependence in healthy subjects*. *Physiological measurement*. 33, 1289-1311. 0967-3334.

- Vuckovie, A.; Popovie, D. (2002): Artificial neural network for detecting drowsiness from EEG recording, *IEEE Transactions on Monitoring Driver Vigilance. Real-Time Imaging*, 8, p. 357-377.
- Watkins, J.K.; Phillips, V.J. (1973): Speech scrambling by the matrixing of amplitude samples. *Radio and Electronic Engineer*, vol: 43, Issue: 8, p. 459 - 470
- Webster, J.G. (1992): *Medical instrumentation. Application and design*. Boston: Houghton Mifflin Company, 1992.
- Wickens, C.D. (1980): The structure of attentional resources. En R. Nickerson y R. Pew (Eds.), *Attention and performances VIII* (p. 239-257). Hillsdale, N. J.: Erlbaum.
- Wilkinson, R.T. (1992): The measurement of sleepiness, In R. J. Broughton and R. D. Ogilvie (Eds.), *Sleep, arousal, and performance: a tribute to Bob Wilkinson*. Boston: Birkhäuser. (p. 254-265), 1992.
- Williams, G.W. (1963): Highway hypnosis: an hypothesis. *The International Journal of Clinical and Experimental Hypnosis*, XI (3), 143-151.
- Williamson, A.M.; Feyer, A.M.; Coumarelos, C.; Jenkins, A. (1992): Strategies to combat fatigue in the long distance road transport industry. Stage 1: The industry perspective, *Report No CR 108*, Canberra, Australia: Federal Office of Road Safety.
- Williamson, A.M.; Feyer, A.M.; Friswell, R. (1996): The impact of work practices on fatigue in long distance truck drivers, *Accident Analysis and Prevention*, vol. 28, núm. 6, p. 709-719.
- Williamson, A.M.; Feyer, A.M.; Mattick, R.P.; Friswell, R.; Finlay-Brown, S. (2001): Developing measures of fatigue using an alcohol comparison to validate the effects of fatigue on performance. *Accid Anal Prev*. 2001 May; 33 (3): 313-26.
- Witten, I.H.; Frank, E. (2000): *Data Mining: Practical machine learning tools with Java implementations*. Morgan Kaufmann, San Francisco, 2000.

- Wylie, C.D.; Shultz, T.; Miller, J.C.; Mitler, M.M.; Mackie, R.R. (1996): Commercial Motor Vehicle Driver Fatigue and Alertness Study, Technical summary. Montreal: Transport Development Centre, 1996.
- Xhyheri, B.; Manfrini, O.; Mazzolini, M.; Pizzi, C.; Bugiardini, R. (2012): *Heart Rate Variability Today*. Progress in cardiovascular diseases. 55, 321-331. 0033-0620.
- Yoshizaki, T.; Kawano, Y.; Tada, Y.; Hida, A.; Midorikawa, T.; Hasegawa, K.; Mitani, T.; Komatsu, T.; Togo, F. (2013): *Diurnal 24-Hour Rhythm in Ambulatory Heart Rate Variability during the Day Shift in Rotating Shift Workers*. Journal of biological rhythms. 28, 227-236. 0748-7304.
- Zhao, C.L.; Zhao, M.; Liu, J.P.; Zheng, C.X.; (2012): *Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator*. Accident analysis and prevention. 45, 83-90. 0001-4575.

## 8. ANEXOS

### 8.1. Anexo 1: Participantes y Autorizaciones

#### 8.1.1. Participantes

Los participantes se seleccionaron según dos criterios básicos: Por un lado que fueran sanos y no tuvieran ninguna enfermedad relacionada con trastornos del corazón, y por otro lado fueran de grupos de edad diferente. Resaltar que fue especialmente complicado seleccionar participantes con edad superior a los 55 años, sobre todo teniendo en cuenta el primer criterio y que fueran conductores en activo.

Según esto, se obtuvo un grupo de 38 personas diferentes, todas ellas de sexo “Varón”, las cuales fueron informadas del objetivo del experimento, autorizando la utilización de la información recogida según autorizaciones firmadas por los mismos. La Tabla 27 muestra su identificador (Id) dentro de este estudio, así como información resumida de los mismos:

<b>Id</b>	<b>Nombre:</b>	<b>Apellido1</b>	<b>Apellido2</b>	<b>Fecha Nacimiento</b>
1	Jesus	Murgoitio	Larrauri	07/1963
2	Manuel	Marín	Andrés	01/1963
3	Martin	Etxebarria	Ezpeleta	04/1962
4	Txomin	Rodriguez	Villa	06/1965
5	Juan	Nevado	Durán	01/1959
6	Eneko	Nevado	Murgoitio	09/1988
7	Alberto Demetrio	DeDiego	Ibáñez	04/1975
8	Jose Ignacio	Beitia	Sertucha	11/1963
9	Alberto	DeDiego	Campos	08/1946
10	Jose Ignacio	Rivada	González	04/1959
11	Juan Antonio	Mentxaka	Alberdi	09/1957
12	Javier	Markaida	Markaida	08/1964
13	Jon Koldo	Barcenilla	Sagardui	11/1964
14	Joseba	Arruza	Elguezabal	11/1963
15	Juan Antonio	Diaz	Nevado	10/1968
16	Manu	Jauregui	Larringan	11/1949
17	Jose Luis	Zabaleta	LasHeras	08/1960
18	Javier	Olazar	Gangoiti	07/1964
19	Pedro	Artetxe	Atucha	02/1939
20	Jose Maria	Arieta-araunabeña	Bustinza	02/1953
21	Josu	Barcenilla	Sagardui	02/1963
22	Aitor	Arin	Bilbao	10/1970
23	Josu	Jaka	Aspe	08/1964
24	Alfredo	Madina	Jugo	08/1956

<b>Id</b>	<b>Nombre:</b>	<b>Apellido1</b>	<b>Apellido2</b>	<b>Fecha Nacimiento</b>
25	Asier	Arambarri	Beldarrain	07/1987
26	Ekain	Aguirre	García	03/1987
27	Xabier	Razkin	Euba	03/1987
28	Jon Iker	Arambarri	Beldarrain	05/1981
29	Alain	Bilbao	Abaunza	12/1986
30	Juan Jose	Amarika	Sagarminaga	04/1951
31	Victor	Sagardui	Larrondo	02/1944
32	Pedro	Intxauralde	Etxebarria	05/1939
33	Jose Antonio	Larringan	Uriarte	01/1943
34	Benito	Nevado	Durán	02/1952
35	Francisco Javier	Saez	Martínez	07/1954
36	Gaspar	Nevado	Durán	05/1956
37	Jon Andoni	Murgoitio	Larrauri	03/1977
38	Santos	Goti	Eguiguren	07/1937

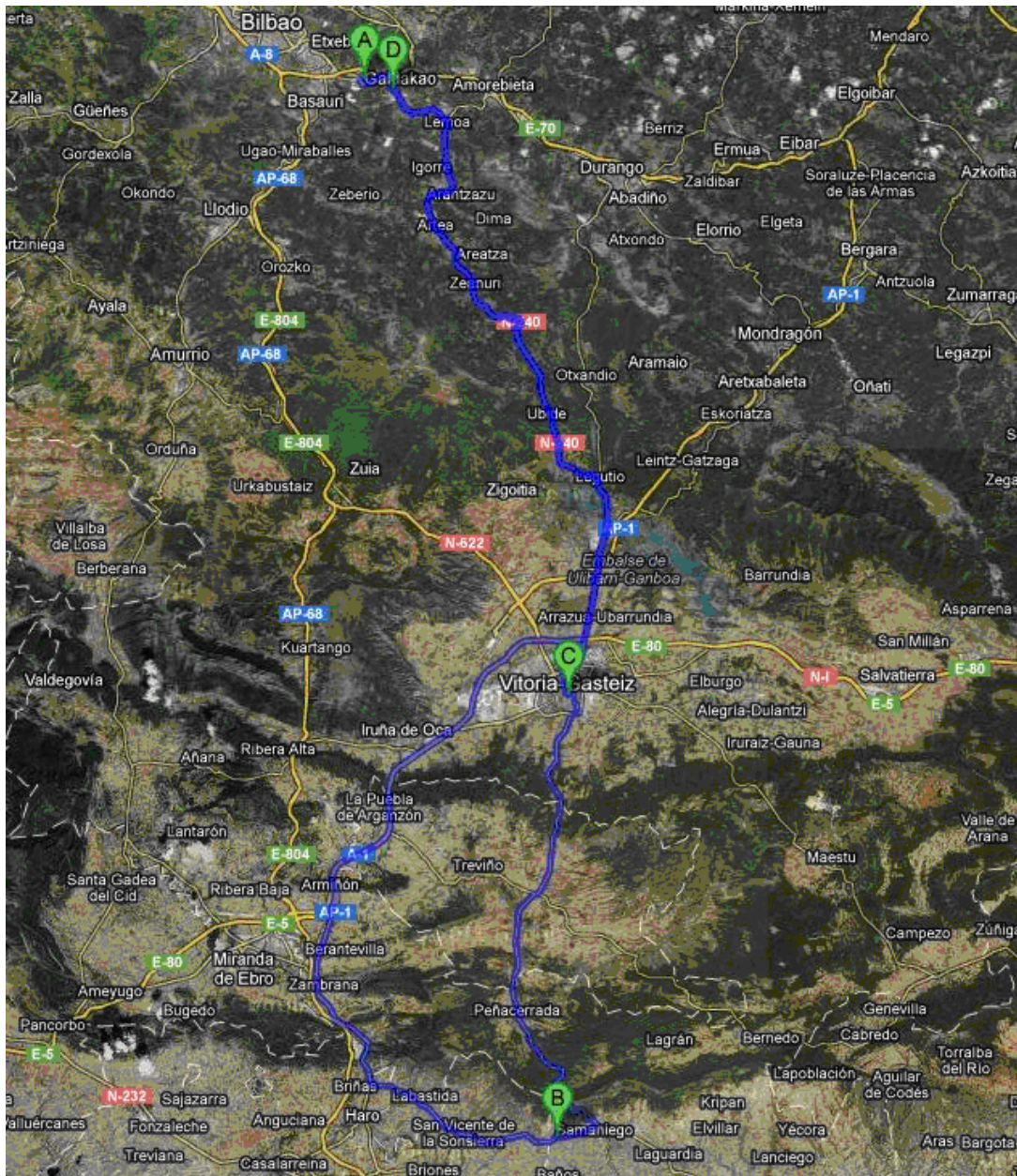
**Tabla 27: Identificadores de los participantes**

### **8.1.2. Autorizaciones**

El objetivo de este estudio es conocer el comportamiento del registro ECG durante la conducción en trayectos que sobrepasen las dos horas y media. Este estudio forma parte de la tesis doctoral de investigación "NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES", que cuenta con la participación de la Universidad de Deusto y la corporación tecnológica TECNALIA.

Como parte de esta investigación, se han tomado previamente algunos datos personales, y principalmente se registrará la actividad cardiaca, además de datos proporcionados por un grupo de sujetos voluntarios mientras conducen un vehículo. Para ello se les colocarán los electrodos necesarios. Los sujetos irán siempre acompañados por un miembro del equipo investigador, que se encargará de controlar el buen desarrollo del estudio y de registrar algunos datos durante la conducción.

El vehículo que conducirán los sujetos es en su mayoría el suyo propio, que cuenta con el seguro obligatorio. El recorrido se iniciará y concluirá en la zona denominada "Plazakoetxe" del municipio de Galdakao, según el recorrido de la Figura 86 y sin realizar paradas de descanso.



**Figura 86: Recorrido del experimento**

Los participantes declaran acceder a las pruebas de forma voluntaria y gratuita. Los datos particulares que se obtengan para cada sujeto son absolutamente confidenciales y sólo tendrán acceso a ellos los investigadores involucrados que están desarrollando este estudio. De hecho, nuestro interés no reside de ninguna forma en casos individuales, sino en obtener datos promedio que representen lo que le ocurre a la mayoría de los conductores, por lo que los datos recogidos para cada sujeto se añadirán a los de los otros sujetos para analizarlos estadísticamente, de forma conjunta. No obstante, si algún participante desea tener información sobre sus datos, ésta le será proporcionada.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 1:**

**Jesus Murgoitio Larrauri**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 2:**

**Manuel Marín Andrés**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 3:**

**Martin Etxebarria Ezpeleta**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 4:**

**Txomin Rodriguez Villa**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 5:**

**Juan Nevado Durán**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 6:**

**Eneko Nevado Murgoitio**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 7:**

**Alberto Demetrio DeDiego Ibáñez**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 8:**

**Jose Ignacio Beitia Sertucha**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 9:**

**Alberto DeDiego Campos**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 10:**

**Jose Ignacio Rivada González**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 11:**

**Juan Antonio Mentxaka Alberdi**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 12:**

**Javier Markaida Markaida**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 13:**

**Jon Koldo Barcenilla Sagardui**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 14:**

**Joseba Arruza Elguezabal**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 15:**

**Juan Antonio Diaz Nevado**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 16:**

**Manu Jauregui Larringan**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 17:**

**Jose Luis Zabaleta LasHeras**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 18:**

**Javier Olazar Gangoiti**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 19:**

**Pedro Artetxe Atucha**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 20:**

**Jose Maria Arieta-araunabeña Bustinza**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 21:**

**Josu Barcenilla Sagardui**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 22:**

**Aitor Arin Bilbao**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 23:**

**Josu Jaka Aspe**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 24:**

**Alfredo Madina Jugo**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 25:**

**Asier Arambarri Beldarrain**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 26:**

**Ekain Aguirre García**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 27:**

**Xabier Razkin Euba**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 28:**

**Jon Iker Arambarri Beldarrain**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 29:**

**Alain Bilbao Abaunza**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 30:**

**Juan Jose Amarika Sagarminaga**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 31:**

**Victor Sagardui Larrondo**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 32:**

**Pedro Intxauralde Etxebarria**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 33:**

**Jose Antonio Larringan Uriarte**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 34:**

**Benito Nevado Durán**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 35:**

**Francisco Javier Saez Martínez**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 36:**

**Gaspar Nevado Durán**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 37:**

**Jon Andoni Murgoitio Larrauri**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

**AUTORIZACIÓN DE PARTICIPANTE CON ID = 38:**

**Santos Goti Eguiguren**

mayor de edad y en pleno uso de sus facultades, declara que participó voluntariamente y según lo indicado anteriormente como sujeto experimental en la investigación "**NUEVO MÉTODO BASADO EN EL HRV PARA LA EVALUACIÓN DE HMIs Y SISTEMAS ITS PARA TRANSPORTE POR CARRETERA INTEGRANDO FACTORES PERSONALES, TEMPORALES Y AMBIENTALES**" realizada por la Universidad de Deusto y la corporación tecnológica TECNALIA, que era conocedor de todas las pruebas en las que participó y de las condiciones en las que se llevaron a cabo tales pruebas, según se describe arriba, y que asumía toda la responsabilidad de su participación.

Fdo.: .....

Bilbao, a ...**30**..... de .....**Junio**.....de **2014**.

## 8.2. Anexo 2: Cuadro resumen de resultados.

	DT		DF		DC	
	CL1	CL2	CL1	CL2	CL1	CL2
<b>Caso de estudio 1</b>						
Nº Variables:	2	2	2	2	2	2
Variables:	SD1 SD2	SD1 SD2	LF_Power HF_Power	LF_Power HF_Power	SD1 LF_Power	SD1 LF_Power
Nº Clases (n):	2	2	2	2	2	2
Instancias (n):	2.689	2.689	2.689	2.689	2.689	2.689
ROC área (%):	85,0	<b>93,2</b>	91,7	<b>94,4</b>	86,7	93,9
Correctly Classified (%):	84,0	86,3	84,7	87,1	85,5	87,5
Kappa statistic:	0,6818	0,7268	0,6947	0,7420	0,7093	0,7503
Mean absolute error:	0,1288	0,1176	0,1061	0,0987	0,1196	0,1026
<b>Caso de estudio 2</b>						
Nº Variables:	3	3	3	2	2	2
Variables:	Media SD1_Nor SD1/SD2	Media SD1 <b>SD1/SD2</b>	VLF_PeakY HF_Power HF_XC	<b>HF Power</b> HF_XC	SD1/SD2 HF_Power	SD1/SD2 HF_Power
Nº Clases (n):	2	2	2	2	2	2
Instancias (n):	2.705	2.705	2.705	2.705	2.705	2.705
ROC área (%):	81,3	<b>97,1</b>	81,3	<b>85,5</b>	78,5	82,6
Correctly Classified (%):	74,7	90,8	76,2	77,7	69,9	73,8
Kappa statistic:	0,4962	0,8151	0,5239	0,5544	0,3974	0,4780
Mean absolute error:	0,2190	0,1207	0,2191	0,2284	0,2387	0,2425

Tabla 28: Tabla resumen de los rendimientos de los clasificadores (Casos de estudio 1 y 2)

	DT		DF		DC	
	CL1	CL2	CL1	CL2	CL1	CL2
<b>Caso de estudio 3</b>						
Nº Variables:	2	2	3	3	2	3
Variables:	Media_Norm SD1_Norm	Media_Norm SD1_Norm	VLF_Power_No LF_Power_No HF_Power_No	VLF_Power_No LF_Power_No HF_Power_No	SD1_Norm HF_Power_No	Media_Norm SD1_Norm HF_Power_No
Nº Clases (n):	2	2	2	2	2	2
Instancias (n):	1.854	1.854	1.854	1.854	1.854	1.854
ROC área (%):	74,4	<b>79,0</b>	67,0	<b>80,4</b>	68,8	85,5
Correctly Classified (%):	68,0	71,8	61,4	72,2	64,1	77,9
Kappa statistic:	0,3619	0,4398	0,2315	0,4445	0,2845	0,5592
Mean absolute error:	0,1288	0,1314	0,1461	0,1272	0,1429	0,1147
<b>Caso de estudio 4</b>						
Nº Variables:	3	3	3	3	2	4
Variables:	Media_Norm SD1_Norm SD2_Norm	Media_Norm SD1_Norm SD2_Norm	VLF_Power_No LF_Power_No HF_Power_No	VLF_Power_N LF_Power_No HF_Power_No	SD1_Norm LF_Power_No	Media_Norm SD1_Norm LF_Power_No HF_Power_No
Nº Clases (n):	6	6	6	6	6	6
Instancias (n):	5.887	5.887	5.887	5.887	5.887	5.887
ROC área (%):	77,7	<b>87,5</b>	66,9	<b>78,2</b>	78,6	91,4
Correctly Classified (%):	40,3	55,9	29,7	44,9	41,8	66,8
Kappa statistic:	0,2753	0,4632	0,1429	0,3303	0,2908	0,5981
Mean absolute error:	0,2325	0,2102	0,2587	0,2423	0,2267	0,1804

**Tabla 29: Tabla resumen de los rendimientos de los clasificadores (Casos de estudio 3 y 4)**

## 8.3. Anexo 3: Equipamiento – Especificaciones técnicas.

### 8.3.1. Netbook: DELL – Inspiron 1110: Especificaciones

<b>Procesador</b>	
Tipo	Intel® Celeron® Intel Pentium®
Frecuencia de bus externa	800 MHz
Caché L2	1 MB o 2 MB
<b>Memoria</b>	
Conectores	Un zócalo SO-DIMM accesible internamente
Tipo de memoria	DDR2 a 800 MHz
Posibilidad de configuración de memoria	1 GB, 2 GB y 4 GB
<b>Información del equipo</b>	
Conjunto de chips del sistema	Intel GS45
Amplitud del bus de datos	64 bits
Amplitud del bus de DRAM	Buses de 64 bits de canal único
Amplitud del bus de direcciones del procesador	32 bits
EPROM rápida	1 MB
<b>Comunicaciones</b>	
Adaptador de red	tarjeta de sistema LAN Ethernet 10/100
Inalámbrico	Minitarjeta WLAN interna con WIFI Minitarjetas Bluetooth® y WWAN opcionales
<b>Vídeo</b>	
<i>UMA (asignación de memoria unificada):</i>	
Tipo de vídeo	Integrado en la placa base
Controladora de vídeo	UMA Intel GS45 integrada
Memoria de vídeo	Dinámica basada en la capacidad de memoria del sistema
Interfaz LCD	Señal diferencial de bajo voltaje (LVDS)
<b>Audio</b>	
Tipo de audio	Audio de alta definición de 4 canales
Controladora de audio	Realtek ALC272
Interfaz interna	Audio Intel de alta definición
Altavoz	Dos altavoces de 1 vatio
<b>Puertos y conectores</b>	
Audio	Un conector para micrófono, un conector para altavoz estéreo/auriculares mono
Minitarjeta	Una ranura para minitarjeta de tamaño completo y una ranura para minitarjeta de medio tamaño, admite interfaz

	doble	
Adaptador de red	Un conector RJ45	
USB	Dos conectores de 4 patas compatibles con USB 2.0 Un conector de 4 patas compatibles con USB 2.0	
HDMI	Un conector HDMI	
<b>Pantalla</b>		
Tipo	HD+ WLED de 11,6 pulgadas con TrueLife	
Dimensiones:		
Altura	144 mm (5,7 pulg.)	
Anchura	256,12 mm (10,1 pulg.)	
Diagonal	293,83 mm (11,6 pulg.)	
Resolución máxima	1366 x 768 con una escala de colores de 262 K	
Frecuencia de actualización	60 Hz	
Ángulo de funcionamiento	De 0° (cerrada) a 135°	
Luminancia	180 min. 200 cd/m <sup>2</sup> (normal)	
Ángulos de percepción:		
Horizontal	±40° mín	
Vertical	+15°/-30° mín	
Separación entre píxeles	0,1875 x 0,1875 mm	
Controles	El brillo puede controlarse mediante las teclas de acceso rápido del teclado (para obtener más información, consulte la Guía tecnológica de Dell en su unidad de disco duro o en <a href="http://support.dell.com">support.dell.com</a> )	
<b>Teclado (retroiluminación)</b>		
Número de teclas	86 (EE. UU.); 87 (RU/Brasil); 90 (Japón)	
Diseño	QWERTY	
<b>Cámara</b>		
Píxel	1,3 megapíxeles	
Resolución de vídeo	640 x 480 a 30 marcos por segundo (máximo)	
Ángulo de visión en diagonal	66°	
<b>Superficie táctil</b>		
Resolución de posición X/Y (modo de tabla de gráficos)	240 cpp	
Tamaño:		
Anchura	Área activa del sensor 84,45 mm (3,32 pulg.)	
Altura	41,6 mm (1,64 pulg.) rectangular	
<b>Batería</b>		
Tipo	Iones de litio de 3 celdas	Iones de litio de 6 celdas
Atributos	3 celdas	Seis celdas
Profundidad	23,1 mm (0,91 pulg.)	45,62 mm (1,80 pulg.)
Altura	39,04 mm (1,54 pulg.)	45,7 mm (1,80 pulg.)
Anchura	204,14 mm (8,04 pulg.)	214,0 mm (8,42 pulg.)
Peso	0,165 kg (0,36 lb)	0,350 kg (0,77 lb)
Voltaje	11,1 VCC	11,1 VCC
Tiempo de carga cuando el equipo está apagado (aproximado)	3 horas (de 3 celdas)	4 horas (de 6 celdas)
Tiempo de funcionamiento	El tiempo de funcionamiento de la batería depende de las condiciones de funcionamiento y puede disminuir de manera significativa en determinadas condiciones de consumo intensivo.  Para obtener más información, consulte la Guía tecnológica de Dell en la unidad de disco duro o en <a href="http://support.dell.com">support.dell.com</a> .	

Duración (aproximada)	300 ciclos de carga/descarga
<b>Intervalo de temperatura:</b>	
En funcionamiento	De 0 a 35 °C (de 32 a 95 °F)
En almacenamiento	De -40 a 65 °C (de -40 a 149 °F)
Batería de tipo botón	CR-2032
<b>Adaptador de CA</b>	
Voltaje de entrada	100-240 VCA
Corriente de entrada (máxima)	1,5 A
Frecuencia de entrada	50-60 HZ
Potencia de salida	65 W
Intensidad de salida	3,34 A (continua)
Voltaje nominal de salida	19,5 ± 1 V CC
<b>Dimensiones:</b>	
Altura	16 mm (0,63 pulg.)
Anchura	66 mm (2,6 pulg.)
Longitud	127 mm (5 pulg.)
Peso (sin cables)	0,25 kg (0,55 lb)
<b>Intervalos de temperatura:</b>	
En funcionamiento	De 0 a 35 °C (de 32 a 95 °F)
En almacenamiento	De -30 °C a 65 °C (de -22 °F a 149 °F)
<b>Aspectos físicos</b>	
Altura	De 16,4 mm a 25,8 mm (de 0,64 pulg. a 1,01 pulg.)
Anchura	292 mm (11,49 pulg.)
Profundidad	214 mm (8,42 pulg.)
<b>Peso (con batería de 3 celdas):</b>	
Configurable a menos de	1,35 kg (2,98 lb)
<b>Entorno del equipo</b>	
<b>Intervalo de temperatura:</b>	
En funcionamiento	De 0 a 35 °C (de 32 a 95 °F)
En almacenamiento	De -40 a 65 °C (de -40 a 149 °F)
<b>Humedad relativa (máxima):</b>	
En funcionamiento	Del 10 % al 90 % (sin condensación)
En almacenamiento	Del 5 % al 90 % (sin condensación)
<b>Vibración máxima (utilizando un espectro de vibración aleatoria que simula el entorno del usuario):</b>	
En funcionamiento	0,66 GRMS
Sin funcionamiento	1,3 GRMS
<b>Impacto máximo (en estado operativo: medido con Dell Diagnostics en ejecución en el disco duro y 2 ms de pulso de media onda sinusoidal; en estado no operativo: medido con los cabezales de la unidad de disco duro en posición de reposo y 2 ms de pulso de media onda sinusoidal):</b>	
En funcionamiento	110 GRMS
Sin funcionamiento	160 GRMS
<b>Altitud (máxima):</b>	
En funcionamiento	De -15,2 a 3.048 m (de -50 a 10.000 pies)
En almacenamiento	De -15,2 a 10.668 m (de -50 a 35.000 pies)
Nivel de contaminación atmosférica	G2 o menos de acuerdo con ISA-571.04-1985

## 8.3.2. Sistema de adquisición: NI – DAQ 6009



Ventas  
España  
91 640  
0085begin\_of\_the\_skype\_highlighting  
91 640  
0085 end\_of\_the\_skype\_highlighting  
ni.spain@ni.com

### NI USB-6009

#### DAQ Multifunción de Bajo Costo de 14 Bits, 48 kS/s

- 8 entradas analógicas (14 bits, 48 kS/s)
- 2 salidas analógicas (12 bits a 150 S/s), 12 E/S digitales; contador de 32 bits
- Energizado por bus para una mayor movilidad, conectividad de señal integrada
- La versión OEM está disponible
- Compatible con LabVIEW, LabWindows/CVI y Measurement Studio para Visual Studio .NET
- El software de NI-DAQmx y software interactivo NI LabVIEW SignalExpress LE para registro de datos



### Información General

El USB-6009 de National Instruments brinda funcionalidad de adquisición de datos básica para aplicaciones como registro de datos simple, medidas portátiles y experimentos académicos de laboratorio. Es accesible para uso de estudiantes y lo suficientemente poderoso para aplicaciones de medida más sofisticadas. Para usuarios de Mac OS X y Linux, descargue el software NI-DAQmx Base y programe el USB-6009 con LabVIEW o C.

Para cursos suplementarios con experimentos prácticos sobre teoría de simulación, medidas y automatización, NI desarrolló un Paquete de Estudiante USB-6009 que incluye una copia de LabVIEW Edición de Estudiante. Estos paquetes son exclusivamente para estudiantes, proporcionándoles una herramienta de aprendizaje potente, práctica y a bajo costo. Visite la página de productos académicos de NI en <http://www.ni.com/academic/measurements.htm> para más detalles.

Para un muestreo más rápido, medidas más precisas y mayor número de canales, considere los dispositivos de adquisición de datos de alto rendimiento energizados por bus USB, NI USB-6210 y NI USB-6211.

Cada dispositivo USB de adquisición de datos incluye una copia de NI LabVIEW SignalExpress LE así usted puede adquirir, analizar y presentar datos rápidamente sin programación. Además de LabVIEW SignalExpress, los módulos de adquisición de datos por USB son compatibles con las siguientes versiones (o posteriores) de software de aplicación - LabVIEW 7.x, LabWindows™/CVI 7.x o Measurement Studio 7.x. Los módulos USB de adquisición de datos también son compatibles con Visual Studio .NET, C/C++ y Visual Basic 6.

### Especificaciones

#### Documentos con Especificaciones

- Especificaciones (2)
- Hoja de Datos

#### Resumen de Especificaciones

General

Producto	USB-6009
Familia de Productos	DAQ Multifunción
Formato Físico	USB
Sistema Operativo/Objetivo	Windows , Linux , Mac OS , Pocket PC
Familia de Productos DAQ	<span class="help" onmouseover="showTitleNIFN(event,'GLOSSARY_DEF','bser','ESA');" onmouseout="hideTitleNIFN();">Serie B</span>
Tipos de Medida	Voltaje
Compatibilidad con RoHS	Sí
<b>Entrada Analógica</b>	
Canales	8 , 4
Canales de una sola terminal	8
Canales Diferenciales	4
Resolución	14 bits
Velocidad de Muestreo	48 kS/s
Rendimiento	48 kS/s
Máx. Voltaje de Entrada Analógica	10 V
Rango de Voltaje Máximo	-10 V , 10 V
Precisión Máxima del Rango de Voltaje	138 mV
Rango de Voltaje Mínimo	-1 V , 1 V
Mínima Precisión del Rango de Voltaje	37.5 mV
Número de Rangos	8
Muestreo Simultáneo	No
Memoria Interna	512 B
<b>Salida Analógica</b>	
Canales	2
Resolución	12 bits
Máx. Voltaje de Salida Analógica	5 V
Rango de Voltaje Máximo	0 V , 5 V
Precisión Máxima del Rango de Voltaje	7 mV
Rango de Voltaje Mínimo	0 V , 5 V
Mínima Precisión del Rango de Voltaje	7 mV

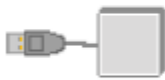
Razón de Actualización	150 S/s
Capacidad de Corriente Simple	5 mA
Capacidad de Corriente Total	10 mA
<b>E/S Digital</b>	
Canales Bidireccionales	12
Canales de Entrada Únicamente	0
Canales de Salida Únicamente	0
Número de Canales	12 , 0 , 0
Temporización	Software
Niveles Lógicos	TTL
Entrada de Flujo de Corriente	Sinking , Sourcing
Salida de Flujo de Corriente	Sinking , Sourcing
Filtros de Entrada Programables	No
¿Soporta Estados de Encendido Programables?	No
Capacidad de Corriente Simple	8.5 mA
Capacidad de Corriente Total	102 mA
Temporizador Watchdog	No
¿Soporta Protocolo de Sincronización para E/S?	No
¿Soporta E/S de Patrones?	No
Máximo Rango de Entrada	0 V , 5 V
Máximo Rango de Salida	0 V , 5 V
<b>Contadores/Temporizadores</b>	
Número de Contadores/Temporizadores	1
Operaciones a Búfer	No
Eliminación de Rebotes	No
Sincronización GPS	No
Rango Máximo	0 V , 5 V
Frecuencia Máxima de la Fuente	5 MHz
Entrada Mínima de Ancho de Pulso	100 ns
Generación de Pulso	No
Resolución	32 bits

Estabilidad de Tiempo	50 ppm
Niveles Lógicos	TTL
<b>Especificaciones Físicas</b>	
Longitud	8.51 cm
Ancho	8.18 cm
Altura	2.31 cm
Conector de E/S	Terminales de tornillo
<b>Temporización/Disparo/Sincronización</b>	
Disparo	Digital
Bus de Sincronización (RTSI)	No

## Precios

### NI USB-6009 Paquete Completo

Cada NI USB-6009 requiere:



NI USB-6009



Software

Mueva los íconos de arriba para aprender por qué necesita cada producto en el paquete.

### NI USB-6009 y Accesorios

Subtotal de Hardware: € 269

NI USB-6009 - 779026-01begin_of_the_skype_highlighting 779026-01 end_of_the_skype_highlighting	Cant.	1	€ 269 cada uno
Juego de Accesorios para NI USB-6008/6009 - 779371-01begin_of_the_skype_highlighting 779371-01 end_of_the_skype_highlighting	Cant.	0	€ 29 cada uno
Accesorio de la Serie NI USB-6000 para Generación de Prototipos - 779511-01begin_of_the_skype_highlighting 779511-01 end_of_the_skype_highlighting	Cant.	0	€ 29 cada uno

### Software

Subtotal Software: € 0

LabVIEW Signal Express Cant. for Windows (Inglés) -		0	€ 799 cada uno
--	--	---	----------------

779037-35begin\_of\_the\_skype\_highlighting  
779037-35 end\_of\_the\_skype\_highlighting

#### Cambiar Software

Aprenda sobre otras opciones de software, SO y lenguaje compatibles.

Días estimados para el envío:

Precio del Paquete: € 269

[Añadir a Carro >>](#)

Garantíe su precio por 30 días (aprenda más)Cotización al Instante

[Save for Later](#)

Usted ha seleccionado **Spain** como el país donde los productos serán utilizados (cambiar).

## Servicios

---

### Garantía Extendida

National Instruments diseña y fabrica todos los productos para reducir fallas, pero sabemos que pueden ocurrir fallas inesperadas. La garantía extendida le proporciona un precio económico fijo al momento de comprar su sistema que cubre cualquier costo de reparación por hasta tres años. Además, ofrecen los siguientes beneficios:

- Ahorro significativo en comparación con incidentes que requieren reparación individual.
- NI lleva a cabo localización de fallas, diagnósticos y reparación en cualquier momento si el sistema del producto falla.
- Todas las partes y costos por mano de obra están cubiertos, así como cualquier ajuste necesario para restaurar completamente las especificaciones de manufactura del hardware

Para mayor información sobre sus opciones de garantía:

- Aprenda Más sobre los Servicios de Garantía [<http://www.ni.com/services/esa/warranty.htm>]
- Hable con un Experto Sobre Garantías Extendidas [[javascript:openCallMeWindowCTA\(document.referrer,'%20US'\)](http://www.ni.com/services/esa/warranty.htm#callme)]
- Vea las Políticas de la Garantía de Reparación [[http://www.ni.com/services/warranty\\_repair\\_policies.htm](http://www.ni.com/services/warranty_repair_policies.htm)]

### Calibración

NI reconoce la necesidad de mantener los dispositivos calibrados adecuadamente para poder obtener medidas de alta precisión. Proporcionamos procedimientos de calibración manual, servicios para calibrar sus productos y software de calibración automatizada para calibrar varios de nuestros productos de medida.

- Aprenda Más sobre los Servicios Calibración [<http://www.ni.com/services/calibration.htm>]

### Capacitación

La capacitación de NI es el camino más rápido y acertado a la productividad con las herramientas de NI y al desarrollo exitoso de aplicaciones.

- Aprenda más acerca de la Capacitación y Certificación de National Instruments [<http://www.ni.com/training/esa/>]
- Vea información y horarios de cursos [<http://sine.ni.com/apps/utf8/nisv.custed>]

### Servicios de Reparación

Usted puede devolver su producto registrado bajo garantía sin costos adicionales por mano obra y de partes. NI ofrece localización de fallas, diagnósticos y reparación en cualquier momento si el sistema falla así como cualquier ajuste necesario para restaurar completamente las especificaciones de manufactura del hardware.

- Aprenda Más sobre los Servicios de Reparación [<http://www.ni.com/services/esa/warranty.htm>]
- Comuníquese con nosotros para obtener un formulario de Return Material Authorization (RMA) e instrucciones de envío [[http://sine.ni.com/apps/utf8/nicc.call\\_me](http://sine.ni.com/apps/utf8/nicc.call_me)]
- Vea el estado de su petición de RMA en línea [<http://www.ni.com/support/servicioereq/>]
- Registre su producto [<http://www.ni.com/register>]

### Soporte Técnico

## Recursos

---

### Información Adicional de Productos

- Manuales (4)
- Dibujos Dimensionales (2)
- Certificación de Producto

### Información Relacionada

- Adquisición de Datos USB para OEM
- Descargue Controladores de Adquisición de Datos
- Software Interactivo para Registro de Datos NI LabVIEW SignalExpress

### 8.3.3. Sensor ECG: EKG-Flex/Pro: SA9306M



## Thought Technology ECG/EKG Sensor (SA9306M)

TTL T9306M  
Price: \$275.00  
Free Shipping to US & Canada

The Thought Technology pre-amplified electrocardiograph (ECG/EKG) sensor (T9306M) for the Procomp Infiniti Systems or Procomp+. Used for directly measuring electrical activity of the heart.

The T9306M EKG Sensor connects via extender cables for a single channel hook up.

This cable is approximately 48 in (122.5 cm) in length and is terminated with a TTL T8720M sEMG Sensor Extender cable (included), with 3 snap style receptacles that accept many popular disposable sensors.

Skin surface sensors are optional, BMI GS26, or GS-27 are recommended.

**Operating Principle:**  
Use the Sensor Extender Cable (TTL T8720M). The surface electrodes monitor electrical activity from the heart muscle. Place the active electrodes along the central axis of the heart and the reference electrode on the opposite side of the chest.

### How it works

EKG (Electrocardiography) is similar to sEMG (surface Electromyography): The sensor detects and amplifies the small electrical voltage that is generated by the heart muscle when it contracts.

The actual EKG signal is measured in microvolts ( $\mu\text{V}$ ). Most often, though, the clinically useful measures are those that are computed from the raw EKG: the heart rate (HR) and its inverse, the inter-beat interval (IBI). HR is measured in beats per minute (B/min or Bm) and IBI, in milliseconds (ms).

The standard EKG sensor placement requires that the negative sensor be placed on the right shoulder, the positive sensor be placed on the lower center or left side of the chest (xyphoid process) and the ground sensor, on the left shoulder.

When a chest placement is impossible or too invasive, it is possible to place the sensors on the forearms. For a forearm placement, the negative lead should be placed on the right forearm, while both the positive and ground leads are placed on the left forearm.

EKG measures muscle activity by detecting and amplifying the tiny electrical impulses that are generated by muscle fibers when they contract. Since all the muscle fibers within the recording area of the surface sensor contract at different rates, the signal detected by the sensor is a constantly varying difference of potential between its positive and negative sensors. The number of muscle fibers that are recruited during any given contraction depends on the force required to perform the movement. Because of this, the intensity (amplitude) of the resulting electrical signal is proportional to the strength of contraction.

## Includes

- EKG-Flex/Pro Sensor with Approx. 48" Cable
- EKG Sensor Extender Cable - 3 Snap Lead (Approx. 27" long)
- EKG Sensors (9/pkg)

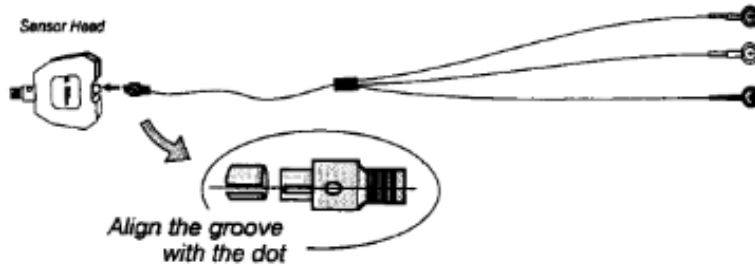
## Specifications

- Channel Bandwidth: 0.50Hz - 1kHz
- Accuracy: +/- 3uV RMS, +/-5% of reading @ 25 to 30 degrees C
- Signal Input Range: 0 - 12mV RMS
- Signal Output Range: 0 - 600mV RMS
- Input/Output Gain: 50
- Sensitivity: <0.1mV RMS
- Input Impedance: 1,000,000 Mohms in parallel with 10pF
- CMMR: >130dB
- Size (Approx): 1.45" x 1.45" x 0.45" ( 37mm x 37mm x 12mm)
- Weight (Approx): 1oz ( 25g )
- Current Consumption: <1.5mA

**Thought Technology Ltd.**  
**EKG SENSOR EXTENDER CABLE - 3 LEADS**  
**ITEM #T8710M**

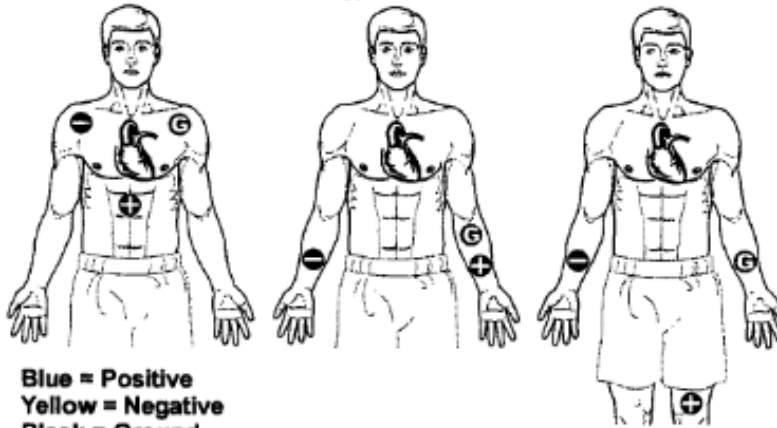
**To connect:**

Insert the extender cable in the sensor head, making sure to align the guiding dot on the cable connector with the groove on the sensor head:



**To place:**

Use Thought Technology's Uni-Gel electrodes.



**Blue = Positive**  
**Yellow = Negative**  
**Black = Ground**

*\* 90 day warranty void if damage is incurred through misuse of the equipment.*

**Thought Technology LTD**  
**PINOUTS AND CONNECTIONS FOR**


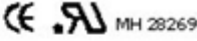
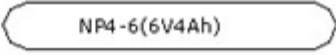
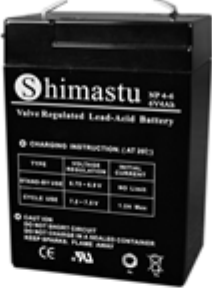
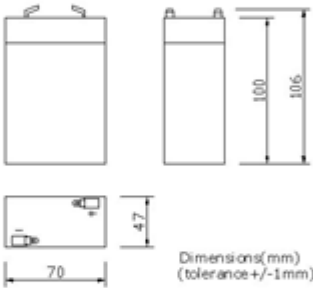

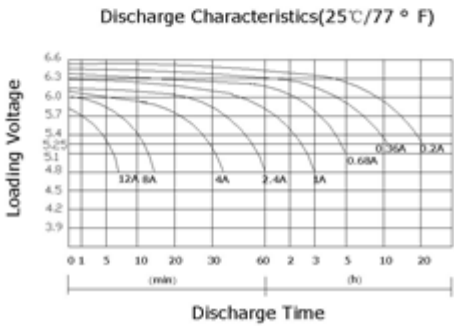
EKG Pro/Flex SA9306



Pin1 =Ground  
 Pin2 =Sensor ID Output  
 Pin3 = Signal  
 Pin4 = Power (+7.26 V)  
 Signal Range (Pin3) 2.8V±1.696V

### 8.3.4. Otros equipos utilizados

#### 8.3.4.1. Fuente de alimentación para sensor ECG

			 MH 28269			
<b>PHYSICAL SPECIFICATION</b>						
Nominal Voltage		6V				
Nominal Capacity (20HR)		4Ah				
Dimension	Length	70±1mm (2.76 inches)				
	Width	47±1mm (1.85 inches)				
	Container Height	100±1mm (3.94 inches)				
	Total Height	106±1mm (4.17 inches)				
Weight		Approx. 0.67Kg (1.48lbs)				
Standard Terminal		Fasten Tab No.187E				
 <p style="text-align: center;">Dimensions(mm) (tolerance +/-1mm)</p>				F1(Fasten Tab No.187E) Standard Terminal  <p style="text-align: right;">(in mm)</p>		
<b>ELECTRICAL SPECIFICATION</b>						
<b>Characteristics</b>				<b>Discharge curves @ 25°C (77 ° F)</b>		
Capacity	20 hour rate (0.2A)	4Ah		 <p style="text-align: center;">Discharge Characteristics(25°C/77 ° F)</p>		
	10 hour rate (0.36A)	3.8Ah				
Capacity affected by temperature	5 hour rate (0.68A)	3.4Ah				
	3 hour rate (1A)	3Ah				
	1hour rate (2.4A)	2.4Ah				
Capacity affected by temperature		40°C (104 °F)	102%			
		25°C (77 °F)	100%			
		0°C (32 °F)	85%			
Internal Resistance		Full charged battery (25°C,77 °F) 25mΩ				
Constant Voltage Charge	Cycle	Initial Charging Current less than 1.2A Voltage 7.20~7.50V at 25°C (77°F) Temperature Coefficient -15mV/°C				
	Standby	No limit on Initial Charging Current Voltage 6.75~6.90V at 25°C (77°F) Temperature Coefficient -10mV/°C				

### 8.3.4.2. Electrodo



## Kendall / Tyco ARBO Disposable Electrodes 50pkg

BRD H124SG  
Price: \$15.00

These high quality disposable electrodes are to be used to measure EEG, ECG and EMG. They are to be used once and are very handy because of integrated gel. They adhere very well to the skin and are clean to use. One bag contains 50 electrodes. Ø 24 mm. Electrode type: H124SG

### Features

- Type: H124SG
- Ref: 31.1245.21
- Ag/AgCl(=silver silver chloride) sensor ø 24 mm
- Latex free
- Patented gel formula
- (AAMI standards)
- Short-term Monitoring
- ECG, EMG, ERP and EEG

This ECG/EMG/EEG electrode is ideal for a variety of purposes and is frequently used for Neurofeedback and Biofeedback.

The H124SG has a unique, patented pre-gelled adhesive side with non-irritating gel, especially developed to prevent allergic reactions. These foam electrode is latex free and therefore suitable for every skin type.

The snap-on connector can easily be pushed on or removed from the electrode lead. Therefore you have optimal user friendliness as you dispose this electrode after every session! No more greasy electrodes as the H124SG will cover the surface of the electrode. This extends the life of your electrodes.

#### **8.4. Anexo 4: Manual de usuario de la aplicación.**

**NUEVO MÉTODO BASADO EN EL HRV  
PARA LA EVALUACIÓN DE HMIs Y  
SISTEMAS ITS PARA TRANSPORTE  
POR CARRETERA INTEGRANDO  
FACTORES PERSONALES,  
TEMPORALES Y AMBIENTALES**

**Manual de Usuario**

**VERSION 2015\_01**

## **Contenido**

- 1. Introducción**
- 2. Instalación, configuración y arranque de la herramienta**
  - 2.1. Instalación
  - 2.2. Configuración
  - 2.3. Arranque
- 3. Patrón**
- 4. Proyecto**
- 5. Evaluación**

# 1. Introducción

El presente anexo pretende ser una guía de instalación y uso de la herramienta desarrollada para extraer e interpretar, de forma ágil, el mayor conocimiento posible de los datos relacionados con el nivel de alerta y la fatiga, obtenidos a partir de un sistema de información basado en cuatro parámetros normalizados: Media, SD1, PSD de LF y PSD de HF.

Las principales características tecnológicas de esta herramienta son:

- **Basada en el lenguaje de programación Java:** esta característica otorgará a la herramienta la posibilidad de ejecutarse en diferentes Sistemas Operativos (multiplataforma). También dadas las cualidades de este lenguaje de programación, propiciará su fácil mantenimiento y evolución. Actualmente la herramienta soporta versiones de la plataforma superiores a Java 8.
- **Motor de aprendizaje automático y minería de datos basado en WEKA:** Weka (*Waikato Environment for Knowledge Analysis*, en español «entorno para análisis del conocimiento de la Universidad de Waikato») es una plataforma de software para el aprendizaje automático y la minería de datos escrito en Java y desarrollado en la Universidad de Waikato. Weka es software libre distribuido bajo la licencia GNU-GPL.

## 2. Instalación, configuración y arranque de la herramienta

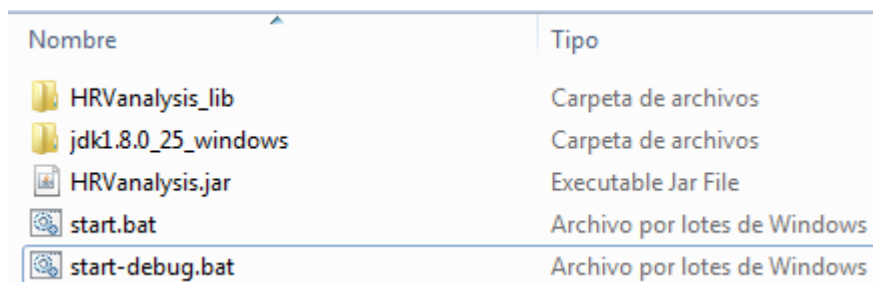
El siguiente apartado resume el proceso de instalación, configuración y arranque de la herramienta.

### 2.1. Instalación

La herramienta se distribuye bajo un único paquete comprimido, el proceso de instalación será tan sencillo como descomprimirlo bajo el directorio deseado.

Descomprimiendo el paquete de instalación se crearán varios directorios y archivos, que se detallan a continuación:

- **jdk1.8.0\_25\_windows**: máquina virtual Java para Windows. (Si se desea instalar la herramienta bajo un Sistema Operativo Linux será necesario instalar la máquina virtual Java adecuada a la arquitectura del sistema)
- **HRVanalysis\_lib**: librerías de terceros empleadas para la implementación de la herramienta. En este directorio se encontrará el motor de WEKA.
- **HRVanalysis.jar**: paquete ejecutable por la máquina virtual Java. Representa el software desarrollado para implementar la herramienta.
- **start.bat**: archivo ejecutable que lanzará la herramienta.
- **start-debug.bat**: archivo ejecutable que lanzará la herramienta en modo “*debug*”, para poder depurar cualquier “bug” que contenga el software.



Nombre	Tipo
HRVanalysis_lib	Carpeta de archivos
jdk1.8.0_25_windows	Carpeta de archivos
HRVanalysis.jar	Executable Jar File
start.bat	Archivo por lotes de Windows
start-debug.bat	Archivo por lotes de Windows

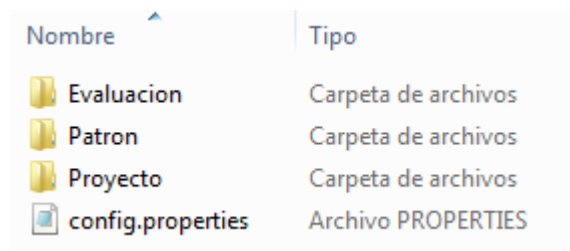
Figura 1: Directorio de instalación de la herramienta

### 2.2. Configuración

Al arrancar el software por primera vez se creará en el directorio del usuario (c:\users\UsuarioX) una nueva carpeta llamada **ITS\_Assessment\_HRV\_Analysis**. Este nuevo directorio será la carpeta de

trabajo de la herramienta. En él se encontrarán los siguientes ficheros y directorios:

- **config.properties**: fichero de propiedades clave-valor donde se configuran los umbrales establecidos por defecto en la pestaña *Evaluación* (ver apartado 4).
- **Patron**: directorio donde se guardarán todos los patrones generados con la herramienta y los ficheros de datos ARF a ser procesados.
- **Proyecto**: directorio donde se guardarán todos los proyectos asociados a los patrones generados previamente y los ficheros de datos ARF a ser procesados.
- **Evaluación**: directorio donde se encontrarán los informes generados en las distintas parametrizaciones de evaluación de proyectos.



Nombre	Tipo
Evaluacion	Carpeta de archivos
Patron	Carpeta de archivos
Proyecto	Carpeta de archivos
config.properties	Archivo PROPERTIES

Figura 2: Directorio de trabajo de la herramienta

## 2.3. Arranque

Para arrancar la herramienta simplemente será necesario ejecutar el archivo **start.bat** en Windows o ejecutar directamente el paquete **HRVanalysis.jar** en el caso de sistemas Linux.

Tal y como se muestra en la Figura 3, la herramienta se compone principalmente de tres pestañas: Patrón, Proyecto y Evaluación. El uso de cada una de estas tres pestañas se detallará en las siguientes secciones.

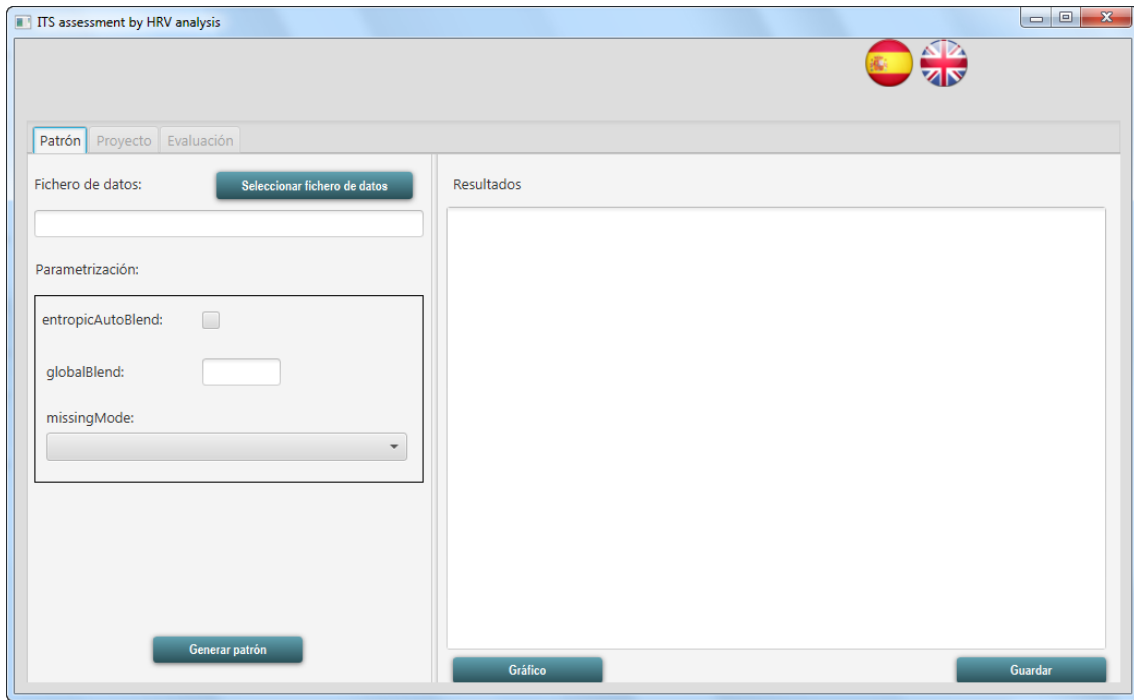


Figura 3: Estado inicial de la herramienta

### 3. Patrón

El primer paso para poder realizar una evaluación es generar un patrón a partir de la información contenida en un archivo “.arff” donde se encuentre la información con las variables “Media, SD1, PSD de LF y PSD de HF” normalizadas obtenidas a partir de los registros ECG cuyas condiciones sirvan de referencia.

Para generar un patrón será necesario seguir los siguientes pasos:

1. **Seleccionar fichero de datos:** archivo “.arff” donde se encontrará la información con las cuatro variables mencionadas en la introducción a este anexo. Estos ficheros deben de ubicarse en el directorio de trabajo (sección 2.2) bajo en subdirectorío “Patrón”.
2. **Parametrización:** el método de clasificación de los datos podrá ser ajustado mediante parámetros. Por defecto, la herramienta se ejecutará con una parametrización estándar.
3. **Generar patrón:** lanza el proceso de entrenamiento y clasificación del patrón. Al finalizar se mostrarán los resultados en el panel lateral, habilitándose también la pestaña de **Proyecto**.

TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
0.804	0.051	0.767	0.804	0.785	0.957	A
0.7	0.076	0.678	0.7	0.689	0.908	B
0.686	0.074	0.68	0.686	0.683	0.914	C
0.68	0.071	0.684	0.68	0.682	0.911	D
0.752	0.061	0.731	0.752	0.741	0.934	E
0.617	0.015	0.8	0.617	0.697	0.961	F
Weighted Avg.	0.714	0.062	0.715	0.714	0.713	0.928

=== Confusion Matrix ===						
	a	b	c	d	e	f
a	653	92	37	13	14	3
b	107	609	76	49	22	7
c	54	113	597	89	10	7
d	20	63	113	586	69	11

Figura 4: Generar patrón

Al pulsar en **Generar patrón**, aparecerán dos botones nuevos:

- **Gráfico**: visualiza un gráfico basado en la matriz de confusión y el concepto de “Tendencia de los errores” manejado en durante el análisis del caso de estudio 4 (línea azul). Las líneas verde y roja distinguen el comportamiento de la matriz de confusión por encima (línea roja) y por debajo (línea verde) de la diagonal principal. La suma de las ordenadas de las líneas verde y roja debe ser el valor de la correspondiente ordenada en la línea azul (figura 5).
- **Guardar**: permite guardar en el directorio de trabajo diferentes parametrizaciones y sus resultados para un mismo fichero de datos ARF. Si existen patrones ya generados para un mismo fichero ARF, se mostrará un menú desplegable donde se podrán volver a cargar con su configuración y sus resultados previamente generados (figura 6).

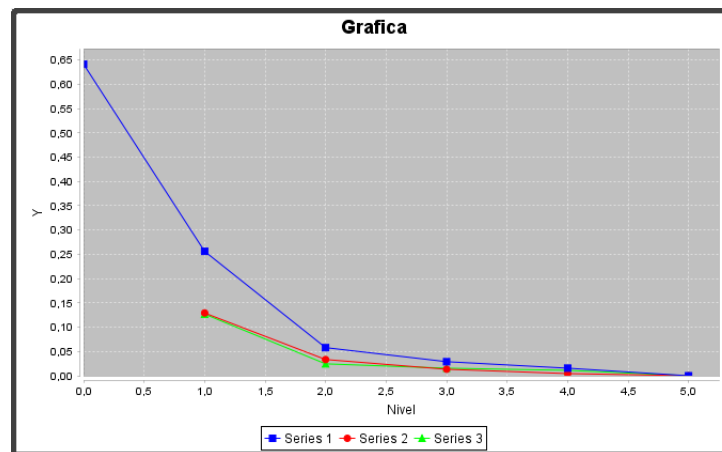


Figura 5: Gráfico patrón

Patrón Proyecto Evaluación

Fichero de datos:  Ejemplo\_10P - PEQUENIO.arff

Parametrización:

entropicAutoBlend:

globalBlend:

missingMode:

Nombre modelo:

Resultados

Kappa statistic  
Mean absolute error

=== Detailed Accuracy By

TP Rate	FP Rate
0.828	0.067
0.6	0.095
0.511	0.067
0.622	0.136
0.6	0.051
0.706	0.022

Weighted Avg. 0.641 0

=== Confusion Matrix ===

a	b	c	d	e	f	<- classifi
77	12	0	2	2	0	a = A
14	54	7	13	2	0	b = B
3	13	46	24	2	2	c = C
2	6	15	51	8	0	d = D
5	3	0	9	36	7	e = E
0	0	2	2	6	24	f = F

Figura 6: Patrones previamente generados

## 4. Proyecto

El último paso previo a la evaluación es generar un proyecto a partir de la información contenida en un archivo “.arff” y empleando como “modelo entrenado” el patrón generado en el paso anterior (sección 3).

Para generar un proyecto será necesario seguir los siguientes pasos:

1. **Seleccionar fichero de datos:** archivo “.arff” donde se encontrará la información con las cuatro variables mencionadas en la introducción a este anexo. Estos ficheros deben de encontrarse en el directorio de trabajo (sección 2.2) bajo en subdirectorío “**Proyecto**”.
2. **Generar proyecto:** lanza el proceso de clasificación de resultados gracias al patrón entrenado previamente (sección 3). Al finalizar se mostrarán los resultados en el panel lateral, habilitándose también la pestaña de **Evaluación**.

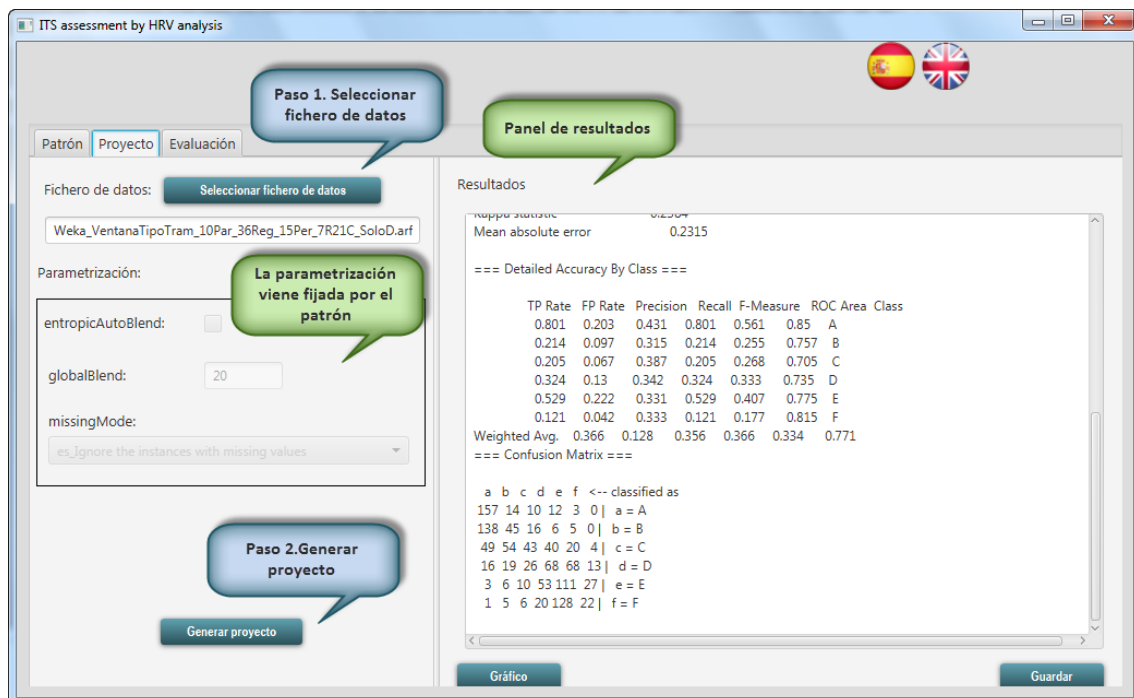


Figura 7: Pestaña de Proyecto

Al pulsar en **Generar proyecto**, aparecerán dos botones nuevos:

- **Gráfico:** visualiza un gráfico basado en la matriz de confusión y el concepto de “Tendencia de los errores” manejado en durante el análisis del caso de estudio 4 (línea azul). Las líneas verde y roja distinguen el comportamiento de la matriz de confusión por encima (línea roja) y por debajo (línea verde) de la diagonal principal. La suma de las ordenadas de las líneas verde y

roja debe ser el valor de la correspondiente ordenada en la línea azul (figura 8).

- **Guardar:** permite guardar en el directorio de trabajo el resultado dado un archivo ARF de proyecto y un patrón. Si para el patrón y el archivo ARF seleccionados existen resultados previamente generados, estos se cargarán automáticamente, teniendo la opción de volver regenerar dichos resultados.

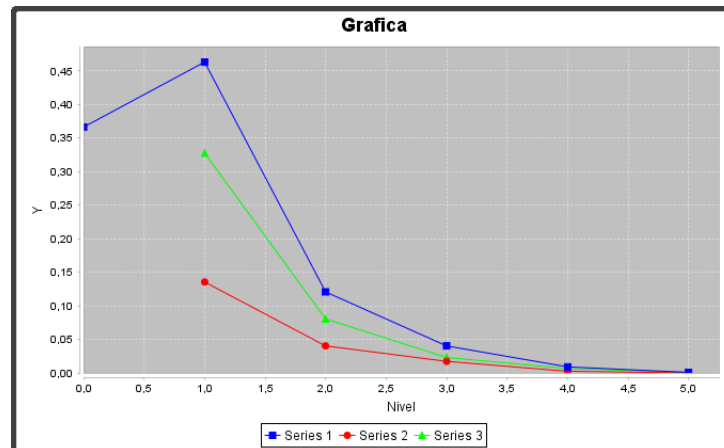


Figura 8: Gráfico proyecto

## 5. Evaluación

Es en la última pestaña donde será posible evaluar la desviación de los resultados del proyecto frente al patrón, junto con la “calidad o fiabilidad” del propio patrón empleado.

La pestaña de **Evaluación** se divide en 4 zonas diferenciadas (figura 9):

- **Evaluación del patrón:** evalúa el propio patrón. Para considerar un patrón correcto (semáforo verde) tiene que cumplir las siguientes condiciones:
  1. Su valor *Kappa* tiene que superar un umbral establecido por defecto o por el propio usuario (Los valores por defecto de los umbrales de evaluación son configurados mediante un fichero de propiedades ubicado en el directorio de trabajo, tal y como se mencionó en la sección 2.2).
  2. La suma de la diagonal principal (nivel 0) de la matriz de confusión y las diagonales inmediatamente superior e inferior (nivel 1) han de superar un umbral establecido por defecto o por el usuario (Los valores de los diferentes niveles serán expresados en porcentajes. Por ejemplo, nivel 0 = Suma diagonal principal / número de instancias).
  3. El ratio de los niveles 0 y 1 de la matriz de confusión han de superar un umbral establecido por defecto o por el usuario.
  4. La diferencia entre los valores por encima de la diagonal principal de la matriz de confusión y los valores ubicados por debajo, no ha de superar un umbral establecido por defecto o por el usuario.
- **Evaluación del proyecto:** evalúa el proyecto frente al patrón. Para considerar un proyecto correcto (semáforo verde) tiene que cumplir las siguientes condiciones:
  1. El sumatorio de los niveles seleccionados ha de superar un umbral establecido por defecto o por el usuario.
  2. El sumatorio de las diferencias de los niveles superiores e inferiores del patrón y proyecto seleccionados ha de superar un umbral establecido por defecto o por el usuario.
- **Evaluación final:** esta zona solo estará visible si las evaluaciones del patrón y proyecto han sido satisfactorias. Se lanzará al pulsar el botón **Evaluar**. Para considerar una evaluación final correcta (semáforo verde) tiene que cumplirse que el valor calculado de la condición 2 de la evaluación del proyecto se encuentre entre un rango establecido por defecto o por el usuario.
- **Gráfica:** en esta zona se presenta a modo resumen las gráficas del patrón y proyecto superpuestas.

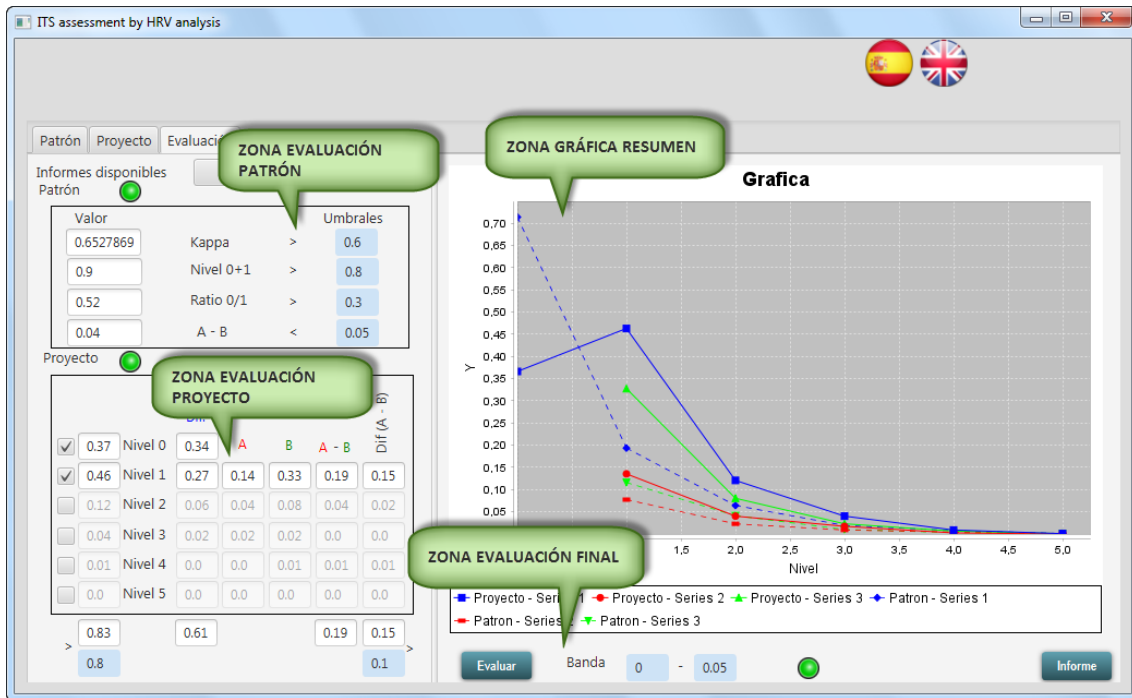


Figura 9: Zonas pestaña de Evaluación

Finalmente será posible generar un informe PDF con el resumen de las tres pestañas. Al pulsar el botón **Informe** se mostrará el PDF y se guardará la parametrización de la evaluación. Los distintos informes generados se visualizarán en el menú desplegable ubicado en la parte superior de la pantalla:

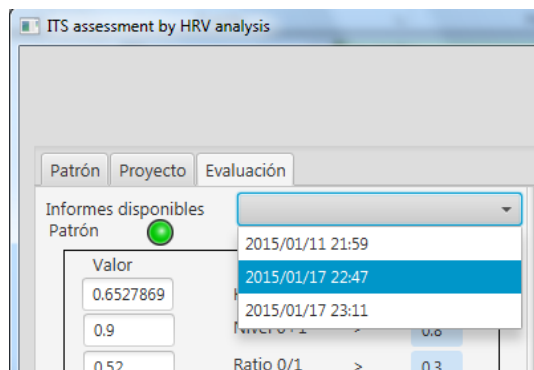


Figura 10: Informes de evaluación disponibles

# Informe de "Evaluación"

**PATRÓN**.....

**ATRIBUTOS:**

VentMediaNorm  
VentLF\_Power\_No  
VentHF\_Power\_No  
Vent\_SD1\_Norm  
VentanaTipoTram

--- Classifier model (full training set) ---

KStar Beta Verion (0.1b).

Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).

Java port to Weka by Abdelaziz Mahoul (am14@cs.waikato.ac.nz).

KStar options : -B 20 -M d

Time taken to build model: 0 seconds

Correctly Classified Instances	3333	71.3857 %
Incorrectly Classified Instances	1336	28.6143 %
Kappa statistic	0.6528	
Mean absolute error	0.1669	

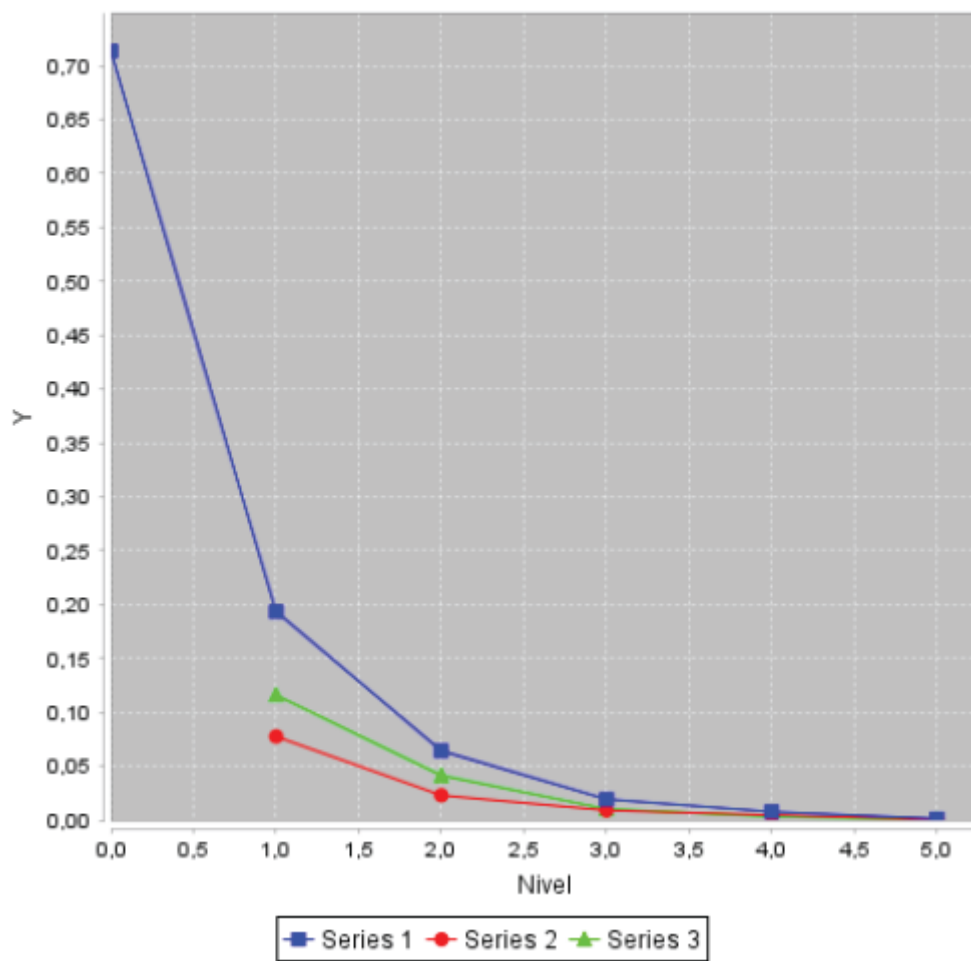
--- Detailed Accuracy By Class ---

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.804	0.051	0.767	0.804	0.785	0.957	A
	0.7	0.076	0.678	0.7	0.689	0.908	B
	0.686	0.074	0.68	0.686	0.683	0.914	C
	0.68	0.071	0.684	0.68	0.682	0.911	D
	0.752	0.061	0.731	0.752	0.741	0.934	E
	0.617	0.015	0.8	0.617	0.697	0.961	F
Weighted Avg.	0.714	0.062	0.715	0.714	0.713	0.928	

--- Confusion Matrix ---

	a	b	c	d	e	f	<-- classified as
a - A	653	92	37	13	14	3	
b - B	107	609	76	49	22	7	
c - C	54	113	597	89	10	7	
d - D	20	63	113	586	69	11	
e - E	14	20	47	91	632	36	
f - F	3	1	8	29	118	256	

## Grafica



**PROYECTO**.....

**ATRIBUTOS:**

VentMediaNorm  
VentLF\_Power\_No  
VentHF\_Power\_No  
Vent\_SD1\_Norm  
VentanaTipoTram

--- Classifier model (full training set) ---

KStar Beta Verion (0.1b).

Copyright (c) 1995-97 by Len Trigg (trigg@cs.waikato.ac.nz).

Java port to Weka by Abdelaziz Mahoul (am14@cs.waikato.ac.nz).

KStar options : -B 20 -M d

Time taken to build model: 0 seconds

Correctly Classified Instances	446	36.6174 %
Incorrectly Classified Instances	772	63.3826 %
Kappa statistic	0.2384	
Mean absolute error	0.2315	

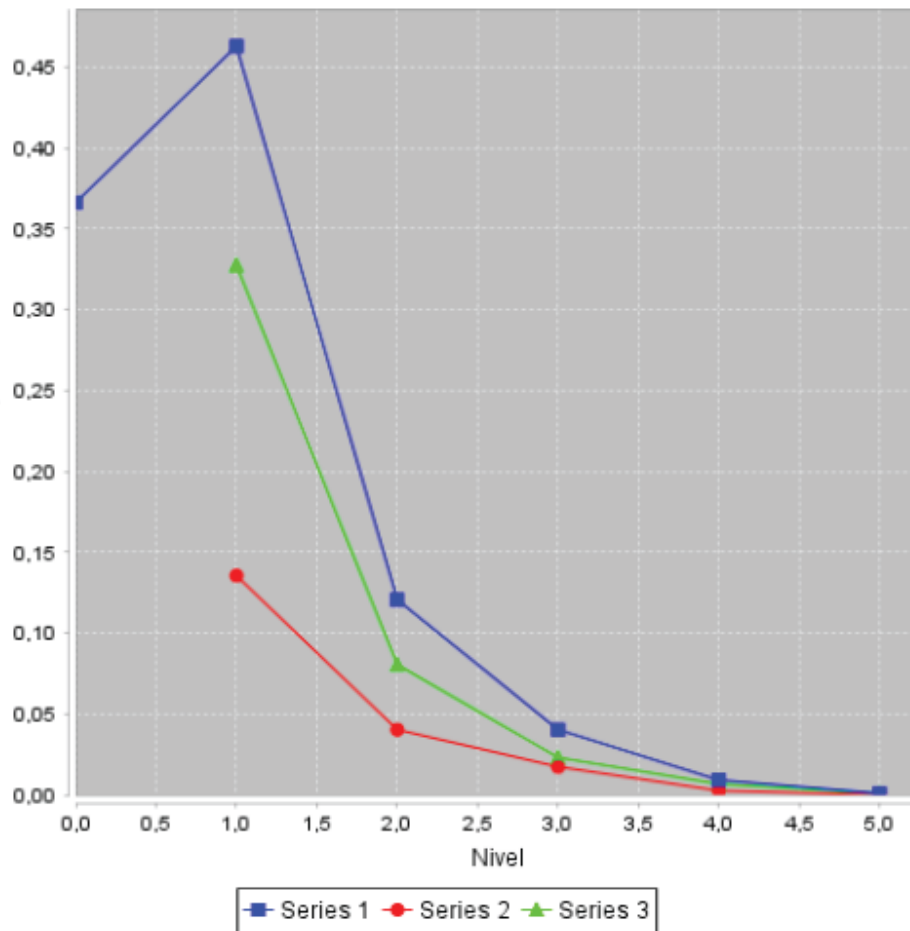
--- Detailed Accuracy By Class ---

	TP Rate	FP Rate	Precision	Recall	F-Measure	ROC Area	Class
	0.801	0.203	0.431	0.801	0.561	0.85	A
	0.214	0.097	0.315	0.214	0.255	0.757	B
	0.205	0.067	0.387	0.205	0.268	0.705	C
	0.324	0.13	0.342	0.324	0.333	0.735	D
	0.529	0.222	0.331	0.529	0.407	0.775	E
	0.121	0.042	0.333	0.121	0.177	0.815	F
Weighted Avg.	0.366	0.128	0.356	0.366	0.334	0.771	

--- Confusion Matrix ---

	a	b	c	d	e	f	<-- classified as
a - A	157	14	10	12	3	0	
b - B	138	45	16	6	5	0	
c - C	49	54	43	40	20	4	
d - D	16	19	26	68	68	13	
e - E	3	6	10	53	111	27	
f - F	1	5	6	20	128	22	

## Grafica



## EVALUACIÓN

Patrón
Proyecto
Evaluación

Informes disponibles

Patrón ●

Valor	Kappa	Umbrales
0.6527869	>	0.5
0.9	>	0.8
0.52	>	0.4
0.04	<	0.5

Proyecto ●

0.23843 Kappa > 0.01

	Dif.	A	B	A - B	Dif(A - B)
<input checked="" type="checkbox"/> 0.37 Nivel 0	0.34				
<input checked="" type="checkbox"/> 0.46 Nivel 1	0.27	0.14	0.33	0.19	0.15
<input type="checkbox"/> 0.12 Nivel 2	0.06	0.04	0.08	0.04	0.02
<input type="checkbox"/> 0.04 Nivel 3	0.02	0.02	0.02	0.0	0.0
<input type="checkbox"/> 0.01 Nivel 4	0.0	0.0	0.01	0.01	0.01
<input type="checkbox"/> 0.0 Nivel 5	0.0	0.0	0.0	0.0	0.0

> 0.83
< 0.61
> 0.19
< 0.15

0.80
0.70
0.01
0.10

**Grafica**

Nivel	Proyecto - Series 1	Patrón - Series 1	Proyecto - Series 2	Patrón - Series 2	Proyecto - Series 3	Patrón - Series 3
0.0	0.37	0.70				
1.0	0.46	0.46	0.14	0.14	0.33	0.33
2.0	0.12	0.12	0.04	0.04	0.08	0.08
3.0	0.04	0.04	0.02	0.02	0.02	0.02
4.0	0.01	0.01	0.00	0.00	0.00	0.00
5.0	0.00	0.00	0.00	0.00	0.00	0.00

Evaluar
Banda 0 - 0.05
●
Informe