Universidad de Deusto
Deustuko Unibertsitatea
University of Deusto

Promoting the Perception of Emerging Technologies in Work Environments through Edge Computing and Hybrid Intelligence

PhD dissertation by Oihane Gómez Carmona

within the doctoral program Engineering for the Information Society and Sustainable Development

Candidate	Advisors	
Oihane Gómez Carmona	Dr. Javier García Zubia	Dr. Diego Casado Mansilla

A mi ama, mi aita y a mi hermana por creer en mi
A mi aitona Carlos por inculcarme la electrónica desde niña
Y sobretodo a mi sobrino Oihan, por alegrarme cada día desde que nació.

Abstract

The advent of the Internet of Things and the massive growth of everyday devices connected to the internet are reshaping modern societies and contemporary human lifestyles. However, some intelligent environments such as the workplace are especially challenging scenarios, as the rhythm to adopt them varies due to different factors. In those spaces, privacy and trust concerns regarding data protection are critical issues that affect the perception of emerging technologies. Moreover, when deploying intelligent systems, people are usually treated as mere passive receivers of technologies and services. These are designed without considering their engagement and proactive interaction, which increases the reluctance of individuals to be part of this transformation.

In that sense, promoting a smarter collaboration between intelligent systems and people is essential to contribute to a more favorable perception of emerging technologies in work environments. This involves understanding the influence of privacy concerns and user involvement from a human-centric and -driven standpoint. With this objective, this dissertation brings together machine intelligence and human intelligence and explores the synergies between the Internet of Things, energy-efficient Artificial Intelligence techniques, and Hybrid Intelligence to provide a suitable framework for creating safer spaces. That is, spaces where technological adoption does not imply compromising the privacy of the collected information and where users are given full control over their data and the system.

For that, in this dissertation, a *privacy-by-design* approach through Edge Computing is proposed. The presented approach aims to optimize Machine Learning tasks by making them lightweight enough to be performed at the local Edge instead of the Cloud and preserve the privacy of the collected information. Moreover, the role of the end-users in those spaces is addressed from a Human-in-the-Loop perspective. That is, this dissertation suggests giving them complete control over their data and prompting them to personalize and improve the learning system according to their changing needs and participation willingness. As a result, this enables intelligent spaces that are aware of users' preferences and demands for privacy and control. The quantitative and qualitative experimental results show the potential of the proposed approach to contribute to the creation of privacy-preserving and participatory smart workplaces where positive interactions between the technology and its users can be promoted.

Resumen

La llegada del Internet de las Cosas y el crecimiento masivo de los dispositivos cotidianos conectados a internet están remodelando las sociedades modernas y transformando el estilo de vida contemporánea de los seres humanos. Sin embargo, algunos entornos inteligentes como el lugar de trabajo son escenarios especialmente desafiantes, ya que su ritmo de adopción varía por diferentes factores. En estos espacios, las preocupaciones sobre la privacidad y la confianza en la protección de los datos son cuestiones críticas que afectan a la percepción de las tecnologías emergentes. Además, al desplegar sistemas inteligentes, las personas suelen ser tratadas como meros receptores pasivos de tecnologías y servicios. Estos se diseñan sin tener en cuenta su potencial compromiso y sus expectativas de interacción proactiva, lo que aumenta la reticencia de las personas a formar parte de esta transformación.

En este sentido, promover una colaboración más estrecha entre los sistemas inteligentes y las personas es esencial para contribuir a una percepción más favorable de las tecnologías emergentes en los entornos de trabajo. Esto implica comprender la importancia de los requisitos de privacidad e implicación del usuario desde un punto de vista centrado y dirigido por el ser humano. Con este objetivo, esta disertación aúna la inteligencia de las máquinas y la inteligencia humana y explora las sinergias entre el Internet de las Cosas, las técnicas de Inteligencia Artificial de bajo consumo y la Inteligencia Híbrida para proporcionar un marco adecuado en la creación de espacios más seguros. Es decir, espacios en los que

la adopción tecnológica no implique comprometer la privacidad de la información recogida y en los que los usuarios tengan pleno control sobre sus datos y el sistema.

Para ello, en esta tesis se propone un enfoque de *privacidad-por-diseño* a través del concepto de Computación Frontera. El enfoque presentado tiene como objetivo optimizar las tareas de aprendizaje automático haciéndolas lo suficientemente ligeras como para ser realizadas en el borde local de la red, en lugar de en la nube, con el fin de preservar la privacidad de la información recogida. Además, se aborda el papel del usuario en estos espacios desde una perspectiva Humano-en-el-Bucle. Es decir, esta disertación sugiere darles un control total sobre sus datos e incitarles a personalizar y mejorar el sistema de aprendizaje según sus necesidades cambiantes y su voluntad de participación. Como resultado, esto permite diseñar espacios inteligentes donde se abordan las preferencias y demandas de privacidad y control de los usuarios. Los resultados experimentales, tanto cuantitativos como cualitativos, muestran el potencial del enfoque propuesto para contribuir a la creación de lugares de trabajo inteligentes que preserven la privacidad y sean participativos, en los que se puedan promover interacciones positivas entre la tecnología y sus usuarios.

Acknowledgements

No sería posible estar escribiendo estas líneas sin el apoyo de todas las personas que me han acompañado durante este camino. En primer lugar, me gustaría agradecer a mi madre, mi padre y mi hermana por estar ahí desde el principio. Gracias por haberme dado siempre la libertad para elegir lo quería que hacer, facilitarme las cosas para que lo hiciera y estar siempre orgullosos de mí. Por otra parte, a mi sobrino Oihan que, aunque aún no sea consciente, ha sido la motivación que necesitaba durante esta etapa final.

Como no podía ser menos, a mis directores de tesis, Javier y Diego. Al primero, principalmente por apoyarme desde el inicio y animarme a que me embarcara en esta aventura. Al segundo, por estar en el día a día tanto de la tesis como de otros proyectos que compartimos. Gracias a ambos por acompañarme durante esta larga travesía y por los consejos me habéis dado todos estos años.

A todo el grupo de MORElab, empezando por Diego I., por facilitar que esta tesis fuera posible. Mención especial a mis compañeros predocs, con quienes he compartido tantas alegrías y penas. En primer lugar, a mi compañera de tesis Ane, por entender mejor que nadie todos los altibajos de estos años y estar siempre ahí. A Adrián, por alejarme del pozo con su positividad y su inagotable energía. También a Irene, por enseñarme siempre otro punto de vista de las cosas. A todos los compañeros con los que he coincidido y compartido esta experiencia: Koldo, Unai Z. Rubén S., Aritz, Aitor, Unai A. Mikel E. David, Enrique, Rubén M, Gorka, Jesús, Josu, Oihane K., Mikel S., Unai L., Iván, y un largo etcétera.

También a todas las personas que tengo la grandísima suerte de tener a mi alrededor y que me han estado recordando que la tesis no lo es todo: Berrio, Sonia, Maite, Nai, Iñaki, Markel, Iban, Dani, Alaitz, Irene G. y tantos otros. En la vida hay gente mágica y momentos mágicos.

I would also like to thank the people at the NTNU, with whom I was lucky to share this stage of my life for almost four months. To Frank, who was essential for the core of this dissertation, for his valuable advice, involvement, and guidance. To all the colleagues from the IIK department for their warm welcome, especially Faiga, who was the best support I could have in such an experience.

Por último, al Gobierno Vasco por la beca predoctoral gracias a la cual he podido centrarme en el trabajo de esta tesis, así como a DEUSTEK por el apoyo financiero para publicar y realizar la estancia doctoral.

Eskerrik asko
Oihane Gómez Carmona

Table of Contents

List of Figures

List of Tables

Acronyms

ANN Artificial Neural Network

AVE Average Variance Extracted

CFA Confirmatory Factor Analysis

CR Composite Reliability

DL Deep Learning

DT Decision Tree

GDPR General Data Protection Regulation

HAR Human Activity Recognition

HCI Human Computer Interaction

HitL Human-in-the-Loop

IML Interactive Machine Learning

IoP Internet of People

IoT Internet of Things

KNN K-Nearest Neighbors

LG Logistic Regression

ML Machine Learning

MLP Multi-layer Perceptron

NB Naive Bayes

OHM Office Monitoring Hydration

PR Privacy Risks

RF Random Forest

SR Security Risks

SVM Support Vector Machines

"If we knew what it is we were doing, it would not be called research. Would it?".

Albert Einstein

CHAPTER 1

Introduction

THE exponential growth of the Internet of Things (IoT) is redefining the concept of intelligent environments (Zhu et al., 2006). That is, novel ubiquitous technologies are emerging, creating an interconnected world. However, despite this continuous progress in smart services and technology, human lifestyles are not evolving at the same pace as technology does. Such a fact often leads to users' reluctance and rejection towards emerging technologies (Hong et al., 2020). Among the different factors for the aforementioned reluctance, the aversion to being continuously tracked, concerns associated with who may access the collected data, or the loss of civil liberties are the most prominent ones (Paul et al., 2020; Zheng et al., 2018).

In those spaces where smart technologies have been already adopted (i.e. the so-called intelligent environments), such as some new workplaces, trust and privacy concerns are critical issues that affect the technology acceptance and the future adoption from users (Zweig and Webster, 2002). Hence, it is essential to find a trade-off between respecting user's involvement/privacy and the deployment of IoT devices in a given environment irrespective of their final objective. Moreover, traditional Artificial Intelligence (AI) approaches have left users apart, creating a barrier between them and their surrounding technology instead of promoting collaborative spaces where technology is adapted to them and not the other way around (Casado-Mansilla et al., 2019b). In this sense, future smart spaces must both guarantee the privacy of the captured information and empower the role of end-users by giving them increased control over the intelligent system and their data, creating more human-centric interactive spaces (Sicari et al., 2015).

Privacy can be promoted in IoT environments by moving the processing and computational power closer to the data source (i.e., to a local stage), boosting what is called *privacy-by-design* (Pape and Rannenberg, 2019). The previous definition is related to the concept of Edge Computing, which can also provide users with local control as a suitable means to reduce the reluctance towards emerging technologies' adoption. However, Edge architectures have to cope with their own flaws. As an example, while Cloud Computing platforms can easily scale up if temporary demand for resources requires it, Edge Computing platforms are typically resource-constrained and cannot resort to using more computational power. Moreover, when human-centric Edge solutions need to be adaptively adjusted to better fit users expectations (e.g., to include user-dependent data to personalize Machine Learning (ML) models), the computational requirements of AI tasks increases, and so does the technological challenge of its deployment. Consequently, the computational need of AI tasks is a critical factor to consider when deploying collaborative solutions in local devices. For this reason, this dissertation will propose new strategies that are devised (i) to favour the inclusion of complex computational tasks in such constrained settings and (ii) to better involve users, creating spaces that engage them to actively collaborate with the system to adapt it tho their requirements and preferences.

The remainder of this chapter is structured as follows: Section 1.1 explains the context and the motivation behind this research; Section 1.2 formulates the hypothesis, the goals to achieve, and the scope of the work; Section 1.3 describes the methodology to achieve these goals; Section 1.4 enumerates the contributions of this dissertation; and Section 1.5 presents a scheme of the dissertation.

1.1 Context and Motivation

Technological advancements are starting to accelerate the evolution of future smart environments. Now, this concept goes much further than implementing technology to achieve a digital transformation and aims to create interactive spaces where people and technology collaborate. These spaces, called smart environments, sense the physical world, give meaning to the obtained information, and trigger suitable reaction to transform human lifestyles. The IoT paradigm is the main enabler of this new vision of the surrounding space that can enhance health (Dang et al., 2019), wellbeing (Papa et al., 2020), or promote sustainable practices (Salam, 2020). Under this context, the overriding presence of technology can play a relevant role in addressing new existing societal challenges and bring added value to the technology in a way never imagined before (Majchrzak et al., 2016).

However, one of the current main challenges of integrating technology solutions in those domains is lowering the barriers between individuals and their interactions with everyday devices and smart objects (Abdelghani et al., 2019). In the context of the IoT, these devices and objects can be called ”things”. Each of them is connected, resulting in significant privacy threats, as the sensed data collected by those ”things” may contain private information concerning their users (Sicari et al., 2015). Generally speaking, as has been introduced at the beginning of this chapter, people in modern societies are averse to be continuously surveilled (i.e., monitored) by a digital entity that they do not trust in, without knowing which data they are sharing (Agaku et al., 2013). Understanding up to what extent humans are confident with the data or service offered by a ”thing” is a pivotal factor to consider for the success of any emerging technology (Voas et al., 2018).

Beyond the common privacy and security concerns that every smart space has to cope with, one specific scenario poses additional threats: the work environment. In those spaces,the concept of privacy acquires new dimensions involving a social component, as this data can be associated with the image given to third parties or with the perception of productivity and work performance (Bhave et al., 2020). In these spaces, the adherence to the technology and the potential outcomes of its deployment largely depend on employees' acceptance of the technology. In addition, the lack of understanding about the behaviour of smart technologies and services, together with the sense of disengagement with systems that end-users usually do not understand, increase the technological gap and make employees reduce their confidence and perceived value toward those services and technologies (Jacobs et al., 2019). Despite these additional barriers, the IoT can be considered an emerging paradigm to mediate the relationship between employees and technology, providing insights into the transformations that the workplace could undergo to become socially engaging places that respond to workers' needs (Casado-Mansilla et al., 2018). For this reason, it is essential not to leave these environments behind in the modernization and digitization of smart spaces.

1.1.1 From the Internet of Things to the Internet of People

Under those circumstances, successfully integrating emerging technologies in such challenging workplace scenarios is only possible if there is a better and smarter collaboration among augmented devices and people (Irizar-Arrieta et al., 2020). Some scholars demand more human-centric approaches for sensing the physical world that enable safe spaces where human and machine intelligence ally (Miranda et al., 2015). This corresponds to the paradigm of the Internet of People (IoP) that seeks to improve how people are integrated with the IoT in order to create more trustworthy spaces (Conti et al., 2017; Chen et al., 2017).

However, most of the state-of-the-art IoT ecosystems rely on remote Cloud data centers to analyze the sensed data and give meaning to the obtained information. This means that, despite being obtained locally, the information needs to travel long distances and reach remote data centers where the processing capabilities are placed (Hou et al., 2016). While this might be the typical architecture to implement, it increases the difficulty of adopting IoT due to privacy threats it entails, such as losing control over personal data (Pearson and Benameur, 2010).

To overcome this first pitfall, the overall body of knowledge should be revisited in order to reduce this gap and improve the interweaving between human beings and technology. When the confidence in the technology is highly dependent on the privacy of the information it captures, it becomes more logical to place at a local stage as much analytical power as possible since data is generated at the sensors close to the user (Zhou et al., 2019). This way, data still needs to be captured and processed to create contextual services, but the physical proximity of users and their data can reduce the critical barriers regarding the acceptance of intelligent environments. From a human perspective, this ensures that the collected sensitive information remains on personal devices and not on third-party servers, which favors privacy, confidence in technology, and acceptance. Therefore, the adoption of emerging technologies increases if personal data privacy is favored and users' control over the information is promoted, keeping it close to the place where it is produced (Nappi and de Campos Ribeiro, 2020).

In this direction, Edge Computing proposes a paradigm shift from the legacy systems, that have traditionally relied on remote Cloud data centers to carry out the most fundamental parts of data analysis, to the concept of Edge Intelligence (Deng et al., 2020). However, this transition from the Cloud to the Edge is not straightforward. The main reason is that complex computation, including highly demanding applications such as ML or AI techniques, is usually served in powerful CPUs and GPUs in big distant computers. Therefore, bringing them to the Edge poses new challenges in increasing the computation resources of Edge devices and reshaping the algorithms and applications to the existing and available resources (Samie et al., 2019).

For this reason, performing ML tasks directly on local embedded platforms at the Edge, or attached to sensors, is emerging as a growing area of research within the IoT domain, which is aligned with the vision of the IoP, as it brings computation closer to the user. This falls within the scope of Embedded ML, which seeks to bring ML tasks to resource-constrained devices and thereby break the traditional barrier with Cloud machine intelligence. Thus, the goal of Embedded ML is to create an intelligent and stand-alone concept of Edge that surrounds users , preserving the externalization of their data.

Nonetheless, despite the possibilities offered by modern hardware platforms to host such applications, the limitations of the computational capabilities of Edge and end devices hinder the execution of ML techniques on them (Dhar et al., 2019; Adegbija et al., 2017). As the technologies close to the user (i.e., Fog, Edge, and users' devices) are not powerful enough to perform their own operations by themselves, the technological challenges of integrating such demanding tasks at the Edge involve optimizing classification tasks to be embedded on resource-constrained devices rather than having them in larger platforms in the Cloud (Sakr et al., 2020).

This challenge becomes even more complicated when this local concept is extended to both the training and inference phases of the ML process in interactive learning solutions. Usually, the training stage of ML algorithms is relegated to the Cloud or distributed devices. For instance, model training from scratch (or even re-training) is outsourced to powerful external devices, while inference is performed in tiny equipment at the Edge (Wang et al., 2018; Zhou et al., 2019). Although we acknowledge that this solves the computational problems, in the vision of this new human-centric IoP, externalizing personal and sensitive data may compromise the privacy of users. In contrast, integrating the model training stage at the Edge can answer the need for privacy of new smart spaces while increasing users' perceived level of trust in the system (Zheng et al., 2019).

1.1.2 **From the Internet of People to the Human-in-the-Loop concept**

Edge Computing and Embedded ML represent a step forward in the effort to create intelligent smart solutions with their computational burden at a local stage. Nonetheless, making the Edge truly intelligent goes beyond this concept. For Garcia Lopez et al. (2015), apart from facing privacy concerns, increasing technology perception also involves including humans in the system, making them part of the decision process and giving them increased control over the management of the information they generate. This idea is in line with the enhanced convergence between the user's physical world and the cyber world of the Internet that the IoP paradigm promotes. Consequently, the role of the user needs to be carefully considered to reduce their reluctance to be part of this new scenario (Casado-Mansilla et al., 2019a). Thus, such spaces demand an increased users' control over the system and their data, preventing the threat of intelligent environments taking control over them (Kaasinen et al., 2013).

For this reason, human-centric approaches should also engage users to actively collaborate with new emerging smarts solutions, and overall with those based on AI (Amershi et al., 2014). This concept falls within the scope of Hybrid Intelligence, whereby humans are included in the loop of AI (Dellermann et al., 2019). This idea is extended to the Human-in-the-Loop (HiTL) concept, a new paradigm in which users are engaged to improve and personalize automatic AI-based solutions by iteratively interacting with the learning systems (Schirner et al., 2013). In essence, HiTL provides two key aspects: a) users feel more involved in the process of understanding the AI behind the decisions taken depending on their willingness, and b) users may educate the system rather than relying exclusively on what the machines decide for them. As a result, machine intelligence is aware of the users' needs and demands, i.e., the technology adapts to the users and not the other way around, increasing acceptance and adoption.

An example of this collaboration can be found in the personalization of Human Activity Recognition (HAR) models with user-dependent data. This collaboration is intended to improve the classification of human movements and actions using sensor data. Traditionally, systems based on AI and ML techniques start from a base model trained with generic data. However, this model may not be good enough to generalize based on new data captured in real contexts. Furthermore, it may not fit appropriately to new users which data was ever unseen before (Lin and Marculescu, 2020). Consequently, this base model may not be developed considering the particularities of the end-user. For instance, a model trained with data from right-handed people could give poor performance if used by a left-handed person, which can worsen the user experience for the end-user, who sees that the system is not able to detect their actions correctly.

To this aim, Interactive ML (IML) solutions propose a more active role of end-users, that are involved in helping the learning system to improve its detection capabilities. This collaboration is driven by providing new data, in the form of correct labels, for some of the predictions the system will make (Fails and Olsen Jr, 2003). This idea resembles the Active Learning concept, a sub-field of IML that incorporates user feedback in model training phases to give meaning to the most informative samples of the initial data. In Active Learning approaches, the system is allowed to dynamically pose queries during the training process, usually in the form of unknown data instances, to be labeled by what is called an oracle, which often is a human annotator (Settles, 2012). As such, active Learning is one of the most prominent examples of the success of the HiTL paradigm.

In this work, this collaboration is extended to create a continuous and lasting interaction that elongates beyond the training period. With this collaboration, users are included in the control loop and iteratively help the system to create personalized models that are adapted to fit their particularities *on-the-fly*. Hence, this action can improve both the model's performance and user's confidence in the intelligent system (Robert et al., 2016). In fact, confidence management has been ignored in the research agenda, and an improvement of the user experience is required to achieve the actual settlement of IoT systems in our society (Yan et al., 2014).

Nevertheless, in the same way that users should not be left aside, it is also pivotal to consider their role in these spaces and not to take their involvement level for granted. Therefore, IML solutions should be flexible enough to adapt their functionality in terms of human goals, contexts, and preferences to make ML more useful and usable for every particular individual (Gillies et al., 2016). Similarly, intelligent models should also consider an evolving scenario in which user interactivity may vary along with context or time. For example, users may decide to decrease or increase the number of interactions with the system depending on their experience of use, how they perceive the outcome of their efforts according to the system performance, or if they found those interactions too distracting.

While these characteristics are always valuable in every interactive system, they become essential in the workplace context. According to Orhan et al. (2021), in those spaces, an inadequate interaction with the technology might affect employees' performance and create a distrustful atmosphere. For this reason, different end-users may have different expectations and motivations to adapt and personalize the system when it comes to HAR models training and validation processes. Hence, end-users should be given the ability to provide feedback to the system to personalize it to their preferences, avoiding the "*one-size-fits-all*" problem.

1.1.3 Closing the loop: Hybrid Intelligence at the Edge

Therefore, to solve the aforementioned challenges for the integration of emerging technologies in the work environment, this dissertation proposes to bring together machine intelligence and human intelligence in an Edge Computing approach. Key emerging technologies, namely, IoT and AI, are applied from a human-centric perspective for HAR applications. Due to the introduced additional barriers of smart workplaces, the context and motivation of this work focus on such spaces to understand how the addressed barriers can be overcome through a more human-centric deployment of the technology. In particular, this work puts the emphasis on the specific use case of promoting healthier behaviors through the technology in those spaces, as healthcare and wellness promotion is one of the most common use cases for IoT and smart technologies (Stoyanova et al., 2020). In congruence with this, the workplace is also gaining increasing significance on overall health promotion (WHO, 2010).

The proposed Edge Computing approach envisages scenarios where positive interactions through *privacy-by-design* might be promoted. That is, to contribute to the perception of technology in those areas, while involving end-users in the process. Under this context, the IoP paradigm should significantly contribute to boosting human involvement through new privacy-aware services and, hence, minimize their reluctance to be part of smart environments (Kim et al., 2017). For that, this dissertation first presents an optimization strategy that contributes to developing more efficient and private solutions,

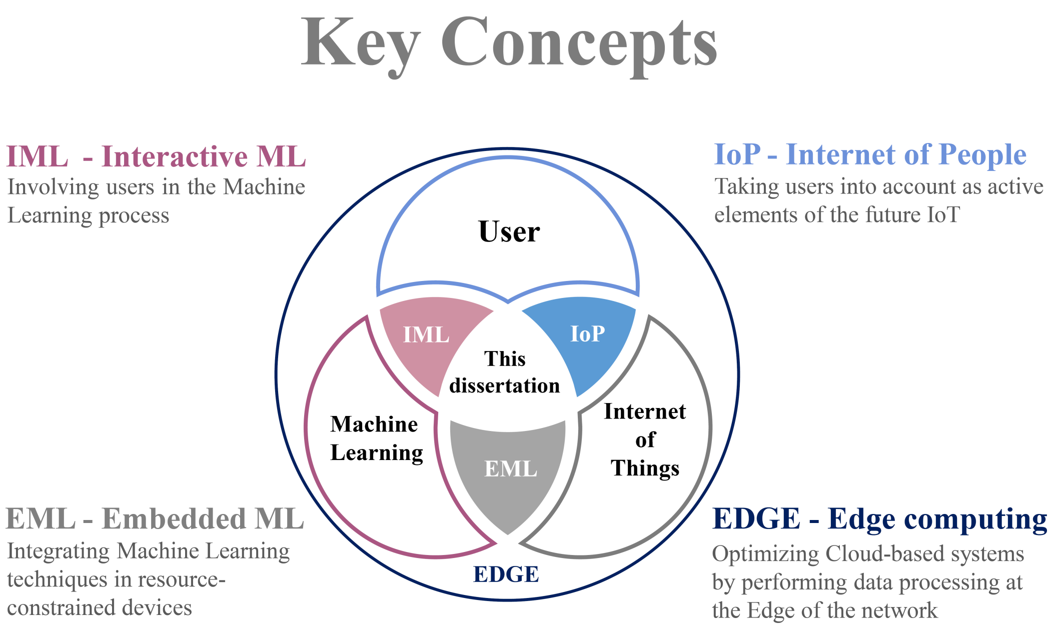

Figure 1.1: Venn diagram associating all the key concepts included in this dissertation as well as the intersection between them.

preventing data from being outsourced. The proposed strategy integrates both the inference and training stages of time series-based HAR applications into resource-constrained devices hosted at the Edge of the network.

The primary objective of this approach is to keep personal data closer to users and, at the same time, to give them control over their information. Then, once this privacy is promoted, an interactive proposal is presented to allow them to collaborate with the system and locally adapt and personalize ML models according to their needs using their own data. Consequently, the implications of hybrid and interactive intelligent systems and the requirements and expectations of users involved in collaborative smart workplaces are also evaluated. For that, a hydration monitoring system prototype to classify office employees' hydration patterns is introduced, serving as an example of how IoT technology can be integrated into those spaces.

Figure 1.1 illustrates the already mentioned factors and concepts around which this dissertation revolves, namely the user, ML, and IoT, as well as the intersections between them. Those intersection requires understanding users' expectations concerning their role within hybrid and interactive intelligent spaces while overcoming the technological challenges that demand the need for rapid model updates in resource-constrained Edge infrastructures. For this reason, the objective of creating both efficient and interactive systems hosted at the local Edge involves interdisciplinary elements derived from studying users' concepts of privacy, confidence, adoption, and technology appropriation factors. Under this context, intelligent solutions need to be optimized and reduced in size and complexity without significantly compromising their functionality and, at the same time, adapting their behaviour to users' preferences and expectations. In essence, those factors are considered altogether to offer full-data sovereignty to end-users with *privacy-by-design* approach in an interactive smart workplace scenario.

This dissertation aims to make significant contributions to the research field of intelligent environments and contribute to a more human-centered view of IoT, particularly for the workplace context. The synergy between the concept of IoT, energy-efficient AI techniques, and Hybrid Intelligence provides a suitable framework for creating more confident smart spaces where technology adoption does not involve compromising the privacy of the collected information, and where a more active role of the user in such spaces is promoted. For that, the main principle that articulates this dissertation is that people should not be treated as mere users of networked technologies and services, but their behaviour should become one of the key factors for designing technologies, turning them into real "smart users" that participate in the digital sphere that the IoT creates.

1.2 Hypothesis, Objectives and Scope

Based on the reviewed barriers, pitfalls and limitation of the current State of the Art in IoT approaches and their implications in intelligent work environments, the hypothesis of the present dissertation is:

Hypothesis. By integrating local processing and Hybrid Intelligence at the Edge, it is possible to create a system that (i) promotes the privacy of personal data and (ii) helps users to retain the control of the system and their associated personal data to improve their perception of emerging technologies in work environments.

Hence, this dissertation sets the following goal in order to validate the aforementioned hypothesis:

Goal. To design and implement strategies for a system to be hosted at the local Edge, within resource-constrained devices, that entails Hybrid Intelligence through a Human-in-the-Loop perspective to (i) promote the privacy of personal data and (ii) help users to retain the control of the system and their private data to improve the perception of emerging technologies in work environments.

This general goal can be achieved by addressing the following more specific and measurable objectives:

1. To study the current start of the art on IoT approaches applied to smart workplaces as well as Embedded ML and Hybrid Intelligence solutions.
2. To identify the main barriers and challenges of the integration of technology in such spaces and the role that privacy concerns and users' control over the information they generate may have to support this integration.
3. To design and implement suitable strategies to incorporate and optimize the training and inference stages of classification techniques for HAR applications in local Edge devices, preserving the external accessibility of data through a *privacy-by-design* concept.
4. To identify an appropriate evaluation methodology for the optimization strategies and quantitatively validate the results obtained with and without the use of the proposed strategies.

5. To extend the scope of Hybrid Intelligence, adapting the proposed *privacy-by-design* approach to include human intelligence in an interactive scenario where users have control over both the intelligent system (i.e., personalizing it) and their data.

6. To discover qualitative insights regarding the interaction between human and machine intelligence in work environments, the potential willingness of users to be committed to helping the system, and how to engage them to do so.

The resulting optimized and interactive system should also fulfill the following requirements:

1. Environment independence: Despite the particular focus on work environments, the proposed *privacy-by-design* strategy is also applicable for other contexts revolving around the user, in where privacy concerns are an issue for the deployment of IoT solutions, but is more relevant to workplaces.

2. Optimization capabilities: The studied mechanisms to integrate and optimize the computation at a local stage should be able to enhance the performance of classification techniques hosted on embedded devices for different contexts and domains, as long as they involve times-series data and continuous signals such as the ones studied for HAR applications.

3. Applicability: The different use cases analyzed through this dissertation are selected to provide an evaluation framework. The proposed approaches are independent of the selection of activities that are used as measurable examples.

Beyond that, the work presented in this dissertation does not deal with the following conditions:

1. In the context of this dissertation, the concept of privacy is approached from a human-centric perspective. For this reason, privacy is related to the psychological concepts of trust and confidence that can be generated

by avoiding externalizing users' data. That is, keeping it in local devices where this information is preserved. However, the concept of privacy and security in the IoT encompasses a broader range of technical aspects to keep IoT systems safe that are not covered in this dissertation.

2. We assume the technological limitations of the Edge of the network by selecting resource-constrained devices that may work as personal devices for data collection and processing.

3. Due to the technical limitation and implications of this dissertation, we focus our study on those ML techniques that offer the best balance between their performance and results. Therefore, Deep Learning (DL) architectures are out of the scope of this dissertation as they are unfeasible to be trained efficiently in the most limited resource-constrained devices.

4. Following the previous point, improving the state-of-the-art results of the evaluated datasets is out of the scope of this dissertation. Instead, the significance of this work relies on the optimization capabilities of the proposed approaches and the comparison of the results obtained when applying them.

1.3 Research Methodology

In order to achieve the statement and derived goals presented in the previous Section 1.2, the following strategy has been defined:

1. **Exploratory phase**: explore the literature related to the research field to build a solid theoretical background on which to support the rest of the research process. While this task of revising the literature is presented as an initial step of the research work, it is understood as an incremental and continuous process that needs to be done throughout the entire research process.

2. **Definition of the scope and validation scenario**: after the first stage of the revision of the State of the Art and based on the existing knowledge, the scope, context and motivation of this work are defined. Then, the scenario in which the project will be developed will be defined in accordance with the objectives set and knowing the limitations and advantages of the research proposal. Defining a robust initial framework is pivotal to guide future research towards the analysis that forms its basis.

3. **Specification, design and development of the solution**: at this stage, the acquired previous knowledge will be used to determine the solution that better fits the identified requirements. According to those requirements, the solution that is considered to obtain the best results is designed and developed, including the obtained inputs of previous stages and the potential iterations of this research process.

4. **Evaluation of the solution**: after the design and development phase, the implemented solution is evaluated. This evaluation requires assessing if the system does fulfill the specified requirements and verifying its validity and applicability in real environments and contexts.

5. **Final conclusions, dissemination and writing the PhD dissertation**: the final task of the research process will focus on the analysis of the obtained results, the derived conclusions and the contributions of this process. Those conclusions and contributions are expected to be innovative and of great significance and to contribute to both academic and social fields when disseminated. Finally, the research work will be completed and refined with the development of the present dissertation and with its subsequent submission and defence.

Image 1.2 illustrates this research methodology, specifying the different phases and task that it encompasses and the iterative process followed to refine the initial conceptualization of the implemented solutions.

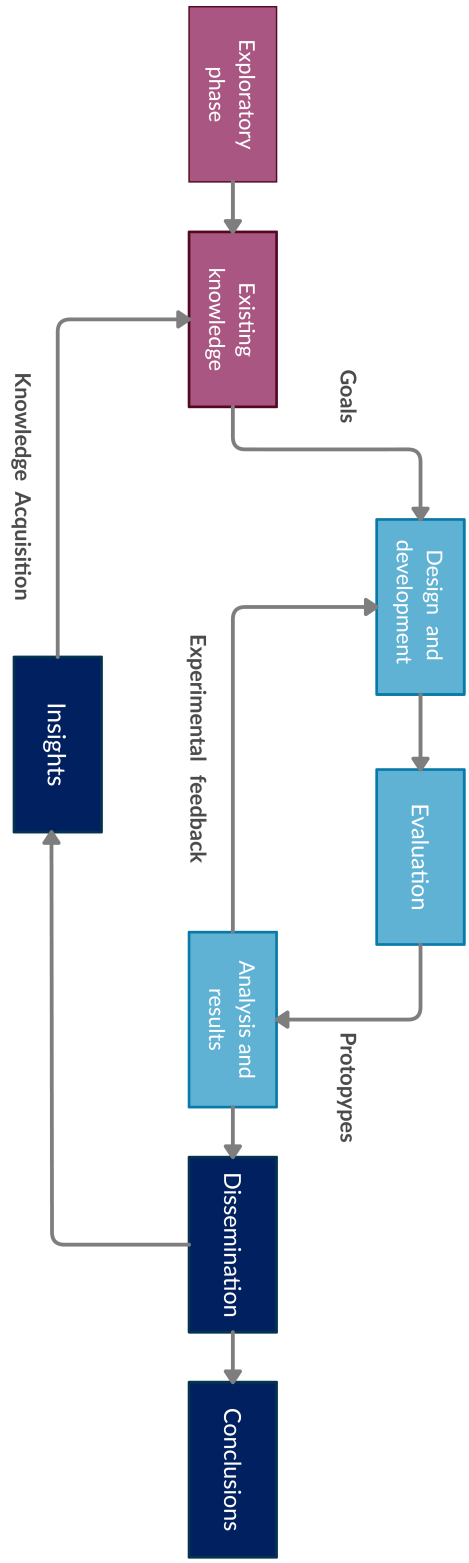

Figure 1.2: Schematic representation of the research methodology followed in this dissertation.

1.4 Main Contributions

The following scientific contributions can be found in this dissertation:

- An analysis of the implications of privacy concerns and data control when deploying IoT systems in smart workplaces. First, the related State of the Art is reviewed in Chapter 2, with a special focus on health-promoting solutions. Then, aiming to complement the conclusions drawn from the literature, the results obtained from a conducted study, consisting of an online survey completed by 524 respondents, are included in Chapter 3.

- The design and implementation of a novel strategy to integrate and optimize inference and training stages of HAR applications at the local Edge to promote data privacy. This approach optimizes the cost–accuracy trade-off of ML techniques to create stand-alone Edge solutions. The analyzed factors, system structure and the empirical evaluation of the proposed approach are shown in Chapter 4.

- A method for improving the performance of complex classification tasks using model ensemble techniques in resource-constrained devices. This method proposes a discriminative model cascade that adapts the resources to be used according to the task's difficulty to simplify the complexity of supervised learning algorithms and bring more complicated classification systems to the Edge of the network. The description of the strategy, its application and its empirical evaluation are shown in Chapter 5.

- A novel interactive approach that integrates hybrid intelligence at the Edge and allows users to determine their participation level. The previous cascade proposal is adapted to combine performance optimization strategies with IML techniques to modulate the role of users and the customization ability of the system. The description of this implementation, as well as a quantitative evaluation of the implemented use case, are presented in Chapter 6.

- An exploration of the potential implication of hybrid intelligence approaches in the workplace and the users' requirements to be involved in collaborative scenarios. This is assessed through an evaluation consisting of semi-structured interviews with 12 participants. The procedure, main findings, and insights obtained from the qualitative evaluation of the implemented use case are also included in Chapter 6.

- A dataset containing 1000 instances of labeled data for classifying office employees' hydration patterns. The Office Hydration Monitoring (OHM) dataset includes 25 variations of different interactions that could be made with liquid containers grouped into three classes. This dataset is described in Chapter 4 and can be found at (Gómez-Carmona and Casado-Mansilla, 2021).

1.5 Thesis Outline

The thesis is structured in seven chapters:

The current section, Chapter 1, introduces this dissertation, including the context and motivation, the hypothesis, objectives and contributions, as well as the research methodology followed to achieve them.

Chapter 2 presents an analysis of the State of the Art in those concepts that compose the theoretical foundation of this dissertation. Those factors are categorized into: (i) IoT technologies for work environments (with a particular focus on health and wellness promotion), (ii) Edge Intelligence, including the EML techniques and ensemble learning techniques for optimization, and (iii) Hybrid Intelligence and HiTL approaches for interactive human-machine collaboration.

Chapter 3 describes the conducted evaluation to understand the privacy and control requirements of office environments when deploying IoT solutions. The main obtained findings, together with the analysis of the State of the Art included in Chapter 2, provide the motivation for the proposal of this dissertation.

Chapter 4 presents the holistic optimization approach that aims to enhance all the stages of the classification pipeline and integrate both inference and training at the Edge. An extensive evaluation is carried out using a fluid intake recognition use case and a public dataset containing such activity, providing the necessary background to understand the potential of this strategy. Then, to validate the generalization of the system, various experiments with two additional datasets are performed to evaluate the presented strategies.

Chapter 5 describes the ensemble learning-based model cascade approach to enhance the integration of complex classification models in constrained settings. The suitability of this approach is widely validated with five datasets and several experiments.

Chapter 6 presents an adaptation of the model cascade approach for interactive scenarios in which user involvement and participation requirements are addressed. This chapter firstly starts by providing a quantitative evaluation of the effect of the interactive strategy proposed. For that, the same fluid intake recognition use-case is employed, using this time a novel dataset created for the purpose of this work, with the idea of having an activity dataset that resembles real-world scenarios and incorporates high variance between users and classes. Secondly, a qualitative conducted evaluation is presented. This study provides insights regarding participants opinions towards the proposed collaborative hybrid scenario.

Chapter 7 summarizes the main findings and conclusions of this dissertation and outlines the future research lines related to it.

The science of today is the technology of tomorrow.

Edward Teller

CHAPTER 2

Related work

THE wide scope of this dissertation makes it necessary to introduce concepts from several research fields to provide the needed background for the sake of understanding the rest of the work. For this reason, this chapter will present a compilation of the existing literature that, in one way or another, address similar aspects that this dissertation revolves around. Under the lenses of each of their disciplines, all the reviewed pieces of research have, either as a means or as an end, the goal of posing a more user-centric view of IoT. This goal also includes how integrating the processing of the data around the user can improve the perception of intelligent environments.

Following the Venn diagram shown in Figure 1.1, this section will be articulated around three main concepts; (i) the integration of IoT solutions and tracking technologies in intelligent workplaces, with a particular focus on those works related to wellness promotion in office environments, (ii) Edge Computing and Edge Intelligence as a tool for promoting *privacy-by-design* in AI applications, and (iii) the role of the user in those ntelligent environments through HiTL approaches.

Therefore, Section 2.1 of this chapter will cover the integration of IoT technologies in office environments and its main barriers and concerns. Then, Section 2.2 will review current efforts in Embedded ML and Section 2.3 will introduce the IML and Hybrid Intelligence concept. Finally, we make a summary of the chapter in 2.4.

2.1 The IoT concept and the idea of smart workplaces

Several fields and domains ranging from healthcare (Baker et al., 2017) to Industry 4.0 (Ghobakhloo, 2020), including transportation (Hars, 2015), education (Collins and Halverson, 2018), or business (Berman, 2012), are exploiting the never-ending advances of the IoT to create smart spaces. People and technology are called to collaborate in those spaces to enhance user experience, habits, and behaviors. Among them, the workplace is a particular environment to put the focus on in this digital transformation. In fact, the effect of the routines conducted in those spaces on individuals' physical or physiological health and wellness is a renowned problem in which technology can be part of the solution (Waddell and Burton, 2006). Thereby, work environments offer an opportunity to identify unhealthy behaviors associated with this space and correct these practices (Sparks et al., 2001). For this reason, these spaces are gaining increasing significance on overall health promotion (WHO, 2010).

As a consequence, the academia is more concerned than ever about the importance of bringing well-being into the workplace, making the promotion of wellness an objective tackled from an interdisciplinary point of view (Newton, 2012). The role of the IoT in workers' well-being, as is understood in this section, represents how technology can contribute to enhancing peoples' health perception (which is a predictor of health promotion). Under this context, a workplace augmented with IoT can detect and classify unhealthy habits like bad postures or inadequate hydration and notify those harmful practices to end-users to eventually enhance health (Nguyen et al., 2017) and wellness (Cahill et al., 2019). Therefore, those spaces can contribute to cover the lack of awareness of individuals about the impact of the habits on their life (Young et al., 2015).

Several attempts have been made to design enhanced smart workplaces by adopting the IoT paradigm. From occupational risk assessment and ensuring safety in the workplace (Maman et al., 2017), different solutions are proposed to reach large audiences. Promoting more active behaviours in those spaces, as sedentary behaviors stands out as one of the leading health-related problems of moderns society that tend to be less active. In office environments, its influence becomes more significant as the long periods of inactivity and sitting times are increased (Buckley et al., 2015). Thus, the attempts to design and implement interventions in such spaces have evolved a more holistic approach that includes workers' lifestyle changes within such spaces (Organization et al., 2017).

However, the literature also illustrates the difficulty of integrating those technologies in the work environment, which often derives from users' reluctance and aversion to the proposed interventions, limiting their actual scope and outcomes. In those spaces, the acceptance of the IoT technology is conditioned by the perceived value and risks of the proposed interventions (Nappi and de Campos Ribeiro, 2020). In general, work environments are especially challenging scenarios in which additional barriers regarding the privacy (Schall Jr et al., 2018) and the ethical concerns of data collection (Bowen et al., 2017) must be considered.

In this section, we review those technologies designed to correct unhealthy behaviors associated with work environments (e.g. inactivity or musculoskeletal disorders due to inappropriate sitting postures at the office) while raising the wellness concept and promoting changes that persist over time. In particular, this collection of evidence focuses on office environments, where its inherent sedentary nature is directly related to a decrease in the workers' health (De Croon et al., 2005). Then, we explore the human factor behind these interventions and how people and the devices that populate smart workplaces can cooperate towards bringing health awareness to the workplace by increasing their perception of technology, which is one of the main pillars of this dissertation.

2.1.1 Health promotion in office environments

The review of the consulted literature shows that assessing occupational sedentary behavior in the workplace stands out as one of the most addressed health-related concerns (Owen et al., 2010). Its deleterious impacts on health include cardio-respiratory, metabolic or cardio-metabolic risks (Tremblay et al., 2010). When considering the office environment, its influence becomes bigger, driven by the health outcomes of long periods of inactivity and sitting times (Buckley et al., 2015). Obesity corresponds to another identified problem that also correlates to decremented productivity (Shrestha et al., 2016). At the same time, ergonomic-related problems may result from long working hours without physical inactivity. Besides, harmful postures can contribute to musculoskeletal disorders (Waddell and Burton, 2001). Other associated risks are carpal tunnel syndrome due to continuous typing at work (Franklin and Gray, 2017) or the development of computer vision syndrome because of the exponential screen time exposure (Randolph, 2017).

Beyond understanding the potential risks of occupational health in those spaces, a key factor when designing and implementing programs to promote new habits in the workplace is to study specific methods to identify the main problems that expose the worker to those risks. Then, to carry out useful strategies to solve them (Ilvesmaki, 2007). In this regard, as has been addressed, Transforming the quality of the workplace experience implies monitoring which habits need to be changed and providing information about the consequences of these habits. IoT-based solutions allow us to physically or digitally interact with our surroundings to obtain data that can be transformed into information and, in the end, knowledge about the daily routines of the workers. Based on this knowledge, context-aware guidance can be provided to influence the users and change their behaviours. Thus, IoT solutions can be considered appropriate drivers to promote wellness in the workplace. In particular, its purpose is to understand how technology can enhance peoples' health perception and self-awareness, studying how technology-based solutions such as wearable devices can be considered appropriate means for preventing indirect risks associated with these spaces.

To this end, several efforts have been made to pursue the objective of developing IoT applications for the workplace. Some scholars have approached wellness interventions through digital technologies. For this reason, a wide range of solutions been proposed for reducing sedentary behaviours (Huang et al., 2019), as well as to increase energy expenditure and promote more active working periods (Pedersen et al., 2014; Howarth et al., 2018). In this direction, Taylor et al. (2018b) reviewed the existing literature addressing interventions designed to reduce sitting time and analyzed the importance of the organizational culture. Positive trends when applying those interventions were found. These results are in line with the ones presented by Stephenson et al. (2017), which concluded that interventions using a computer, mobile, and wearable technologies could be useful in reducing sedentary behaviours. Similarly, the PEROSH initiative (Holtermann et al., 2017) studied how wearable devices could be part of wellness promotion interventions. It elaborated a decision support framework for selecting useful sensors and proper data collection strategies for avoiding sedentary behaviours, neglecting data privacy issues. Other works modelled physical fatigue in the workplace using wearable sensors (Maman et al., 2017), dealt with worker's cognitive load in interruption management for occupational risk assessment (Schaule et al., 2018), proposed a non-intrusive monitoring system to avoid lower back injuries (Zhao et al., 2017), or designed an ecosystem of wearable elements for occupational risk assessment (Bernal et al., 2017).

HealthOffice presented an automated mood recognition system to benefit employees' health and productivity through wearable technology (Zenonos et al., 2016). It consists of a framework and a mobile app that register vital signals such as the pulse and heart rate or the skin temperature through a prototype wearable sensor and correlates them with the mood and stress levels of the user. In fact, symptoms of stress also constitute an essential challenge for enhancing worker psychological health (Murphy, 1996), and some initiatives have presented systems for measuring the physical symptoms reflecting the stress level with wearable devices, such as the Stess@Work system (Bakker et al., 2012) or the work presented by Han et al. (2017). In parallel, commercial solutions such as Comfy[1] are committed to providing a virtual link between the digital workplace and the physical environment by means of a Cloud-based platform able to collect users data. These data might be used as well to monitor user activity (Cheng et al., 2013) or even suggest the most appropriate time intervals to take a break considering the user's focus state (Kaur et al., 2020). Furthermore, sensors data has also been used for improving the user's sitting position and promote better ergonomics using a smart chair (Roossien et al., 2017; Hu et al., 2020).

In the same way, Jimenez and Bregenzer (2018) presented some guidelines for workplace health promotion using electronic and mobile health tools to provide easier administration for campaign proposers while considering data privacy from a technical and psychological point of view. However, no specific IoT architectures are proposed to conduct these processes. Additionally, some efforts have been made in the direction of unifying the methods for measuring the effects of these interventions in different domains, such as the association between health interventions and productivity (Maheronnaghsh et al., 2018) while novel technologies such as 5G in IoT domains have been devised to boost comfort (Ahmed et al., 2016) and safety (Podgorski et al., 2017) in working environments.

In (Lennefer et al., 2020), authors conducted an experiment to evaluate how an increased activity level could prevent burnout and improve health perception. By means of activity trackers and a behavioural approach, the conducted evaluation showed a significant increase in health perception as well as a decrease in the body mass index. However, the intervention was not effective in reducing burnout. These results are in line with the ones obtained by Glance et al. (2016), in which, thanks to the proposed intervention, participants increased their activity level and maintained at least 10,000 steps a day during the study. Besides, it resulted in measurable changes in lipid profile, blood glucose, renal function, blood pressure or weight, which were measured using a health and wellbeing assessment. In Gorm and Shklovski (2016), a nation-wide workplace step-counting campaign was presented.

[1] https://www.comfyapp.com/

Despite the popularity of the campaign and the initial excitement of the participants, their data bear no evidence of healthier practices continuing beyond the three weeks of the campaign. Thus, they concluded that the promoted campaign was successful as a form of health promotion rather than a behavior change effort.

This last work is an example of the difficulties of making strong validations of the direct influence of technology on workers' health for the long term. The work presented by Malik et al. (2014) evaluated the impact of health interventions on activity levels. Contrary to some of the previously reviewed works, they concluded that, although some evidence of efficacy was found, an equal volume of successful and non-successful studies was found, and overall results were inconclusive. Nevertheless, considering the positive outcomes, they found a strong argument for pursuing research efforts to design physical activity interventions for promoting wellness in workplaces. Similarly, Yu et al. (2017) evaluated its impact in terms of health outcomes (i.e. body mass index -BMI-, total cholesterol and blood pressure). In this study, the authors detected a slight decrease in BMI associated with individuals changing their exercise behaviors, but no significant changes in cholesterol and blood pressure levels. Other work reviewed systematically those interventions designed to reduce sitting time among office workers and the role of organizational culture (Taylor et al., 2018b). Recurrently, not conclusive data could be extracted. This shows a tendency in the body of the literature that tends to be careful overstating conclusions from inconclusive data or short-term studies.

In this regard, the lack of adherence of the target audience has been identified as a key element for the low-to-moderate success of those interventions. This also influences their potential long-term outcomes (Howarth et al., 2018). Furthermore, according to Jacobs et al. (2019), the adherence and outcomes of technology deployment largely depend on employee acceptance of the technology. Therefore, understanding the factors and barriers behind its adoption is essential to designing engaging technologies to be integrated into the smart workplace. In the following subsection, we will study in detail those barriers found in the literature.

2.1.2 Barriers and challenges for the adoption of IoT technologies in the workplace

The previous section illustrates how the advent of the IoT and the massive growth of devices connected to the internet has the potential to reshape the workplace concept. Integrating monitoring technologies in these spaces involves ensuring employees' perception and acceptance of technology. However, this integration often derives into users' reluctance and aversion to technology, which hinders its adoption. In fact, the analysis of the effectiveness of workplace health promotion programs shows that interventions fail when heterogeneous and large audiences are targeted (Rongen et al., 2013). In this line, insufficient benefit or low perceived value have also been pointed as a reason for the failure of such interventions (Person et al., 2010). According to Street and Lacey (2018), the absence of positive enough results deviates from the definition and design of these interventions. This definition tends to fall into a "one-fits-all" problem in which the different perceptions of these initiatives and the perspectives of employees are not fully considered. In addition, understanding the potential commitment of every user to be part of these interventions is also crucial factors when workday duties involve the total daily routine (Helfrich et al., 2018).

Beyond that, another reason for this lack of adherence is the reluctance of employees to be monitored in smart workplaces. In fact, the workplace needs to overcome additional barriers regarding privacy concerns of the collected data (Carpenter et al., 2018). For this reason, various authors have studied the main determining factors that could predict the employee acceptance of technology in the workplace, with a particular focus on privacy and trust concerns. The review presented by Nappi and de Campos Ribeiro (2020) evaluated the perceived risks and benefits associated with the integration of IoT technology in the workplace and identified primordial aspects that companies must consider to improve employees' attitudes toward it. They concluded that the awareness of the benefits and risks are key for its adoption. Those risks and benefits are included in Table 2.1.

Benefits	Risks
Real-time data with speedy information retrieval	Employee discomfort in sharing his/her data, threats to employee private
Access to a new range of employees' data	Employee limited control of the information shared with the company and coworkers
Can help to improve employees' health	Employee biometric surveillance
Allows the identification of unexpected associations (for example, face-to-face interaction and employee productivity)	The economic value of the data can motivate the reindividualization of the employee (IoT) data

Table 2.1: Benefits and the risks associated with the use of IoT technology in office environments. From the work of Nappi and de Campos Ribeiro (2020).

As a consequence, concerns about user privacy are growing at the same time the options for collecting, processing, distributing and using personal information are increasing. In their work, Teebken and Hess (2021) adapted the general dimensions of information privacy to the workplace context. The fist concept is Errors, which represents user's concern that their stored personal information contains deliberate or accidental errors. The rest of the dimensions have to do with the employee's concern about the amount of collected data (Collection), who has access to it (Unauthorized Access) or how it is used (Awareness and Secondary Use). The remaining factor, Control, is related to the ability of the user to have control over their data or even opt-out. The dimensions that represent every aspect of privacy are included in Figure 2.1.

The literature on integrating technology in smart workplaces has analyzed the importance of these dimensions in its success, either directly or indirectly. Masson et al. (2016) conducted a study to understand users perceptions about the integration of activity trackers devices in the workplaces. According to their findings, all the participants have any privacy concern about sharing their data. In fact, most of the users had hesitations giving researchers access to their data, even though it was only an experiment and not a real scenario where their employers would collect these data. In (Chung et al., 2017), authors conducted a study to understand the employees' perception of existing

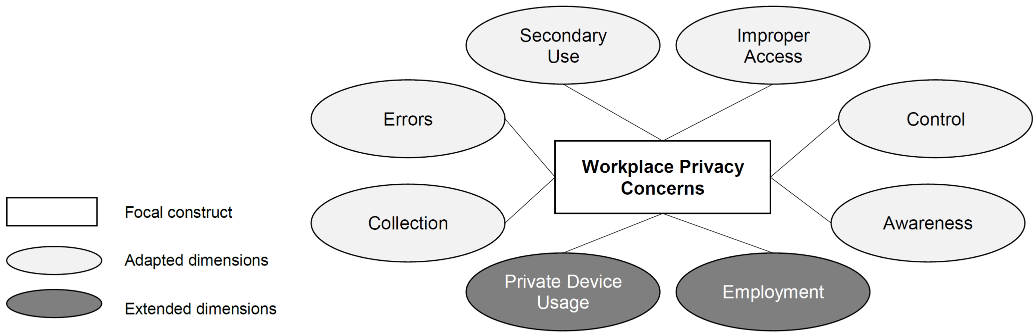

Figure 2.1: Dimensions of privacy concerns. From the work of Teebken and Hess (2021).

wellness programs. Although, contrary to the previous work, they did not find major problems with privacy, they recognize the importance it can have. They also stress the importance of transparency about the policies used to manage the data collected. Furthermore, they also addressed an additional concern: while many companies are transparent about data collection, there is an issue with the fact that third-party platforms are used for this.

This probes that, in general, users are hesitant to share their own data. This data is seen as a threat when it can give a negative image of the user and give contextualized information that can invade their privacy or refer to their health issues. Furthermore, fitness and health records can be combined to present a general medical profile of the employee that could be disclosed to the employer, which causes rejection by end-users (Bhave et al., 2020). This shows that people do not really trust their companies measuring what they are doing (i.e. schedules or work performance) which creates a sense of surveillance in office environments that contrasts with the ethical concerns of personal data collection (Ajunwa et al., 2017; Bowen et al., 2017). Moreover, beyond privacy concerns, there are doubts about the accessibility of data by users, since many times, the data generated by IoT infrastructures do not belong to them. As Moore and Robinson (2016) argue, without a transparent use of the data, the benefit of quantification falls on employers rather than employees, as they are the ones in control of both the data and the devices. Therefore, if the system does not give the user control of their own data, people are usually confused about how their data are used by these systems (Ackerman, 2000). Hence, greater transparency about the use of employees' data is the key to their acceptance of IoT technology in office settings (Khakurel et al., 2018). Mechanisms and privacy laws such as the European General Data Protection Regulation (EU GDPR) help to ensure that procedures. However, the implementation of technology in those spaces may not be transparent and straightforward enough with employees about the purpose of collecting, using, and storing data.

The analyzed studies show that employee awareness of the benefits and risks associated with IoT technology is only partial. The use of personal data within an organizational context raises this kind of concern. It highlights the importance of addressing the issues of acceptance and ethics related to privacy, given their influence on the refusal to share data and the lack of understanding of the data collection process. While privacy concerns have been clearly identified as one of the main factors that affect employees' beliefs towards technology, no study raises questions about users opinions regarding data ownership. According to Nappi and de Campos Ribeiro (2020), ensuring that all employees can control the information they share is a crucial element for adopting IoT technologies in the workplace. For this reason, through this dissertation, we stress the relevance of the concept of data control and ownership as a mechanism to create more confident and trustworthy spaces where the acceptance of the technology is promoted. Thus, we cope with the lack of studies addressing how data control and ownership may affect the perception of the technology. For that, in the first part of this dissertation (Chapter 3), we will address not only the concept of privacy related to the possible use of the captured data but also how their perception would change if the control of this information belonged solely to the owner of this data, that is, the user.

2.2 Edge Intelligence: bringing processing capabilities closer to the user

The idea of combining Edge Computing and AI is an emerging research area, which is particularly interesting in those cases where keeping data close to the source (e.g., the user) is of vital importance to ensure their confidence in the system. For this reason, it can be considered a suitable mechanism to create a sense of privacy and trust towards IoT technologies in the workplace. This concept aims to make intelligent applications less dependent on the centralized Cloud infrastructures and bring them closer to the users (Xu et al., 2020). Embedded ML techniques, also coined as TinyML, make it possible for low-end devices to perform ML inference techniques directly on the Edge or attached to sensors (Soro, 2021). In fact, ML frameworks are being adapted to this new Edge scenario. For instance, Tensorflow (Abadi et al., 2016) has incorporated a variant called Tensorflow Lite that improves on-device inference and provides tools to include pre-trained AI models in resource-constrained devices (David et al., 2020). Other available resources correspond to development platforms such as Edge Impulse [1], that provide several features and tools to facilitate its integration of ML on Edge device.

However, scaling those applications to accommodate greater processing requirements and create more proficient solutions efficiently is now emerging as an open challenge in the IoT domain. In this regard, a suitable integration of those computational complex techniques in such limiting scenarios implies identifying how the available resources could be used more effectively. This goes through finding the optimal balance between the cost and the accuracy (Dhar et al., 2019). In (Jensen et al., 2016), authors referred to this characterization as the accuracy–cost conflict, in which embedded classification systems must not only consider the classification rate but also be able to balance the cost of each classifier. In this line, what is needed is the development of techniques that optimize the execution of the ML tasks. In the following, we will address some of the major advances for bringing computation to a

[1] `https://www.edgeimpulse.com/`

local stage while improving the performance of classification tasks. Then, we will include some works that have approached ensemble learning techniques from an optimization perspective.

2.2.1 **Optimizing Embedded Machine Learning**

Early research studies for bringing computation to the Edge of the network have mainly focused on verifying the feasibility of Edge devices to run ML applications, identifying execution bottlenecks and providing some initial insights about how models could be scaled down (Lane et al., 2015). Challenged by the constraints on memory or computation, Banbury et al. (2020) highlighted the importance of creating an appropriate frame of reference for benchmarking and analyzing the capabilities of devices that are integrated into the Edge. In this line, other investigations have approached embedded intelligence evaluating the suitability of devices such as the Raspberry Pi platform to implement ML inference tasks on the Edge (Yazici et al., 2018). Similarly, Desraches et al. (2018) considered the simplification of the classification problems to shed light on the feasibility of embedding simple yet accurate probabilistic models on constrained devices.

When it comes to reducing the gap between the computation complexity and the available resource capacity, the existing literature follows two main approaches: (i) implementing more efficient and lightweight representations of the learning methods and (ii) optimizing different parts of the classification pipeline to lower the amount of processed data. This latter reduces the complexity of the classification systems and the final models.

The first approach corresponds to adapting classifiers to the capabilities of the target hardware or creating lightweight implementations of the ML algorithms (Neto et al., 2018). For that, Banbury et al. (2020) analyzed the importance of systematically comparing, evaluating and improving the performance of the embedded hardware for ML workloads. Whilst presented efforts are focus on inference at the Edge, it remarks the importance of comparison of the available hardware and software tools for continuous progress and stability of the research fields. The different learning methods differ from each other in terms of computation requirements, memory size, or accuracy properties. For this reason, Alam et al. (2016) analyzed the applicability of some of the most common ML models for the IoT. They concluded that some traditional ML algorithms, such as Naive Bayes (NB), Support Vector Machines (SVM), Linear Regression, and Decision Trees (DT), usually generate a relatively low footprint over resources. For this reason, several approaches have been proposed for those classifiers; for instance, optimizing SVM (Haigh et al., 2015), k-Nearest Neighbor (KNN) (Gupta et al., 2017) and implementing new DT-based algorithms for efficient prediction on IoT devices (Kumar et al., 2017). In the same line, the EmbML Tool (da Silva et al., 2019), implements a pipeline for converting classifiers from the most popular software packages to C++ to embedded them into low-cost systems. Libraries as Sklearn-porter (Morawiec, 2016), MicroML [1] or Emlearn (Nordby, 2019) also provide tools for generating C code and bring ML algorithms to embedded platforms.

All the previous approaches have something in common: they propose the use of lightweight, non-neural network based techniques. Other proposals follow a different approach focusing on DL techniques, and propose to simplify the complexity of classification systems through limiting the size of the input and the number of layers of those models (Liu et al., 2020). Various works have also dealt with the optimization of DL models and the ability to deploy of ANN on tiny devices (Lane et al., 2015; Andrade et al., 2018). In this regard, Scheidegger et al. (2019) analyzed how ANN can run efficiently on IoT devices and proposed an automatic way to design DL models that satisfy user-defined constraints. Moreover, Wang et al. (2019) optimized the performance of Multilayer Perceptron (MLP) methods on low power embedded devices.

The second approach comprises diverse optimization strategies that, in general terms, seek to save further resources on the different stages of the transformation of data into knowledge (Shafique et al., 2018; Fafoutis et al., 2018). To do that, those strategies address the resource overhead by reducing the data input size or the final model complexity, aiming to achieve a similar classification accuracy. In this respect, feature selection can reduce the cost

[1] https://github.com/eloquentarduino/micromlgen

associated with the extraction of features set (Cilla et al., 2009). However, since a smaller number of features could lead to lower classification accuracy, it is essential to profile this trade-off. This has already been addressed for HAR systems in embedded and wearable devices (Elsts et al., 2018; Zalewski et al., 2020). Similarly, Elsts et al. (2020) proposed an energy-efficient framework for designing activity recognition applications for wearable platforms. Since extracting and sending the features is less energy expensive than sending the raw data, their approach was based on finding groups of features with good energy-accuracy trade-offs to send them to external platforms in charge of the classification step.

The sampling frequency is another aspect that deserves attention to improve the final performance of Edge intelligent systems (Khan et al., 2016). Together with saving resources at the Edge, decreasing the sampling frequency can also be potentially interesting to reduce end-device energy consumption as well as communication bandwidth. A work from Kraemer et al. (2019) evaluated the effects of the sampling intervals on both energy consumption and accuracy in order to save energy and costs on the network infrastructure. Similarly, in (Yan et al., 2012) an adaptive approach for mobile devices was proposed and in (Qi et al., 2013) there was presented a novel algorithm to determine the optimal sampling rates according to the requirements of a multi-activity classification and a single activity event detection. Zalewski et al. (2020) presented a framework for increasing the device energy lifetime by reducing the duty-cycle of the radio and the gyroscope of wearable devices. Their argumentation revolves around accepting a small reduction of short-term accuracy to save energy to sustain higher performance in the long run. In the same way, the number of sensors in a multi-sensor setting can be optimized to improve the trade-off between them (and/or signal components) and the classification accuracy. In this line, Kang et al. (2009) proposed a dynamic selection of a small subset of sensors. Finally, Gordon et al. (2012) evaluated the importance of every sensor and compared the trade-off between the number of sensors and the detection accuracy.

DL approaches can also benefit for an optimization process. This effort is geared towards improving the memory impact and computational workload of over-parameterized DL models (Shafique et al., 2018). This optimization is achieved by means of changing the data representation strategy to reduce the number of bits that are used to represent a number (Sun et al., 2020; Ghamari et al., 2021), or removing the least essential parameters of the final model (Lane et al., 2017). In addition, model compression (Kim et al., 2015) and pruning techniques (Molchanov et al., 2016) have proved to be suitable tools to simplify ANN topologies with a small impact on the accuracy of the classification.

When it comes to addressing privacy concerns as well as giving the user total control of the captured data, integrating the training stage in the Edge can answer the need for privacy of new smart spaces while increasing users' perceived level of trust in the system (Yu et al., 2018). Despite the promising advances reviewed through this section, the research on this topic still focuses on the inference at the Edge, relying on centralized servers or distributed approaches for training the models (Eshratifar et al., 2019). In fact, model training is often the most time-consuming part of the model development process. Thus, training stages at the Edge are limited by the performance of embedded devices. This limitation makes this specific part of the classification pipeline a key area of focus on for bringing all the phases of the ML process to the local stage. Initial attempts to evaluate the suitability of embedded platforms to intragrate training tasks have been conducted (Liu et al., 2019; Chandakkar et al., 2017). Besides, other approaches have presented a novel solution for on-device training by leveraging incremental online learning (Bhat et al., 2018; Ren et al., 2021). However, on-device training and optimizing its performance continues to be understudied in the body of the literature. Hence, with this dissertation, we aim to contribute to this open challenge by improving not only ML inference at the Edge but also the training efficiency. In contrast to most of the reviewed works, we present a holistic strategy, i.e., the implemented system is globally analyzed, taking into account the contribution of each part of the ML solution towards the final pipeline.

2.2.2 Ensemble learning techniques for optimization

Ensemble learning techniques have been historically well covered in ML research. Approaches like Bagging (Breiman, 1996), Boosting (Schapire, 1990) or Stacking (Wolpert, 1992) are well-known methods that have traditionally been used to improve the general performance of classification systems by combining several base models in a final predictive meta-model (Zhang and Ma, 2012). Cascading is another particular example of ensemble learning. Firstly introduced by Viola and Jones (2001), it consists of the concatenation of several classifiers to minimize the generalization error rate. HAR applications (Xu et al., 2019) or the classification of rare events in face detection (Wu et al., 2003) are examples of domains in which cascade ensemble learning can be applied.

Even though these strategies improve the predictive performance of a single model, they also tend to increase the computational complexity of training the ensemble and predicting new instances (Sagi and Rokach, 2018). An interesting alternative is to use the combination of models of cascading approaches in a cost-sensitive manner. In this regard, the complexity of the ensemble methods is faced by implementing several increasingly complex classifiers that divide the computation into different stages and classify the input data with the stage that fits the most to its difficulty (Rokicki and Drozda, 2017). This novel idea was presented by Kaynak and Alpaydin (2000), who proposed a multistage cascading scheme that relies on the concatenation of several small classifiers. The objective of this methodology was to reduce computational cost yet not losing considerably in terms of accuracy. The authors assessed their proposal in a two-stage approach. The former classifier was based on single-layer and MLP models that are considered lighter methods for prediction. The latter classifier was based on KNN. To improve the classification results, successive models were trained with a self-modified training set, prioritising non covered patterns or data that the previous models overlooked. Additionally, confidence thresholds to define if the classification was acceptable were defined based on previous probabilities rates.

In (Silva et al., 2017), authors proposed a classification approach also based on a two-stage combination of monolithic and ensemble classifiers. The rationale behind their proposal was to deal with the majority of unclassified instances using the monolithic classifier leaving the most challenging instances for the ensemble classifier systems. The latter classifier was based on pool generation methods. However, in that case the complexity of every stage of the system is determined by the inherent complexity of the generation methods. In particular, they applied KNN and SVM methods for the initial model and different ensemble techniques (i.e., Bagging, Boosting, and Random Subspace Selection) for the final model. Nevertheless, since the authors compared their approach against other multiple classifiers systems, it remains an open question whether such a method outperforms in optimization a single monolithic classifier. In addition, other approaches have proposed additional simplification techniques by understanding the cascading techniques as a tree topology case (Xu et al., 2014; Chen et al., 2012). In this case, different topologies can be combined to construct a tree of classifiers that determine how many features are needed to predict new unseen data.

The reviewed methods for optimizing the classification of data are often done at the cost of losing accuracy. This highlights the need for additional adjustments to compensate for this loss. As introduced above, the work from Kaynak and Alpaydin (2000) coped with accuracy loss by specifically tuning each of the successive models and thresholds to better match with previously uncovered patterns. Other authors usually consider the influence of additional/side factors, such as the rejection margin, and fit the models accordingly (Fumera et al., 2000). In this case, to keep the loss in accuracy to a minimum, it is important to analyze the reasons for accepting or rejecting the instances classified by each of the models that comprise the cascade. In (Oliveira et al., 2005) authors studied the optimum class-related reject threshold to provide a better error-rejection in cascading classifiers. The proposed methodology proved to improve the performance of a system based on a set of MLP models for handwritten digits recognition systems. In (Zhang et al., 2007) the recognition reliability was increased by using a double-check mechanism to weight and verify the Artificial Neural Network (ANN) confidence values and to correct the errors. Nevertheless, the computational cost was not considered in their work.

Regarding ANN, Wang et al. (2020a) analyzed how cascade models could outperform in computational efficiency other ensemble-based architectures. Similarly, adapting the size and the number of layers of those models has been a subject of deep research. One approach consists in defining a combination of big/little models (Park et al., 2015; Tian et al., 2019), while another approach proposed creating a cascade based on the depth of the layers in ANN techniques (Leroux et al., 2017). Similarly, Taylor et al. (2018a) faced the intrinsic complexity of the DL models by presenting an adaptive scheme to determine which model to use for a given input. Considering the desired accuracy and inference time, their approach employed a KNN predictive model to select those pre-trained models that best suited a given input and the optimization constraint. In parallel, other approaches leave open the possibility to adapt their work in an optimized way (Lefakis and Fleuret, 2010).

Going in another direction with respect to the body of knowledge reviewed, this dissertation targets IoT architectures and Edge Computing approaches. Therefore, it aims to shed light on the actual optimization opportunities that a discriminative cascade of models has in constrained settings. Hence, to provide an empirical demonstration of the potential computational cost saving in resource-constrained Edge devices. For that, this dissertation introduces a model cascade approach for optimizing the performance of complex classification application, Here, the complexity of the cascade approach is determined by the number of characteristics of the input data. The accuracy loss is faced by setting up confidence thresholds that are more restrictive for those potentially weaker classifiers. This way, we evaluate the influence of the simplicity of the models in terms of feature computation since the computation time required for feature extraction can increase drastically in those applications that heavily rely on the characteristics obtained from the signal (Gómez-Carmona et al., 2019). In the following of this dissertation, we will present our approach in detail and examine its suitability for a HAR context.

2.3 Human-in-the-Loop Machine Learning and Hybrid Intelligence

The ML approaches described in the previous section have one thing in common: all of them are constituted by pre-trained models using generic data. Therefore, none of them takes advantage of human knowledge to improve their capabilities, nor are they adapted to the particularities of end-user. In contrast, IML makes it possible for end-users to interact with the learning process to tune the system according to their preferences, or even to feed it with more data (Bernardo et al., 2017). This corresponds to a HiTL approach in which enhanced interactions between end-users and the learning process are created. In essence, HiTL embraces the idea of Hybrid Intelligence, defined as an intelligent model that requires human interaction (Zheng et al., 2017).

Users' involvement with ML can take many roles in classification tasks, including data curation, learning algorithm tuning, and labeling. The latter is one of the most common approaches, where the final user provides the learning system with sources of truth to personalize it. Additionally, users can participate through their feedback on the definition and fitting of the model. Thus, machine intelligence can be adaptively adjusted in terms of human goals. In supervised approaches, this means that users can create personalized HAR systems by providing new annotation data (Miu et al., 2015) or serving as an oracle in Active Learning approaches (Tegen et al., 2020a). In semi-supervised ones, the model can also be fed with raw time-series data (Cardoso and Moreira, 2016). Special attention must also be given to those approaches that integrate the human in the learning process as they benefit the most from the domain knowledge. This is the case of healthcare in which the integrated expertise in IML can be suitable to solve problems that deal with complex data and/or rare events (Holzinger, 2016).

Although IML is a promising candidate for further improving the knowledge discovery process, a major problem that academia needs to address in future research is what questions should be posed to humans and how should they be asked (Robert et al., 2016). That is, humans are involved in providing data, but previous ML research mostly neglects the question of how humans can provide this data and how they cope with an inaccurate model. For end-users, algorithms are similar to a black box, and there is a lack of research analyzing the challenges and opportunities of a human-algorithm interaction. Thus, optimal integration of human and computer input is of great importance to obtain the best possible result. For this reason, in (Porter et al., 2013) authors break down the IML process for human centered ML. The central idea of this definitions is to assess the needed aspects that enable experts to customize their application:

- Task decomposition: The level of coordination between humans and machines. In other words, the number of tasks in which the work is divided and that are distributed among the computer and the user.

- Training vocabulary: Interactions between humans and machines during training and the different inputs that users can provide the machine.

- Training dialogue: The interactive dialogues during training related to the number and frequency of interaction between the end-user and the learner.

This brings into attention the need for increased collaboration form a AI/ML and Human Computing Interaction (HCI) design research fields and Hybrid Intelligence approaches. In order to do so, (Inkpen et al., 2019) examined how better integrate artificial and human intelligence in the decision making, addressing the different human roles and relationships in human AI systems. Users engagement has been also addressed from the HCI perspective evaluating such engagement practices to conversational AI agents (Muralidharan et al., 2019) or defining a Human-centred thinking approach to apply IML methods that aims to support the iterative design, development, and dissemination process of learning systems (Mathewson, 2019).

Even son, bringing closer the gap between automated systems and users also has to do with understanding the level of trust users have in the decision making process of automated ML approach (Yu et al., 2018). In fact, the lack of user trust may deter the user from following these suggestions and be detrimental to the uptake of the system in intelligent systems (Eiband et al., 2021). In this regard, (Amershi et al., 2014) studied how IML can result in better user experiences and more effective learning systems. They concluded that further evaluation of novel interaction methods is needed to determine whether the user influence over the model does or not result in better systems. In fact, (McCallum and Fiebrink, 2019) evaluated how users could be supported to contribute to the model. They examined the use case of a system in which a user-driven feature representation for new ML problems needed to be assessed. The conducted evaluation demonstrates that users' have a low perception of their actions. Hence, better mechanisms are needed to examine the impact of their actions interactively. The reliability of the user providing new data to the interactive model was also studied by (Tegen et al., 2020b).

The reviewed works demonstrate the advantages of including the human in the learning process of AI algorithms. However, from a user-centric point of view, the interest of IML lies in how this process can help the human. To this end, it is essential to understand how people actually interact (and want to interact) with the ML system. For this reason, this dissertation will consider the role of the user in interactive smart workplaces. Following the IML process decomposition presented by Porter et al. (2013), this part of the dissertation will focus on the training dialogue between users and IML solutions. More particularly, proposing customizable and flexible interaction techniques to meet the requirements and expectations of end-users.

2.4 Summary and Conclusions

Throughout this chapter, the State of the Art of emerging IoT ecosystems and Hybrid Intelligence at the Edge has been analyzed. For that, a compilation of works that deal with the concepts around which this dissertation revolves has been presented. To begin with, the main concerns and barriers detected in the literature that affect the integration of technology in intelligent spaces have been addressed. In particular, the focus has been set on the work environment, where the concerns about data privacy have proven to be an impediment to technology acceptance and possible future adoption. Then, in response to these privacy and adherence requirements, mechanisms that seek to turn smart spaces into user-centric environments were subsequently analyzed. Firstly, those Edge Computing and Embedded ML approaches that aim to incorporate into local Edge devices the computational power needed to convert data into information. Secondly, it has been analyzed how the Hybrid Intelligence concept can create an interactive environment that reinforces this user-centric approach.

We have identified and extracted the requirements across physical domains and application contexts that we seek to address through this dissertation. In this regard, the State of the Art agree on highlighting privacy concerns as the main obstacle in the perception of the technology in intelligent workplaces. However, aspects such as data appropriation, who has the control over the captured data or to whom it belongs, have gone more unnoticed by scholars. Consequently, this dissertation proposes a complete local approach in which data is not outsourced at any time. In this scenario, integrating the training stage at the Edge can answer the need for privacy of new smart spaces, while increasing users' perceived level of trust in the intelligent systems. We address that by avoiding any data externalization in interactive environments. Furthermore, we provide users with the control over their information, as it is only stored in a local device under their ownership. Beyond data privacy and control, understanding the willingness of the user to interact with AI systems is also crucial for improving trust and technological perception. Therefore, based on the requirements for an enhances adaptation of the intelligent system, this work will propose an interactive strategy adaptable and flexible to the involvement level of the user.

Table 2.2 summarizes the main findings and insights obtained throughout this chapter. We also propose future research gaps and lines for the development of those research areas that are pivotal to better integrate emerging technologies in work environments. By addressing all these gaps it is expected to contribute to the State of the Art of intelligent environments by creating more confident spaces, where users are willing to embrace IoT ecosystems, with *privacy-by-design* and interactive approaches in mind.

Finding	Current Status	Gaps and future lines
Low perceived value and privacy concerns affect the adoption of emerging technologies in the office environment	IoT solutions should promote greater transparency and understanding about the use of data to increase users perception of emerging technologies in office settings	The lack of understanding about users' limited control over the information that is shared and the doubts about the ownership of the captured data remain as an open gap
The adoption of emerging technologies in the workplace increases if personal data privacy and control is ensured by keeping it close to the place where it is produced	Edge Computing approach envisages scenarios where local interactions through *privacy-by-design* might be created to contribute to the perception of technology in those spaces	Edge Computing needs to overcome the technological challenges that arise when deploying complex IoT solutions in resource-constrained devices
Bringing complex computation, including ML algorithms, to embedded devices increases the responsiveness and privacy of ML applications	Embedded ML classification techniques make it possible for resource-constrained devices to perform ML inference directly on the Edge or attached to sensors	Edge approaches still rely on centralized servers for some parts of the classification pipeline. The optimization of complex ML systems and on-device training are understudied in the literature
Successful integration of emerging technologies in such challenging scenarios is only possible if there is a better and smarter collaboration among augmented devices and people	Hybrid Intelligence and HiTL engage users to improve and personalize automatic AI-based technologies by interacting with the learning system	HiTL solutions should be flexible enough to rethink their functionality in terms of human goals, contexts, and participation preferences

Table 2.2: Summary of the insights obtained from this chapter regarding the current status, found gaps, and guidelines for the future lines of research that articulates this dissertation.

Big brother is watching you.

Orson Welles - 1984

CHAPTER

3

Data control and privacy in smart workplaces

THE review presented in the previous chapter and the obtained insights in relation to those works that deal with the integration of technology in the workplace, demonstrate that privacy concerns play a critical role in these spaces. Privacy is linked to the use that someone (i.e., a third party) does of the data captured through the deployed devices, the information that can be extracted from it, and, above all, to whom this information is visible or shared. Thus, addressing privacy concerns implies preventing data from being disclosed to anyone and understanding how this issue can affect employees' beliefs and acceptance of the technology. However, to the best of our knowledge, there is no prior evidence in the reviewed body of knowledge about users' perceptions of data ownership or the importance of who is in control of the information that is shared in workplaces.

Therefore, existing literature for the workplace context can benefit from an extended view of the privacy that elicits how employees' perception would change if the control of this information belonged solely to them as owners of the data. For this reason, this chapter aims to understand better the role of data control and ownership, data privacy, and its correlation with the perception of the value offered by IoT technologies in work environments. For that, a study based on an online questionnaire was conducted to collect information about the concepts of privacy and control among employees across European countries. A total of 524 participants completed this questionnaire. In the following, the procedure and the main findings extracted from this study will be stated.

The remaining of this chapter is structured as follows: Section 3.1 introduces the scope and objectives of this study. Section 3.2 describes its procedure and the questionnaire. Section 3.3 presents the obtained results. Finally, Section 3.4 concludes this chapter with a summary of the main obtained findings.

3.1 Objectives related to the dissertation's hypothesis

As outlined in the introduction of this chapter, it is essential to focus on threats to informational privacy and examine other related aspects when deploying technology in smart environments. Therefore, in this chapter, we aim at analyzing the implications of data control and data ownership in the workplace and how it affects the general perception of the technology by employees. For that, we created a questionnaire that presents a speculative scenario where a company designs an intervention to promote physical activity at work to promote healthier habits. To run this campaign, the fictional company provides devices (e.g. a smartwatch or smart sensors) that monitor how much time employees spend without getting up, how much they walk, how much they hydrate or how many coffees they drink per day. In addition, the fictional company also decides to replace work chairs with smart chairs that provide employees insights into their ergonomics. During the study and the design of the questionnaire and/or the speculative scenario, we deliberately avoid creating concerns about privacy issues in an explicit way to void potential bias when addressing those aspects of the system (Braunstein et al., 2011).

Hence, we did not use words such as "supervision", "surveillance" or explicitly explaining the misuse that could be made of this data or that it could be associated with people's working performance.

The following section explains how, through this scenario, we aimed to understand the implications of the deployment of IoT technologies in the workplace and what barriers to their integration are perceived by participants, always focusing on the concept of privacy.

In addition to obtaining initial insights into what users expect from the integration of technology in workplaces, this study has a further twofold objective. On the one hand, to find out what is the perceived value that users place on the information that would be obtained with IoT devices. For instance, how private they consider the information retrieved from them and how they feel about the possible use of this information by external peers, companies or third parties. On the other hand, we aimed to understand how their perception would change if the control of this information belonged only to them. To this end, two research questions have been proposed based on the needs identified in the literature regarding data control and ownership when deploying technology in those spaces.

> **Research Question 1:** *Which is the relationship between the perceived value of the data that employees generate at the workplace and which are the main concerns about its potential use?*

The main concerns about data being captured in the workplace relate to how much employees feel that IoT devices invade their privacy, how contextualised the information that can be obtained is, and how possible misinterpretation of data can affect one's image (or others' perception about him/her). Furthermore, knowing the risks that employees declare when it comes to sharing their data with third parties and how this might affect their attitude towards the technology and the enterprise is also important to mitigate those privacy concerns. Transparency about the policies used to manage the data collected and doubts about how the system uses their data can lead to a conflict of interest between the employee and the employer (Chung et al., 2017). For this reason, there is a need to understand the perceived benefits of using such data (e.g. improving their workplace habits) and the concerns that arise about data sharing, the collection of personal information, and the ability of this data to be used by third parties without the user's knowledge or consent (Nissenbaum and Patterson, 2016).

To address the influence of these factors, throughout the questionnaire that we prepared on top of the speculative scenario, users were asked to rate the features they consider more valuable in the proposed system, including options relating to its usefulness. Their attitudes towards the captured information and the entities they would be willing to share this information with were also asked to understand better how this was related to the perceived (potential) advantages and benefits of using a system like the one described in the workplace.

Research Question 2: *How does an increased control over employees' data and a higher transparency about its use affect the employees' perception of technology at the workplace?*

When deploying technology, there is a gap when it comes to addressing the ownership of the captured data, as has been reviewed in the previous chapter. Moreover, most IoT solutions do not give the user control over their data beyond allowing them to browse and explore over their data collected. In the context of the workplace, it is not clear whether the captured information should be owned by the user, the employer or even the platforms being used for such data capturing (Lupton, 2014). Under those circumstances, users may decrease their privacy concerns regarding the use of smart technologies if this control over both the data and the devices was granted to them. In fact, according to Khakurel et al. (2018), the lack of user control on the accessibility of this information and the fact that this data is being obtained by platforms that do not belong to them can hinder their trust in the technology.

By means of this research question, the designed questionnaire is intended to shed light on the participants' perceptions about the ownership and responsibility of the data they produced in the workplace context. In other words, if the data belong to the prevention manager (who are the ones who supply the devices), or if the data belong to the users (who are the ones who use the devices and produce data with them). Based on these initial findings, we delved into how these notions of control increases or decreases the perception of the technology and whether participants believe that the control requirements would be met if the information is not outsourced. That is, it belongs only to them, and they are the only controllers. The following section will introduce the questionnaire designed to obtain evidence to the aforementioned research questions.

3.2 Method and sample

We designed an online study to understand the effect of previously addressed factors in the deployment of technology in the workplace. In the following, the details about this experiment are provided.

3.2.1 Experimental design and sample

This study consists of an online survey designed to answer the research questions and the complementary objectives described in the previous section. As long as ethical considerations are taken into account (Buchanan and Hvizdak, 2009), online survey tools are a relevant method for data gathering that offers a flexible way to collect both quantitative and qualitative information (Evans and Mathur, 2005).

The designed questionnaire contains 4 main sets of questions that include Likert–scale an multiple-choice questions, as well as a final open question to obtain qualitative answers regarding participants general opinion of the use-case presented to them. To begin the questionnaire, participants were provided with a brief study description. Then, they were asked to respond to an initial set of questions aiming to obtain initial insights regarding the integration of smart technologies in the workplace, the perceived compatibility with those smart technologies and how privacy and security risks may affect them. The items for measuring those factors were drawn from previously validated studies (Nikou, 2019; Wang et al., 2020b). This set of questions was repeated at the end of the survey after describing a fictional use case scenario that illustrates a technology-based system for promoting health in the workplace. Thus, the idea was to compare the initial responses against the ones given when finalizing the survey. With that, to analyze whether describing a fictional use case to the participants and making them reflect upon privacy-related aspects could induce changes in their initial answers (Braunstein et al., 2011).

After this initial part of the survey, respondents were asked to fill to demographic and socio-economic questions such as gender, nationality, or education. Then, more specific questions for the scope of this study were provided. These were related to the number of hours spent being sit down daily, owned smart devices, their technological background, and their knowledge about Data Protection Regulation (GDPR) and/or other related data privacy policies in their countries. After these two initial parts of the survey, a speculative scenario was presented. This part of the study follows a user-centred approach through speculative design, a research method for exploring imaginary configurations of the world as a way to interrogate questions about them (Galloway and Caudwell, 2018). In this case, the rationale behind this approach is to understand the users' perceptions and concerns regarding a fictional scenario in where participants are placed. The imaginary situation described introduces the idea of a health promotion campaign orchestrated by the managers of the participant's company, designed to improve his/her habits during the working routine. The following fictional description was provided to the participants:

> *"Imagine that you work in a company that decides to launch a voluntary campaign to promote physical activity at work and healthier habits during working hours. To do so, the managers of the campaign (belonging to your organization) provide you with an intelligent system. This system is comprised of several smart devices (e.g. a smartwatch or water intake tracker) that collect data about your daily routine and monitor those habits and behaviours that could impair your health and wellbeing. By means of those devices, the proposed intervention seeks to shed light on how much time you spend without getting up from your chair, how much you walk or move during working hours, how good or bad your hydration habits are or how many coffees you drink per day. They also change the*

traditional office chairs for smart chairs to improve your ergonomics and, hence, alert you when an adjustment of your body posture is advisable."

After reading the speculative scenario, the respondents had to provide a series of answers related to the privacy requirements of such a system, the different options they foresaw for data storage, their preferred level of control over the data, and other general aspects related to such a perception of emerging technology in the workplace. Then, as introduced, the initial sub-set of questions about integrating smart technologies in the workplace was repeated at the end of the study to observe potential changes potentially due to the self-reflection that the scenario sparks. All the statements and items in every question were randomized to change the order in which they are presented for every participant to remove order bias (Malhotra, 2008). The final survey designed for this study can be found in Appendix A.

We recruited a total of 524 participants, both disseminating the survey among colleagues of the university, by snowball effect, in Twitter, and using Prolific Academic [1], a crowd-sourcing platform that recruits participants for academic studies (Palan and Schitter, 2018). This platform is widely validated by the research community and has a worldwide participant pool with more than 150,000 reliable potential respondents (Peer et al., 2017). Participation was voluntary and all respondents were anonymous. Each participant was remunerated for completing the survey in accordance with the platform's policy. The recruitment was restricted to participants from the European Union and the United Kingdom, although some participants indicated different nationalities. Furthermore, we applied a screening process in which participants had to be actively working in office-based workplaces (e.g. in tertiary buildings or service sector).

The obtained responses were analyzed and reviewed to gather the most relevant issues and aspects. In the following, we provide an analysis of the extracted information. We begin our analysis with the demographic information of the respondents included in this section. Then, in Section 3.3 a detailed analysis of the obtained responses is included to finish this chapter with the extracted conclusions.

[1] `https://www.prolific.co/`

3.2.1.1 Respondent profiles

Table 3.1 provides an overview of the demographic profiles of the respondents who participated in the present study. As the results below show, more males participated in the questionnaire (58.2 %) than females (41.4 %). Besides, most of the participants were 22–40 years old, followed by the 41-52 years-old group. Thus, we see that most of the respondents to this survey were relatively young or middle-aged. In terms of educational achievement, the majority of the participants have a high educational level, being University degree (bachelor or equivalent) the most frequent answer, followed by Post-graduate (master or equivalent). People from 35 countries participated in the questionnaire, mainly from Europe. Portugal (24.0 %), United Kingdom (22.3 %), Poland (14.1 %), Italy (12.6 %) and Spain (7.8 %) are the top-5 countries with the highest number of participants. The last question included in this table is related to the number of hours, on average, that respondents estimate they spend at their work desk in a day. The most common answer was 5-8 hours (51.5 %). This may be an indicator of the type of work they develop, and, considering the nature of the responses, most of the participants seems to have a sedentary working routine.

This part of the survey also includes three additional questions regarding the technological background of the respondent. In the first one, respondents were asked to select which smart devices they own or use regularly. Between the selection of answers, it was deliberately avoided to include options such as smartphone, laptops, tablets or similar gadgets. Instead, only those devices that can be associated with assisting the user in their daily routine or promoting new practices were addressed. Some examples for each device were given, and participants were free to select as many answers as they considered. The 53.10 % of the respondents selected wearable devices or apps for health tracking (fitness trackers, nutrition diaries, etc.), followed by the 42.20 % who marked voice assistants (Alexa/Google home or similar devices). The 33.80 % selected smart home devices (smart meters, smart security cameras or any other connected equipment),% of them, and any other system that guides them in their everyday decision making (e.g., driver-assisted car) by

		Frequency	Percentage
Gender			
	Male	305	58.2 %
	Female	217	41.4 %
	Other	2	0.4 %
Age group			
	<21	2	0.4 %
	22-40	426	81.3 %
	41-52	75	14.3 %
	53-71	21	4.0 %
Educational achivement			
	None	1	0.3 %
	High-school/secondary	46	8.8 %
	Post-secondary (non-university)	62	11.8 %
	University degree (bachelor or equ.)	213	40.6 %
	Post-graduate (master or equ.)	173	33.0 %
	Doctoral degree (PhD or equ.)	29	5.5 %
Nationality			
	Portugal	126	24.0 %
	United Kingdom	117	22.3 %
	Poland	74	14.1 %
	Italy	66	12.6 %
	Spain	41	7.8 %
	France	17	3.2 %
	Ireland	12	2.3 %
	Rest of the world	71	13.7 %
Sitting time/Day			
	None	24	4.4 %
	1-2 horus	49	9.4 %
	2-5 hours	93	17.7 %
	5-8 hours	270	51.5 %
	More than 8 hours	88	16.8 %

Table 3.1: Descriptive statistics of the demographic information of the respondents. N = 524.

the 15.10 %. The percentage of participants that did not select any of the previous options was 24.20 %.

Lastly, for this demographic section, a five-point Likert scale was used to measure respondents' self-reported knowledge about technology and their knowledge about the EU GDPR and/or other data privacy policies of their countries. Table 3.2 summarizes the obtained results, where 1 point corresponds to None knowledge and 5 points to Expert. The 67.9 % of the participants rated their technological background with 4 or 5 points. However, des-

		Ratings (1-5 likert scale)				
		1	2	3	4	5
Technological knowledge						
	Frequency	1	28	139	235	121
	Percentage	0.2 %	5.3 %	26.5 %	44.8 %	23.1 %
Knowledge on data privacy policies						
	Frequency	38	120	204	137	25
	Percentage	7.3 %	22.90%	38.9 %	26.1 %	4.80%

Table 3.2: Self-perceived technological background and knowledge about GDPR or other data privacy policies. N = 524.

pite the wider adoption of technology that this response illustrates, together with the number of devices respondents' own or use, the score obtained for the GDPR question is lower than the reported technological background. In the case of data privacy policies, most of the participants expressed a medium knowledge (38.9 %) and only the 30.9 % of them rated their knowledge with 4 or more points. That is, this percentage is significantly lower than those that refereed themselves as a high-knowledgeable or expert in technology. This illustrates how even the self-reported most technologically literate people are unaware of the regulations concerning privacy and individual rights regarding their personal information.

3.3 Descriptive analysis: questionnaire results

This section presents the results obtained after evaluating the 524 collected responses of the previous survey. Before doing so, we firstly analyze whether the responses for the repeated set of questions present remarkable variances in their responses before and after describirng the speculative scenario (Pre-Post scenario description). Then, we focus on the evaluation of the questions concerning the speculative scenario.

3.3.1 Pre-Post scenario questions

To begin analysing the participants' responses, we tested the possible influence of the speculative scenario on the initial thoughts about privacy, security risks

Constructs	Item	Statement
Privacy Risks	PR1	If I use a smart work device, I would lose control over the privacy of my personal data.
	PR2	My personal information will be less confidential if I use a smart work device.
Security Risks	SR1	The security systems built into smart work devices are not strong enough to protect my information.
	SR2	Internet hackers might take control of my information if I use a smart work device.
Compatibility	COMP1	I feel that the smart work equipment fits my lifestyle.
	COMP2	I feel that the smart work equipment is compatible with my day-to-day needs.
	COMP3	I think that the smart work equipment will fit well into my work environment.
	COMP4	I think that the smart work products and applications are useful for the tasks I do at work.

Table 3.3: Items and constructs for the pre and post scenario evaluation. Those constructs were extracted from previously validated studies (Nikou, 2019; Wang et al., 2020b).

and compatibility. For that, this section evaluates the obtained results of the questions included in the pre and post parts of the survey. Each of these parts includes eight items that respondents are asked to rate between 1 and 5 with a Likert scale according to their agreement level with each of the item. Then, as in the original questionnaires from which these questions have been extracted, those observed variables (items) are grouped into three unobserved latent variables (constructs): Privacy Risks (PR), Security Risks (SR) and Compatibility (COMP). These items and constructs are included in Table 3.3.

The average results for every item in the pre and post scenario are shown in Figure 3.1 and Table 3.4. According to the results, the large majority of the participants perceive using smart devices positively. Concerning their perceived risks regarding privacy and security, participants were more concerned about losing control over stored personal data than security breaches. When compared, it can be observed that perceived benefits of smart devices for the workplace (compatibility) are above perceived risks (security and privacy) ac-

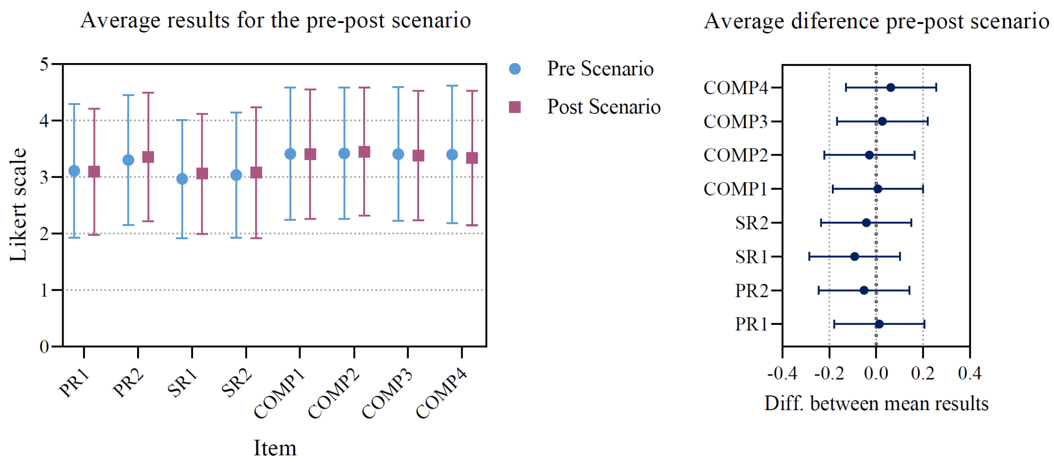

Figure 3.1: Mean values for each of the factors included in PR, SR and COMP for the pre and post scenario questions (left plot) and their difference (right plot).

	Pre scenario		Post scenario		N	Difference
	Mean	Std	Mean	Std		
PR1	3.109	1.184	3.095	1.114	524	-0.0134
PR2	3.302	1.149	3.353	1.137	524	0.0515
SR1	2.968	1.044	3.059	1.060	524	0.0916
SR2	3.036	1.103	3.078	1.157	524	0.0420
COMP1	3.410	1.172	3.403	1.144	524	-0.0076
COMP2	3.420	1.162	3.448	1.132	524	0.0286
COMP3	3.406	1.182	3.380	1.146	524	-0.0267
COMP4	3.399	1.216	3.336	1.187	524	-0.0630

Table 3.4: Comparison between the mean and standard deviation of the pre and post speculative scenarios responses regarding the Privacy Risks, Security Risks and Compatibility constructs.

cording to the responses. Nonetheless, it is also remarkable the magnitude of the privacy constructor, particularly the item PR2, indicating the extent to which participants are aware of the implications and the value of the information captured by these systems.

Furthermore, the average results reveal no notable differences regarding the pre and post results. This lack of significant differences means that the participants did not change their opinion significantly before and after being situated in the speculative scenario.

To verify it, the associations between pre-post changes with regard to the mean results of the items were explored using Pearson correlation coefficient (r), obtaining a high correlation value ($r = 0.97,\ p < 0.0001$). This shows that the pairing was effective, so we conducted a paired t-test to check whether the difference between the pre-and post-test results of the average of all constructs is statistically significant. By conventional criteria, the obtained results ($t(7) = 0.7347,\ p = 0.4864$) show that this difference is considered to be not statistically significant. This comparison indicates that the described scenario did not influence users. The rest of the survey addressed the issue of information control and privacy without raising awareness about it and, therefore, without influencing the user's answers.

3.3.2 Speculative scenario

After presenting the speculative scenario in which participants are placed in a fictional situation, respondents were inquired about their initial thoughts towards participating in the described intervention. The findings reveal that 36.3 % (N=190) of the respondents answering this question preferred to discuss first with their colleagues and peer and then decide how to proceed. Furthermore, a similar number (N=174, 33.2 %) would join the campaign immediately, while 17.6 % (N=92) would only join after a while (i.e., after observing others using the technology). Finally, only 13 % of the respondents would decide to not participate in such fictional intervention, that is, 68 out of the 524 participants.

For the second question of this part of the survey, participants were asked to select from a list the top three aspects that they consider the most valuable for them in the described system (i.e. the provided smart devices for health tracking). The wide range of proposed options and the percentage of participants that selected each of the options are summarized in Table 3.5.

A closer look at the responses shows that, when considered alongside with other options, users place importance primarily on the perceived value of the system, i.e., that the system is helpful to them (54.00 %). In line with Nappi and de Campos Ribeiro (2020) and Nissenbaum and Patterson (2016), this

I Value appreciate...	Freq.	%
That the system is actually beneficial to me or my health.	283	54.00 %
That the system is secure and privacy-aware (i.e your personal information will not be compromised).	236	45.00 %
That I can understand how the system works, what kind of information it collects, which is the use of the collected data and who has access to them.	232	44.30 %
That my personal data is only used for the purpose for which it is collected.	227	43.30 %
That the collected data does not cause any harm to me, neither personally nor professionally.	224	42.70 %
That I can decide which of the data collected by the system is public or private, or to whom it is shared.	188	35.9 %
That the system is efficient enough without requiring much attention or time from me.	182	34.70 %

Table 3.5: Obtained responses to the question *"What do you value most when it comes to using such a system? (i.e. the provided smart devices)"*. Respondents were asked to select three options out of the seven given.

indicates that it is necessary to understand the perceived benefits of using such a system and even the reason for using the captured data (e.g. improving workplace habits). A significant relevance is also given to privacy, in the sense that the information collected will not be compromised in any way or used for any other purpose (45.00 %). Transparency and understanding the purpose of the collected information are also considered essential aspects (44.30 %). Transparency is linked to the policies used to manage the collected data and how this data is used, which can lead to a conflict of interest between the employee and the employer (Chung et al., 2017).

This conflict of interest is closely linked to the meaning given to this data and the possible malicious use of the extracted information. These factors become a paramount concern in the context of the workplace, where every captured data can be related to the productivity or performance within the workplace (Neff and Nafus, 2016). For these reasons, participants also selected as a valuable aspect that the personal data is only used for its original purpose (43.30 %) and that this information cannot cause any harm neither personally nor professionally to them (42.70 %). In this case, choosing whether the data is public or private, or with whom it is shared, was the second least chosen option (35.90 %). This response could be conditioned by those participants who do not consider as an option the possibility of making their data public. Finally, those aspects related to the functionality of the system and its efficiency from the outset were ranked as the least important among these options (34.70 %).

From the reported data, it can be concluded that the top selected statements are closely related to the balance between the risks that users perceive when sharing their data with third parties and the perceived benefits and the added value of the system that captures this data. Thus, obtained responses correlate to the concerns about data sharing, the collection of personal information, and the possibility of this data to be used by third parties without the user's knowledge.

For this reason, the third question of the survey addressed specifically who should have the final responsibility for the data. In essence, who should be the owner of this data and to whom it should be visible. Several statements were provided, and the participants' opinion from 1 (completely disagree) to 5 (completely agree) was asked. The averages of their responses are included in Table 3.6.

When comparing those results, it is highly remarkable that participants lean towards being the only ones who should be granted privileges to own and control their data, deciding which data can be shared and to whom. For this reason, 240 respondents completely agreed with the statement *All data should remain under my ownership, and I will be the only person that can grant access to third parties of certain data I produce.* On the contrary, they mostly disagree with the idea of their data remaining under the ownership of their company. A total number of 155 respondents completely disagreed with this idea, which obtained an average result of 2.485 (SD 1.262), compared to the 3.996 (SD 1.150) of the previous statement. When the ownership of the data

Statement	MEAN	SD
All data should remain under my ownership, and I will be the only person that can grant access to third parties of certain data I produce.	3.996	1.150
All data that I produce should remain under my ownership but visible to interested entities, as long as this data can not be linked to me and only for the sake of my health.	3.403	1.251
All data that I produce should remain under the ownership of my company/organization, as long as these data can not be linked to me.	3.149	1.240
All data that I produce should remain under my ownership but visible to the healthcare company of the organization where I work.	3.147	1.226
All data that I produce should remain under my ownership but visible to my company/organization.	3.094	1.239
All data that I produce should remain under the ownership of my company/organization since they are the health campaign promoters.	2.485	1.262

Table 3.6: Average results to the question *In your opinion... Who should be responsible for the storage and security of the data that you produce in the workplace, and that is collected by the intelligent system?"*. Each item was rated from 1 (completely disagree) to 5 (completely agree).

belongs to them, participants were less reluctant to make it visible to healthcare companies (3.147, SD=1.226) and their organization (3.094, SD=1.239). Another point of interest is the disparity when the concept of data anonymization is included. In such a case, the number of participants who would agree with their companies' data being owned by their companies increases (3.149 SD=1.240), only if this data can not be linked with them. When this privacy is ensured, more participants agree with the idea of sharing this data with interested entities for the sake of their health (3.403 SD=1.251). This is in line with the literature and the general doubts that the sensed data may contain private information concerning personal or productivity-related aspects (Nissenbaum and Patterson, 2016). A more detailed visualization of the results of this part of the survey is included in Figure 3.2.

The conclusions for this question are twofold. On the one hand, participants perceive positively that the captured information can be visible by

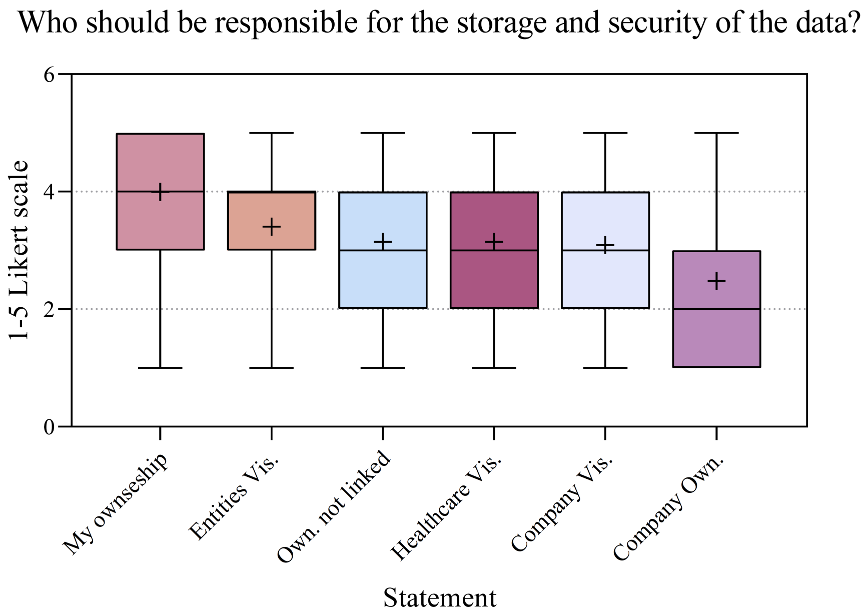

Figure 3.2: Boxplot representation of the responses to the question *In your opinion...Who should be responsible for the storage and security of the data that you produce in the workplace, and that is collected by the intelligent system?"*. Each item was rated from 1 (completely disagree) to 5 (completely agree).

other entities or the health companies only if this information still belongs to them or is not linked with them. This way, having a final purpose for the data increases the perceived value of using such a system. However, on the other hand, respondents value more positively that they can have control over this data instead of being managed by their companies. Thus, according to participants' responses, the concept of privacy and visibility of data by third parties is conditional on their capabilities to have control over this data.

The last question of this survey regarding the speculative scenario put the focus on investigating those aspects that respondents consider more valuable for using the described IoT system. We asked respondents to rate various statements from 1 to 5, according to their level of agreement or disagreement with each of them. The obtained average results are included in Table 3.7 and, visually, in Figure 3.3.

The findings reveal that the most valuable aspect is the system's customization capabilities, which obtained the highest mean score, 4.063 (SD=0.893).

Statement	MEAN	SD
The intelligent system must have the ability to be customized according to my specific needs.	4.063	0.893
The intelligent system can be located locally or remotely, as long as the protection of the collected data is regulated or I have control over them.	3.836	0.995
The intelligent system must be able to interact with me in a tailored manner, e.g. only when I have availability.	3.706	0.935
The intelligent system must be located close to me, in a local device I own, so that the collected data are not stored on external devices (e.g. cloud servers my company own) that I can not control.	3.448	1.126
The intelligent system must be always on, collecting my data at all times.	2.531	1.271
The intelligent system must be remotely located so that the collected data are stored on external devices (e.g. cloud servers my company own) that I can not control.	2.378	1.124

Table 3.7: Average results for rating the included items from 1 (completely disagree) to 5 (completely agree).

Another highly scored aspect related to the system's usability was the possibility to tailor the way the system interacts with the user, obtaining a mean score of 3.707 (SD=0.935). Assessing whether the system should always collecting data was perceived very negatively (2.531 SD=1.271).

The convenience of the data storage location was measured using three questionnaire items. Obtained results show that participants resoundingly disagree with the idea of externalizing their data to be allocated in remote servers that they can not control. A total number of 143 participants completely disagreed with this idea, with obtained the lowest average score of this question (2.378, SD=1.124). On the contrary, as expected based on previous answers, hosting the intelligent system and the data in a local device was very positively valued (3.448 SD=1.126). This is also reflected in the high number of participants who would be in favor of the system being located remotely or locally, if and only if they keep control of the data and the system. The

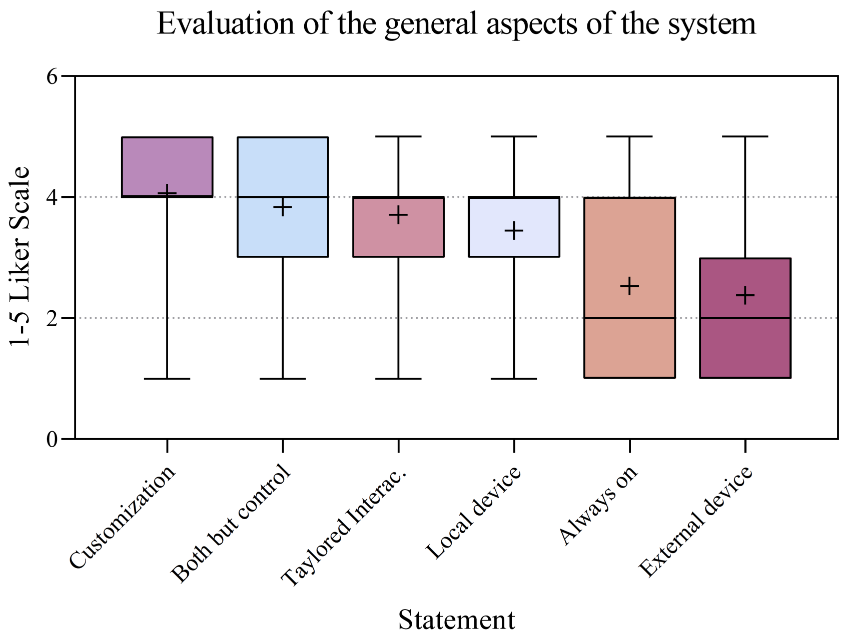

Figure 3.3: Boxplot representation for the obtained responses when rating each item from 1 (completely disagree) to 5 (completely agree).

convenience of this approach obtained even a higher score (3.836 SD=0.995). These results are in line with the insights obtained from the previous question and highlights how the concept of being the only ones in charge of managing their data emerges as one of the most valued point. In general terms, it can be concluded that there is agreement among the vast majority of participants in this respect.

To conclude with the questionnaire, participants were asked to voluntarily provide their feedback regarding the proposed system or the questionnaire itself. A relatively small number of participants (N=80) decided to participate in this optional question. In the following, we include a small selection of opinions that are relevant for the purpose of this study:

Some of the participants emphasized their fears about privacy and the use of their data:

"If such a device would exist I would find it terrifying and invasive of privacy/personal boundaries and I would feel judged and watched"

"Having data like how long am I sitting down or how many coffees I

drank is not data that should be easily accessible to my company. That kind of control is honestly scary and I don't want to live in a world like that."

"I believe any kind of information it tracks is prone to leaking. Hackers can break into any cloud storage. So these days nothing is safe. But still, these advancements are very useful."

Other participant associated his/her personal experiences in the past with the proposed system:

"I love the idea to integrate a smart work device in mi day to day but I´m always afraid it will be hacked and my information shared as it has happened in the past with personal information such as my telephone number"

Also, some of them expressed their concerns about their companies having access to this data:

"My problem with the concept didn't come up in the questions. I would be concerned that instead of being used to promote my health it would be used to track how many hours I spent logged in and working e.g. the opposite of the purported intention."

"I don't fear hackers or "outsiders" getting their hands on my personal data. I fear the company/owners/makers of the program I use will sell my information in one way or another. Even if they say they don't sell my information and even if it is against the law I am confident they will still circumvent this somehow every time - no matter what company, and with no exceptions."

Lastly, one participant raised the dilemma between the functionality of such a system and the potential risks it may entail:

"I think we can all agree on the comfort and practicality of having these devices. They fit into our lives like a glove, but it doesn't mean they can't be harmful."

3.3.3 Measuring the effect of privacy and security risks

So far, we have identified with statistical results the opinions of the participants with regard to the integration of technology at work. Based on this, we have addressed which aspects they consider more positive or negative when it comes to evaluating data privacy, data ownership or system's performance. In the following, an exploratory analysis of these results is detailed to provide a first approximation to the effect of privacy and security, concerning the aspects previously explained, on the compatibility with smart technology for the workplace.

To do so, we propose a conceptual model that shows the study's relationships between those concepts by means of different indicators and latent variables, which, in the following, will also be called constructs. The different constructs and their influence on the perception of the IoT system are assessed through Partial Least Squares Structural Equation Modeling (PLS-SEM). PLS-SEM is a modelling method that allows the estimation of complex cause-effect models to estimate and test causal relationships between latent variables (Hair Jr et al., 2016).

However, the completeness and number of nuances of the reflective items included in the speculative part of this survey hinder the possibility of easily categorizing these items into constructs to be evaluated through the PLS-SEM method. For this reason, we want to highlight the exploratory nature of this part of the section before delving into the results of this procedure. Furthermore, it is essential to note that this study is only intended to analyze further the correlation between the different constructs of the addressed concerns and obtain insights that could indicate a potential effect between them. Therefore, the results presented in this section should not be overestimated and are not intended in any case to constitute a universal integrated model or a measurement instrument.

3.3.3.1 The proposed structural equation model

We have defined 4 new constructs that are added to the PR, SR and COMP ones included in Table 3.3. Those new constructs are: Control (Increased

Constructs	Item	Statement
Control	CTR1	The intelligent system must be located close to me, in a local device I own, so that the collected data are not stored on external devices (e.g. cloud servers my company own) that I can not control.
	CTR2	All data should remain under my ownership, and I will be the only person that can grant access to third parties of certain data I produce
Transference	TR1	All data that I produce should remain under the ownership of my company/organization, as long as these data can not be linked to me
	TR2	All data that I produce should remain under the ownership of my company/organization since they are the health campaign promoters.
	TR3	The intelligent system must be remotely located so that the collected data are stored on external devices (e.g. cloud servers my company own) that I can not control.
Visibility	VI1	All data that I produce should remain under my ownership but visible to my company/organization
	VI2	All data that I produce should remain under my ownership but visible to the healthcare company of the organization where I work.
Conditional	CN1	All data that I produce should remain under my ownership but visible to interested entities, as long as this data can not be linked to me and only for the sake of my health.
	CN2	The intelligent system can be located locally or remotely, as long as the protection of the collected data is regulated or I have control over them

Table 3.8: New items and constructs, extracted from the designed questionnaire, for the analysis of the structural model.

control requirements are positively considered), Transference (Reduced control requirements and the transference of the responsibility of the data are positively considered), Visibility (granting access to the data and making it visible to third parties is positively considered), and Conditional (Control and Visibility are positively considered only if they are conditional on another factor). The items of the questionnaire grouped under each of these constructs are listed in Table 3.8. To better associate those new constructs with the initial ones (PR, SR and COMP), we will consider for this part of the evaluation only the responses to the post-scenario data, as those responses were obtained once participants were familiar with the described speculative scenario.

Figure 3.4 shows the proposed structural model. This model hypothesizes that the perceived privacy and security risks condition the self-reported com-

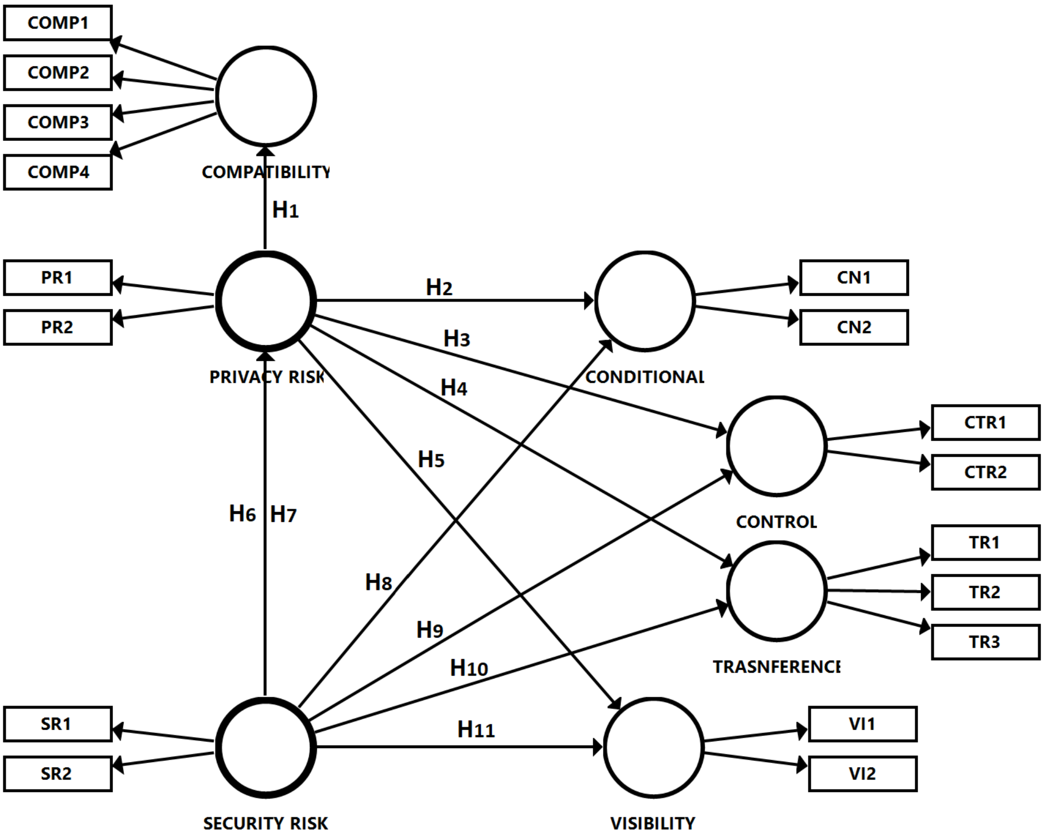

Figure 3.4: The proposed structural model for the exploratory analysis of the effect of PR and SR on the perceived COMP and the preferences regarding Control, Transference, Visibility and Conditional aspects.

patibility with smart technologies for the workplaces and the participants' preferences regarding Control, Transference, Visibility and Conditional aspects. Each of the potential effects is hypothesized through the relationships between the constructs. In other words, this analysis aims to determine the potential influence of PR and SR on participants preferences over the characteristics of the smart systems that have been grouped as Control, Conditional, Visibility and Transference constructs, as well as the influence of PR and SR on self-reported compatibility with the described smart system. We also consider the hypothesized relationship between PR and SR both directly and indirectly to evaluate their effect. In the following, the measurement and the structural model evaluated using SmartPLS software (Ringle et al., 2015) are

	1	2	3	4	5	6	7
1. Compatibility	0.910						
2. Conditional	0.172	0.750					
3. Control	-0.120	0.045	0.800				
4. Privacy Risks	-0.300	-0.023	0.320	0.909			
5. Security Risks	-0.189	-0.084	0.287	0.685	0.898		
6. Transference	0.255	0.149	-0.287	-0.167	-0.106	0.746	
7. Visibility	0.198	0.247	-0.107	-0.121	-0.075	0.380	0.855

Table 3.9: Discriminant validity - correlation between constructs and AVE square root on diagonal.

illustrated and discussed.

We first evaluate the measurement model. To begin with, the discriminant validity of the constructs was obtained to determine whether those constructs that should not be related are actually unrelated. Then, this discriminant validity, whose results are included in Table 3.9, was confirmed using the (Fornell and Larcker, 1981) criterion. This means checking that the square root of the Average Variance Extracted (AVE) included on the diagonal exceeds all correlations between that construct and the rest of them.

Then, we examine the internal consistency of the models. A Confirmatory Factor Analysis (CFA) was performed for all constructs depicted in Figure 3.4. The obtained results are included in Table 3.10.

The first measured factor of the CFA is the factor loading, which represents the strength of the relationship between constructs and items. According to Hair et al. (2009), factor loading values higher than 0.5 are acceptable and practically significant for exploratory studies, while 0.7 or higher are desirable. We meet the first requirement in all the factor loadings except one, CN1. Given the completeness of the items included in the Conditional constructor, it could be expected that such a complex item may not have enough relevance in explaining this construct. The traditional procedure in such cases is to remove this item. However, as the AVE and Composite Reliability (CRs) values of Conditional construction meet the requirements for exploratory research (Fornell and Larcker, 1981), we decided to maintain it in our model. We

		Mean	SD	Factor L.	Cronbach's α	CR	AVE
Conditional	CN1	3.403	1.249	0.367	0.384	0.676	0.558
	CN2	3.836	0.994	0.991			
Compatibility	COMP1	3.403	1.143	0.919	0.931	0.951	0.829
	COMP2	3.448	1.13	0.923			
	COMP3	3.38	1.145	0.914			
	COMP4	3.336	1.186	0.884			
Control	CTR1	3.448	1.125	0.679	0.463	0.777	0.639
	CTR2	3.996	1.149	0.905			
Privacy Risks	PR1	3.095	1.113	0.913	0.790	0.905	0.826
	PR2	3.353	1.136	0.905			
Security Risks	SR1	3.059	1.059	0.910	0.761	0.893	0.806
	SR2	3.078	1.156	0.885			
Tranference	TR1	3.149	1.239	0.585	0.617	0.781	0.551
	TR2	2.485	1.261	0.919			
	TR3	2.378	1.123	0.684			
Visibility	VI1	3.094	1.237	0.767	0.647	0.841	0.727
	VI2	3.147	1.224	0.930			

Table 3.10: CFA results to examine the internal consistency of the model.

acknowledge the potential limitations for the consistency of this model that this may entail for the validation of this study.

Cronbach's alpha (α) values, which measure how related are the items as a group, were all above the acceptable threshold of 0.60 for exploratory research (Hair et al., 2017), except the Conditional and Control constructs. It should be noted that Cronbach's alpha needs to be interpreted in context and with respect to the number of items that constitute the construct. A small number of items (N = 2 in some of the constructs in our case) may affect this value. This possibility was described by van Griethuijsen et al. (2015). In that case, authors justified continuing with their analysis by arguing that, as the Spearman-Brown prediction formula shows, slightly increasing the number of items would lead to acceptable values for Cronbach's alpha. We adopt the same strategy, underlying again the possible limitations of this circumstance.

H	Path relationship	β	t-value	p-Values	Validation
H_1	**Privacy Risk ->Compatibility**	**-0,299**	**6.834**	**<0.001**	**Supported**
H_2	Privacy Risk ->Conditional	0.066	0.851	0.395	Not significant
H_3	**Privacy Risk ->Control**	**0.231**	**4.291**	**<0.001**	**Supported**
H_4	Privacy Risk ->Transference	-0.181	2.974	0.003	Not supported
H_5	Privacy Risk ->Visibility	-0.134	2.001	0.0456	Not supported
H_6	**Security Risk ->Compatibility**	**-0.205**	**6.451**	**<0.001**	**Supported**
H_7	**Security Risk ->Privacy risk**	**0.685**	**22.198**	**<0.001**	**Supported**
H_8	Security Risk ->Conditional	-0.085	1.318	0.18770	Not significant
H_9	**Security Risk ->Control**	**0.287**	**6.844**	**<0.001**	**Supported**
H_{10}	Security Risk ->Transference	-0.106	2.185	0.029	Not supported
H_{11}	Security Risk ->Visibility	-0.075	1.436	0.151	Not significant

Table 3.11: Total effects β between the different constructors of the structural model. Not significant if p > 0.05 or t-value < 1.96. Not supported if $\beta < |0.2|$.

The internal consistency reliability of the construct is measured through CR. CR values higher than 0.60 are acceptable in exploratory research, and all the constructs meet this minimum requirement Hair et al. (2009). Finally, according to Fornell and Larcker (1981), AVE values above 0.50 shows that both construct and individual variables have high validity. Again, our study meets the minimum value for this indicator in all the constructs.

Having evaluated the model through CFA with the aforementioned limitations, a path analysis for hypothesis testing is performed to study the size of a relationship between two constructs. That is, how strong the effect of one construct is on another construct. This relationship can be negative or positive. The estimated effect size is greater than 0 if hypothesized as positive and less than 0 if hypothesized as negative, with higher numbers (between −1.0 to 1.0) showing a more robust association (Mclntosh and Gonzalez-Lima, 1994). Path coefficient should be larger than $|0.2|$ and have a t-value over 1.96 in order to demonstrate its significance in the predicted direction (Wong, 2013).

Table 3.11 includes all the hypothesized relationships, the obtained total effects given by the standardized path coefficient, and the indirect effect. It also includes whether this hypothesis is supported or not. The structural model results reveal that PR has a negative and statistically significant influence on Compatibility (β=-0.299, t=6.834, p<0.001). Similarly, SR shows a negative indirect effect with Compatibility (β=-0.205, t=6.451, p<0.001).

Thus, perception of privacy and security risks were found to contribute negatively to the potential compatibility of the smart system, which validates H_1 and H_6. This indicates that the compatibility with the smart system might decrease when individuals find the perceived privacy and security risk high.

A strong direct positive effect was also found between SR and PR (β=0.685, t=22.198, p<0.001), meaning that the perception of SR directly affects the perceptions of PR, supporting H_7 and having, thus, an influence on Compatibility. The last construct whose path relationships with PR and SR is strong enough to be supported is Control. For this reason, respondents perceived requirements for Control are directly influenced by the Privacy Risks (β=0.231, t=4.291, p<0.001) and the Security Risks (β=0.287, t=6.844, p<0.001). Therefore, the control intentions might become more positive for those participants with increased privacy and security concerns, validating H_3 and H_9.

Finally, The influence of PR and SR on Conditional, Visibility or Transference was not strong enough to be supported. Therefore, H_2, H_4, H_5, H_8, H_{10} and H_{11} are all rejected by the model. Even though being rejected, their results are in line with those already mentioned. Privacy and Security Risks report a weak but negative influence on the Transference and Visibility constructs, meaning that individuals are more reluctant to grant access to their data to third parties when they have a higher perception of the privacy and security risks.

In summary, the exploratory analysis described above allows us to establish specific influences between the different concepts discussed. Firstly, the perception of compatibility with the proposed intelligent system is conditioned mainly by the privacy and security risks perceived by the participants. Therefore, these concerns are a burden for seeing this system as a suitable tool for personal use in these spaces. Secondly, there are indications of the influence of the perception of privacy and the importance given to the ability to control data. Consequently, users who perceive these risks more strongly tend to choose those options that allow them to be in charge of this data as a precautionary measure. This is in line with the conclusions obtained from the descriptive analysis performed over the questionnaire results.

3.3.4 Limitations

Before outlining the main conclusions extracted from the gathered information, the limitations of this study need to be addressed. To begin with, the conclusions derived in this study are based on the statistical data obtained from a fictional scenario. That is, participants' responses are given considering the described speculative scenario and their opinion is not based on a real system that all of them had the chance to test in their work premises. Consequently, their opinion may be conditional on the technology they use on a daily basis or on their previous experiences with other similar devices or systems.

Moreover, despite the broad range of profiles that this survey covers (included in Section 3.2), this evaluation does not ensure equal coverage of profiles. Thus, the second limitation concerns the generalization of the findings (i.e., external validity). To begin with, those findings are based on users and mostly reported a medium-to-high level of technological background. Thus, the findings may not be applicable to those who are less technologically advanced and may lean towards heavy IoT technology users. Secondly, demographics data mainly covers European Union countries as well as the United Kingdom. Moreover, most of the responses are grouped into 5 countries (Portugal, United Kingdom, Poland, Italy and Spain). The perception of the concept of sensitive information or the associated privacy-related behaviors may depend on the cross-cultural differences of the population. Those cultural differences and data regulatory policies of such countries are not considered in this study. Hence, the conclusions derived in this study are based on statistical data that can not be universally generalized.

Finally, carefulness needs to be taken in order to avoid overstating the conclusions of the second part of this section (i.e., internal validity). The addressed statistical limitations of the structural model hinder the task of providing an impactful analysis of the obtained results of that part of the study. Those hypothesized relationships are based on constructors created by grouping the items included in the questionnaire by the researchers. Some of these items are sufficiently complex that categorizing them into these constructs may not be robust enough to validate the defined model, even though two investigators validated the grouping, ensuring the inter-rater reliability (Gwet, 2014). Thus, although the evaluated model meets most of the reliability and validity criteria (with some pertinently reported exceptions), they are based on exploratory rule of thumb metrics that are suitable for early stages of research but may not be consistent enough. Further significance in these criteria and a deeper analysis of the model's robustness would be necessary to validate it. For this reason, this analysis can not be considered confirmatory nor to provide empirical evidence of those findings. Instead, it should be taken as a complementary approach to understanding better the conclusions obtained from the initial descriptive analysis of the questionnaire data.

3.4 Conclusions and implications for this dissertation

Throughout this chapter, a study to understand in detail the implications of the integration of IoT technologies in the workplace has been introduced. To this end, an online survey was conducted, collecting 524 responses from people of different locations in Europe. In this survey, respondents were confronted with a speculative scenario in which their fictional company was the promoter of a campaign to foster healthier habits at work. To that end, IoT devices would be in charge of recognizing employees habits and give them advice on how to improve them. Based on this described fictional situation, participants were asked to provide their opinion regarding their requirements for adopting such a system. The obtained responses have been quantitatively analyzed and the limitations of this study clearly addressed. Even considering these limitations, both the descriptive analysis of the results and the findings obtained from the complementary structural model show potential to obtain the final remarks that conclude this chapter.

To begin with, it can be observed how privacy and security risks have a negative effect on the perception of IoT systems. However, according to the results shown in Table 3.4, the self-reported compatibility of the participants with the technology has proven to be higher than the perception of these risks, indicating an initial positive predisposition of the participants to adopt it in work environments. Within the perception of these risks, users are most concerned about the potential use of the captured data. Therefore, as Table 3.5 illustrates, in addition to the fact that the system is beneficial for them, having the evidence that their data will not be compromised or used for other purposes appear among the most valued characteristics in relation to the system described.

In response to the Research Question 1, the obtained results reflect that respondents' perception of data collection in those spaces is related to the possible cost-benefit ratio of generating such data. This cost is associated with the misinterpretation that can be made of this data. Once participants are confident that they are the only ones who decide how to proceed with their data, they feel more akin to making their data available to others. The same applies when the data can not be associated with them by any means. In those cases, sharing their data with the healthcare company of the organization they work for or with interested entities is perceived more positively as there is added value with the process of collecting this data (see Table 3.6).

Concerning Research Question 2, the main problem arises with the reluctance of participants to let their own companies be in charge of storing and controlling the data captured, as this may entail a limited control of the information shared with third parties. Consequently, when inquired about the storage of data, having complete control and ownership over the collected data and deciding to whom it can be shared emerges as the most agreed factor. Besides, as included in Table 3.7, they completely disagree with the idea of storing their information in remote devices that they cannot control. On the contrary, they feel more confident with the idea of locating the storage of this data in devices closer to them to ensure that this control is granted. However, they could also opt for remote storage if they are entirely sure that they are the only administrators of their information.

Those identified requirements for control are in line with the findings obtained from the evaluation of the different factors proposed in the analyzed structural model. The privacy concerns of the participants directly influence respondents' perceived requirements for control. At the same time, those privacy risks have an active influence on the self-reported compatibility with smart systems in the workplace. Thus, this provides insights that indicate that the perception of IoT technologies in the workplace increases if users' control over the information is ensured. Finally, the importance of involving the user in the personalization of the system was also reflected in this question, as the most agreed statement was related to having the ability to customize the system according to their needs.

In summary, the combination of the insights obtained from Chapter 2 and the results of the present study provide us with a solid theoretical background and motivation to fully embrace an Edge Computing approach in this work. Through a local approach, we aim to let users always have control over their information. Thus, the privacy of the data is promoted by preserving its externalization. Furthermore, in our vision of a human-centric IoT approach, this control is also extended to the possibility of interacting with intelligent systems to personalize it to their expectations and requirements. In the following, this dissertation will introduce several approaches and proposals that aim to contribute to embrace this privacy-enabling and interactive smart workplace that participants demand.

Everything should be made as simple as possible, but no simpler.

Albert Einstein

CHAPTER 4

Cost-accuracy trade off for optimizing local data processing

In response to users' demand for privacy, trust and control over their data, executing ML tasks at the Edge of the system has the potential to preserve the privacy of capture data and create positive interactions between users and the IoT through a *privacy-by-design* approach. This implies moving complex computational tasks to a local stage. Specifically, computational tasks related to both inference and training stages of the classification process, in which Edge devices must balance the computational cost of the ML techniques to meet the available resources.

The main difficulty that will be observed throughout this chapter arises when high-demanding applications compromise the computational capability of the Edge devices. Some scholars have pointed out that pre-trained models are easily deployable in constrained nodes and, hence, inference can run on them. However, processing new data and training new models are computationally intensive tasks that are still insufficiently investigated in Edge devices. Thus, if a model needs to be retrained to integrate new data, the usual answer to overcome this drawback is to move the training stages to Cloud servers. However, this involves externalizing the captured data, which is against the benefits of Edge Computing in terms of privacy. Hence, the next stage of an intelligent Edge is to find new ways to integrate the whole classification pipeline at the Edge to perform on-site training and decrease data exposure.

For this reason, this chapter aims to analyze all the factors affecting the classification process. Then, to empirically evaluate their impact in terms of performance and cost, addressing together the training and inference stages of the ML pipeline. From there, further evaluation of the trade-off between classification accuracy and computational cost can optimize the performance of the system. To this end, a multi-objective Pareto optimization analysis will be presented to find the subset of solutions with the best cost-accuracy trade-off. We focus our analysis on HAR systems, representing a standard type of classification problems in human-centered IoT applications. In essence, this chapter contributes with empirical data to demonstrate the opportunities that the optimization of the classification pipeline can provide to (i) match the available resources of Edge devices for a balanced cost-accuracy trade-off and; (ii) be able to integrate both inference and training at the Edge instead of the Cloud, avoiding extra threats on compromising private data. These objectives are intended to create spaces where human and machine intelligence could collaborate while data control and privacy are promoted.

The chapter comprises various sections, starting from Section 4.1, which describes the different variables and factors that take part in the classification pipeline. Section 4.2 explains the procedure and the experimental setup followed to measure the performance of the activity recognition system. Section 4.3 describes the influence of the factors described in this chapter and their impact on the deployment of similar applications in resource-constrained devices. The proposed strategy is also applied to evaluate the potential optimizations that can be obtained in an automatic way (i.e., setting a threshold for the maximum classification loss). Finally, the conclusions of the chapter are presented in Section 4.4. Most of the results described in this chapter have been published as a journal article (Gómez-Carmona et al., 2020).

4.1 A holistic approach to optimize the classification pipeline

Detecting human activities through the classification of sensor data involves an ensemble of computationally intensive tasks. For this reason, in the following, we want to better understand the computational cost of each of those tasks and how these are connected to the resulting accuracy of the classification results. Thus, in this section, we identify and break down from a methodological point of view the different factors that take part in all the stages of the ML process, considering together both training and inference phases of the ML pipeline.

To that end, we have identified five factors that affect the computational cost of HAR applications in IoT contexts: (i) the platform on which the classification is performed and the set-up, (ii) the sampling frequency in which raw data is captured, (iii) the number of components of the signal (for accelerometer data), or the number of sensors that can be involved, (iv) the number of features extracted from the data, and, finally, (v) the ML algorithm used. These factors are depicted in the center of Figure 4.1.

When calculating the cost, we should differentiate between two main stages in the ML process. On the one hand, creating a new model (training stage, also known as model fitting), and, on the other hand, deploying this trained model to make predictions (inference stage). Both stages, as can be observed

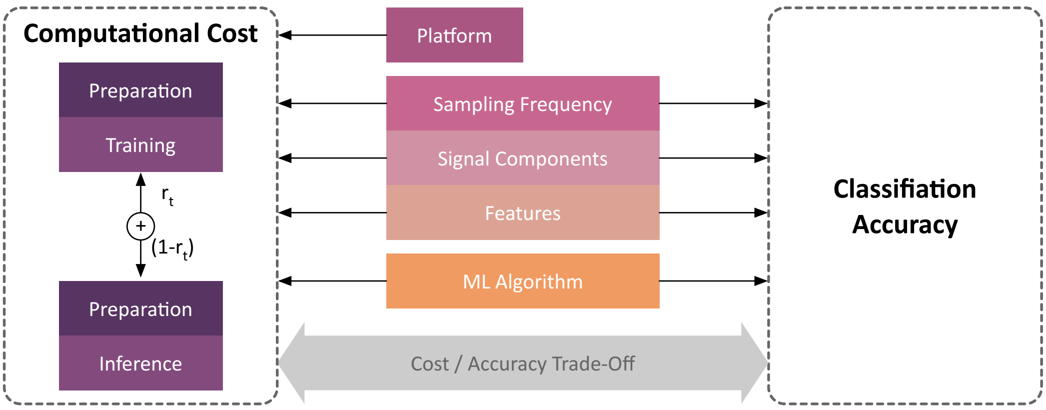

Figure 4.1: Schematic representation of the different factors affecting the cost-accuracy trade-off.

in Figure 4.1, involve data pre-processing (preparation), which corresponds to analyzing the raw data and transform them into useful inputs for ML algorithms. Such preparation usually relies on features calculated from the processed data that represent the main characteristics of the initial data.

As explained in the introduction of this chapter, we foresee executing the training stage at the Edge and not only inference. In this way, training also contributes to the sum of the total computational cost we will calculate for the whole pipeline. Depending on the ML technique, the training cost for a model can be several orders of magnitude higher than its inference associated cost. However, it is true that training new models is usually done much less frequent than inference. For example, a HAR system could infer movements from data several times per minute, while retraining only happens every few days. We capture this by the training ratio r_t; $0 < r_t < 1$, so that the total computational cost c_{total} is calculated with Equation 4.1:

$$c_{total} = r_t * c_{training} + (1 - r_t) * c_{inference} \tag{4.1}$$

We will later study the influence of this ratio on the feasibility of different solutions.

Apart from affecting the computational cost, all factors of Figure 4.1, except the platform, also influence the accuracy of the classification. Thus, all these factors condition the reliability of ML algorithms when it comes to detecting human activities from the sensor data. This interdependence implies a critical trade-off in the design and operation of the proposed Edge system. After the choice of the platform, which is done during deployment and determines the available hardware, the other parameters can be adjusted during runtime. Thus, these factors provide the possibility to adjust computational cost and accuracy, which will be studied in the remainder of this chapter. Based on this analysis, we present a holistic optimization strategy through data reduction that can mitigate the limitations of resource-constrained devices.

4.2 Procedure and Methodology

To evaluate the performance of several ML algorithms, we propose a HAR classification problem where sequences of accelerometer data are used to relate the inertial signals to previously labeled actions. Even though this optimization process can be extrapolated to any classification problem, we will initially focus our study on one of the target activities related to wellness monitoring in office environments: hydration habits. Therefore, to test the suitability of this approach with a validated dataset, we first focus on the methodology to design a drinking activity detection system using a publicly available dataset that includes this targeted activity, namely the *ADL Recognition with Wrist-worn Accelerometer Dataset* (Bruno et al., 2014). Then, we will address how this strategy can be applied to create an optimization pipeline that reduces the initial data to the minimum search space that meets a pre-set requirements that get a good cost-accuracy trade-off.

This second stage will introduce a novel dataset for monitoring office hydration patterns, explicitly created to fit into the context of this dissertation. This novel dataset serves as a use-case in which the surrounding working environment is enhanced to incorporate IoT technologies, e.g., placing a sensor in the usual bottle or mug used by each worker. More specifically, this dataset was elaborated with the idea of having an activity dataset that resembles real-world scenarios and incorporates high variance between users and classes, ensuring the robustness of this approach to fit into the particularities of final end-users. Moreover, we will also evaluate the proposed method with a public dataset for recognizing larger activity patterns. To sum up, in this part of the work, we consider three datasets related to sensors that represent a typical use-case scenario for HAR applications and our specific context. This enhances the evaluation procedure and the generalization capabilities of the system, minimizing potential biases in the validation of the proposed approach.

4.2.1 Selected datasets

In the following, we describe the datasets aforementioned that will be used for evaluation purposes through this chapter. Starting with the two datasets obtained from the literature, we will afterwards introduce the previously mentioned novel dataset that we have created for this dissertation.

4.2.1.1 Public datasets

- *ADL Recognition with Wrist-worn Accelerometer Dataset (Bruno et al., 2014)*

  To evaluate in detail the impact of those factors described in 4.2, we have selected, from the existing HAR datasets, one that includes the drinking activity, namely, the ADL Recognition with Wrist-worn Accelerometer Data Set. It is publicly available in the UCI ML Repository (Dua and Graff, 2017). This dataset contains 14 labeled activities of daily living in 839 recorded trials, or instances: brush teeth, climb stairs, comb hair, descend stairs, drink glass, eat meat, eat soup, get out of bed, lie down in bed, pour water, sit-down in a chair, stand-up from a chair, use telephone and walking. These activities were recorded by 16 volunteer participants (11 men and 5 women aged between 19 and 81 years) wearing a right-wrist worn tri-axial accelerometer with an output rate of 32 Hz. The accelerometer measurements are provided as a series of data points indexed in time order so that activities are classified based on the signal behavior over a period of time. As such, this is a time series classification problem.

  With the idea of better fitting to the scope of this work, this dataset was binarized to represent two classes: drink gesture (100 instances) and the rest of the activities (739 instances). Even though this might create an unbalanced distribution of classes, the available data still provide a suitable framework to perform the comparative analysis of the classification solution. Thus, to maintain the classification pipeline as standard as possible for evaluation purposes, no specific balancing method was applied to the data. Note that performing any balancing method (under-sampling or over-sampling strategies) would have an impact on the computational cost, which should also be considered when balancing the cost-accuracy trade-off. Therefore, we did not introduce such a phase in our proposal.

- *The Daily and Sports Activities Data Set (Altun et al., 2010)*

  Despite being out of the scope of this dissertation, we have also selected one more general and widely validated dataset that includes a large variety of different activities. This dataset was explicitly chosen to support our findings of the optimization of resource-constrained devices since it is one of the datasets recommended for measuring the performance of Edge devices for classification purposes (Banbury et al., 2020). This dataset contains 19 different activities. (sitting, standing, lying on the back and on the right side, ascending and descending stairs, standing in an elevator still, moving around in an elevator, walking in a parking lot, walking on a treadmill with a speed of 4 km/h, running on a treadmill with a speed of 8 km/h, exercising on a stepper, exercising on a cross-trainer, cycling on an exercise bike in horizontal and vertical positions, rowing, jumping, and playing basketball). A total of 8 subjects (4 female, 4 male) participated in the data collection, performing each of the activities for 5 minutes. Then, the 5-min signals were divided into 5-sec segments. Data was obtained through gyroscopes and magnetometers located in different positions. In this work, only the torso and right arm sensors have been considered.

4.2.1.2 The Office Hydration Monitoring dataset

In this subsection, this novel dataset, specifically created for this work, will be presented in detail. The OHM dataset focuses on classifying office employees' hydration patterns (e.g., drinking water, tea or coffee) based on a wearable sensor placed on different liquid containers (e.g., mug, bottles or glasses). Figure 4.2 shows an example of the setup for collecting the data for one of the used bottles. The Office Hydration Monitoring (OHM) dataset is publicly

Figure 4.2: The IoT sensor used to collect motion signals placed on one example of the employed liquid containers.

available at (Gómez-Carmona and Casado-Mansilla, 2021), where more details about it can be found.

This dataset contains 1000 recorded sequences of time series data performed by 10 different subjects. These instances include 25 variations of different interactions that could be made with liquid containers, included in the taxonomy presented in Table 4.1. Each of the 25 variations was repeated 4 times for each volunteer (6 male, 4 female, all right-handed), and 3-axial accelerometer and gyroscope signals were recorded to categorize the movement of the liquid container.

The dataset provides labeled recorded executions of these variations grouped into three main classes: (1) drinking from a bottle, (2) drinking from a cup and, (3) other kinds of interactions, e.g., inspect or shake the glass/cup or the bottle). Alternatively, a binary representation of the classes can be considered

Dataset Taxonomy				
Container	**Label**	**Posture**	**Description**	**Number**
Glass/ Mug	Drink	Sitting	Grab glass, drink, leave glass	1
			Grab glass, sip, leave glass	2
		Standing	Grab glass, drink, leave glass	3
			Grab glass, drink, walk	4
			Grab glass, sip, leave glass	5
		Walking	Drink	6
	Other	Sitting	Grab and leave the glass	7
			Raise and lower	8
			Move the glass (from one point to another)	9
			Grab the glass and stand up	10
			Inspect the glass	11
			Shake the glass	12
		Standing	Grab the glass and walk	13
			Sit and leave the glass	14
		Walking	Walk	15
			Walk and leave the glass	16
Bottle	Drink	Sitting	Grab bottle, drink, leave bottle	17
			Grab bottle, open, drink, leave	18
		Standing	Grab bottle, drink, leave bottle	19
			Grab bottle, open, drink, leave bottle	20
			Garb bottle, open, drink, walk	21
		Walking	Open bottle and drink	22
	Other	Sitting	Grab the bottle and close it	23
			Pour the bottle of water into sth	24
			Inspect the botlle	25

Table 4.1: The taxonomy of the different activities performed by each subject during data collection. Each variation of the 25 variations was repeated 4 times.

with ”Drink” and ”Other”. This dataset was created with the idea of having a semi-controlled activity dataset that resembles real-world scenarios. Therefore, the interaction to be recorded was intentionally described very vaguely to the volunteer and no detailed instructions were given to guide their movements. Moreover, all of them had their own containers (no instructions were given detailing which container should be used apart from the bottle and mug/glass categories).

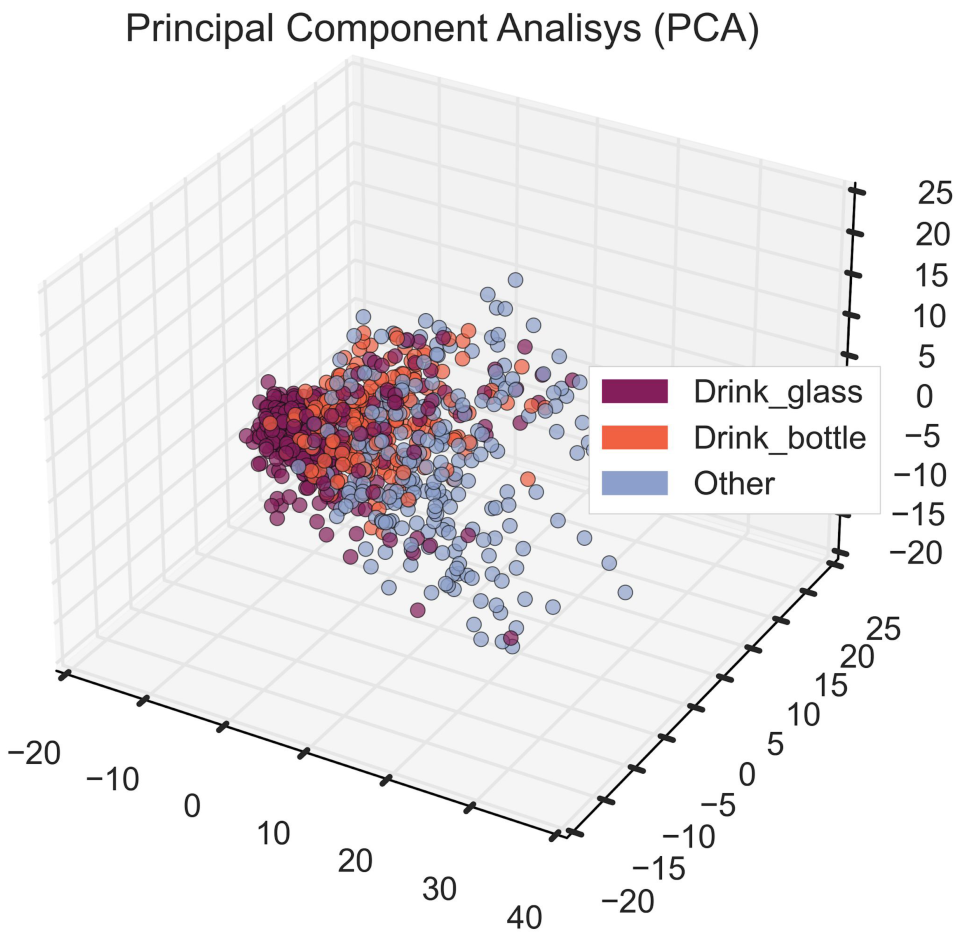

Figure 4.3: 3-dimensional representation of the Principal Component Analysis for the OHM dataset and the three-class scenario.

Besides, the placement of the sensor around the glass, mug or bottle was not fixed, and only the component of the signal perpendicular to the plane pointed the same direction in every case (i.e., volunteers could rotate the water container with the sensor attached, and the initial orientation was not fixed). Thus, this induces a high variance in the recorded data, as the reference system for the accelerometer and gyroscope signals can vary.

To illustrate this variance and the internal structure of the data, a Principal Component Analysis is included in Figure 4.3 for the three-class scenario. This representation of the data indicates a clear overlapping of classes in the 3-dimensional feature space.

In summary, the main strengths of this novel dataset are:

- (To the best of our knowledge) There is no public dataset of a ”Smart Bottle” or similar where the sensor is placed in the container.
- Variable position of the accelerometer: The position of the accelerometer on the components perpendicular to the plane is variable. That is, it has not been controlled to be always in the same direction.
- Large variance within classes: Taxonomy that contemplates a high number of possible interactions with the device in different situations.
- Use of different containers: Traditional mugs, Bamboo cups, Stainless steel bottle, Plastic bottles or Thermos-type bottles, among others.
- Variable weight of the liquid: Together with the shape and weight of the container, the amount of water can influence the movement. For example, using a plastic water bottle with a small quantity of water generates vibrations when left on the table. The sensor registers those vibrations.
- Semi-structured movements: Mechanical or staged movements are avoided. Instructions were given about each subclass, but no specific way to perform each movement was indicated. Instead, each participant was given the freedom to execute them as the movement seemed most natural to them.

4.2.2 The design of the classification system

In the following, we provide a detailed description of all the steps for designing the whole classification pipeline using the previously introduced ADL Dataset. These steps can be extrapolated to any other applications with similar nature and include an exploratory analysis of the initial data, which is pivotal for the optimization of the whole classification pipeline.

4.2.2.1 Data preprocessing

The design of the HAR systems starts with the preparation phase. As a first step, we applied a 3-point median filter to smooth the signal to avoid big spikes and anomaly values induced by noise. This filter is a simple yet effective method to reduce high-frequency noise that usually is combined with low-pass filters (Arias-Castro et al., 2009). In this case, after checking that there were no significant improvements in classification results, no frequency-domain filter was implemented to reduce the computational complexity. Thereafter, every filtered sequence of data goes through a segmentation process. The entire window of the data sequence is divided into five segments (or sub-windows) of equal length without overlapping.

Algorithms usually rely on features calculated from raw data for classification. This study primarily focuses on time-domain statistical features that exhibit better cost-benefit properties than frequency-domain features (Dargie, 2009). The selected set of features includes mean, min, max, standard deviation, median, kurtosis, skewness, variance and mean absolute deviation. This selection represents the most commonly used features for activity recognition problems according to the literature (Janidarmian et al., 2017). These nine statistical features are calculated for every component of the signal (X, Y, and Z), for each of the five segments in which the signal is divided and also for the entire sequence. This sums a total of 162 initial features that are vectorized to characterize every sequence of the dataset. Then those features are scaled into a [0-1] interval using max-min normalization.

4.2.2.2 Feature selection

Feature selection is a discriminating process to find important features that have more weight in the model. It consists of reducing the dimension of the feature matrix, removing the irrelevant features, and obtaining the subset that contributes the most to the prediction. Thus, this stage of feature selection is essential when simplification is needed to optimize the classification process and reduce training and inference time. The feature selection strategy can be chosen based on considerations such as simplicity, stability or classification

Position	Feature	Score
1	Min X 3	86.121
2	Min X	80.993
3	Mean X 3	66.046
4	Median X 3	62.846
5	Min X 2	44.452
6	Mean X 2	36.978
7	Mean X	36.236
8	Median X 2	34.839
9	Median X	32.655
10	Mean X 4	29.103

Table 4.2: The most representative features and their correlation score values indicating the strongest dependency.

accuracy, and it may entail substantial differences on the final classification results (Chandrashekar and Sahin, 2014). In this work, we apply a Chi2 filtering method that evaluates the correlation between variables (features and target classes) and ranks them according to their contribution to the prediction (Liu and Setiono, 1995). This method is a computationally light strategy that provides a good balance between its potential results and its simplicity (Suto et al., 2017).

The 10 most representative features are listed in Table 4.2. Every feature is named by its statistical property, its accelerometer component, and the time-series sequence segment to which the feature corresponds. As observed, in this feature selection method, the top-10 features were captured from acceleration data in X, which proved to be the most representative one for the drinking activity. The higher scores of the top-ranked features indicate how relevant this short subset of characteristics is in the classification results. This makes it feasible to reduce the number of features without an entailed substantial loss of accuracy on the detection.

We obtained this reduced subset of characteristics by applying the feature selection process to the whole dataset. However, to evaluate the accuracy of algorithms, the most representative features for each model are repeatedly calculated within the validation process using only the training data.

4.2.3 Experimental setup

According to Dhar et al. (2019), understanding how ML algorithms can contribute to Edge Intelligence is a crucial challenge for on-device training. Traditional methods are especially interesting for building Edge learning capabilities, mainly when the computation power is low and the memory is limited. Hence, there is a need to explore the traditional ML approaches for implementing on-device training. For this reason, this study mainly focused on various supervised ML methods that are implemented for the classification problem: Logistic Regression (LG), Random Forest (RF), KNN, NB - Gaussian, Linear SVM, MLP, and DT. Default parameters were used in most of the algorithms. Only in some cases any of the parameters were changed. For example, in KNN classifier, the number of Neighbors was set to three for simplicity. Those parameters will be maintained for the rest of the experiments performed in this dissertation. That is, no parameter tuning will be performed to improve classification results. Thus, obtained results are presented to compare the effect of the proposed strategy, not to get the best of any of the evaluated datasets.

In terms of hardware, a stand-alone Edge approach is proposed, where Edge devices are in charge of retraining the initial model and inferring new knowledge. This specific application aims to be part of an intelligent workplace that collects data from different sensors of the workers and analyses the gathered information to improve their habits in these spaces, taking into account the privacy of the collected information in all the stages of its curation.

For this reason, we evaluate four different platforms and compare them against a laptop computer that works as a reference device. The specifications of this laptop include an Intel Core i7-9750H processor with a base frequency of 2.60 GHz, equipped with 16 GB of RAM. For the IoT devices, one Nvidia Jetson Nano and three Raspberry Pi Foundation low-cost single board minicomputers were chosen: the Raspberry Pi 4 model B, the Raspberry Pi 3 model B+, and the Raspberry Pi Zero W. The Nvidia device includes a quad-core Cortex A-57 64-bit-based System on Chip (SoC) at 1.44 GHz and 2 Gb of RAM memory. In the case of the Raspberry foundation devices, the RPi4 also includes a quad-core ARM 64-bit processor, the A72 at 1.43 GHz in this case, with 4 Gb of Ram. These two first devices have higher computational capabilities than the Rpi 3 and the Rpi Zero. The former incorporates a quad-core A53 (ARMv8) 64-bit-based System on Chip (SoC) at 1.4GHz and 1 Gb of RAM. The latter (which is more limited in terms of hardware) includes a single-core ARM11 (ARMv6) 32-bits-based SoC at 1 GHz and 512 Mb of RAM. Even though Raspberry devices belong to the same family, judging by their characteristics and performance (in terms of processing power and memory), there are significant differences between them. In fact, taken into account that all the devices are compatible with the same software tools and, thus, they are evaluated under the same conditions, this work provides an unbiased comparison between several distinct enough devices.

In the following, we will base our analysis on the two more constrained devices (the Raspberry Pi 3 and Zero) in order to understand the potential performance optimization in those platforms with more limitations and that, in theory, have more difficulties to deal with demanding applications. Then, for the second part of this analysis, we will also apply the described strategy with the less limited devices (Rpi4 and Jetson Nano). For the sake of fairness, all experiments are executed solely on the CPU, with no other application running at the same time. It is also worth mentioning that those experiments do not take advantage of any additional processing units of the evaluated devices (e.g., the Maxwell GPU included in the Jetson Nano).

The classification solution software relies on the Python library Scikit-learn (Pedregosa et al., 2011), which includes efficient versions of the most common algorithms. The experimental process includes the training and the inference phase, as well as the validation of the models for each of the selected ML algorithms. The materials used for the evaluation experiments are publicly available, where additional details about the classifiers parameters or the performed pre-processing stages can be found [1].

[1] https://github.com/OihaneGomez/Exploring_Computational_Cost_ML_IoT

4.3 Understanding the potential of the strategy

In this section, we present an empirical demonstration of the effect of each of the factors previously described. In addition to their impact in terms of classification results and timing results, we present a selection of cost-accuracy optimal solutions obtained through the process of data analysis and reduction combined with a multi-objective optimization strategy. Thus, the results presented in this section can be divided into three different categories: the evaluation of the classification accuracy, the computational cost of ML techniques (measured as the necessary time to perform the classification process), and the trade-off between them. In both cases, the impact of the described factors (sampling frequency, number of signal components, number of features and the algorithm) are measured for each of the selected hardware platforms. An autonomous reduction of the initial data size based on this analysis is included to finish with this section. First, the initial demonstration is performed over the ADL dataset and using the Raspberry Pi 3 and Zero. Then, we extend the scope of this detailed analysis by applying the reduction strategy to two additional datasets (OHM Dataset and Sports Dataset) and including two more devices (Raspberry Pi 4 and Nvidia Jetson Nano).

4.3.1 Classification accuracy

This subsection evaluates the accuracy of the ML models according to the validation procedure explained in Figure 4.4. We perform this evaluation through a 5-fold-cross-validation procedure. Such validation was repeated 100 times to obtain more robust results.

Four metrics are used to measure the predictive performance of algorithms: *Precision* (Equation 4.2), which measures the true positives among all positive results; *Recall* (Equation 4.3), which computes correctly labeled positives based on all the correct positive and negative events; *F1-Score* (Equation 4.4), the harmonic mean of precision and recall metrics; and *Accuracy (Acc)* (Equation 4.5), the ratio of correct prediction over the total number of predictions.

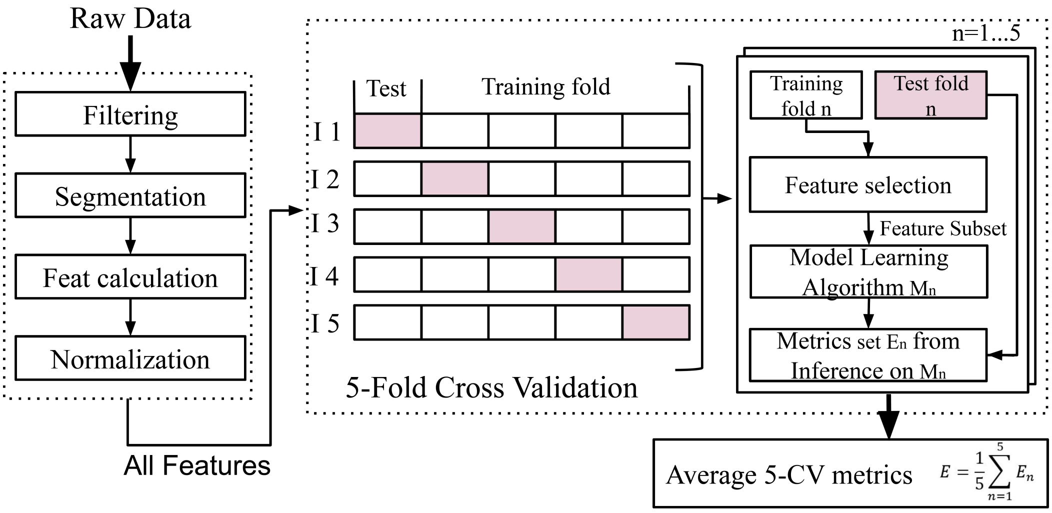

Figure 4.4: Schematic representation of the model training phase and its evaluation through a 5 cross-validation process.

Macro average results are provided for the first three metrics.

$$Precision = \frac{TruePositive}{TruePositive + FalseNegative} \tag{4.2}$$

$$Recall = \frac{TruePositive}{TruePositive + FalsePositive} \tag{4.3}$$

$$F1 - Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{4.4}$$

$$Acc = \frac{TruePositive + TrueNegative}{TruePositive + FalseNegative + TrueNegative + FalsePositive} \tag{4.5}$$

The first step in the optimization process is to quantify the influence of reducing the data input on the classification capabilities. Thus, we consider the three factors upon which the accuracy depends, and we evaluate them for each of the ML algorithms. First, the influence of the sampling frequency is analyzed by downsampling the original data to lower frequencies. For the number of components, we compare the performance of the models that use all

	All Features	10 Features	3 Features
32 Hz (Initial)	95.60 (SD=1.86)	94.83 (SD=3.72)	92.55 (SD=4.06)
16 Hz	95.60 (SD=1.28)	94.07 (SD=2.24)	88.66 (SD=2.90)
8 Hz	94.91 (SD=1.84)	93.52 (SD=2.32)	88.31 (SD=2.54)
4 Hz	94.87 (SD=1.73)	93.52 (SD=1.90)	86.45 (SD=3.48)

Table 4.3: Average F1 results for several sampling frequencies.

the accelerometer signal components (XYZ) with models that only receive the most representative one (X). This component was selected according to the ranking of features (see Table 4.2) that revealed X as the most representative component. Finally, we analyze the effects of the number of features on the classification output, comparing the results for all the original features set (162), a subset of the top 10, and the 3 top-ranked features.

Concerning the sampling frequency, Table 4.3 resumes the average F1 results for the sum of all the learning methods. These tables compared initial results with those obtained by reducing the original 32 Hz sampling rate to 16 Hz, 8 Hz and 4 Hz. It can be noted how lower frequencies decrease the classification performance but still maintain acceptable detection rates. In this particular application, this means that it is feasible to lower the frequency up to 4Hz. However, these results could be biased for the stationary nature of the target class (drinking) and may not be extrapolated to standard classification problems.

Thus, we have decided to cut off this reduction and keep 16 Hz as the reference downsampled ratio since this frequency belongs to the 15 Hz–20 Hz range, which is considered appropriate for HAR systems (Twomey et al., 2018). Therefore, in Tables 4.4 and 4.5 we summarize the performance of all the analyzed supervised ML approaches for the original sampling data (32 Hz) and the downsampled data (16 Hz). In both cases, the results for XYZ and X (the most representative component) are included. As can be observed, the classification rates yield average values over 90 % in almost all cases.

	All Components (XYZ)											
	All 162 Features				10 Features				3 Features			
	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy
LG	94.73	96.36	95.40	98.10	90.20	93.95	91.75	96.70	82.87	91.03	86.09	94.77
RF	95.66	98.14	96.78	98.70	96.19	96.89	96.46	98.53	93.18	94.54	93.72	97.41
KNN	97.78	96.40	**97.01**	98.72	98.07	96.70	**97.32**	98.85	95.61	96.33	**95.89**	98.29
NB	92.91	76.58	81.26	89.38	94.89	81.00	85.77	92.51	96.22	86.48	90.36	95.33
SVM	95.51	94.67	94.99	97.87	92.99	94.33	93.49	97.30	92.10	93.11	92.47	96.88
MLP	95.40	95.48	95.34	98.04	94.45	94.43	94.31	97.60	94.57	94.17	94.26	97.56
DT	94.62	93.89	94.10	97.48	95.31	96.15	95.62	98.18	93.19	92.82	92.84	96.97
	Most Representative Components (X)											
	All 54 Features				10 Features				3 Features			
	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy
LG	93.66	95.07	94.22	97.62	90.20	93.95	91.75	96.70	82.87	91.03	86.09	94.77
RF	96.05	97.70	96.78	98.68	96.19	96.89	96.46	98.53	93.18	94.54	93.72	97.41
KNN	97.33	96.85	**97.02**	98.74	98.07	96.70	**97.32**	98.85	95.61	96.33	**95.89**	98.29
NB	94.98	83.85	88.05	94.03	94.89	81.00	85.77	92.51	96.22	86.48	90.36	95.33
SVM	95.35	94.73	94.94	97.86	92.99	94.33	93.49	97.30	92.10	93.11	92.47	96.88
MLP	95.43	95.17	95.20	97.97	94.45	94.43	94.31	97.60	94.57	94.17	94.26	97.56
DT	94.74	94.84	94.68	97.76	95.31	96.15	95.62	98.18	93.19	92.82	92.84	96.97

Table 4.4: Performance (F1 macro results) of the supervised ML algorithms for the original sampling frequency (32Hz).

	All Components (XYZ)											
	All 162 Features				10 Features				3 Features			
	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy
LG	94.95	96.47	95.57	98.18	88.28	93.47	90.43	96.25	79.90	89.00	83.32	93.84
RF	95.84	97.89	96.76	98.68	95.71	96.24	95.89	98.29	88.75	90.77	89.53	95.72
KNN	97.77	95.99	**96.80**	98.62	97.35	96.48	**96.85**	98.66	90.02	89.91	89.78	95.69
NB	92.98	76.72	81.42	89.50	94.46	79.81	84.63	91.76	94.94	82.63	87.13	93.45
SVM	96.16	95.02	95.48	98.07	92.23	94.87	93.34	97.29	91.42	90.55	90.80	96.08
MLP	95.65	95.97	95.71	98.20	94.23	94.68	94.32	97.62	92.69	89.79	**91.03**	96.08
DT	93.49	93.40	93.29	97.17	93.39	94.03	93.56	97.31	87.48	87.95	87.48	94.76
	Most Representative Components (X)											
	All 54 Features				10 Features				3 Features			
	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy	Recall	Precision	F1	Accuracy
LG	93.66	95.46	94.42	97.71	88.28	93.47	90.43	96.25	79.90	89.00	83.32	93.84
RF	95.82	97.45	96.54	98.58	95.71	96.24	95.89	98.29	88.75	90.77	89.53	95.72
KNN	97.58	96.45	**96.94**	98.69	97.35	96.48	**96.85**	98.66	90.02	89.91	89.78	95.69
NB	95.11	84.37	88.46	94.27	94.46	79.81	84.63	91.76	94.94	82.63	87.13	93.45
SVM	93.68	93.26	93.33	97.18	92.23	94.87	93.34	97.29	91.42	90.55	90.80	96.08
MLP	94.29	94.68	94.35	97.64	94.23	94.68	94.32	97.62	92.69	89.79	**91.03**	96.08
DT	93.41	94.07	93.58	97.33	93.39	94.03	93.56	97.31	87.48	87.95	87.48	94.76

Table 4.5: Performance (F1 macro results) of the supervised ML algorithms for half the sampling frequency (16 Hz).

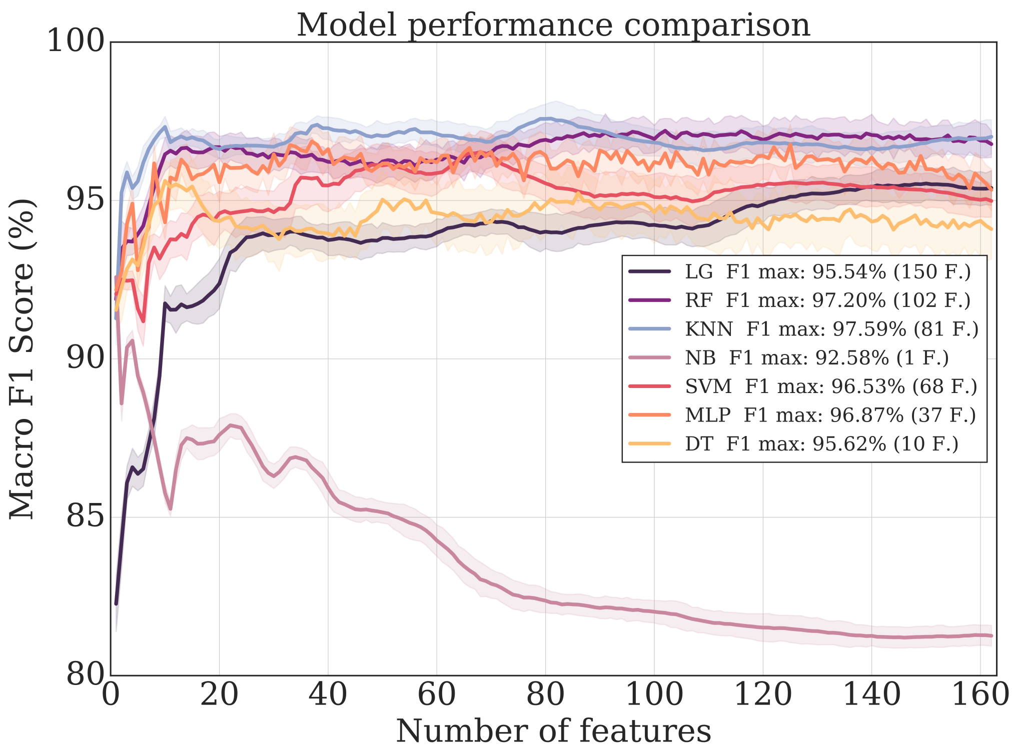

Figure 4.5: Model performance evolution, measured by the F1-Score, according to the selected number of features.

Figure 4.5 illustrates the influence of the number of features on the accuracy for the regular sampling. In particular, this figure highlights how the detection rates increase quickly at the beginning of the graph and reach values over 92 % with the first 20 features. This rapid increase indicates that a good balance between the number of characteristics and the classification accuracy can be obtained with few features irrespective of the classification method. Furthermore, as can be observed, adding more features to the models can improve the performance in some cases while decreasing it in others since the feature selection process is independent of the algorithm.

A thorough analysis shows that KNN performs slightly better than the other algorithms when the classification depends on all the 162 features. However, NB scores are worse for a larger number of features than for a small number of them. This can indicate that the features are not independent enough by themselves, something opposite to the inherent modelling assump-

	All Features	10 Features	3 Features	Decrease
XYZ	95.60	94.83	92.55	3.199 %
X	95.47	94.83	92.55	3.058 %
% Decrease	0.135 %	0 %	0 %	

Table 4.6: Average F1 results comparison for the original sampling (32 Hz). XYZ and the most representative component (X) are included.

	All Features	10 Features	3 Features	Decrease
XYZ	95.60	94.07	88.66	7.259%
X	94.86	94.07	88.66	6.535 %
% Decrease	0.774 %	0 %	0 %	

Table 4.7: Average F1 results comparison for half the original frequency (16 Hz). XYZ and the most representative component (X) are included.

tion of NB, which assumes that the presence of a specific feature is entirely unrelated to the presence of any other (Domingos and Pazzani, 1997). For this reason, we still include it for the analysis of the cost-accuracy trade-off, but we consider it an outlier for average calculations of Tables 4.6 and 4.7.

In these Table 4.6 and Table 4.7, we show the final impact of every of the evaluated factors compared to the baseline performance (original sampling frequency, XYZ signal components, and 162 features). Those tables also include the total percentage decrease when comparing the difference between all the signal components and the most representative component. The percentage decrease measures the percentage change between two values, according to Equation 4.6:

$$\%D = \frac{InitialValue - FinalValue}{InitialValue} * 100 \tag{4.6}$$

In this case, $\%D$ measures the relative decrease of each of the proposed approach's results ($FinalValue$) when compared against the reference model ($InitialValue$).

These results indicate minor differences when decreasing sampling frequency to 16 Hz. Moreover, the classification results are not drastically penalized when the number of features is reduced to 10 and even 3. However, decreasing the number of features has a more significant impact on lower sampling frequencies. Using only the top 3 features for the original sampling frequency penalized the results in a 3.19 %. For half of the sampling rate, the decrease was 7.259 %. Furthermore, considering only the most representative component of the signal (X) does not affect the overall results significantly: 0.135 % for the regular sampling and 0.774 % for the downsampled data. The results are the same when considering only the top 10 and top 3 features in classification. This means that only X-related features are on this top

4.3.2 Computational cost

This subsection evaluates how the design parameters affect the model in its two stages: training and inference. Figure 4.6 represent the experimental sequence in both cases. We measure the computational cost of each process as the necessary time to accomplish all the tasks it encompasses.

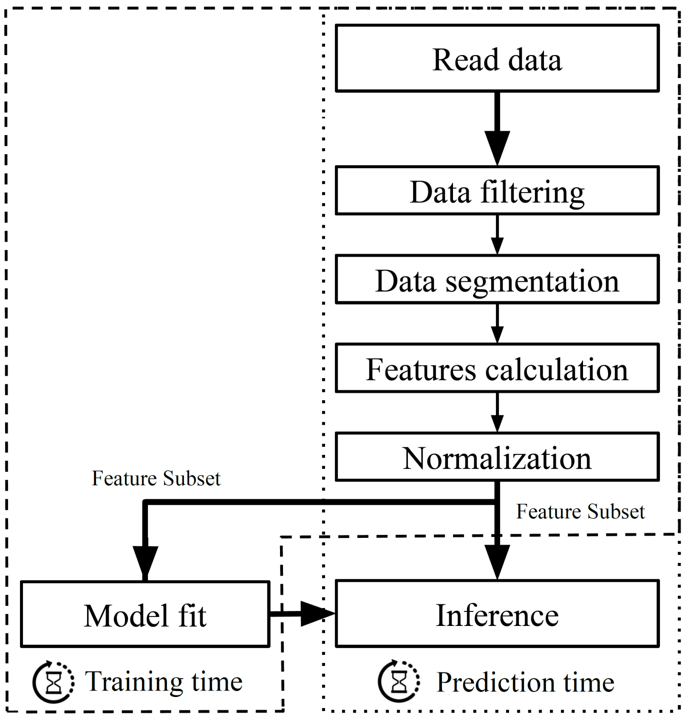

Figure 4.6: Representation of the training and inference phases.

4.3.2.1 Training process

Training a ML model corresponds to one of the most computationally intensive processes within an activity recognition application. The main reason is that it involves processing large amounts of data. The training process can be divided into two parts, the common pre-processing steps applied to the whole dataset and the algorithm-dependent model fitting task. Table 4.8 presents the elapsed time for the initial pre-processing stage when all the dataset information is read and the features are calculated. These results were obtained after repeating the process at least 15 times (3 runs of 5 loops) and calculating the average running times. They are also displayed in Figure 4.7, which compares the evolution of the processing time according to the number of features.

At first sight, it can be noted how the computational time shows a quadratic growth when the data input (the number of features) increases. This is particularly relevant for the scalability of the classification system and may heavily penalize the performance of less powerful devices. For this reason, a more in-depth analysis of the percentage reduction shows that more significant performance improvements are obtained in those platforms. As expected, the laptop computer outperformed the other devices in absolute terms. In relative

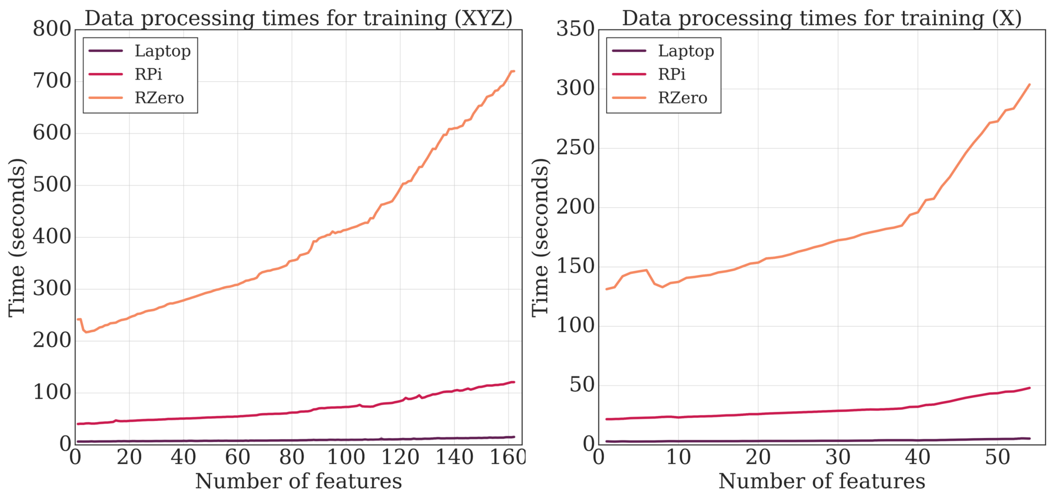

Figure 4.7: Evolution of the elapsed time for processing all the instances of the dataset for training.

	XYZ			X		
	Laptop	Rpi	Rzero	Laptop	Rpi	Rzero
All Feat.	$16s$	$2min\ 5s$	$12min\ 32s$	$5.3s$	$52s$	$5min\ 5s$
10 Feat.	$7.1s$	$43.6s$	$4min\ 6s$	$3.1s$	$25s$	$2min\ 18s$
3 Feat.	$7s$	$41.4s$	$3min\ 56s$	$3.0s$	$22.5s$	$2min\ 9s$
% Decrease	56.32 %	66.94 %	68.61 %	44.05 %	56.78 %	57.73 %

Table 4.8: Computational cost of processing all the dataset for the training process using the original sampling frequency (32 Hz).

	XYZ			X		
	Laptop	Rpi	Rzero	Laptop	Rpi	Rzero
All Feat.	$15.98s$	$2min\ 5s$	$12min\ 19s$	$5.38s$	$49.34s$	$5min10s$
10 Feat.	$7.72s$	$42.92s$	$4min5s$	$3.13s$	$24.44s$	$2min\ 18s$
3 Feat.	$7.43s$	$40.79s$	$3min\ 44s$	$3.01s$	$22.1s$	$2min\ 8s$

Table 4.9: Computational cost of processing all the dataset for the training process and half the original sampling frequency (16 Hz).

terms, the laptop decreased its running times by 56.32 % for XYZ, while the same data reduction achieved a decrease of 68.61 % on the execution times for the Raspberry Zero platform. Moreover, from the baseline parameters (all features and all the signal components) to the final simplified stage (3 features and only the most representative component) the time reduction was 81.24 % for the laptop (from 16 seconds to 3 seconds), 82.05 % for the Raspberry Pi (from 2 min and 5 seconds to 22.5 seconds) and 92.32 % for the Raspberry Zero (from 12 min and 32 seconds to 2 min and 9 seconds).

The effect of the remaining factor, the sampling frequency, is evaluated in Table 4.9, which includes a comparison of the pre-processing times for 16 Hz. In this case, contrary to what intuitively could be expected, a lower sampling frequency did not substantially affect the time measurements. A more detailed analysis shows that only minor differences are found for the largest number of features compared to Table 4.8. In some cases, the performance was even slightly worse because of the non-deterministic nature of the timing measurements. This lack of improvement can be a consequence of the baseline resource overhead associated with data management. Another possible reason is the length of the data window, which is not large enough to have noticeable improvements when it is reduced.

Therefore, in our evaluated case of study, this may indicate that the number of computations (e.g., how many characteristics of the signal are obtained) has a greater impact on the pre-processing stage than the effect of the data length. However, this particular conclusion cannot be generalized to similar applications. Thus, it would be necessary to look at the effect of the sampling frequency in applications with larger data sets, longer data streams, more complex filtering stages or frequency-domain features. In the following, we will continue the analysis using only the original sampling frequency (32 Hz) since the model created with the downsampled data is more sensitive to decreasing the number of features.

The second step of the training stage is to feed the model with the pre-processed data for the fitting task. This task depends on the different classifiers and their parameters, together with the amount of data used to build it. Figure 4.8 represents the average computational cost of each algorithm for the model fitting task after executing it 1000 times (100 runs of 10 loops). Furthermore, it shows the time impact caused by creating the model with an increasing number of features. RF and MLP techniques are excluded from these plots since their duration, significantly greater than those represented, compromise the visibility of the figure. Instead, their results are presented in Table 4.10 and Table 4.11.

For MLP, increasing the number of features can affect positively on the timing results. In this case, from 3 to 54 features, the efficiency of the fitting task was increased. As the MLP is iteratively trained, depending on how fast it converges, the time to fit the data can vary. Consequently, a larger data input size may then improve the fitting process of the model.

The main conclusion that can be extracted from those results is that fitting task times are significantly lower than the time needed to process all the initial data and create the input for the model. Contrary to the pre-processing stage, the algorithm and its internal parameters play an important role in creating the model. In general terms, the fitting task shows a behavior close to

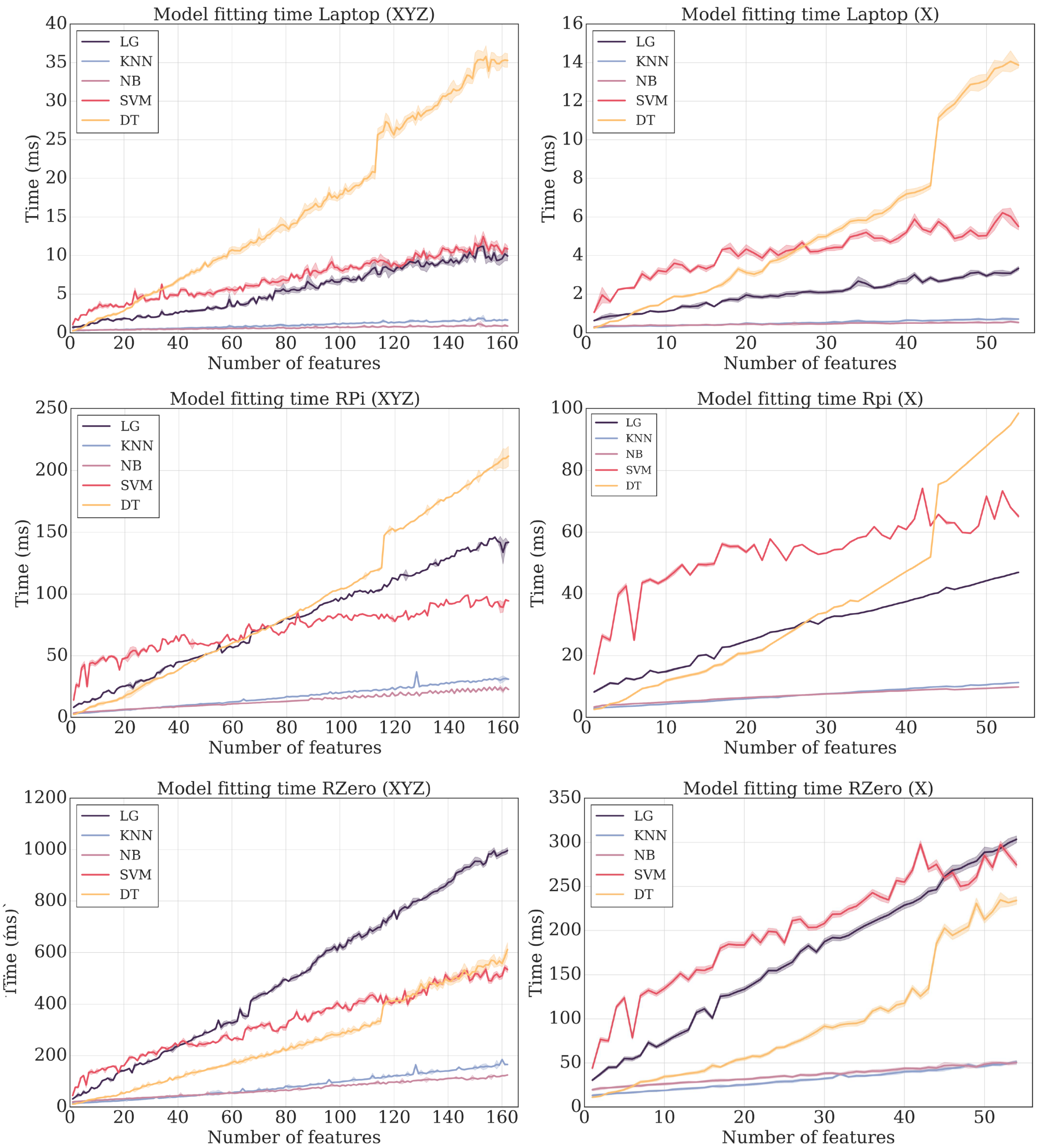

Figure 4.8: Comparative between the number of features and the elapsed time for fitting the model. MLP and RF are out of this representation since they have a significantly greater duration than those included.

linear when increasing the input size. This involves a baseline computational complexity of $O(nm)$, being m is the number of instances and n the number of features. For this reason, with a small number of features, all the algorithms last similarly, but when increasing the number of features, the computational cost is augmented.

	Laptop			Raspberry Pi			Raspberry Zero		
Features	162	10	3	162	10	3	162	10	3
RF	$181.9ms$	$96.4ms$	$80.5ms$	$1.6s$	$1.1s$	$970ms$	$7.8s$	$6.2s$	$6.1s$
MLP	$782.3ms$	$456.9ms$	$622.2ms$	$3.2s$	$2.3s$	$2.4s$	$20.2s$	$14.9s$	$17.6s$

Table 4.10: Elapsed time fitting the model for MLP and RF considering XYZ.

	Laptop			Raspberry Pi			Raspberry Zero		
Features	54	10	3	54	10	3	54	10	3
RF	$136.0ms$	$96.3ms$	$80.6ms$	$1.3s$	$1.0s$	$991.3ms$	$6.8s$	$6.5s$	$6.3s$
MLP	$19.2ms$	$452.1ms$	$624.3ms$	$2.2s$	$2.5s$	$2.4s$	$13.2s$	$14.3s$	$15.7s$

Table 4.11: Elapsed time fitting the model for MLP and RF considering X.

On top of that, the increments in time (the slope) of each learning algorithm depend on the complexity of the learning method itself or the different parameters and estimators that can be adjusted on the fitting phase. For this reason, the complexity of the ML techniques can be reduced by modifying those fitting parameters to simplify the model but decreasing its accuracy in return. Still, considerable improvements in the performance can be obtained when adjusting the internal values that affect the data model. This is precisely the case for RF and MLP, which are the most resource-demanding algorithms in this comparison.

4.3.2.2 Prediction process

The prediction process is tested to evaluate how long it takes to read and process a data sequence and compare it with the already trained model. The experiment consists of measuring the elapsed time in the process of classifying a sample of 14 instances of the dataset. Therefore, this test resembles an online classification system where a new sequence of data needs to be predicted without an entailed excessive processing time. Again, the results are divided into two parts, the common process of reading and processing the new data and the inference task.

Table 4.12 shows the computational cost of reading and processing the 14 instances of new data for a variable number of features, computed after

	XYZ			Most representative component		
	Laptop	Rpi	Rzero	Laptop	Rpi	Rzero
All Features	$317ms$	$2.62s$	$15.9s$	$125ms$	$1.07s$	$6.57s$
10 Features	$140ms$	$838ms$	$5.48s$	$71.2ms$	$557ms$	$3.61s$
3 Features	$134ms$	$781ms$	$5.06s$	$62.5ms$	$500ms$	$3.24s$
% Decrease	58.04%	70.20%	68.25%	49.19%	53.14%	50.65%

Table 4.12: Time needed to process 14 instances of new data for the prediction process.

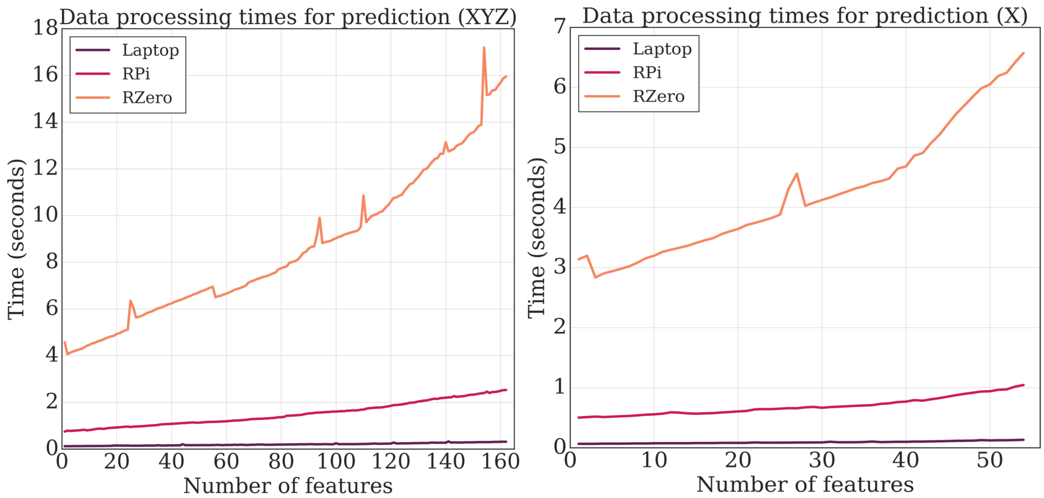

Figure 4.9: Evolution of the computational cost of processing 14 instances of the dataset for inference.

executing the experiment 30 times (10 runs of 3 loops). Again, it illustrates how processing times benefit from the feature reduction step and the selection of the most representative components of the signal. This is especially relevant in the case of the Raspberry Pi and Zero. The total time reduction is 80.28 % for the Laptop PC, 80.93 % for the Raspberry Pi and 97.96 % for the Raspberry Zero. In particular, the Raspberry Zero reduced its total processing times from around 16 seconds to 3 seconds. This improvement is very noticeable and makes a difference when designing online classification systems for low-powered devices. Figure 4.9 reflects the computational cost tendency that shows a quadratic growth again.

These slopes are also affected by the platform on which this training pro-

	Laptop			Raspberry Pi			Raspberry Zero		
Feat.	162	10	3	54	10	3	54	10	3
RF	$5.6ms$	$5.5ms$	$5.5ms$	$46.1ms$	$39.9ms$	$39.4ms$	$261.5ms$	$260.9ms$	$260.7ms$
KNN	$2.7ms$	$0.7ms$	$0.6ms$	$23.9ms$	$3.7ms$	$2.6ms$	$100.4ms$	$15.9ms$	$13.4ms$

Table 4.13: Prediction times for RF and KNN considering XYZ.

	Laptop			Raspberry Pi			Raspberry Zero		
Feat.	54	10	3	54	10	3	54	10	3
RF	$5.6ms$	$5.6ms$	$5.5ms$	$39.6ms$	$39.0ms$	$38.7ms$	$279.8ms$	$279.0ms$	$278.8ms$
KNN	$1.2ms$	$0.7ms$	$0.6ms$	$7.8ms$	$4.1ms$	$2.6ms$	$41.3ms$	$17.2ms$	$14.5ms$

Table 4.14: Prediction times for RF and KNN considering X.

cess is carried out. In general, under 20 features, all the algorithms' times are within the first segment of the Y-axis for the three platforms, being 200 ms the worst case for the Raspberry Pi Zero W.

Compared to pre-processing and training, the inference task is several orders of magnitude faster on all platforms (see Figure 4.10). The prediction task needs less than 0.3 ms for the laptop device, 1.4 ms for the Raspberry Pi, and less than 8 ms for the Raspberry Pi Zero (mean values of 1000 runs). RF and KNN are out of this representation. Instead, Table 4.13 and Table 4.14 include the results of both algorithms.

In this case, reducing the number of features does not significantly affect the time required for inference. Few differences can be found among all algorithms in terms of performance, except the mentioned RF and KNN that took a slightly longer time than the rest of them to compare the new data with the model. Furthermore, Naive and SVM times show a steady increase, while the rest of the algorithms show a sustained response. This is a consequence of the constant computational complexity of the prediction task in some of the evaluated algorithms, which are less sensitive to the data input size. Nevertheless, the number of features is still a relevant factor for the prediction phase according to the computational cost of pre-processing new data.

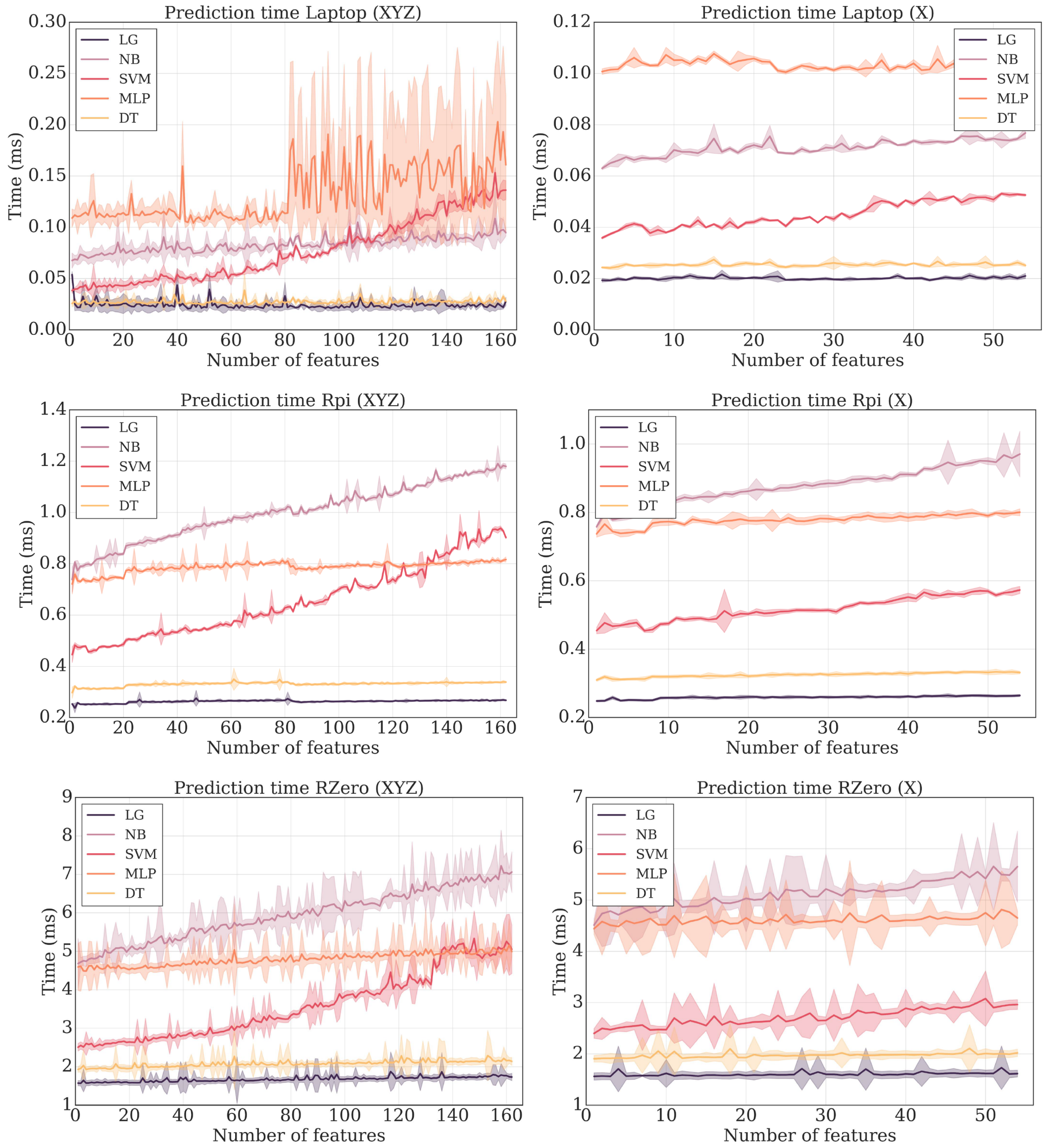

Figure 4.10: Comparison between the number of features and the elapsed time for predicting a sequence of new data. KNN and RF are out of this representation since they have a significantly greater duration than those included.

4.3.2.3 Cost-accuracy trade-off: a Pareto approach

The correlation between the model size and the accuracy studied in this section leads us to draw an initial conclusion: it is feasible to maintain detection rates over a 90 % when the data input is reduced. In our case, the studied simplification strategy reduced the data input size to a small subset of one representative component and its top-10 features, which outperforms the baseline substantially. Moreover, the time improvements explored in the previous subsection produce an enhanced performance in most of the classification tasks when the data input size is reduced, regardless of the devices' resources. This allows bringing more complex applications to a local stage that could not be executed in the Edge otherwise. Based on the obtained results, it is expected that, in some applications, optimizing the early stages of the pipeline can improve the performance of the system more than the specific learning method.

From that point, further evaluation can optimize the trade-off between classification accuracy and computational cost. To that end, we conduct a Pareto multi-objective optimization analysis to find the small-sized subset of solutions with the best performance (Qian et al., 2015). Pareto analysis is a statistical tool for decision-making that selects a limited number of Pareto-optimal solution that illustrates the best trade-off between two objectives, cost and accuracy in our case. Since we seek to evaluate the efficiency of the model for both training and inference stages, we include the role of each of these stages through the train/inference ratio (explained in section 4.1) in the Pareto analysis.

Figure 4.11 represents the relationship between the total cost (the sum of the prediction and training times without considering data processing) and the accuracy of the system (F1 values). Every point corresponds to a certain number of features for each of the evaluated algorithms running on the laptop device. This figure also illustrates the effect of this weight correction on the total cost when different train/inference ratios are applied. This ratio is a design choice that should be made according to the context of the classification problem. As explained in Section 4.2, it gives more importance to the predominant task and sets the frequency in which a model may be retrained compared to the estimated number of times new data would be classified. We have selected a weight correction of r=1 (baseline ratio, no weight correction is applied), r=100, r=1000 and r=10000 to show the differences for comparative purposes. This ratio means that the training stage may occur once every 100 or 1000 times or, in other words, giving 100, 1000 or even 10000 times more

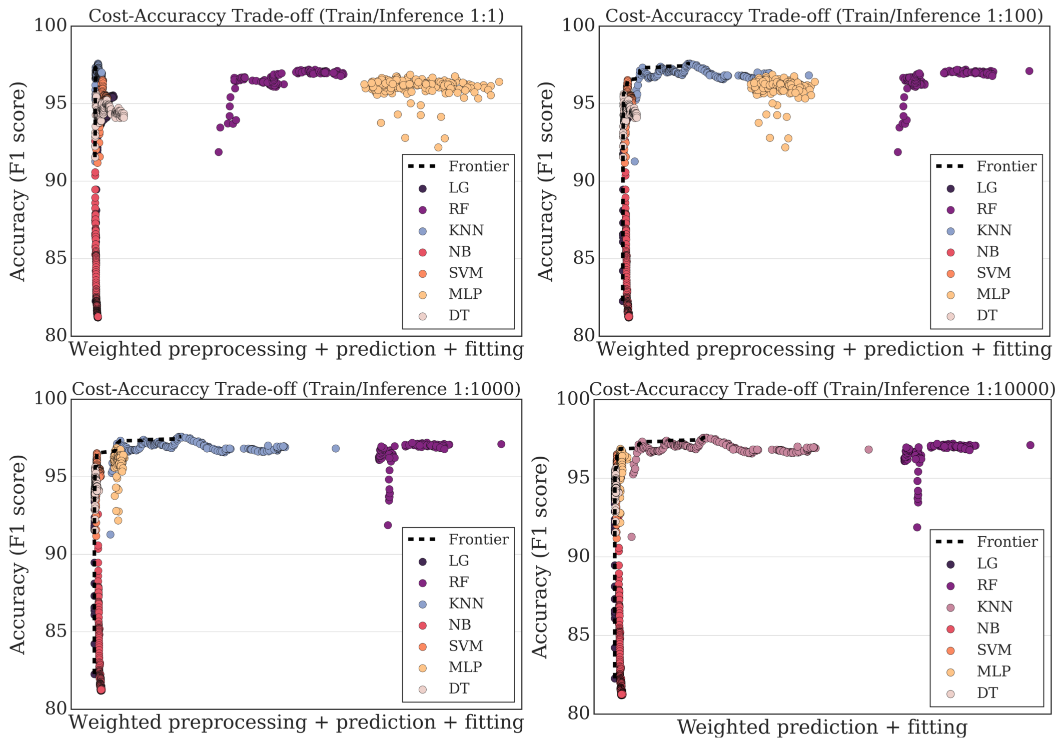

Figure 4.11: Cost-accuracy trade-off for the balanced computational cost between prediction and fitting times (upper left) and applying the train inference ratio (rest of the plots). The Pareto frontier represents the most optimal solutions.

importance to the prediction stage over the training phase. As we explained in initial sections, the weighted computational cost c_{total} is calculated with Equation 4.7:

$$c_{total} = r_t * c_{training} + (1 - r_t) * c_{inference} \tag{4.7}$$

This figure also includes the Pareto frontier, representing the most optimal solutions and marking the points with better accuracy (highest y-axis values) and lower computational costs (smaller values on the x-axis).

The train/inference ratio is particularly relevant for those algorithms with an unbalance performance between training and inference tasks, such as KNN and MLP. The former has a great performance on training (See Figure 4.8) while the inference times are worse than the average (Tables 4.13 and 4.14).

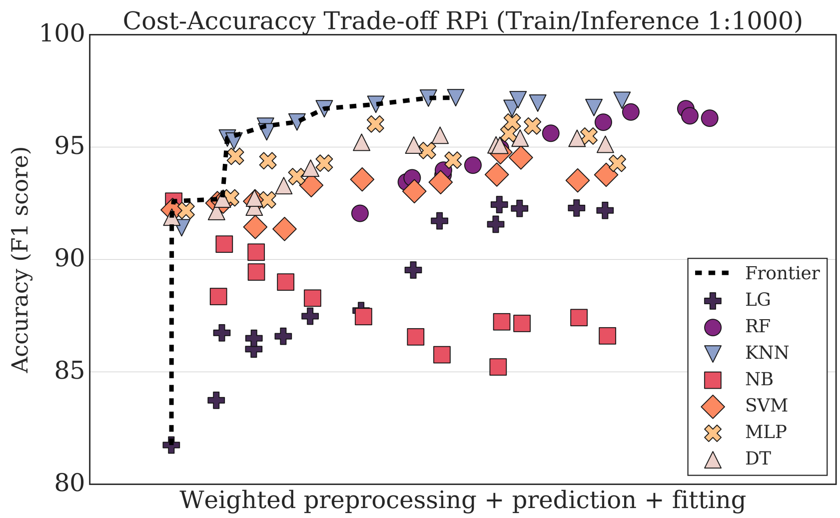

Figure 4.12: Cost-accuracy trade-off and Pareto optimal points for the reduced data input achieved during the optimization process on the RP3 device.

MLP, on the contrary, performs well for inference (Figure 4.10), but its complexity demands higher times to create the model (Table 4.10 and Table 4.11). As the train/inference ratio grows, we can observe how KNN is penalized, increasing its total weighted computational cost, and MLP moves to the left, closer to the Pareto frontier. Higher train/inference ratio values would increase this difference in those applications where training times are considerable compared to the inference time (e.g., when dealing with larger datasets).

Figure 4.12 shows the Pareto analysis performed only for the final reduced solution (top-10 features and X component) when applying a 1:1000 ratio and for the Raspberry Pi 3 device. In this case, the computational cost also included the pre-processing times for training and inference. Again, the dashed line highlights the Pareto optimal points, where the small set of options that have the most efficient balance between the classification rates and the computational cost are included.

Finally, Table 4.15 includes the values associated with each of these points, where total times are the sum of data processing for training and inference together with the fitting or prediction tasks times. The final step in the optimization process is to select the one that better meets the initial design requirements from this selection, either prioritizing accuracy or performance.

Alg	N Features	F1	Total inference time (s)	Total training time (s)
LG	1	81.73	0.503349	21.680870
NB	1	92.58	0.503824	21.676447
SVM	1	92.20	0.503571	21.686039
DT	1	91.88	0.503432	21.675802
KNN	2	95.40	0.514883	21.811192
KNN	3	95.93	0.522636	22.030610
DT	4	92.69	0.513096	22.578788
KNN	6	96.11	0.528342	22.926632
KNN	7	96.70	0.533877	23.088214
KNN	8	96.90	0.544137	23.593250
KNN	9	97.18	0.554919	23.781464
KNN	10	97.19	0.561269	23.153416

Table 4.15: Final optimal Pareto solutions with the best cost-accuracy trade-off for the Raspberry Pi 3.

4.3.3 Applying the optimization strategy

Throughout this section, we have evaluated how a comprehensive reduction of the initial data input can lead to a more optimized cost-accuracy trade-off. The steps for this reduction involve making a detailed analysis of each of the factors described in Section 4.1 and checking their influence on the classification results. Beyond the exploratory analysis studied, now we apply this strategy to see the potential optimizations that can be obtained in the rest of the datasets and devices included in Section 4.2. That is, OHM Dataset, Sports Dataset, and Raspberry Pi 4 and Nvidia Jetson Nano for the hardware devices. In this part of the work, the OHM dataset is considered a binary dataset in which only ”Drink” and ”Other” classes are targeted.

In this case, we base this part of the study on obtaining the reduced search space that leads to selecting the most optimal solutions through the Pareto strategy. For that, we will apply an automatic analysis of the effect that each of the factors has on the classification results. In order to perform this analysis, it is necessary to set a pre-set requirement, which corresponds to the maximum accuracy loss that can be considered acceptable for the classification task. Then, considering this constraint, the optimization pipeline checks for each factor the minimum selection of elements (sampling frequency, number

	OHM DATASET	SPORTS DATSET
Initial Data	9 signals 486 Features	18 signals 162 Features
Reduced Data	3 Signals 10 Features	3 Signals 20 Features

Table 4.16: The number of signals and features for the initial and reduced data input of each of the datasets evaluated inf this part of the chapter.

of components/signals, and number of features) that meets the established requirement.

We evaluate the optimization potential of this preliminary reduction in the most limiting scenario, the data processing task in the training phase. For these experiments, we set the requirement of maintaining the classification results above 90 points in at least one of the potential optimal solutions. However, this requirement could be extended to other situations (e.g., that all the ML algorithms are above this threshold). Having applied this automatic optimization pipeline, Table 4.16 includes the final number of signals and features that characterize the reduced subset of options obtained with this automatic strategy.

Table 4.17 includes the F1-macro classification results for both the initial and the reduced data input. The OHM dataset decreased its classification capabilities by a 4.08 % on average, compared to the 5.80 % decrease percentage of the Sports dataset. These results illustrate that, for the latter, the penalty for including a smaller input data size is more significant, with more than 5 points of difference between the average initial and reduced results. On the contrary, the OHM Dataset's room for improvement is broader, in such a way that only 3.74 points are lost when improving its performance. In both cases, these differences are more significant than those found in the analysis of binary classification scenario with the ADL dataset, in which it was possible to reduce the initial data to a smaller number of features and signals (1 signal and 3 features) with a 3.18 % decrease percentage.

Based on this reduction, the computational time of processing this reduced data input is measured. For comparative purposes, we use the time results

	OHM DATASET		SPORTS DATASET	
	Initial	Reduced	Initial	Reduced
LG	93.88 (SD 0.40)	84.52 (SD 0.34)	97.64 (SD 0.03)	87.99 (SD 0.20)
RF	94.14 (SD 0.21)	91.14 (SD 0.25)	99.13 (SD 0.06)	96.79 (SD 0.15)
KNN	90.15 (SD 0.32)	90.50 (SD 0.56)	97.78 (SD 0.05)	96.53 (SD 0.15)
NB	83.09 (SD 0.27)	80.23 (SD 0.26)	93.61 (SD 0.13)	79.45 (SD 0.47)
SVM	93.86 (SD 0.55)	89.16 (SD 0.47)	98.37 (SD 0.08)	94.73 (SD 0.09)
MLP	94.81 (SD 0.49)	90.97 (SD 0.28)	98.02 (SD 0.14)	93.93 (SD 0.27)
DT	90.35 (SD 0.73)	87.62 (SD 0.60)	96.92 (SD 0.16)	92.52 (SD 0.22)
MEAN	91.47 (SD 4.14)	87.73 (SD 4.05)	97.35 (SD 1.78)	91.71 (SD 6.16)
% D	**4.08 %**		**5.80 %**	

Table 4.17: Performance (F1-macro results) of the supervised ML algorithms for the initial and reduced data input.

concerning the most limiting task (i.e., training data processing times) as a reference to measure the potential time improvements. Even though code optimization is out of the scope of this study, when possible (i.e., when the number of cores allows it), a multi-processing approach has been followed for parallelizing the data processing tasks and compensate for the increased complexity of these datasets.

Table 4.18 includes the obtained results. As can be observed, newly added platforms, Raspberry Pi 4 and Jetson Nano, still benefit from the optimization strategy proposed. All platforms except for the RZero (that has a single-core processor) took advantage of all the available cores of their respective CPUs for these experiments. In fact, this platform was unable to process the initial data for the Sports dataset as the device run out of memory during this task. Thanks to the optimization process, it was possible to perform it in 44 minutes. On average, these platforms obtained a decrease percentage of 91.94 % of the OHM Dataset and 82.48 % for the Sports Dataset. These results are in line with those obtained in the previous analysis.

When comparing those performance improvements of Table 4.18 against the F1 difference included in Table 4.17, it becomes evident how the OHM Dataset provides a greater optimization room. This dataset also shows a smaller detriment in its results than Sports Dataset when the input data size

		OHM Dataset	% D	Sports Dataset	% D
Laptop	Reduced Initial	$0.95s$ $14.5s$	93.44	$16.7s$ $1min\ 32s$	81.84
Jetson Nano	Reduced Initial	$6.86s$ $1min\ 19s$	91.31	$1min\ 8s$ $6min\ 52s$	83.49
Rpi 4	Reduced Initial	$9.5s$ $1min\ 32s$	89.67	$1min\ 33s$ $9min\ 20s$	82.11
Rpi 3	Reduced Initial	$18.3s$ $2min\ 59s$	89.77	$3min\ 4s$ $18min\ 4s$	83.02
RPi Zero	Reduced Initial	$3min\ 30s$ $55min\ 57s$	95.53	$44min\ 34s$ *	*
Mean			**91,94%**		82.48%

** Raspberry Pi Zero device reported an out-of-memory error in this experiment*

Table 4.18: Data processing times for the data processing task of training phase of the classification pipeline.

is reduced. The OHM dataset also benefits from the higher number of initial features available, making it easier to reduce its number if some of these features are not very significant for the classification model. However, it can also be observed how the trade-off between cost and accuracy becomes less and less beneficial when dealing with the optimization of complex classification systems. For the multi-class Sports dataset, reducing the computational time to the same extent as in other evaluated datasets involves an accuracy loss close to 6 %. This drop in the classification results is almost twice as high as the one for the binary ADL Dataset. In this sense, complex datasets, as multi-class ones, impair the optimization capabilities of this approach and the cost-accuracy trade-off.

Even considering the specific particularities of the different classification applications, the obtained results highlights the importance of finding the right balance between the computational cost and the final accuracy of the classification model. To that end, this analysis seeks to assist in providing a strategy for optimizing similar problems and bringing advanced processing capabilities to the Edge of the network.

4.3.4 Limitations

Before concluding based on the experimental results, we want to address the possible limitations of this study. The main goal of this work is to analyze the different factors involved in the ML pipeline and optimize their influence on the cost-accuracy trade-off. Thus, the significance of this work relies on the comparison of the *relative* results obtained when each of these factors is evaluated. Additional elements affecting the *absolute* results such as the ML framework, the programming language, the libraries' version or the hyper-parameter optimization are out of the scope of this comparative analysis. For example, the way the different tasks are implemented may have an impact on the execution of the system. For this reason, optimizing general characteristics of the code, such as the way the information is processed (e.g., vectorizing computations over data structures (Walt et al., 2011)) can improve the general performance of the system. Moreover, a specific multiprocessing scenario may also influence the final results and performance of the system. We take this into account in the evaluated solution, but we do not benchmark code optimization mechanisms.

Besides, there are other aspects in the pipeline, such as the data filtering method or the feature selection algorithms; the comparison of different strategies for those aspects is out of the scope of this work. Therefore, we explain the decision criterion to choose them. However, we do not delve into the analysis of the consequences that these decisions may have in the classification results. We note that it would be interesting for future work to evaluate the potential improvements that alternative strategies could entail.

Finally, we evaluated the computational cost in terms of the elapsed time for each of the different stages of the classification problem. Despite running these tasks in isolation, they can be sensitive to internal processes or routines of the device and hence be non-deterministic. To overcome this issue, we took a large number of measurement samples and provided average results based on a large number of iterations.

4.4 Summary and Conclusions

This chapter has presented a methodology to meet a satisfactory cost-accuracy trade-off on the execution of the whole pipeline of ML techniques. This is motivated by the challenge of improving the feasibility of resource-constrained devices to perform inference and training tasks at the Edge, addressing data privacy concerns and contributing to a more human-centric IoT. In our vision of a more human-centric IoT, the privacy of the obtained information is preserved by ensuring that none of the classification tasks needs to be outsourced, neither inference nor training. Furthermore, since the data is kept in a local environment close to the user, users can retain the control over their data. Therefore, with a localized approach, both privacy and control of the captured information are enhanced, fulfilling the requirements for increased technological perception in intelligent workplaces identified in Chapter 3.

The presented case of study deserves attention to illustrate the potential impact of the optimization strategy that addresses this piece of research. In its evaluation, the experimental results suggest that, in some applications, the specific processing times for each algorithm can be practically negligible concerning the time required to capture the data and process it. This emphasizes the relevance of understanding the characteristics of the data and its impact on the performance and cost, particularly for the training phases. We stress that highlighting that the selection of the ML algorithm is not the only important factor to consider in the classification process at the Edge. Analyzing the initial data and optimizing the process from the early stages of the classification pipeline is vital to meet the cost-accuracy requirements.

The results presented in this chapter show that there is a highly non-linear trade-off to make between the computational cost, in terms of processing time, and the achieved classification accuracy. This highlights the optimization opportunities that understanding those factors provides when this cost-accuracy trade-off can be considered. As observed in Section 4.3 it was possible to substantially reduce the baseline processing time of the ML techniques while still obtaining acceptable detection rates. For the initial case of study, the proposed strategy has been able to reduce by more than 80 % the baseline processing time of the ML techniques in all the evaluated platforms. Furthermore, acceptable detection rates are maintained, assuming a decline of classification accuracy of only 3 %. However, complementary experiments illustrate that as the complexity of the classification problem increases the room for improvement becomes smaller. Thus, this trade-off becomes less beneficial. For those scenarios, in the following, we will complement this approach with a new method that seeks to compensate for the potentials errors of having a unique reduced model. With this part of the work and the following chapter, we continue studying optimizations techniques that simplify the complexity of ML techniques, intending to provide the necessary tools for bringing advanced processing capabilities to the Edge of the network instead of the Cloud, avoiding extra threats on compromising private data.

Simplicity is the ultimate sophistication.

Leonardo do Vinci

CHAPTER

5

Ensemble learning techniques for simplifying computation

BLANCING the computational cost and the accuracy of ML techniques being executed upon IoT is highly dependent on the complexity of the classification problem to optimize. As observed in the previous chapter, in some contexts, having a single reduced model may not be sufficient to maintain acceptable detection rates efficiently in Edge-based intelligent systems. For this reason, in this chapter, we present an optimization alternative to overcome the potential limitations of the previous approach when dealing with more complex classification problems (e.g., multi-class classification problems) in constraint settings.

Therefore, this chapter will explore an ensemble learning strategy to ease the optimization and distribution of the burden on data processing tasks. In particular, a discriminative model cascade is implemented. In this approach, complex ML systems are divided into successive layers of simpler models that filter input elements at each stage of the cascade based on the confidence of the prediction at that level. Three variations of the training setup for the cascade strategy are presented and analyzed. The balance of this strategy each of the cascade variations in terms of accuracy and potential computational savings is explored. Besides, the suitability of some of the most common classic ML algorithms to be part of a model cascade, and the final performance of the system are empirically evaluated. This chapter complements the optimization strategy presented in Chapter 4. Thus, it aims to facilitate the integration of a broader range of HAR application in a privacy-preserving local stage, regardless of their difficulty. Furthermore, due to the interest in comparing the proposed methods in HAR contexts, the proposal of this chapter will continue being evaluated for action classification in several multi-class datasets.

The chapter is divided into several sections, starting from Section 5.1, which presents the model cascade approach and includes three proposals to leverage the outcome of this strategy. Section 5.2 shows the experimental methodology, similar to that of Chapter 4, followed for the evaluation of this proposal across the rest of the chapter. The obtained results are shown in Section 5.3. Finally, the conclusions of the chapter are presented in Section 5.4.

5.1 The model cascade optimization approach

As has been introduced during the previous chapters, in this dissertation, we are studying how privacy concerns and data control requirements can be satisfied through an Edge Computing approach. In the previous chapter, we have presented new mechanisms to improve the feasibility of resource-constrained devices to perform inference and training tasks at the Edge and keep data locally and closer to the user. To that end, we embraced a holistic approach based on understanding the different factors that take part in the whole classification process (from data acquisition to the final fitting and prediction stages) to improve the performance of the whole classification system by reducing the complexity of the final model. However, new challenges arise when using a single reduced model is insufficient to meet a good trade-off between the system's accuracy and performance.

For this reason, in this chapter, we propose a complementary approach to perform more complex classification task more efficiently. In particular, we describe a multistage classification approach for constrained contexts. This multistage scheme implements an adaptive ensemble learning technique through a discriminative model cascade. This way, the model cascade adapts the resources to be used according to the task's difficulty. Contrary to the previous proposal, complex ML systems are divided into successive layers of simpler models (equivalent to the ones used in the previous chapter). Thus, instead of having a single model that classifies all the instances of new data, an ensemble of models with increasing complexity is proposed in this approach. This cascade modulates which stage classifies the input data based on the confidence of the prediction at that level.

In particular, in this approach, the complexity of the models depends on the number of features that are considered at each level. Given a particular input (e.g., a new data stream captured by an IoT sensor), the proposed multistage ensemble strategy adjusts the complexity of the systems and the optimal selection of features according to the new instance's difficulty. Optimization comes from having simple classifiers at the cascade's initial levels to increase the system's efficiency when classifying the most straightforward instances. Then, more complex classifiers in the final levels of the cascade are considered for a refinement of the system's accuracy in hard to classify instances.

Through this section, we will first describe in detail the model cascade proposal for inference. Then we will introduce three variations of the training process, which seek to find the most suitable strategy to distribute the discovery patterns across the included models.

5.1.1 **The proposed model cascade for inference**

The introduced cascade method integrates a discriminative model ensemble technique with a probabilistic approach. Based on the confidence level of the prediction, it modulates the predictive system according to the new data. This method is comprised of a series of successive N discriminative models. Based

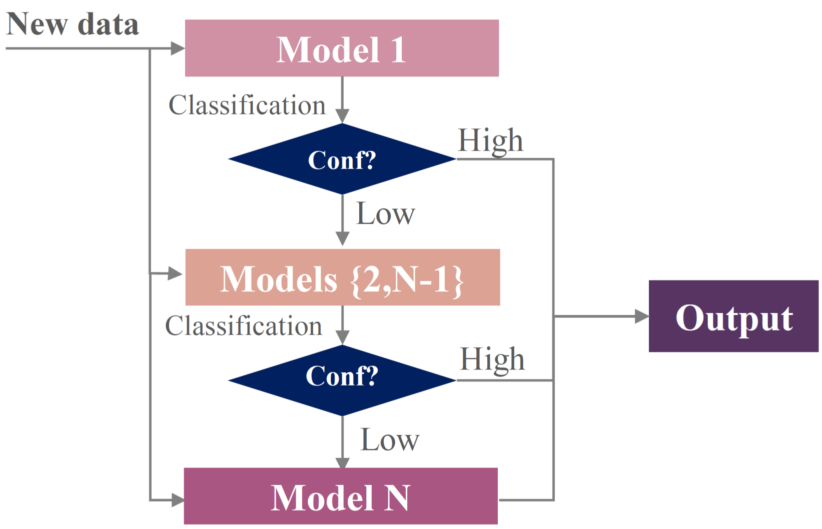

Figure 5.1: Schematic representation of the workflow for classifying new samples with the proposed model cascade.

on the confidence of the prediction at each level i, a particular data input may requiere a more fine-grained prediction with the following model at level $i+1$, which is more complex in terms of features, given some confidence thresholds t (hyper-parameter to be tuned). If the prediction of a sample at level i is considered reliable, the sample is removed from the pipeline and classified using the ith model's prediction. If none of the different $N-1$ models is able to provide a reliable prediction within these thresholds t_i, the final Nth model in charge of classifying the remaining instances. Notice that the latter model is created using all the available features to have the maximum amount of information possible for the task. A schematic representation of this system is presented in Figure 5.1.

With this strategy, simpler models act as filters that make it less likely for the data to reach the complex ones, allowing for a reduced computational burden and time to process the samples through the $N-i$ models. Therefore, the complexity of the final application is reduced by choosing from an initial cascade of pre-trained models the configuration that fits the most to the complexity of the prediction. The confidence values are determined by the prediction probabilities, which summarize the likelihood (or uncertainty) of samples belonging to each of the available classes (in our case, human activities). In this context, prediction probabilities represent the confidence a model has in the predictions, i.e., in samples belonging to the different classes. If the confidence of a sample being of class c is high, the model is confident the sample should be classified as class c. In this specific proposal, this value is calculated through a multi-class log loss metric (also referred as categorical cross-entropy). This metric calculates the negative log-likelihood for the probability predictions made by the classification model. Log loss values (those obtained from the aforementioned metric) are always positive, being bounded between 0 and infinity. Thus, a log loss of 0 represents the minimum divergence between the prediction probability and its corresponding actual/true value. In other words, the probability indicates no uncertainty and the model can be considered to be entirely confident in the prediction.

In this approach, we consider this second interpretation of this concept. For this reason, instead of using log loss for evaluating the predicted probabilities and assessing the performance of a classification problem, we use this metric as an estimator of the reliability of a prediction in which the actual class is unknown. Hence, we calculate the multi-class log loss to obtain the loss coefficient by comparing the expected probability vector, constituted by the predicted probabilities of all the classes, against the predicted label. The obtained value would be given by Equation 5.1, which calculates the loss for each class label per observation and then sums the result:

$$L(p, y_{pred}) = -\sum_{c=1}^{M} y_{o,c} log(p_{o,c}) \tag{5.1}$$

where M is the number of classes, y_{pred} is a discrete binary value (0 or 1) of the predicted class c of the observation o, and p is the predicted probability of that observation belonging to class c.

This $L(p, y_{pred})$ value is compared against the numeric threshold set at each level; If the confidence $L(p, y)_i$ at stage i is above a specific threshold value t_i, the prediction is marked as valid. In this case, the lower the threshold, the more stringent the confidence requirements of the model, i.e., it is easier that a sample has to go through more stages until it is successfully classified with a confidence higher than t_i. From the optimization point, we avoid specific samples to be processed by more complex models than what is required for classifying that difficulty. Besides, the data processing stages of the signals and the feature extraction phases are improved. That is, just the features needed for each of the levels are computed, and the remaining features are only computed at the following levels of the cascade if required. Recall that the presented cascade approach may have all the new samples labeled with the initial and simpler classifiers without requiring the computation of all the features and just a sub-set of them.

5.1.2 Different training methods

To better understand the flexibility of the proposal, we have defined three different implementations of the system's training strategy. The difference among those implementations is determined by the input data used to train each of the layers. Resembling model ensemble techniques, we can differentiate those implementations in which all the models are trained with all the available data (parallel implementation) from the ones in which the successive layers of models are trained only with previously unclassified data (sequential implementation). We also present a mixture of both methods that seeks to combine the potential advantages of both approaches (hybrid implementation).

5.1.2.1 Parallel implementation

In the parallel training all the models are trained using all the available data to, presumably, be able to generalize better on unseen samples. Therefore, all the levels are independent in terms of input. A schematic representation of the training set-up is shown in Figure 5.2. This training approach is similar to the one followed by the stacking model ensembling technique (Wolpert, 1992), in which several heterogeneous models learn in parallel with the whole training dataset. However, contrary to the stacking method, in the cascade of models, there is no final stage of creating a meta-predictor or meta-model that combines the output of the different trained models.

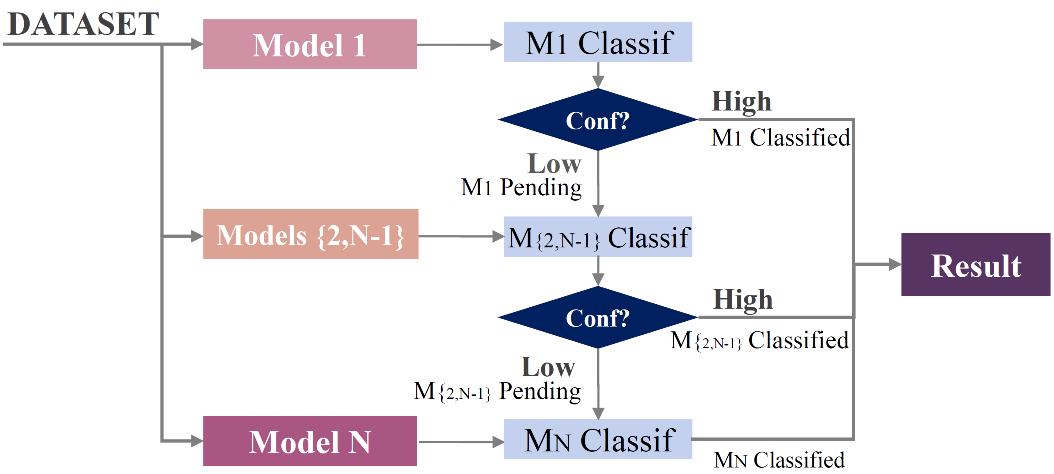

Figure 5.2: Schematic representation of the workflow for the parallel training scheme of the proposed model cascade.

5.1.2.2 Sequential implementation

The second implementation uses the whole dataset to train the initial model. However, the successive models are trained using only a fraction of this initial dataset. At stage i, this fraction consists of those instances that do not have a reliability/confidence higher than t_i in the previous phase $i-1$. That is, only those instances that fail to pass all of these $N-1$ confidence tests are used in the construction of the following exception-learner model, as shown in Figure 5.3. This way, the data input of the successive models is conditioned by the previous predictions. This allows having $N-1$ models forced to learn suitable patterns for hard to classify samples, covering the flaws of the initial model. Thus, it is designed to better discover non-covered patterns of the data. This strategy follows a similar approach to the Boosting model assembling technique (Schapire, 1990), where the successive homogeneous models are created based on the errors of the past ones and random sampling techniques. In contrast to this, in the presented cascade approach, new data does not need to go through all the trained models to be classified, which improves the performance of the final systems.

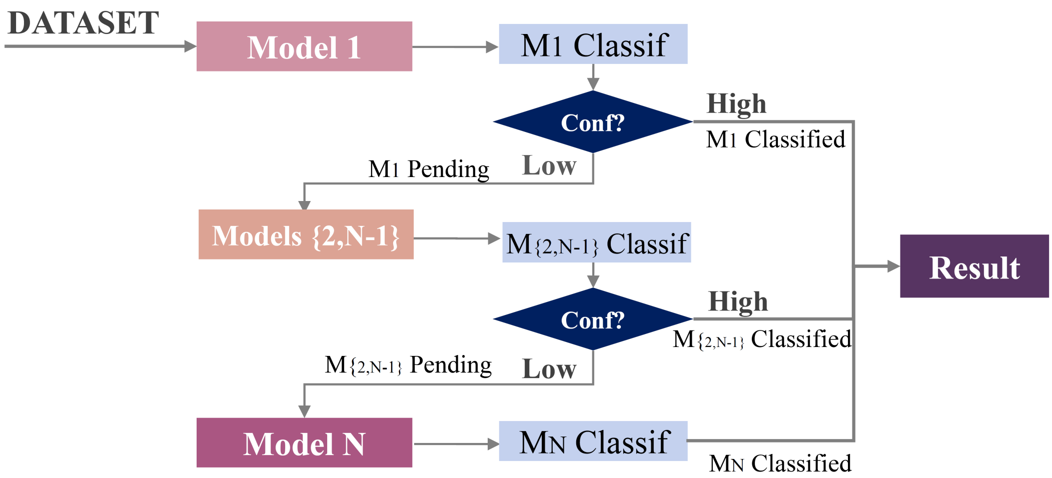

Figure 5.3: Schematic representation of the workflow for sequential training.

5.1.2.3 Hybrid implementation

It presents a mixed approach in which the idea of parallel training is followed in all the stages except in the last stage of the cascade, as Figure 5.4 illustrates. The objective of this implementation is to balance the generalization of the initial cascade stages (trained with all the available input data) with a greater particularization in the final model (trained only with only a fraction of this input data). This way, the last model is intended to specialize only in those patterns that have been left uncovered by the rest of the $N-1$ models and it is forced to be trained with few but hard to classify examples.

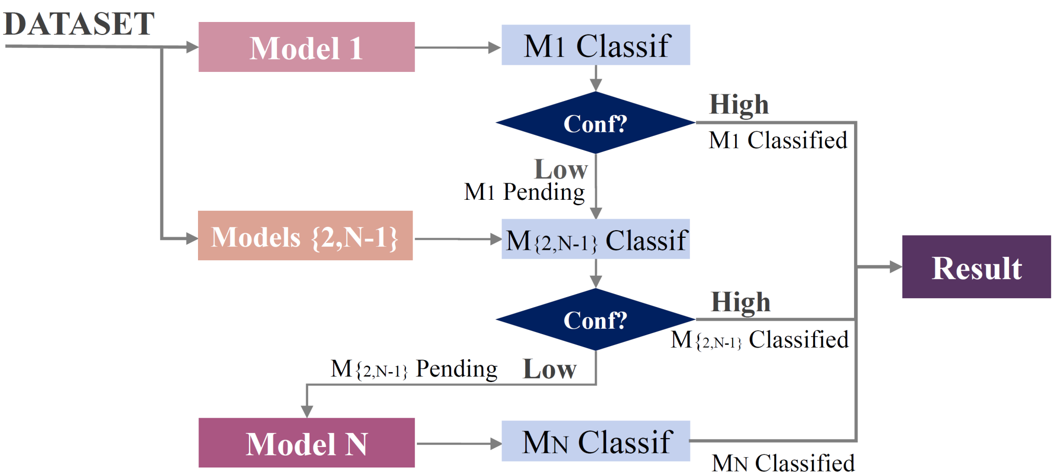

Figure 5.4: Schematic representation of the workflow for hybrid training.

5.2 Procedure and Methodology

In this section, we continue using a HAR problem to analyze the suitability of the introduced approach for continuous time-series data (e.g., data from accelerometers). For this reason, we evaluate the performance of the different variations presented in Section 5.1 through several public HAR datasets in a multi-class classification context.

5.2.1 Selected datasets

In this part of the work, we incorporate two new public datasets into the datasets used in the previous chapter. To the collection of three already introduced datasets (ADL Dataset, Sports Dataset and the novel OHM dataset), we have included two additional ones to complement the experiments of the inference stage. Those datasets are the *Human Activity Recognition Using Smartphones Dataset* (Anguita et al., 2013) (HAR Dataset) and *Office Activity Recognition using accelerometers and gyroscopes located on the forearms Dataset* (Avellaneda Gonzalez, 2019) (OFFICE Dataset). The former is one of the most common datasets for HAR applications using time-series data, while the second one is specifically focused on the context of this work. With those additions, this chapter evaluates five datasets related to sensor data and time-series data, which represent a typical use-case scenario for HAR applications and our specific context.

In the following we will introduce in detail the characteristics of these two additional detasets:

- *Human Activity Recognition Using Smartphones Dataset (Anguita et al., 2013)*

  This dataset contains 6 activities (Walking, Walking upstairs, walking downstairs, sitting, standing and laying) obtained from the accelerometer and gyroscope sensor of a mobile device located on the waist. These activities were recorded by 16 volunteer participants (11 men and 5 women, ages between 19 and 81 years). The signal of the embedded tri-axial accelerometer and gyroscope were recorded at a rate of 50Hz.

Those signals were filtered and sampled in fixed-width sliding windows of 2.56 sec and 50% overlap, constituting 128 readings per window. For this work, the pre-calculated features provided by the datasets' authors are used. It contains 561 features with time and frequency domain variables obtained for the 10299 sequences of data.

- *Office activity recognition using accelerometers and gyroscopes located on the forearms Dataset (Avellaneda Gonzalez, 2019)*

  It contains information about 5 different activities that can be performed in office settings (Use the computer, drink a beverage, have a phone conversation, take notes-handwriting and look at the smartphone). It contains 230 trials of 3 minutes of continuous data captured from sensors placed on both right and left forearms. 7 users, whose ages were not described, participated in this dataset. Tri-axial accelerometer and gyroscope signals were recorded. To perform the evaluation of this dataset, a pre-processing stage was carried out in which data was divided into segments of 255 points window with a 50% overlap. 272 time-domain features were obtained for a total of 7551 sequences of data.

As a reference for further comparison, a summary of the number of instances, classes, and the initial set of features for the total of 5 datasets analyzed in this chapter is shown in Table 5.1. This final selection of five dataset aims to cover different contexts in human activities classification in intelligent environments with data captured by motion sensors. Due to the

Dataset	Nº Instances	Nº Classes	Nº Initial Feat.
OHM Dataset	1000	3	486
ADL Dataset	10299	6	561
HAR dataset	839	13	162
Office Dataset	7551	5	272
Sports Dataset	9120	19	162

Table 5.1: Main characteristics of the selected datasets.

specific context of this dissertation, it also includes a particular focus on office environments and the use-case of monitoring hydration habits, evaluated in Chapter 4. However, this approach can be extended to other solutions or datasets as long as they contain time-series continuous data.

5.2.2 Experimental setup and design

In this part of the dissertation, we will continue with the same selection of classifiers as the previous chapter. This selection offers a good efficiency for resource-constrained environments when classifying sequences of continuous data from IoT sensors (Dhar et al., 2019). Only DT was removed for the experimental setup as this classifier's intrinsic nature opposed the probabilistic approach presented. Thus, this selection includes LG, RF, KNN, Gaussian NB, Linear SVM, and MLP. Furthermore, it is essential to remark that not all the included algorithms are probabilistic. In some cases, such as KNN or SVM, the confidence of the prediction is measured according to other parameters. For example, these parameters correspond to the number of nearest neighbors from a new instance in KNN, or to the distance from the hyperplane in SVM. Moreover, RF is for itself an ensemble learning method, consisting of an ensemble of DT algorithms. This variety of methods enriches the scope of the proposal by not limiting it to specific probability-based algorithms and provides a broader scenario for evaluation purposes.

Once more, Edge devices are the target hardware platforms for inferring new knowledge through the different proposed strategies. For that, we maintain the same selection of resource-constrained devices previously considered. Thus, we evaluate the performance of 4 different platforms (the Raspberry Pi model 4 B, Raspberry Pi 3 model B+, and the Raspberry Pi Zero W). With the idea of recapitulating their main characteristics, Table 5.2 includes a summary of the main technical specifications of the selected platforms.

All the included devices are compatible with the same software tools. Besides, they are evaluated under the same conditions to provide an equitable comparison between each device. Again, for the sake of fairness, all experiments are executed solely on the CPU, with no other application running at

Device	Processor	Architecture	CPU Freq	Cores	RAM
Laptop	i7-9750H	Intel x86	2.60 GHz	6	16 Gb
Nvidia Jetson Nano	Cortex-A57	ARMv8 64	1.43 GHz	4	2 Gb
Raspberry Pi 4 B	Cortex-A72	ARMv8 64	1.50 GHz	4	4 Gb
Raspberry Pi 3 B+	Cortex-A57	ARMv8 64	1.40 GHz	4	1 Gb
Raspberry Pi Zero	ARM11	ARMv6 32	1 GHz	1	512 Mb

Table 5.2: Selected Edge devices and their main technical specifications.

the same time. The classification solutions are powered by the ML framework Scikit Learn (Pedregosa et al., 2011). The materials used for the evaluation experiments are publicly available [1].

5.2.3 Model cascade configuration

As the objective of this work is to evaluate the capabilities of the proposed approach in an algorithm and parameter independent and agnostic manner, we made a deliberate choice to maintain the same cascade configuration for all the experiments in the comparison. Furthermore, no hyperparameter tuning has been done to improve the ML algorithms' performance and adapt them to a specific scenario. Hence, the algorithm parameters remain the same for all the evaluated datasets. Thus, choosing the optimal combination of parameters and hyperparameters for a learning algorithm that maximises its classification capabilities is out of the scope of this evaluation procedure. This avoids the impact of other factors that could distort the overall vision of the strategy being evaluated.

In the proposed discriminative model cascade, the complexity of the different models is given by the number of features selected for creating the model, as it is a crucial factor for its computational cost (Gómez-Carmona et al., 2019). For this reason, for every stage i a feature selection process is performed. This process reduces the feature matrix's dimension by removing irrelevant features, obtaining the subset of them that contributes the most to the prediction. From the exiting feature selection strategies. We apply a

[1] `https://github.com/OihaneGomez/Model_Cascade_Optimization`

Chi2 filtering method (Liu and Setiono, 1995), which was introduced in the previous chapter in Section 2.2.

We opt for a cascade configuration based on 3 layers of models that are trained with an increasing number of features. We maintain the same learning algorithm in all the models of the cascade to better compare it against the reference model. Again, for comparison purposes, this configuration remains the same, independent from the dataset or the classifier. The first level i_1 corresponds to a reduced model constituted only by 10 features. With it, the most straightforward data sequences are processed in a computationally lighter way. The second level i_2 increases the number of features up to 50. In this case, the model can deal with more complex data sequences that the previous model was not able to classify with the required confidence. Finally, the last model i_3 is in charge of classifying all the remaining instances, and it is trained using all the available baseline features, included in Table 5.1.

The agnostic nature of this setup also applies to the thresholds t_i applied at each stage of the cascade. We have selected two different thresholds, one more restrictive than the other. Even though the log loss value is dependent on the number of classes and the balance of the dataset, those thresholds remain the same in each of the evaluated datasets and algorithms. For the sake of comparison, the first threshold t_1 (from model 1 to 2) is set at 0.15, which is a highly restrictive value, as it is closer to 0. The second one t_2 (from model 2 to 3) is set at 0.40, being laxer than the previous one. These parameters represent the maximum loss value for considering the output of the corresponding model as acceptable.

The final configuration of the evaluated setup is shown in Figure 5.5. As aforementioned, we fixed this selection of models and thresholds for all our experiments. Thus, this setup is only one example of the many possibilities that the proposed approach's flexibility allows. While tuning the models is out of the scope of this work, we hypothesize that both the selected number of features and the applied thresholds could provide a good cost-accuracy balance for our evaluation purposes. However, we leave the possibility of better adjusting these factors open for further implementations of this strategy

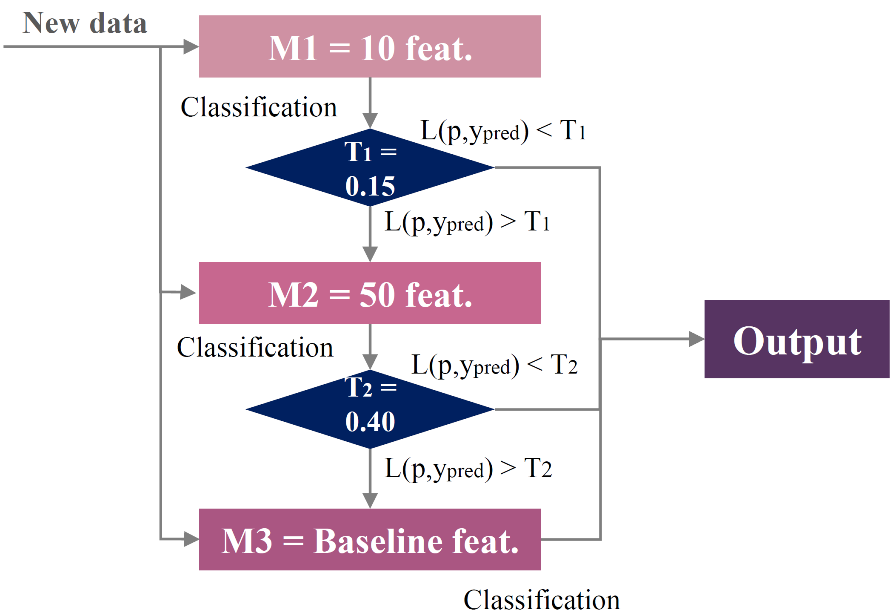

Figure 5.5: The selected configuration of the model cascade for the evaluation of the different training methods.

in where parameter fitting or thresholds adjustments could be studied for the specific task and data.

In the following, we will provide the results, both in terms of classification performance and efficiency, for the described setup.

5.3 Analysis and Results

Several experiments were conducted to analyze the potential benefits of the presented approach regarding classification performance and efficiency. For this reason, the results presented in this section can be divided into three main categories: (i) the evaluation of the classification accuracy for the presented approach, (ii) the comparison of the three variations of the proposal described in Section 5.1, and (iii) the computational cost, in terms of execution time, of the selected variant of the presented strategy when performing classification tasks in various Edge devices.

5.3.1 Classification results

This subsection evaluates the classification capabilities of the model cascade according to the procedure described in Section 5.2. For evaluation purposes, we have adapted the training schemes explained in Section 5.1 for each variation of the cascade. To ensure the reliability of our experiments, a stratified 5 fold-cross-validation (CV) procedure was applied for each dataset. The CV process is applied at each stage of the cascade independently.

In the parallel approach, the training set (composed of 4 folds of the dataset) remains the same at every level. At each stage, only the samples of the test set (the remaining fold) whose log loss is above t_i go to the next stage. The evaluation metrics for the cascade are computed using the samples of the test folds. For each sample, the prediction obtained at the stage in which the sample got a log loss below t_i is considered to obtain the final classification score. That is, if the k^{th} sample gets a log loss above t_1 at the first level and a log loss below t_2 at the second level, only the prediction of model 2 would be considered.

For the sequential approach, the CV process of the first level is performed over the whole training set. For the next stage, at each iteration, only the samples of the test fold that got a log loss above t_1 are used. Then, a new 5-fold CV is performed with all the data that remain unclassified from each test folds of the previous levels. This process is repeated with models 2 and 3. In the case of the hybrid approach, the parallel approach is used between the first and second models, while the sequential approach is used between the second and the third one. To increase the robustness of the evaluation, the results are provided as the average of 10 runs for each of the variations.

Finally, we calculate the macro-weighted *F1-score* metric for all the experiments. It averages both the true positives among all positive results (precision) and the correct labeled positives based on all the correct positive and negative events (recall).

We compare the three variations of the model cascade (sequential, parallel, and hybrid) against the reference model and another well-known model ensembling technique (i.e., Stacking technique) to evaluate their suitability.

The reference model corresponds to a single model created using the maximum initial number of features available, according to Table 5.1. Thus, this is the most complex single model that can be created using the available features. Stacking is an ensemble modeling technique in which a stack of different estimators is used to train a meta-predictor based on the outputs of all the individual estimators. Despite their differences, this is the closest example of model cascade among the rest of the model ensembling techniques. We use the same setup for the Stacking method and the cascade, using 10, 50, and the total available features to train each level. In Stacking, both the single models and the final meta-predictor are trained using the same algorithm.

Table 5.3 shows the obtained results for the five methodologies (the reference model, the stacking technique, and the 3 variations of the cascade) in terms of F1 macro results. Besides, Table 5.4 presents the relative percentage of instances classified with the two initial models of the cascade (10 and 50 features). This table provides an absolute representation of the performance of the different evaluated scenarios for all the datasets and algorithms combined.

A closer look at the algorithm-dependent results indicates that LG, RF, and SVM are the classifiers that show the most stable performance in all the evaluated datasets, presenting uniform F1 results. Ensemble techniques obtained worse results in MLP and KNN compared to the reference model. They also show high variances compared to the rest of the algorithms, particularly for the ADL and HAR datasets. Still, they maintain reasonable classification rates. NB obtains the worst performance. Not only the reference results for NB are the worst but applying NB within the cascade penalizes its results. A reason for this circumstance can be found in Table 5.4. Here, NB obtains, by far, the highest rates of classified instances using the simplest models (first levels of the cascade), in which almost all the instances do not reach the last level. This has an impact on the classification results and may indicate that the applied confidence thresholds are not well calibrated for this algorithm. Similarly, MLP and, to a lesser degree, KNN also classify a high number of instances at the early stages of the cascade. As it can be observed, for the last two algorithms, there is an indirect effect between the correlation of classified

DATASET	Threshold	Features	Method	Algorithm					
				LG	RF	KNN	NB	SVM	MLP
OHM	T1 = 0.15 T2 = 0.40	486 (all)	REFERENCE	**87.69 (SD 0.708)**	**89.44 (SD 0.368)**	80.03 (SD 0.596)	**78.76 (SD 0.122)**	87.35 (SD 0.542)	**87.75 (SD 0.407)**
		10-50-486	Stacking	87.58 (SD 0.900)	88.27 (SD 0.775)	72.45 (SD 1.817)	78.23 (SD 0.401)	**87.93 (SD 0.551)**	86.23 (SD 0.888)
			Casc. sequencial	86.70 (SD 0.561)	88.36 (SD 0.865)	81.66 (SD 0.742)	69.74 (SD 0.905)	85.76 (SD 0.727)	84.38 (SD 0.851)
			Casc. parallel	85.94 (SD 0.599)	88.58 (SD 0.425)	**82.52 (SD 0.668)**	66.87 (SD 0.709)	86.40 (SD 0.682)	83.45 (SD 0.950)
			Casc. hybrid	86.93 (SD 0.806)	87.90 (SD 0.473)	81.93 (SD 1.334)	66.82 (SD 0.646)	85.96 (SD 0.692)	82.70 (SD 0.965)
HPM	T1 = 0.15 T2 = 0.40	162 (all)	REFERENCE	83.35 (SD 1.093)	85.92 (SD 0.948)	**86.45 (SD 1.166)**	**74.74 (SD 1.357)**	**88.96 (SD 0.951)**	**83.24 (SD 1.118)**
		10-50-162	Stacking	79.09 (SD 1.160)	83.66 (SD 1.111)	72.98 (SD 1.742)	53.89 (SD 3.183)	85.30 (SD 1.614)	77.73 (SD 1.451)
			Casc. sequencial	84.39 (SD 0.922)	86.37 (SD 1.605)	79.24 (SD 1.001)	63.71 (SD 1.671)	86.98 (SD 1.075)	78.91 (SD 1.826)
			Casc. parallel	**84.48 (SD 1.000)**	85.78 (SD 0.663)	84.55 (SD 0.798)	65.41 (SD 1.709)	88.36 (SD 1.155)	80.27 (SD 1.790)
			Casc. hybrid	84.21 (SD 1.434)	**86.68 (SD 1.274)**	80.73 (SD 1.306)	65.41 (SD 1.709)	87.17 (SD 1.145)	77.56 (SD 2.299)
HAR	T1 = 0.15 T2 = 0.40	561 (all)	REFERENCE	97.96 (SD 0.065)	97.85 (SD 0.076)	**96.93 (SD 0.106)**	**73.90 (SD 0.357)**	**98.67 (SD 0.069)**	**98.19 (SD 0.137)**
		10-50-561	Stacking	**98.01 (SD 0.075)**	**98.01 (SD 0.075)**	95.67 (SD 0.245)	70.13 (SD 0.604)	98.66 (SD 0.066)	97.91 (SD 0.150)
			Casc. sequencial	95.71 (SD 0.072)	95.77 (SD 0.138)	90.46 (SD 0.235)	73.87 (SD 1.131)	93.77 (SD 0.134)	92.73 (SD 0.342)
			Casc. parallel	95.45 (SD 0.079)	96.16 (SD 0.107)	92.12 (SD 0.199)	73.00 (SD 0.393)	94.06 (SD 0.101)	93.27 (SD 0.193)
			Casc. hybrid	95.50 (SD 0.077)	96.33 (SD 0.099)	91.49 (SD 0.203)	73.35 (SD 0.453)	93.67 (SD 0.072)	92.87 (SD 0.223)
OFFICE	T1 = 0.15 T2 = 0.40	272 (all)	REFERENCE	87.52 (SD 0.145)	**97.37 (SD 0.083)**	**94.93 (SD 0.107)**	63.86 (SD 0.059)	**93.37 (SD 0.169)**	**92.34 (SD 0.378)**
		10-50-272	Stacking	84.58 (SD 0.154)	94.23 (SD 0.296)	72.73 (SD 2.792)	**65.84 (SD 0.156)**	89.33 (SD 0.211)	84.15 (SD 0.684)
			Casc. sequencial	**88.66 (SD 0.149)**	96.60 (SD 0.135)	92.94 (SD 0.188)	57.65 (SD 0.277)	92.48 (SD 0.186)	91.02 (SD 0.750)
			Casc. parallel	87.25 (SD 0.153)	96.81 (SD 0.102)	93.95 (SD 0.162)	54.65 (SD 0.213)	92.28 (SD 0.160)	92.19 (SD 0.284)
			Casc. hybrid	88.65 (SD 0.167)	96.47 (SD 0.108)	93.29 (SD 0.144)	54.85 (SD 0.213)	91.81 (SD 0.127)	91.56 (SD 0.397)
SPORTS	T1 = 0.15 T2 = 0.40	162 (all)	REFERENCE	**97.58 (SD 0.051)**	**99.13 (SD 0.037)**	**98.01 (SD 0.070)**	**94.06 (SD 0.108)**	**98.43 (SD 0.049)**	**98.05 (SD 0.122)**
		10-50-162	Stacking	97.47 (SD 0.062)	98.88 (SD 0.073)	93.85 (SD 1.779)	80.37 (SD 0.856)	98.08 (SD 0.050)	97.39 (SD 0.224)
			Casc. sequencial	97.35 (SD 0.060)	98.57 (SD 0.089)	96.18 (SD 0.212)	77.10 (SD 1.531)	97.09 (SD 0.130)	96.16 (SD 0.177)
			Casc. paralel	97.30 (SD 0.058)	98.65 (SD 0.048)	96.64 (SD 0.192)	74.11 (SD 1.435)	97.26 (SD 0.105)	96.36 (SD 0.150)
			Casc. hybrid	97.40 (SD 0.041)	98.53 (SD 0.070)	96.68 (SD 0.210)	74.05 (SD 1.432)	97.14 (SD 0.137)	95.97 (SD 0.336)

Table 5.3: F1 macro results for the reference model, the stacking method, and the three presented variations of the model cascade.

DATASET	Threshold	Features	Method	Algorithm					
				LG	RF	KNN	NB	SVM	MLP
OHM	T1 = 0.15 T2 = 0.40	10-50-486	Casc. sequencial	**70.42 (SD 1.981)**	**77.07 (SD 0.691)**	**80.31 (SD 0.917)**	**99.68 (SD 0.218)**	76.77 (SD 1.673)	**97.84 (SD 0.480)**
			Casc. paralel	70.21 (SD 0.670)	76.36 (SD 0.649)	79.98 (SD 1.006)	99.39 (SD 0.225)	**78.30 (SD 0.841)**	97.28 (SD 0.543)
			Casc. hybrid	70.21 (SD 0.670)	76.36 (SD 0.649)	79.98 (SD 1.006)	99.39 (SD 0.225)	**78.30 (SD 0.841)**	97.28 (SD 0.543)
HPM	T1 = 0.30 T2 = 0.50	27-50-162	Casc. sequencial	**51.47 (SD 0.415)**	61.60 (SD 1.484)	74.80 (SD 1.098)	99.32 (SD 0.266)	61.07 (SD 1.337)	88.02 (SD 1.801)
			Casc. paralel	51.44 (SD 0.539)	**65.42 (SD 0.727)**	**79.30 (SD 0.734)**	**99.52 (SD 0.380)**	**62.69 (SD 0.862)**	**91.63 (SD 0.777)**
			Casc. hybrid	51.44 (SD 0.539)	**65.42 (SD 0.727)**	**79.30 (SD 0.734)**	**99.52 (SD 0.380)**	**62.69 (SD 0.862)**	**91.63 (SD 0.777)**
HAR	T1 = 0.15 T2 = 0.40	10-50-561	Casc. sequencial	68.12 (SD 0.128)	74.76 (SD 0.199)	82.09 (SD 0.391)	**97.29 (SD 1.844)**	77.63 (SD 0.474)	84.23 (SD 3.179)
			Casc. paralel	**69.46 (SD 0.104)**	**81.67 (SD 0.171)**	**85.80 (SD 0.217)**	93.96 (SD 0.722)	**81.37 (SD 0.160)**	**86.37 (SD 0.544)**
			Casc. hybrid	**69.46 (SD 0.104)**	**81.67 (SD 0.171)**	**85.80 (SD 0.217)**	93.96 (SD 0.722)	**81.37 (SD 0.160)**	**86.37 (SD 0.544)**
OFFICE	T1 = 0.15 T2 = 0.40	10-50-272	Casc. sequencial	**65.40 (SD 0.065)**	87.18 (SD 0.204)	92.64 (SD 0.307)	**99.73 (SD 0.032)**	74.13 (SD 0.214)	90.59 (SD 3.742)
			Casc. paralel	65.27 (SD 0.156)	**87.41 (SD 0.244)**	**93.66 (SD 0.21)**	99.41 (SD 0.058)	**76.02 (SD 0.313)**	**92.72 (SD 0.497)**
			Casc. hybrid	65.27 (SD 0.156)	**87.41 (SD 0.244)**	**93.66 (SD 0.21)**	99.41 (SD 0.058)	**76.02 (SD 0.313)**	**92.72 (SD 0.497)**
SPORTS	T1 = 0.15 T2 = 0.40	10-50-162	Casc. sequencial	**81.17 (SD 1.071)**	94.67 (SD 0.123)	96.51 (SD 0.165)	**99.76 (SD 0.144)**	96.38 (SD 0.127)	**97.58 (SD 0.172)**
			Casc. paralel	74.99 (SD 0.192)	**94.79 (SD 0.075)**	**97.79 (SD 0.132)**	99.41 (SD 0.135)	**96.64 (SD 0.084)**	97.23 (SD 0.261)
			Casc. hybrid	74.99 (SD 0.192)	**94.79 (SD 0.075)**	**97.79 (SD 0.132)**	99.41 (SD 0.135)	**96.64 (SD 0.084)**	97.23 (SD 0.261)

Table 5.4: Percentage (%) of instances classified with the two initial models of the cascade for the three presented variations.

	OHM		ADL		HAR		OFFICE		SPORTS	
	F1	% D	F1	% D	F1	% D	F1	% D	F1	% D
REFERENCE	**86.45**		**85.58**		**97.92**		**93.11**		**98.24**	
Stacking	83.45	2.02	75.44	9.95	93.07	0.91	81.81	7.28	94.34	3.28
Sequential	82.77	2.82	79.93	4.59	90.39	3.76	86.56	1.90	93.74	3.90
Parallel	82.29	3.38	81.48	2.75	90.68	3.45	86.19	2.32	93.39	4.26
Hybrid	82.04	3.68	80.29	4.16	90.54	3.60	86.11	2.41	93.30	4.36

Table 5.5: Average F1 results for all the classifiers and their decrease percentage (% D) when compared against the reference results.

	F1	% D
REFERENCE	**89.73 (SD 5.85)**	
Stacking	85.62 (SD 7.97)	4.69 (SD 3.80)
Sequential	86.68 (SD 5.58)	3.39 (SD 1.04)
Parallel	86.80 (SD 5.18)	3.23 (SD 0.74)
Hybrid	86.45 (SD 5.50)	3.64 (SD 0.75)

Table 5.6: Average F1 results for all the classifiers and datasets, and their decrease percentage (% D) when compared against the reference results.

instances and the F1 values, meaning that the higher this percentage is, the lower the classification result may be.

Table 5.5 and Table 5.6 show a summarized comparison between the different methodologies (the reference model, the Stacking technique, and the three variations of the cascade). The former averages the algorithms' results, while the latter averages the results across all the algorithms and datasets. They also include the averaged percentage decrease of comparing the stacking and the three cascade approaches against the reference model. As explained in the previous chapter, the percentage decrease measures the percentual change between two values, according to Equation 4.6.

In this regard, the single (and more complex) model that acts as the reference method outperforms the rest of the candidates (Stacking and cascade strategies). That is, Stacking and cascade strategies obtain, on average, lower results than the reference model. Moreover, cascade variations show similar or even higher classification results than stacking techniques across algorithms. In fact, when compared against the reference model, the average percentage

	OHM	ADL	HAR	OFFICE	SPORTS
Sequ.	**83.68 (SD 12.1)**	72.71 (SD 18.2)	80.69 (SD 9.9)	84.95 (SD 12.7)	94.34 (SD 6.7)
Parallel	83.59 (SD 11.9)	**75.00 (SD 18.4)**	**83.10 (SD 8.1)**	**85.75 (SD 12.8)**	**93.47 (SD 9.2)**
Hybrid	83.59 (SD 11.9)	**75.00 (SD 18.4)**	**83.10 (SD 8.1)**	**85.75 (SD 12.8)**	**93.47 (SD 9.2)**

Table 5.7: Average percentage (%) of instances classified with the two initial models of the cascade.

decrease when averaging the 3 cascade variations is 3.42 % (SD 0.81 %), compared to the 4.69 % (SD 3.8 %) shown by Stacking techniques. Nonetheless, based on the standard deviation, stacking techniques present higher performance variations depending on the dataset. On the contrary, cascade methods present more consistent results.

However, it is paramount to emphasize that the importance of cascade approaches lies in the optimization possibilities they enable and their potential efficiency, not in their capabilities to improve the classification results. Therefore, its interest lies in the relative comparison of how much computational improvement can be obtained without substantially impairing the system's performance. With the cascade approach, the potential computational gains are given by the percentage of instances classified with the simplest models, summarized in Table 5.7. The first noticeable fact is that parallel and hybrid approaches classify the same number of instances with the two first stages of the model cascade. This is attributed to both of them sharing the same structure for those levels. Beyond that, some differences can be observed depending on the dataset, but, on average, the 83.88 % (SD 13.17 %) of the instances can be classified within these two stages.

When combined, the average F1 (Table 5.5) and the average percentage of classified instances with the initial levels of the cascade (Table 5.7) illustrate the capabilities of the model cascade system and the benefits of setting a confidence threshold. This is reflected in the comparison of cascading variations with stacking techniques in Table 5.6. In the cascade method, setting a confidence threshold compensates for the potential errors induced by the lighter models. Thus, even mainly using simple models belonging to the first levels of the cascade, it is possible to achieve similar or even better results than using a much heavier and complex model such as the one created in the stacking technique. In our case, the studied cascade strategies are able to reduce by more than 80% the number of instances to be classified with the last level of the cascade and maintain detection rates that only decrease on approximately 3.4 %. This means that assuming a decrease in accuracy lower than 4 %, we can have a system in which more than 80 % of the instances can be classified with much simpler models than the reference one.

5.3.2 Selecting the best cascade strategy

Selecting which of the three variations of the cascade approach is better requires choosing between maximizing the classification results or the number of instances classified with the simplest models (recall that efficiency in this piece of work is given by the number of instances that are classified with the initial levels of the cascade). Depending on this choice, the performance of the system is prioritized in terms of the classification outcomes or its efficiency, one being conditional on the other. Having two different optimization objectives, the best strategy consists of analyzing which of the three variations provides a better balance between both factors (i.e., which of the three variations optimize in a better balance between them).

By comparing Table 5.5 and Table 5.7 some initial conclusions can be obtained in this regard. At least for the evaluated datasets, the hybrid approach does not provide any advantage over the rest of the methods. Not only the parallel and sequential approaches obtain, in general, better classification results but the models with the parallel approach classify the same percentage of instances as models with the hybrid approach, as they maintain the same structure for the cascade's initial stages. From this point, further evaluation can optimize this trade-off to find which is the model cascade variation that better fulfils those two objectives (the best classification results and highest number of initially classified instances).

To that end, as we did in the previous chapter in Section 4.3, we conducted a Pareto multi-objective optimization analysis for decision making that selects a limited number of Pareto-optimal solutions. These solutions correspond to the Pareto front, which maximizes two objectives: accuracy and percentage

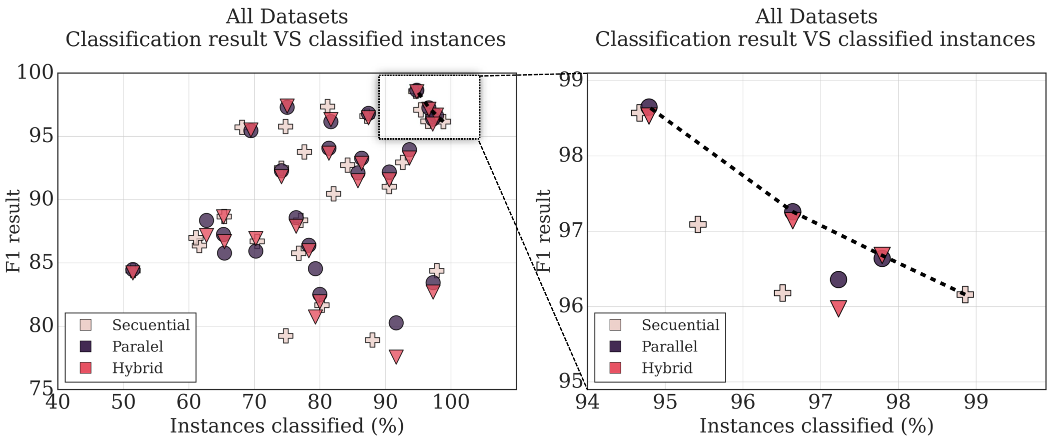

Figure 5.6: The balance between the classification results and the number of instances classified by the first levels of the cascade for all the datasets together (left plot), with special focus on the best accuracy-classified percentage region of the latter (right plots).

of classified instances in the first stage of the cascade. We apply this strategy for all the combinations of algorithms and datasets we are considering in this work. Additionally, we have specifically targeted the OHM Dataset as we will focus the rest of the computational experiments on it.

Figure 5.6 represents the relationship between the accuracy (F1 result) and the percentage of classified instances for the combination of all the dataset. A zoom-in of the upper right corner of the middle plot is shown in the right plot for the sake of visualization. This region represents the best balance between the two analyzed factors. Figure 5.7 represents the same relationship with a specific focus on the OHM Dataset. Those plots also include the Pareto frontier as a dashed line representing the optimal solutions that maximize the accuracy (highest y-axis values) and the classification percentage (largest values on the x-axis).

Finally, Table 5.8 list each of these optimal solutions for both all and OHM datasets. It includes the name of their corresponding variation of the cascade approach, F1 value and the percentage of initially classified instances. These Pareto optimal solutions show that the sequential and parallel approaches are predominant. The latter, in particular, is the most frequently listed variation. Although it is not possible to generalize based on the evaluation carried out

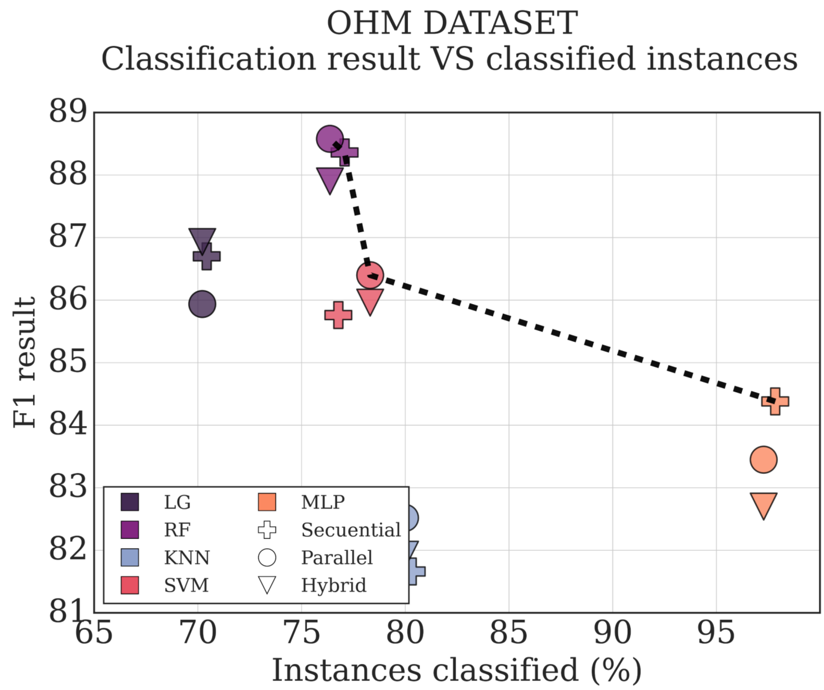

Figure 5.7: The balance between the classification results and the number of instances classified by the first levels of the cascade for the OHM dataset.

All Datasets			OHM Dataset		
Type	F1	%	Type	F1	%
Parallel	98.65	94.79	Parallel	88.58	76.36
Parallel	97.26	96.64	Sequential	88.36	77.07
Hybrid	96.68	97.79	Parallel	86.40	78.30
Sequential	96.16	98.86	Sequential	84.38	97.84

Table 5.8: Final optimal Pareto solutions that showed thee best accuracy-percentage trade-off.

and the five selected datasets, it does allow us to obtain certain indications of the potential strengths of this approach. For this reason, we carry out the computational experiments using this variation, as it obtains the most balanced results. Either way, considering how close the average classification percentage is for the sequential and parallel approaches in Table 5.7 (83.68 % vs 83.59 %), no substantial performance differences are expected due to this decision.

5.3.3 Timing results

The results presented in this subsection seek to shed light on the potential of the cascade strategy to save time and resources in the classification task. To evaluate its impact, we compare the performance of both the reference model and the parallel implementation of the model cascade strategy using the OHM dataset. We measure the computational cost as the time it takes to process and classify new instances of data.

We have randomly divided the dataset into two parts to perform this experiment, maintaining the 80 % - 20 % proportion used in the 5 fold CV process. Thus, 800 out of the 1000 instances that the dataset contains are used for training and the remaining group of data, consisting of 200 instances, are used for prediction. Therefore, we compare the elapsed time in classifying these 200 instances of the dataset with the reference model and the cascade strategy. To ensure the reliability of the results, the averaged elapsed time is computed (i.e., the experiment is executed 30 times -10 runs of 3 loops-).

The results obtained in the experiments with the five evaluated devices are included in Table 5.9. The obtained results show that, in the evaluated scenario, our approach is more efficient than the reference model when classifying new instances of HAR data. This also means that classifying part of these instances with simpler models offsets the possible overhead of some of them going through the different stages of the cascade until the confidence requirement is met. Table 5.10 summarizes the order of these time improvements, reflected in the decrease percentage (%D) between the reference times and the optimized ones. As can be observed, in most cases, the metrics improve more than the 50 %. As expected, the obtained results for every algorithm are highly reliant on the percentage of classified instances of the OHM dataset showed in Table 5.4. Consequently, those algorithms that could deal with a more considerable amount of data at the early stages of the model cascade show higher performance improvements. In the case of RF, the intrinsic time complexity of this method penalizes the efficiency of the cascade approach, obtaining improvements below the ones that would correspond to its classification percentage. Even so, it still improves the reference model's performance.

		OHM Dataset					
		LG	**RF**	**KNN**	**NB**	**SVM**	**MLP**
Laptop	Reference	$7.69s \pm 156ms$	$9.46s \pm 181ms$	$7.96s \pm 198ms$	$7.29s \pm 122ms$	$7.79s \pm 54.8ms$	$7.19s \pm 331ms$
	Cascade	$4.1s \pm 198ms$	$5.38s \pm 57.4ms$	$3.08s \pm 193ms$	$1.21s \pm 33.3ms$	$3.12s \pm 34.3ms$	$1.76s \pm 59.6ms$
Jetson Nano	Reference	$39s \pm 242ms$	$46.5s \pm 131ms$	$39.6s \pm 14.5ms$	$38.7s \pm 7.34ms$	$38.8s \pm 1.34ms$	$38.8s \pm 15.8ms$
	Cascade	$18.9s \pm 20.9ms$	$27.3s \pm 5.27ms$	$14.6s \pm 3.83ms$	$6.12s \pm 2.35ms$	$14.7s \pm 6.57ms$	$4.91s \pm 1.33ms$
RPi 4B	Reference	$35.7s \pm 14.4ms$	$42.4s \pm 63.4ms$	$36.8s \pm 7.44ms$	$36s \pm 3.72ms$	$36s \pm 4.93ms$	$35.8s \pm 5.14ms$
	Cascade	$17.1s \pm 3.27ms$	$23.9s \pm 31.2$	$13.3s \pm 3.63ms$	$5.37s \pm 142s$	$13s \pm 6.26ms$	$8.6s \pm 2.88ms$
RPi 3B+	Reference	$1min\ 10s \pm 301ms$	$1min\ 17s \pm 121ms$	$1min\ 11s \pm 152ms$	$1min\ 10s \pm 164ms$	$1min\ 11s \pm 233ms$	$1min\ 11s \pm 233ms$
	Cascade	$39s \pm 86.6ms$	$44.3s \pm 161ms$	$29.6s \pm 243ms$	$14.5s \pm 35.9ms$	$31.4s \pm 280ms$	$11.9s \pm 11.8ms$
RPI Zero	Reference	$6min\ 55s \pm 584ms$	$7min\ 46s \pm 265ms$	$7min\ 1s \pm 73.5ms$	$6min\ 57s \pm 100ms$	$6min\ 58s \pm 49.6ms$	$6min\ 56s \pm 524ms$
	Cascade	$4min\ 13s \pm 76ms$	$4min\ 14s \pm 156ms$	$2min\ 44s \pm 38.8ms$	$1min\ 16s \pm 186ms$	$2min\ 47s \pm 185ms$	$2min\ 17s \pm 79.6ms$

Table 5.9: Elapsed time for classifying 200 new data instances with the reference model and the parallel cascade method.

OHM Dataset						
	LG	**RF**	**KNN**	**NB**	**SVM**	**MLP**
Laptop	46.68	43.13	61.31	83.40	59.95	75.52
Jetson Nano	51.54	41.29	63.13	84.19	62.11	87.35
Rpi 4B	52.10	43.36	63.86	85.08	62.22	75.98
RPi 3B+	44.29	42.47	58.31	79.29	55.77	83.24
Rpi Zero	48.19	45.49	61.05	81.77	60.05	67.07

Table 5.10: Decrease Percentage (% D) showing the time improvement when applying the cascade method for each algorithm.

When averaging all the results included in Table 5.10, the time improvements were 61.67 % (SD 15.71 %) for the laptop and 64.93 % (SD 18.01 %) for the Jetson Nano board. For the Raspberry foundation devices, the average Pi 4 results were 60.56 % (SD 17.23 %), the Pi 3 obtained 60 % (SD 13.21 %), and the Pi Zero 62.31 % (SD 13.21 %). Therefore, the parallel model cascade explored in this subsection produced an enhanced performance in all the classification tasks regardless of the devices' resources. This was evaluated for a continuous HAR classification system. Furthermore, the time decreased on average 62.31 % (SD 1.96 %) when averaging all the devices' mean results. This illustrates the potential of this cascade strategy for deploying more complex HAR applications in a local stage that, otherwise, could not be efficiently executed at the Edge.

5.3.4 Limitations

Before concluding this chapter, we want to address the possible limitations of this study. The obtained results are dependent on the number of levels and the selection of the models (according to the number of features in each stage) that constitute the cascade. Additionally, they are also dependent on the probabilistic thresholds that rule the transition from one stage to another. In these experiments, the same parameters are applied for every algorithm and dataset to provide better insights into the effect of the different strategies in the classification rates. Therefore, the importance of the performed comparison lies in the relative observable results between the different strategies, not in their absolute values. It is important to note that the results obtained, both in the reference model and in the strategies evaluated, are not the best possible ones, nor do they seek to be compared against the state of the art of their respective datasets.

Additionally, the confidence boundary of the evaluated approach has been set using an adaptation of the log loss metric applied to the prediction label. As this log loss depends primarily on the probability of the respective label (as observed in Equation 5.1, it also depend on well-calibrated class conditional probabilities to estimate model uncertainty. For this reason, it is sensitive to class imbalances (Wallace and Dahabreh, 2012), and further analysis should be done to calibrate the confidence of those probabilities (Enomoto and Eda, 2021). Additionally, other metrics could be explored to measure the reliability of the predictions.

Finally, as addressed in the limitations of Chapter 4, the obtained results are not deterministic and can be sensitive to the device's internal processes or routines (despite calculating these times iteratively and running solely on the device). To increase the reliability of the measurements, we took a large number of measurement samples and provided average results.

5.4 Summary and Conclusions

In this chapter, a method to optimize and improve the performance of classification tasks using model ensemble techniques in resource-constrained devices has been presented. This is motivated by the challenge of bringing more complex classification tasks to the Edge of the network and facilitate the integration of classification problems that, due to their complexity, might not reach a satisfactory balance between their computational cost and their classification results. Thus, it contributes to enhancing the capabilities of the devices located at the Edge of the network and creating privacy-preserving spaces where data is maintained locally and closer to the user. To that end, we have analyzed the suitability of three variations of a discriminative model cascade strategy that combine different classification stages adjust the system to the complexity of the input data. In addition, we have empirically analyzed and compared the performance of the presented approach concerning multi-class HAR applications within four resource-constrained Edge platforms.

Compared with the full (and more complex) reference model, the proposed cascade strategy maintains acceptable detection rates with an average decline of only 3.42 % while classifying around 80 % of the instances at the early stages of the cascade. In the evaluated case of study, the cascade reduced by more than 60 % the baseline processing times of the prediction task for a set of 200 instances in all the evaluated platforms. Besides, on average, the proposed model cascade strategy has the potential to outperform the results of ensemble methods, i.e., the model stacking technique. The outcome of this evaluation shows the potential of the proposed approach to substantially improve the efficiency of classification tasks without drastically compromising their results.

As a conclusion of this chapter, we believe that this approach can complement the optimization process presented in Chapter 4 by presenting a more flexible system that adapts its complexity according to the new instance's difficulty. With this strategy, we seek to cover the potential limitations of the previous approach, in which a single optimized model is proposed to cover all the classification needs. Thus, those instances that do not require greater complexity can be classified by the first level of the cascade in the same way as a single reduced model would classify them. More complex models are available for the rest of the instances if needed, optimally adjusting its configuration for the best trade-off between results and performance. This way, additional steps, in the form of a sequence of incrementally complex models, give this part of the dissertation an additional depth and broaden the possibilities to carry out the classification tasks in an efficient and optimized manner. However, the obtained benefits come at a price. This cascade strategy improves inference tasks, but results in a less efficient training stage than the one presented in Chapter 4, in which a unique and straightforward classifier was targeted. This particular approach would involve re-training at least two models instead of one single one. Even so, training a reduced model such as the ones included at the early stages of the cascade would still be less computationally expensive than training the reference model.

This proposal's flexibility and optimization capabilities enable an increased landscape of efficient applications that make the Edge truly intelligent. In consequence, to preserve the external accessibility of the data thanks to its local integration. For this reason, it contributes to alleviating privacy concerns in a human-centric work environment. Moreover, this strategy could eventually enable the possibility to port models with lower computational requirements from the Edge to the end device and allocate part of the cascade system there, directly to the data source.

The difference between involvement and commitment is like ham and eggs. The chicken is involved; the pig is committed.

Martina Navratilova

CHAPTER

6

Hybrid Intelligence: a collaborative interactive approach at the Edge

ADAPTING intelligent systems in terms of users' goals and involvement levels is an essential factor to consider when designing trustworthy and human-centric IoT solutions. In some intelligent environments, where sensitive information is collected, two main challenges arise: (i) preserving the external access to the critical data and (ii) interactively modulating users' participation. To address them, in this chapter, a hybrid IML approach is presented. In this approach, users participate in the AI's model fitting by providing new annotated data when inquired by the system. This process helps to personalize the model according to their personal inputs (e.g., personal gestures) and create tailored AI models that can evolve as the user does. While most IML systems may assume that participants are always willing and ready to be involved in this interactive process, this is not always the case. In this work, the fact that different end-users may have different involvement levels is also contemplated to accommodate the participation willingness of every individual at any time.

For this reason, to face the aforementioned challenges, the proposal that will be presented in this chapter combines performance optimization techniques with a HiTL approach. Furthermore, to better adapt the system to users' preferences, this proposal allows them to determine their involvement level and willingness to participate in the learning process of the model. Thus, in this chapter, the model cascade optimization technique presented in Chapter 5 is adapted to an interactive scenario that modulates the number of times the user may be queried to label new instances of data. For that, this approach matches the confidence thresholds that rule the classification cascade with the user's desired participation level. This way, depending on these thresholds, the cascade could provide the most probable prediction for a new sequence of data or ask the user about its corresponding label. Those annotated sequences of data can be used to retrain the initial model and personalize it to every user.

We evaluate our strategy both from a quantitative and a qualitative point of view. For that, this chapter first provides a quantitative experimental evaluation to explore the performance of the cascade approach when applying the interactive strategy. Then, the results of a conducted qualitative evaluation are presented. This evaluation consists in semi-structured interviews with 12 participants in which an interactive prototype of a hydration monitoring system was presented and their opinion regarding such a system gathered. Our preliminary analysis provides insights into the potential of this approach to customize the users' role in interactive work environments and create more flexible and effective learning systems.

The rest of the chapter comprises various sections, starting from Section 6.1, which analyzes the motivations for designing flexible IML strategies. Then, in 6.2 the proposed interactive strategy's effect on the model cascade scheme is evaluated. Section 6.3 describes the Smart drink monitoring system, designed as a use-case prototype for the qualitative evaluation. Then, in section 6.4 we summarize the main finding obtained through the conducted evaluation. Finally, we conclude this chapter in Section 6.5, which includes the final remarks and insights gathered throughout this part of the work.

6.1 Revisiting the importance of Interactive ML approaches

Transforming the captured raw data into useful information is a complex task (Guo et al., 2013). Consequently, IoT devices usually raise questions about the accuracy of the collected data or how well that data can be associated with the phenomena it measures (Masson et al., 2016). In this regard, Yang et al. (2015) analyzed the users' perception about various tracking devices, studying how the accuracy of the same device is perceived very differently by each user depending on his/her expectations or characteristics. Authors pointed to the difficulty of the models to adapt their performance to real scenarios as one of the options for those differences. Moreover, they addressed that users' lack of knowledge of how their devices work, what data is collected, or how the measured phenomena are calculated can also impair this perception. In the same direction, they also discovered that users are likely to expect to customize the device to suit their individual needs better. In fact, users may exhibit frustration when activity monitoring devices malfunction or are not suitable enough to classify their patterns, leading to future disengagement (Fausset et al., 2013; Shih et al., 2015). This fact highlights the importance of considering the specific circumstances of each individual when designing IoT-based systems involving tracking devices. That is, allowing the user to have greater control over the performance of the system and the device while increasing the degree of customization possible. In consequence, the design process of the activity classification system should take advantage of enhanced model calibration and personalization process by end-users. This personalization can increase the accuracy of the classification system and helps to address a wider diversity of user.

Under this context, user involvement can be crucial for improving system performance while creating engaging experiences. HiTL approaches examine how this connection between the user and the system can be exploited to increase the perception of technology. In those approaches, users' collaboration with AI techniques can take many roles. Whilst in classic ML the assumption is that the training phase will be performed only once, in a HiTL scenario, models can be continuously retrained thanks to the user feedback. For example, users may be integrated into the learning process and provide the system with sources of truth (i.e., new labeled sequences of data) to improve its detection capabilities. However, the level of involvement needed to obtain those sources of truth appears to be a barrier for motivating the user to participate in model improvement over a long time, especially when there is no offset with the perceived value produced. For instance, in the experimental evaluation performed by Masson et al. (2016) regarding users involvement with AI systems, participants initially took an active role in data capture. However, eventually, most of them ended up quitting because of the time involved and the poor accuracy and perceived reward.

For this reason, it is essential to have a good trade-off between the efforts that are made and the perceived value of such efforts. Additionally, users may not be willing to get highly involved in the learning process and, thus, it is important not to force them to interact more than necessary. Therefore, the particularities of each individual must be taken into account to fit better the expectations of a broader range of users. Similarly, their willingness to interact with the deployed system may vary along with context and task (Ramos et al., 2019). Thus, continuously adapting the interactive system in terms of users current goals and involvement levels is a pivotal factor to consider when designing trustworthy solutions.

This chapter introduces how the cascade optimization approach described in Chapter 5 can be transformed into a machine-human collaborative strategy. This strategy aims to modulate the participation requirements of interactive processes according to each user's level of involvement. In essence, this strategy combines the optimization needs of a privacy-preserving local scenario with a flexible user integration in the interactive learning process. In the following, we analyze the suitability of this approach from a two-fold perspective: (i) exploring how different involvement levels can contribute to improving the classification rates of the system and (ii) performing a qualitative evaluation to collect insights about users initial thoughts and opinions toward the collaborative system described.

6.2 The cascade approach for Interactive ML: quantitative analysis and results

This section presents a flexible proposal that can be considered an intermediate approach between a classic ML system (with no user interaction) and a fixed interactive system (with a pre-set interaction level). This intermediate approach is based on the cascade strategy presented in Chapter 5 and takes advantage of the confidence level of the prediction to differentiate several involvement levels. Hence, to modulate users' participation in the classification process. In this adaptation of the cascade strategy, the final stage of the cascade (that in the previous chapter corresponded to the most complex model) is replaced by an interactive stage. This way, if the series of successive N discriminative models that constitute the cascade is not able to predict with the required confidence a new sequence of data, the user could be inquired about the activity to which the captured data corresponds. This adaptation of the cascade technique to this new interactive scenario is presented in Figure 6.1.

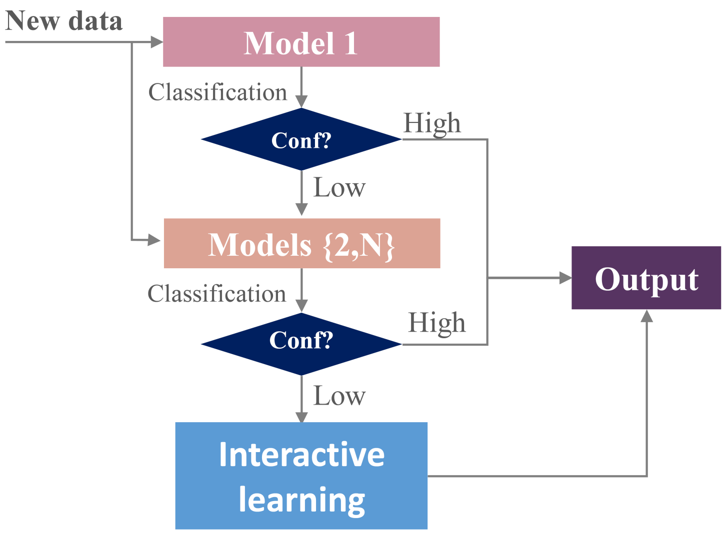

Figure 6.1: Schematic representation of the interactive model cascade approach, adapted to incorporate a final interactive stage in which the system may inquire the user if none of the N levels of the cascade can provide a reliable prediction for a new sequence of data.

	Confidence Threshold	Expected interaction	Classification results
Higher involvement	High	High	Better
Lower involvement	Low	Low	Worse

Table 6.1: The expected outcomes of the proposed interactive cascade strategy depending on the selected involvement level of the user.

The main difference between the previous implementation of the cascade and this new proposal is that the confidence thresholds t are not static parameters. Here, those thresholds are internally linked with the participation preferences of the user. This way, these thresholds determine how reliable a prediction needs to be to determine whether to accept it or, eventually, interact with the user to obtain the true label for the predicted data sequence. In other words, this interactive cascade balances the interaction frequency of the user (i.e., how many times the system may inquire him or her) and the classification accuracy.

Lower involvement levels are linked to less restrictive confidence thresholds, whereas higher involvement levels are related to more restrictive confidence thresholds. Therefore, if the confidence requirement is high, it is more likely that a complex data sequence can not be classified by any of the cascade models with the required reliability, being necessary to query the user about its true label. Thus, it is expected to have a greater number of interactions when higher confidence requirements are set. On the contrary, with low confidence requirements, less reliable predictions are admitted as valid. As a result, more data sequences are classified with the cascade models, reaching the final interactive stage less often. Therefore, this adaptive configuration optimizes the classification performance for every new sequence of data depending on its complexity. At the same time, it adapts the interactive strategy depending on the user's participation willingness. Table 6.1 illustrates the expected outcomes of this interactive strategy and how the system is expected to behave depending on the selected involvement level of the user.

Once the user is involved in the process by providing new labels for the hardest to classify instances, this new information can be used to improve the

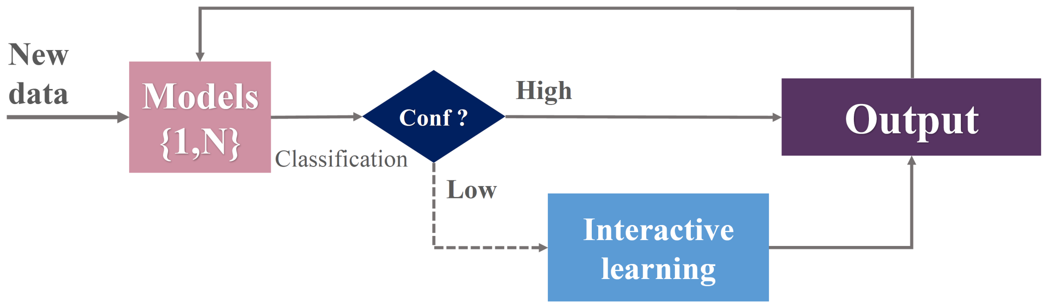

Figure 6.2: Schematic representation of the interactive retraining loop. User labeled data is incorporated into the learning process to personalize the model using those instances of data.

classification system by retraining the model with this new data. It is expected that this retraining phase has an impact on future classifications results. In this case, the more involved the user is, the more data will provide to the new model. This way, retraining the system according to new annotated data ultimately results in a more personalized system capable of being adapted to the particularities of each user. A simplified schematic of this described process is shown in Figure 6.2. This idea is similar to the Active Learning strategy. An oracle (in this case, the user) is asked to label some specific instances of the dataset to include them in the final model.

In the following, we continue the experiments started in Chapter 5 and analyze how this interactive strategy may affect the classification performance of the system using the novel OHM Dataset. To better evaluate how an interactive strategy would influence the personalization of a certain model, it is essential to know first how a non-interactive reference model behaves when classifying data from a new user whose examples have not been part of the training process. In our context, this reference model corresponds to the one trained using all the available baseline features. For this reason, we begin our experimentation by performing a leave-one-subject-out evaluation with this reference model. Then, we evaluate the interactive model cascade strategy when classifying data from an unseen user.

6.2.1 Leave-one-subject-out evaluation of the reference model

In this subsection, we obtain the classification results of the reference model for all the classifiers in a user-dependent form. In order to do so, a leave-one-subject-out cross-validation evaluation procedure will be followed. This evaluation procedure is an iterative method that uses the available observations of all subjects except one, which is excluded for validation purposes, to train the model. The remaining subject is used to test the prediction capabilities of the trained model. This procedure allows knowing the individualized performance of the model in relation to each of the users. In this case, since the OHM Dataset consists of 10 independent subjects (or users) and 1000 instances of data, the classification model is trained with the data of 9 subjects (900 instances) and tested with the remaining one (100 instances). This process is repeated 10 times (one subject at a time) to obtain the average classification rates of the system.

The obtained results for this leave-one-subject-out method for all the classifiers (LG, RF, KNN, NB - Gaussian, Linear SVM, MLP, and DT) are shown in Table 6.2 and Table 6.3. The former includes the classification results according to the macro *F1-Score* evaluation metric, while the latter corresponds to the obtained Accuracy results. Both tables show the results for each of the 10 users included in the OHM dataset. It also includes the average results for all the classifiers and users. It is worth noting how the obtained results vary according to each user. For example, user 9 presents the highest-ranking result in both F1 (85.73 SD = 5.54) and accuracy (87.71 SD = 4.60). In contrast, user 4 obtains the worst results of the comparison, with an F1 of 69.52 (SD = 3.50) and an accuracy of 78.29 (SD = 3.14). From the best-case scenario in the classification (user 9) to the worst-case scenario (user 4), there is a loss of 16.21 points for F1 and 9.43 points for accuracy. The variance observed in the obtained results is indicative of the dispersion of the dataset and the freedom of execution in the actions that every user performed when recording the three target activities (Drink from a mug, Drink from a bottle and Other). Moreover, each of them brought their own liquid container for

User	Algorithm							MEAN (SD)
	LG	RF	KNN	NB	SVM	MLP	DT	
User 1	85.90	89.79	75.51	86.30	81.99	88.56	79.37	**83.92 (5.18)**
User 2	78.28	82.78	73.08	71.39	82.41	80.24	65.13	**76.19 (6.56)**
User 3	77.26	86.08	70.02	80.20	76.00	78.04	66.65	**76.32 (6.42)**
User 4	73.33	75.42	68.53	65.64	67.43	67.93	68.39	**69.52 (3.50)**
User 5	87.74	81.94	71.65	76.92	89.16	83.58	75.78	**80.97 (6.47)**
User 6	85.02	82.03	73.14	78.23	84.32	87.20	79.78	**81.39 (4.77)**
User 7	79.59	82.55	78.03	75.52	82.50	80.86	68.98	**78.29 (4.81)**
User 8	87.78	92.44	68.44	70.41	87.85	90.26	81.01	**82.60 (9.67)**
User 9	88.69	90.87	79.41	79.05	90.83	90.02	81.24	**85.73 (5.54)**
User 10	84.77	82.95	71.22	73.65	82.64	88.14	56.05	**77.06 (11.07)**
MEAN (SD)	**82.84 (5.31)**	**84.69 (5.14)**	**72.90 (3.76)**	**75.73 (5.81)**	**82.51 (6.79)**	**83.48 (6.86)**	**72.24 (8.48)**	**79.20 (4.72)**

Table 6.2: Macro F1 results for the leave-on-subject-out evaluation of the OHM dataset.

User	Algorithm							MEAN (SD)
	LG	RF	KNN	NB	SVM	MLP	DT	
User 1	87.00	90.00	78.00	87.00	84.00	90.00	80.00	**85.14 (4.70)**
User 2	83.00	86.00	77.00	77.00	86.00	84.00	74.00	**81.00 (4.89)**
User 3	80.00	88.00	74.00	83.00	80.00	81.00	73.00	**79.86 (5.14)**
User 4	81.00	84.00	76.00	75.00	77.00	78.00	77.00	**78.29 (3.14)**
User 5	89.00	85.00	77.00	80.00	90.00	86.00	79.00	**83.71 (5.08)**
User 6	87.00	84.00	76.00	80.00	87.00	90.00	83.00	**83.86 (4.74)**
User 7	82.00	85.00	81.00	78.00	85.00	83.00	74.00	**81.14 (3.97)**
User 8	90.00	93.00	74.00	73.00	89.00	92.00	83.00	**84.86 (8.39)**
User 9	90.00	92.00	83.00	81.00	92.00	91.00	85.00	**87.71 (4.60)**
User 10	88.00	85.00	77.00	76.00	86.00	91.00	65.00	**81.14 (9.00)**
MEAN (SD)	**85.70 (3.83)**	**87.20 (3.36)**	**77.30 (2.83)**	**79.00 (4.11)**	**85.60 (4.50)**	**86.60 (4.90)**	**77.30 (6.02)**	**82.67 (2.82)**

Table 6.3: Accuracy results for the leave-on-subject-out evaluation of the OHM dataset.

the recording session of the dataset, which also affects the variance of data and, thus, the presented results.

One remarkable aspect of the obtained results concerns the observable differences between F1 and Accuracy values. On the one hand, accuracy measures the number of errors concerning the total number of classified instances.

F1-score, on the other hand, is the harmonic mean of Precision and Recall. Thus, while the former is a combined measure of all the correctly identified cases, the latter measures the trade-off between the type I error (false positive) and type II error (false negative). In addition, the reported F1 value is more sensitive to class imbalance, as it considers how the data is distributed and provides macro averaged results. In this specific evaluation, the essential aspect for the remainder of the experimentation is understanding how the ability to classify new data improves when applying an interactive strategy. Therefore, its relevance lies in providing an understandable representation of how many errors can be avoided with an interactive strategy and the system's ability to classify any instance correctly. For this reason, from this point on, the provided results will be referring to the accuracy metric. Although the accuracy results (included in Table 6.3) are slightly higher than those of F1 (included in Table 6.2), since they will be compared on an equal basis with those obtained when applying the different strategies, this relative comparison is equitable with respect to the metric used.

6.2.2 The cascade strategy for an interactive scenario

Having understood how user-dependent evaluation affects model performance, this subsection will evaluate the impact of replacing a single complex reference model (trained using all the available baseline features) with a cascade of simple models. To do so, we will initially evaluate a cascade analogous to the one presented in Chapter 5, using the parallel cascade approach presented there. The final configuration of this cascade is shown in Figure 6.3. Two levels are included in this setup. The first level i_1 corresponds to a reduced model of 10 features. The second level i_2 increases the number of features up to 50. Lastly, the final model is replaced by an interactive stage in which the user may be inquired about the true label of the data if none of the two previous models can provide a prediction that meets the confidence requirements.

The confidence requirements of each i level depend on the T_i thresholds. Those thresholds are compared against the multi-class loss value of the pre-

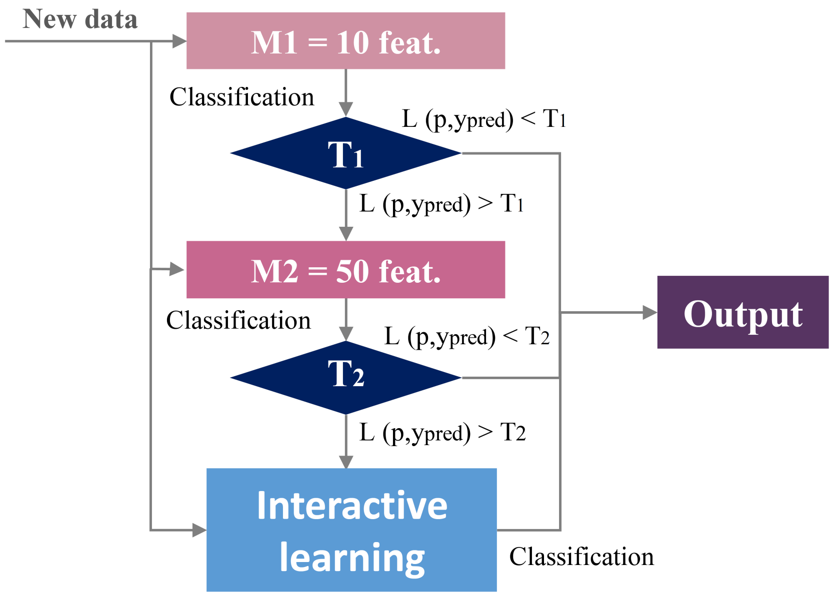

Figure 6.3: The evaluated model cascade's configuration including the final interactive stage. T_1 and T_2 are defined according to the participation willingness of the user.

diction $L(p, y_{pred})$ defined in Chapter 5 with Equation 6.1:

$$L(p, y_{pred}) = -\sum_{c=1}^{M} y_{o,c} log(p_{o,c}) \tag{6.1}$$

As this value tends to infinity for progressively worse scores, lower loss values involve higher confidence levels on the model's prediction. In contrast, higher loss values are associated with less reliable predictions. In this implementation of the cascade, those confidence requirements T_1 and T_2 are not fixed thresholds. Instead, they depend on the potential involvement level of the user. For this reason, three levels of stringency have been defined: low, intermediate, and permissive. Those levels are linked to the $L(p, y_{pred})$ threshold values and intrinsically related to the reported participation willingness of each user. So that, the lower the threshold, the more likely it is that the interactive stage will be reached. Therefore the user will be queried by the system about a specific data sequence.

Those three levels are calculated by setting a confidence factor that mul-

Involvement	Confidence level	Threshold value
High	Restrictive	$T_1/1.25$ — $T_2/1.5$
Medium	Intermediate	T_1 — T_2
Low	Permissive	$T_1 * 1.25$ — $T_2 * 1.5$

Table 6.4: The correlation between the users' involvement requirements in the interactive strategy, the associated confidence levels, and the threshold values for each of the cascade levels.

tiplies or divides the baseline thresholds. For the sake of comparison, the first threshold T_1 (from model 1 to 2) is set at 0.15. The second one T_2 (from model 2 to 3) is set at 0.40. The confidence factor for each level has been set at 1.25 and 1.50, respectively. This maintains a scale according to which higher cascade levels have more restrictive thresholds than successive ones. That is, the first T_1 threshold will become more demanding in terms of confidence on the prediction than T_2. It should be remarked that both T_1, T_2, and the confidence factors 1.25 and 1.50 could be tuned to improve the classification performance of the strategy. However, we have set these values only for comparative purposes, and no other fitting process is performed. Table 6.4 summarizes the correlation between the concepts of involvement, the confidence level associated with the prediction, and the corresponding threshold value once these confidence factors are applied.

The obtained accuracy results for the leave-one-subject-out cross-validation process of the parallel cascade are included in Table 6.5. For comparative purposes, this table includes the reference results (from Table 6.3). It also incorporates the results obtained concerning those instances classified by the first two levels of the cascade (Acc M1 + M2) as well as the number of instances (out of the 100 sequences of data that form the test set) that would reach the interactive stage (Pend Inst.). Those pending instances would not have been automatically classified by the early stages of the cascade (M1 and M2) and would be inquired to the user.

As can be observed, setting a confidence threshold compensates for the potential errors induced by the lighter models. In fact, it increases the accuracy of all the classifiers for those instances that are classified by one of

	Invol.	Threshold	Acc Reference	Acc M1+M2	Pend. Inst.
	High	T1/1.25 - T2/1.5		94.156	39.80
LG	Medium	T1 - T2	85.70	90.703	29.20
	Low	T1*1.25 - T2*1.5		85.531	15.30
	High	T1/1.25 - T2/1.5		94.969	34.30
RF	Medium	T1 - T2	87.20	92.701	26.10
	Low	T1*1.25 - T2*1.5		88.839	15.00
	High	T1/1.25 - T2/1.5		87.016	21.40
KNN	Medium	T1 - T2	77.30	87.015	21.40
	Low	T1*1.25 - T2*1.5		83.077	2.50
	High	T1/1.25 - T2/1.5		72.966	3.70
NB	Medium	T1 - T2	79.00	72.916	2.80
	Low	T1*1.25 - T2*1.5		71.913	2.30
	High	T1/1.25 - T2/1.5		93.879	34.90
SVM	Medium	T1 - T2	85.60	89.818	23.90
	Low	T1*1.25 - T2*1.5		85.440	9.40
	High	T1/1.25 - T2/1.5		81.223	3.90
MLP	Medium	T1 - T2	86.60	80.415	2.80
	Low	T1*1.25 - T2*1.5		80.447	0.80

Table 6.5: Accuracy results for the leave-on-subject-out evaluation according to several threshold levels. Low corresponds to a lower confidence requirement on the prediction to accept it as valid, while High relies on higher confidence rates.

the two initial models (M1 or M2) except NB and MLP. Similar to what was observed in Chapter 5, both classifiers obtain the worst performance, and the cascade approach does not improve the reference results. This may indicate that the applied confidence thresholds are not well calibrated for those algorithms. Still, the rest of the classifiers benefit from the cascading strategy, improving substantially the results of the classified instances.

These results also illustrate how the different thresholds impact the classification results. The effect of the level of stringency in the confidence requirement is very apparent when comparing the obtaining results for each classifier. It can be noted how, as the degree of involvement rises and the confidence requirement increases, the classification results also improve. Consequently, fewer data sequences are classified with the required reliability by Model 1 and Model 2, and more of them reach the interactive stage. This means that users would be inquired more often to provide a label for those instances. In the following subsection, we will evaluate the effect on the classification performance of retraining the model with those instances labeled by the user.

6.2.3 Retraining the model with user annotated data

As stated previously, integrating the user in the control loop can personalize the classification system once new labeled instances are used to retrain the initial model. For this reason, in this subsection, we perform an experiment in which an interactive simulated scenario is analyzed. For this evaluation, a specific splitting process of the original dataset is needed. A representation of the process followed to perform this evaluation is included in Figure 6.4.

In this process, the initial 100 instances of each subject of the dataset, which were previously used as a single test set (*Initial Test*), have been divided into two blocks of 60 (*Test*) and 40 (*Validation*) instances each. Thanks to this division, it is possible to split the evaluation process into two stages to see how the classification results evolve when new data is included in the model. In the first stage, the model is trained with the 900 instances included

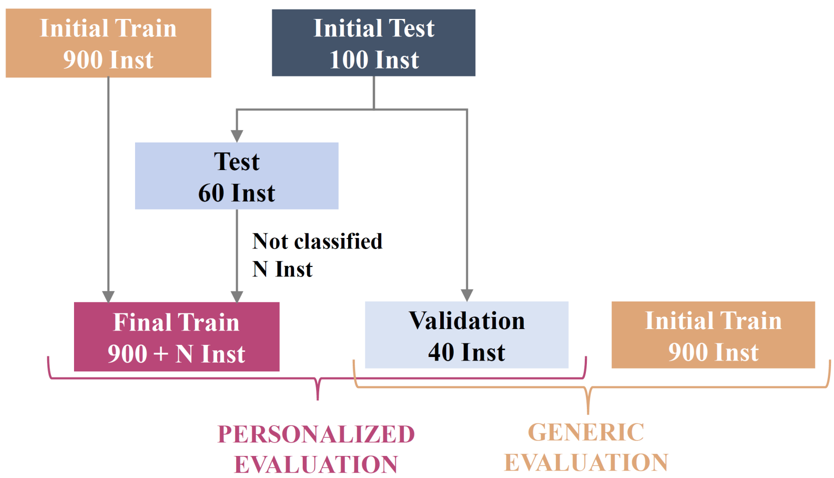

Figure 6.4: Dataset division to simulate an interactive scenario and evaluate the effect of retraining a generic model with user-dependent data.

in *Initial Train data*. Next, the 60 instances of *Test* data for each target user are used to determine the number of instances of these 60 that would remain without being classified by the two initial levels Model 1 and Model 2 of the cascade described above. These remaining *N Instances* are assigned their corresponding correct label as if a user was labeling them. Then, those labeled instances are incorporated into the initial *Initial Train*.

Thus, for the second stage of the evaluation process, the initial model is retrained including the user-specific data. The final training set *Final Train* that will be used at this second stage corresponds to the sum of the initial 900 instances of the *Initial Train* and the remaining *N instances* of the previous stage. To evaluate the influence of incorporating those *N instances* into the model, we compare the classification performance of the initial model and the new personalized one. We use the *Validation* data (the remaining 40 instances of the original *Initial Test* set) for that. We compare the obtained results when these *Validation* instances are classified by the initial model (*Generic evaluation*), and when they are classified with the retrained model that includes the new included *N instances* (*Personalized evaluation*).

The obtained results of the described evaluation are included in Table 6.6. This table appends both the average reference results (from the leave-on-subject-out evaluation included in Table 6.3) and the average number of pending instances (*N Inst*) that remained unclassified from the *Test set*. Those *N Inst* are the ones that are included in the *Final Train* set to retrain the model with user-dependent data. Moreover, this table illustrates the influence in the accuracy results of including those pending instances in the final model. For that, the obtained accuracy for the generic model (the results for the initial model trained with the *Initial Train* data) and personalized model (the results for the retrained final model with the *Final Train* which includes user-dependent data) are shown for comparative purposes.

In both cases, two columns of results are included, encompassing two different scenarios. On the one hand, the interactive scenario is evaluated. Coined as *M1+M2* in the table, those results correspond to those instances classified in the first two levels of the cascade (Model 1 and Model 2). In this case, the interactive strategy remains at the final stage of the cascade. On the other

Alg	Invol.	Threshold	Reference	N Inst	Generic		Personalized	
					M1+M2	All M3	M1+M2	All M3
LG	High	T1/1.25 - T2/1.5	85.70	23.01	95.74	85.56	96.50	88.28
	Medium	T1 - T2		17.00	92.31	85.15	92.97	87.54
	Low	T1*1.25 - T2*1.5		8.04	87.54	83.34	88.70	85.28
RF	High	T1/1.25 - T2/1.5	87.20	22.14	95.42	86.92	96.69	89.22
	Medium	T1 - T2		16.76	93.74	86.84	94.78	88.60
	Low	T1*1.25 - T2*1.5		8.71	89.19	85.66	90.58	87.19
KNN	High	T1/1.25 - T2/1.5	77.30	19.45	88.59	81.88	89.28	83.59
	Medium	T1 - T2		15.33	88.59	81.88	89.28	83.59
	Low	T1*1.25 - T2*1.5		7.30	83.96	82.68	84.04	82.83
NB	High	T1/1.25 - T2/1.5	79.00	16.02	76.87	76.05	77.07	76.35
	Medium	T1 - T2		12.88	76.41	76.17	76.61	76.27
	Low	T1*1.25 - T2*1.5		5.85	76.33	76.05	76.45	76.18
SVM	High	T1/1.25 - T2/1.5	85.60	15.27	94.06	86.27	94.25	88.60
	Medium	T1 - T2		11.99	90.98	85.74	92.55	87.96
	Low	T1*1.25 - T2*1.5		5.64	86.29	84.02	88.13	85.47
MLP	High	T1/1.25 - T2/1.5	86.60	14.81	86.85	84.18	89.01	86.63
	Medium	T1 - T2		11.35	85.73	83.92	87.66	85.35
	Low	T1*1.25 - T2*1.5		5.315	84.61	83.40	84.71	83.53

Table 6.6: Average accuracy results and the number of instances used to retrain the model for an interactive scenario. Generic results correspond to the validation set classified with the original model. Personalized results correspond to the evaluation of the model trained with the modified train set that includes user annotated data.

hand, a non-interactive scenario is evaluated. In this case, the last interactive stage at the end of the cascade is replaced by Model 3, which corresponds to the reference model (the one trained with all the available features). Those results are coined as *All M3* in the table. This approach resembles to the cascade presented in the previous Chapter 5 and evaluates the effect of stopping the collaboration with the learning system.

When comparing the effect of the involvement levels, once again, it can be noted that a higher involvement level implies better results for the classification of new instances by the early stages of the cascade (*Acc M1+M2*). Nevertheless, a slight effect can be observed when incorporating a third model (*All M3*). Besides, adding user-specific data and retraining the model positively influences the classification results. This can be observed when comparing the generic versus the personalized results. For this reason, this table

Alg	Invol.	Acc M1+M2 (% Incrase)	Acc All M3 (% Increase)
LG	High	0.79	3.17
	Medium	0.72	2.81
	Low	1.33	2.32
RF	High	1.33	2.65
	Medium	1.11	2.02
	Low	1.55	1.78
KNN	High	0.78	2.09
	Medium	0.78	2.09
	Low	0.10	0.18
NB	High	0.26	0.38
	Medium	0.25	0.13
	Low	0.16	0.17
SVM	High	0.21	2.70
	Medium	1.72	2.59
	Low	2.13	1.72
MLP	High	2.49	2.91
	Medium	2.25	1.70
	Low	0.12	0.16

Table 6.7: Percentage increase between the accuracy results of the initial generic model and the personalized one for the interactive (*Acc M1+M2*) and non-interactive scenario (*Acc All M3*).

Invol.	Nº Instan	Acc M1+M2 (% Increase)	Acc All M3 (% Increase)
High	18.45 (SD 3.60)	0.98 (SD 0.85)	2.32 (SD 1.01)
Medium	14.22 (SD 2.46)	1.14 (SD 0.73)	1.89 (SD 0.95)
Low	6.81 (SD 1.41)	0.90 (SD 0.89)	1.05 (SD 0.99)

Table 6.8: Number of instance incorporated to the personalized models and the percentage increase between the accuracy results of the initial generic model and the personalized one for the interactive (*Acc M1+M2*) and non-interactive scenario (*Acc All M3*), averaged for all the algorithms.

also illustrates how a personalized model consistently outperforms the generic model.

Table 6.7 and Table 6.8 summarize the percentage increase between the accuracy results of the initial generic model and the personalized one. The percentage increase measures the percentage change between two values, according to Equation 6.2:

$$\%Increase = \frac{FinalValue - InitialValue}{InitialValue} * 100 \tag{6.2}$$

In this case, $\%Increase$ measures the relative increase of retraining the model ($FinalValue$) when compared against the generic model ($InitialValue$).

In particular, Table 6.8 summarizes the results of 6.7, averaging all the evaluated algorithms. It also includes the average number of instances that were used to train the personalized model. Results show how the interactive strategy improves the classification capabilities of the cascade, particularly in an eventual scenario in which user intervention is removed (*Acc All M3*). Moreover, on average, higher involved users contribute the most to improve the final results of the cascade under that circumstance. In that scenario, the classification rates improve by up to 2.32 % in the case of the most participate scenario and by 1.05 % when only an average of 6 new sequences of data are included to retrain the model.

Therefore, in general terms, a higher level of collaboration with the interactive system leads to better classification results. This improvement is more relevant when the user is ultimately replaced by a third model (*Acc M3*). Considering a scenario in which the interactive stage of the system is eliminated allows us to anticipate how a personalized model with user-specific data would behave if an individual decided to reduce or stop their collaboration with the learning process. However, the accuracy improvement is moderate in the case of continuing with an interactive cascade strategy (*Acc M1 + M2*). In this case, the difference between a more engaged user and a less engaged user is more subtle, although, on average, an improvement of around 1 % is obtained. In the case of maintaining an interactive approach, in which the user continues to participate in the classification process, it is to be expected that subsequent training cycles will continue with this improvement of the classification results in a progressive manner. That is, the more annotated data is included, the better results are obtained for every retraining cycle.

In summary, the results obtained from this evaluation show that an interactive strategy has the potential to improve the model performance. This improvement becomes significant when the participation of a user is higher. Furthermore, it highlights how most involved users could eventually reduce or even stop their participation in the learning process, obtaining a noticeable improvement of the system's performance in return for their initial efforts. Although the provided results may not represent a remarkable improvement in some cases, it should be noted that the number of instances that are added to the personalized model is very modest compared to the original 900 instances that make up the *Initial Train* set of the dataset. Thus, those results are conditioned by the reduced size of the *Test* and *Validation* set available.

6.3 The Smart drink monitoring system

The previous section analyzed the impact that user participation can have on the outcomes of a classification system. Now, in this section, we want to progress in setting up bi-directional channels through which this participation is enabled. In order to do so, user-machine interfaces and interaction mechanisms mediating between IoT and end-users should be devised to dispose of with an ability to provide feedback to enhance future uses of the system.

Following the use case of detecting hydration patterns in office spaces, in this section, we present a prototype of an interactive IoT smart system that will be used to validate user opinion regarding human-machine collaborative spaces. This prototype, coined as Smart drink monitoring system, consists of two parts: (i) a sensor device that monitors users' actions and provides quick feedback, and (ii) two applications through which the user can customize the system, modulate its degree of involvement in the learning process, and provide new data examples to customize the model.

6.3.1 The Smart drink IoT device

This IoT device can be placed in any bottle, cup, or drink container. Its primary function is to record all the actions that users make with the container in which it is placed. Therefore, it is a practical example of how an

Figure 6.5: A sample of the messages that are prompted through the device's display to interact with the user.

everyday object can be augmented with technology to provide new functionalities, in this case detecting people's hydration patterns. For this prototype, the commercial IoT development device M5-StickC was used.

The inertial sensors of this IoT device register movement in terms of acceleration and orientation. When any movement is detected, the captured data is sent wirelessly to an Edge device (e.g., a Raspberry Pi device), where it can be used to discern between the drinking gesture or any other movement.

The main particularity of this device is its interactivity. Following the IML strategy described during this chapter, with this device, it is possible to notify or inquire the user about a performed movement through the included screen and buttons. An example of the range of messages that are shown trough the screen and these interactions is included in Figure 6.5. This way, if the captured data was classified by the system, a message will inform the user about the detected activity. On the contrary, if the activity could not be predicted with the required confidence, the user will be prompted about the gesture he/she has just performed. The user will be able to respond to this query via the available buttons, and his/her response will be recorded.

6.3.2 User management system and Model personalization engine

The User management system and the Model personalization engine are conceived as two interfaces through which users can personalize the system and include their own examples of data in the model. Through them, it is expected to promote an enhanced user integration and voluntary participation, which ultimately encourages user adherence through personalization.

The User management system is the primary application to customize the desired system preferences. Its interface is shown in Figure 6.6. In order to better fit a wider diversity of users, through this interface, it is possible to define the desired level of participation in the proposed interactive scenario. Three levels of involvement are defined: *High*, *Low* or *Intermediate*. This selection determines the number of times the IoT device may ask the user about a specific activity performed. It also includes the *None* option, which involves eliminating the interactive component of the system functionality. The system is converted into a non-interactive or classical approach with no user involvement in the learning process by selecting this option.

In addition, specific to this use case and the OHM Dataset, it is possible to define whether the user wants the system to discern between the type of container used, Bottle or Cup/Mug. This allows to make the classification options more flexible and adapt the system to the user's expectations. Users can always use the same type of container (for example, a bottle from which they always drink water) or vary between the bottle aforementioned bottle and a cup or a mug (where they can drink tea or coffee). This is another example of the flexibility provided by the system. Finally, all the selected options are saved in association with the user name, and their personalized preferences are updated accordingly.

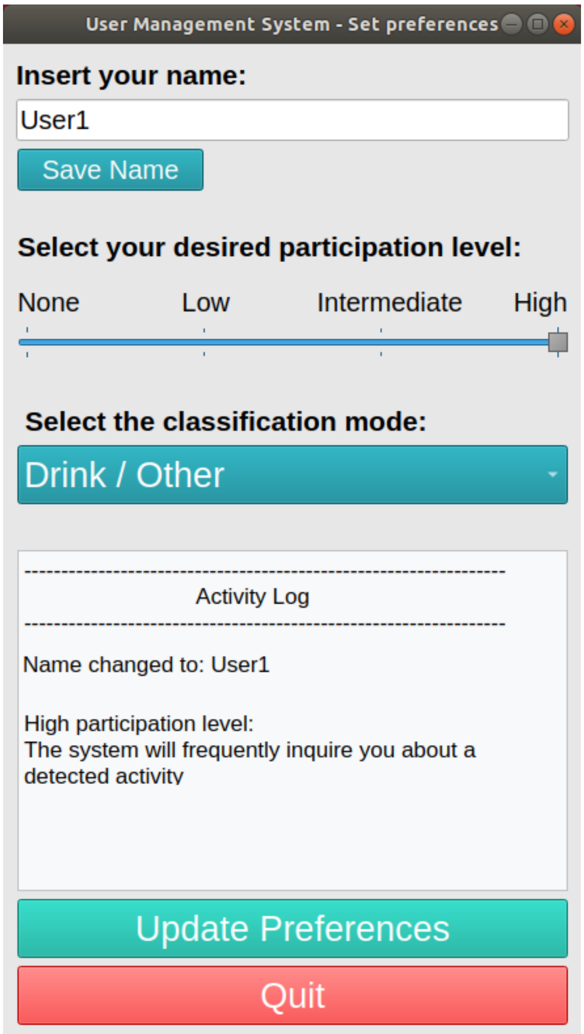

Figure 6.6: The user management application, designed to set the preferences of the system and modulate the desired participation level.

The Model personalization engine is an application designed to help users in the process of personalizing the model. Its interface allows them to easily and visually provide examples of their own movement patterns to retrain the model. Figure 6.7 includes the interface of this application.

To this end, the developed application displays the signals that are being captured by the sensor. This way, users can perform a movement and select the part of the captured signal that corresponds to the drinking gesture to be saved and used later to retrain the base model with user-specific data.

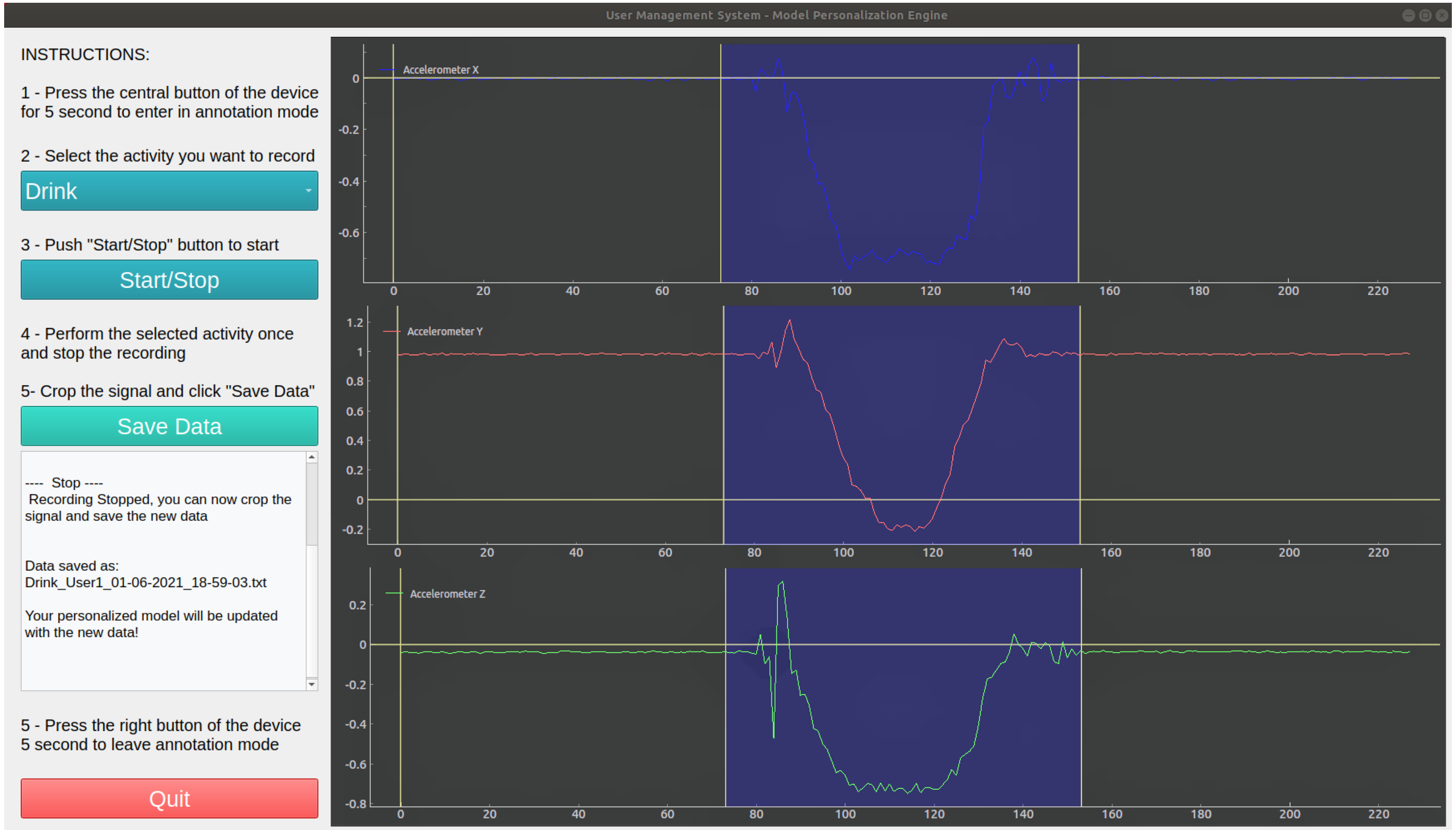

Figure 6.7: The model personalization engine. Through this interface, users can record their own examples of data and save them for further retraining processes of the model.

This idea is aligned with the concept of transparency and explainability of the ML process, as it can provide a clear vision of the kind of data the device captures. In this way, there is direct and visual feedback between the container movements and the signal that the user can see on the screen, which promotes a better understanding of the system's inner workings and how it can identify activities through the signal patterns that are captured.

Following the Edge Computing approach, both applications are designed to run on Linux-based Edge devices, which function as a gateway or concentrator of the information received from the sensor device. The source code for both tools is available in Github [1].

6.4 Qualitative evaluation of the interactive system

The proposed interactive system enables the possibility of collaborating with the system to customize the activities recognition model and, therefore, to improve it. To do so, an interaction between the user and the system is needed. For this reason, the previously presented system was conceived as a tool to evaluate the proposed interactive approach with the presented use case, which is related to the smart workplace context that serves as the basis for this entire dissertation. For that. this evaluation poses a session in which the user is introduced to the Smart drink monitoring system. From this point on, we define this evaluation and present the obtained results.

6.4.1 Evaluation objectives

This study aims to get insights into the degree to which users would be predisposed to be involved in the learning process and the customization of the system. Furthermore, it also aims to address the most suitable mechanisms for motivating those interactions. To this end, two research questions have been proposed. Those are based on the needs identified in the literature regarding the customization and adaptation of IML systems.

[1] https://github.com/OihaneGomez/Smart_Drink_Monitoring

> **Research Question 1:** *Which is the willingness of the user to be involved in the process of obtaining more accurate predictions?*

According to Yang et al. (2015), users are likely to expect to be able to customize the device to suit their individual needs better. A fine-tuning of the models by the end-users (e.g., through a process of re-training and debugging of models/classifiers) may increase its performance. Moreover, including user-specific data also improves the detection capabilities of the model for each user. Thus, it helps to address a wider diversity of profiles and expectations (Kulesza et al., 2015). However, this task may be burdensome for some users if this collaboration is not adequately presented. For this reason, this research question seeks to analyze whether users would be willing to be engaged in improving and customizing a system, to what extent, and the motivations for doing so.

> **Research Question 2:** *Which is considered a good trade-off effort between the improvements and the time commitment?*

According to Meyer et al. (2014), there is a relationship between the quality of the data, in terms of accuracy and reliability, and the user's effort to acquire it and collaborate with the learning system. This effort makes it necessary to consider the trade-off between the user's involvement and the possible improvement results. If this effort is higher than the perceived value of the system and no tangible benefits are obtained, it may not be considered as rewarded (Masson et al., 2016). In addition, the commitment to maintain this collaborative dynamic over time can become arduous for some users. In consequence, greater flexibility is needed to better adapt the interactive system to the context and the involvement of every individual (Chung et al., 2017). Thus, with this question, we aim to understand better what incentives users need to maintain their commitment over time and what they would expect to obtain in return for doing so. In addition, we address how the flexibility of modulating their participation at any given moment is considered favorable to contribute to their involvement in the learning process.

6.4.2 Method and procedure

As has been previously introduced, this study focuses on the participants' perceptions regarding their role in the described interactive learning process. The study protocol includes several tasks to familiarize participants with the interactive system and a semi-structured interview on the experience (Longhurst, 2003). A total of 12 participants were recruited through convenience sampling to participate in this study (7 men, 5 women). While all the participants are middle-aged and their works involve medium to high contact with the technology, not all of them were familiar with AI techniques nor the concept of IML.

Sessions were designed to take 20 minutes and were conducted presently with audio recording. First, participants were introduced to the concept of IML and the designed interactive system using the interaction tools included in Section 6.3. Then, the interviewer described each personalization option and asked participants to practice sample interactions. Those interactions included a data sampling task to familiarize users with the system's interactive capabilities and an exploration of the available preferences to illustrate the available customization options. For evaluation purposes, all data gathering tools were fully functional, but for those interactions that involved identifying a specific gesture, a *Wizard-of-Oz* testing was followed to improve the outcome of the evaluation session (Dahlbäck et al., 1993). On those occasions, participants were placed in a series of scenarios where they could interact with the designed system in a controlled way, that is, being directed by the session moderator.

We evaluated two interaction mechanisms for the data sampling task: a reactive approach and a proactive approach. Participants had the chance to test both approaches, starting with the reactive one. In the reactive approach, participants were invited to interact with the IoT device using the simplest interaction mechanisms (i.e., pushing a button) to confirm the correct label for a performed movement. Moreover, the User management system was presented. By means of this application, participants could set their desired level of participation from four levels (None, Low, Intermediate, High) that determine the degree of involvement that the user would have in the process of annotating new data. For that, two different procedures were followed to avoid the learning bias when interacting with the system. Thus, 6 participants were introduced to the IoT device before setting their desired preferences, while the rest were asked to set their preferences after using the device. The first group had the option to change their initial decision after interacting with the device.

Then, the proactive approach was presented. At that point, participants were invited to record new data examples using the Model personalization engine. In this approach, we evaluate the participants' willingness to use a system that allows them to proactively give examples that can be helpful to customize the detection of activities (e.g., repeating a gesture or marking a starting position). Their thoughts regarding these data-gathering mechanisms and their preferred interaction mode (reactive or proactive) were collected.

During the whole experimentation, participants were invited to respond and discuss the questions that were asked about their perceptions and preferences regarding the system. To finish with, they had the chance to make any additional comments.

6.4.3 Main findings and insighsts

The objective of analyzing the provided feedback is to find out the initial predisposition of participants to be involved in the customization of the system and their preferred interaction mechanisms to do so. Moreover, we aim to obtain some insights into the users' opinions regarding the flexibility given to decide how they want to participate in the learning process and whether this possibility may change their initial perceptions toward the proposed interactive scenario.

In order to evaluate it, we have organized the empirical findings of the described experimental evaluation around three themes: (i) initial perceptions of the interactive system, (ii) the involvement modulation and device's feedback, and, finally, (iii) the proactive interactive strategy. The first two themes refer to the Smart drink IoT device and the Users management system. That is, the interaction with the device itself for the reactive learning process and the possibility of changing the frequency of this interaction. For the last theme, the proactive approach was introduced through the Model personalization engine and was compared to the previous interaction mechanism. To finish with, the main findings of the evaluation are summarized and analyzed. These will be related to the introduced research questions in the final conclusion of this chapter.

Initial perceptions of the interactive system:

After describing the purpose of the interactive system, participants were questioned about their general perception of it. All participants agreed on having a positive or very positive attitude towards the possibility of interacting with the system. Participants highlighted its adaptability and the idea of personalizing the recognition of the activities. Some participants described their initial thoughts as positive because they were interested in the system being able to recognize their own movements. In this regard, they expressed no doubts about the possibility of the system learning from them. One participant addressed that he/she was confident in the system's ability to learn from any user and be adapted to them. In fact, this increased the reported usefulness of the interactive prototype.

Those who had the option to interact with the device before being introduced to the management interface focused their initial opinions on those aspects related to the device's purpose or usability. For example, one participant expressed finding the device beneficial because of its health-related purpose. Two participants emphasized their interest in using it on a daily basis or even in the possibility of taking it to meetings, considering that the feedback mechanisms were not intrusive. In some cases, as they were not yet familiar with the option to modulate their potential participation in the interactive proposal, the system's learning process raised doubts about an excessive frequency and number of interactions needed to provide data to the system. For example, one of the participants found the idea of annotating new data very interesting, as long as it does not force users to do it every time a movement is detected or when they are busy.

In contrast, those participants who were initially introduced to the management interface focused their preliminary opinions on the idea of being able to modulate the level of participation. In fact, this was one of the most positively perceived characteristics of the system. In the following, we explore in detail this characteristic of the system.

Involvement modulation and device's feedback:

When asked about the possibility of setting their system's preferences, participants reinforced their positive perception of being able to adapt the system to their personal needs. In line with the first impressions gathered in the previous point, they were particularly into the idea of customizing and modulating their involvement level. Two participants described this option as the most valuable characteristic of the system. At the same time, another one expressed his/her interest in the possibility of modulating it at any time, while another participant suggested that he/she would like to have always the option to change it.

Moreover, to explore whether participants wanted to be involved in the system's learning process, they were asked to select their desired participation level in the management interface, choosing one option from the four available (None, Low, Intermediate, High). The most common response was "Intermediate", selected by 7 participants. "High" was selected by 3 participants, while "None" and "Low" were set by 1 participant each. When analyzing those responses on the basis of the two types of intervention carried out, no clear pattern is observed in these responses between those who were first introduced to the device or the management interfaces. Those who were invited to select this level before interacting with the device had the chance to change it afterward. None of them changed their selection.

Those who chose a high participation level indicated that they were likely to lower it over time depending on the system's response. Therefore, highly motivated participants would be willing to collaborate as much as possible with the system at the beginning. Then, progressively, they would decrease their involvement once the system has sufficient data to learn from them. In some cases, the choice for the intermediate level was motivated by the prudence of waiting to find out what each of the options entailed in the daily interaction with the device. On the other side, the participant who decided to choose a low frequency said that he/she did not want to be asked more than twice a day. However, despite this initial reluctance to collaborate with the learning system, he/she was very much in favor of having the chance to make this choice. Finally, the person who opted for having none participation claimed that, although he/she does not find intrusive the way of interacting with the devices and receiving notifications from the system, he/she does not like notifications of any kind. Hence, he/she tries to reduce them.

Two other participants shared this last opinion regarding the system's feedback. In this regard, one of these participants highlighted not being obliged to respond to the prompted interactions. The rest of the participants considered that the way the system interacts with them via on-screen notifications, as well as using device buttons to respond to those interactions, was an appropriate mechanism that they would not find intrusive. Furthermore, two participants argued that the system's queries were not disruptive from their tasks since they are integrated as an additional step linked to the fluid intake action.

The proactive interactive strategy:

When introduced to the possibility of recording new data samples to retrain the model using the Model personalization engine, most of the participants declared being willing to use it. Only two users showed no interest in using it. The first one claimed to have too many things to do daily to spend time on one more. The second said he/she did not want to make any extra effort. In both cases, they declared that the first option (the reactive interaction through the IoT device) was more convenient for them. Among those willing to use this application, two of them considered that the reactive approach was better for their routine but agreeing on using this application if necessary. The rest of the participants opted to provide initial data to the system proactively instead of waiting for the system to ask them to do so.

The reasons behind this decision are diverse. In this regard, one participant indicated that it was understandable that each person may have a different opinion, highlighting the provided ability to choose among them. Several participants uniformly reported that they preferred giving more examples at the beginning and then fewer throughout their use. In fact, those who were willing to do this did not rule out continuing the interaction using the reactive method. In this regard, some participants stated that they would like to use it initially to provide some examples of data and create a personal database to personalize the model. Then, they would continue involved in the collaborative process to increase this collection of examples progressively. Another argument in favour of the Model personalization engine was the convenience of this proactive approach. As one participant stated, recording examples of data is a task that only takes a moment, and that can be done several times when someone has time to do it. One participant also justified selecting the proactive approach correlating it with his/her interest in learning more about how the system works.

Additionally, during this part of the interview, some ideas emerged on how to improve this proactive data annotation process. To this end, one participant suggested simplifying the process as a way to avoid other people discarding this option. Another participant commented that, although he prefers this form, he understands that it may be less user-friendly for some people. Therefore, he/she suggested gamifying the process or giving additional motivation to the user to extend this collaboration over time.

6.4.4 **Summary of findings**

Participants reported a medium-to-high initial willingness to participate in the learning process. In general terms, the idea of generating their own data to improve the outcome of the activity detection system was positively received. Thus, participants showed a remarkable interest in collaborating with the learning process. This interest is reflected in the summary of the selected participation level included in Figure 6.8. Participants were also keen on the purpose of the system, as they saw an added value in the health-related problem that the presented use-case addresses. Consequently, we found overall positive experiences elicited after being introduced to the smart device for the

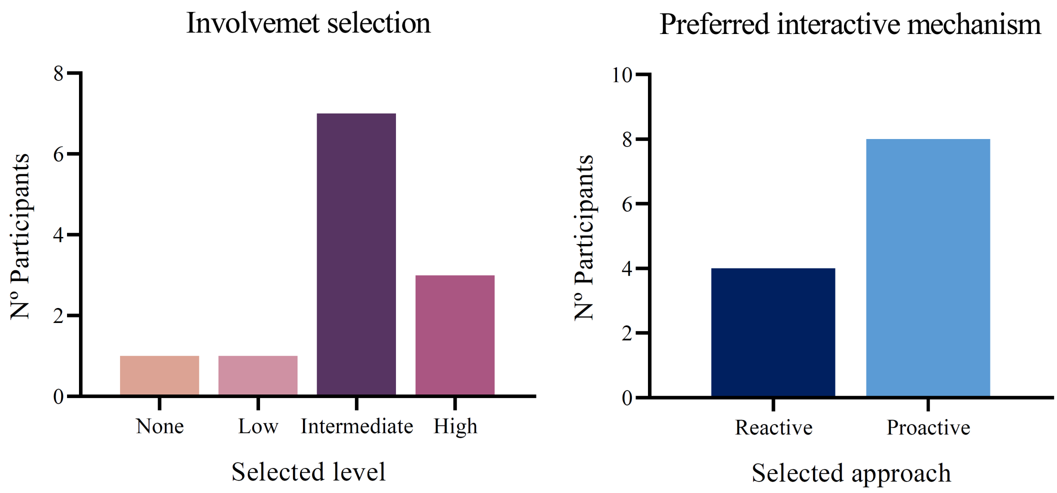

Figure 6.8: Participants' responses when inquired about their initial involvement level in the learning process and their preferred interactive strategy.

first time. This positive experience, together with the health-related purpose of the prototype, also motivated their interest in the collaborative scenario that the learning capabilities of the system enable. Hence, high satisfaction levels about the possibility of modulating the desired participation level were also reported. This flexibility stands out as one of the most highlighted characteristics of the introduced system.

Participants were comfortable with the idea of being part of the learning process. Furthermore, the presented alternatives for personalizing the model were perceived as adequate to improve the performance of the system. Which one was preferred, on the other hand, showed a split opinion. For this reason, when it comes to the most suitable approach to promote this collaboration, there is no uniformity in their choices, as can be observed in Figure 6.8. That is, between the choice of the proactive approach (voluntarily recording a set of data using the Model personalization engine) or the reactive approach (waiting for the system to ask you about the data through the Smart drink IoT device). In general terms, most participants found more suitable for their routine the proactive approach. Furthermore, those users willing to give an initial data input to the system were also willing to continue collaborating with the system through the reactive method. In this case, they saw the proactive strategy as an excellent way to accelerate the learning process, with the idea of reducing their further participation once they see an improvement in its detection capabilities.

The fact that there is a wide variety of opinions among participants highlights the importance of allowing users to adapt the system to their preferences and daily routine. This enhances the perceived usefulness of the technological solution and eventually could prevent future disengagement, as users can adapt the system's functionality to their own context and the evolution of their emotions towards the system.

6.4.5 **Limitations**

Before concluding this chapter, the limitations of the conducted evaluation need to be addressed. The analyzed qualitative opinions correspond to participants' initial thoughts when the system was presented to them. For this reason, the obtained conclusions can not be extrapolated to determine the adoption of the proposed system over time. Hence, they only can serve as a predictor of the predisposition that may exist to do so. Additionally, the sample of this experimentation is not wide enough to cover a broad population of users. All participants were aged between 20 and 50 years, and some of them had a high background in technology and AI techniques. Even though no specific differences in the provided responses could be found in those keener on technology, the presented findings should not be extrapolated to every individual and, particularly, to other profiles that may have a lower technological background. Further research needs to be done to understand how those profiles would like to be involved in the collaborative scenario proposed by this interactive system. However, this approach provided promising insights, thanks to its flexibility, to facilitate this involvement.

6.5 Summary and Conclusions

This chapter has evaluated both quantitatively and qualitatively the suitability of an interactive approach to enhance the recognition of human activities in collaborative work environments. For that, the model cascade proposal introduced in Chapter 5 has been adapted to incorporate the willingness of the user to be involved in this collaborative scenario. This is done by adjusting the confidence thresholds that rule the classification process according to desired participation level of the user. This way, the preferences of the user, who is in charge of providing new data to the model, are included as a parameter in the interactive classification process. This parameter modulates the number of times the system may prompt users to provide the true label for a specific detected movement.

For the presented use-case, the quantitative data obtained throughout the experimental evaluation included in Section 6.2 validates the capabilities of this strategy to (i) modulate the frequency in which the system may inquire users to label new data and (ii) improve the classification capabilities of the system by means of user-annotated examples of data. Furthermore, even though those improvements are modest, these results provide insights into the positive effect of a continuous interaction between the user and the system.

Then, the suitability of the proposed IML approach and the potential convenience of a flexible implementation of such an idea in work environments was evaluated. For that, we have presented in Section 6.3 a prototype of an interactive system. This system comprises an IoT smart device and two applications through which users can set their preferences, modulate their degree of involvement and proactively create new samples of data to personalize the model. By means of this prototype, 12 participants took part in an evaluation study consisting in a 20 minutes semi-structured session where the smart interactive system was introduced and their opinions collected. Finally, on the basis of the research questions formulated in Section 6.4, the outcomes from the conducted evaluation have been reviewed.

In response to the defined Research Question 1, we have explored the participants' expected role within the collaborative scenario described. In this regard, participants were willing to participate in the learning process of the system and reported a positive attitude towards the idea of improving the classification results with their own data. On top of this, the characteristic of the system that was found most valuable was the flexibility to adapt this collaboration to their preferences and routine. These customization capabilities, brought together with the possibility of defining how much every individual wants to be involved in the process, aim to adapt a system's functionality to increase its relevance to an individual. According to Mugge et al. (Mugge et al., 2009), by personalizing a product or service, users direct time, energy, and attention to it. In this sense, users can create custom solutions they feel more attached to and conserve a sense of continuous novelty. This latter feature is paramount to enhance the perceived usefulness of using the system throughout the time and prevent future disengagement.

To analyze the potential commitment of the participants to dedicate time and effort in the learning process, we evaluated two independent approaches for providing new data to the system. One of them was integrated into the interaction mechanism of the Smart drink IoT device. The other one explicitly involved recording new samples of data using the available software tools. This part of the evaluation is related to Research Question 2. Although participants showed some division of preferences, in general terms, most of them were likely to spend their time voluntarily providing new samples of data to the system. With it, they expected to improve its detection capabilities thanks to this data and reduce the number of interactions afterward. Nevertheless, even after this initial effort, they continued to agree with the idea of responding to the system's interactions, as long as the frequency of those interactions was progressively reduced. This participation is motivated by the perceived usefulness of the system itself and the learning process, and it is related to the concept of technology appropriation and the sense of technology ownership (Kirk et al., 2015). This concept stresses the importance of empowering and involving the user in maintaining and improving a digital system to boost self-efficacy and satisfaction. In this regard, making individuals participate in the learning process can contribute to developing feelings of psychological ownership concerning digital technology itself, where individual appropriation is experienced as a subjective form of customization.

In summary, in this chapter, we have provided three contributions: (i) the evaluation of the interactive model cascade and its suitability to modulate users interactions, (ii) a set of tools to support users in their collaboration with IML approaches, and (iii) a qualitative evaluation to gather participant insights about the designed intelligent system and their role within this collaborative scenario. This study revealed that participants are primarily satisfied with the idea of adapting the system to their needs. Thus, the expectations of end-users and their role in participatory approaches are pivotal factors to consider when planning interactive strategies that require additional user involvement. On this subject, we propose conceiving flexible solutions according to those expectations to fit better the needs of every individual that populate smart environments and, in particular, smart workplaces.

AI don't know where I am going, but I am on my way.

Voltaire

CHAPTER

7

Conclusions and Future Work

THIS chapter seeks to provide an overview of the described work and the contributions this dissertation has generated. Therefore, the hypothesis and objectives introduced in Chapter 1 are analyzed to evaluate to what extent they have been achieved. Some of the results of this research have been disseminated to validate their contribution with the research community. That is, the accomplishment of those objectives are objectively support by the publications reviewed in this chapter. To conclude with this dissertation, future lines of work and ideas are proposed to continue the presented research line, and some final remarks in the form of discussion are presented.

The rest of the chapter is structured as follows: Section 7.1 summarises the work done in this dissertation and the conclusion obtained throughout it, Section 7.2 lists the contributions from this work, Section 7.3 validates the hypothesis and the objectives that were set in Chapter 1, and Section 7.4 includes the scientific manuscripts published during the development of this dissertation. To finish with, Section 7.5 introduces the future work, while Section 7.6 draws some final remarks.

7.1 Summary of work and conclusions

Current IoT architectures and frameworks usually neglect the degree of confidence and involvement that humans require to adopt these technologies in their everyday lives. For this reason, the reviewed IoP concept represents an evolution of the traditional idea of the IoT, placing the role of the user in these smart spaces as the central element of the intelligent system. However, IoT deployments still defy legal norms and surveillance policies due to the close monitoring and collection of personal information and the possibility of this data to be used by third parties without explicit consent. This is of vital importance in an environment such as the workplace, where even seemingly innocuous and apparently meaningless data, such as the number of steps an employee takes during the working hours, can serve to infer personal information or be linked with the employee's productivity or work performance. Therefore, creating confident and safe scenarios for integrating technologies in the workplace is necessary to improve the employees' perception of emerging technologies when used for beneficial purposes such as promoting healthier habits.

7.1.1 Addressing the workplace problematic: barriers and facilitators

From the revision of the State of the Art carried out in Chapter 2, we have discovered the importance of understanding the implication of privacy requirements and user involvement. The most relevant challenges that should be taken into account before deploying an IoT solution have also been reviewed. We found that the existing body of knowledge has widely addressed privacy concerns as one of the main obstacles to the acceptance of emerging technologies in smart spaces. This way, the potential of wearable devices in the workplace to improve employees' physical well-being and reduce work-related injuries has been validated. However, the study of the literature also reveals that the technological and social challenges hinder its total adoption. In this regard, IoT technology could be considered more effective in promoting healthier habits when they play a supportive role in which employees do not share their collected data with their companies or external organisms.

However, aspects such as who controls the data being captured or to whom it belongs seem to have gone more unnoticed in the literature. Moreover, traditionally IoT implementations usually do not address those issues. Instead, they usually centralize data processing capabilities in external servers, increasing the gap between the user and the technology. We stress this circumstance in Chapter 3,in which relevant insights in this regard have been identified from a conducted quantitative survey among 524 people. From the obtained results, the actual scope of the barriers regarding data gathering that arise in work environments have been discovered, as well as the potential implications of these factors. Our findings indicate how privacy concerns negatively affect participants' attitude toward integrating smart technologies in the workplace. For this reason, respondents' perception regarding technology and data collection in those spaces are related to the possible cost-benefit of using such systems and generating such data. When participants are placed closer to their data, knowing that they are the only ones granted with privileges to own and control the generated information, they feel more confident with both the technology and the final purpose of data collection.

Those findings are in line with the outcomes extracted from the existing literature. However, our contribution extends the concept of privacy to the ability of the user to be the owner of the information they generate and to determine how, when, and to whom this personal information can be disclosed. Furthermore, the conducted research stresses the importance of addressing the users' role within the system and the control and customization options users expect to have. These are significant factors to consider to fit their needs and expectations in those intelligent environments. Thus, this study has contributed to the body of knowledge by better understanding the individual perception towards the data control and ownership concept in the work environment. Furthermore, our findings settle the conceptual basis for the rest of the work that we have presented in this dissertation.

For this reason, this thesis document highlights the importance of considering first data privacy and control concerns, together with user's involvement, when deploying IoT devices in a human-centric work environment. To this aim, we have divided our research in two parts: (i) the integration of local processing in resource-constrained devices at the Edge to increase the privacy and control of the personal data; and (ii) the adoption of HiTL approaches to involve the user in the learning process of the system through Hybrid Intelligence and IML. In essence, we have proposed an Edge Computing approach in which interactive systems are hosted closer to the user, in devices they can own, to foster better user perception and trustworthiness in IoT environments.

7.1.2 ***Privacy-by-design*** through Edge Computing and embedded intelligence

Following this setting, the dissertation first proposes applying this hybrid-intelligence approach in a use case scenario where ML algorithms try to make decisions related to the individual's data and adapt to their preferences and needs, creating flexible systems where the user feels more involved. All of this, while maintaining the requirements of proximity that applies to the concept of embedded intelligence. Hence, Edge devices were not only endowed with the ability to analyze new data and perform inference tasks but also to retrain and personalize models according to new data provided by the user. Data gathering, recognizing human activities, model training, and user involvement were the main ingredients in such a process that aims to fit deployed intelligent solutions to the context and users' requirements.

For all the aforementioned reasons, this dissertation has proposed new ways to integrate the whole classification pipeline at the Edge of the network to perform on-site local training and decrease data exposure. In this regard, Chapter 4 has contributed with an optimization strategy for improving the efficiency of HAR applications in resource-constrained devices. Optimizing its performance is a matter of finding the optimal point in which the balance between the resource requirements and the predicting capabilities of the system can pay off. The proposed method has proven to be suitable to substantially improve the performance of classification tasks and highlights the importance of understanding the initial data and studying the most relevant characteristics of the signal. In fact, the results of the conducted evaluation show that there is a highly non-linear trade-off to make between the computational cost, in terms of processing time, and the achieved classification accuracy. This implies that small reductions in detection rates can lead to significant improvements in classification times and resources. However, the presented approach decreases its efficiency when complicated classification problems are targeted (e.g., multi-class classification). That is, a single optimized model may not be not enough in all use cases and scenarios to obtain a satisfactory cost-accuracy trade-off in complex multi-class scenarios.

For this reason, Chapter 5 covers the potential limitations of the preceding approach and facilitates the integration of classification problems that, due to their complexity, might not reach a satisfactory balance between their computational cost and their classification results with the previous strategy. To that end, this chapter has validated the suitability of an ensemble learning strategy, based on a discriminative cascade of models, to ease the optimization of data processing in Edge-based intelligent systems. This cascade combines different classification stages ruled by probability thresholds to better fit the system to the complexity of the input data and save the available resources. The performed evaluation demonstrates that the computational cost of classification tasks can be reduced without a significant accuracy loss in the evaluated multi-class scenarios. Furthermore, the results of more complex ensemble techniques (i.e., model stacking techniques) can be matched or even improved with less computational expensive approaches.

To sum up, the previously reviewed chapters validate how, through optimization techniques, it is possible to increase the capacity of local devices to carry out the fundamental parts of data analysis in a more efficient way and replace remote Cloud data centers. This approach ensures that the collected sensitive information remains on local devices and not on third-party servers. Thus, it favors privacy and users' control over the information, keeping it close to the place where it is produced even when the personalization of models using ML techniques involves retraining them.

7.1.3 Hybrid Intelligence: the role of the user in interactive approaches

When data classification capabilities are brought closer to the user, the next step is to create spaces where human and machine intelligence collaborate. That is, making possible the personalization of models. Placing the human in the loop means that appropriate participative mechanisms are devised to help users managing their interactions with intelligent systems. Consequently, Hybrid Intelligence addresses the importance of the role of the users in creating a participatory scenario where both the classification capabilities of the system and the user's involvement are improved. For that, the challenge is to provide intelligent solutions with enough flexibility to meet the different needs and expectations of individuals in participatory environments.

In this sense, Chapter 6 proposes new schemes personalizing ML models while improving the classification results. In this context, the user is involved in the learning process by providing new annotated data when inquired by the system. In order to do so, we have categorized the role of the user according to their involvement level in the interactive system. Then, we have adapted the cascade model of Chapter 5 to include this willingness to participate in this learning process as a variable to modulate the system behavior. Thus, the combination of IML and the optimization proposal of the cascade model seek to increase the role of the users while still preserving the external accessibility of the data. The initial quantitative results of the proposed strategy show that, for the evaluated use-case, using user-dependent data to retrain the cascade can create models that outperform the baseline pre-trained ones. Although the size of the sampling data limited the obtained results, this evaluation shows the potential improvement that successive retraining phases can entail for the model. Moreover, it also illustrates how the more involved the user is, the bigger this improvement becomes.

For this reason, it is necessary to understand to what extent users are willing to be committed to helping the system in this learning process and how to engage them to do so. Aiming to better understand those factors, a qualitative study has been conducted to determine the users' requirements to be involved in such collaborative scenarios in the workplace. This evaluation used a prototype of an interactive IoT smart system to collect participants' opinions regarding human-machine collaborative approaches for work environments. The implemented prototype consists of an intelligent system for tracking hydration habits in office environments that incorporates the option to create data examples to improve its detection capabilities.

The insights obtained from the analysis of the qualitative insights show the potential benefit of an interactive approach. Participants found favourable the option of customizing the system and improving the detection rates using their own data. Moreover, the possibility of defining their involvement level was the most valuated aspect of the provided prototype. This shows that just as intelligent solutions should not bypass the user and left them apart, neither should the degree to which users want to be part of those solutions be assumed. Both the quantitative and qualitative results obtained from Chapter 6 validate the importance of considering the particularities of every individual and their role in intelligent solutions to create flexible systems that can be adapted to fit their expectations and perceptions.

In essence, and considering the research process described above, we can conclude that this dissertation contributes to improving the perception of IoT technologies in the workplace by addressing the main concerns identified both from the body of the literature and the study performed in Chapter 3.

7.2 Contributions

As illustrated in previous chapters of this dissertation, the knowledge and findings obtained from this thesis are diverse. As a consequence of those findings, the scientific and technical contributions generated from this dissertation are presented in this section. Those contributions are linked to their respective objectives, described in Section 1.2.

7.2.1 Scientific contributions

- **An analysis of the implications of privacy concerns and data control when deploying IoT systems in smart workplaces.** Through the study presented in Chapter 3, it has been assessed the perceived risks in sharing personal data and the value of giving users full control and ownership over their information in such spaces. This study analyzed 524 responses to an online questionnaire specifically designed for this purpose. The obtained conclusions indicate that placing greater emphasis on the importance of data control and ownership concepts is essential when addressing the barriers and concerns for the integration of IoT technologies in the workplace. This analysis contributes to Objective 2 of this dissertation and provides a broader view than the studies reviewed in Objective 1, in which we examined the State of the Art. Besides, the insights exposed define the design requirements that support the following contributions.

- **The design and implementation of a novel strategy to integrate and optimize inference and training stages of HAR applications at the local Edge to promote data privacy.** The approach presented in Chapter 4 analyzes all the factors affecting the classification process and proposes a novel approach to create stand-alone Edge solutions by addressing the cost-accuracy trade-off of ML techniques. This approach relates to Objective 3 of this dissertation and contributes to the open challenge of improving ML inference at the Edge and the training efficiency. The latter has been scarcely addressed in the revised existing literature to the best of our knowledge. Various datasets and Edge devices are considered for the empirical evaluation of this proposal. The set of conducted experiments addresses Objective 4 of this dissertation. Based on this evaluation framework described, it validates the suitability of proposed approaches to improve the processing capabilities of both stages in resource-constrained devices when compared against reference classifiers. In fact, this highlights that understanding the initial data and optimizing the process from the early stages of the classification pipeline is crucial for improving the efficiency of embedded intelligent systems.

- **A method for improving the performance of complex classification tasks using model ensemble techniques in resource-constrained devices.** This contribution also relates to Objective 3 and seeks to enhance the previous approach when dealing with the inference stage of more complex classification problems. The proposed method, presented in Chapter 5, describes a discriminative model cascade strategy that separates complex models into multiple simpler classifiers. This cascade has been designed to better fit the system to the complexity of the input data and exploit the computational capabilities of Edge devices more efficiently. Following Objective 4, the performed evaluation proves the potential of this strategy to improve the performance of the execution of HAR applications in constrained local equipment. Together with the previous contribution, both approaches aim to validate the part of the hypothesis that stresses how enhancing local processing can preserve the privacy of the capture data. That is, maintain it closer to the user by avoiding any externalization of the data processing stages. Therefore, hosting all the data on a local device to which only the employee has access fosters that the data belongs to them (data ownership) and that no other can manage their information (data control and appropriation).

- **A novel interactive approach that integrates hybrid intelligence at the Edge and allows users to determine their participation level.** This contribution is presented in Chapter 6 and combines the privacy-preserving strategy described in Chapter 5 (that corresponds to the previous contribution) with an IML approach in which the user is involved in the learning process by providing new annotated data when inquired by the system. Furthermore, as a novelty from traditional IML approaches, the participation willingness of the user is included as a variable for the cascade system modulating the frequency the system may interact with them. This is actually in line with Objective 5. The evaluation results support the potential of this participatory approach to create personalized models that improve the system's classification capabilities.

- **An exploration of the potential implication of hybrid intelligence approaches in the workplace and the users' requirements to be involved in collaborative scenarios.** After defining the interactive approach that fulfils Objective 5, a qualitative study has been conducted to understand the main challenges of IML solutions in the workplace. A total of 12 participants took part in semi-structured sessions where they were invited to interact with a proof of concept of an intelligent system. Their opinions with regard to this IoT solution were collected and analyzed. A parallel technical contribution arises from this evaluation since an interactive prototype for evaluating the use case of detecting hydration patterns in office spaces was implemented. Such a system consists of two parts: (i) a sensor device in charge of capturing motion data and providing quick feedback to the user, and (ii) the interfaces through which the user can customize its operation, modulate its degree of involvement and provide new data to the system. This evaluation concludes with Chapter 6. The obtained findings are related to the role of the user in those spaces and their willingness to be involved in such a collaborative scenario, improving the system with their own data. From the obtained insights, we set to contribute to the existing literature by better understanding the requirements for envisioning more human-centric and confident spaces where human and machine intelligence collaborate. This is stressed in Objective 6. In this regard, more positive interaction can be created in the workplace if IML solutions are designed with enough flexibility to adapt their interaction requirements to the context and the level of user involvement. Based on those findings, the presented approach in which the user can personalize and improve the system while modulating their participation highlights the relevance of providing user control over all aspects of the intelligent system.

- **A dataset containing 1000 instances of labeled data for classifying office employees' hydration patterns.** The Office Hydration Monitoring (OHM) Datasets includes a taxonomy of 25 variations of different interactions that could be made with liquid containers. These interactions are grouped into three classes (drink from a glass, from a bottle or other kind of interactions), recorded from 10 different subjects using a wearable sensor placed on the liquid container. This dataset was created with the idea of contributing to the community with a semi-controlled activity dataset that resembles real-world scenarios.

7.2.2 Technical contributions

The supporting materials for this dissertation have been released in Github: For Chapter 4, the supplementary materials for its related publication "Exploring the Computational Cost of Machine Learning at the Edge for Human-Centric Internet of Things" are publicly available [1]. For Chapter 5, the materials for the evaluation of the proposed model cascade are also available [2]. Finally, for Chapter 6, the source code of the IML software tools presented in that chapter has been published [3].

7.3 Hypothesis and objective validation

At the beginning of the dissertation, in section 1.2, the following hypothesis was formulated:

> **Hypothesis.** By integrating local processing and Hybrid Intelligence at the Edge, it is possible to create a system that (i) promotes the privacy of personal data and (ii) helps users to retain the control of the system and their associated personal data to improve their perception of emerging technologies in work environments.

[1] https://github.com/OihaneGomez/Exploring_Computational_Cost_ML_IoT
[2] https://github.com/OihaneGomez/Model_Cascade_Optimization
[3] https://github.com/OihaneGomez/Smart_Drink_Monitoring

Then, in order to validate it, the following goal was proposed:

> **Goal.** To design and implement strategies for a system to be hosted at the local Edge, within resource-constrained devices, that entails Hybrid Intelligence through a Human-in-the-Loop perspective to (i) promote the privacy of personal data and (ii) help users to retain the control of the system and their private data to improve the perception of emerging technologies in work environments.

To achieve this goal, more specific objectives were defined. Although these objectives have been introduced in the review of the contributions of this dissertation, the following is a more specific description of how each sub-objective has been addressed in this dissertation:

1. *To study the current start of the art on IoT approaches applied to smart workplaces as well as Embedded ML and Hybrid Intelligence solutions.* Chapter 2 explores the most relevant works related to the three main concepts that articulate this dissertation. Moreover, to conclude this study, Section 2.4 presents a table including the identified research gaps and future lines to better integrate emerging technologies in work environments. In this section, Table 7.1 correlates those gaps with the objectives addressed during this dissertation.

2. *To identify the main barriers and challenges of the integration of technology in such spaces and the role that privacy concerns and users' control over the information they generate may have to support this integration.* This objective is met in Chapter 3, in which the conducted survey extends the scope of the privacy concerns in the work environment to include the concept of data ownership and control. This analysis complements the conclusions extracted in the previous objective. The results of the conducted study are described in Section 3.3. A summary of the extracted conclusions is included in Section 3.4.

3. *To design and implement suitable strategies to incorporate and optimize the training and inference stages of classification techniques for HAR applications in local Edge devices, preserving the external accessibility of data through a privacy-by-design concept* This objective has been tackled throughout Chapter 4 and Chapter 5. In Section 4.1, a novel optimization strategy for improving the performance of training and inference stages is proposed. Moreover, in Section 5.1, a discriminate cascade of models is presented, complementing the efficient integration of classification tasks in resource-constrained devices.

4. *To identify an appropriate evaluation methodology for the optimization strategies and quantitatively validate the results obtained with and without the use of the proposed strategies.* For the evaluation of the previous objectives, the procedure and methodology followed is described in Section 4.2 and Section 5.2. Every aspect of the definition of the experiment (e.g., selected datasets and devices, or initial data preprocessing stages) is described. Then, in Section 4.3 and Section 5.3 the results of the performed experiments are shown, quantitatively validating the optimization capabilities of both approaches.

5. *To extend the scope of Hybrid Intelligence, adapting the proposed privacy-by-design approach to include human intelligence in an interactive scenario where users have control over both the intelligent system (i.e., personalizing it) and their data.* This objective is included in Chapter 6. In that chapter, the adaptation of the cascade optimization approach to an interactive scenario is described and evaluated in Section 6.2. Then, in Section 6.3, the interaction mechanisms designed to mediate between IoT devices, intelligent systems and end-users are presented.

6. *To discover qualitative insights regarding the interaction between human and machine intelligence in work environments, the potential willingness of users to be committed to helping the system, and how to engage them to do so.* Following the designed system presented in Section 6.3, Chapter 6 continues describing the insights and main findings obtained

from the conducted qualitative evaluation in Section 6.4. The extracted conclusions are summarized in Section 6.5.

Gaps and future lines	Related obj.
The lack of understanding about users' limited control over the information that is shared and the doubts about the ownership of the captured data	Objective 2
Edge Computing needs to overcome the technological challenges that arise when deploying complex IoT solutions in resource-constrained devices The optimization of complex ML systems and on-device training are understudied in the literature	Objectives 3 & 4
HiTL solutions should be flexible enough to rethink their functionality in terms of human goals, contexts, and participation preferences.	Objectives 5 & 6

Table 7.1: Correlation between the objectives of this dissertation and the current gaps and future research lines obtained from the analysis of the literature conducted in Chapter 2.

Taken these previous bullets into account, it can be claimed that all the previous objectives have been accomplished. As it can be observed, Chapter 3 was essential to define the requirements of the designed approaches and strategies. Furthermore, that chapter, together with the main findings of the related work in Chapter 2, identified the barriers and limitation that future IoT solutions should overcome to be part of a human-centric vision of smart workplaces. In this regard, Chapter 4 and Chapter 5 were pivotal for addressing how a more efficient integration of data analysis in local Edge devices can contribute to promoting the privacy, ownership and control of personal data. Finally, Chapter 6 extended this concept to integrate the user in a participatory HiTL scenario and contribute to providing increased control over the system to users.

To finish with, through accomplishing these objectives and the main goal, the hypothesis set at the beginning of this dissertation has been validated. The quantitative results obtained in Chapter 4 and Chapter 5 empirically support the suitability of such strategies to perform complex classification tasks efficiently in constrained settings. Thus, preserving any externalization of users' data and giving them total control over their information. Moreover, Chapter 6 highlights the potential outcomes of personalizing the system and providing mechanisms to modulate the user's role in this learning process. Together with the qualitative insights obtained both in Chapter 3 and in Chapter 6, it can be certified that those previous approaches contribute to creating more confident and interactive solutions that address the privacy, appropriation and control concerns of the users as a way to improve their perception of technology in work environments. Nonetheless, this is validated under the constraints of Section 1.2.

7.4 Relevant publications

Through this dissertation, several scientific manuscripts have been published in order to validate with the research community the advances reflected in this work. The included publications fall into two categories: (i) the ones that are directly related to the specific context of this work and its contributions, and (ii) those complementary works that extend this context to study the role of the user in IoT-based intelligent environments. We also include in this section those supplementary materials that are released and that contribute to the body of knowledge.

7.4.1 International JCR Journals and Book Chapters

A holistic optimization approach for meeting the requirements of ML-based classification tasks in resource-constrained Edge devices was presented in the following journal publication:

- Oihane Gómez-Carmona, Diego Casado-Mansilla, Frank Alexander Kraemer, Diego López-de-Ipiña, Javier García-Zubia. (2020) "Exploring the Computational Cost of Machine Learning at the Edge for Human-centric Internet of Things" In Future Generation Computer Systems. vol. 112.

p. 670-683. DOI: 10.1016/j.future.2020.06.013. JCR Impact Factor (2020): 7.187, Q1. Published in November 2020.

An analysis of the different factors affecting technology perception in the workplace, with a special focus con health-related interventions and privacy concerns, was published in the form of a book chapter:

- Oihane Gómez-Carmona, Diego Casado-Mansilla, Javier García-Zubia. (2019) "Opportunities and Challenges of Technology-based Interventions to Increase Health-awareness in the Workplace" In Transforming Ergonomics with Personalized Health and Intelligent Workplaces. vol. 25. p. 14. DOI: 10.3233/978-1-61499-973-7-33. Published in September 2019.

A co-authored publication explaining the principles of a conceptual architecture to enable *privacy-by-design* solutions with HiTL approaches was published in the following journal:

- Fatima Zohra Benhamida, Joan Navarro, Oihane Gómez-Carmona, Diego Casado-Mansilla, Diego López-de-Ipiña, Agustín Zaballos. (2021) "PyFF: A Fog-Based Flexible Architecture for Enabling Privacy-by-Design IoT-Based Communal Smart Environments" In Sensors 21, 3640. DOI: 10.3390/s21113640. JCR Impact Factor (2020): 3.576, Q1. Published in May 2021.

A parallel work focused on the Smart Cities concept and the importance of understanding the interactions between users and IoT-based objects was published in the following journal:

- Oihane Gómez-Carmona, Juan Sádaba, Diego Casado-Mansilla. (2019) "Enhancing street-level interactions in smart cities through interactive and modular furniture" In Journal of Ambient Intelligence and Humanized Computing. DOI: 10.1007/s12652-019-01577-8. JCR Impact Factor (2019): 4.548, Q1. Published in November 2019.

Moreover, a review on the human factors concerning technology-based interventions for the promotion of people's behaviour change towards energy-efficiency in work environments was presented in the following co-authored journal publication:

- Ane Irizar-Arrieta, Oihane Gómez-Carmona, Aritz Bilbao Jayo, Diego Casado-Mansilla, Diego López-de-Ipiña, Aitor Almeida. (2020) "Addressing Behavioural Technologies Through the Human Factor: A Review" In IEEE Access. vol. 8. p. 52306-52322. DOI: 10.1109/ACCESS.2020.2980785. JCR Impact Factor (2020): 3.367, Q2. Published in March 2020.

7.4.2 International Conferences

An analysis of the optimization opportunities of sensor-based classification systems at the Edge was presented. This paper was recognized with an Honorable Mention as one of the best papers of the conference:

- Oihane Gómez-Carmona, Diego Casado-Mansilla, Diego López-de-Ipiña, Javier García-Zubia. (2019) "Simplicity is Best: Addressing the Computational Cost of Machine Learning Classifiers in Constrained Edge Devices" In Proceedings of the 9th International Conference on the Internet of Things - IoT 2019. p. 1-8. DOI: 10.1145/3365871.3365889. Published in October 2019.

An exploratory analysis containing some initial thoughts on the role of the IoT in wellness promotion in work environments:

- Oihane Gómez-Carmona, Diego Casado-Mansilla, Javier García-Zubia. (2018) "Health Promotion in Office Environments: A Worker-centric Approach Driven by the Internet of Things" In Intelligent Environments 2018: Workshop Proceedings of the 14th International Conference on Intelligent Environments. vol. 23. p. 355 - 363. DOI: 10.3233/978-1-61499-874-7-355. Published in June 2018.

A first approach for outlining the contributions of this dissertation and its possible scope was presented in a Doctoral Consortium and as a Poster:

- Oihane Gómez-Carmona, Diego Casado-Mansilla, Javier García-Zubia. (2018) ”Towards Healthy Office Environments: A worker-centric Internet of Things Approach” In Proceedings of the 12th EAI International Conference on Pervasive Computing Technologies for Healthcare – Demos, Posters, Doctoral Colloquium. DOI: 10.4108/eai.20-4-2018.2276414. Published in May 2018.

A preliminary research work that provided initial insights about the potential of incorporating IoT objects in the workplace:

- Oihane Gómez-Carmona, Diego Casado-Mansilla. (2017) ”SmiWork: An interactive smart mirror platform for workplace health promotion”. In 2017 2nd International Multidisciplinary Conference on Computer and Energy Science (SpliTech) (pp. 1-6). IEEE. Published in August 2017.

A co-authored contribution related to this dissertation analyzing the role of Fog/Edge Computing approaches for addressing the privacy concerns of work environments:

- Fatima Zohra Benhamida, Diego Casado-Mansilla, Oihane Gómez-Carmona, Joan Navarro, Diego López-de-Ipiña, Agustín Zaballos. (2019) ”Smart-Workplace: A Privacy-based Fog Computing Approach to Boost Energy Efficiency and Wellness in Digital Workplaces” In Proceedings of the 1st Workshop on Cyber-Physical Social Systems. vol. Vol-2530. p. 7. Published in October 2020.

A parallel work presenting an enhanced solution for improving smart services in public spaces was presented:

- Oihane Gómez-Carmona, Diego Casado-Mansilla, Diego López-de-Ipiña. (2018) ”Multifunctional Interactive Furniture for Smart Cities” In Proceedings. vol. 2. p. 1212. DOI: 10.3390/proceedings2191212. Published in December 2018.

7.4.3 Datasets

- OHM Dataset: A novel dataset containing 1000 labeled sequences of data for classifying hydration patterns in office environments (Gómez-Carmona and Casado-Mansilla, 2021).

7.5 Future work

The following research lines and future work have been identified inspired by the observations and limitations found throughout this dissertation:

- *Analyze the suitability of other training methods and reducition approaches for Edge devices*: As exposed throughout this dissertation, an intelligent and human-centric IoT solution should rethink its functionalities and adapt itself to the preferences and particularities of the user. In the proposed IML solution, this involves re-training the model to include the provided user-dependent data. Chapter 4 deals with those requirements in local settings and propose optimization mechanisms to improve the time and resources needed by Edge devices to perform this re-training step. In the evaluated implementation of the training process, the entire model is re-trained when a significant number of new labelled instances are available, adding those new observations to the already existing dataset. This technique is selected as a borderline case to test the capability of the Edge devices and the real potential of the optimization strategy. While this is a flexible yet adequate approach to improve the classification performance of the model, new mechanisms can be studied to incorporate new observations to the already trained model incrementally. This progressive learning concept falls under the scope of Incremental and Online Learning methods (Gepperth and Hammer, 2016) and its particular application for personalizing models (Siirtola et al., 2019). For this reason, exploiting the particularities of this type of learning could be an interesting alternative to consider in the deployment of dynamically intelligent systems. In the same line, selecting a training data subset that contains a meaningful representation of the complete set could also be a potential approach for reducing the training stage's computational burden. Thus, future research may point out the optimization opportunities of identifying this reduced subset using Active Learning techniques.

- *Study the feasibility of DL techniques to be applied in the same contexts and whether it is worthwhile compared to traditional methods*: In Section 1.2, the constraints and premises of this dissertation were addressed. One of those constraints is related to the enormous challenge of integrating training phases of DL models in Edge devices. For this reason, traditional ML techniques have been proposed as a suitable and efficient alternative in this dissertation that provide good performance on inference, as shown by their applications within the TinyML concept in low-powered devices (Banbury et al., 2020). Throughout this work, several approaches have been proposed to optimize Edge resources. One example is the model cascade presented in Chapter 5 and adapted in Chapter 6, which focused on improving inference results for more complex classification systems (e.g. multi-class problems). Inspired by the work of Marco et al. (2020), a possible future research line could be to study a configuration where both classical ML or DL techniques are combined in the discriminative cascade to improve classification results. In this scheme, only those levels related to the ML algorithms studied during this work would be retrained, and DL techniques could be eventually used as a support for prediction purposes. To do so, it would be necessary to study the trade-off between classification results and system performance and see the potential advantages DL techniques present over the approach followed in this work.

- *Explore if this concept can be extended to other applications and domains*: In this dissertation, the workplace has been chosen as the central element that motivates this work. Under this context, this dissertation focuses con HAR applications and, in particular, how IoT technology can be more suitable for identifying habits and promoting better practices. Since the importance of the presented approach relies upon making the local Edge intelligent and privacy-preserving, it would be interesting to study how this concept could be adapted to tackle other kinds of applications for IoT Edge devices. By way of illustration, personalized healthcare or smart home domains are some other examples where the adaptation of environments and resources should be in line with individual human needs. Moreover, another relevant future research line would be to provides some guidelines suitable for any other sensor data available within the IoT paradigm.

7.6 Final remarks

This dissertation aimed to make significant contributions to the research field of intelligent environments, but, especially, to contribute to a more human-centric vision of the IoT, particularly for the workplace context. We firmly believe the synergy between IoT, energy-efficient AI techniques, and Hybrid Intelligence provides a suitable framework for creating more confident spaces where technological adoption does not imply compromising the privacy of the collected information, and the user is given complete control over his/her data and the personalization of the system.

Bibliography

Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., et al. (2016). Tensorflow: A system for large-scale machine learning. In *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, pages 265–283.

Abdelghani, W., Zayani, C. A., Amous, I., and Sèdes, F. (2019). User-centric iot: challenges and perspectives. In *UBICOMM 2018: The Twelfth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, pages 27–34.

Ackerman, M. S. (2000). The intellectual challenge of cscw: The gap between social requirements and technical feasibility. *Human–Computer Interaction*, 15(2-3):179–203.

Adegbija, T., Rogacs, A., Patel, C., and Gordon-Ross, A. (2017). Microprocessor optimizations for the internet of things: A survey. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 37(1):7–20.

Agaku, I. T., Adisa, A. O., Ayo-Yusuf, O. A., and Connolly, G. N. (2013). Concern about security and privacy, and perceived control over collection and use of health information are related to withholding of health information from healthcare providers. *Journal of the American Medical Informatics Association*, 21(2):374–378.

Ahmed, E., Yaqoob, I., Gani, A., Imran, M., and Guizani, M. (2016). Internet-of-things-based smart environments: state of the art, taxonomy, and open research challenges. *IEEE Wireless Communications*, 23(5):10–16.

Ajunwa, I., Crawford, K., and Schultz, J. (2017). Limitless worker surveillance. *Cal. L. Rev.*, 105:735.

Alam, F., Mehmood, R., Katib, I., and Albeshri, A. (2016). Analysis of eight data mining algorithms for smarter internet of things (iot). *Procedia Computer Science*, 98:437–442.

Altun, K., Barshan, B., and Tunçel, O. (2010). Comparative study on classifying human activities with miniature inertial and magnetic sensors. *Pattern Recognition*, 43(10):3605–3620.

Amershi, S., Cakmak, M., Knox, W. B., and Kulesza, T. (2014). Power to the people: The role of humans in interactive machine learning. *Ai Magazine*, 35(4):105–120.

Andrade, L., Prost-Boucle, A., and Pétrot, F. (2018). Overview of the state of the art in embedded machine learning. In *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 1033–1038. IEEE.

Anguita, D., Ghio, A., Oneto, L., Parra, X., and Reyes-Ortiz, J. L. (2013). A public domain dataset for human activity recognition using smartphones. In *Esann*, volume 3, page 3.

Arias-Castro, E., Donoho, D. L., et al. (2009). Does median filtering truly preserve edges better than linear filtering? *The Annals of Statistics*, 37(3):1172–1206.

Avellaneda Gonzalez, J. A. (2019). Data for: Office activity recognition using accelerometers and gyroscopes located on the forearms.

Baker, S. B., Xiang, W., and Atkinson, I. (2017). Internet of things for smart healthcare: Technologies, challenges, and opportunities. *IEEE Access*, 5:26521–26544.

Bakker, J., Holenderski, L., Kocielnik, R., Pechenizkiy, M., and Sidorova, N. (2012). Stess@ work: From measuring stress to its understanding, prediction and handling with personalized coaching. In *Proceedings of the 2nd ACM SIGHIT International health informatics symposium*, pages 673–678.

Banbury, C. R., Reddi, V. J., Lam, M., Fu, W., Fazel, A., Holleman, J., Huang, X., Hurtado, R., Kanter, D., Lokhmotov, A., et al. (2020). Benchmarking tinyml systems: Challenges and direction. *arXiv preprint arXiv:2003.04821*.

Berman, S. J. (2012). Digital transformation: opportunities to create new business models. *Strategy & Leadership*, 40(2):16–24.

Bernal, G., Colombo, S., Al Ai Baky, M., and Casalegno, F. (2017). Safety++: Designing iot and wearable systems for industrial safety through a user centered design approach. In *Proceedings of the 10th International Conference on PErvasive Technologies Related to Assistive Environments*, pages 163–170. ACM.

Bernardo, F., Zbyszynski, M., Fiebrink, R., and Grierson, M. (2017). Interactive machine learning for end-user innovation. In *2017 AAAI Spring Symposium Series*, pages 1–7.

Bhat, G., Deb, R., Chaurasia, V. V., Shill, H., and Ogras, U. Y. (2018). Online human activity recognition using low-power wearable devices. In *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, pages 1–8. IEEE.

Bhave, D. P., Teo, L. H., and Dalal, R. S. (2020). Privacy at work: A review and a research agenda for a contested terrain. *Journal of Management*, 46(1):127–164.

Bowen, J., Hinze, A., Griffiths, C., Kumar, V., and Bainbridge, D. (2017). Personal data collection in the workplace: Ethical and technical challenges. In *British Human Computer Interaction Conference,(BHCI)*.

Braunstein, A., Granka, L., and Staddon, J. (2011). Indirect content privacy surveys: measuring privacy without asking about it. In *Proceedings of the Seventh Symposium on Usable Privacy and Security*, pages 1–14.

Breiman, L. (1996). Bagging predictors. *Machine learning*, 24(2):123–140.

Bruno, B., Mastrogiovanni, F., and Sgorbissa, A. (2014). A public domain dataset for adl recognition using wrist-placed accelerometers. In *the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 738–743. IEEE.

Buchanan, E. A. and Hvizdak, E. E. (2009). Online survey tools: Ethical and methodological concerns of human research ethics committees. *Journal of Empirical Research on Human Research Ethics*, 4(2):37–48.

Buckley, J. P., Hedge, A., Yates, T., Copeland, R. J., Loosemore, M., Hamer, M., Bradley, G., and Dunstan, D. W. (2015). The sedentary office: a growing case for change towards better health and productivity. expert statement commissioned by public health england and the active working community interest company. *Br J Sports Med*, pages bjsports–2015.

Cahill, J., Portales, R., McLoughin, S., Nagan, N., Henrichs, B., and Wetherall, S. (2019). Iot/sensor-based infrastructures promoting a sense of home, independent living, comfort and wellness. *Sensors*, 19(3):485.

Cardoso, H. L. and Moreira, J. M. (2016). Human activity recognition by means of online semi-supervised learning. In *2016 17th IEEE International Conference on Mobile Data Management (MDM)*, volume 2, pages 75–77. IEEE.

Carpenter, D., McLeod, A., Hicks, C., and Maasberg, M. (2018). Privacy and biometrics: An empirical examination of employee concerns. *Information Systems Frontiers*, 20(1):91–110.

Casado-Mansilla, D., Garaizar, P., Irizar-Arrieta, A., and López-de Ipiña, D. (2019a). On the side effects of automation in iot: Complacency and comfort vs. relapse and distrust. *arXiv preprint arXiv:1911.08657*.

Casado-Mansilla, D., Garaizar, P., and López-de Ipiña, D. (2019b). User involvement matters: The side-effects of automated smart objects in pro-environmental behaviour. In *Proceedings of the 9th International Conference on the Internet of Things*, pages 1–4.

Casado-Mansilla, D., Garaizar, P., López-de Ipina, D., et al. (2018). Design-insights for devising persuasive iot devices for sustainability in the workplace. In *2018 Global Internet of Things Summit (GIoTS)*, pages 1–6. IEEE.

Chandakkar, P. S., Li, Y., Ding, P. L. K., and Li, B. (2017). Strategies for re-training a pruned neural network in an edge computing paradigm. In *2017 IEEE International Conference on Edge Computing (EDGE)*, pages 244–247. IEEE.

Chandrashekar, G. and Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28.

Chen, M., Xu, Z., Weinberger, K., Chapelle, O., and Kedem, D. (2012). Classifier cascade for minimizing feature evaluation cost. In *Artificial Intelligence and Statistics*, pages 218–226. PMLR.

Chen, S., Liu, T., Gao, F., Ji, J., Xu, Z., Qian, B., Wu, H., and Guan, X. (2017). Butler, not servant: A human-centric smart home energy management system. *IEEE Communications Magazine*, 55(2):27–33.

Cheng, J., Zhou, B., Sundholm, M., and Lukowicz, P. (2013). Smart chair: What can simple pressure sensors under the chairs legs tell us about user activity. In *UBICOMM13: The Seventh International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies*, pages 81–84.

Chung, C.-F., Gorm, N., Shklovski, I. A., and Munson, S. (2017). Finding the right fit: understanding health tracking in workplace wellness programs. In *Proceedings of the 2017 CHI conference on human factors in computing systems*, pages 4875–4886.

Cilla, R., Patricio, M. A., Berlanga, A., and Molina, J. M. (2009). Creating human activity recognition systems using pareto-based multiobjective optimization. In *2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance*, pages 37–42. IEEE.

Collins, A. and Halverson, R. (2018). *Rethinking education in the age of technology: The digital revolution and schooling in America*. Teachers College Press.

Conti, M., Passarella, A., and Das, S. K. (2017). The internet of people (iop): A new wave in pervasive mobile computing. *Pervasive and Mobile Computing*, 41:1–27.

da Silva, L. T., Souza, V. M., and Batista, G. E. (2019). Embml tool: Supporting the use of supervised learning algorithms in low-cost embedded systems. In *2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI)*, pages 1633–1637. IEEE.

Dahlbäck, N., Jönsson, A., and Ahrenberg, L. (1993). Wizard of oz studies—why and how. *Knowledge-based systems*, 6(4):258–266.

Dang, L. M., Piran, M., Han, D., Min, K., Moon, H., et al. (2019). A survey on internet of things and cloud computing for healthcare. *Electronics*, 8(7):768.

Dargie, W. (2009). Analysis of time and frequency domain features of accelerometer measurements. In *2009 Proceedings of 18th International Conference on Computer Communications and Networks*, pages 1–6. IEEE.

David, R., Duke, J., Jain, A., Reddi, V. J., Jeffries, N., Li, J., Kreeger, N., Nappier, I., Natraj, M., Regev, S., et al. (2020). Tensorflow lite micro: Embedded machine learning on tinyml systems. *arXiv preprint arXiv:2010.08678*.

De Croon, E., Sluiter, J., Kuijer, P. P., and Frings-Dresen, M. (2005). The effect of office concepts on worker health and performance: a systematic review of the literature. *Ergonomics*, 48(2):119–134.

Dellermann, D., Ebel, P., Söllner, M., and Leimeister, J. M. (2019). Hybrid intelligence. *Business & Information Systems Engineering*, 61(5):637–643.

Deng, S., Zhao, H., Fang, W., Yin, J., Dustdar, S., and Zomaya, A. Y. (2020). Edge intelligence: the confluence of edge computing and artificial intelligence. *IEEE Internet of Things Journal*, 7(8):7457–7469.

Desraches, F., Kamara-Esteban, O., Casado-Mansilla, D., and Borges, C. E. (2018). Forecasting the usage of appliances of shared use: an analysis of simplicity over complexity. In *2018 Energy and Sustainability for Small Developing Economies (ES2DE)*, pages 1–6. IEEE.

Dhar, S., Guo, J., Liu, J., Tripathi, S., Kurup, U., and Shah, M. (2019). On-device machine learning: An algorithms and learning theory perspective. *arXiv preprint arXiv:1911.00623*.

Domingos, P. and Pazzani, M. (1997). On the optimality of the simple bayesian classifier under zero-one loss. *Machine learning*, 29(2-3):103–130.

Dua, D. and Graff, C. (2017). UCI machine learning repository.

Eiband, M., Buschek, D., and Hussmann, H. (2021). How to support users in understanding intelligent systems? structuring the discussion. In *26th International Conference on Intelligent User Interfaces*, pages 120–132.

Elsts, A., McConville, R., Fafoutis, X., Twomey, N., Piechocki, R. J., Santos-Rodriguez, R., and Craddock, I. (2018). On-board feature extraction from acceleration data for activity recognition. In *EWSN*, pages 163–168.

Elsts, A., Twomey, N., McConville, R., and Craddock, I. (2020). Energy-efficient activity recognition framework using wearable accelerometers. *Journal of Network and Computer Applications*, 168:102770.

Enomoto, S. and Eda, T. (2021). Learning to cascade: Confidence calibration for improving the accuracy and computational cost of cascade inference systems. *arXiv preprint arXiv:2104.09286*.

Eshratifar, A. E., Abrishami, M. S., and Pedram, M. (2019). Jointdnn: an efficient training and inference engine for intelligent mobile cloud computing services. *IEEE Transactions on Mobile Computing*.

Evans, J. R. and Mathur, A. (2005). The value of online surveys. *Internet research*.

Fafoutis, X., Marchegiani, L., Elsts, A., Pope, J., Piechocki, R., and Craddock, I. (2018). Extending the battery lifetime of wearable sensors with embedded machine learning. In *2018 IEEE 4th World Forum on Internet of Things (WF-IoT)*, pages 269–274. IEEE.

Fails, J. A. and Olsen Jr, D. R. (2003). Interactive machine learning. In *Proceedings of the 8th international conference on Intelligent user interfaces*, pages 39–45.

Fausset, C. B., Mitzner, T. L., Price, C. E., Jones, B. D., Fain, B. W., and Rogers, W. A. (2013). Older adults' use of and attitudes toward activity monitoring technologies. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 57, pages 1683–1687. SAGE Publications Sage CA: Los Angeles, CA.

Fornell, C. and Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18(1):39–50.

Franklin, G. M. and Gray, Z. (2017). Work-related carpal tunnel syndrome: What's important? *Muscle & nerve*.

Fumera, G., Roli, F., and Giacinto, G. (2000). Reject option with multiple thresholds. *Pattern recognition*, 33(12):2099–2101.

Galloway, A. and Caudwell, C. (2018). Speculative design as research method. *Undesign: critical practice at the intersection of art and design. Routledge*, pages 85–96.

Garcia Lopez, P., Montresor, A., Epema, D., Datta, A., Higashino, T., Iamnitchi, A., Barcellos, M., Felber, P., and Riviere, E. (2015). Edge-centric computing: Vision and challenges. *ACM SIGCOMM Computer Communication Review*, 45(5):37–42.

Gepperth, A. and Hammer, B. (2016). Incremental learning algorithms and applications. In *European symposium on artificial neural networks (ESANN)*.

Ghamari, S., Ozcan, K., Dinh, T., Melnikov, A., Carvajal, J., Ernst, J., and Chai, S. (2021). Quantization-guided training for compact tinyml models. *arXiv preprint arXiv:2103.06231*.

Ghobakhloo, M. (2020). Industry 4.0, digitization, and opportunities for sustainability. *Journal of Cleaner Production*, 252:119869.

Gillies, M., Fiebrink, R., Tanaka, A., Garcia, J., Bevilacqua, F., Heloir, A., Nunnari, F., Mackay, W., Amershi, S., Lee, B., et al. (2016). Human-centred machine learning. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, pages 3558–3565.

Glance, D. G., Ooi, E., Berman, Y., Glance, C. F., and Barrett, H. R. (2016). Impact of a digital activity tracker-based workplace activity program on health and wellbeing. In *Proceedings of the 6th International Conference on Digital Health Conference*, pages 37–41.

Gómez-Carmona, O., Casado-Mansilla, D., Kraemer, F. A., López-de Ipiña, D., and García-Zubia, J. (2020). Exploring the computational cost of machine learning at the edge for human-centric internet of things. *Future Generation Computer Systems*, 112:670–683.

Gómez-Carmona, O., Casado-Mansilla, D., López-de Ipiña, D., and García-Zubia, J. (2019). Simplicity is best: Addressing the computational cost of machine learning classifiers in constrained edge devices. In *Proceedings of the 9th International Conference on the Internet of Things*, pages 1–8.

Gordon, D., Czerny, J., Miyaki, T., and Beigl, M. (2012). Energy-efficient activity recognition using prediction. In *2012 16th International Symposium on Wearable Computers*, pages 29–36. IEEE.

Gorm, N. and Shklovski, I. (2016). Steps, choices and moral accounting: Observations from a step-counting campaign in the workplace. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, pages 148–159.

Guo, F., Li, Y., Kankanhalli, M. S., and Brown, M. S. (2013). An evaluation of wearable activity monitoring devices. In *Proceedings of the 1st ACM international workshop on Personal data meets distributed multimedia*, pages 31–34.

Gupta, C., Suggala, A. S., Goyal, A., Simhadri, H. V., Paranjape, B., Kumar, A., Goyal, S., Udupa, R., Varma, M., and Jain, P. (2017). Protonn: Compressed and accurate knn for resource-scarce devices. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1331–1340. JMLR. org.

Gwet, K. L. (2014). *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.* Advanced Analytics, LLC.

Gómez-Carmona, O. and Casado-Mansilla, D. (2021). Office hydration monitoring (ohm) dataset.

Haigh, K. Z., Mackay, A. M., Cook, M. R., and Lin, L. G. (2015). Machine learning for embedded systems: A case study. *BBN Technologies: Cambridge, MA, USA*.

Hair, J., Anderson, R., and Babin, B. (2009). *Multivariate Data Analysis.* Prentice Hall, 7 edition.

Hair, J., Hollingsworth, C. L., Randolph, A. B., and Chong, A. Y. L. (2017). An updated and expanded assessment of pls-sem in information systems research. *Industrial Management & Data Systems*.

Hair Jr, J. F., Hult, G. T. M., Ringle, C., and Sarstedt, M. (2016). *A primer on partial least squares structural equation modeling (PLS-SEM).* Sage publications.

Han, L., Zhang, Q., Chen, X., Zhan, Q., Yang, T., and Zhao, Z. (2017). Detecting work-related stress with a wearable device. *Computers in Industry*, 90:42–49.

Hars, A. (2015). Self-driving cars: The digital transformation of mobility. In *Marktplätze im Umbruch: Digitale Strategien für Services im Mobilen Internet*, pages 539–549. Springer Berlin Heidelberg, Berlin, Heidelberg.

Helfrich, C. D., Kohn, M. J., Stapleton, A., Allen, C. L., Hammerback, K. E., Chan, K., Parrish, A. T., Ryan, D. E., Weiner, B. J., Harris, J. R., et al. (2018). Readiness to change over time: change commitment and change efficacy in a workplace health-promotion trial. *Frontiers in public health*, 6:110.

Holtermann, A., Schellewald, V., Mathiassen, S. E., Gupta, N., Pinder, A., Punakallio, A., Veiersted, K. B., Weber, B., Takala, E.-P., Draicchio, F., et al. (2017). A practical guidance for assessments of sedentary behavior at work: A perosh initiative. *Applied ergonomics*, 63:41–52.

Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics*, 3(2):119–131.

Hong, A., Nam, C., and Kim, S. (2020). What will be the possible barriers to consumers' adoption of smart home services? *Telecommunications Policy*, 44(2):101867.

Hou, L., Zhao, S., Xiong, X., Zheng, K., Chatzimisios, P., Hossain, M. S., and Xiang, W. (2016). Internet of things cloud: architecture and implementation. *IEEE Communications Magazine*, 54(12):32–39.

Howarth, A., Quesada, J., Silva, J., Judycki, S., and Mills, P. R. (2018). The impact of digital health interventions on health-related outcomes in the workplace: A systematic review. *DIGITAL HEALTH*, 4:2055207618770861.

Hu, Q., Tang, X., and Tang, W. (2020). A smart chair sitting posture recognition system using flex sensors and fpga implemented artificial neural network. *IEEE Sensors Journal*, 20(14):8007–8016.

Huang, Y., Benford, S., and Blake, H. (2019). Digital interventions to reduce sedentary behaviors of office workers: Scoping review. *Journal of medical Internet research*, 21(2):e11079.

Ilvesmaki, A. (2007). Drivers and challenges of personal health systems in workplace health promotion. In *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, pages 5878–5881. IEEE.

Inkpen, K., Chancellor, S., De Choudhury, M., Veale, M., and Baumer, E. P. (2019). Where is the human? bridging the gap between ai and hci. In *Extended abstracts of the 2019 chi conference on human factors in computing systems*, pages 1–9.

Irizar-Arrieta, A., Gómez-Carmona, O., Bilbao-Jayo, A., Casado-Mansilla, D., López-De-Ipiña, D., and Almeida, A. (2020). Addressing behavioural technologies through the human factor: A review. *IEEE Access*, 8:52306–52322.

Jacobs, J. V., Hettinger, L. J., Huang, Y.-H., Jeffries, S., Lesch, M. F., Simmons, L. A., Verma, S. K., and Willetts, J. L. (2019). Employee acceptance of wearable technology in the workplace. *Applied ergonomics*, 78:148–156.

Janidarmian, M., Roshan Fekr, A., Radecka, K., and Zilic, Z. (2017). A comprehensive analysis on wearable acceleration sensors in human activity recognition. *Sensors*, 17(3):529.

Jensen, U., Kugler, P., Ring, M., and Eskofier, B. M. (2016). Approaching the accuracy–cost conflict in embedded classification system design. *Pattern Analysis and Applications*, 19(3):839–855.

Jimenez, P. and Bregenzer, A. (2018). Integration of ehealth tools in the process of workplace health promotion: proposal for design and implementation. *Journal of medical Internet research*, 20(2):e65.

Kaasinen, E., Kymäläinen, T., Niemelä, M., Olsson, T., Kanerva, M., and Ikonen, V. (2013). A user-centric view of intelligent environments: User expectations, user experience and user role in building intelligent environments. *Computers*, 2(1):1–33.

Kang, S., Lee, J., Jang, H., Lee, Y., Park, S., and Song, J. (2009). A scalable and energy-efficient context monitoring framework for mobile personal sensor networks. *IEEE Transactions on Mobile Computing*, 9(5):686–702.

Kaur, H., Williams, A. C., McDuff, D., Czerwinski, M., Teevan, J., and Iqbal, S. (2020). Optimizing for happiness and productivity: Modeling opportune moments for transitions and breaks at work. In *ACM Conference on Human Factors in Computing Systems (CHI)*, page 1–15.

Kaynak, C. and Alpaydin, E. (2000). Multistage cascading of multiple classifiers: One man's noise is another man's data. In *ICML*, pages 455–462.

Khakurel, J., Melkas, H., and Porras, J. (2018). Tapping into the wearable device revolution in the work environment: a systematic review. *Information Technology & People*.

Khan, A., Hammerla, N., Mellor, S., and Plötz, T. (2016). Optimising sampling rates for accelerometer-based human activity recognition. *Pattern Recognition Letters*, 73:33–40.

Kim, Y., Park, Y., and Choi, J. (2017). A study on the adoption of iot smart home service: using value-based adoption model. *Total Quality Management & Business Excellence*, 28(9-10):1149–1165.

Kim, Y.-D., Park, E., Yoo, S., Choi, T., Yang, L., and Shin, D. (2015). Compression of deep convolutional neural networks for fast and low power mobile applications. *arXiv preprint arXiv:1511.06530*.

Kirk, C. P., Swain, S. D., and Gaskin, J. E. (2015). I'm proud of it: consumer technology appropriation and psychological ownership. *Journal of Marketing Theory and Practice*, 23(2):166–184.

Kraemer, F. A., Alawad, F., and Bosch, I. M. V. (2019). Energy-accuracy tradeoff for efficient noise monitoring and prediction in working environments. In *Proceedings of the 9th International Conference on the Internet of Things*, IoT 2019, New York, NY, USA. Association for Computing Machinery.

Kulesza, T., Burnett, M., Wong, W.-K., and Stumpf, S. (2015). Principles of explanatory debugging to personalize interactive machine learning. In *Proceedings of the 20th international conference on intelligent user interfaces*, pages 126–137.

Kumar, A., Goyal, S., and Varma, M. (2017). Resource-efficient machine learning in 2 kb ram for the internet of things. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pages 1935–1944. JMLR. org.

Lane, N. D., Bhattacharya, S., Georgiev, P., Forlivesi, C., and Kawsar, F. (2015). An early resource characterization of deep learning on wearables, smartphones and internet-of-things devices. In *Proceedings of the 2015 international workshop on internet of things towards applications*, pages 7–12. ACM.

Lane, N. D., Bhattacharya, S., Mathur, A., Georgiev, P., Forlivesi, C., and Kawsar, F. (2017). Squeezing deep learning into mobile and embedded devices. *IEEE Pervasive Computing*, 16(3):82–88.

Lefakis, L. and Fleuret, F. (2010). Joint cascade optimization using a product of boosted classifiers. In *Advances in neural information processing systems*, pages 1315–1323.

Lennefer, T., Lopper, E., Wiedemann, A. U., Hess, U., and Hoppe, A. (2020). Improving employees' work-related well-being and physical health through a technology-based physical activity intervention: A randomized intervention-control group study. *Journal of occupational health psychology*, 25(2):143.

Leroux, S., Bohez, S., De Coninck, E., Verbelen, T., Vankeirsbilck, B., Simoens, P., and Dhoedt, B. (2017). The cascading neural network: building the internet of smart things. *Knowledge and Information Systems*, 52(3):791–814.

Lin, C.-Y. and Marculescu, R. (2020). Model personalization for human activity recognition. In *2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 1–7. IEEE.

Liu, D., Kong, H., Luo, X., Liu, W., and Subramaniam, R. (2020). Bringing ai to edge: From deep learning's perspective. *arXiv preprint arXiv:2011.14808*.

Liu, H. and Setiono, R. (1995). Chi2: Feature selection and discretization of numeric attributes. In *Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence*, pages 388–391. IEEE.

Liu, J., Liu, J., Du, W., and Li, D. (2019). Performance analysis and characterization of training deep learning models on nvidia tx2. *arXiv preprint arXiv:1906.04278*.

Longhurst, R. (2003). Semi-structured interviews and focus groups. *Key methods in geography*, 3(2):143–156.

Lupton, D. (2014). Self-tracking cultures: towards a sociology of personal informatics. In *Proceedings of the 26th Australian computer-human interaction conference on designing futures: The future of design*, pages 77–86.

Maheronnaghsh, S., Santos, J., Marques, A. T., and Vaz, M. (2018). Methods for measuring association between intervention for increasing movement and productivity: systematic review. *U. Porto Journal of Engineering*, 4(1):27–41.

Majchrzak, A., Markus, M. L., and Wareham, J. (2016). Designing for digital transformation: Lessons for information systems research from the study of ict and societal challenges. *MIS quarterly*, 40(2):267–277.

Malhotra, N. (2008). Completion time and response order effects in web surveys. *Public opinion quarterly*, 72(5):914–934.

Malik, S. H., Blake, H., and Suggs, L. S. (2014). A systematic review of workplace health promotion interventions for increasing physical activity. *British journal of health psychology*, 19(1):149–180.

Maman, Z. S., Yazdi, M. A. A., Cavuoto, L. A., and Megahed, F. M. (2017). A data-driven approach to modeling physical fatigue in the workplace using wearable sensors. *Applied ergonomics*, 65:515–529.

Marco, V. S., Taylor, B., Wang, Z., and Elkhatib, Y. (2020). Optimizing deep learning inference on embedded systems through adaptive model selection. *ACM Transactions on Embedded Computing Systems (TECS)*, 19(1):1–28.

Masson, C. B., Martin, D., Colombino, T., and Grasso, A. (2016). "the device is not well designed for me" on the use of activity trackers in the workplace? In *COOP 2016: Proceedings of the 12th International Conference on the Design of Cooperative Systems, 23-27 May 2016, Trento, Italy*, pages 39–55. Springer.

Mathewson, K. W. (2019). A human-centered approach to interactive machine learning. *arXiv preprint arXiv:1905.06289*.

McCallum, L. and Fiebrink, R. (2019). Supporting feature engineering in end-user machine learning.

Mclntosh, A. and Gonzalez-Lima, F. (1994). Structural equation modeling and its application to network analysis in functional brain imaging. *Human brain mapping*, 2(1-2):2–22.

Meyer, J., Simske, S., Siek, K. A., Gurrin, C. G., and Hermens, H. (2014). Beyond quantified self: Data for wellbeing. In *CHI'14 Extended Abstracts on Human Factors in Computing Systems*, pages 95–98.

Miranda, J., Mäkitalo, N., Garcia-Alonso, J., Berrocal, J., Mikkonen, T., Canal, C., and Murillo, J. M. (2015). From the internet of things to the internet of people. *IEEE Internet Computing*, 19(2):40–47.

Miu, T., Missier, P., and Plötz, T. (2015). Bootstrapping personalised human activity recognition models using online active learning. In *2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing*, pages 1138–1147. IEEE.

Molchanov, P., Tyree, S., Karras, T., Aila, T., and Kautz, J. (2016). Pruning convolutional neural networks for resource efficient inference. *arXiv preprint arXiv:1611.06440*.

Moore, P. and Robinson, A. (2016). The quantified self: What counts in the neoliberal workplace. *new media & society*, 18(11):2774–2792.

Morawiec, D. (2016). sklearn-porter. Transpile trained scikit-learn estimators to C, Java, JavaScript and others.

Mugge, R., Schoormans, J. P., and Schifferstein, H. N. (2009). Emotional bonding with personalised products. *Journal of Engineering Design*, 20(5):467–476.

Muralidharan, D., Kao, J., Yang, X., Li, L., Viswanathan, L., Ibrahim, M. S., Luikens, K., Pulman, S., Garg, A., Kothari, A., et al. (2019). Leveraging user engagement signals for entity labeling in a virtual assistant. *arXiv preprint arXiv:1909.09143*.

Murphy, L. R. (1996). Stress management in work settings: a critical review of the health effects. *American Journal of Health Promotion*, 11(2):112–135.

Nappi, I. and de Campos Ribeiro, G. (2020). Internet of things technology applications in the workplace environment: a critical review. *Journal of Corporate Real Estate*.

Neff, G. and Nafus, D. (2016). *Self-tracking*. MIT Press.

Neto, A. R., Soares, B., Barbalho, F., Santos, L., Batista, T., Delicato, F. C., and Pires, P. F. (2018). Classifying smart iot devices for running machine learning algorithms. In *45º Seminário Integrado de Software e Hardware 2018 (SEMISH 2018)*, volume 45, pages 1–12. SBC.

Newton, J. T. (2012). Interdisciplinary health promotion: a call for theory-based interventions drawing on the skills of multiple disciplines. *Community dentistry and oral epidemiology*, 40(s2):49–54.

Nguyen, H. H., Mirza, F., Naeem, M. A., and Nguyen, M. (2017). A review on iot healthcare monitoring applications and a vision for transforming sensor data into real-time clinical feedback. In *2017 IEEE 21st International Conference on Computer Supported Cooperative Work in Design (CSCWD)*, pages 257–262. IEEE.

Nikou, S. (2019). Factors driving the adoption of smart home technology: An empirical assessment. *Telematics and Informatics*, 45:101283.

Nissenbaum, H. and Patterson, H. (2016). Biosensing in context: Health privacy in a connected world. *Quantified: Biosensing technologies in everyday life*, 79.

Nordby, J. (2019). emlearn: Machine learning inference engine for microcontrollers and embedded devices.

Oliveira, L. S., Britto, A., and Sabourin, R. (2005). Improving cascading classifiers with particle swarm optimization. In *Eighth International Conference on Document Analysis and Recognition (ICDAR'05)*, pages 570–574. IEEE.

Organization, W. H. et al. (2017). Promoting a healthy lifestyle at the workplace-be the change.

Orhan, M. A., Castellano, S., Khelladi, I., Marinelli, L., and Monge, F. (2021). Technology distraction at work. impacts on self-regulation and work engagement. *Journal of Business Research*, 126:341–349.

Owen, N., Healy, G. N., Matthews, C. E., and Dunstan, D. W. (2010). Too much sitting: the population-health science of sedentary behavior. *Exercise and sport sciences reviews*, 38(3):105.

Palan, S. and Schitter, C. (2018). Prolific. ac—a subject pool for online experiments. *Journal of Behavioral and Experimental Finance*, 17:22–27.

Papa, A., Mital, M., Pisano, P., and Del Giudice, M. (2020). E-health and wellbeing monitoring using smart healthcare devices: An empirical investigation. *Technological Forecasting and Social Change*, 153:119226.

Pape, S. and Rannenberg, K. (2019). Applying privacy patterns to the internet of things'(iot) architecture. *Mobile Networks and Applications*, 24(3):925–933.

Park, E., Kim, D., Kim, S., Kim, Y.-D., Kim, G., Yoon, S., and Yoo, S. (2015). Big/little deep neural network for ultra low power inference. In *2015 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ ISSS)*, pages 124–132. IEEE.

Paul, C., Scheibe, K., and Nilakanta, S. (2020). Privacy concerns regarding wearable iot devices: how it is influenced by gdpr? In *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Pearson, S. and Benameur, A. (2010). Privacy, security and trust issues arising from cloud computing. In *2010 IEEE Second International Conference on Cloud Computing Technology and Science*, pages 693–702. IEEE.

Pedersen, S. J., Cooley, P. D., and Mainsbridge, C. (2014). An e-health intervention designed to increase workday energy expenditure by reducing prolonged occupational sitting habits. *Work*, 49(2):289–295.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: Machine learning in python. *the Journal of machine Learning research*, 12:2825–2830.

Peer, E., Brandimarte, L., Samat, S., and Acquisti, A. (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, 70:153–163.

Person, A. L., Colby, S. E., Bulova, J. A., and Eubanks, J. W. (2010). Barriers to participation in a worksite wellness program. *Nutrition research and practice*, 4(2):149.

Podgorski, D., Majchrzycka, K., Dabrowska, A., Gralewicz, G., and Okrasa, M. (2017). Towards a conceptual framework of osh risk management in smart working environments based on smart ppe, ambient intelligence and the internet of things technologies. *International Journal of Occupational Safety and Ergonomics*, 23(1):1–20.

Porter, R., Theiler, J., and Hush, D. (2013). Interactive machine learning in data exploitation. *Computing in Science & Engineering*, 15(5):12–20.

Qi, X., Keally, M., Zhou, G., Li, Y., and Ren, Z. (2013). Adasense: Adapting sampling rates for activity recognition in body sensor networks. In *2013 IEEE 19th Real-Time and Embedded Technology and Applications Symposium (RTAS)*, pages 163–172. IEEE.

Qian, C., Yu, Y., and Zhou, Z.-H. (2015). Subset selection by pareto optimization. In *Advances in Neural Information Processing Systems*, pages 1774–1782.

Ramos, G., Suh, J., Ghorashi, S., Meek, C., Banks, R., Amershi, S., Fiebrink, R., Smith-Renner, A., and Bansal, G. (2019). Emerging perspectives in human-centered machine learning. In *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–8.

Randolph, S. A. (2017). Computer vision syndrome. *Workplace health & safety*, 65(7):328–328.

Ren, H., Anicic, D., and Runkler, T. (2021). Tinyol: Tinyml with online-learning on microcontrollers. *arXiv preprint arXiv:2103.08295*.

Ringle, C. M., Wende, S., Becker, J.-M., et al. (2015). Smartpls 3. *Boenningstedt: SmartPLS GmbH.*

Robert, S., Büttner, S., Röcker, C., and Holzinger, A. (2016). Reasoning under uncertainty: Towards collaborative interactive machine learning. In *Machine learning for health informatics*, pages 357–376. Springer.

Rokicki, M. and Drozda, M. (2017). Evaluating trade-offs in energy-efficient error detection. *International Journal of Communication Systems*, 30(7):e3028.

Rongen, A., Robroek, S. J., van Lenthe, F. J., and Burdorf, A. (2013). Workplace health promotion: a meta-analysis of effectiveness. *American journal of preventive medicine*, 44(4):406–415.

Roossien, C., Stegenga, J., Hodselmans, A., Spook, S., Koolhaas, W., Brouwer, S., Verkerke, G., and Reneman, M. F. (2017). Can a smart chair improve the sitting behavior of office workers? *Applied ergonomics*, 65:355–361.

Sagi, O. and Rokach, L. (2018). Ensemble learning: A survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 8(4):e1249.

Sakr, F., Bellotti, F., Berta, R., and De Gloria, A. (2020). Machine learning on mainstream microcontrollers. *Sensors*, 20(9):2638.

Salam, A. (2020). Internet of things for environmental sustainability and climate change. In *Internet of Things for Sustainable Community Development*, pages 33–69. Springer.

Samie, F., Bauer, L., and Henkel, J. (2019). From cloud down to things: An overview of machine learning in internet of things. *IEEE Internet of Things Journal*.

Schall Jr, M. C., Sesek, R. F., and Cavuoto, L. A. (2018). Barriers to the adoption of wearable sensors in the workplace: A survey of occupational safety and health professionals. *Human factors*, 60(3):351–362.

Schapire, R. E. (1990). The strength of weak learnability. *Machine learning*, 5(2):197–227.

Schaule, F., Johanssen, J. O., Bruegge, B., and Loftness, V. (2018). Employing consumer wearables to detect office workers' cognitive load for interruption management. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(1):32.

Scheidegger, F., Benini, L., Bekas, C., and Malossi, A. C. I. (2019). Constrained deep neural network architecture search for iot devices accounting for hardware calibration. In *Advances in Neural Information Processing Systems*, pages 6054–6064.

Schirner, G., Erdogmus, D., Chowdhury, K., and Padir, T. (2013). The future of human-in-the-loop cyber-physical systems. *Computer*, 46(1):36–45.

Settles, B. (2012). Active learning. *Synthesis lectures on artificial intelligence and machine learning*, 6(1):1–114.

Shafique, M., Theocharides, T., Bouganis, C.-S., Hanif, M. A., Khalid, F., Hafız, R., and Rehman, S. (2018). An overview of next-generation architectures for machine learning: Roadmap, opportunities and challenges in the iot era. In *2018 Design, Automation & Test in Europe Conference & Exhibition (DATE)*, pages 827–832. IEEE.

Shih, P. C., Han, K., Poole, E. S., Rosson, M. B., and Carroll, J. M. (2015). Use and adoption challenges of wearable activity trackers. *IConference 2015 proceedings*.

Shrestha, N., Pedisic, Z., Neil-Sztramko, S., Kukkonen-Harjula, K. T., and Hermans, V. (2016). The impact of obesity in the workplace: a review of contributing factors, consequences and potential solutions. *Current obesity reports*, 5(3):344–360.

Sicari, S., Rizzardi, A., Grieco, L. A., and Coen-Porisini, A. (2015). Security, privacy and trust in internet of things: The road ahead. *Computer networks*, 76:146–164.

Siirtola, P., Koskimäki, H., and Röning, J. (2019). Personalizing human activity recognition models using incremental learning. *arXiv preprint arXiv:1905.12628*.

Silva, E. J., Britto, A. S., Oliveira, L. S., Enembreck, F., Sabourin, R., and Koerich, A. L. (2017). A two-step cascade classification method. In *2017 International Joint Conference on Neural Networks (IJCNN)*, pages 573–580. IEEE.

Soro, S. (2021). Tinyml for ubiquitous edge ai. *arXiv preprint arXiv:2102.01255*.

Sparks, K., Faragher, B., and Cooper, C. L. (2001). Well-being and occupational health in the 21st century workplace. *Journal of occupational and organizational psychology*, 74(4):489–509.

Stephenson, A., McDonough, S. M., Murphy, M. H., Nugent, C. D., and Mair, J. L. (2017). Using computer, mobile and wearable technology enhanced interventions to reduce sedentary behaviour: a systematic review and meta-analysis. *International Journal of Behavioral Nutrition and Physical Activity*, 14(1):105.

Stoyanova, M., Nikoloudakis, Y., Panagiotakis, S., Pallis, E., and Markakis, E. K. (2020). A survey on the internet of things (iot) forensics: challenges, approaches, and open issues. *IEEE Communications Surveys & Tutorials*, 22(2):1191–1221.

Street, T. D. and Lacey, S. J. (2018). Employee perceptions of workplace health promotion programs: Comparison of a tailored, semi-tailored, and standardized approach. *International journal of environmental research and public health*, 15(5).

Sun, X., Wang, N., Chen, C.-Y., Ni, J., Agrawal, A., Cui, X., Venkataramani, S., El Maghraoui, K., Srinivasan, V. V., and Gopalakrishnan, K. (2020). Ultra-low precision 4-bit training of deep neural networks. *Advances in Neural Information Processing Systems*, 33.

Suto, J., Oniga, S., and Sitar, P. P. (2017). Feature analysis to human activity recognition. *International Journal of Computers Communications & Control*, 12(1):116–130.

Taylor, B., Marco, V. S., Wolff, W., Elkhatib, Y., and Wang, Z. (2018a). Adaptive deep learning model selection on embedded systems. *ACM SIGPLAN Notices*, 53(6):31–43.

Taylor, W., Suminski, R., Das, B., Paxton, R., and Craig, D. (2018b). Organizational culture and implications for workplace interventions to reduce sitting time among office-based workers: A systematic review. front. *Public Health*, 6:263.

Teebken, M. and Hess, T. (2021). Privacy in a digitized workplace: Towards an understanding of employee privacy concerns. In *Proceedings of the 54th Hawaii International Conference on System Sciences*, page 6661.

Tegen, A., Davidsson, P., and Persson, J. A. (2020a). Activity recognition through interactive machine learning in a dynamic sensor setting. *Personal and Ubiquitous Computing*, pages 1–14.

Tegen, A., Davidsson, P., and Persson, J. A. (2020b). The effects of reluctant and fallible users in interactive online machine learning.

Tian, Y., Li, M., and Xu, Q. (2019). Lightweight prediction based big/little design for efficient neural network inference. In *Proceedings of the 4th ACM/IEEE Symposium on Edge Computing*, pages 356–356.

Tremblay, M. S., Colley, R. C., Saunders, T. J., Healy, G. N., and Owen, N. (2010). Physiological and health implications of a sedentary lifestyle. *Applied Physiology, Nutrition, and Metabolism*, 35(6):725–740.

Twomey, N., Diethe, T., Fafoutis, X., Elsts, A., McConville, R., Flach, P., and Craddock, I. (2018). A comprehensive study of activity recognition using accelerometers. In *Informatics*, volume 5, page 27. Multidisciplinary Digital Publishing Institute.

van Griethuijsen, R. A., van Eijck, M. W., Haste, H., den Brok, P. J., Skinner, N. C., Mansour, N., Gencer, A. S., and BouJaoude, S. (2015). Global patterns in students' views of science and interest in science. *Research in science education*, 45(4):581–603.

Viola, P. and Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001*, volume 1, pages I–I. IEEE.

Voas, J., Kuhn, R., Laplante, P., and Applebaum, S. (2018). Internet of things (iot) trust concerns. *NIST, Tech. Rep.*

Waddell, G. and Burton, A. K. (2001). Occupational health guidelines for the management of low back pain at work: evidence review. *Occupational medicine*, 51(2):124–135.

Waddell, G. and Burton, A. K. (2006). *Is work good for your health and well-being?* The Stationery Office.

Wallace, B. C. and Dahabreh, I. J. (2012). Class probability estimates are unreliable for imbalanced data (and how to fix them). In *2012 IEEE 12th international conference on data mining*, pages 695–704. IEEE.

Walt, S. v. d., Colbert, S. C., and Varoquaux, G. (2011). The numpy array: a structure for efficient numerical computation. *Computing in Science & Engineering*, 13(2):22–30.

Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., and Chan, K. (2018). When edge meets learning: Adaptive control for resource-constrained distributed machine learning. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pages 63–71. IEEE.

Wang, X., Kondratyuk, D., Kitani, K. M., Movshovitz-Attias, Y., and Eban, E. (2020a). Multiple networks are more efficient than one: Fast and accurate models via ensembles and cascades. *arXiv preprint arXiv:2012.01988*.

Wang, X., Magno, M., Cavigelli, L., and Benini, L. (2019). Fann-on-mcu: An open-source toolkit for energy-efficient neural network inference at the edge of the internet of things. *arXiv preprint arXiv:1911.03314*.

Wang, X., McGill, T. J., and Klobas, J. E. (2020b). I want it anyway: consumer perceptions of smart home devices. *Journal of Computer Information Systems*, 60(5):437–447.

WHO, W. H. O. (2010). Healthy workplaces: a model for action: for employers, workers, policy-makers and practitioners.

Wolpert, D. H. (1992). Stacked generalization. *Neural networks*, 5(2):241–259.

Wong, K. K.-K. (2013). Partial least squares structural equation modeling (pls-sem) techniques using smartpls. *Marketing Bulletin*, 24(1):1–32.

Wu, J., Rehg, J. M., and Mullin, M. D. (2003). Learning a rare event detection cascade by direct feature selection. Technical report, Georgia Institute of Technology.

Xu, D., Li, T., Li, Y., Su, X., Tarkoma, S., and Hui, P. (2020). A survey on edge intelligence. *arXiv preprint arXiv:2003.12172*.

Xu, S., Tang, Q., Jin, L., and Pan, Z. (2019). A cascade ensemble learning model for human activity recognition with smartphones. *Sensors*, 19(10):2307.

Xu, Z., Kusner, M. J., Weinberger, K. Q., Chen, M., and Chapelle, O. (2014). Classifier cascades and trees for minimizing feature evaluation cost. *The Journal of Machine Learning Research*, 15(1):2113–2144.

Yan, Z., Subbaraju, V., Chakraborty, D., Misra, A., and Aberer, K. (2012). Energy-efficient continuous activity recognition on mobile phones: An activity-adaptive approach. In *2012 16th international symposium on wearable computers*, pages 17–24. Ieee.

Yan, Z., Zhang, P., and Vasilakos, A. V. (2014). A survey on trust management for internet of things. *Journal of network and computer applications*, 42:120–134.

Yang, R., Shin, E., Newman, M. W., and Ackerman, M. S. (2015). When fitness trackers don't'fit' end-user difficulties in the assessment of personal tracking device accuracy. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 623–634.

Yazici, M., Basurra, S., and Gaber, M. (2018). Edge machine learning: Enabling smart internet of things applications. *Big Data and Cognitive Computing*, 2(3):26.

Young, W., Davis, M., McNeill, I. M., Malhotra, B., Russell, S., Unsworth, K., and Clegg, C. W. (2015). Changing behaviour: successful environmental programmes in the workplace. *Business Strategy and the Environment*, 24(8):689–703.

Yu, J., Abraham, J. M., Dowd, B., Higuera, L. F., and Nyman, J. A. (2017). Impact of a workplace physical activity tracking program on biometric health outcomes. *Preventive medicine*, 105:135–141.

Yu, K., Berkovsky, S., Conway, D., Taib, R., Zhou, J., and Chen, F. (2018). Do i trust a machine? differences in user trust based on system performance. In *Human and Machine Learning*, pages 245–264. Springer.

Zalewski, P., Marchegiani, L., Elsts, A., Piechocki, R., Craddock, I., and Fafoutis, X. (2020). From bits of data to bits of knowledge—an on-board classification framework for wearable sensing systems. *Sensors*, 20(6):1655.

Zenonos, A., Khan, A., Kalogridis, G., Vatsikas, S., Lewis, T., and Sooriyabandara, M. (2016). Healthyoffice: Mood recognition at work using smartphones and wearable sensors. In *2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*, pages 1–6. IEEE.

Zhang, C. and Ma, Y. (2012). *Ensemble machine learning: methods and applications*. Springer.

Zhang, P., Bui, T. D., and Suen, C. Y. (2007). A novel cascade ensemble classifier system with a high recognition performance on handwritten digits. *Pattern Recognition*, 40(12):3415–3429.

Zhao, W., Lun, R., Gordon, C., Fofana, A.-B. M., Espy, D. D., Reinthal, M. A., Ekelman, B., Goodman, G. D., Niederriter, J. E., and Luo, X. (2017). A human-centered activity tracking system: Toward a healthier workplace. *IEEE Transactions on Human-Machine Systems*, 47(3):343–355.

Zheng, M., Xu, D., Jiang, L., Gu, C., Tan, R., and Cheng, P. (2019). Challenges of privacy-preserving machine learning in iot. In *Proceedings of the First International Workshop on Challenges in Artificial Intelligence and Machine Learning for Internet of Things*, pages 1–7.

Zheng, N.-n., Liu, Z.-y., Ren, P.-j., Ma, Y.-q., Chen, S.-t., Yu, S.-y., Xue, J.-r., Chen, B.-d., and Wang, F.-y. (2017). Hybrid-augmented intelligence: collaboration and cognition. *Frontiers of Information Technology & Electronic Engineering*, 18(2):153–179.

Zheng, S., Apthorpe, N., Chetty, M., and Feamster, N. (2018). User perceptions of smart home iot privacy. *Proceedings of the ACM on Human-Computer Interaction*, 2(CSCW):1–20.

Zhou, Z., Chen, X., Li, E., Zeng, L., Luo, K., and Zhang, J. (2019). Edge intelligence: Paving the last mile of artificial intelligence with edge computing. *Proceedings of the IEEE*, 107(8):1738–1762.

Zhu, K., Dong, S., Xu, S. X., and Kraemer, K. L. (2006). Innovation diffusion in global contexts: determinants of post-adoption digital transformation of european companies. *European journal of information systems*, 15(6):601–616.

Zweig, D. and Webster, J. (2002). Where is the line between benign and invasive? an examination of psychological barriers to the acceptance of awareness monitoring systems. *Journal of Organizational Behavior: The International Journal of Industrial, Occupational and Organizational Psychology and Behavior*, 23(5):605–633.

APPENDIX

A

Questionnaire on technology perception in the workplace

A.1 Initial information regarding the online survey

Survey Description

The goal of this survey is to collect your ideas on your perception of smart workplaces and the use of smart devices in those spaces. To do so, we describe a speculative scenario related to a technology-led wellness promotion campaign. It is designed to promote healthier behaviours by monitoring those habits of your working routine that could have a negative impact on your well-being.

This survey is comprised of the following aspects:

1. General questions about your perception of the use of technology at the workplace. (1 min)

2. Demographic information. (1 min)

3. Your opinion regarding the speculative scenario described. (3 min)

4. Wrap up feedback. (1 min)

Finally, your participation in this survey is voluntary. You have the right to withdraw at any point during the survey, for any reason, and without any justification. If you would like to contact the members of the research team for any reason, please do it by sending an email to Oihane Gómez-Carmona (oihane.gomezc@deusto.es) or Diego Casado-Mansilla (dcasado@deusto.es), both researchers at the University of Deusto. By clicking the "I consent" button below, you acknowledge that your participation in the survey is voluntary, you are at least 18 years old, and that you are aware that you may choose to terminate your participation in the study at any time and for any reason.

- I consent

- I do NOT consent

A.2 Adoption of smart devices and monitoring equipment in the workplace

To begin with, we would like to know your opinion about integrating smart technologies in the workplace (e.g. activity trackers). These are open questions. Try to answer them by thinking about your general feelings regarding the potential use of new technologies in those spaces.

Please, rate the following statements related to the use of smart devices in the workplace from 1 (completely disagree) to 5 (completely agree):

1. If I use a smart work device, I would lose control over the privacy of my personal data.

2. My personal information will be less confidential if I use a smart work device.

3. The security systems built into smart work devices are not strong enough to protect my information.
4. Internet hackers might take control of my information if I use a smart work device.
5. I feel that the smart work equipment fits my lifestyle.
6. I feel that the smart work equipment is compatible with my day-to-day needs.
7. I think that the smart work equipment will fit well into my work environment.
8. I think that the smart work products and applications are useful for the tasks I do at work.

A.3 Socio-Demographic information

Age group:

- <21
- 22-40
- 41-52
- 53-71
- 72+

Nationality

- Dropdown menu containing all county options.

Education

- None

- High-school /secondary
- Post-secondary (non-university)
- University degree (bachelor or equivalent)
- Post-graduate (master or equivalent)
- Doctoral degree (PhD or equivalent)

How much of your time, on average, do you spend at your work desk in a day?

- None
- 1-2 hours
- 2-5 hours
- 5-8 hours
- More than 8 hours

Which of the following smart devices do you own or use regularly?

- Voice Assistants (Alexa/Google home or similar devices)
- Wearables or apps for health tracking (fitness trackers, nutrition diaries, etc.)
- Smart home devices (smart meters, smart security cameras or any other connected equipment)
- Any other system that guides you in your everyday decision making (e.g., driver-assisted car)
- None

How do you rate your technological knowledge and/or background from 0 (None) to 5 (tech-savvy)?

- 1 (None)
- 2
- 3
- 4
- 5 (Expert)

How do you rate your knowledge from 0 (None) to 5 (Expert) on the EU General Data Protection Regulation (GDPR) and/or other data privacy policies of your country?

- 1 (None)
- 2
- 3
- 4
- 5 (Expert)

A.4 Speculative Scenario

Scenario description:

Imagine that you work in a company that decides to launch a voluntary campaign to promote physical activity at work and healthier habits during working hours. To do so, the managers of the campaign (belonging to your organization) provide you with an intelligent system. This system is comprised of several smart devices (e.g. a smartwatch or water intake tracker) that collect data about your daily routine and monitor those habits and behaviours that could impair your health and wellbeing.

By means of those devices, the proposed intervention seeks to shed light on how much time you spend without getting up from your chair, how much you walk or move during working hours, how good or bad your hydration habits

Figure A.1: An illustrative representation of speculative scenario to illustrate the described concept.

are or how many coffees you drink per day. They also change the traditional office chairs for smart chairs to improve your ergonomics and, hence, alert you when an adjustment of your body posture is advisable.

Based on this speculative scenario described above, please answer the following questions:

As an employee, what would you do if you were invited to participate in such an intervention?

- Join immediately
- Discuss first with colleagues and then decide
- Join after a while
- Do not participate

Taking into account the smart environment described above: What do you value most when it comes to using such a system? (i.e. the provided smart devices). Please, select only the top 3 statements you think fit the most with your preferences:

- I value/appreciate...that the system is efficient enough without requiring much attention or time from me.
- I value/appreciate...that the system is secure and privacy-aware (i.e your personal information will not be compromised).
- I value/appreciate...that I can understand how the system works, what kind of information it collects, which is the use of the collected data and who has access to them.
- I value/appreciate...that I can decide which of the data collected by the system is public or private, or to whom it is shared.
- I value/appreciate...that the system is actually beneficial to me or my health.
- I value/appreciate...that the collected data does not cause any harm to me, neither personally nor professionally.
- I value/appreciate...that my personal data is only used for the purpose for which it is collected.

In your opinion... Who should be responsible for the storage and security of the data that you produce in the workplace, and that is collected by the intelligent system? Please provide your opinion from 1 (completely disagree) to 5 (completely agree):

1. All data should remain under my ownership, and I will be the only person that can grant access to third parties of certain data I produce.
2. All data that I produce should remain under my ownership but visible to my company/organization.
3. All data that I produce should remain under my ownership but visible to the healthcare company of the organization where I work.

4. All data that I produce should remain under my ownership but visible to interested entities, as long as this data can not be linked to me and only for the sake of my health.

5. All data that I produce should remain under the ownership of my company/organization since they are the health campaign promoters.

6. All data that I produce should remain under the ownership of my company/organization, as long as these data can not be linked to me.

Given that: 1) the fictional intelligent system can provide you with health advice based on your activity at work, and 2) the fictional intelligent system can be supervised by you in order to improve the feedback that it gives…Then, please, evaluate from 1 (completely disagree) to 5 (completely agree) the following statements:

1. The intelligent system must be located close to me, in a local device I own, so that the collected data are not stored on external devices (e.g. cloud servers my company own) that I can not control.

2. The intelligent system must be remotely located so that the collected data are stored on external devices (e.g. cloud servers my company own) that I can not control.

3. The intelligent system must be always on, collecting my data at all times.

4. The intelligent system must be able to interact with me in a tailored manner, e.g. only when I have availability.

5. The intelligent system must have the ability to be customized according to my specific needs.

6. The intelligent system can be located locally or remotely, as long as the protection of the collected data is regulated or I have control over them.

A.5 Wrap-up

To conclude, please rate the following statements related to the use of smart devices in the workplace from 1 (completely disagree) to 5 (completely agree)::

1. If I use a smart work device, I would lose control over the privacy of my personal data.
2. My personal information will be less confidential if I use a smart work device.
3. The security systems built into smart work devices are not strong enough to protect my information.
4. Internet hackers might take control of my information if I use a smart work device.
5. I feel that the smart work equipment fits my lifestyle.
6. I feel that the smart work equipment is compatible with my day-to-day needs.
7. I think that the smart work equipment will fit well into my work environment.
8. I think that the smart work products and applications are useful for the tasks I do at work.

Anything you want to share with us? Final insights or feedback (optional)

- Long text response space

Declaration

I, Oihane Gómez Carmona, herewith declare that this dissertation is my own original work, carried out as a doctoral student at the University of Deusto. All assistance received and notions from other sources have been identified as such, acknowledging their correspondent contributions and citing them properly.

This work contains no material which has been presented in identical or similar form to any examination board, except where due acknowledgement is made in the dissertation.

This dissertation was finished writing on June 14^{th}, 2021.