Universidad de Deusto
University of Deusto
Deusto

QUANTUM MACHINE LEARNING APPROACHES APPLIED TO HEALTHCARE AND WELL-BEING

Doctoral thesis by
Ubaid Ullah

Supervisors
PhD. Maria Begoña Garcia-Zapirain Soto
PhD. Aurora Madariaga Ortuzar

Faculty of Engineering

University of Deusto

February 2025

Universidad de Deusto
University of Deusto
Deusto

QUANTUM MACHINE LEARNING APPROACHES APPLIED TO HEALTHCARE AND WELL-BEING

Doctoral thesis by
Ubaid Ullah

Supervisors:
PhD. Maria Begoña Garcia-Zapirain Soto
PhD. Aurora Madariaga Ortuza

Doctoral Student

ULLAH UBAID - Y8813638Y

Digitally signed by ULLAH UBAID - Y8813638Y
DN: c=ES, serialNumber=IDCES-Y8813638Y, givenName=UBAID, sn=ULLAH, cn=ULLAH UBAID - Y8813638Y
Date: 2025.02.27 15:01:12 +01'00'

Supervisors

AURORA MADARIAGA ORTUZAR

Firmado digitalmente por AURORA MADARIAGA ORTUZAR
Fecha: 2025.02.27 18:10:34 +01'00'

Bilbao, February 2025

Dedication

This doctoral thesis is dedicated with profound gratitude to my beloved parents, Mr. Purdil Khan and Mrs. Razia Begum, whose boundless love and strength have been my guiding light. I also extend my heartfelt appreciation to my brother, Kalim Ullah, and all my sisters, whose steadfast support has been invaluable throughout this journey. Finally, I dedicate this work to my devoted wife, Wafa Ibadat, and to my cherished daughter, Amna Ubaid, my little princess, who continually inspire me with their love and joy.

Declaration

I hereby declare that, with the exception of explicit references to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification at this or any other university. Except as mentioned in the text and Acknowledgements, this dissertation is entirely my own work.

Ubaid Ullah
February 2025

Acknowledgments

The completion of this thesis would not have been possible without the invaluable support, encouragement, and guidance of many individuals and institutions. First and foremost, I am profoundly grateful to my supervisors, Prof. Begoña Garcia-Zapirain and Dr. Aurora Madriaga Ortuza, for their steadfast support, insightful feedback, and unwavering motivation throughout my PhD journey. Their mentorship has been fundamental in cultivating my growth as an independent thinker and researcher, always pushing the boundaries of my curiosity and dedication to learning.

I would also like to extend my deepest gratitude to DEIKER for providing me with the opportunity to pursue my doctoral research at the University of Deusto through the Marie Curie Cofund grant. This opportunity has been instrumental in shaping my academic career.

I also wish to express my sincere thanks to the eVIDA research group and DeustoTech for fostering a stimulating research environment and providing the necessary resources to bring this work to fruition. I am equally indebted to my colleagues and friends, whose camaraderie and encouragement made this challenging journey a more enjoyable and enriching experience.

Finally, I am forever grateful to my parents, family members, and friends for their unconditional love, support, and belief in my abilities. Their constant encouragement has been my guiding force, providing me the strength to overcome every challenge I faced. I also dedicate all my accomplishments to them, and I will always strive to make them proud.

Ubaid Ullah
Bilbao, February 2025

Abstract

Quantum computing (QC) differs from traditional computing by using quantum bits (qubits) and leveraging quantum phenomena like superposition, entanglement, and interference. These properties enable quantum computers to perform parallel computations, vastly surpassing classical systems. Quantum machine learning (QML) combines QC with machine learning to enhance data processing and classification. This thesis explores the theoretical advantages and practical applications of QML in healthcare. It begins with a summary of recent advancements, emphasizing QML's significance in the healthcare sector, followed by three key case studies demonstrating its practical use. The study adopts a systematic literature review approach, including research question formulation and article quality assessment using specific metrics. Of 2,038 records collected, 468 duplicates and 1,053 non-related to healthcare were excluded, leaving 55 articles from 2019 to 2024 for evaluation. The analysis provides a foundation for future research, promoting innovative solutions to healthcare challenges at the intersection of quantum computing and machine learning.

The first case study examines two QML algorithms Enhanced Quantum Support Vector Machine (E-QSVM) and Quantum Random Forest (QRF) applied to COVID-19 and influenza datasets. After data cleaning, normalization, and scaling, dimensionality is reduced by selecting the top 10 features using Chi-Square, ANOVA, and classical models for quantum efficiency. E-QSVM employs a parameterized quantum circuit with a ZZ feature map to capture complex data relationships, while QRF, implemented via Pennylane, uses two parallel QPUs, selecting the most accurate output. Results show E-QSVM and QRF outperform classical counterparts by 1% and 2%, respectively, with E-QSVM and QRF achieving up to 6% improvement over recent quantum models for the COVID-19 dataset.

The second case study introduces a quantum version of the Fully Convolutional Neural Network (FCQ-CNN) for Ischemic Heart Disease (IHD) classification. The model integrates quantum and classical convolutional layers, a pooling layer, and a fully connected layer. Parameters are optimized using the Adam optimizer with a learning rate of 0.001 and a batch size of 32, while binary cross-entropy is used as the loss function over 100 epochs. Dimensionality reduction is achieved using SVM with Recursive Feature Elimination (SVM-RFE). Testing on 20% of the dataset yields an accuracy of 84.6% with a loss of 0.28. When compared to classical models (Optimized-

CNN and FCNN) and previously published quantum models, the FCQ-CNN shows accuracy improvements of 8.6%, 12.6%, 3.5%, and 1.8%, respectively.

The third case study explores two quantum models: Quantum K Nearest Neighbor (QKNN) and Amplitude Encoding Variational Quantum Classifier (AE-VQC). Both models utilize amplitude encoding to transform input data into quantum state amplitudes and were evaluated on datasets with 8, 16, and 32 features, selected through a feature selection process. QKNN uses the quantum K minimum-finding algorithm to identify the nearest neighbors, measuring distances between quantum states using fidelity and applying a swap test for statistical assessment. AE-VQC comprises a feature map, variational quantum circuit, measurement circuit, and COBYLA optimizer. The experimental results demonstrate that both models outperform their classical counterparts, with QKNN achieving 3%, 3%, and 0.5% higher accuracy across set of features. AE-VQC also shows improved accuracy of 0.3% and 0.5% in classifying sarcopenia disease, while reducing computational complexity.

In summary, this thesis evaluates various QML models, including data pre-processing and state preparation, across three diffrent case studies involving COVID-19, cardiovascular disease, and sarcopenia. The models show improved accuracy over classical machine learning techniques, highlighting QML's potential in addressing complex healthcare challenges efficiently. Despite these challenges like noise and limited resources, the findings contribute to the growing intersection of QML and healthcare, offering promising avenues for future research.

Contents

List of Figures

List of Tables

Abbreviations

QC Quantum Computing

QML Quantum Machine Learning

AI Artificial Intelligent

E-QSVM Enhanced Quantum Support Vector Machine

QRF Quantum Random Forest

FCQ-CNN Fully Connected Quantum Convalutional Neural Network

IHD Ischemic Heart Disease

MRMR Maximally Relevant Minimally Redundancy

QKNN Quantum K Nearest Neighbor

AE-VQC Amplitude Encoding Variational Quantum Classifier

SOs Specific Objectives

IBM International Business Machines

WHO World Health Organization

EHR Electronic Healthcare Record

UCI University of California Irvine

MNIST Modified National Institute of Standards and Technology database

CC Classical Computing

DL Deep Learning

QNN Quantum Neural Network

QLSTM Quantum Long Short Term Memory

QPCA Quantum Principal Component Analysis

DNA Deoxyribonuclease Acid

HQMLP Hybrid Quantum Multi Layer Perception

QDT Quantum Decision Tree

QKD Quantum K Distribution

QAOA) Quantum Approximate Optimization Algorithm

QRAC Quantum Random Access Coding

NISQ Noisy Intermediate Scale Quantum

MRI Magnetic resonance imaging

EEG Electroencephalogram

EMG Electromyography

QA Quantum Annealing

ARDS Acute Respiratory Distress Syndrome

ESC European Cardiology Society

CAD Coronary Artery Disease

RFE Recursive Feature Elimination

DEXA Dual Energy X-ray Absorptiometry

COBYLA Constrained Optimization By Linear Approximation optimizer.

MHR Medical Healthcare Record

QPUs Quantum Processing Units

Chapter 1

Introduction

1.1 Introduction

In physics, the smallest possible discrete unit of a physical quantity is known as quantum. The term "quantum" in physics refers to the smallest indivisible unit of a physical quantity. Quantum particles exhibit dual properties, acting as both waves and particles and are described by quantum theory, which focuses on determining the position of a quantum particle at a specific location in space [1]. Quantum information science leverages quantum phenomena, such as entanglement, to manage and process information. Utilizing these principles, quantum computers can address computational challenges at unprecedented speeds [2]. Quantum computing (QC) has gained significant attention in the fields of Artificial Intelligence (AI) and information science, inspiring computer scientists, engineers, and physicists because of its potential applications [3]. Quantum ML and quantum DL is a subset of quantum AI, which, utilizing the principles of quantum mechanics [4], can rapidly solve complex problems and process and transmit information simultaneously, as shown in figure 1.1.

Recently, quantum computing and machine learning (ML) have attracted considerable attention due to their potential to greatly enhance the efficiency of various tasks compared to classical methods. For instance, quantum computers can solve specific computational, optimization, and search problems much faster than classical computers. They can also simulate the behavior of certain physical systems with greater accuracy, which is particularly useful in fields like healthcare, industrial, communication and materials science [5]. The integration of ML theory with the principles of quantum computing has led to the emergence of a new research sub-discipline known as Quantum Machine Learning (QML) [6]. QML focuses on developing quantum applications for various ML algorithms, leveraging the computational power of quantum computers alongside the scalability and learning capabilities of ML algorithms. The general overview of the QML working process is shown in figure 1.2, which contains on input dataset, state preparation, quantum model, measurement circuit and pre-

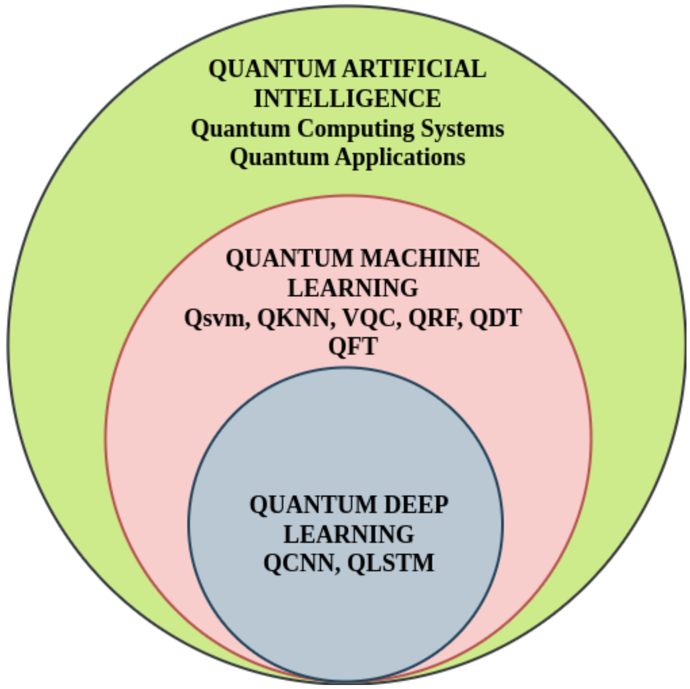

Figure 1.1: Quantum artificial intelligence is a superset of quantum machine learning and quantum deep learning.

dicted output. This dissertation focuses on the use of various QML algorithms for the processing of healthcare data. The main work contains three different case studies, where case study 1 focuses on Covid-19 disease classification. The work presents the two QML algorithms, i.e., Enhanced Quantum Support Vector Machine (E-QSVM) and Quantum Random Forest (QRF), applied to Covid-19 and influenza datasets, which were collected from different private hospitals. The experimental results show that the proposed models outperform in terms of accuracy. The competency of the models is obtained by comparing them with classical models and recently published quantum models. Case Study 2 focuses on the evaluation and classification of Cardiovascular disease, where I propose a quantum version of the Fully Convolutional Neural Network (FCQ-CNN), which evaluates the quantum circuit-based technique that was inspired by convolutional neural networks, a very successful machine learning model. Case study 3 deals with the evaluation of sarcopenia disease, which is the condition in which older individuals experience a gradual loss of muscle mass and function. Here, I used two quantum models: Quantum K Nearest Neighbor (QKNN) and Amplitude Encoding Variational Quantum Classifier (AE-VQC). The models were tested using three sets of experiments with 8, 16, and 32 features, and a feature selection mechanism was used to determine the most important set of features. The major working principle for all QML models is shown in figure 1.2.

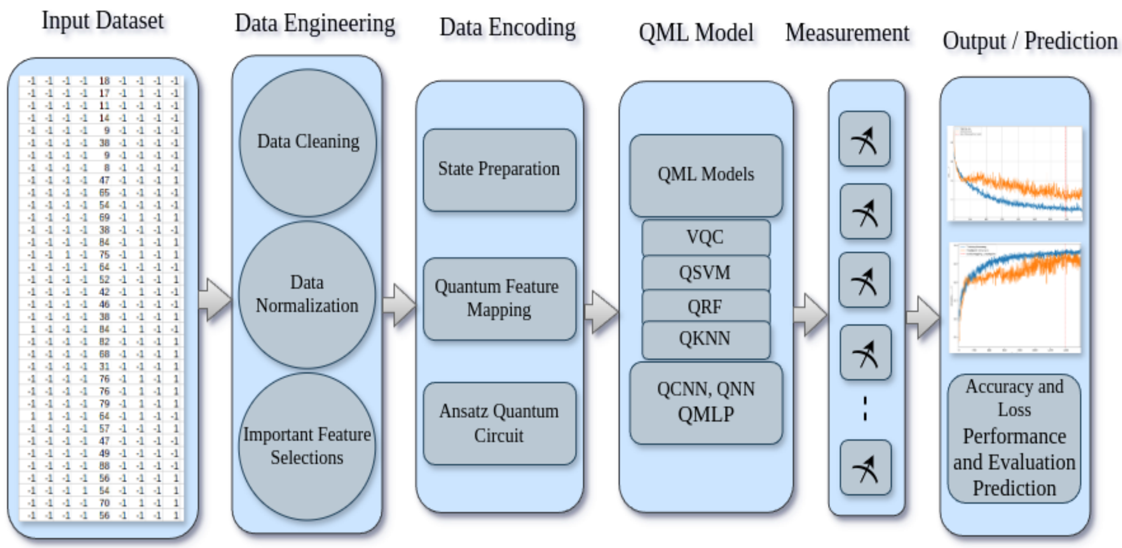

Figure 1.2: Working principle of QML model.

1.2 Research Hypothesis and Objectives

This section presents the hypothesis and objectives of the thesis. The hypothesis are develop based on a thorough analysis of the challenges discussed in the previous section and is outlined as follows.

> Quantum machine learning is able to outperform classical machine learning in disease early prediction, including Covid-19, cardiovascular disease, and sarcopenia, by utilizing quantum properties, a new learning paradigms to enhance accuracy, and leading to significant advancements in personalized healthcare and treatment optimization.

In light of the hypothesis mentioned above, the primary aim of this dissertation is to demonstrate the advantages of quantum computing integrated with machine learning (QML) for addressing healthcare challenges by comparing it to traditional machine learning models. To achieve this overarching goal, the following specific objectives (SOs) have been established.

> **SO 1:** To establish the current state-of-the-art in quantum computing and quantum machine learning, explore their healthcare applications, and examine the diverse types of medical datasets utilized.

For the above objective, a concise summary of the latest advancements in quantum computing and quantum ML, drawn from recently published articles in the healthcare domain, has been compiled. This review adheres to the essential guidelines for conducting a systematic literature review, which involves formulating research questions and assessing the quality of articles using defined metrics. Initially, 2,038 records were collected from multiple databases, with 468 duplicates and 1,053 records unrelated to healthcare being excluded. An additional 258, 68, and 39 records were discarded based on title, abstract, and full-text criteria, respectively. Ultimately, 49 articles remained for evaluation, offering a comprehensive overview of recent literature and enhancing the current understanding of quantum machine learning algorithms and their applications in healthcare.

Case Study 1: Covid-19.

SO 2: To determine the pre-processing and data analysis techniques applied to the Covid-19 dataset.

SO 3: To check the performance of QML models with it's counterpart classical models for Covid-19 cases classification.

In this case study, addressing the first objective, a brief analysis of the Covid-19 pandemic has been conducted, emphasizing the societal risks posed by the virus and exploring pandemic-related statistics. The target dataset was then examined, with inconsistencies removed, followed by data cleaning, balancing, and normalization.

For the second objective, two QML models, QRF and QSVM models, were employed. The QRF model was executed using both the PennyLane and IBM simulators. At the same time, the QSVM was implemented with the Qiskit library and run on the IBM simulator. Both quantum models were compared to their classical counterparts, and the results were thoroughly evaluated.

Case Study 2: Cardiovascular.

SO 4: To analyze and understand cardiovascular disease, data types, and the significance of feature dimensionality reduction.

SO 5: To design and validate a QCNN model that outperforms its counterpart CNN in the classification of cardiovascular diseases.

For the first objective, cardiovascular disease was briefly studied, and statistics were gathered from sources such as the WHO and the European Society of Cardiology to highlight the significance of addressing this disease. The data was collected from a hospital's EHR of patients. Preprocessing involved a series of steps to clean, transform, normalize, and standardize the dataset. Subsequently, feature selection techniques were applied to improve the predictive performance of the model.

The second objective is addressed by designing a quantum version of a convolutional neural network. Initially, the data is encoded using a feature map with various gate configurations. A quantum circuit is then employed to create the convolutional layer, which is connected to a pooling layer. Finally, a measurement circuit is used at the output to decode the predicted observable values.

Case Study 3: Sarcopenia.

SO 6: To analyze the data pre-processing, and state preparation for QML models.

SO 7: To develop and validate amplitude-encoded QKNN and VQC models that outperform their classical NN and KNN counterparts for the prediction of sarcopenia.

The process involves data analysis, data pre-processing, and state preparation for the study of sarcopenia disease. Data analysis is focused on extracting meaningful insights and identifying patterns within the dataset. Pre-processing ensures that the data is cleaned, transformed, normalized, and standardized to improve the quality and accuracy of the analysis. In the context of QML validation, state preparation refers to encoding the pre-processed data into quantum states for further processing and modeling.

For this objective, the QML models, VQC and QKNN, were implemented. Initially, the model's performance was evaluated using basis encoding, followed by amplitude encoding. Various sets of key features were utilized for both models, and their results were compared to those of a classical model to demonstrate improvements in prediction accuracy.

1.3 Scientific, Social Impact, and Contribution

Following the introduction of the dissertation framework, this section focuses on underscoring the social and scientific implications of the research conducted. QML techniques are believed to improve the diagnosis of various health conditions by leveraging different data types and modalities. The main issue with classical approaches is the dimensionality reduction techniques like PCA and face significant challenges with large datasets due to high computational complexity ($(O\left(n^3\right)$ and $O\left(n^2\right)$, receptively), and limited ability to capture complex, non-linear correlations. QML addresses these issues by utilizing amplitude encoding, allowing massive datasets to be represented in quantum states and processed simultaneously through quantum parallelism. Quantum models leverage quantum kernels and variational circuits to map data into high-dimensional Hilbert spaces efficiently, capturing intricate correlations with reduced computational overhead. Additionally, quantum distance and similarity computations are exponentially faster, enabling more efficient clustering and classification while preserving essential data structures, making QML a powerful alternative for handling complex, high-dimensional data in applications like healthcare analytics. Another important property of the quantum models to explore is quantum supervision, which allows qubits to exist in a combination of both states 0 and 1 simultaneously. In contrast, classical bits can only exist in a single state, either 0 or 1, at any given time. This fundamental difference provides quantum models with significant advantages over classical models, especially in computational and data processing tasks.

This study proposes three quantum-based methodologies to address various healthcare challenges efficiently. The first approach focuses on the binary classification of Covid-19 patient data using E-QSVM and QRF models. By integrating both classical and quantum algorithms, this method enhances computational speed, providing physicians and caregivers with faster and more accurate diagnostic results. The second approach utilizes electronic healthcare records (EHRs) of cardiovascular patients, employing a QCNN to solve a binary classification problem. The study highlights that not all variables were necessary, as an effective pre-processing technique was applied to optimize predictions. This dimensionality reduction allows healthcare providers to focus on key factors affecting patients, leading to more accurate diagnoses and personalized treatment plans. The third methodology involves an experimental analysis of sarcopenia disease. Here, amplitude encoding is used for dimensionality reduction in VQC and QKNN models. Three distinct subsets of important features were tested to evaluate the performance of quantum and classical models. The results demonstrate superior efficiency and accuracy in quantum models when compared to recent studies employing similar classical approaches.

1.3.1 Contribution

The primary objective of this thesis is to evaluate complex, real-world medical datasets using advanced QML and Quantum Deep Learning (QDL) algorithms. The proposed quantum models demonstrate significant improvements in performance, validated through comprehensive comparisons with both classical counterparts and recently published state-of-the-art models. These comparisons primarily focus on accuracy metrics, showcasing the enhanced predictive capabilities and efficiency of quantum approaches in addressing complex healthcare challenges. This thesis is based on the following published journal articles and conference papers. Furthermore, the thesis clearly outlines the individual contributions to each publication, providing a detailed account of the original research, methodological advances, and key findings attribution.

Contribution 1: Ullah, Ubaid, Alain García Olea Jurado, Ignacio Diez Gonzalez, and Begonya Garcia-Zapirain. "A fully connected quantum convolutional neural network for classifying ischemic cardiopathy." IEEE Access 10 (2022): 134592-134605.

Contribution 2: Ullah, Ubaid, Danyal Maheshwari, Cristian Castillo Olea, and Begonya Garcia Zapirain. "Sarcopenia risk prediction and feature selection by using quantum machine learning algorithms." Quantum Machine Intelligence 6, no. 2 (2024): 80.

Contribution 3: Ullah, Ubaid, and Begonya Garcia-Zapirain. "Quantum machine learning revolution in healthcare: a systematic review of emerging perspectives and applications." IEEE Access (2024).

Contribution 4: Ullah, Ubaid, Danyal Maheshwari, Hanna Helene Gloyna, and Begonya Garcia-Zapirain. "Severity classification of COVID-19 patients data using quantum machine learning approaches." In 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME), pp. 1-6. IEEE, 2022.

1.4 Research Methodology

This section delineates the methodology employed in the research study presented within this thesis. This dissertation synthesizes insights from three distinct research investigations, each characterized by unique designs and methodologies with different pre-processing mechanisms. Consequently, each study inherently possesses its own contextual framework, materials, and procedural approaches. Nonetheless, the processes described here in figure 1.3 have been adapted to accommodate the specific requirements of each individual research endeavor.

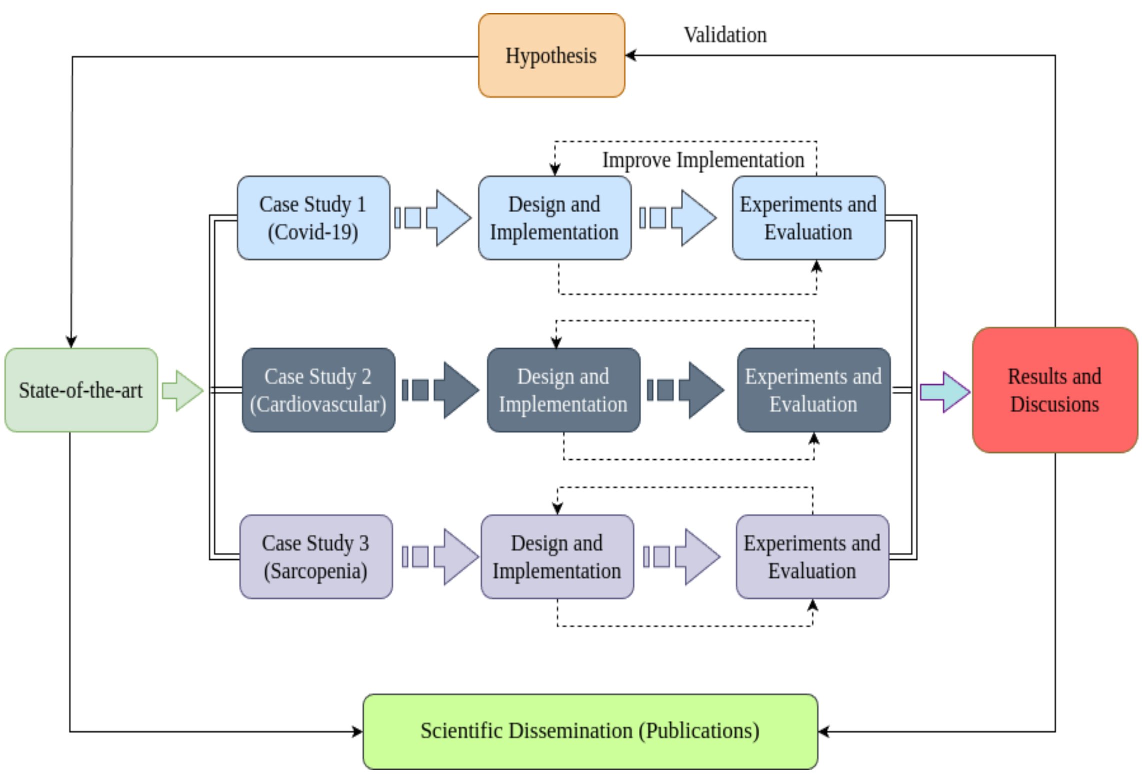

Figure 1.3: Research methodology followed for all three study cases.

* **Sate-of-the-art:** The state-of-the-art review encompasses an examination of the current literature, the development of various algorithms, diverse QML models, and

the available datasets within the healthcare domain. The knowledge acquired during this phase will inform the formulation in shaping the hypothesis and facilitating a subsequent comparison with the empirical results obtained.

* **Design and Implementation:** Following the analysis of the current state-of-the-art and the pre-processing of the knowledge acquired in the preceding step, this phase involves the design and implementation of the various components of the system.

* **Experiments and Evaluation:** At this stage, comprehensive evaluation methods are established to rigorously assess the system's performance. Based on the outcomes of the validation datasets, the system undergoes iterative redesigns to optimize functionality and address any deficiencies.

* **Result and discussion:** This step is dedicated to comparing the obtained results with the state-of-the-art benchmarks in the field. This comparative analysis is crucial for validating the research findings and assessing the robustness of the established hypothesis.

1.5 Thesis Outline

This section provides a detailed overview of the structure and content of the chapters included in this dissertation. The dissertation follows the PhD by publication format, meaning that several chapters are presented as articles that have been published in peer-reviewed journals with impact factors. Each chapter is designed to contribute to the overall research narrative, showcasing significant findings and discussions that align with the central theme of the dissertation.

Chapter 1 - Introduction: This first chapter provides the main concept of the thesis as well as the technique of the defined specific research. Its key objective is to have clear and simple principles of the carefully designed as well as a broad evaluation of the carried research study and system development. Furthermore, the hypothesis and particular desired outcomes provided in this section will perform an essential role in the entire completion of this study.

Chapter 2 - Literature Review: The second chapter of this dissertation provides a comprehensive overview of the current state of development in the primary areas relevant to this research work. It begins with an extensive analysis of quantum computing, including its algorithms and QML models, with a particular emphasis on quantum supervised learning models, which constitute the core focus of this dissertation. Following this, the chapter explores various types of healthcare data, such as imaging, biomedical signals, and MHR, with a primary focus on MHR datasets. The thorough review of the literature in these two domains is crucial for informing the final system design and enhancing the overall performance of the research study.

Chapter 3 - Case Study1 - Covid-19: This chapter focuses on the implementation of QML models using a real Covid-19 dataset, offering a comprehensive comparison between quantum and classical models. Additionally, it delves into the innovative pre-processing and feature selection techniques employed for model evaluation. These methodologies provide critical insights into the performance and efficacy of the models, highlighting the advantages and limitations of both quantum and classical approaches in the context of Covid-19 data analysis.

Chapter 4 - Case Study2 - Cardiovascular: This section details the contributions of the technical implementation of classical and quantum algorithms to cardiovascular disease classification. It explores how these algorithms are utilized to analyze cardiovascular data, enhancing the understanding and management of the disease. It also emphasizes the comparative effectiveness of classical and quantum methods, showcasing their respective roles in improving diagnostic accuracy, prognostic assessments, and therapeutic strategies in cardiovascular healthcare.

Chapter 5 - Case Study3 - Sarcopenia: Chapter 5 presents Case Study 3, which provides an evaluation of sarcopenia disease using QML models. It offers a concise overview of the assessment process, highlighting the significance of quantum state preparation and the application of amplitude encoding in the development and optimization of quantum models. This chapter underscores the importance of these quantum techniques in enhancing the accuracy and efficacy of sarcopenia diagnosis and analysis.

Chapter 6 - Conclusions: The 6th and final chapter of this dissertation presents the insights and conclusions derived from the comprehensive evaluation of the research work. This chapter revisits the specific objectives outlined in Section 1.1, assessing the extent to which they were achieved throughout the research process. Additionally, it offers a discussion on future research directions and potential developments based on the findings of this study, providing a forward-looking perspective on the implications and applications of the work conducted.

Chapter 2

Literature Review

2.1 Introduction

Machine Learning (ML) is a rapidly emerging computer field that is fuelled by massive amounts of data sent, collected, and analyses every day [7]. There are numerous ML applications and implementations with QC in the real world. The science of QC is exciting and has many practical applications which cover a wide range of topics [8]. The recent advancement of QC technology has made the processing of large data scale feasible, with QC demonstrating that it is able to solve challenging tasks much more quickly than classical computing [9] [10]. Quantum methods for computing offer new concepts and strategies to the field of machine learning and the development of computer-based technologies grounded in the principles of quantum theory [11]. The nature and behavior of energy and matter at a quantum level are described by quantum theory, where QC constitutes a collection of bits that work together to solve problems [12]. In the field of quantum machine learning, classical machine learning and quantum physics are combined. A symbiotic link exists between the construction of quantum versions of machine learning algorithms using quantum computers and the analysis of quantum systems using machine learning algorithms [13]. In the past few years, there has been notable growth in deep learning and machine learning, with these advanced technologies having found extensive application across various industries, ranging from the military, aerospace, and agriculture to banking and healthcare. Models developed through these methods have been widely adopted in these sectors [14, 15].

Recently, the fields of machine learning and QC have been gaining prominence and expanding opportunities in the field of Artificial Intelligence (AI) [16], and numerous researchers have successfully applied diverse quantum algorithms to real healthcare datasets. Among these applications QML stands out as a highly promising field, with multiple research teams actively engaged in its research [17]. Particularly significant is the exploration of novel machine learning approaches that leverage the advantages of QC within the healthcare domain. In this context, supervised learning has emerged as a noteworthy QML model that has garnered considerable attention from both academia and in healthcare. In the case of classification and diagnosing problems, exponential experimental improvements have been achieved through various contributions [18]. These include the development of QSVM, VQC, hybrid algorithms, quantum deep learning models, and error minimization algorithms, as well as pre-processing techniques. The design and implementation of Quantum Artificial Neural Networks (QANN) have provided a gateway for the application of additional algorithms in quantum states [19]. Furthermore, there are several methods for encoding classical data into quantum states that offer advantages such as reduced exploratory costs in terms of resources and the incorporation of non-linearity in the data [20]. Kernel-based methodologies have proven useful in achieving data linearity for linear classifiers. Similar to classical machine learning algorithms, researchers are also now concentrating their efforts on establishing comprehensive quantum algorithms capable of solving classification problems specific to the healthcare industry. QC techniques have been merged with classical ML for various applications, with these algorithms using different private and publicly available datasets such as the University of California Irvine (UCI) ML repository, Iris, and MNIST dataset [21] [22].

QML has the potential to revolutionize healthcare by leveraging the power of QC in order to analyses complex healthcare data and make more accurate predictions [23]. It can process large and diverse healthcare datasets, including electronic health records (EHRs), medical imaging data, genomic data, and sensor data, with a view to improving disease diagnosis and prognosis [24]. EHRs contain a wealth of patient information, including medical history, diagnoses, treatments, laboratory results, and so on, and these comprehensive datasets offer great potential for advancing healthcare research and in improving patient outcomes [25]. QML techniques can be further used to leverage the power of QC and extract valuable insights from EHRs. Moreover, QML holds tremendous potential in the field of medical imaging, where it can make a significant impact. Medical imaging techniques, such as X-rays, CT scans, MRI scans, and ultrasound, play a critical role in diagnosing and treating various diseases and conditions. However, these imaging processes often require more computational power and can be time-consuming, leading to delays in diagnosis and treatment [26]. Moreover, the advent of QML brings the promise of speeding up image processing and analysis, thus revolutionizing the field of medical imaging. By harnessing the principles of quantum mechanics, quantum algorithms can tackle complex computational tasks more efficiently than classical algorithms, with this advancement opening up new possibilities for doctors and medical professionals in providing faster and more accurate diagnoses. By analyzing these datasets, QML algorithms can identify patterns, correlations, and bio-markers that traditional machine learning algorithms might miss, leading to more accurate predictions of diseases and their progression [27] [28]. For simplicity, the hierarchy of the entire chapter is shown in figure 2.1, which further contain on various sections and subsections.

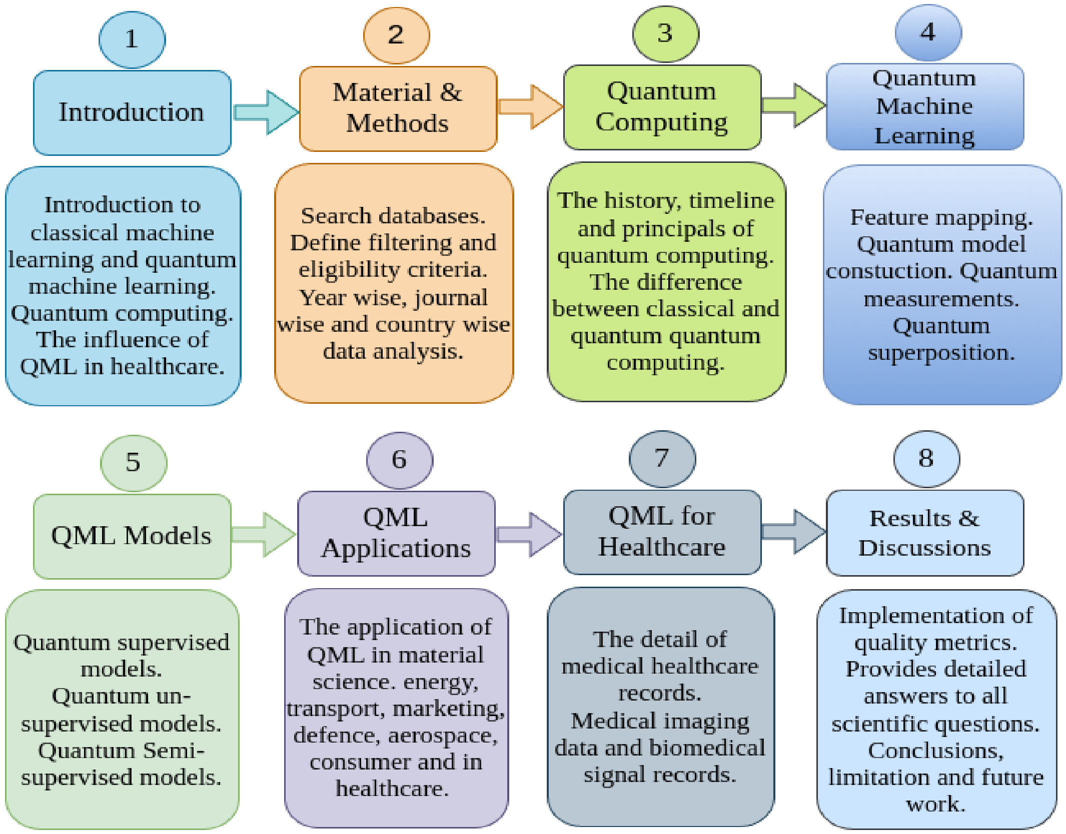

Figure 2.1: Hierarchy model for the entire chapter.

Beside the healthcare, QML has shown immense potential across various industries. For an instance, in the financial sector, quantum computing is utilized for portfolio optimization, risk assessment, and fraud detection [29]. By leveraging quantum algorithms such as the Quantum Approximate Optimization Algorithm (QAOA) [30] and Monte Carlo methods, it enhances computational efficiency and accuracy [31]. In cybersecurity, quantum cryptography and Quantum Key Distribution (QKD) [32] have been extensively researched to strengthen secure communications, ensuring robust data protection against both classical and quantum cyber threats. Additionally, quantum-resistant encryption techniques are being developed to safeguard sensitive information from potential quantum attacks [33]. QML models are also widely applied in material science, utilizing quantum technologies to solve complex challenges in material design, discovery, and characterization. These models efficiently simulate electronic structures and material properties at the atomic level [34] [35]. For example, Variational Quantum Circuits (VQC) and QAOA assist in determining the ground-state energy of new materials, accelerating Density Functional Theory (DFT) calculations used in material design, and enabling defect detection in semiconductors, superconductors, and nano materials by analyzing quantum states [36]. Furthermore, QML has been used in addressing climate change challenges through advanced data analysis and predictive modeling. Recent studies have explored various applications of QML in this field, demonstrating its ability to enhance the accuracy of climate models. Techniques such as VQC, QBoost, and Quantum Support Vector Classification (QSCV) have been applied to process complex climate data more efficiently, leading to improved climate pattern forecasting [37]. In flood prediction, the integration of QML with classical machine learning models has shown promising results [38]. Research conducted on daily flood events along Germany's Wupper River indicated that QML models offer competitive training times and enhanced prediction accuracy, underscoring their potential in climate change adaptation strategies.

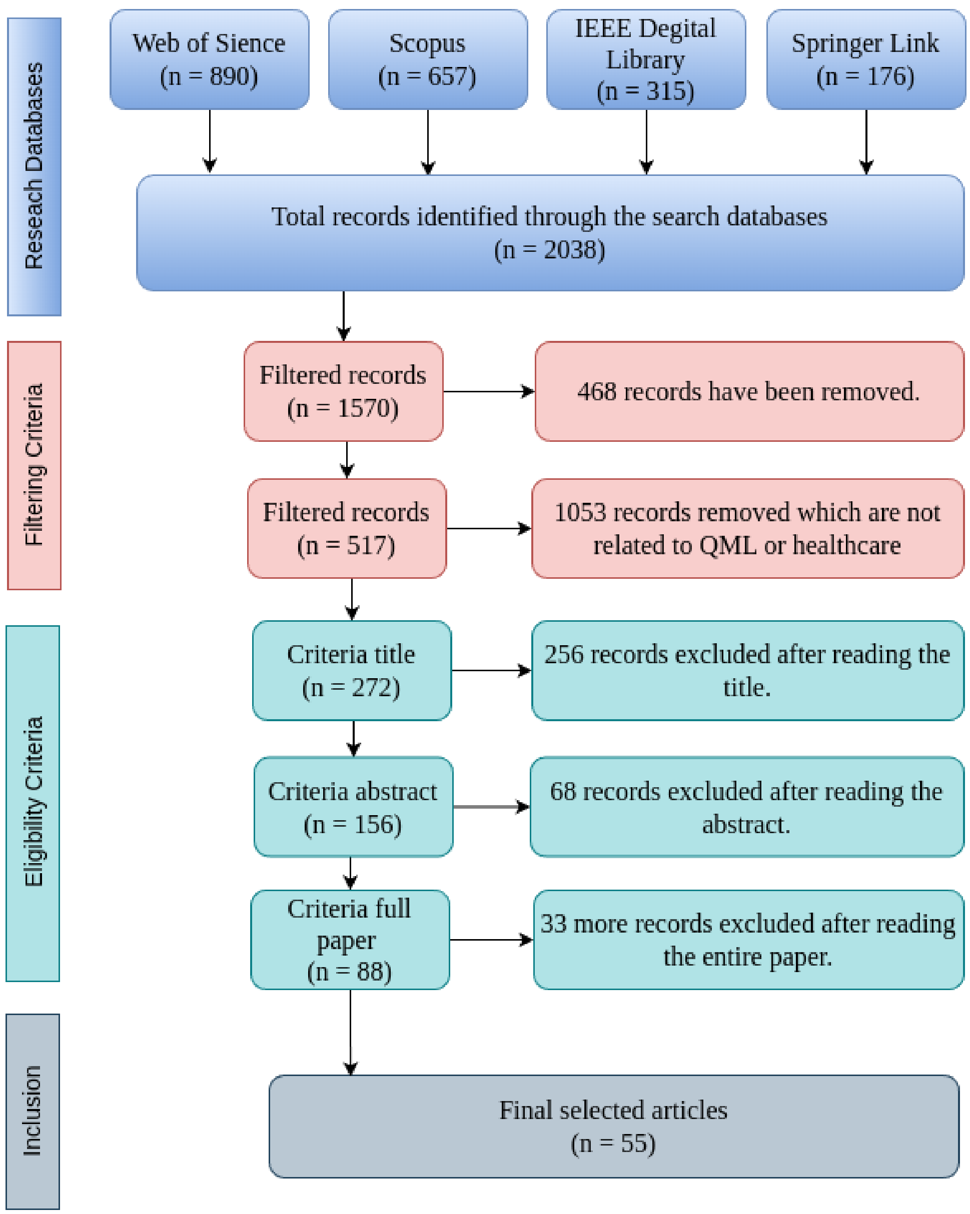

Figure 2.2: Prisma diagram for the records selected.

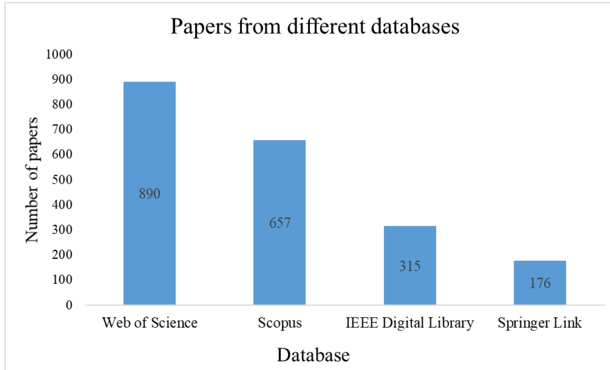

Figure 2.3: Number of papers from different data bases.

2.2 Materials and Method

This review aims to demonstrate solutions and advancements for the development and adaptation of various quantum processing methods, quantum machine learning models, and quantum simulation tools in the healthcare domain. The work serves as an introduction to this emerging area, along with a review of recent developments and a presentation of the issues that have yet to be resolved in earlier survey studies [39–42]. In comparison to classical computing (CC), QC is an emerging field. I have focused on the developments and applications of the previous seven years as a result, and have contrasted the various approaches used as well as the challenges.

2.2.1 Search Databases

Papers that fit the criteria for this review include those that (a) concentrate on learning patient representations, (b) use patient data, such as EHR, images, and ECG signals, and (c) employ quantum machine learning and quantum deep learning models. In contrast to other techniques, ML and DL are a practical data-driven solution to build robust representation simultaneously from a variety of resources, and it is thus noteworthy that I consider research utilizing QC techniques as one criterion. This is because QC techniques have unique characteristics that make them a viable data-driven solution. Keyword searching queries for literature database searches were generated by applying these criteria, and articles from the past five years from 2019 to 2024 were extracted from four distinct databases, including Web of Science, Scopus, IEEE Explore Digital Library, and Springer Nature.

2.2.2 Search String

A systematic search strategy was implemented using a combination of keywords and Boolean operators (AND, OR) to retrieve relevant articles on Quantum Computing (QC) and Quantum Machine Learning (QML) in healthcare applications. The search queries were constructed using various terms and their combinations. For instance, keywords such as "Quantum Computing" AND "Quantum Machine Learning" OR "Quantum Neural Network" were used to capture general research on QC and QML. Additionally, healthcare-specific terms were incorporated using ("Cancer" OR "Oncology" OR "Tumor") AND ("Machine Learning" OR "Quantum Computing"), as well as ("Diabetes" OR "Pancreas") AND ("Quantum Machine Learning") to refine disease-specific results. Other relevant terms, such as "Cardiovascular Disease" OR "Stroke" OR "Epilepsy" OR "Alzheimer" OR "Omics" OR "Genomics", were also combined using Boolean operators to enhance search precision and coverage.

2.2.3 Filtering Criteria

In terms of publication characteristics, we restricted our search to English-language academic papers and concentrated only on conference proceedings and peer-reviewed journals (posters and pre-prints were excluded). Only study designs relevant to creating a quantum ML or quantum DL based on any type of health-related dataset were taken into account as study features. The study results should also be used to downstream clinical prediction evaluations. It is important to note that review papers, proceeding summaries, and QML articles not connected to healthcare applications were excluded from the study. By using a snowballing strategy [43], duplicate records were eliminated and additional records were added, which were gathered based on references in the papers and personal readings. From out of the total 2038 records collected from four open sources database, 468 articles were deleted in the filtering criteria stage which were most likely duplicated. Additionally, from 1570 records, a further 1034 were eliminated after filtering again, because most of them were not related to QML or considered to be other data for model evaluation purposes, as shown in figure 2.2.

2.2.4 Eligibility Criteria

Eligibility criteria include three essential reading techniques: reading the title, reading the abstract, and reading the entire article. Following filtering criteria, 272 records were considered for further processing in order to check their eligibility. I obtained 272 records following the filtering stage, from which 116 more records were eliminated at the title screening stage because the topic under consideration did not meet our specific requirements. Furthermore, 68 additional records were excluded after reading the abstract. Any articles that use classical ML instead of quantum ML, use text datasets, or are related to industrial, material science, or transportation applications were removed. From the remaining 88 records, 33 records were further excluded in accordance with the criteria of the full text screening stage, where each study was researched in full detail to ensure the relevant records. Finally, 55 records were selected based on three categories: the Electronic Healthcare Record (EHR) dataset, imaging dataset records and biomedical signal record, which are the most appropriate for the purpose of our proposed work.

2.2.5 Inclusion Criteria

Finally, based on the inclusion criteria, 55 records were selected, all of which are related to QML, QDL, and QC in healthcare applications. The selection also includes articles that present solutions implemented through software or those that can be simulated or executed on quantum devices or simulators. Additionally, only records published between 2019 and 2024 in peer-reviewed journals or conference proceedings were considered.

2.2.6 Exclusion Criteria

Papers that did not explicitly incorporate QML methods or did not utilize healthcare-related data were excluded from this review. Additionally, studies that focused solely on classical machine learning techniques without integrating quantum computing principles were not considered. Furthermore, articles lacking biomedical applications or those primarily discussing theoretical aspects of QML without practical implementation in healthcare were also filtered out to ensure relevance and specificity in the review.

2.2.7 Data Analysis

This section contains data scrutiny and data categorization for the records selected. All the data was analyzed by taking three different strategies into consideration, and the corresponding complete texts of 55 records that met the necessary requirements were deemed eligible for review were then evaluated. As a consequence, the following information was obtained.

2.2.7.1 Year Wise Evaluation

The records selected were assessed based on their respective years with a view to examining the level of interest among scholars in working with QML over the past few decades. As shown in figure 2.4, notable contributions in the field of QML for healthcare began to emerge after 2018. In 2019, there were 9 records, followed by 10 in 2020, 9 in 2021, 12 in 2022, 9 in 2023, and up to the present date in 2024 6 articles in total have been considered in this study.

2.2.7.2 Country Wise Evaluation

Additionally, I also conducted a country-wise analysis of the records selected, as illustrated in figure 2.5, and this evaluation generated further interest in the subject by

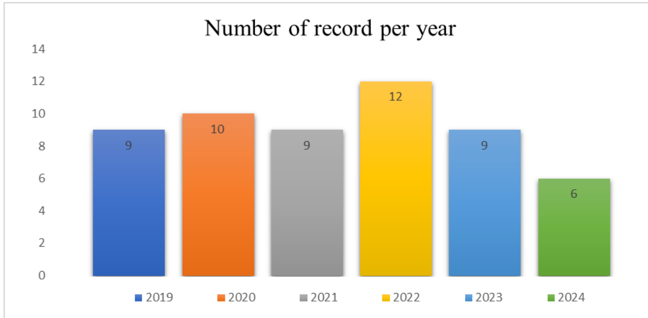

Figure 2.4: Number of papers per year.

taking both the publication year and the different geographical locations into consideration. Among the research studies examining the adoption of QML in the healthcare industry, India stands out as the leading contributor with 15 papers, followed by USA as the second-highest contributor with 8 articles to date. Spain has 7, China and France has 3, Saudi Arabia, Columbia, Pakistan, Turkey, Iraq and Austria has 2 articles and so on.

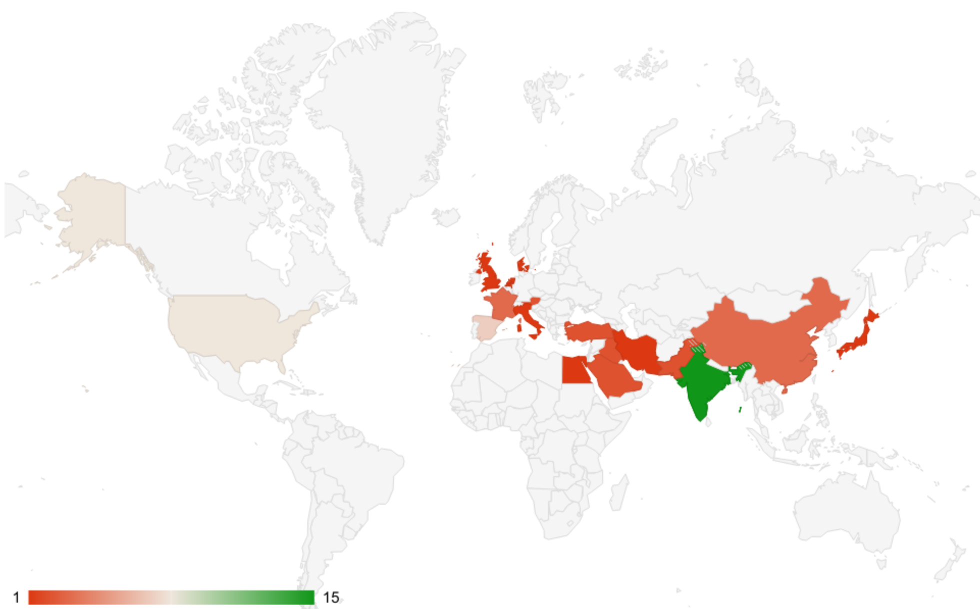

Figure 2.5: Mapped the published article in quantum machine learning and healthcare domain as per country-wise.

2.2.7.3 Type of Selected Articles

Moreover, the articles that were chosen underwent a thorough evaluation based on type of article, with figure 2.6 illustrating the various types of articles considered in this study. Out of total records, it was found that 37 records, accounting for approx-

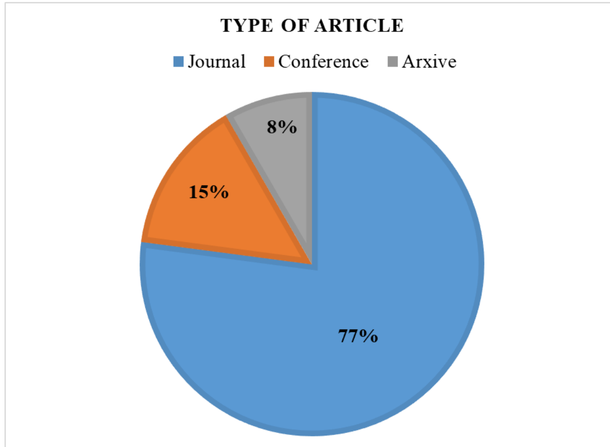

Figure 2.6: Types of selected published articles.

imately 77%, originated from international journals. Furthermore, 7 records, constituting approximately 15% of the total, were sourced from international conferences. Finally, 4 records, making up approximately 8% of the total, were selected from arXiv.

Furthermore, following evaluation and analysis of the desired records selected, I defined a couple of research questions and proposed a point-oriented quality based metric as shown in table 2.1 and table 2.2 respectively.

2.3 What is Quantum Computing

The concept of QC was introduced in the early 1980s by Benioff and Feynman, who stated that quantum-based computers outperformed their classical equivalents when it came to solving particular problems [44]. Feynman suggested that quantum mechanics could be used to solve computational problems by simulating complex quantum systems with standard quantum systems. This approach allows problems to be solved that classical computers are unable to solve [45]. Feynman's idea had a direct impact on the advancement of QC, and gave rise to the notion of a quantum Turing machine, also leading to the theoretical proof of the existence of universal models based on quantum mechanics [46]. Quantum algorithms can also be used to ascertain the processing power of classical computers, with development of one of the earliest quantum algorithms that offered a speed advantage over its classical counterparts being provided in [47]. The algorithm was designed to probabilistically determine whether a two-bit function is balanced or constant using just one function call. Researchers have proposed subsequent quantum algorithms in [48] [49] that demonstrate the superiority of QC over classical computers in solving specific problems. However, these problems are manually designed, and their practical applications are limited.

QC relies on the principles of quantum mechanics, which employs observable quantum phenomena such as quantum entanglement and quantum superposition and quan-

Table 2.1: The point guided quality metrics proposed for selected record evaluation.

Type of Metrics	S.No	Overview	Merit	Points
Flow of the paper (50 points)	1	A well-explained precise abstract was provided - a short overview of the proposed model and datasets used; the solution to the problem or an improvement in experimental results was provided.	0 to 10	10
	2	The highlighted problem and background to it were briefly explained in the introduction.	0 to 10	10
	3	Simulation or optimisation of the proposed QML model was explained in detail, with mathematical modelling and diagrams.	0 to 20	20
	4	The challenges and limitations of the work were discussed, and the methodology adopted to overcome these challenges explained.	0 to 10	10
Materials presented in the paper (30 points)	5	Experimental results were discussed in detail in the form of plots or confusion metrics; results were compared to recently published models.	0 to 10	10
	6	Materials used were briefly discussed, and the type of dataset used and its availability clearly described.	0 to 10	10
	7	The code for the proposed work is publicly available.	0 to 5	5
	8	Originality or innovation of the proposed research work.	0 to 5	5
Outcome presentation (15 points)	9	Determining system in terms of accuracy, loss function, and error.	0 to 15	15
Quality Measures (5 points)	10	Number of Citations.	0 to 5	5

tum interference [50]. The quantum entanglement property is a non-intuitive phenomenon, famously referred to by Einstein as "spooky action at a distance," whereby an entangled pair of electrons always spin in opposite directions and influence each other through time and space, even when not physically connected. This property provides quantum algorithms with significantly more power than conventional ones. The quantum superposition is the property of an electron in which its position cannot be precisely determined at any given time. Instead, the electron's position is described by a probability distribution, where it has a chance to exist in all possible locations simultaneously, with varying probabilities. The quantum interference property refers to the

Table 2.2: Possible research questions.

Question	Purpose
Question1: How can quantum machine learning algorithms be utilised to enhance medical data analysis, such as improving disease segmentation, classification, or anomaly detection?	Explore the potential of QML models, such as QSVM, QKNN, QRF or quantum neural networks, such as QNN, QCNN in extracting meaningful insights from electronic healthcare records for diagnosis, treatment planning, or personalised medicine.
Question2: How can quantum machine learning models be integrated with classical machine learning approaches to leveraging the strengths of both in healthcare data analysis?	Research into hybrid approaches that combine classical and quantum machine learning techniques to exploit the computational power of quantum algorithms while leveraging the robustness and interpretability of classical models, enabling enhanced analysis and interpretation of healthcare data across various domains.
Question3: What are the fundamental limitations and advantage of quantum machine learning models in handling healthcare data compared to classical machine learning models?	Explore the unique properties of quantum machine learning models, such as quantum superposition and entanglement, and assess their potential benefits and limitations when applied to healthcare data analysis tasks, taking factors such as interpretability, scalability, and the computational resources required into consideration.
Question4: What are the considerations and methodologies for evaluating the robustness and generalisability of quantum machine learning models when applied to diverse healthcare datasets, including data from different hospitals, regions, or demographic groups?	Ensure that the models can perform reliably and accurately across different settings, populations, and data sources. This evaluation helps identify any biases, limitations, or performance variations that may arise and enables the development of more effective and equitable healthcare applications using quantum machine learning techniques.
Question5: Which types of data can be used for adoption of a quantum predictive model, and is it feasible to use open access datasets for evaluating such models? Additionally, what specific quantum computing devices are applicable for evaluating QML models using healthcare data?	How can different types of datasets, including private and open-access datasets, be evaluated for quantum machine learning models? What are the options available for accessing quantum simulators, quantum real devices, and quantum processing units (QPUs) in order to assess the performance of these models with healthcare data?

ability of an individual particle, such as a photon, to interfere with its own trajectory and alter its path's direction. The technology used for constructing qubits, which are the fundamental units of quantum computers, is rapidly advancing. A quantum computer leverages an unusual observation from quantum physics in which a single bit can exist in both states of '1' and '0' at the same time, this unique bit being referred to as a quantum bit or qubit [51]. By utilizing these principles it builds a highly efficient computing system that can process multiple data pieces concurrently, resulting in the ability to handle massive amounts of information in a real-time scenario [52]. The basic concept of QC is to research into the inherent challenges of data analysis, data storage and its processing [53], and quantum mechanical systems are established by encoding information, which is commonly known as quantum information in terms of the state of a quantum system [54]. The QC development in terms of increasing qubits is shown in figure 2.7.

There are several differences between CC and QC in terms of capabilities, computational power and error rates. In CC, information is processed using bits that represent 0 or 1, while QC employs quantum bits or qubits capable of representing 0, 1, or even both states simultaneously through superposition. CC relies on multiple transistors to create logical switches and gates, whereas QC utilises quantum dots and supercon-

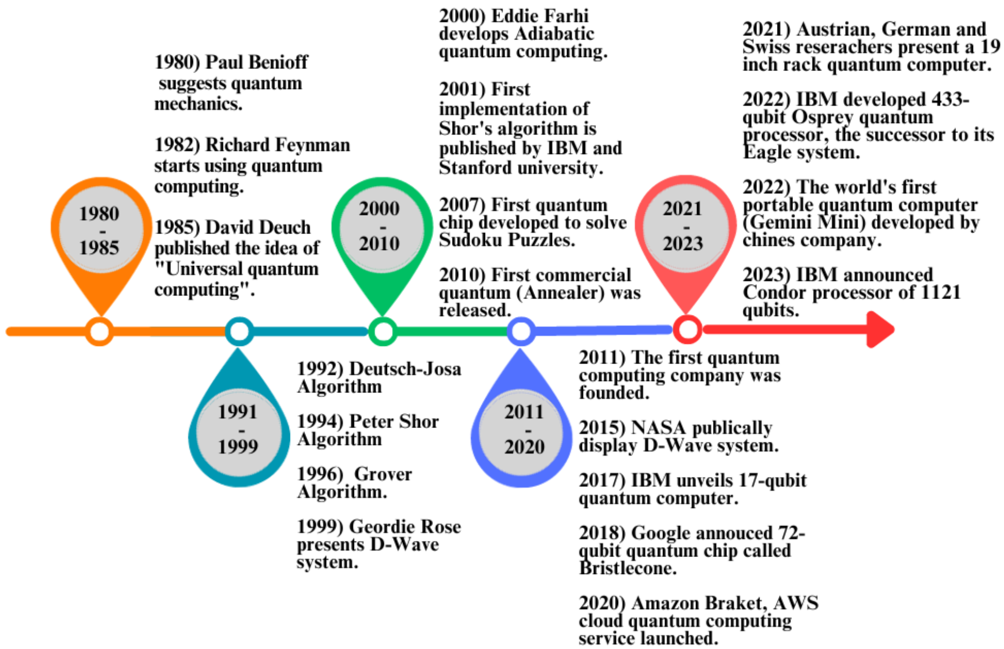

Figure 2.7: Quantum development timeline till to date.

ducting loops to create qubits, with several qubits forming a logical qubit. In terms of scalability, CC's computing power increases linearly with the addition of more transistors, while QC's potential grows exponentially with the inclusion of more qubits. Moreover, CC operates at room temperature with relatively low error rates and finds applications in general-purpose computing. Conversely, QC operates under extremely low temperatures, has a higher error rate, and specialists in tasks such as factoring, optimisation, and complex processing, as shown in figure 2.8.

2.4 What is Quantum Machine Learning

In quantum machine learning, quantum algorithms have been developed to address fundamental machine learning challenges while using the computational capability of QC. The standard approach to achieving this is to modify classical algorithms or their subroutines so that they may function on a hypothetical quantum computer [55]. Generally, QC and machine learning is combined in four different ways, i.e, classical data quantum model, quantum data classical model, classical data quantum model, and quantum data quantum model, as shown in figure 2.9. In such a scenario, our focus is on a third "quantum enhanced-machine learning" approach which is most frequently employed for analysis of classical data using quantum models [56]. In order to im-

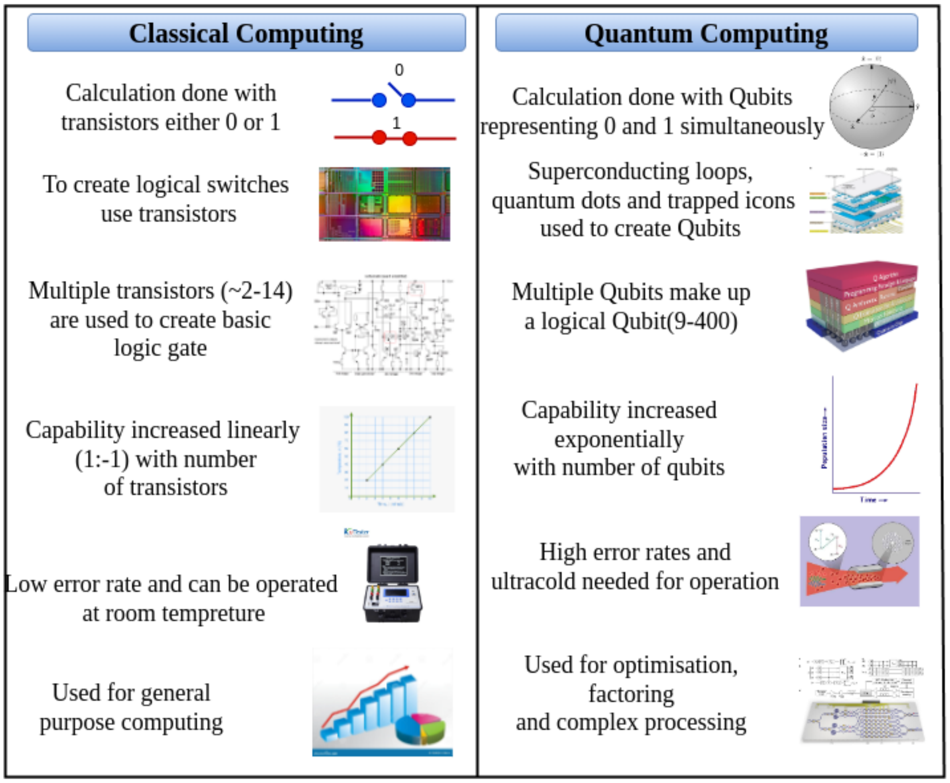

Figure 2.8: Classical computing vs quantum computing.

prove computing speed and data storage, qubits and quantum operations are utilised in the QML algorithm, and computationally challenging subroutines are handed off to a quantum device in hybrid approaches that combine classical machine learning approaches with QC, as shown in figure 2.10. On a quantum computer, such processes may be more complex and run faster [57–59].

The workflow of QML typically involves several steps, including system backend, state preparation, feature mapping, unitary gate operation, and quantum measurement, as illustrated in figure 2.11. Feature mapping is the process involving encoding classical data into a quantum state representation. In classical machine learning, features are usually represented as vectors or matrices, while in quantum machine learning, these classical features are transformed into quantum states by mapping them onto qubits. There are various techniques that can be used for feature mapping, depending on the specific problem and the quantum algorithms available, with one common approach being to use quantum circuits to transform classical features into quantum states. Once the classical data is mapped to a quantum representation, a quantum model is then constructed in order to process and analyse the quantum states. The quantum model is typically composed of quantum gates, which are operations that act on the qubits and manipulate their quantum states. Different quantum models can be employed depending on the specific learning task for example, a quantum neural network can be constructed using layers of quantum gates to perform computations

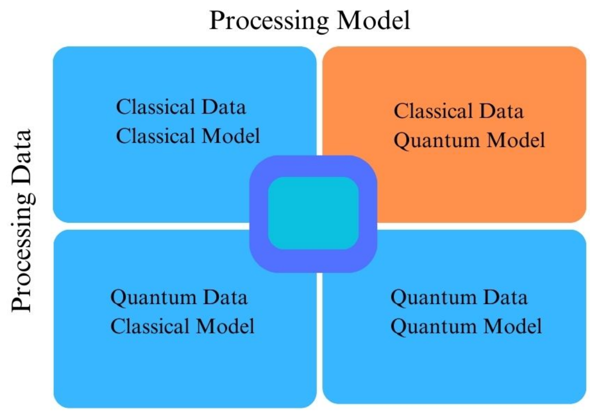

Figure 2.9: Matrix of various QML algorithms.

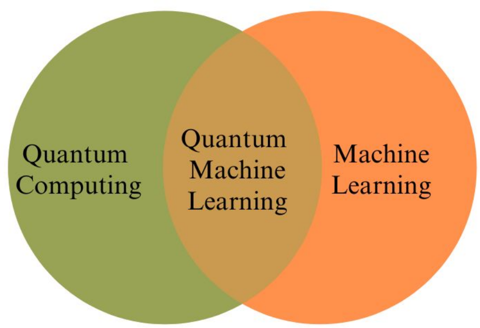

Figure 2.10: Quantum machine learning intersection.

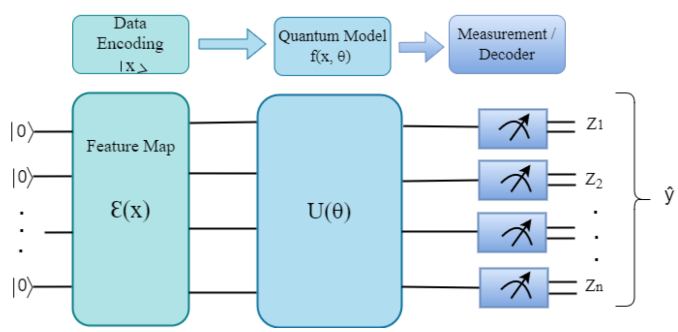

Figure 2.11: Basic steps of qml model.

analogous to classical neural networks.

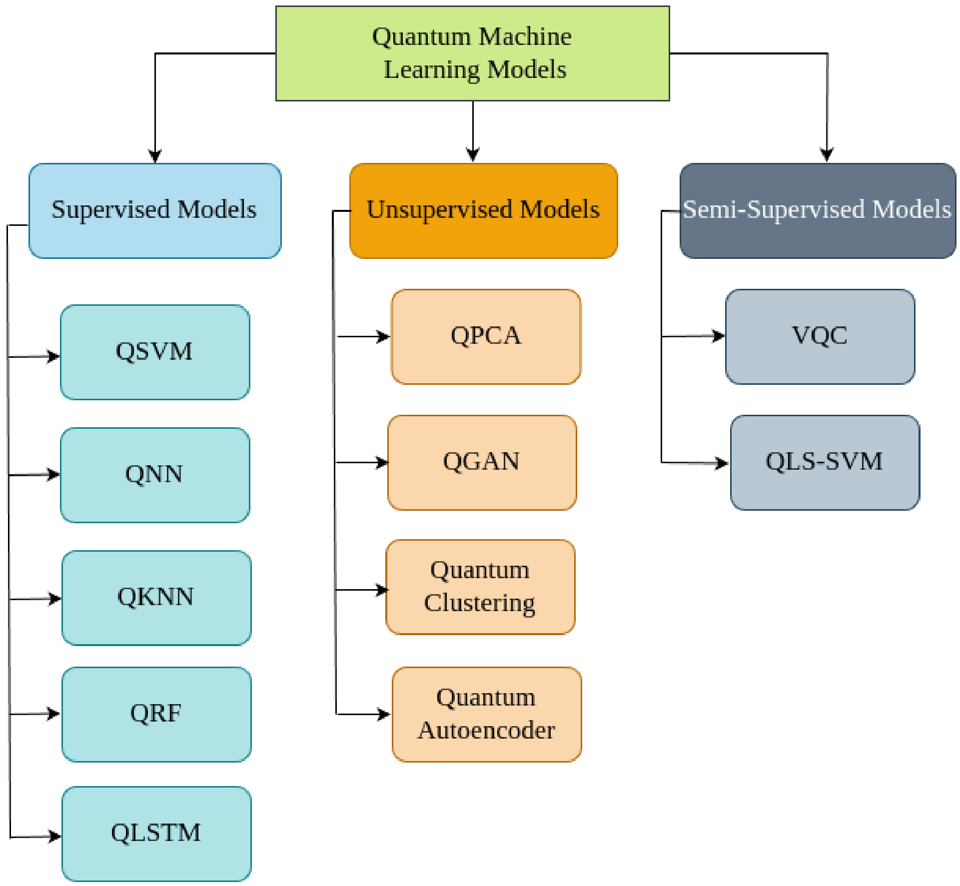

Figure 2.12: Various type of QML models.

The parameters of the quantum gates in the model are adjusted during the training process to optimise the model's performance. Likewise, there is the quantum support vector machine (QSVM), which utilises quantum circuits and quantum algorithms to perform classification tasks. QSVM maps the input data to quantum states, performs quantum computations on these states, and uses quantum measurements to classify new instances.

Eventually, once the quantum model has processed the data, a quantum measurement is then taken to obtain classical output results. In quantum mechanics, when a quantum system is measured, its superposition collapses into a specific state with a certain probability, and in the context of quantum machine learning, the quantum measurement operation extracts classical information from the quantum states and provides the output for the model. The measurement outcome is typically a classical bit string representing the result of the computation. Statistical techniques are often applied to analyse the measurement outcomes and extract relevant information or make predictions. For instance, in classification tasks, the measurement outcome can correspond to a class label. It is important to note that the measurement operation inherently introduces randomness due to the probabilistic nature of quantum systems, and so consequently, quantum machine learning models may require repeated measurements or statistical sampling in order to obtain reliable results.

2.5 Quantum Machine Learning Models

QML models are categorised into three basic categories: supervised model, unsupervised and semi-supervised, as shown in figure 2.12 and as follows:

2.5.1 Quantum Supervised Models

Supervised ML is a type of ML where a model learns from labelled training data to make predictions or decisions. According to this approach, the model is provided with input data along with corresponding output labels or target values. There are various supervised machine learning models available, and the choice of model depends on the specific problem you are trying to solve and the nature of the data. Here are a few commonly used supervised learning models:

2.5.1.1 Quantum Support Vector Machine (QSVM)

A QSVM is an ML algorithm that combines the principles of QC with the concept of a classical SVM. It is designed to perform classification tasks on quantum data or to exploit quantum effects for improved classical SVM training. For its part, an SVM is a supervised learning algorithm used for classification and regression tasks. Given a set of labeled training data, the purpose of the SVM is to find a hyperplane in a high-dimensional feature space that best separates the data into different classes. The hyperplane is chosen to maximise the margin, which is the distance between the hyperplane and the nearest data points of each class. QC, on the other hand, leverages quantum phenomena such as superposition and entanglement to perform computations on quantum bits or qubits. Qubits can exist in multiple states simultaneously, allowing for parallel processing and potentially more efficient algorithms. In a QSVM, the classical SVM is enhanced using quantum techniques, and the main idea is to map the input data to a quantum feature space, where the QSVM can perform classification, and this mapping is done using a quantum kernel function [60]. In classical SVM, a kernel function computes the similarity or distance between data points in the input space, while in the case of QSVM, the quantum kernel computes the similarity or distance between data points mapped to a quantum feature space. The input classical data is encoded into quantum states by using various quantum algorithms, and the quantum data is then fed into a quantum circuit that applies the quantum kernel function. This circuit takes advantage of quantum operations and entanglement to perform computations on the quantum data. As shown in figure 2.13, in the case

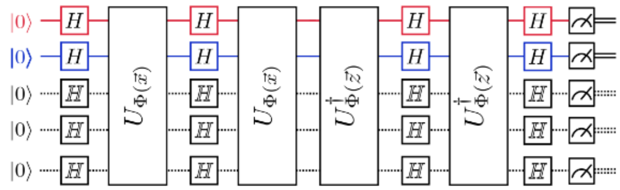

Figure 2.13: QSVM circuit diagram
[61].

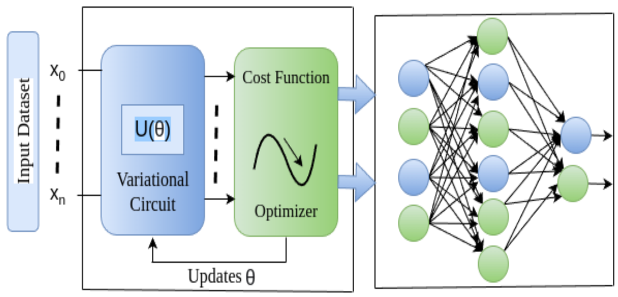

Figure 2.14: QNN model.

of quantum kernel computation, measurements are taken on the quantum circuit in order to obtain classical information and the measurement circuit then extracts the relevant features from the quantum state.

2.5.1.2 Quantum Neural Network (QNN)

A computational neural network model based on the principle of quantum mechanics is known as a QNN. In 1995, Subhash Kak [62] and Ron Chrisley [63] independently published the initial concepts regarding quantum neural computation. Traditional computers store and process information as binary bits, which can represent either a 0 or a 1. In contrast, quantum computers use quantum bits, or qubits, which can exist in a superposition of both 0 and 1 states simultaneously. This allows quantum computers to perform certain calculations much faster than classical computers in the case of certain types of problems. In QNN, the basic building block is the addition of a quantum layer, which is analogous to the artificial neuron in a classical neural network as shown in figure 2.14. The quantum neuron processes and transmits information using quantum operations. Typically, a qubit is used to represent the state of the quantum neuron, with the state being manipulated using quantum gates. The activation function in a classical neural network is replaced by a quantum gate, which transforms the state of the qubit based on the input. Various types of quantum gates can be used, such as the Hadamard gate, controlled-phase gate, or any other gate that is able to manipulate the quantum state, and these gates introduce quantum effects, such as superposition and entanglement, into the computation. Training a quantum neural network involves adjusting the parameters of the quantum gates to optimise the network's performance

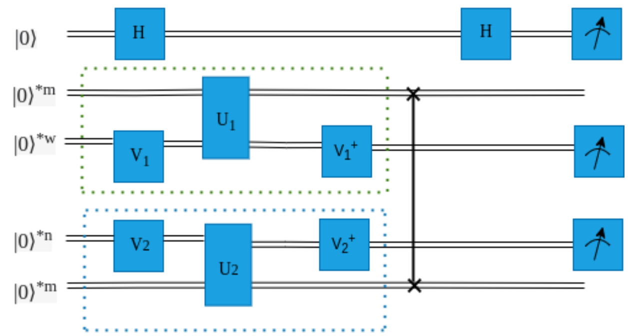

Figure 2.15: QKNN circuit diagram [64].

in a specific task. One of the potential advantages of quantum neural networks is their ability to process and analyse large amounts of data simultaneously, thanks to quantum superposition. This can be especially useful for tasks such as pattern recognition, optimisation, and machine learning, where large-scale parallelism can provide computational advantages.

2.5.1.3 Quantum K Nearest Neighbor (QKNN)

Quantum KNN is also a supervised ML algorithm, which uses the quantum properties (superposition and parallelism) to process the classical KNN classification algorithm. The QKNN algorithm introduces quantum mechanics principles, specifically quantum superposition and quantum interference, to improve the performance of the KNN algorithm. It utilises a quantum representation of data and exploits quantum parallelism in order to perform the distance calculations more efficiently, while the input training and testing dataset can be converted into vectors. The classical CC method requires n bit registers for the desired dataset, while the QC needs n qubit superposition quantum states. All of the binary numbers in the datasets can be converted into n qubit states with the appropriate probability, which reduces storage capacity. As shown in figure 2.15, the five registers contain five qubits, with the first auxiliary qubit being operated by Hadamard gate. The second and third registers are used for testing the dataset, while the fourth and fifth registers are used for training it. In the QKNN algorithm, quantum parallelism is leveraged to calculate the distances between the encoded new data point and all the encoded training data points simultaneously, with this parallel distance calculation being obtained using quantum gates and quantum circuits. Finally, the class label predicted is determined based on the outcome of the measurement. The QKNN algorithm assigns the class label that is most prevalent among the K nearest neighbors in the collapsed quantum state.

2.5.1.4 Quantum Random Forest (QRF)

A quantum RF is a variant of the classical random forest algorithm that incorporates principles and techniques from QC. It leverages the power of quantum properties, such as superposition and entanglement, to enhance performance and capabilities of the random forest model. In a classical random forest, multiple decision trees are created and trained on different subsets of the training data, and each decision tree is built using a random selection of features and employs a majority voting scheme to make predictions. The final prediction of the random forest is determined by aggregating the predictions of all the individual decision trees. In a quantum random forest, the underlying decision trees are replaced by quantum decision trees, which are quantum circuit representations of decision trees, and these quantum decision trees utilise quantum gates and qubits instead of classical bits and logic gates, as shown in figure 2.16. A qubit is the fundamental unit of quantum information, and can exist in a superposition of states, representing both 0 and 1 simultaneously. To form a quantum random forest, an ensemble of quantum decision trees is created through a process called quantum bootstrap aggregating or quantum bagging. Multiple copies of the training data are created, and each copy is randomly perturbed to introduce diversity. Quantum decision trees are then built on these perturbed copies, and the final prediction is obtained by aggregating the predictions of all the quantum decision trees through majority voting.

2.5.1.5 Quantum Long Short Term Memory (QLSTM)

LSTM networks are a type of recurrent neural network (RNN) designed to handle long-term dependencies in sequential data. They are particularly effective in tasks such as speech recognition, language translation, and time series analysis. LSTMs are composed of cells that maintain an internal memory state, allowing them to remember information over extended time intervals. Each cell is equipped with three main components: an input gate, a forget gate, and an output gate, and these gates regulate the flow of information into, out of, and within the cell. By replacing the classical neural networks in the LSTM cells with VQCs, the classical LSTM can likewise be transformed into the quantum state [65]. The QLSTM schematic architecture shown in figure 2.17 consists of three layers - the data encoding layer, the variational layer, and the quantum measurement. The input classical vector is converted into a quantum state by the data encoder layer and the variational layer is the learnable component, where an optimisation technique is used to update the circuit parameters. Eventually, the quantum measurements are employed in order to obtain the values required for further processing.

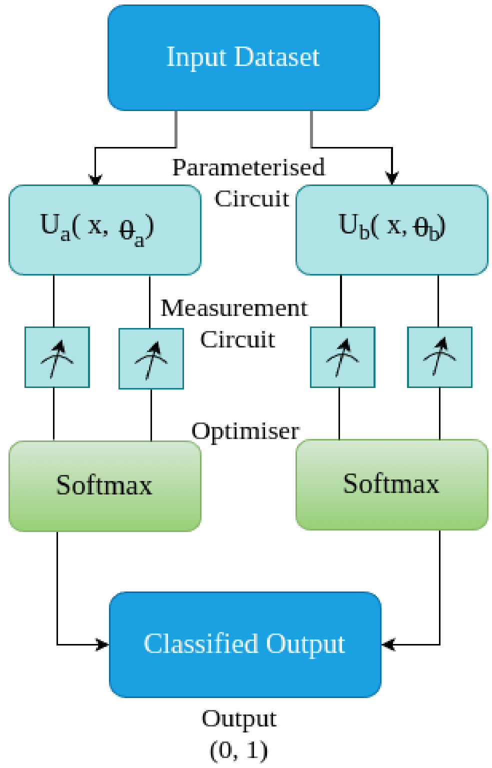

Figure 2.16: QRF diagram [85].

2.5.2 Unsupervised Models

Unsupervised learning analyses and clusters unlabeled information using machine learning algorithms to identify data groupings or hidden patterns without requiring human involvement.

2.5.2.1 Quantum Principal Component Analysis (QPCA)

Quantum PCA is a QC algorithm that is an extension of the classical PCA technique used in machine learning and data analysis. PCA is a dimensionality reduction method whose purpose is to find the most significant features or patterns in a dataset by projecting it onto a lower-dimensional subspace. The input dataset is transformed into quantum states by encoding it in the amplitudes of a set of quantum bits (qubits). The number of qubits required depends on the size of the dataset and the desired precision. The encoded quantum states are then prepared on a quantum computer using quantum gates and operations, and this step initializes the quantum system in order to represent the input data. For its part, the quantum PCA algorithm employs quantum phase estimation to estimate the eigenvalues of the covariance matrix, with this step being critical in determining the principal components. The eigenvalues are estimated from the results of the quantum phase estimation, which provides information about the variance captured by each principal component. Estimated eigenvalues are used to

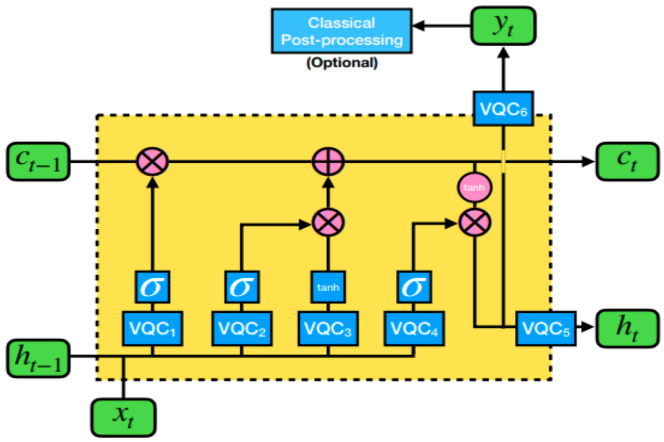

Figure 2.17: Schematic architecture of QLSTM [65].

identify the principal components, while the eigenvectors corresponding to the largest eigenvalues represent the principal components of the dataset. The quantum states representing the principal components are then measured to obtain classical information, although additional classical post-processing may be required in order to finalise the output and interpret the results.

2.5.2.2 Quantum Generative Adversarial Network (QGAN)

The quantum GAN use two neural networks a generator and a discriminator that are simultaneously trained are used to generate data that is identical to the original data used in training. The generator generates fake data that mimics the actual training dataset, while the discriminator works like a detective, aiming to differentiate between actual and fake data. The QGAN model is based on the patch approach [66], which employs a number of quantum generators, each of which is in charge of creating a small patch of the final output, as shown in figure 2.18. The generator is a quantum variational circuit comprising alternating layers of single qubit rotation gates and two-qubit entanglement gates. Additionally, a first layer of Hadamard gates provides a superposition of all computational basis states with equal weight. In the context of training, the generator and discriminator are alternately optimised to create a probability that closely resembles the target distribution, while the discriminator input contains continuous scalar real data and discrete integral fake data.

2.5.2.3 Quantum Clustering

Clustering is the process of grouping similar data points together based on their characteristics or proximity. Traditional clustering methods, such as k-means, partition data into clusters based on classical distance measures. In quantum clustering, the data points are represented as quantum states instead of using classical distance mea-

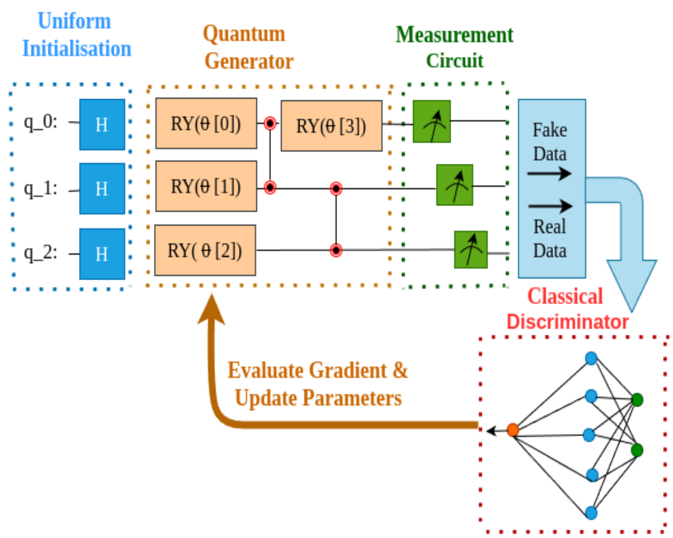

Figure 2.18: Schematic diagram of QGAN model [67].

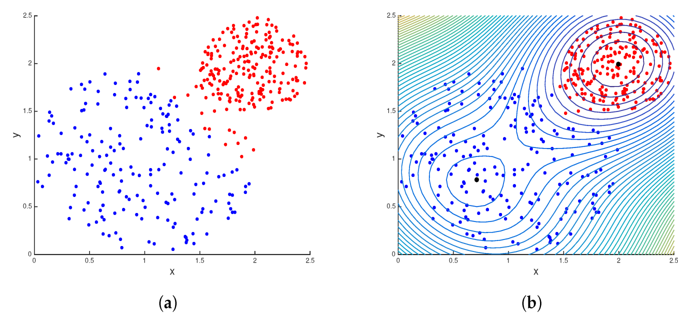

Figure 2.19: K mean clustering vs quantum clustering [68].

sures, and these quantum states can be created using techniques such as quantum superposition and quantum entanglement. By encoding the data into quantum states, quantum clustering allows for the exploration of complex relationships and interactions among data points. The quantum clustering algorithm operates by applying quantum gates and measurements to the quantum states representing the data. These operations can manipulate the quantum states to reveal underlying patterns and structures in the data, the goal being to find clusters that have close intra-cluster similarity and low inter-cluster similarity. figure 2.19 indicates the key advantages of quantum clustering by handling the high-dimensional and complex data more effectively than classical clustering methods. Quantum systems can simultaneously process multiple states and explore a much larger search space, potentially leading to more accurate clustering results.

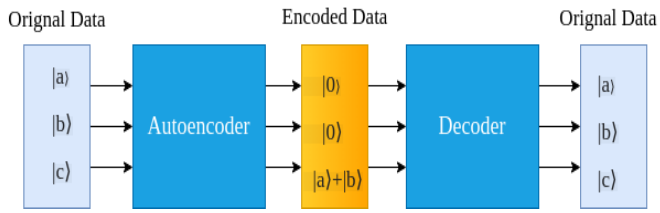

Figure 2.20: Structural diagram of quantum autoencoder [68].

2.5.2.4 Quantum Autoencoder

A quantum autoencoder is a quantum machine learning algorithm that leverages the principles of quantum mechanics in order to perform data compression and feature extraction. A quantum autoencoder follows a similar principle, but instead of using classical bits, operates on quantum bits or qubits. Qubits can exist in superpositions of 0 and 1, allowing for more complex representations and computations. Quantum autoencoders are typically implemented using quantum circuits and quantum gates, and their architecture consists of two main components: the quantum encoder and the quantum decoder, as shown in figure 2.20. The quantum encoder maps the input data, represented as quantum states, into a lower-dimensional code. This is achieved by applying a series of quantum gates and transformations to the input qubits, with the resulting code qubits capturing the essential features of the input data. Once the input data is encoded, the quantum decoder then performs the reverse process, transforming the code qubits back into an output state that approximates the original input data. The aim of the decoder is to reconstruct the data with minimal error, while the reconstruction process involves applying a series of quantum gates and operations that are the reverse of those used in the encoder. Quantum autoencoders have the potential to offer advantages over classical autoencoders in certain scenarios, while quantum systems can capture complex relationships and correlations that may be challenging for classical systems. Additionally, quantum autoencoders can exploit the properties of quantum entanglement and superposition in order to represent and process information in more powerful ways.

2.5.3 Semi-Supervised Models

Semi-supervised learning is a type of machine learning that builds models using both an enormous quantity of unlabeled data and a small amount of labeled data.

2.5.3.1 Variational Quantum Classifier (VQC)

VQC is a semi-supervised QML approach that enables the use of NISQ devices to acquire experimental results without the use of extra error-correction methods. This technique is a hybrid approach whereby the parameters are updated and optimized in a

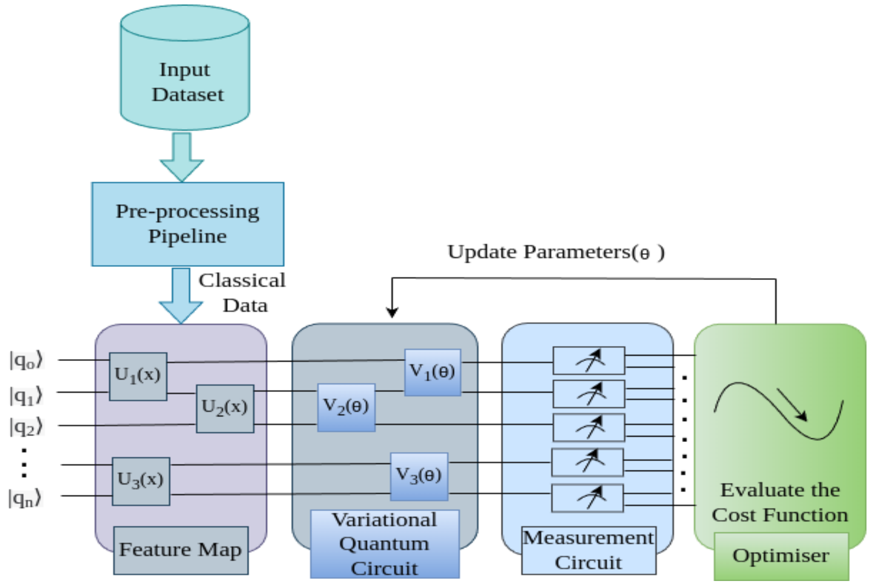

Figure 2.21: Block diagram of VQC.

traditional computer, allowing for optimization without improving the coherence times required. The device's iterative measurements are used to calculate the cost function, which is based on a system aimed at reducing errors by integrating noisy measurements in optimization calculations [69]. The amplitude encoding, which maps features to quantum states, is an appropriate solution for data pre-processing when utilizing the VQC model, as demonstrated in [70] [71]. The feature map is one of the major components which transform data into a quantum system's potentially higher-dimensional Hilbert space, making it possible to efficiently compute across non-linear fundamental functions on the feature space. As shown in figure 2.21, a VQC model contains 3 main components: a feature map that converts classical data into quantum states; a variational circuit with layers of short depth unitary circuits and θ-parameters that is iteratively tuned and has been trained by minimizing a cost function in a classical device; and at the end of the VQC model, a measurement circuit is used which returns the quantum variable decoded into classical output.

2.5.3.2 Quantum Least Square SVM (QLS-SVM)

A QLS-SVM is a variant of the traditional SVM algorithm. The aim of the QLS-SVM is to find a linear or non-linear hyperplane that best separates classes in a dataset by minimizing the squared error rather than maximizing the margin as in the standard SVM. Like classical SVM, QLS-SVM operates in a high-dimensional feature space. The input data consists of labeled examples, whereby each example is represented by a feature vector and belongs to a specific class, and uses a kernel function to implicitly map the input data into a high-dimensional feature space. This transformation allows the algorithm to find non-linear decision boundaries in the original input space, formulating the classification problem as a constrained optimization problem. The goal is to minimize the sum of squared errors between predicted outputs and desired outputs, subject to some constraints. The desired outputs are typically represented as +1 or -1 in the case of binary classification tasks. To solve the optimization problem, QLS-SVM employs a technique known as Lagrange multipliers - by introducing these, the problem is transformed into a dual problem that can be solved more efficiently. The solution involves finding the support vectors, which will be the data points that lie closest to the decision boundary. Once the QLS-SVM model is trained, it can be used to make predictions on new, unseen data, with prediction being based on the sign of the learned function, which determines the class label. If the output is positive, the sample is classified as one class, and if negative, it belongs to the other class.

2.6 QML Applications

Healthcare industry has made significant advancements with QC interms of data management, clinical studies, disease diagnosing, EHR, and medical device inspections. Moreover, QML is widely been using in various healthcare applications,i.e, molecular simulation, medical precision, radiotherapy, drug development, clinical trials and diagnose assistant as shown in figure 2.22. It also offers a significant increase in processing capacity, which resulted in improvements in the healthcare industry. Personalized treatment is provided using deoxyribonucleic acid (DNA) [72] sequencing with QC, and advanced therapies and medications are created using systematic modeling. QC addresses complex optimization challenges, such as devising effective radiation plans to eradicate specific cancer cells while minimizing harm to healthy organs and body parts [73] [74]. Qubit processing allows the quick sequencing and analysis of genomes, and cloud-based hospital infrastructure migration makes it possible to forecast chronic disease issues and protect medical data. Integrating QC into healthcare systems brings a possible advantages, including improving patient management, and medical professional experiences, reduced costs, and superior patient treatment outcomes [75] [76].

2.6.1 Diagnosis Assistance

QML algorithms are utilized to provide early-stage disease detection, leading to reduced healthcare costs and improved diagnosis and treatment. For example, early detection of cancer and Covid-19 can significantly decrease treatment expenses. While diagnosis tools such as X-rays, MRIs, and CT scans are expensive computer-aided devices that are evolving rapidly [77], and still facing challenges related to noise, quality, replicability, and safety. In this regard, QC aids diagnosis by examining medical images through edge detection and enhancing the diagnosis more quickly. Moreover, categorizing cells based on biochemical and physical characteristics requires ample space due to the abundance of predictor variables [78]. These challenges can be addressed by

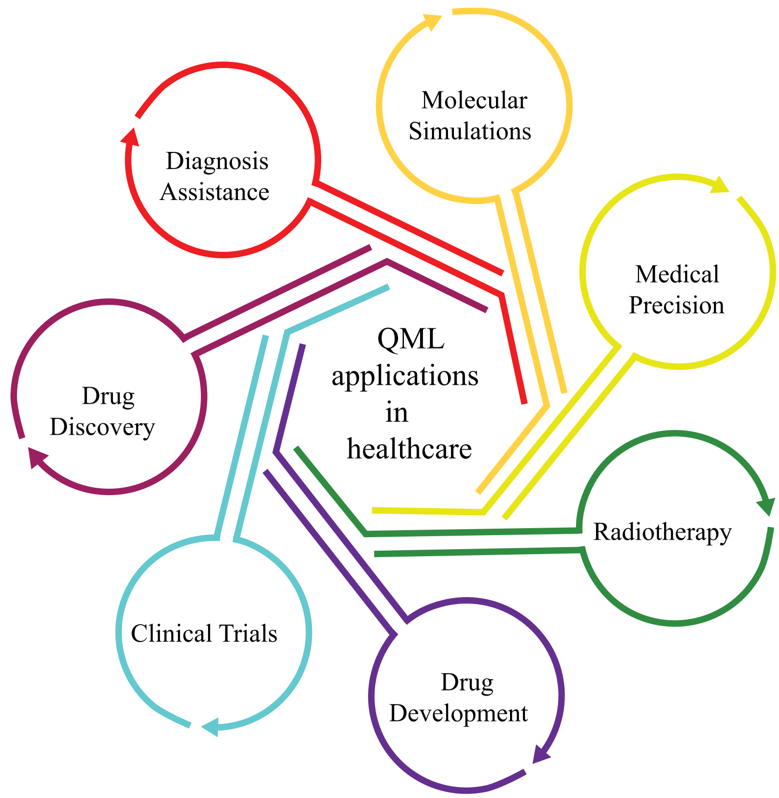

Figure 2.22: QML applications in the healthcare industry.

leveraging quantum-enhanced ML techniques such as quantum vector space, which in turn enhances single-cell diagnosis. The integration of QC enables repetitive diagnosis and treatment to be avoid and allows for regular monitoring of individual health.

2.6.2 Molecular Simulations

Quantum computers offer a fundamentally different approach to data processing compared to classical computing, which relies integrated circuits for speed. Quantum computers utilize qubits and leverage quantum entanglement, allowing for the development of quantum algorithms that exploit quantum phenomena. In the healthcare industry, quantum computers can be utilized to enhance ML techniques, optimization, AI techniques for complex simulations. This is particularly valuable in modeling complex correlations and dependencies among highly interconnected elements, such as molecular structures where multiple electrons may interact. QML provides an efficient means of analyzing healthcare processes. Additionally, QC enables complex simulations to be undertaken that would otherwise be challenging for classical algorithms due to scaling limitations. As the size of the problem increases, classical algorithms often face exponential increases in resource requirement, while QC offers a potential solution to managing these challenges effectively.

2.6.3 Medical Precision

The aim of medical precision is to provide personalized treatment to individual patients based on their specific diseases. Consequently, in the coming years, patient-focused medical care will be crucial in dealing with the intricate biological systems of human beings. However, medical expenses constitute only 10%–20% of healthcare costs, with the remaining 80%–90% associated with various other factors such as socio-economic conditions, environmental influences, and challenges related to health behavior. Additionally, conventional drug-based therapies may not yield effective results in the case of certain individuals and could even lead to fatal drug reactions. Therefore, early intervention and preventive measures can improve healthcare goals and reduce expenses through QML and QC techniques. Traditional approaches are reasonably effective in predicting future risks of disease while rendering noise, low data quality, small size, and high complexity. Therefore, quantum-based ML techniques enhance the accuracy of early disease detection. Healthcare practitioners can leverage medical devices to facilitate disease identification and manage risks through continuous monitoring and providing appropriate treatment. As a result, QML can enhance patient-centered treatment by continuously streaming data from medical devices, enabling uninterrupted services to patients to be provided.

2.6.4 Radiotherapy

Radiotherapy is a cancer treatment method that utilities electromagnetic energy to destroy cancer cells and prevent their growth. However, it is a delicate technique that is essential for meticulous calculations in order to accurately target the disease without causing harm to healthy organs. The administering of radiotherapy involves the use of highly precise devices that require complex optimization problems to be solved using a high level of precision. Thus, ensuring precise radiographic procedures entails conducting multiple accurate and sophisticated simulations in order to find an effective solution. The adoption of QC enables various simulation recommendations to be implemented, allowing for the concurrent execution of numerous simulations and facilitating enabling an effective solution to be determined faster.

2.6.5 Drug Development

Medical professionals can now model intricate chemical interactions at an atomic level thanks to QC, which is essential for medical research, i.e, disease diagnosing, and treatment. The ability to encode proteins in the human genome and simulate their interactions with current medication has been made possible by developments in QC, and applying AI methods to assist in patient diagnosis is becoming more and more popular. The vast majority of ML techniques currently in use are pattern recognition techniques, where various ML models are trained utilizing a large amount of patient data to create computer-assisted diagnosis systems. QC significantly enhances the processing of information far more efficiently than conventional computing methods. In such a scenario, the objective may involve leveraging the aforementioned comparisons in order to facilitate accurate diagnosis.

2.6.6 Clinical Trials

QC and QML can be used to revolutionize clinical trials in several ways, i.e, quantum computers can process vast amounts of clinical data much faster, which allows more complex and comprehensive data analysis, identifying patterns and correlations. Furthermore, QML can optimize clinical trial designs by accurately predicting patient outcomes and treatment efficacy. This leads to more efficient trials with a higher chance of success. QC can enhance medical imaging techniques, providing clearer and more detailed images, which are crucial for accurate diagnosis and treatment planning. QML can process real-time data from clinical trials, allowing for immediate adjustments and interventions, thereby increasing the safety and effectiveness of trials. QC facilitates the sharing and analysis of clinical data across different geographical locations, promoting global collaboration in medical research and trials.

2.6.7 Drug Discovery

Novel opportunities for medication discovery are presented by QML and QC. They facilitate accurate drug-biotarget interaction modeling by emulating molecular interactions at the quantum level. Compared to conventional approaches, this sophisticated modeling aids in the more precise prediction of pharmacological features including toxicity and efficacy. The identification of interesting chemicals, the optimization of medication formulations, and the customization of treatments to individual genetic profiles can all be greatly accelerated by quantum algorithms. The pharmaceutical business could undergo a transformation thanks to this technology, which offers the promise of personalized therapy quicker, and more affordable in medical development.

2.7 The Emergence of QML in Healthcare

The promise of QML in healthcare is becoming more and more apparent. As healthcare organizations work to enhance patient outcomes and reduce expenses, QML offers great potential. QC is used by QML, a type of artificial intelligence (AI), to analyse and predict enormous amounts of data. It relies on the concepts of quantum mechanics, which perform considerably faster data manipulation than is possible with conventional computing. Considering the complexity and size of the data in healthcare applications, QML is especially well suited to such applications. Identifying patterns in data that might normally be too challenging can be found easily using QML - for instance, it can be used to identify possible drug targets, find early disease symptoms, and predict how well treatments would work. Additionally, it can be utilized to examine patient records and provide individualized therapies.

2.7.1 Medical Healthcare Record (MHR)

Medical Healthcare Records (MHRs) play a vital role in the field of healthcare, providing a comprehensive and digitized collection of patients' medical information. With the advent of cutting-edge technologies like QML models, the importance of MHRs has become even more pronounced in predicting and improving patient outcomes. QML models have the potential to revolutionize healthcare by offering enhanced predictive capabilities and personalized treatment strategies. MHRs act as a repository for patient data, containing a variety of data including medical history, diagnosis, medication, lab results, imaging studies, and other disease-related parameters. Obtaining longitudinal patient data is one of the main benefits of using MHRs for QML model predictions, and MHRs offer a comprehensive overview of a patient's medical history throughout time, allowing physicians to track the occurrence of diseases and spot patterns that might otherwise be difficult to pinpoint. Furthermore, MHRs also include a significant amount of both structured and unstructured data. Unstructured data consists of clinical notes, medical imaging reports, and pathology reports, whereas structured data consists of information such as demographics, vital signs, and lab values.

Recently, there has been growing interest in utilizing QML models for various healthcare tasks. These models, such as QSVM, QRF, QNN, QCNN, VQC, QKNN, and Quantum Decision Trees (QDT), have the potential to reform the healthcare sector by offering enhanced predictive capabilities and personalized treatment strategies. In [79], the authors the two QML model,i.e, QSVC and VQC for chronic disease prediction. The models are compared with the previous published studies and claim the accuracy of 82%. For classifying heart disease, an ensemble learning techniques by using QNN and QSVM model has been presented in [80]. The cleveland heart disease dataset has used and obtained a good results for bothe models. Moreover, to classify cardio disease [81] an optimized QSVM and Hybrid Quantum Multi Layer Perception (HQMLP) models were combined using a privet MHR cardio dataset, while feature dimensions were reduced using the wrapper and filter method. The proposed models are compared to their classical version and their competency ensured, and the distinction between classical and quantum ML shown in [82]. To classify mellitus diabetes, the classical SVM and DT models are compared to the Qboost classifier by taking the kernel PCA method into consideration. The quantum model improves accuracy by 10%-15% by correctly observing the pattern for the desired dataset. In QML, ensemble learning is a popular ML approach that combines several QML models in order to make more precise predictions, while in [83], three QML models - QSVC, QNN and VQC - are combined to undertake ensemble learning in order to classify heart disease. The

UCI Cleveland dataset is used and the results are compared to the classical ensemble model, with results demonstrating that the bagging ensemble model is effective in improving quantum classifiers with a view to ensuring accurate prediction. Breast cancer, which is discussed in [84], is one of the leading causes of death in women worldwide. A VQC model with amplitude encoding along with an EfficientSU2 circuit made up of a single qubit layer using CX entanglements was considered for such purpose. The publicly available breast cancer dataset with 2 feature and 4 feature datasets obtains the greatest accuracy - 95% and 94% - in the case of the quantum VQC model. Moreover, T. Shahwar *et.al* [85] Present a hybrid classical–quantum machine learning by considering a hybrid transfer learning and VQC model for Alzheimer detection. The hybrid technique proves fruitful in QML for tackling some classification problems, while the QML framework with classical DL was described in [86], where a PIMA Indian diabetes dataset was used for the VQC model, and pre-processing and exploratory-data analysis was explored and found essential for the purpose of robust prediction. Another hybrid QNN and hybrid QRF for early heart disease detection using the Cleveland and Statlog heart disease dataset was also described in [87], with the models having been evaluated by adopting 10-fold cross validation with different qubits and different layers, and where the results show the robustness of the models. A systematic investigation of empirical quantum advantage has been introduced in [88], where the biological therapies dataset has been used for classification.

Moreover, a quantum-inspired approach with ensemble learning was modeled in [89], where the Pima Indian dataset was considered for decision-making purposes. The quantum encoding contains object encoding and an RF classifier is used for ensemble learning, which results in accurate prediction. A VQC model for heart disease and breast cancer prediction which takes different encoding techniques into consideration is also described in [238]. During pre-processing a quantum random access coding (QRAC) was used to map the discrete future, which outperforms in terms of accuracy compared to the ZZ-FeatureMaping technique. Another contribution using amplitude encoded VQC and QSVM models for diabetes classification was modeled in [290], whereby the feature dimension was reduced in pre-processing by taking the intersection of RF, LR and SVM outputs. [90] demonstrated the importance of the preprocessing technique for the quantum model and claimed that it reduces model complexity. The quantum version of SVM was used after reducing the dimension of the breast cancer dataset, which improves prediction accuracy. For its part, the practical implementation of a kernel based QSVM algorithm to tackle a classification problem using breast cancer Wisconsin datasets is shown in [91], which demonstrates that quantum-driven ML could deliver a quantum speedup in solving many challenging tasks. The results claimed that the quantum computer is much faster than the classical computer. The hybrid feature selection approach was used in [92], taking quantum Oracle operation, amplitude estimation and amplitude amplification into consideration for the breast cancer dataset, with results indicating the strong performance and major searchability, with a high level of efficiency and less complexity than the classical method. Most QML models are based on use of the classical model, as discussed in [93], where various models are used to create a new voting model results in ensemble learning. Furthermore, the diabetes dataset contains the MHR records of various patients used to evaluate the model; in comparison to classical models, the new quantum model is 55 times computationally faster. The importance of state preparation for the VQC model is also researched by the authors [94], and the diabetes dataset was mapped using stock parameters to solve the Noisy Intermediate Scale Quantum device (NISQ) classification problem. The Poincaré sphere representation is used with two experiments for 2 qubit and 3 qubit, obtaining the greatest accuracy of 70% and 72% respectively, while the classical quantum hybrid algorithm for supervised binary classification task was modeled in [95], where a VQC model was used to classify the UCI ML breast cancer dataset. The highest testing accuracy of 2 and 3 qubits with 2000 shots was recorded as 91% and 73% respectively. The VQC is now garnering a lot of interest, since it can be used with upcoming quantum computers - the framework, whereby the classical ML is merged with the VQC model for classification of MNIST breast cancer dataset, is described in [96], and in the case of back propagation, various optimizers were used to reduce the loss. The healthcare sector has been further revolutionized using unsupervised models for disease classification. Additionally, three quantum distances prototypes-based clustering models are analyses and compared in [97], where K-means outperform other clustering approaches. The work concluded that the quantum version K-means models take logarithmic time while the classical one takes polynomial time complexity. A quantum-inspired neural network based on fuzzy logic was also published in [98]; Fuzzy C-Means (FCM) clustering is used in learning, and neurons are added to the hidden layer in order to constructively develop NN architecture. Lastly, various healthcare datasets - i.e, breast cancer, diabetes, liver disorder, and heart disease - are taken into consideration, and an overview of medical healthcare records is shown in table 2.3

2.7.2 Medical Imaging Data

QC provides exponentially higher computational power, which enables large-scale medical imaging datasets to be processed and analyses with increased speed and efficiency. Complex algorithms for tasks such as image reconstruction, feature extraction, and classification can be executed more quickly, allowing for faster and more accurate diagnoses. For effective processing and analyzing of imaging data, quantum algorithms can enhance the accuracy of disease detection, image segmentation, and classification, which can lead to more precise diagnoses and personalized treatment plans. Furthermore, feature selection is a critical step in medical image analysis. Quantum machine learning algorithms can efficiently explore vast feature spaces and identify the most relevant and informative features for the purposes of diagnosis, which leads to im-

Table 2.3: Overview of medical history record articles.

Year Published	Author Citation	Methodology Used	Dataset Used	Task Targeted	Result	Country
2024	M. Munshi et al. [79]	Consider the two QML models, QSVC and VQC, use a data pre-processing and dimensionality reduction techniques.	Health Failure dataset	Classification	Obtained 83% accuracy for QSCV and 79% for VQC model.	India
2024	H. Ghazi Enad et al. [80]	Consider the two QML models, QNN and QSVM, use a data cleaning, transformation and normalization for the input data.	Cleveland HD dataset	Classification	Obtained 77% accuracy for QNN and 85% for QSVM model.	Iraq
2023	D.Maheshwari et al. [81]	Wrapper and filter technique used for pre-processing. The wrapper technique consists of RFE and logistic regression, while mRMR, an optimized QSVM and a hybrid QML are taken into consideration in the case of the filter technique.	Ischemic health disease dataset	Classification	Obtained 94% accuracy in the case of the IHD dataset with OQSVM and 93% in that of the HQMLP model.	Spain
2023	B.Prakash et al. [82]	Ensemble method used, Qboost classifier, XGboost, quantum model compared to classical Kernel PCA, DT, SVM and Bayesian network.	Diabetic mellitus dataset	Classification and regression	Obtained 73.35% accuracy for Qboost, 68.34% for Kernel PCA-SVM, 69.9% for Bayesian network and 57.27% for DT model.	India
2023	G.Abdulsalam et al. [83]	Three QML models, QSVC, QNN and VQC combined to undertake ensemble learning to classify heart disease; an ensemble model is formed	UCI heart disease dataset	Binary Classification	Obtained maximum accuracy of 88.52% for QSVC, 86.84% for QNN, and 86.89% for VQC respectively.	Saudi Arabia
2023	S.Diaz-Santos et al. [84]	A VQC model with amplitude encoding is used; an EfficientSU2 circuit made up of a single qubit layer and CX entanglements used	Breast cancer dataset	Classification	Obtained 95% accuracy for 2 VQC features and 94% for 4 features respectively.	Spain
2022	T.Shahwar et al. [85]	Present a hybrid classical–quantum machine learning model, Hybrid classical–quantum transfer learning, used a VQC model	Alzheimer dataset	Binary Detection	Obtained training accuracy 99.1% and classification accuracy 97.2% .	Pakistan
2022	H.Gupta et al. [86]	A hybrid quantum classical technique with VQC model is taken into consideration; xploratory data analysis explored	PIMA Indian diabetes dataset	Binary classification	Obtained 95% and 74% accuracy for the proposed model.	India
2022	H.Heidar et al. [87]	Hybrid quantum NN (HQNN) and hybrid quantum RF (HQRF) models used.	Cleveland and Statlog HD dataset	Detection	For Cleveland dataset AUC 96.6% accuracy obtained, for HQNN model 94.3%, and for HQRF and statlog model 97.7% and 90.5 respectively%.	Iran
2022	K. Zoran et al. [88]	Used an empirical quantum advantages, used classical SVM with function kernel and quantum models with custom kernels using an IBM, QSVM model	biological therapies dataset	Classification	Clained framework represents progress toward a prior identification of data sets.	Switzerland
2021	MS.Ishwarya et al. [89]	Quantum inspired approach with ensemble learning used.and also object encoding with RF.	PIMA Indian diabetes dataset	Classification	Obtained testing accuracy of 90.5% for proposed model.	India
2021	H.Yano et al. [238]	QRAC is used to reduce the number of qubits for pre-processing; k-fold cross validation with VQC model and ZZ-featureMaping is taken into consideration.	Heart disease dataset	Classification	Obtained testing accuracy of 85.1%, 83.3% and 84.2% respectively in the case of various pre-processing techniques.	Japan
2021	D.Maheshwari et al. [290]	Important features selected using the intersection of RF, LR and SVM; a VQC model with amplitude encoding and QSVM Model used.	Diabetes dataset	Classification	Obtained maximum accuracy of 68.7% for VQC and 74.1% for QSVM model.	Spain
2021	D.Pomarico et al. [90]	10-fold cross validation used for pre-processing; a QSVM model used to reduce complexity	Breast cancer dataset	Classification and detection	Obtained maximum accuracy 69.5% for VQC and 69.2% for QSVM model.	Italy
2020	S.Saini et al. [91]	A QSVM and a kernel-based QSVM algorithm used; a quantum simulator and realtime quantum process used.	Breast cancer Wisconsin dataset	Classification	Obtained maximum accuracy of 85% for KQSVM and 80% for QSVM model.	India
2020	S.Chakraborty et al. [92]	Hybrid feature selection technique used for quantum oracle operation, amplitude amplification, and amplitude estimation.	Breast cancer dataset	Classification	Obtained maximum accuracy and F1-score of 95%, Recall 94% and precision score 95.3% for HQFSA model.	India
2020	D.Maheshwari et al. [93]	Ensemble learning using various models DT, RF, Adaboost, Qboost voting model 1 and voting model 2.	Type 2 diabetes mellitus dataset	Classification	Obtained maximum accuracy. Recall and specificity 69% .	Spain
2020	D.Sierra-Sosa et al. [94]	The importance of state preparation for VQC model researched; stock parameters used to solve the NISQ device problem	Diabetes dataset	Classification	Obtained maximum accuracy of 70% and 72% for 2 qubit and 3 qubit respectively.	USA
2019	S.Mardirosian et al. [95]	Classical quantum hybrid algorithm used; VQC model with 2 qubit and 3 qubit is taken into consideration.	UCI breast cancer dataset	Classification	Obtained maximum accuracy of 91% and 73% for 2 qubit and 3 qubit respectively.	Netherlands
2019	ZY.Chen et al. [96]	VQC model used; various optimisers used to reduce loss for back propagation.	MNIST breast cancer dataset	Classification	Obtained maximum accuracy 87% with -1 loss.	China
2019	K.Benlamine et al. [97]	Three quantum distance prototype-based clustering models analysed, where K-means outperforms than other clustering approaches.	UCI breast cancer dataset	Classification		France
2019	OM.Parkash-Patel et al. [98]	A quantum inspired neural network based on fuzzy logic is taken into consideration; Fuzzy C-Means (FCM) clustering also used for learning.	UCI breast cancer, diabetes, heart dataset	Classification	Obtained maximum accuracy of 99.85%, 98.1% and 94.44%.	France

proved accuracy by reducing noise, irrelevant data, and the risk of overfitting. Image reconstruction techniques, such as CT scan or MRI, can benefit from quantum machine learning algorithms, while quantum-enhanced optimization algorithms can improve the quality of reconstructed images by solving complex inverse problems more effectively. It can also enhance image segmentation - the process involving delineating different regions or structures within medical images as discuss in [99].The study focus on the use of quantum kernel and the implementation of hybrid quantum-classical neural network. Th covid-19 lungs image data has been considered and compared the results with classical SVM. Quantum algorithms can optimize the segmentation algorithms, leading to more accurate and efficient delineation of organs, tumors, and other pathological regions. A Hybrid Quantum-Classical Convolutional Neural Network (HQC-CNN) has been used in [100] for brain MRI images segmentation and classification. For image segmentation it consider bath wise Q-Means clustring and sued Max-cut approach with adiabatic quantum computation for tumor region extraction. In such a scenario, neuro-degenerative disease may be diagnosed as in [101], where the classical AlexNet is used with a VQC model to form a hybrid by taking two MRI datasets from PPMI and ADNI into consideration. The results indicate that the classical model helps the quantum model by securing the greatest accuracy of 97% and 96% respectively. Another improved hybrid model including additional explainable strategy is described in [102]. To diagnose abnormal activities in breast cancer and knees, a local interpretable model explanation is merged with quantum K-means algorithm using MRI images, and this improved model performed well in terms of accuracy with 92.6% and 93.7% in the case of breast cancer and knee MRI datasets. A 3D self supervised quantum inspired NN for diagnosing medical data is introduced in [103], whereby MR brain images and liver tumor data is used, and a quantum fuzzy logic processes the information of low level and high level features of the local image in order to form accurate segmentation of the 3D medical data and obtain the greatest accuracy of 99% and 98.9% respectively. For cervical cancer classification, further quantum inspired weed optimization with DL is described in [104]. A Gabor filtering is used during preprocessing, whereby features are extracted using a DCNN based on the SqueezeNet approach, and a deep variational autoencoder is used and maximum accuracy of 99.07 obtained%. Furthermore, research into feasibility utilizing QML on a real healthcare dataset is described by S.Moradi et al. [105]. Using IBM hardware the two QML algorithms - i.e, quantum distance classifier (qDS) and simplified quantum kernel SVM (sqkSVM) - were proposed, and the models were tested using the breast cancer dataset, bone narrow transplant dataset and heart failure dataset with maximum accuracy of 91% and 87%. For the classification and prediction of COVID-19, [106] provides the hybrid quantum-classical convolutional neural network (HQ-CNN) model, whereby the main model was divided into two categories: quantum and classical. The quantum section used a quantum convolutional layer, whereas the traditional network included two convolutional layers, three maximum pooling layers, and two fully connected layers.

In the case of the Covid-19 X-ray image dataset, extensive trials showed that current CNNs perform better. Another 2-qubit quantum CNN model referred to as Javeria was described in [107], where the encrypted brain tumor data was obtained using the SHA-256 algorithm. The model consists of two dense layers and a Keras layer with Softmax activation function. Four brain image datasets - BraTS2018 with 191 samples, BraTS2019 with 335 samples, BraTS2020 with 335 samples and BraTS2021 with 1251 samples were taken into consideration and maximum accuracy obtained of 98%, recall 99%, F1-score 98%, and precision 99%. The article [108] describes a quantum-inspired deep probabilistic learning ordinal regression model for diagnosing medical images that makes use of the representational strength of deep learning and the inherent ordinal information of disease. Two distinct medical image analysis tasks are used to gauge the model's performance. Using eye fundus images, prostate cancer is diagnosed and the degree of diabetic retinopathy estimated, with the model rendering promising results. A quantum clustering technique is described in [109] to identify prostate cancer using an MR image brain tumor dataset. For classification purposes, a prostate radiation protocol with 100 to 170 slice images per patient, minimum pixel sizes between 256×256, a resolution plane of less than 1 mm, and slice thickness ranging from 1 mm to 2 mm were taken into consideration. [110] used conventional methods to undertake image processing and feature extraction operations on the data, and then train a VQC model with a 4-qubit quantum processor in order to recognize CT images of COVID-19-healthy and infected patients. The results demonstrated that the quantum computer provided a competitive edge in COVID-19 images with classification accuracy of 90.9% - 97.7%. In order to expand the size of the dataset and improve accuracy, J.Amin et al. [210] first used a Conditional-GAN (CGAN) to create CT images, whereby two models - CML and QNN were described for classification of Covid-19, with different layers, sets of parameters and activation functions. Both the models performed well and obtained maximum accuracy of 93% and 80% in the case of the UCSD-AI4H and POF hospital dataset. [111] used quantum circuits to train a classical neural network termed quantum-assisted NN (qNN), where QNN was executed in parallel by a quantum circuit and was discovered to be significantly faster. Furthermore, orthogonal weight matrices were developed to train quantum orthogonal networks (qOrthNN) in order to avoid gradient explosion and improve accuracy. Chest X-ray images were also used from PneumoniaMNIST and RetinaMNIST and maximum accuracy of 86% and 78% obtained respectively. For its part, a locally orderless tensor network model (LoTeNet) with fewer computational resources using few model hyper parameters and GPU memory is described by R. Selven in [112], with accuracy of the experimental results for binary classification experiments on the PCam and LIDC datasets of 94.3% and 87.4% being obtained, respectively. [113] developed a quantum based on variational algorithms in order to categorise data by employing quantum feature mapping with a significantly smaller number of training parameters. The breast cancer datasets were used to train the classifier, and the results (93.7%)

showed that it performed better in binary and multi classification than the conventional neural network model, whether using linear or non-linear separable data. QNN implementation is the best solution for the purpose of identifying some specific diseases. [114] demonstrated that there are two distinct components of QML: 1) adding quantum data to neural networks; and 2) utilising hybrid MIA technology to complete information about disease identification. For classification and detection of medical CT Scan images maximum accuracy of 94.3% and 87% was obtained, respectively. In order to conduct quantum feature selection, [115] suggested a hybrid quantum feature selection algorithm (HQFSA) that made use of graph theory. The suitability for dimensionality reduction was ascertained, and then an important set of features extracted using the Grover algorithm and amplitude amplification method, using the UCI breast cancer dataset and with the greatest accuracy of 96% being obtained. [116] used BQ-CNN, a quantum particle swarm optimisation technique that evolves CNN structure based on binary coding for image classification tasks. On the MNIST images dataset, experiments showed that the technique improved performance with an accuracy of 96% and 85% in the case of the CS and MDRBI dataset. The quantum SVM algorithm, which is claimed to offer exponential speedup for least squares SVM (LS-SVM), was described with a view to addressing the big-data challenge [117], and the classification task was accelerated exponentially, demonstrating the algorithm's feasibility. The proposed model was evaluated by taking two privet datasets into consideration with an accuracy of 84.9% and 91.4% respectively. Practical applications for NISQ computers are also possible thanks to hybrid quantum-classical algorithms. Additionally, the data drive quantum circuit learning (DDQCL) approach is provided in [118] and can be used to train shallow circuits for generative applications as well as help characterize the quantum devices. The three Synthetic datasets with 1000 data points sample has been taken into consideration and get good results other than architectural circuit design obtained, while a new quantum-based autonomous perception model (APM) was developed by A. Sagheer et al. [119] to address categorization issues and boost learning effectiveness. Furthermore, for pattern classification purposes, research was conducted into synthetic and breast cancer datasets, which demonstrated benefits in terms of computing speed and maximum accuracy of 99% and 98% respectively, and described in 2.4.

2.7.3 Biomedical Signal Datasets

In the healthcare industry, biosignals play an important role in diagnosing some disorders. The term "biosignal" refers to electrical impulses generated by brain neurons, tissues and muscles, and monitored by biomedical sensors [121]. The brain's activity can be monitored by a computer with the use of a BioSignal interface, which comprises hardware and software [122]. The four main components of the BioSignal system are filter, control devices, amplifiers and sensors. These signals come from the body and are

Table 2.4: Overview of imaging articles in healthcare.

Year Published	Author Citation	Methodology Used	Used Dataset	Task Targeted	Result	Country
2024	S. Naguleswaran et al.. [99]	Quantum kernal method, hybrid quantum classical neural network, SVM.	Covid-19 lungs dataset	Classification	Obtained 85% accuracy in the case of HQCNN, and 80% for SVM.	Australia
2024	S. Kumar Roy et al.. [100]	Hybrid Quantum Classical Convolutional Neural Network (HQCCNN), Q-Means clustering and Max-cut approach with adiabatic quantum computation.	Brain tumor MRI dataset	Classification, Segmentation	Obtained 98.3% accuracy with precision 99%, recall 97.7% and F1-score 98.3.	India
2023	N.Alsharabi et al.. [101]	Classical AlexNet used with quantum VQC model to form a hybrid model in order to diagnose neuro-degenerative disease.	PPMI MRI, ADNI MRI dataset	Classification	Obtained 97% accuracy in the case of Parkinson's disease, and 96% in that of Alzheimer's disease.	Yemen
2023	S.Deshmukh et al. [102]	Improved hybrid model used including additional explainable strategy; a local interpretable model explanation is merged with quantum K-means algorithm.	Breast cancer MRI, Knee MRI dataset	Quantum clustering	Obtained 92.6% accuracy in the case of breast cancer, and 93.7% in that of Knee disease.	India
2023	D.Konar et al. [103]	3D self supervised quantum inspired NN is taken into consideration; a quantum Fuzzy logic technique used	Brain MR and liver tumor dataset	Segmentation	Maximum accuracy of 99% and 98.9% obtained for both datasets respectively.	India
2023	AK.Mishra et al. [104]	Gabor filtering is taken into consideration for pre-processing while feature extraction using DCNN is based on the SqueezeNet approach, with a deep variational autoencoder being used	Cervical cancer image dataset	Classification	Obtained 99.07% accuracy.	India
2022	S.Moradi et al. [105]	Two QML models used: a quantum distance classifier (qDS) and simplified quantum kernel SVM (sqkSVM),	Breast cancer and Heart failure dataset	Binary classification	Obtained maximum accuracy of 91% and 87% for both datasets.	Austria
2022	EH.Houssein et al. [106]	A hybrid quantum-classical convolutional neural network (HQ-CNN) model, quantum layer, two convolutional layers, three maximum pooling layers, and two fully connected layers used	Covid-19 X-ray dataset	Classification	Obtainedmaximum 98.6% accuracy, 99% and recall respectively in the case of the proposed model.	Egypt
2022	J.Amin et al. [107]	Proposed 2 qubit QCNN model, data encrypted and decrypted is held using the SHA-256 algorithm, with two dense layers, a Keras layer and Softmax activation function	BraTS2019, BraTS2020, BraTS2021 brain MRI	Semantic segmentation	Obtained greatest accuracy of 98%, recall 99%, F1-score 98%, and precision 99%.	Pakistan
2022	S.Toledo-Cortes et al. [108]	Quantum-inspired deep probabilistic learning ordinal regression model, using Deep Quantum Ordinal Regressor (DQOR)	Eye fundus, Prostate cancer images	Diagnosis	Obtained testing accuracy of 58.7%, MAE 69.5%	Colombia
2021	J.Reyes Bruno et al. [109]	Quantum clustering technique used, with prostate radiation protocol with 100 to 170 slice images per patient being taken into consideration.	Brain tumor MRI	Diagnosis	Obtained good results in terms of diagnosing brain tumor.	Colombia
2021	E.ACAR et al. [110]	Conventional method used for feature extraction, then a VQC model trained with a 4-qubit quantum processor, and quantum transfer learning used.	COVID-19 image dataset	Classification	Obtained testing accuracy of 90.9% and 97.7% respectively.	Turkey
2021	J.Amin et al. [210]	Conditional-GAN (CGAN) used to create CT images, then two models, CML and QNN, were described using different sets of parameters and activation functions.	Covid-19 CT Scan Images	Classification	Maximum accuracy of 93% and 80% obtained for two Covid-19 datasets.	Pakistan
2021	N.Mathur et al. [111]	Quantum circuits used to train a classical neural network termed quantum-assisted NN (qNN); orthogonal weight matrices were developed for training of quantum orthogonal networks (qOrthNN).	PneumoniaMNIST, RetinaMNIST Chest X-ray	Classification	Obtained maximum accuracy of 86% and 78% for both datasets respectively.	France
2020	R.Selven et al. [112]	Locally order-less tensor network model used (LoTeNet) to reduce computational resources with GPU memory	PCam and LIDC thoracic CT, MRI	Classification	Obtained maximum accuracy of 94.3% and 87.4% for both datasets.	Denmark
2020	S.Adhiksry et al. [113]	Quantum based on variational algorithm developed for data categorisation, using quantum FeatureMapping to reduce training parameters.	Breast cancer dataset	Binary and Multi Classification	Obtained maximum accuracy of 93.7%.	India
2020	V.Dutt et al. [114]	QNN model implemented, using hybrid MIA technology to complete information about disease identification.	Medical CT Scan Images	Classification and Detection	Obtained maximum accuracy of 94.3% and 87% respectively .	Spain
2020	S.Chakraborty et al. [115]	Hybird quantum feature selection algorihm (HQFSA) and quantum graph theory used; features extracted using the Grover algorithm and amplitude amplification method	UCI Breast cancer	Classification	Obtained maximum accuracy of 96%.	India
2019	LI.Yangyang et al. [116]	CNN and binary encoding strategy used; a quantum particle swarm optimisation technique is also taken into consideration.	CS and MDRBI Images	Classification	Performed well with 96% and 85% accuracy for CS and MDRBI datasets.	China
2019	C.Ding et al. [117]	Quantum SVM taken into consideration for big-data challenges using least squares SVM (LS-SVM) with matrix sampling.	Two privet medical image datasets	Classification	Obtained maximum accuracy of 84.9% and 91.4% respectively.	China
2019	M.Benedetti et al. [118]	Data drive quantum circuit learning (DDQCL) approach implemented; shallow quantum circuit used to reduce the number of training parameters.	Synthetic dataset	Classification	Model confidence of 95% obtained.	UK
2019	A.Sagheer et al. [119]	Quantum-based autonomous perceptron model (APM) developed; non-linear classification model.	Synthetic and UCI Breast cancer	Pattern recognition and Classification	Obtained maximum accuracy of 99% and 98% respectively.	Saudi Arabia

Table 2.5: Overview of biosignal articles in healthcare.

Year Published	Author Citation	Methodology Used	Dataset Used	Task Targeted	Result	Country
2024	M.Kaur Saggi et al. [130]	Develop a pipeline for data processing, important fetaures are selected by using RF, used QNN model.	GDC TCGA Dataset	Classification	Obtained accuracy of 96% with loss of o.13.	USA
2024	S. M Tripathi et al. [131]	Quantum AI-driven Heart Health Framework (QAIHHF), used QLSTM model, used classical SVM, DT,T=RF and XGBoost models.	MIT ECG Dataset	Classification	Obtained the highest accuracy of 81.5% for batch size 512	USA
2023	N.Baygin et al. [132]	LOSCO CV and K-fold CV are taken into consideration for pre-processing, using feature selection techniques based on quantum computing.	Mental Arithmetic EEG Dataset	Classification	Obtained accuracy of 93.40% and 97.88%, with geometric means of 88.44% and 96.42% respectively.	Turkey
2022	M.Sameer et al. [133]	Feature extraction using quantum algorithm; hybrid classical-quantum layers with 1D CNN model; model training hyper parameters reduced.	Bonn EEG Dataset	Binary Classification	Maximum accuracy and specificity obtained of 100%.	India
2022	T.Koike-Akino et al. [134]	Hybrid quantum classical QNN model implemented, using a VQC model with classical DNN model,	Stress, RSVP, MI, EEG, EMG, ECOG Dataset	Signal Processing	Obtained maximum accuracy of 87.23%, 95.12%, and 60.22%, respectively.	USA
2022	A.Padha et al. [135]	Hybrid quantum model used, with learning capacity increased using a parameterised quantum circuit and a classical LSTM Model taken into consideration.	SWELL-KW Stress EEG Dataset	Classification	Obtained maximum accuracy of 87.67%.	India
2022	S.Sridevi et al. [136]	Hybrid quantum classical model taken into consideration, using a quanvolutional neural network and 2D scalogram technique, with desecrate wavelet transformation	MIT-BIH arrhythmia ECG Dataset	Classification	Accuracy and operating curve score of a publicly accessible physio net MITBIH arrhythmia database recorded as 98% and 100% respectively.	India
2021	R.Kumar Nath et al. [137]	Quantum annealing is taken into consideration for important four feature selection; a Pearson correlation coefficient among the attribute variable and target variable used	Respiration, ECG, foot EDA, hand EDA Dataset	Classification	Quantum annealing obtained more promising results than with the classical technique in optimising training phase.	USA
2020	S.Aishwarya et al. [138]	Various QML models taken into consideration, using the VQC model, quantum annealing, hybird quantum classical NN model	Cognitive stae of mind EEG Dataset	State mind Prediction	Obtained greatest validation accuracy of 61.53%.	India
2020	Li.YaoChong et al. [139]	A hierarchic quantum mechanics-based architecture taken into consideration for feature selection, using a random nonlinear kernel from the modified QSVM model.	EEG signal Dataset	Classification	Obtained maximum accuracy of 95.14%	China
2019	S.MR Taha et al. [140]	Hybrid method and auto regressive model used; quantum recurrent neural network (QRNN), with model proposed compared to QNN and Quantum Wavelet NN.	Biomedical EEG signal Dataset	Classification	Obtained maximum accuracy of 88.28% with the fastest processing time of 6 seconds.	Iraq

enabled by the machine and encoded, decoded, and processed by body interfaces. Neurons in the human brain produce signals as a result of both reflexive and voluntarily performed actions [123]. Biosignal accretion methods are further classified as invasive and non-invasive. Invasive procedures such as electrocorticography (ECOG) [124] involve the insertion of sensors into the human body, while non-invasive data which does not involve the skin breaking involves an electrocardiogram (ECG) [125], magnetoencephalography (MEG) [126], electroencephalograms (EEG), and Functional Near-Infrared Spectroscopy (fNIRS) [127]. ECG examines the electrical activity of the heart and diagnoses cardiovascular disease, and EEG is used to identify disorders of the brain, such as Alzheimer [128], while EOG is used to monitors the cornea-retinal of the front and rear of the human eye and conducts ophthalmological and eye movement diagnoses. Furthermore, electromyograms (EMG) [129] assess the electrical activity of prosthetic function as well as the ability of skeletal muscles. For their part, quantum sensors are devices that exploit quantum effects in order to achieve enhanced sensitivity and precision in measuring physical quantities. Integrating quantum sensors with ECG or EEG monitoring systems could enable more accurate and detailed data to be gathered. The high sensitivity and noise reduction capabilities of quantum sensors may contribute to improved biosignal analysis in QML models, mitigated through the use of silico tests, where quantum computers simulate human beings. ECG and EEG data include useful information that can be retrieved and used in ML models as features. One approach in the context of QML is to use quantum algorithms or quantum-inspired techniques to extract features from biosignal data, which can then be supplied to QML models to be analysed and classified. In contrast to MHR and Image datasets, the majority of researchers working on QML also use biosignal data to identify specific diseases. In [130], the authors proposed a study for classifying multi omics lung cancer by using QNN model. The results shws that the quantum algorithm perform well for smaller dataset. Another study propose a framework called the Quantum AIdriven Heart Health Framework (QAIHHF) [131] and use s QLSTM model for MIT ECG dataset. Furthermore, a paradigm for feature engineering based on QC was proposed in [132], by employing LOSO and K-fold cross validation, whereby a publicly available EEG dataset measuring the performance of a mental arithmetic activity including 20-channel EEG signal segments is taken into consideration. The data was collected from 36 healthy right-handed volunteers separated into two groups containing 10 bad counters and 26 good counters. Using LOSO and 10-fold CVs, the model obtained accuracy of 93.40% and 97.88%, with geometric means of 88.44% and 96.42% respectively. A new approach for feature extraction and classification that involves a 1D CNN model using the hybrid classical-quantum layers was proposed in [133]. In the Bonn EEG dataset for the binary classification task, the proposed model obtained maximum accuracy and specificity of 100% while reducing model complexity with the least learning parameters. Additionally, by introducing Gaussian noise into the EEG signal, the proposed technique's robustness was also assessed. [134] offers a hybrid quantum-classical NN model for EEG, EMG, and ECOG analysis that combines a VQC model with a DNN model. Additionally, it also shows that the proposed QNN delivers state-of-art performance while maintaining a small number of trainable parameters for VQC. The model was evaluated using a variety of EEG data sets - stress, RSVP, and MI - and obtained maximum accuracy of 87.23%, 95.12%, and 60.22%, respectively. A parameterized quantum circuit is merged with a traditional LSTM model in [135] to increase the learning capacity and accuracy of predictions, and on a variety of sensor data from the SWELL-KW [136] dataset, the efficiency of the proposed method was evaluated with 87.67% accuracy. The time series dataset includes data from worker information interactions with computers, facial expressions, body postures, heart rates (variability), and skin conductance, which were captured under various working environments. According to S. Sridevi et al. [137], discrete wavelet transform is suggested for decomposing ECG signals, and this is followed by computing a 2D scalogram in order to acquire time frequency features and by applying a convolutional neural network to categories the scalogram images in order to detect arrhythmia. The accuracy and operating curve score of a publicly accessible physio net MITBIH arrhythmia database were recorded as 98% and 100% respectively. Quantum annealing (QA) is an innovative method that was introduced in [138] for important feature selection using physiological signals, whereby four features are extracted from the signal source - respiration, ECG, hand and foot EDAs - for the purpose of stress detection. The Pearson correlation coefficient among the attribute variable and the target variable is used to calculate the bias of feature variable, with results demonstrating the promise of quantum annealing in optimising the training phase of a ML classifier, particularly under situations involving data uncertainty. The cognitive processes of human behavioural outcomes were taken into account by S. Aishwarya et al, [139], taking into consideration different QML classifiers, i.e, VQC, QA classifier and hybrid quantum classical NN. These models were compared, and predictions of upcoming cognitive responses made using EEG data. The preliminary findings of these approaches are shown, and they are quite positive, with up to 61.53% validation accuracy. [140] describes hierarchic quantum mechanics-based architecture for implementing feature extraction and classification in EEG signals, whereby a quantum state was formed via the quantum wavelet packet transformation (QWPT) after the classical EEG signal dataset was created as a quantum state. The random non-linear kernel from the modified QSVM model is used to predict the label of the EEG signal and obtained maximum accuracy of 95.14%. Another novel hybrid method for classifying two types of EEG signals using an auto regressive model was also described in [141], whereby the key features from EEG data were extracted using two distinct element extraction approaches. Back propagation is utilized to train the proposed QRNN model, which is then compared to QNN and Quantum Wavelet NN. As demonstrated in table 2.5, the experimental results show that the proposed model obtained maximum accuracy of 88.28% with a fastest processing time of 6 seconds.

2.8 Discussion

The aim of this section is to address the responses of all the potential research questions outlined in table 2.2. This involves summarising and evaluating the records selected from the state of the art, which will serve as a guideline for future researchers working on QML in the healthcare domain. The quality of each article was ensured after analysing and evaluating each article based on the information provided in table 5.4. The study encompasses a comprehensive analysis of 49 recent articles published between 2018 and 2023, and evaluation was conducted by constructing a set of quality metrics as shown in table 2.6.

RQ1: How QML algorithms can be utilised to enhance medical data analysis, such as improving disease segmentation, classification, or anomaly detection: As discussed briefly in the state of art, QML algorithms have the potential to enhance medical data analysis in several ways, including improving disease segmentation, classification, and anomaly detection. The quantum-enhanced feature selection technique is an important step in medical data analysis, whereby relevant features are selected with a view to building accurate models. Quantum machine learning algorithms can assist in identifying the most informative features from complex medi-

Table 2.6: Content of evaluated research articles taken into consideration based on merit points.

		Content of paper					Additional Quality Measures						
Sr No	Reference	1	2	3	4	5	6	7	8	9	10	Merit points	Paper quality
1	M. Munshi et al. [79]	9	10	18	8	10	9	0	4	15	4	87	Better paper
2	G. Ghazi Enad et al. [80]	8	8	17	9	10	8	0	4	15	4	75	Good paper
3	D. Maheswari et al. [81]	9	10	18	8	10	9	0	4	15	4	87	Better paper
4	B.Prakash et al. [82]	8	9	17	8	10	6	0	4	14	0	76	Average paper
5	G.Abdulsalam et al. [83]	9	9	16	9	8	9	0	4	14	5	83	Good paper
6	S.Diaz-Santos et al. [84]	8	8	16	9	8	9	0	0	14	5	77	Average paper
7	U.Ullah et al. [85]	10	10	16	9	9	9	0	5	13	3	84	Good paper
8	H.Gupta et al. [86]	10	10	18	9	10	10	0	5	14	5	91	Best paper
9	H.Heidar et al. [87]	9	9	15	8	9	10	0	5	14	2	81	Good paper
10	U.Ullah et al. [88]	10	9	19	8	9	10	0	5	14	3	87	Better paper
11	MS.Ishwarya et al. [89]	9	9	19	10	10	10	0	5	15	5	92	Best paper
12	H.Yano [238]	10	10	18	10	10	10	0	5	15	5	93	Best paper
13	D.Maheshwari et al. [290]	10	10	18	9	10	9	0	5	14	5	90	Best paper
14	D.Pomarico et al. [90]	8	9	16	9	8	9	0	5	13	5	82	Good paper
15	S.Saini et al. [91]	9	10	18	10	9	9	0	5	13	5	88	Better paper
16	S.Chakraborty et al. [92]	10	9	17	10	10	10	0	5	14	5	90	Best paper
17	D.Maheshwari et al. [93]	8	8	16	9	8	10	5	5	13	5	87	Better paper
18	D.Sierra-Sosa et al. [94]	9	8	17	8	8	9	0	5	13	4	81	Good paper
19	S.Mardirosian et al. [95]	8	8	16	8	8	9	0	5	13	1	76	Average paper
20	ZY.Chen et al. [96]	9	8	17	9	8	9	5	5	13	5	88	Better paper
21	K.Benlamine et al. [97]	7	8	17	9	8	9	5	5	13	5	85	Better paper
22	OM.Parkash-Patel et al. [98]	8	9	18	8	9	8	0	5	13	5	83	Good paper
23	S.Naguleswaran et al. [99]	8	8	17	9	8	9	5	5	13	5	86	Better paper
24	S. Kumar Roy et al. [100]	8	9	17	8	9	8	0	5	13	5	83	Good paper
25	N.Alsharabi et al. [101]	9	9	17	8	10	8	0	5	12	3	81	Good paper
26	S.Deshmukh et al. [102]	9	9	15	8	9	10	0	5	14	4	83	Good paper
27	D.Konar et al. [103]	10	10	18	9	10	9	0	5	14	5	90	Best paper
28	AK.Mishra et al. [104]	10	10	19	10	10	9	0	5	15	5	93	Best paper
29	S.Moradi et al. [105]	9	10	18	9	9	10	0	5	13	5	88	Better paper
30	EH.Houssein et al. [106]	10	10	20	10	10	10	0	5	15	5	95	Best paper
31	J.Amin et al. [107]	9	10	19	9	10	9	0	5	15	5	91	Best paper
32	S.Toledo-Cortes et al. [108]	10	9	19	9	9	10		5	14	5	90	Best paper
33	J.Reyas Bruno et al. [109]	8	8	16	9	8	10	0	5	13	1	78	Average paper
34	E.Acar et al. [110]	8	9	15	9	8	9	0	5	12	5	80	Average paper
35	J.Amin et al. [210]	10	10	19	10	10	9	0	5	15	5	93	Best paper
36	N.Mathur et al. [111]	9	9	15	8	9	10	0	5	14	5	84	Good paper
37	R.Selven et al. [112]	9	10	18	10	9	9	0	5	13	5	88	Better paper
38	S.Adhiksry et al. [113]	10	10	18	9	10	9	0	5	14	5	90	Best paper
39	V.Dutt et al. [114]	8	9	18	9	10	9	0	5	14	5	87	Better paper
40	S.Chakraborty et al. [115]	10	10	18	10	10	10	0	5	15	5	93	Best paper
41	LI.Yangyang et al. [116]	10	10	20	10	10	10	0	5	15	5	95	Best paper
42	C.Ding et al. [117]	9	10	18	10	10	10	0	5	15	5	92	Best paper
43	M.Benedetti et al. [118]	10	10	20	10	10	10	0	5	15	5	95	Best paper
44	A.Sagheer et al. [119]	9	10	18	9	9	10	0	5	13	5	88	Better paper
45	M.Kaur Saggi et al. [130]	8	9	18	9	9	10	0	5	14	5	87	Best paper
46	S. M Tripathi et al. [131]	8	9	18	9	9	10	0	5	13	5	86	Best paper
47	N.Baygin et al. [132]	10	10	18	9	9	10	0	5	14	5	90	Best paper
48	M.Sameer et al. [133]	8	9	19	9	10	10	0	5	13	5	88	Better paper
49	T.Koike-Akino et al. [134]	8	8	16	8	8	9	0	5	13	2	77	Average paper
50	A.Padha et al. [135]	8	10	17	8	9	9	0	5	14	3	83	Good paper
51	S.Sridevi et al. [137]	9	10	17	8	9	9	0	5	14	3	84	Good paper
52	R.Kumar Nath et al. [138]	10	9	17	8	8	8	0	4	14	3	81	Good paper
53	S.Aishwarya et al. [139]	8	9	18	9	10	9	0	5	14	5	87	Better paper
54	LI.Yao-Chong et al. [140]	9	9	19	8	10	10	0	5	15	5	90	Best paper
55	S.MR Taha et al.. [141]	9	8	13	8	9	10	0	4	14	5	80	Good paper

cal datasets, leading to improved disease segmentation and classification. Moreover, QSVM [141], QRF [142], QKNN [143], QNN [144], VQC [145], and QCNN [146] are the popular algorithms for classification tasks, and these algorithms can leverage the power of QC to enhance classification accuracy, enabling diseases or medical conditions based on patient data to be better identified. Quantum algorithms, such as quantum k-means [147] or quantum spectral clustering [148], can aid in grouping similar data points together, allowing for better disease segmentation, which can be particularly helpful in medical imaging analysis, where accurate segmentation of organs or tissues is crucial for diagnosis and treatment planning. Detecting anomalies in medical data is vital for early diagnosis and treatment of diseases. Quantum machine learning algorithms can assist in identifying abnormal patterns or outliers in large-scale medical datasets, enabling early detection of diseases or unusual patient conditions. By using the unique quantum effects, such as quantum superposition and entanglement, these networks can potentially provide enhanced capabilities for medical data analysis and improve pattern recognition, feature extraction, and classification tasks in medical data analysis.

RQ2: How QML models can be integrated with classical machine learning approaches in order to leverage the strengths of both in healthcare data analysis: Integrating quantum machine learning models with classical machine learning approaches can potentially leverage the strengths of both in healthcare data analysis. A hybrid quantum-classical machine learning model combines classical machine learning algorithms with QC techniques in order to leverage the power of both paradigms. QC has the potential to process information exponentially faster than classical computers in certain scenarios, which makes it an exciting field for various applications, including healthcare and disease diagnosis. The working process for the quantum model uses the same classical classical algorithms, the only difference being the data encoding technique. The classical model uses the classical input information of 0 or 1, known as bits, and processes this data using a classical ML algorithm, providing two exclusive possible states. In the case of the quantum model, the classical one is first converted into a quantum state known as qubits using Feature-Mapping through a unitary gate operation. A quantum feature map utilises a quantum circuit based on the conventional machine learning kernel method in order to represent classical data within the quantum state domain. In order to classify non-linear data by locating distinct hyperplanes, the data is then transformed into a higher-dimensional Hilbert space. The feature map encodes classical input into a quantum variable by employing N unitary gates in order to undertake a ground state transformation. A quantum model contains a quantum layer, quantum circuits, quantum gates, and quantum registers, where the quantum parameters are used for computation. At the output of the quantum model, a measurement circuit is used to decode the quantum variable back into classical data. QML models can be used to optimise complex healthcare systems, such as drug discovery, treatment planning, or resource allocation, by finding the most efficient solutions more quickly than the CC method.

RQ3: What are the fundamental limitations and advantages of QML models are in handling healthcare data compared to classical machine learning models: Researchers often encounter the persistent challenge of isolation, which stems from various factors. Quantum decoherence, triggered by heat and light, poses a significant threat: when qubits are exposed to such conditions, they may lose their quantum properties, including entanglement, resulting in data loss stored within these qubits. Additionally, rotations in logic gates of quantum computers are susceptible to errors. Furthermore, the field of quantum machine learning relies on the utilisation of computers with extended circuit length and error correction, which entails redundancy for each qubit. QML also faces a limitation concerning the utilisation of a limited number of data samples relative to the number of qubits available. To accommodate larger datasets and additional qubits, QC devices necessitate an increased number of logic gates which, in turn, means that the computational cost escalates and prolongs model execution time. These limitations can potentially impact the quantum states, as an incorrect rotation may lead to errors in the final results. Developers of algorithms for quantum computers must pay close attention to the underlying physics - unlike classical algorithms that can be developed following the principles of the Turing machine, designing an algorithm for quantum computers requires a foundation based on the intricacies of raw physics. There are no straightforward formulas that can directly relate it to logical operations, making the development process more nuanced and complex. QML enhances computational speed and facilitates data storage undertaken by algorithms within a programme - it expands learning validation by executing machine learning algorithms on emerging computing devices known as quantum computers. The processing of information relies on the principles of quantum physics, which significantly diverge from traditional computer models.

RQ4: What the considerations and methodologies for evaluating the robustness and generalisability of QML models are when applied to diverse healthcare datasets, including data from different hospitals, regions, or demographic groups: Evaluating the robustness and generalisability of QML models when applied to diverse healthcare datasets, including data from different hospitals, regions, or demographic groups, requires careful consideration of several factors. The first step is data representation, because QML often involves mapping classical data into quantum states. The choice of data representation can impact the model's ability to generalise across diverse datasets, and so it is crucial to select a representation that preserves the relevant features of the data and is robust in terms of variations in data sources. Secondly, to evaluate generalisability, it is essential to gain access to diverse datasets that encompass different hospitals and regions, which ensures that the model's performance is not limited to specific subsets of the data and can handle any inherent variations in healthcare data. A proper pre-processing of healthcare data is then also important, standardising and normalising the data to remove any biases or inconsistencies. Additionally, missing data needs to be handled appropriately to prevent any bias in the model's performance, and ensemble methods employed by combining multiple QML models trained on different datasets or with varied initialisations. Ensemble methods can enhance robustness and generalisability by reducing the impact of individual model biases and errors. The presence of such biases and errors to ensure fairness in predictions across diverse groups also needs to be analysed, and the model's performance evaluated separately for different groups of datasets in order to identify any discrepancies or disparities. Performance of the QML model on any external datasets that were not used during training or model development also needs to be validated, as this helps assess the model's ability to generalise completely unseen data sources. Finally, model interpretability and explainability can be taken into consideration, these being crucial, especially in healthcare applications. Techniques used to explain the model's predictions also have to be developed, providing insights into how the model incorporates different features and influences decision-making.

RQ5: Which types of data can be used for adoption of a quantum predictive model, and is it feasible to use open access datasets for evaluating such models? Additionally, what specific quantum computing devices are applicable for evaluating QML models using healthcare data: Evaluation of QML models can be conducted using various types of healthcare datasets. These datasets can be categorised into two main types: private datasets [81] [85] [88] [290] and publicly available open access datasets [83] [86] [89] [96]. Private datasets are collected from multiple sources such as hospitals, healthcare facilities, and clinical centres, and these datasets comprise healthcare images, electronic healthcare records, or biomedical signals. They contain sensitive and confidential information and are typically not accessible to the general public. On the other hand, publicly available datasets are openly accessible to anyone interested in using them for research or analysis. These datasets also consist of healthcare images, electronic healthcare records, or biomedical signals, but they do not contain sensitive patient information - they are specifically created and made publicly available for research purposes, allowing researchers and developers to test and evaluate their QML models.

There are various QC systems used for executing QML models. For instance, IBM Qiskit [149] is an open-source QC framework developed by IBM, providing a set of tools, libraries, and APIs that allow users to create, compile, and run quantum programs on IBM's quantum systems. IBM has a collection of quantum processors with varying numbers of qubits, which are the basic units of quantum information. For its part, Qiskit supports a range of quantum algorithms and provides a high-level interface for developing quantum machine learning models, being widely used by researchers, developers, and enthusiasts in the QC community. There is also D-Wave quantum annealing [150], which is a specialist approach to QC that focuses on solving optimisation problems, while Google Cirq [151] is an open-source QC framework developed by Google that is designed to create, control, and simulate quantum circuits on both universal and adiabatic quantum systems. Adiabatic QC is another approach to quantum computation, where the system slowly evolves from an initial state to a target state that represents the solution to a problem.

2.9 Conclusions

This work is based on a systematic review, where various QML algorithms that take healthcare datasets into consideration are reviewed and analysed. Initially, I also gathered a total of 2038 articles from four distinct databases - namely, Web of Science, Scopus, IEEE Digital Library, and Springer Link. These articles covered the specific field of study and were published between 2019 and 2023. A meticulous evaluation process was then conducted to ensure impartiality and eliminate any biases. To streamline the research, I employed several criteria to remove duplicate articles and applied various elimination measures and quality assessments. As a result, successfully reduced the initial pool of articles and ultimately identified 49 articles that were deemed highly relevant to our study, and these selected articles served as the primary focus and foundation for our research. All articles that fit the criteria for this review include those that (a) concentrate on learning patient representations, (b) use patient data, such as EHR, images, and ECG signals, and (c) employ quantum machine learning and quantum deep learning models. This systematic review is divided into two parts: creating a potential research questionnaire and defining quality criteria for the records selected, which will serve as a framework for future research and development at the interface of QC and machine learning. Based on the analysis of these papers, I identified several distinct QML designs and implementations, with the primary focus notably seeming to be on applying neural networks within the quantum realm. Among the prominent QML models, I encountered various quantum networks, each serving specific purposes. These included QSVM, QRF, QKNN, QCNN, QLSTM, and VQC algorithms. Interestingly, these quantum algorithms were extensively explored and tested in the context of EHRs, medical images, and biosignal healthcare datasets, which indicates a significant interest in leveraging QC with a view to addressing critical challenges in the healthcare domain, potentially unlocking new avenues for medical advancements and diagnostic improvements.

One of the possible limitations of this study is the difficulty in identifying all the relevant papers that meet our inclusion criteria. The challenge arises from the wide variety of methods used in healthcare, making a comprehensive search using automated keyword queries complex. Furthermore, it is important to highlight the fact that most existing QML algorithms are currently being assessed in classical environments rather than genuine quantum settings. This limitation is attributed to the scarcity of quantum-ready data for conducting QML experiments and the effort involved in converting classical data into quantum data. Consequently, a promising direction for future research involves the development of more efficient encoding methods in order to tackle this issue.

Chapter 3

Case Study I: Covid-19 Disease

3.1 Introduction

In December 2019 Wuhan, China, a new Coronavirus Disease (Covid-19) were discovered [152]. Covid-19 is highly infectious and did spread easily from China to other countries, resulting in a global pandemic [153]. It is an infection with Serve Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) [154], whose typical symptoms are fever, dry cough, fatigue and pneumonia. Serve cases can develop acute respiratory distress syndrome (ARDS), dyspnea, and multi-organ failure additionally [155]. The main risk factors for a serve case are high age and underlying medical conditions including, but not limited to hypertension, obesity, diabetes, cardiovascular disease or immunosupression [156] [158]. Compared to influenza, Covid-19 has similar symptoms although which occur usually at a different time period after infection and often more serve. Furthermore Covid-19 can be accompanied by a loss of smell and taste [159].

In terms of the total number of diagnosed cases, Spain has also been one of the most affected country in the world. The government officially declared state of alarm on March 13, 2020, with restrictions aimed to facilitate diagnosis and ensure the proper treatment of patients. Furthermore, to control the spread of this pandemic, the government also imposes lockdown, population isolation, and mobility restrictions [160]. The number of confirmed cases in Spain in April 2020 were 210773, with 105548 patients were recovered, over 10300 hospitalized to the intensive care unit (ICU), and 23822 fatalities, which was 11.3% of the global fatality rate [161] [162].

Knowledge of the course of an infection influences immensely the applied and necessary treatment. The authors from [163] considered H-score for having a reactive hemophagocytic syndrome, which often associate with serve Covid-19 infections. With medical control values and awareness of risk factors a precise prediction of an infection's course should be possible.

Recently, QML became a popular AI technique for medical diagnosis based on clinical data [164]. A wide variety of applications used data-driven approaches with reasonable results [165], i.e, Image-based diagnosis [166–168] and sentiments analysis [169] are just some examples. In such an era, the current paper presents the two efficient QML algorithms for Covid-19 patients classification. Furthermore, the study discusses the existing data input, the effective data cleaning technique, and hyper-parameters optimization of the E-QSVM and QRF models. In addition, an overview of how to generate a feature ranking by relevance for decision-making and model evaluation is provided. Both the models are run on the quantum processor (Statevector, Qasm simulator) using the identical data samples (2412 samples and 10 features), and the results are compared to previously published models.

3.2 Materials and Methods

This section contains the material used for this experiment, dataset description, and the adopted method, including data cleaning and feature selection for the current experiment.

3.2.1 Materials

The used dataset had 92 features and 6550 entries in total. After removing erroneous entries the remaining 6357 entries were split in a subset with all hospitalized patients and a subset with an equal amount of serve and mild infections, which were used for feature selection. The dataset with hospitalized patients contained 1206 serve and 526 mild cases. For the balanced dataset 680 entries were randomly drawn from non hospitalized patient, which were supposed to have a mild course of infection.

The data were collected from several public hospital from Baja California during the first 8 months of 2020. The population consists of 49,7% women and 50,3% men with an average age of 43 years. The data set contains 4331 infections with a typical pneumonia, 1969 confirmed Covid-19 cases, 46 patients with influenza and other respiratory manifestations as well as 250 records of identified influenza variants, influenza with pneumonia or unidentified viruses. The most common prevalent diseases were a chronic disease with 43%, diabetes and obesity with each 20%.

3.3 Methods

The method section consists of the design and validation of QML models. The section further contains on important features selection, quantum ML models E-QSVM and QRF model, its working principle data encoding, veriational circuit and data decoding.

Table 3.1: 10 most important features from a dataset of hospitalized patients and a balanced dataset.

Rank	Score	Feature by dataset of hospitalised patients	Score	Feature by balanced dataset
1	27	INDICADOR_SOSP_COVID	7	STAY_HOSP_NO HOSP
2	34	RHINORRHEA	26	DIAG_CLIN_PNEUMONIA
3	36	PROSTRATION	46	DESC_TYPE_SAMPLE_1_EXUDADO PHARYNGEAL/NASOPHARYNGEAL
4	37	DIAG_CLIN_PNEUMONIA	48	PNEUMONIA_RADIOGRAPHY
5	41	DESC_RESULT_CONF2_NEGATIVE	61	ESTANICA_HOSP_INTERNAL MEDICINE
6	42	DESC_TYPE_SAMPLE_1_EXUDADO PHARYNGEAL	82	PAIN_THORACIC
7	43	DESC_TYPE_SAMPLE_1_EXUDADO HARYNGEAL/NASOPHARYNGEAL	86	DYSPNOEA
8	58	STAY_HOSP_INTERNAL MEDICINE	94	PRECEDED_DIABETES
9	71	PNEUMONIA_RADIOGRAPHY	94	DIAGNOSIS_FINAL_Covid-19
10	74	AGE	104	OCCUPATION_Retired

3.3.1 Feature Selection

From the 92 features 55 features were dropped in the balanced subset for both the models accordingly. All removed features had either a too low variance, too low entropy, a high Pearson correlation to a kept feature or was equal to or contained by a kept feature. The used threshold for a high Pearson correlation was $|p| > 0.75$ as suggested by Akoglu in [170], whereas a too low variance was considered with 90% of the entries having the same value for a feature. The remaining 39 (respectively 37) features were encoded with a One Hot Encoding and used to determine a ranking of feature importance.

To obtain the desired ranking, a committee is formed with 5 different methods. The f-value of an analysis of variance and χ^2 were calculated and ordered by the assumption that features with a high independence to the target feature are important for prediction. Additionally a random forest, support vector machine and neural network were grid searched, trained and afterwards the feature importance or, when feature importance was not directly applicable, a permuted feature importance determined. Permuted feature importance means, that a baseline metric with all features was calculated. Moreover, for all features each feature was deleted or permuted to calculate the metric again and concluding that a huge difference between the newly calculated metric and baseline metric, means a feature is important for prediction.

After generating 5 rankings by each of the above mentioned methods, a rank was assigned to each feature also considered the score of the desired feature. By adding up the rank of each feature from all rankings and sorting with the lowest score being the best, a final ranking was obtained. This procedure was executed for the dataset with hospitalized patients and the balanced dataset so that 2 rankings for both experiments for both the models were obtained. The 10 most important features are listed in table 3.1. Especially the features from the balanced dataset align with the known risk factors as well as symptoms of serve infections [155] [171].

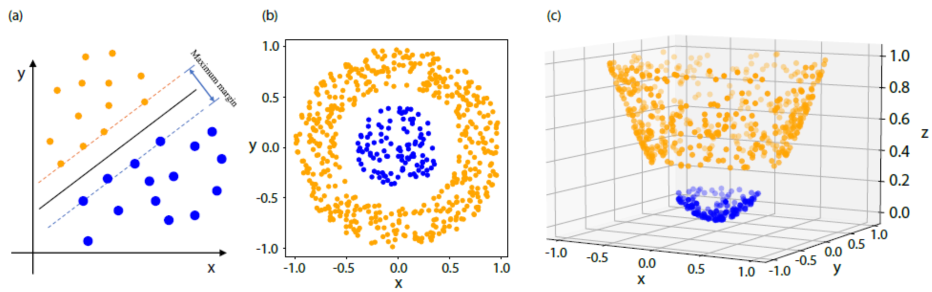

Figure 3.1: (a) SVM schematic diagram (b) Non-linear data. (c) A Kernel based technique [173].

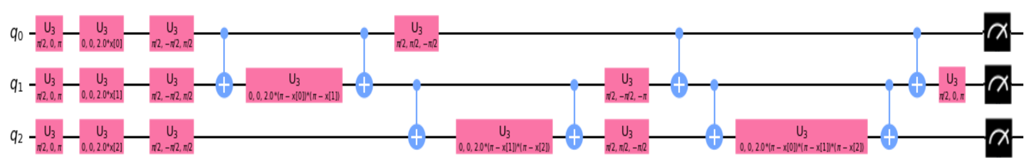

Figure 3.2: Gated circuit diagram of QSVM.

3.3.2 Enhanced Quantum Support Vector Machine

The Support Vector Machine (SVM) is a supervised classical ML model that is used to categorize new testing vectors into classifications. The SVM model learns from training samples to classify incoming samples into negative and positive categories. In case of positive class for training data points $\vec{x}_i$, SVM creates a hyper-plane with $\vec{\omega}.\vec{x}+b=0$, such that $\vec{\omega}.\vec{x}+b \geq 1$ and for negative class considering the same data points $\vec{x}_i$ it creates $\vec{\omega}.\vec{x}+b \leq -1$. This aims to maximize the distance between the two classes during the training phase. Furthermore, to obtain a fine estimation for classification of incoming data samples $\vec{x}_0$, it is reasonable to divide the two classes as much as feasible. The main objective of SVM is to discover hyper-planes that optimize the distance $\frac{2}{||\vec{W}||}$ constraint to $\vec{y}_i(\vec{\omega}.\vec{x}_i+b) \geq 1$. The norm vector $\vec{\omega}_i$ is denoted by $\vec{\omega}_i = \sum_{i=1}^{M} \alpha_i \vec{x}_i \geq 1$, where α_i indicates the wight of i^{th} training vector $\vec{x}_i$. Eventually, the process of selecting the perfect parameters b and α_i is the same as selecting the optimal hyper-plane. The classification of the new vector can be obtained by using the decision function as [172],

$$y_i(\vec{x}_0) = sgn(\vec{\omega}.\vec{x}+b) \tag{3.1}$$

Above the subscripts, $\vec{\omega}$ and $\vec{x}$ indicate the optimal parameters determined from the optimization technique. Once the ideal parameters have been determined, classification will become a linear process. By solving a linear equation using the least-squares approximation of SVM, the optimal parameters may be determined [172].

$$S = (\vec{x}_1, y_1), (\vec{x}_2, y_2), ..., (\vec{x}_M, y_M) \tag{3.2}$$

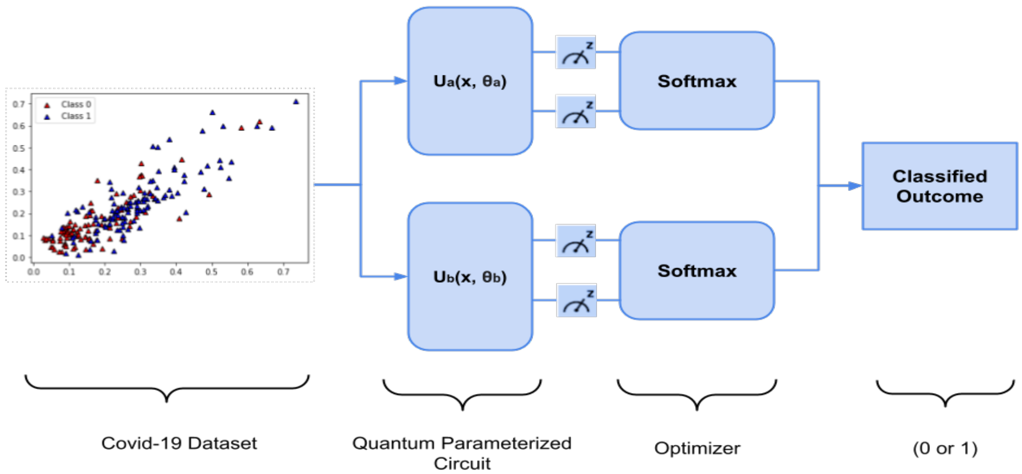

Figure 3.3: Block diagram of QRF, where $U_a(x,\theta_a)$ and $U_b(x,\theta_b)$ are the parameterized circuits and SoftMax is used as an optimizer.

Where M represents the training dataset size with N dimension. To make this simple, the binary classification problem has been used as an example and reference to the labels as $y_i = \pm 1$, figure 3.1(a) [173] illustrates the basic concept in a simplified manner. The Kernel technique considered which helps to generate a hyper-plane in higher dimensional kernel space. Such tactic shift the non-linearly separable data samples into linearly-separable data samples [174].

$$K(X,Z) = \phi(x).\phi(z) \tag{3.3}$$

where the function $\phi(x)$ represents the mapping feature space into kernel space and shifting the data samples into higher dimension space for linear separating as shown in figure 3.1(b) and 3.1(c) [173].

Another supervised ML algorithm is QSVM, also known as the enhancement of quantum and ML technology [174], which functions in a similar manner to classical SVM. It also involves using quantum gates and orthogonal transformations to improve the efficiency of supervised learning algorithms. Moreover, QSVM is also enriched with some additional techniques, i.e, data encoding techniques, feature mapping, different gates spinning, and executed through a quantum processor. The data encoding technique converts classical data sample $\vec{x}$ into a quantum variables $|\Phi(\vec{x})\rangle$ through a unitary gate by spinning the qubits to a given value $U\phi\vec{x}\,|0\rangle$. The measurement operation has been utilized, which is dependent on the generalized quantum circuit $W(\theta)$ as shown in figure 3.2. Furthermore, an parameterized circuit $W(\theta)\; U\phi\vec{x}\,|0\rangle$ is considered to obtain a classified output values of 1 or 0 for each classical input.

3.3.3 Quantum Random Forest

Quantum ML refers to the idea of implementing the quantum algorithm into the ML process. It usually uses a ML algorithm to analyze conventional data and execute it on quantum computer [175]. A random forest algorithm (a sub-category of classical ML) is an ensemble learning technique that generates a series of decision trees using a sample of the random chosen training dataset. It is the most popular AI computational algorithm and has widely been used in classification, regression and other different kinds of applications [176].

$$ni_j = W_J C_J - W_{left(j)} C_{left(j)} - W_{right(j)} C_{right(j)} \tag{3.4}$$

$$\text{fi}_i = \frac{\sum_{j\text{ :node } j \text{ splits on feature i}} ni_j}{\sum_{k\in \text{ all nodes}} \text{ni}_k} \tag{3.5}$$

In above the subscript ni_j indicates the importance of j node, W_J is the node j weighted samples, C_J is the node j impurity value, $left(j)$ and $right(j)$ indicates child node from left and right split on node j respectively. On a decision tree, the relevance of each feature is computed as above, where fi_i shows the importance of feature i and ni_j is the importance of node j. The importance of features can be normalized to a value between 0 and 1, and can be calculated by dividing some of the important values of all features [176].

$$normfi_i = \frac{fi_i}{\sum_{\text{j}\in\text{all features}} fi_j} \tag{3.6}$$

$$RFfi_i = \frac{\sum_{j\in\text{all trees}} normfi_{ij}}{T} \tag{3.7}$$

Table 3.2: Performance evaluation metrics of the proposed models.

Classifiers	Confusion Matrix			Performance Evaluation				Average Values	
	Prediction → Actual ↓	0	1	Precision	Recall	F1 Score	Test Size	Accuracy	F1-Score
SVM	0	172	70	0.81	0.71	0.76	242	77%	77%
	1	40	201	0.74	0.83	0.79	241		
E-QSVM	0	171	71	0.83	0.71	0.76	242	78%	78%
	1	35	206	0.74	0.85	0.80	241		
RF	0	158	79	0.75	0.67	0.71	242	73%	73%
	1	52	194	0.71	0.79	0.75	241		
QRF	0	188	54	0.73	0.78	0.75	242	75%	75%
	1	69	172	0.76	0.71	0.74	241		

Finally, in the Random Forest model, the average of all the trees is the most important property. The total number of trees is divided by the sum of the feature's significance value on each tree. In the above equations, $RFfi_i$ is the importance of calculating feature i from all trees in the random forest model, $normfi_{ij}$ indicates the importance of normalized feature i in tree j and T is the total number of trees.

In QML, to solve the binary classification problem, the QRF is designed whereas two IBM simulators (QASM-Simulator, Statevector Simulator) are combined and executed in a parallel manner. The pennylane qiskit packages are considered, where the function of both the simulators is identical to various trees in the classical RF algorithm. Each simulator produces an individual prediction, and the most confident simulator prediction is used to create an ensemble model. It demonstrates how the ensembling allows the simulators to specialize in distinct classes using a pre-trained model and the PyTorch interface. The diagram depicted in figure 3.3 summarize the model, where two devices of 10 qubits qiskit.aer (QASM, Satetvector) from PenyLane-Qiskit plugin is considered for 10 features of data. R_x gate rotations are used to enter data, and each device has its circuit with its own set of trainable parameters. Both circuits provide a PauliZ measurement for all of the qubits as an output. Furthermore, after passing the parameterized data through a SoftMax optimization function, one dimensional probability vectors corresponding to the binary classes are produced.

Table 3.3: E-QSVM Comparison with Previous State of the Art.

S.No	Authors	Algorithms	Used Methodology	Dataset	Accuracy
1	H. Li et al [177]	GD-QSVM, LS-QSVM	Least Square and Gradient Descent	Kaggle (Blood)	76.8%, 76.6%
2	A. Kariya et al [178]	QSVM	Data Encoding, Used real quantum chip	OCR	62%
3	**Proposed**	**E-QSVM**	**Feature selection, Optimization**	**Covid-19**	**78%**

3.4 Results and Discussions

The experimental results are carried out through a Jupyter notebook with a torch interface. The QSVM in IBM's framework Qiskit version 0.31.0 and quantum random forest from PennyLane are used to classify the data, and both are executed in the given simulator Aer version 0.9.1 respectively. The dataset is balanced and normalized by using standard-scalar, which normalizes the features by removing the mean and then scaling it by multiplying all the values by the standard deviation. The Min-max scalar converts a feature's minimum value to 0, maximum value to 1, and any other values to a decimal between 0 and 1. The dataset is further divided into a training set (80%) and a test set (20%). Each classifier is built using training data, and test samples are used to compare the classification model's prediction to the known test labels.

Table 3.4: QRF Comparison with Previous State of the Art.

S.No	Authors	Algorithms	Used Methodology	Dataset	Accuracy
1	Maheshwari et al [179]	SVM, RF, XGB, and Qboost	Ensemble Learning	Diabetes	69%
2	PennyLane et al [180]	QRF, and Soft-Max	Executed two quantum simulators, Circuit rotations	Iris	72%
3	**Proposed**	**QRF**	**Feature selection, and Optimization**	**Covid-19**	**75%**

3.4.1 Experimental Results of E-QSVM

The performance metrics of the proposed E-QSVM along with the classical SVM are depicted in table 3.2. In the case of the E-QSVM algorithm, the precision measurements for class 0 and 1 are noted at 83% and 74% respectively. Likewise, the recall score values for class 0 and class 1 are recorded at 71%, 85.5%, while the F1 scores for both the classes are 76% and 80% respectively as mentioned. To summarize the overall competency, the proposed E-QSVM achieves the average accuracy for the Covid-19 dataset is noted as 78%, and the classical SVM for the same data sample obtains 77% accuracy for both the classes. The robustness of the model is assured by comparing it with the recently published quantum algorithms mostly by considering the clinical datasets, shown in table 3.3 . For an instant, H. Li et al [177], implemented GD-QSVM and LS-QSVM by considering the blood data samples and obtained the accuracy of 76.8% and 76.6% respectively. Similarly, QRF is also compared with [178], which achieved 72% accuracy by considering the optical character recognition (OCR) dataset. Finally, to summarize, the proposed model is trained with comparatively big data samples and is more accurate as compared to the existing models.

3.4.2 Experimental Results QRF

The experimental metrics of the proposed models RF and QRF for the Covid-19 dataset are shown in table 3.2. In the case of the QRF algorithm, the precision score for class 1 and class 0 are noted at 73% and 76% respectively. The recall and F1 score values for class 1 and class 0 are recorded at 71%, 78%, 74%, and 75% respectively. Our second modified classical RF also obtained tremendous results. To summarize the overall competency, the proposed QRF has an average accuracy of 75% and the classical RF has 73% accuracy for both the classes. The model's robustness is confirmed by comparing it to previously published quantum algorithms, which are premised on existing literature provides. For example, D. Maheshwari et al [179], uses the Diabetes dataset and adopted en ensemble model, and obtained the highest accuracy of 68.38%. Similarly, the proposed work is also compared with another ensemble model developed by the PennyLane dev team [180], which achieved 72% accuracy with a small Iris dataset (150 samples). Eventually, our proposed E-QSVM model has notable results with less complexity as compared to the existing models and obtained 75% accuracy as shown in table 3.4.

3.5 Conclusions and Future Direction

In this study, a pipeline for transforming and pre-processing data has been presented , where QML algorithms may be used to classify it. The proposed E-QSVM, QRF and the design of efficient pre-processing techniques including data balancing and feature selection make it more competent to obtain a better prediction. The data about Covid-19 was obtained from a hospital that included both male and female patients. Here I utilized the E-QSVM model from the qiskit library and modified it with some additional hyper-parameters for optimization. The different entanglement and spinning of the gates ensure better prediction. The proposed QRF is an ensemble model, by executing the two IBM simulators in a parallel manner. Thus every simulator produces its prediction, and the ensemble model is built using the best reliable simulator prediction.

Given the current limitations on coherence time and qubits accessible on current devices, the proposed work currently focused on the development of supervised QML approaches. In the future, the current models can be modified by using different data encoding techniques, such as amplitude encoding and angle encoding.

Chapter 4

Case Study II: Cardiovascular Disease

4.1 Introduction

Ischemic heart disease has become a major health concern for many patients due to its high mortality rate throughout the world [181]. Early diagnosis of cardiac diseases has the potential to save many lives, whereas routine clinical data analysis presents significant challenges in identifying IHD disorders such as heart attacks, strokes, heart failure, etc [182] [183]. The quick rise of coronary heart disease is more common in males than females, especially in middle aged or older persons which poses a serious risk to public health and places a heavy burden on global healthcare systems [184] [185]. Additionally, the children in developed countries have an increasing CAD risk [186]. According to the World Health Organization (WHO), every year globally one third of the population, or around 17.9 million people, die from IHD, with Asia having the highest prevalence [187] [188]. The European Cardiology Society (ESC) reported that each year, 3.6 million adults globally suffer from cardiac disease, and 26 million have been diagnosed with coronary heart diseases [189]. The various unhealthy behaviors, such as diabetes, rising cholesterol, obesity, hypertension, an increase in triglycerides, and others, increase the risk of developing coronary heart disease [13]. Moreover, there are also other symptoms that the American Heart Association lists, such as difficulty sleeping, an irregular heartbeat, difficulties breathing, swollen legs, and occasionally weight gain that happens rapidly, up to 1-2 kg each day [190] [191]. All of these symptoms are indicative of various illnesses, including those that affect elderly people, and are more challenging to diagnose at an early stage.

Currently, echocardiographic stress tests and pharmacological stress tests are used to diagnose IHD by performing a functional evaluation of the coronary artery. A coronary angiography, which provides information on the anatomical structure of the heart, is another diagnostic imaging test for CAD. However, because of the complexity of the coronary anatomy and the structure of the plaque, there is still a chance that the disease will be erroneously diagnosed [192]. Additionally, patients may encounter the possibility of health issues with the use of general anesthesia and contrast material. Estimating the risk of cardiac failure accurately aids in patient protection and the avoidance of severe cardiac strokes [193]. Researchers now have access to massive dataset to create the most precise prediction model thanks to the development of artificial intelligence and machine learning. When trained on relevant data, machine learning algorithms are capable of accurately diagnosing diseases [194]. Recent research, which focuses on cardiac disease in both children and adults, highlights the importance of decreasing mortality from IHD. As a result, it is necessary to investigate the effects of risk factors, their high prevalence in the general population, their significant impact on individual cardiac diseases, and their treatment capability to reduce cardio diseases [195].

Furthermore, the selection of the best features is an essential step to keep in mind when applying machine learning (ML) techniques for the prediction of a certain disease. Initially, the input dataset in their current form are inconsistent and contain missing data, duplicate entries, and other redundant information [196]. In model prediction, it is important to choose the strongest and most important attributes that may be used for the identification or classification of the targeted disease. For the development of accurate prediction, effort should be made to choose the important feature combinations, which help make the machine learning process more accurate and maximize the prediction power of the model [197]. During modeling the predictors for such diseases, different studies consider different risk factors or features as input for their models depending on their feature selection techniques [198] [199]. When implementing the IHD prediction models, various factors are taken into consideration, i.e., high cholesterol, age, chest pain, fasting blood pressure, diabetes, results from a resting electrocardiogram, exercise induced angina, depression, heart status, maximum heart rate, family history, poor diet, obesity, alcohol consumption, and inactivity [200] [201]. These factors play an important role in disease classification and detection.

Recently, numerous quantum enhanced machine learning algorithms have been introduced to speed up some specific ML tasks [202] [203]. The use of quantum deep learning (QDL) and ML models for medical applications has gained significant attention in the past few years [204] [205]. In addition, various researchers have previously used different quantum machine learning algorithms [206] [207] by utilizing real-time healthcare datasets [208] [209]. The academic community and the medical field have both recently shown their interest in supervised learning [210]. The quantum supervised machine learning algorithms, i.e, Variational Quantum Classifier (VQC) [211] [212], Quantum Support Vector Machine (QSVM) [213], the robust pre-processing techniques [214] [215] and error reducing algorithms [216] [217] has a numerous contributions in coping with the classification problems. Convolutional neural network models (CNNs) are considered the most successful machine learning algorithms for pattern recognition and disease prediction. Furthermore, CNNs have shown enhanced performance in early disease diagnosis, which is rooted in the concept of managing the input dataset through numerous representational layers [218] [219]. Many hybrid algorithms use parameterized circuit optimization to reduce the cost function, which usually consists of the expectation value of an output observable, which is the same as classical model parameter optimization [220–222]. From the other side, by using the concept of quantum supremacy, QC has demonstrated its importance for solving problems. The frequent development of QC in the field of ML can naturally handle several problems with complicated input correlations that can be exceptionally difficult for classical computers [223–225]. Moreover, near-term quantum computers have also shown a tremendous impact from quantum computing on ML [226].

In this research, the aim is to construct a fully connected quantum convolutional neural network model, demonstrating the application of various methods for early stage diagnosis of IHD. The proposed model uses a quantum circuit based technique, which reduces the model's complexity by reducing the number veriational parameters. Multiple qubit gates are applied to adjacent qubits in the convolution circuit to find the hidden state, and the pooling circuit reduces the model complexity by considering the CNOT gate operation. The dataset used for this work is private, which is provided by a hospital contains on health record of patients. The desire disease has been classified by using the labels obtained from the input dataset. At first, I tried to explain how disease classification works by utilizing the feature extraction method. Later on, to increase the chances of success, A hybrid model combine with the classical and quantum versions of the convolutional neural network algorithm has been used with the input dataset, which consists of features related to IHD, i.e., hypercholesterolemia, smoking habits, antipsychotic drugs, heart failure, and so on. Comparing the proposed quantum model to a classical Optimized-CNN and FCNN algorithm, the classification results achieved are quite promising. The primary objective of this research is to propose a robust model by combining the benefits of quantum computing with CNNs, and clinical record data is considered a key tool. This results in the least expensive model for early stage IHD diagnosis. The proposed model helps to enhance the CNN performance and ensure the accurate prediction of IHD, and it is also useful to speed up the diagnosis process for heart patients in the early stages.

4.2 Materials

The dataset used for this research work has been provided by Basurto Hospital of the Basque Country and includes patient health record information. For patients suspected of having coronary heart disease, the pretest probabilities should be predicted based on, symptoms, gender, age, BMI, etc. The main patient level data collection tools in healthcare are EHR, which can also be used for research-related purposes. To enhance the pretest IHD risk adjustment more precisely, I examine which EHR features have greater connections with heart failure and cardiac problems. The gathered dataset consists of patient records of both genders collected between 2016 and 2018. This data was gathered for approximately 2,199,711 patients, among those 43,835 individuals who needed treatment for chest pain and thoracic pain within the particular time period. Initially, there were 82 attributes in the dataset, including the targeted CAD variable, and each variable was either categorical or continuous. The dataset has 42 effective features, are described below [227].

- Catheterization: The process in which a flexible tube is guided through a blood vessel to diagnose clogged arteries, irregular heartbeats, etc.
- Ergometry: Ergometry is the measurement of human physical performance.
- Echo Stress: A test that uses ultrasound imaging to show how well the heart muscle works at pumping blood to the body in a stressful situation.
- Echocardiogram: It represents the graphic outline of the heart's movement.
- ECG: The heart's electrical activity is recorded by an ECG. It is a frequent and painless test for identifying heart issues and keeping tabs on heart health.
- Depression: Depression is characterized by a constant feeling of sadness and a lack of interest. It is a medical illness that affects a person's mood and ability to function.
- Alcohol: Alcohol intake causes illness when a person's body develops a dependence on or addiction to it.
- Drug dependence: A biopsychosocial situation whereby a person's functionality is dependent on the consumption, meaning cessation of repeated exposure to a stimulus.
- Anxiety: Mental health condition which carries excessive anxiety or worries, about several things such as personal health, work, and social interactions.
- Dementia: Refers to abnormal brain alterations that affect memory, language, and problem-solving skills in addition to other aspects of thinking.
- Renal insufficiency: Renal failure is a medical condition in which the kidneys stop working less than normal levels.
- Diabetes mellitus: To put it more simply, diabetes mellitus occurs when there is an abnormally high level of sugar flowing in the bloodstream of a person.

- Diabetes type 1: A syndrome when the body is unable to create enough insulin, causing the blood glucose level to rise too high.
- Diabetes type 2: A condition when the blood glucose level is excessively high because the body cannot respond to insulin and frequently produces insufficient insulin.
- Dyslipidemia: The blood's lipid (fat) content is abnormally high.
- Hypercholesterolemia: Hypercholesterolemia is a type of dyslipidemia that is described as a high level of cholesterol in the blood.
- Ventricular fibrillation: Refers to a type of arrhythmia, or irregular heartbeat, affecting the heart's ventricles.
- Auricular Flutter: Refers to another type of arrhythmia which causes the heart to beat faster and inefficiently.
- Heart failure: A disorder when the heart is unable to adequately circulate blood throughout the body.
- Obesity: Overweight and obesity are described as accumulations of abnormal or excessive fat that pose a health risk.
- Age: Number of years a person has lived
- Antithrombotic agents: Antithrombotic agents are drugs that reduce the formation of blood clots.
- Acetylsalicylic acid: To avoid blood clots, strokes, and myocardial infarction (MI), doctors prescribe analgesic drugs that stop platelet aggregation.
- Protonpump inhibitors: Anti acid medications known as proton pump inhibitors act by lowering stomach acid production.
- Diuretics: Diuretics are drugs that work by increasing the amount of water and salt that the body excretes in the form of urine.
- Hormonal Contraceptives: A type of birth control that uses hormones such as estrogen and progesterone by blocking the release of eggs from ovaries.
- Anabolic steroids: A synthetic version of testosterone, which increases protein within cells, is used for growth muscle stimulation.
- Antigout Preparations: Agents that increase uric excretion by the kidney, decrease uric acid production.

- Immunomodulators: A group of drugs that work includes working on the immune system directly, by turning down some proteins and turning up others.
- Antidepressants: Refers to the category of medications used to treat anxiety and depression.
- Antipsychotics: A specific class of psychiatric drug can be acquired with a prescription for treating psychosis.
- Benzodiazepines: Sedative medications like benzodiazepines slow down both physical and mental functions.
- Cardiac therapy: A program that uses exercises, support, and education for people to get recovered from a heart attack, heart surgery, or other heart problem.
- Antihypertensives: A class of medication used for the treatment of hypertension (high blood pressure).
- Vasodilatadores: Drugs known as vasodilators relax the smooth muscles in the blood vessels, allowing them to expand.
- Beta blockers: The medications used lower blood pressure. The blocks the effects of the hormone epinephrine, more often known as adrenaline.
- Calcium channel blocker: A drug that lowers blood pressure by blocking calcium from entering the heart and artery cells.
- Block SRAA: A system used to reduce blood pressure and slow the progression of renal disease.
- Antilipidemics: Lipid-lowering agents are a type of medicine used in the treatment of high levels of fats.
- Smoking habit: Starting a smoking habit quickly leads to both a psychological and a physiological addiction.
- Hypertension: The medical syndrome known as hypertension causes persistently high blood pressure in the arteries.

4.3 Method

The method section contains the process of data cleaning, the selection of important features, and the proposed model as shown in figure 4.1. The data cleaning consists of removing the unwanted information, and feature selection consists of a wrapper and

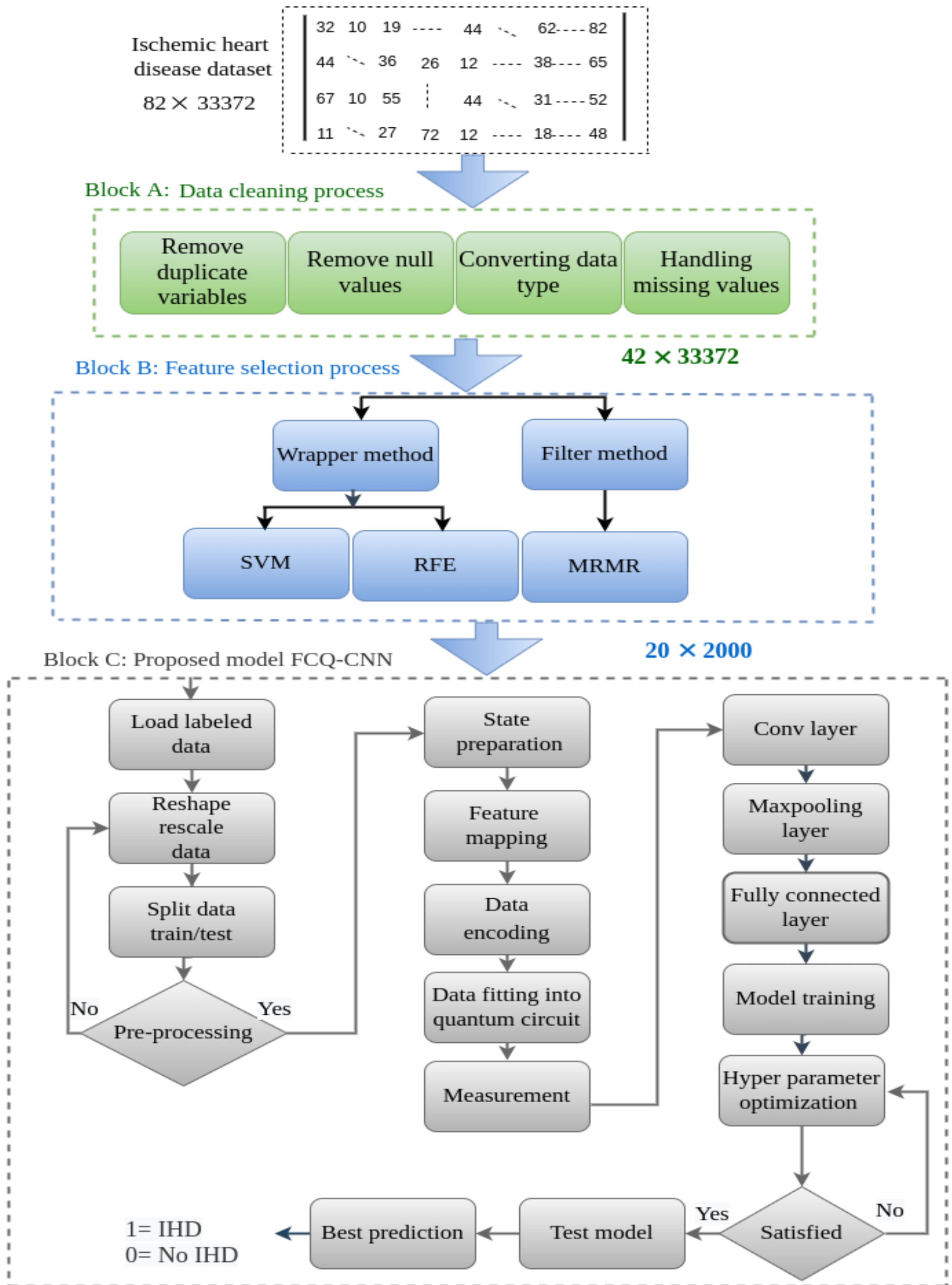

Figure 4.1: The complete block diagram of the proposed methodology.

filter technique, where the wrapper is accomplished by using SVM along with RFE, and then the MRMR filter is used to avoid feature duplication. The proposed model further consists of the quantum state, data encoding and decoding, and different layers, as briefly discussed below.

Table 4.1: Dataset before and after pre-processing.

Parameters	Dataset before cleaning	Dataset after cleaning
Total samples	33372	33372
Features	82	42
Type	Categorical	Categorical
Output	IHD	IHD

4.3.1 Data Cleaning

Data cleaning involves correcting or deleting inaccurate, corrupted, improperly formatted, inoperable, and redundant data from a dataset. When combining different information sources, there are various possibilities for data duplication or labeling errors. In such a context, I have removed 10,463 patients who had no additional diagnostic testing from the dataset, so 5379 (16.1%) out of the 33,372 remaining patients who had CAD during the desired time period are considered. The dataset has been cleaned by removing all the duplicate and irrelevant variables, eliminating nulls and inoperable variables, converting data types, and handling the missing values. Due to null fields, which indicate they do not contain full information for some of the records, the 20 variables that did not match the criterion were deleted.

Similarly, in the elimination of duplicate variables, the 14 variables that have 0 values and make no contribution to the model's prediction were also eliminated. Patient admission and discharge dates, patient IDs, and other irrelevant information were also removed since they were not informative. The essential feature between each pair of variables was chosen using a threshold. Subsequently, the clean dataset contains 42 attributes, including the targeted CAD variable as depicted in table 4.1. The information on the patients who had various therapies, such as cardiac therapy or catheterization, was also included. Moreover, risks associated with various addictions, including alcoholism, certain medicines or pharmaceuticals, including antithrombotic agents, proton pump inhibitors, and anabolic steroids, as well as specific health issues, including diabetes mellitus, obesity, and hypertension, were taken into account. The dataset after the cleaning process, which has 42 effective features, where all the features are categorical, is shown in table 4.2.

4.3.2 Feature Selection

The technique uses a feature selection procedure to choose the minimum possible input parameters while creating a predictive model. It is preferable to decrease the size of input parameters, which minimizes the computational cost of the model and, under certain conditions, improves its performance. In this research, the SVM and RFE were

Table 4.2: The effective 42 features after cleaning.

Name of Feature	**Treatment**	**Variable-type and Values**
Catheterization	Catheterization procedure	Categorical 1/-1
Ergometry	Ergometry study	Categorical 1/-1
Echo stress	Stress Echocardiography	Categorical 1/-1
Echocardiogram	Echo-cardiogram	Categorical 1/-1
ECG	Electrocardiogram	Categorical 1/-1
Depression	Diagnose depression	Categorical 1/-1
Alcohol	Alcoholism	Categorical 1/-1
Drug dependence	Drug dependence	Categorical 1/-1
Anxiety	Anxiety disorder	Categorical 1/-1
Dementia	Dementia	Categorical 1/-1
Renal insufficiency	Renal failure	Categorical 1/-1
Diabetes mellitus	Insulin	Categorical 1/-1
Diabetes type 1	Type 1 diabetes	Categorical 1/-1
Diabetes type 2	Type 2 diabetes	Categorical 1/-1
Dyslipidemia	Dyslipidemia	Categorical 1/-1
Hypercholesterolemia	Hypercholesterolemia	Categorical 1/-1
Fibrillation-palpitation	Ventricular fibrillation	Categorical 1/-1
Flutter	Auricular flutter	Categorical 1/-1
Heart failure	Cardiac insufficiency	Categorical 1/-1
Obesity	Reduced-calorie diet	Categorical 1/-1
Age	Age at notification	Categorical 1/-1
Antithrombotic Agents	Antithrombotic agents	Categorical 1/-1
Acetylsalicylic Acid	Acetylsalicylic acid	Categorical 1/-1
Proton-pump inhibitors	Proton-pump inhibitors	Categorical 1/-1
Diuretics	Diuretics	Categorical 1/-1
Hormonal contraceptives	Hormonal contraceptives	Categorical 1/-1
Anabolic steroids	Anabolic steroids	Categorical 1/-1
Antigout preparations	Antigout preparations	Categorical 1/-1
Immunomodulators	Immunomodulators	Categorical 1/-1
Antidepressants	Antidepressants	Categorical 1/-1
Antipsychotics	Antipsychotics	Categorical 1/-1
Benzodiazepines	Benzodiazepines	Categorical 1/-1
Cardiac therapy	Cardiac therapy	Categorical 1/-1
Antihypertensives	Antihypertensives	Categorical 1/-1
Vasodilatadores	Vasodilators	Categorical 1/-1
Beta blockers	Beta-blockers	Categorical 1/-1
Calcium-channel blockers	Calcium-channel blockers	Categorical 1/-1
Block SRAA	SRAA blockage	Categorical 1/-1
Antilipidemics	Lipid-lowering	Categorical 1/-1
Smoking habit	Smoking	Categorical 1/-1
Hypertension	Hypertension	Categorical 1/-1

Table 4.3: Selected 20 important features.

Ranking	Importance	Name of variables
6	0.040610	Catheterization
22	0.013607	Ergometry
18	0.018645	Echo Stress
10	0.024103	Echo-cardiogram
11	0.021145	ECG
13	0.015723	Diabetes millitus
15	0.018850	Hypercholesterolemia
16	0.015983	Fibrillation-palpitation
17	0.014562	Obesity
1	0.285714	Age
3	0.084274	Antithrombotic agents
16	0.015029	Proton pump inhibitors
13	0.016108	Diuretics
31	0.017277	Benzodiazepines
32	0.285714	Cardiac therapy
35	0.089991	Beta blockers
36	0.016288	Calcium channel blockers
37	0.030682	Block SRAA
38	0.076089	Antilipidemics
39	0.016268	Hypertension

used in combination with a maximally relevant minimal redundancy (MRMR) filter to minimize the input data dimensions. The goal of the MRMR technique is to choose a group of features that are both highly relevant and minimally redundant for identifying different classes of cardiovascular diseases. For an instant, let $D = \{x_{i,k}\}_{n\times K}$ is the expression matrix for the entire dataset, where $x_{i,k}$ is the i_{th} feature of sample k, n represent the total features and K is the total number of samples. So the feature expression for k_{th} sample would be $x_k = (x_{1,k}, x_{2,k}, \ldots, x_{n,k})$, $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,K})$ is i_{th} feature expression, and $G = \{1, 2, \ldots, n\}$ may represent the feature index set. Furthermore, for binary classification, $y_k = \ell \in \{+1, -1\}$ represents the target class values for the k samples by considering +1 and -1, respectively. In a subset $S \subset G$, the relevancy R_s of the features is given as [228],

$$R_S = \frac{1}{|S|} \sum_{\ell} \sum_{i \in S} I(\ell, i) \tag{4.1}$$

$$I(\ell, i) = \sum_{x_{i.}} p(\ell, x_{i.}) \log \frac{p(\ell, x_i)}{p(\ell) p(x_i)} \tag{4.2}$$

Where $I(\ell, i)$ indicates the mutual information among the class label ℓ and feature i. In a subset S, the mutual information can be used to determine the redundancy of feature i with other features and provide the following value [228].

$$Q_{S,i} = \frac{1}{|S|^2} \sum_{i' \in S, i' \neq i} I\left(i, i'\right) \tag{4.3}$$

For feature ranking, SVM-RFE was first presented by Guyon *et. al* [229], which rates the quality performance of the feature subsets for the modeling technique being utilized as a black box evaluator. In essence, the recursive feature elimination technique ranks the features by recursively removing the un-useful features and creating a model based on the remaining ones. The following are the variables of the weight vector w of the SVM that provide the ranking score [230].

$$w = \sum_k \alpha_k y_k x_k \tag{4.4}$$

The MRMR filter may not produce reliable performance when used alone since the classifier operates independently and is not included in the feature selection process. The selected 20 rank wise important features are shown in table 4.3, where the redundancy between the features is not taken into account. By adding an MRMR filter to SVM-RFE, with hope to improve the model's performance by reducing the number of relevant features that are used more than once. This will also make it easier to classify things.

4.3.3 Fully Connected Quantum-Convolution Neural Network (FCQ-CNN)

The Quantum Convolution Neural Network (QCNN) has distinguished itself as one of the machine learning models that perform very well in pattern recognition. The quantum convolution layer can be applied after the data has been encoded into a quantum system. In order to achieve a pure quantum state, a circuit of featuremap has been used for data encoding in the computational basis states, with the number of samples in the probability amplitudes. The subsequent layer, which consists of $O(log(n))$ layers for n input qubits, creates a shallow circuit depth. The desired quantum model structure is capable of avoiding one of the most significant problems connected with parameterized quantum circuit-based algorithms [231]. These architectures seem to be uniquely connected to the tensor network, which offers an easy way to investigate many body physics neural networks, and their interactions. The gradual reduction of the quantum bits is equivalent to the Maxpooling process in CNN. Transnational invariance is a key feature of the QCNN architecture that requires parameterized quantum gate blocks to be similar throughout a layer. The resulting quantum state derived from the I_{th} layer of QCNN can be described as [232].

$$\left|\psi_i(\boldsymbol{\theta}_i)\right\rangle\left\langle\psi_i(\boldsymbol{\theta}_i)\right| = \mathrm{Tr}_{B_i}\left(U_i(\boldsymbol{\theta}_i)\left|\psi_{i-1}\right\rangle\left\langle\psi_{i-1}\right|U_i(\boldsymbol{\theta}_i)^{\dagger}\right) \tag{4.5}$$

The primary objective of the QCNN model is to learn relevant data for accurate prediction of quantum states and their associated labels. At quantum state, the encoded data is denoted by $|x_{in}\rangle$. The features are extracted layer by layer using transformations in parameterized quantum circuits. Quantum measurements on certain qubits are carried out towards the output of the model, which determines the expectation values that represent the classification results as shown in figure 4.2. The proposed FCQ-CNN model is relatively similar to conventional CNNs, the only difference is that it additionally consists of a quantum layer. The complete diagram of the proposed model is shown in figure 4.3, which contains quantum layers, convolutional layers, Maxpooling layers, and a fully connected layer.

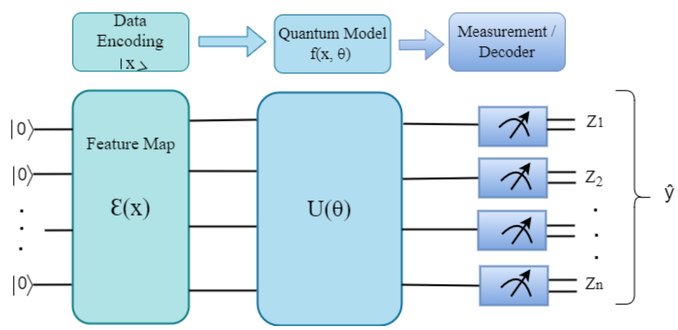

Figure 4.2: Working process of quantum model.

4.3.3.1 Quantum Layer

The quantum layer contains quantum circuits, quantum gates, and quantum registers, where the quantum parameters are used for computation. The quantum layer further contains a parameterized quantum circuit, quantum feature mapping, and a measurement circuit, as briefly discussed below.

4.3.3.1.1 Parameterized Quantum Circuit In a parameterized quantum circuit, real-time classical processing and synchronized quantum operations on quantum data are performed simultaneously. It is composed of several quantum variables, quantum gates, and a measurement circuit that uses the classical data from conventional computers under certain conditions, as shown in figure 4.4. The qubit, which has two fundamental states denoted by $|1\rangle$ and $|0\rangle$ respectively, is the basic information storage unit. Contrary to conventional bits, which can only accept one value at a time, the qubits, can be in any superposition state [280].

$$|\psi\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \in \mathbb{C}^2 \tag{4.6}$$

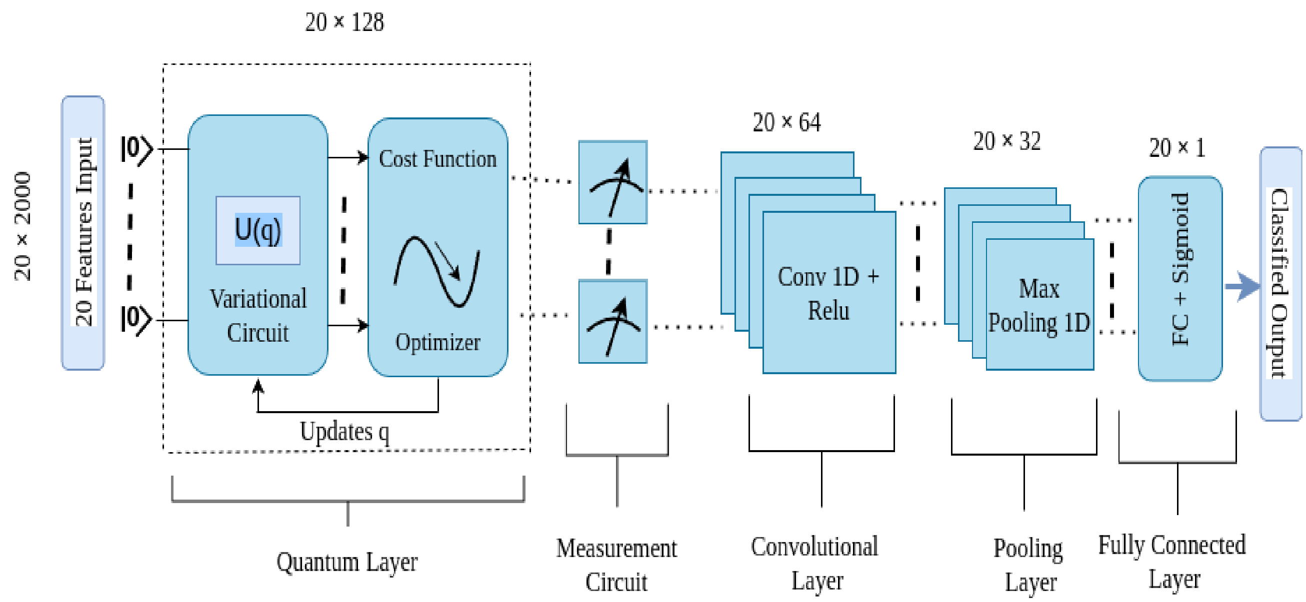

Figure 4.3: The complete FCQ-CNN model's layer-wise overview contains a quantum layer, a convolutional layer, a maxPooling layer, a fully connected layer, and a measurement circuit. The quantum layer further contains a parameterized quantum circuit, quantum feature mapping, and a cost function updated by an optimizer.

In the above equation $\left|\psi^1\right\rangle, \cdots, \left|\psi^n\right\rangle$ can be used to represent the quantum states of n qubits in a quantum circuit. The fundamental of the linear space of any n qubit in a quantum state is as, $\{|00\cdots0\rangle, |00\cdots1\rangle, |11\cdots1\rangle\}$. The quantum state $|\Psi\rangle$ in linear space can be described by using the superposition state [280].

$$|\Psi\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle, \alpha_i \in \mathbb{C} \tag{4.7}$$

where $|i\rangle$ represents the quantum state defined by the binary form of i. Quantum gates can control the states of qubits in quantum computing. A unitary transformation U, where $UU^\dagger = I$ describes the development of a closed system in quantum theory. A quantum gate that performs a unitary transition U for a quantum system in its initial state $\left|\psi_0\right\rangle = \sum_{i=0}^{2^n-1} \alpha_i |i\rangle$, acts like matrix-vector multiplication [232] .

$$U\left|\psi_0\right\rangle = U \sum_{i=0}^{2^n-1} \alpha_i |i\rangle = \sum_{i=0}^{2^n-1} \beta_i |i\rangle \tag{4.8}$$

4.3.3.1.2 Quantum Feature Mapping A quantum feature map uses a quantum circuit derived from the traditional machine learning kernel approach to encode classical data in the quantum state space. To find a separate hyperplane that classifies nonlinear data, the data are mapped onto a higher-dimensional Hilbert space. Feature

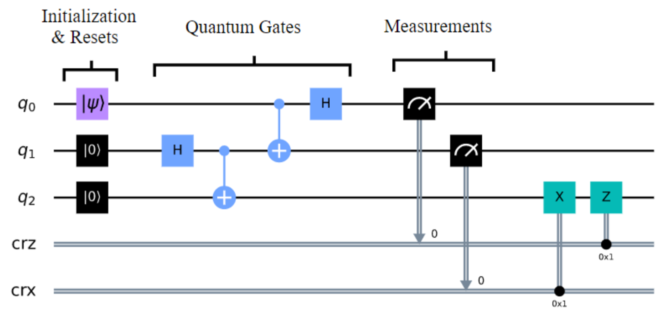

Figure 4.4: Parameterized quantum circuit.

map that encodes classical input x_i into a quantum variable $|\Psi(\mathbf{x}_i)\rangle$ by performing ground state $|0\rangle^n$ transformation using n number of unitary gates [233].

$$u_\phi(\mathbf{x}) = U_{\phi(\mathbf{x})} H^{\otimes n} U_{\phi(\mathbf{x})} H^{\otimes n} \tag{4.9}$$

$$U_{\phi(\mathbf{x})} = \exp\left(i \sum_{S \subseteq [n]} \phi s(\mathbf{x}) \prod_{i \in S} P_i \right) \tag{4.10}$$

Where the Hadamard gate is denoted by H, the subscript $U_{\phi(\mathbf{x})}$ indicates a diagonal gate (unitary gate) on the Paulifeature basis, and (P_i) is the feature space. Finding a separate hyperplane in the new space is necessary since the unitary operations in the initial state pushed the data up to the high dimension of the feature space. The size of the data determines how many qubits are necessary, and unitary gates $U_{\phi(\mathbf{x})}$ are used to encode the data by changing the angle to particular values.

4.3.3.1.3 Quantum Measurement The measurement circuit does a critical evaluation to assess the number of classes. The information about the quantum system is not immediately accessible, so a quantum measurement circuit must be used to obtain it. This procedure is analogous to taking multiple samples from the distribution of feasible cognitive base states, determining the approximate value, and then assessing the class possibilities by taking a final measurement. For instance, executing a projective measurement on a qubit with Z observable in the state $|\phi\rangle = \alpha|0\rangle + \beta|1\rangle$ leads to the generation of -1 and 1, with probabilities of $p(1) = |\alpha|^2$ and $p(-1) = |\beta|^2$. The outcome of the measurement circuit is stochastic, and with each measurement, the quantum state is converted to one of two intermediate states $|0\rangle$ and $|1\rangle$ accordingly. Repeated measurements have been taken to obtain the most precise information possible about the desired outcome. The following is an expression for the expected value of a particular measurement observable output Z of state $|\phi\rangle$ [234].

$$\langle Z \rangle_{|\phi\rangle} \equiv \langle \phi | Z | \phi \rangle = |\alpha|^2 - |\beta|^2 \tag{4.11}$$

Where $Z \equiv \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, such that $\langle Z \rangle \in [-1, 1]$ and $\langle \phi | = (|\phi\rangle)^\dagger$.

4.3.3.2 Convolutional Layer

During the learning process, a convolutional layer is usually known as the most important component of a CNN. It consists of multiple filters (or kernels), and the parameters must be learned throughout the training process. The filter size is usually smaller than the actual dimension of the input data, and the activation map is created by the convolution of filters with the input volume. The convalutional layer output volume is achieved by overlaying the depth dimension along with each filter activation map. On the other side, the layer defined by the quantum convolution layer has properties similar to those of the classical convolution layer. The quantum convolution layer works with filters in the same way that the traditional convolution layer does. It applies a filter to the data that comes in and makes feature maps out of new data.

For an instance, consider n number of layers l, if X^l is the input, K^l is the kernel size, and $f:\mathbb{R}\mapsto[0,C]$ with $C>0$ is a nonlinear function, then the output for l would be $f\left(X^{\ell+1}\right)=f\left(X^{\ell}*K^{\ell}\right)$. The input layer X^l and kernel size K^l stored in the QRAM, and the quantum algorithm is used to obtained a quantum state $\left|f\left(\bar{X}^{\ell+1}\right)\right\rangle$, for $\left\|f\left(\bar{X}^{\ell+1}\right)-f\left(X^{\ell+1}\right)\right\|_{\infty}\leq 2M\epsilon$ for each precision parameter $\epsilon>0$ and $\eta>0$, and can be defined as follows:

$$\begin{cases} \left|\mathscr{X}_j^{\ell+1}-f\left(X_j^{\ell+1}\right)\right|\leq 2\epsilon & \text{if} \quad f\left(\bar{X}_j^{\ell+1}\right)\geq\eta \\ \mathscr{X}_j^{\ell+1}=0 \quad \text{if} \quad f\left(\bar{X}_j^{\ell+1}\right)<\eta \end{cases} \tag{4.12}$$

The execution time of the algorithm for each iteration may be,

$$\widetilde{O}\left(\frac{1}{\epsilon\eta^2}\cdot\frac{M\sqrt{C}}{\sqrt{\mathbb{E}\left(f\left(\bar{X}^{\ell+1}\right)\right)}}\right) \tag{4.13}$$

The above $\mathbb{E}\left(f\left(\bar{X}^{\ell+1}\right)\right)$ is the average value of the quantum state, and $\widetilde{O}$ is the polylogarithmic hide factor that depends on input and kernel size. Furthermore, the kernel size is determined by the Polly logarithmic factor, which allows the QCNN to work with deeper kernels. In forward propagation, the loss function $\mathscr{L}$ is determined by the output of QCNN. For binary classification, the Binary Cross-Entropy has been considered a loss function and may be calculated based on the formula shown below.

$$\mathscr{L}=-\frac{1}{m}\sum_{i}^{m}\left(y_i*\log(p(y_i))+(1-y_i)*\log(1-p(y_i))\right) \tag{4.14}$$

Here, m shows the number of samples starting from i, the probability of true samples, or class 1 is represented by $p(y_i)$ and $p(1-y_i)$ indicates the probability for class 0 respectively. In the case of back-propagation, a quantum circuit can be successfully characterized with a fully connected quantum neural network that has symbolic similarity with the classical neural network, except that the activation function is not assigned to all nodes. Furthermore, the quantum circuit's input may be one of the

2^n (where n is the number of qubits) having an amplitude of 1. Like traditional deep learning, minimizing error back-propagation is also important for quantum circuits. The gradient of the weights is determined via the back-propagation method by using parameter shift and the chain rule [235] [236] of a partial differential to propagate the gradient back from the network output. Back propagation updates the weights in a manner that reduces the loss by providing the nodes with lower error rates and higher weights, and vice versa. This allows it to learn how much of the overall loss each node is accountable for contributing. The quantum gates and quantum state $|\psi\rangle$ are represented by complex values for error back-propagation in the simulation of quantum computing. The output $|\psi_{out}\rangle$ may be stated as follows when $|\psi_{in}\rangle$ is the input for n qubits and $W(\theta)$ is the applied quantum circuit parameter network [232].

$$W(\theta)\left|\psi_{\text{in}}\right\rangle = \sum_{j=0}^{2^n-1} c_\theta^j |j\rangle = \left|\psi_{\text{out}}\right\rangle \tag{4.15}$$

In above, the state $|j\rangle$ amplitude probability is denoted by c_θ^j and $\left|c_\theta^j\right|^2 = p_\theta^j$ shows the observation probability of state $|j\rangle$.

4.3.3.3 Pooling Layer

Pooling layers seeks to continuously reduce the dimensionality of the representation, which reduces the computational cost and number of parameters for the model. In input, the pooling layer scales the dimension of every activation map by using the "MAX" function. The majority of CNNs apply 2-by-2-dimensional kernels in steps of 2 along the spatial dimensions of the input using max-pooling layers. This keeps the depth volume at its actual size while scaling the activation map down to 25% of its original size. Due to the pooling layer's destructive characteristics, there are two frequently utilized methods for max-pooling. The stride and filters of the pooling layer are usually set to 3 and 3, respectively, allowing the layer to extend throughout the maximum range of input spatial dimensions. Additionally, overlapping pooling may be used with a stride and kernel size of 2 and 3, respectively. Because pooling is destructive, if the kernel size is greater than 3, the model's performance will usually go down.

4.3.3.4 Fully Connected Layer

The output from the last pooling or convolutional layer is passed through to the fully connected layer after being flattened. The fully connected layer is composed of a collection of dependent nonlinear functions, each of which is made up of a perceptron or neuron. The neurons perform a linear transformation on the input vector using a weight matrix. After the several convolutional and pooling layers in a quantum environment, the amount of qubits has decreased with the implementation of the FC layer.

The desired qubit circuit has been used in the case of a relatively small system to accomplish the classification. The highly entangling circuits, which include CNOT gates and universal single-qubit gates with basic cells composed of $n(n+1)$ and $n(n-1)$, are considered for the FC layer. Each CNOT has a variable indicating whether it is active or inactive and working as a unitary gate [290]. The two-qubit universal gate operation is shown in figure 4.5.

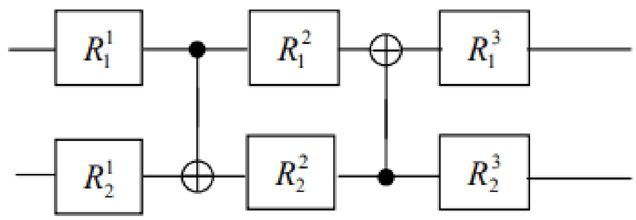

Figure 4.5: Two qubits universal gate operation.

To obtain the prediction values, the FCQ-CNN model last step involves measuring a certain number of output qubits. The measurement circuit converts the quantum variables $|0\rangle$ and $|1\rangle$ back into binary variables, 0 and 1. The mapping of expected values to classification outputs may be carried out in various ways. Formally, the FCQ-CNN model's output for the input xin is expressed as $f(\boldsymbol{\theta}, \boldsymbol{x}_{in})$. In general, measuring one qubit and using the expected value as the output $f(\boldsymbol{\theta}, \boldsymbol{x}_{in}) \equiv \langle Z \rangle$ is a easy way for binary classification tasks. Then, as a result, $\langle Z \rangle \geq 0$ represents the classification of samples into class 1, and $\langle Z \rangle < 0$ denotes classifying the sample into class 0.

4.4 Experimental Results

This section contains the experimental results that demonstrate how the actual quantum convolutional neural network performs when applied to classify diseases using the cardiovascular dataset. The TensorFlow Quantum platform is used for the simulation of the proposed model. Initially, there are 42 attributes in the cardiovascular dataset that are associated with IHD. However, due to resource constraints faced by the quantum computing simulation environment, the features are limited to 20 features, and the total number of samples considered for the desired experiment is 2000. The Torch interface, Qiskit, Keras, and Cirq libraries are used to implement the models in the Python 3 environment. An application programming interface (API) is used to run the proposed FCQ-CNN model on the IBM state-vector simulator. The features are encoded by using a PauliFeature map, where the layer of the Hadamard gate along with entangling blocks maps the classical data into quantum variables. The quantum circuit is rotated with a depth of 10, and a measurement circuit is used to decode the data again into classical 0,1. All models use a Relu activation function for the input and hidden layers, while a Sigmoid function is used for binary classification on the output layer.

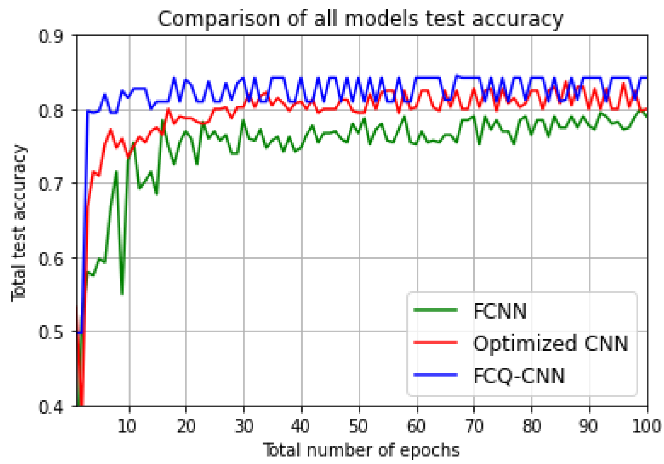

Figure 4.6: Comparison of testing accuracy of all models.

4.4.1 Hyperparameter Tuning

It is possible for neural networks to learn complex correlations among their inputs and outputs. Although many of these correlations will be visible during training but not in actual test data because of sampling noise. This problem could result in overfitting, which would decrease the model's capabilities for prediction. In short, it is quite difficult to avoid the overfitting issue and obtain the generalized prediction of the proposed model. For this work, I applied several optimization techniques for hyperparameters to get the best possible set. The input dataset is first normalized by using min-max and standard scalars, which scale the data in a defined range of 1 and 0. The data are then encoded using different feature maps, such as the PauliFeature map, ZFeature map, and ZZFeature map, with circuit depths of 5, 10, and 15, where the PauliFeature map with a circuit depth of 10 was found to be the best for our model. The model has been evaluated by using various loss functions, and for our binary classification task, Binary Cross-Entropy produced promising results. The effective epochs for the FCNN, Optimized-CNN, and FCQ-CNN models are 100, 60, and 20, respectively. This is because beyond a certain point, the models' accuracy does not increase and remains constant, as shown in figure 4.6. During the training process, Adam (adaptive optimization techniques) is used for 100 epochs, which outperformed stochastic gradient descent (SGD) in terms of optimal performance. The model is initially run using three different learning rates (0.01, 0.001, and 0.0001) with separate batch sizes (16, 32, and 64), and it gets good results by considering learning rate 0.001 and batch size 32. The number of filters for each layer is 128, 64, and 32, respectively, where each filter size is 3. The input and hidden layers are equipped with the Relu activation function, and the

Table 4.4: Set of optimal hyperparameters of each model.

Hyperparameters	FCNN model	Optimized CNN model	Proposed model
Learning rate	0.001	0.001	0.001
Optimizer	Adam	Adam	Adam
Batch size	32	32	32
No of epochs	100	60	20
Loss function	Binary crossentropy	–	–
Activation function for dense layer	Relu	Relu	Relu
Activation function for output layer	Sigmoid	Sigmoid	Sigmoid

output layer uses the Sigmoid function. The model is trained with the training dataset, the weights with the lowest training loss are saved, and the prediction performance on the labeled test dataset is assessed. Finally, the desired optimal weights are used for class prediction of the test dataset. Table 4.4 indicates the optimal hyperparameter values used for each model.

4.4.2 Comparison of FCQ-CNN model with FCNN and Optimized-CNN Models

The competency of the proposed FCQ-CNN model is obtained by comparing it with the classically optimized-CNN and FCNN models by considering the cardiovascular dataset in terms of testing accuracy and testing loss, respectively, as shown in table 4.5. The classification accuracy of each model is obtained by considering the same hyperparameter optimization, which also ensures a fair comparison. It is important to keep a small number of nodes in the input layer of all models to be trained properly with a limited number of parameters. The FCNN model contains one input layer with 20 nodes, two dense layers with 32 and 16 nodes, and an output layer with 1 node for binary classification, respectively. Likewise, the optimized-CNN contains the input layer, two convolutional layers, and two maxpooling layers with the same number of nodes. In the case of FCQ-CNN, after the quantum layer there is a convolutional layer, a maxpooling layer, and a fully connected layer. The input layer, the hidden layer, and the convolutional layer of all models use the Relu activation functions, while the output layer uses the sigmoid function, which converts any input to an output between 0 and 1 for the classification of cardiovascular diseases. The output of the model sigmoid predicts one value for class 0, if $\hat{y} < 0.5$ and for class 1, if $\hat{y} > 0.5$, where $\hat{y}$ is the predicted output of the model. The model is trained with 1600 (80%) samples out of a total of 2000 samples, and 400 (20%) samples are used for testing purposes. An Adam optimizer is used, where the learning rate is considered to be 0.001 for each model. Figure 4.6

shows the testing accuracy of FCNN, Optimized-CNN, and FCQ-CNN models, where the highest testing accuracy achieved by FCQ-CNN is 84.6%, Optimized-CNN accuracy is 80.4% and FCNN has 79.2% accuracy, respectively. Likewise, the calculated loss comparison is indicated in figure 4.7, which also shows that the proposed FCQ-CNN model has the lowest training loss and has a value of 0.24, the optimized-CNN has a loss value of 0.3 and the FCNN model has a calculated loss of 0.4 respectively. In the comparison of the Optimized-CNN and FCNN models, the reason for the highest accuracy and low error rate of Optimized-CNN is due to the selection of the best subset of hyper parameters. Moreover, the Optimized-CNN model seems more robust than the FCNN model because of its shared-weights architecture, comparatively small set of parameters, and characteristics of translation invariance for binary classification problems. In the FCNN model, each node in the dense and flattened layers is connected in the network by utilizing the most trainable parameters, which reduces the model's performance and increases its computational cost. On the other hand, the FCQ-CNN has the highest test accuracy and lowest test loss of the two models (optimized-CNN and FCNN). The FCQ-CNN model loss seems more stable after 30 epochs, this is due to the quantum network's quick learning convergence, which requires a comparatively smaller iteration to achieve the flattening of the loss function and accuracy. The testing and training times cannot be compared, because simulating a quantum network naturally takes a long time and demands the numerical equivalent of quantum computation. In the end, the quantum algorithm does a better job at IHD classification tasks than a classical algorithm and offering advantages for a large amount of data.

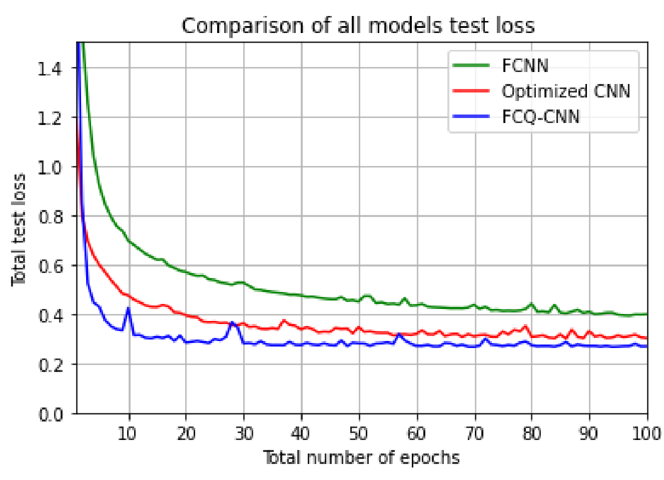

Figure 4.7: Comparison of testing loss of all models.

two quan

Table 4.5: Comparisons of FCQ-CNN with Optimized-CNN and FCNN models.

Parameters	FCNN model	Optimized-CNN model	Proposed model
Dataset	Cardiovascular	Cardiovascular	Cardiovascular
Total samples	2000	2000	2000
Total test samples	400	400	400
Total train samples	1600	1600	1600
No of layers	4	5	5
Testing accuracy	79.2%	80.4%	**84.6%**
Testing loss	0.4	0.3	**0.28**

Table 4.6: Comparisons of FCQ-CNN with previous published quantum models.

Parameters	Maheshwari *et.al*	PennyLane team *et.al*	Yano *at.al*	Kerenidis *at.al*	Proposed model
Dataset	Synthetic, Sonar, Diabetes	Iris	Breast-Cancer, Heart Disease	MNIST	Cardiovascular Disease
Used Model	QSVM, VQC	QRF	VQC, VQC+QRAC	QCNN	FCQ-CNN
Total features	10, 10,13	10	9, 13	8	20
Number of classes	2,2,2	3	2, 2	2	2
Training samples	400, 166, 2121	125	146, 243	60,000	1600
Testing samples	100, 42, 531	25	50, 60	10,000	400
Total size	500, 208, 2651	150	196, 303	70,000	2000
Model accuracy	94%, 75%, 76%, 71%,-74.5%, 69%	72%	69%-73%, 82%, 85%	82.8%	**84.6%**

4.5 Discussion

The proposed FCQ-CNN is also compared with recently published quantum models, such as the Quantum Support Vector Machine (QSVM), the Quantum Random Forest (QRF), the Quantum K-Nearest Neighbor (QKNN), and the Quantum Convolutional Neural Network (QCNN). The table 4.6 shows that the proposed FCQ-CNN is more competent than the existing quantum models in terms of high test accuracy, least complexity, and low computational loss. Numerous studies on quantum-enhanced methods have been carried out to tackle issues with machine learning. For an instant [237], presented various QML approaches, i.e., QSVM, VQC, and amplitude-encoding VQC, by considering the three different datasets. The highest accuracy for the UCI diabetes dataset is achieved at 74.1%, 68.7%, and 74.5% for QSVM, VQC, and amplitude-encoding VQC, respectively. The algorithms are executed by using the IBM simulator and the Qiskit framework [214]. Similarly, comparing our results to those of a literature review of the ensemble learning approach proposed by the Pennylane project team [238]. It uses two different Quantum Processing Units (QPUs), forest.qvm and qiskit.aer, each working as a separate cluster of trees. For identifying three different groups, 150 samples from the iris dataset were employed. A PauliFeature map is used to convert the classical data, and RX gate rotation is used to determine the depth of the circuit. For both the training and testing datasets, the ensemble model predicted results are recorded as 83.2 percent and 72%, respectively. The model's reliability is verified by comparing it with YANO *et.al* [239], which exploits quantum random access coding to effectively map such discrete features into a finite amount of qubits for VQC. The goal of the work is to show that QRAC can speed up VQC training by lowering its settings and minimizing the number of qubits required for the mapping. The models are evaluated using UCI breast cancer and heart disease data with 286 instances and 9 features and 303 instances and 13 features, respectively. The models work well and obtained the maximum accuracy of 72% and 88% for VQC and VQC-QRAC for both datasets. The quantum CNN completely replicates the classical CNN by supporting non-linearities and pooling operations. A new quantum approach with ℓ_∞ norm and new quantitative research algorithms in the context of information processing is introduced by Kerenidis *et.al* [240]. For classification, the MNIST datasets' classification using numerical simulations has been considered, which demonstrates the effectiveness of the QCNN in a real-world scenario. The model achieved the highest test accuracy 82.8% and 74.8% and error rates of 0.51 and 0.77 with $\ell = 0.01, 0.1$ respectively.

In summary, our proposed FCQ-CNN model offers significant results on the cardiovascular dataset and is competitive with recently published different quantum machine learning models [237], [238], [239] and [240].

4.6 Conclusions

Quantum machine learning and data science have interesting future uses thanks to fully parameterized quantum convolutional neural networks. In this paper, a comprehensive benchmark of the FCQ-CNN is presented for handling the classification tasks on classical data. The FCQ-CNN algorithm can be customized using a variety of factors, including the design of parameterized quantum circuits, quantum filters, pooling operators, classical data pre-processing techniques, quantum feature mapping, optimizer, and cost functions. The data used for this experiment comes from public hospital and is based on healthcare records relating to heart disease. In the first stage, the data is cleaned by eliminating any redundant and irrelevant variables, nullifying inoperative variables, converting data types, and handling missing values. Combining a support vector machine (SVM) with recursive feature elimination (RFE) as a wrapper, and a maximally relevant minimal redundancy (MRMR) filter, the important 20 features are selected. The input classical data has been converted into quantum variables using quantum feature mapping (PauliFeature map). Using the IBM state-vector simulator, the model is tested with 20 qubits of QCNN for the binary classification of cardiovascular datasets, which reduces the number of free parameters. The generalized FCQ-CNN architecture put forward in this study has an input layer with 20 nodes and two quantum filters. For the input layer and convolutional layer, a Relu activation function is used, while the output layer has a sigmoid function to accomplish binary classification. A measurement circuit is used at the output layer to decode the quantum variables. The model has been tested using several optimizers and learning rates, and it produced statistically significant results using the Adam optimizer and a learning rate of 0.001. Despite the limited number of free parameters, the FCQ-CNN model achieved the highest testing accuracy of 84.6% in all instances for cardiovascular datasets. The proposed model results were also compared to the optimized CNN and FCNN models, and it was noticed that FCQ-CNN performed slightly better than both classical models under the same training conditions for the same data samples.

The primary limitation of this work is the use of a small number of data samples compared to a small number of qubits. The quantum computing device requires more logic gates to use larger datasets and more qubits, which increases the computational cost and prolongs the model execution time. These limitations might have an impact on quantum states, where a wrong rotation might cause an error in the result. In the future, a universally fault-tolerant quantum computer will be needed to efficiently handle issues like integer factorization and unstructured database search. This computer will need millions of low-error, long-coherence-time qubits.

Chapter 5

Case Study III: Sarcopenia Disease

5.1 Introduction

Sarcopenia is considered to be a dangerous health hazard in terms of social, clinical, and economic aspects. It is characterized by decreased muscle mass and function as a result of nutritional deficiencies and decreased physical activity and is more common in elderly people [241]. The word "sarcopenia," derives from the Greek words "poverty of flesh," was first used by Irwin Rosenberg in 1989 to describe an age-related loss of lean body mass that affects independence, nutritional condition, and mobility [242]. Sarcopenia is becoming more common, which is a result of global population ageing [243]. As reported by the WHO, around 600 million people over the age of 60 were estimated to exist in 2000, and by 2025, this number is expected to rise 1200 million [244]. It also causes a variety of senile decays, including falls, physical disabilities, depression, fractures, poor quality of life, and even death [245]. Sarcopenia has recently been identified as a distinct disease entity and is becoming more common as the world's population increases. According to numerous studies, sarcopenia is related to a variety of chronic and metabolic diseases [246]. Muscle mass falls as levels of physical activity drop for a variety of reasons, eventually resulting in a vicious cycle of sarcopenia-related disease [247]. Therefore, despite addressing other conditions that may induce sarcopenia, a sufficient diet with high protein content, physical activity, and anti-inflammatory drugs are needed. However, the changes in physical activity are asymptomatic, and there is potential for early screening. Therefore, in the framework of preventative and predictive medicine, developing an efficient sarcopenia screening tool for the elderly population is essential to raise the chance of early identification and treatment [248].

Despite the significant clinical importance of sarcopenia, it is still misdiagnosed and poorly treated. Currently, there are no specific diagnostics toll available for screening it, due to the incomplete understanding of the biology of sarcopenia [249]. An earlier observational study revealed that sarcopenia may be correlated with inflammatory indicators such as tumor necrosis $-\alpha$ factor, interleukin-6, and reactive protein [250]. Recent research has raised the possibility that several biomarkers connected to skeletal muscle alterations and neuromuscular junctions may be used to predict the development of sarcopenia [251]. These indicators must be collected by intrusive blood testing, and they are not sufficiently accurate to identify sarcopenia. The majority of patients with cachexia will suffer from the loss of muscle mass and strength typical of sarcopenia, however, sarcopenia patients, such as those with sarcopenic obesity, may not experience cachexia [252]. Frailty occurs from cumulative deficiencies reducing total functional reserve, which can cause hospitalizations, functional reliance falls, weight loss, and other negative effects [253] [254].

Currently, the conventional techniques are used to measure muscle strength and mass including computed tomography (CT), Dual Energy X-ray Absorptiometry (DEXA), and Magnetic Resonance Imaging (MRI) [255–257] which is insufficient for diagnosing sarcopenia. Despite their extensive applications, these clinical exams appear to be underutilized in this environment due to the expensive equipment and lack of accessibility. Additionally, highly skilled medical staff are needed to operate the devices/equipment used in these kinds of clinical tests. Moreover, to overcome this restriction, EHRs [127] have gained popularity recently, enhancing the accuracy of prediction essential for sarcopenia analysis. There are significant possibilities for both epidemiological and clinical research using these routinely created longitudinal databases [258]. It includes details about individual patient health status and is an electronic representation of their medical record [259]. Because of its enormous size, EHR generally have great statistical significance and are used to classify the population. The reliability of the data may be further improved by linking several EHR parameters, which poses a serious obstacle in the healthcare research application. Therefore, it is possible for data elements that could be important for research to be misclassified, inadequately described, or absent, which might result in systematic quantitative error [260].

The capacity of machine learning is quickly approaching the limit of classical computers because of the continuous growth of data. One of the most common computing models for modulating quantum information for computation is quantum computing, which is dependent on the law of quantum mechanics [261]. Due to quantum computing, the field of QML has been investigated. QML is sufficient to solve a number of challenging issues with classical computers because of quantum phenomena, i.e, quantum entanglement, tunneling, quantum superposition [262] [263]. The inclusion of quantum computing into conventional machine learning can increase computational efficacy and decrease complexity [264]. Many quantum-enhanced machine learning algorithms have recently been introduced to accelerate some particular ML tasks [265] [266]. Accurate and quick diagnosing and precise treatment are essential to providing high-quality of care, and QML offers exceptional processing speed [267]. Quantum computing enables healthcare professionals to find correlations, offer diagnoses, and suggest viable therapies [268]. The technology makes it possible to analyze massive amounts of clinical information and provide patient centered care [269]. The

Table 5.1: Physical measurement, type, level and code used in the dataset.

Features	Type	Level	Code
Mobility-score	Normal Mild cognitive impairment Moderate impairment Serious impairment	≥ 24 23-19 14-18 ≤ 14	1 2 3 3
Mini-Nutritional	Normal Malnutritional-risk Malnutrition	24-30 $17.5-23$ ≤ 17	1 2 3
Lawton	Total-dependency Severe-dependency Moderate-dependency Mild-dependency Independent	0-1 2-3 4-5 6-7 8	1 2 3 4 5
Norton	Normal Obvious-risk High-risk	> 14 ≤ 14 ≤ 12	1 2 3
Charlson (Predicts mortality risk)	Low-risk High-risk	< 3 ≥ 3	1 2
Barthel	Independent Mild-DEP DEP-MOD DEP-SEV Total-DEP	100 91-99 61-90 21-60 0-20	5 4 3 2 1

medical industry is being revolutionized by quantum computing, which reduces the cost of uncertain treatments, where a single molecules can be seen using a very accurate representation produced by quantum imaging technology [270].

5.2 Materials

This study is based on data from the geriatrics department of the Tijuana General Hospital and focuses on the prevalence of sarcopenia in elderly people. In 2017, there were 85,259 older residents (aged between 65 and 90 years) in Tijuana, where 65% of them received medical care at the hospital. The hospital served people with limited financial resources [271]. In the descriptive observational research, a group of adults suffering

Table 5.2: The desired risk factor for both sarcopenia and chronic diseases.

Risk factor for sarcopenia	Risk factor for chronic diseases
Physical inactivity	Chronic pain
Low weight by birth	Diabetes mellitus
Female gender	Mood disorder
Genetic predisposition	Heart failure
Lifestyle	Liver failure
Constitutional	Cognitive impairment
Weightlessness	Chronic inflammatory diseases
Low protein intake	Short breathing
Being bedridden	Cancer
Malnutrition	Kidney failure
Inanition	Catabolic effect of drugs
Smoking habit	Osteoarthritis
Living condition	Obesity

from mild to severe sarcopenia were assessed. The required sampling size was determined by factoring in the assumed prevalence by using bioimpedance assessment in older people, which was 17%, and assuming a 5% error margin and 95% confidence range. These parameters determined that 166 individuals were required to achieve the desired outcomes [272]. The dataset contains information on patients from various areas of Baja California, such as Rosarito, Tecate, Tijuana, and Ensenada. Patients who were willing to participate in the investigation and have been diagnosed with mild to severe acute sarcopenia signed an informed consent. The research did not include any patients who were reliant on either psychological or physical huddle. Elderly people do not engage in any physical exercise, which is referred to as being "sedentary. At the Tijuana General Hospital, geriatricians made the sarcopenia diagnosis based on the medical histories of patients. The study used patient electronic health histories while taking into account a variety of factors, including, age, (mini nutritional assessment (MNA) validated test), functional capacity (Lawton and Barthel), socio-demographic and biochemical information, pharmacology, malnutrition, psychological tests, and the weighted body index with specific comorbidity (Charlson). The risk factors for sarcopenia and other chronic illnesses are shown in table 5.2.

Initially, the dataset contained 251 patients EHRs, considering 84 different attributes. Both genders are taken into account, where total womens are 180 (65%) and total mens are 71 (35%) of the total dataset, where the average age of each individual is considered 79 years as shown in table 5.3. Machine learning algorithms which predict the patient's sarcopenia level utilizing the available data about them, were applied to determine the level of sarcopenia in these adults. The physical measurement, type, level and code used in the data set are briefly explained in table 5.1.

Table 5.3: Assessment criteria for both genders.

Assessment	Females	Males
Gender	65%	35%
IMM (Immune mediated myositis)	< 6.1 kg/m^2	< 8.5 kg/m^2
Hand grip strength (lbf)	< 20	< 30
Walking speed (m/s)	< 0.8	< 0.8

5.3 Methods

The method section contains steps for cleaning the raw dataset, important feature selection techniques, and quantum state preparation, and QML models, i.e, QKNN, AE-VQC as briefly discussed below.

5.3.1 Data Cleaning

The dataset cleaning involves a set of procedures that are used to ensure that the data is accurate, consistent, and complete as shown in figure 5.1. Here are the general steps involved in the cleaning process for our sarcopenia dataset.

- Inspect the data: This step involves examining the data to understand its structure, content, and quality. This will help you identify any issues that need to be addressed.
- Handle missing data: Missing values are handled by removing the rows or columns with null values.
- Remove duplicates: All the redundant and duplicated information has been removed.
- Handle inconsistencies: Check the inconsistencies in the data, such as typos error, or inconsistent formatting. Standardize the data to a common format to ensure consistency.
- Normalize the data: The data has been normalized to make all the elements lie between 0 and 1 thus bringing all the values of numeric columns in the dataset to a common scale.

5.3.2 Feature Selection

The important features have been selected by using Random Forest (RF) with Recursive Feature Elimination (RFE) and Logistic Regression (LR) with RFE and finally, an intersection has been used to select the three most important subsets of features as

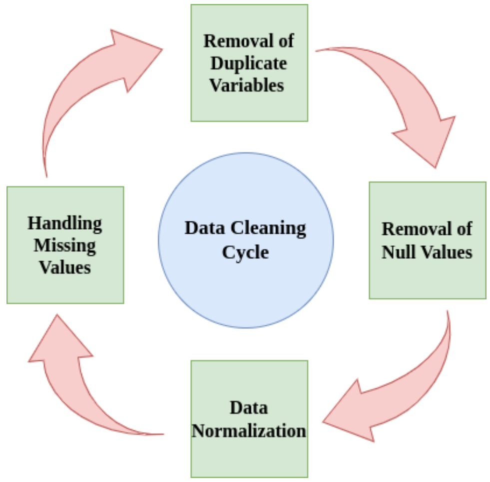

Figure 5.1: Data cleaning process.

shown in figure.5.2. The RF classifier generates a large number of decision trees to rank the significance of each predictor that makes up a model. The best predictor is selected and split by each node of a tree, which considers a different subset of randomly chosen predictors. The best predictor is determined by lowering the impurity of the node and measuring the estimated response variance. In the range implementation of RF, this criteria is regarded as the default technique applied to regression trees. Each tree is created using a randomly selected bootstrap sample, which is used as a training set to predict the data in the second sample of testing set. This sample is made up of around two-thirds of the total observations. The predictions for each variable are combined across all trees, and the remaining samples' mean square error (MSE) is calculated. To assess each RF performance, the remaining samples MSE and percentage of variance are used. I have compared the importance scores obtained after running RF once and after running it repeatedly in order to determine whether RF with RFE was superior to RF alone. The RFE, as described by Guyon et al [273], is essentially a backward selection of the predictors. It computes a relevance score for each predictor by creating a model using the whole set of predictors. The RF-RFE approach involved using RF to establish the initial significance scores, recursively deleting the variables with the lowest importance scores from the data set, then ranking the remaining variables in order of highest importance score accordingly. After removing the least significant predictors, the model is rebuilt, and importance scores are calculated again. In reality, this determines the size and number of each feature subset to be evaluated. Based on the importance rankings as show in figure 5.3, the important features are chosen using the subset size that optimizes the performance requirements.

In parallel, logistic regression with RFE is also considered for important feature selection. The goal of logistic regression is to determine the probability that the desired observation belongs to a specific class. In logistic regression, it's important to choose all the desired features before attempting to predict the target variable. These

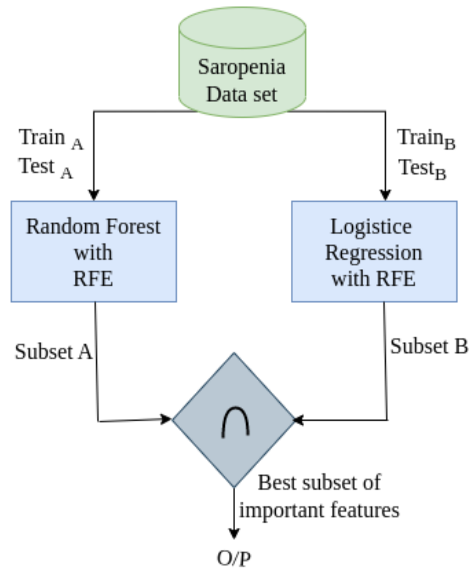

Figure 5.2: Important feature selection process.

Table 5.4: Dataset before and after pre-processing.

Parameters	Raw dataset	Clean dataset	Balanced dataset
Total samples	251	231	162
Features	53	41	41
DataType	Categorical	Categorical	Categorical
Output	Binary	Binary	Binary

methods involve iteratively fitting the model with different subsets of the input features and evaluating the performance of the model on a validation set. The features that contribute the least to the model's performance are then removed, and the process is repeated until a desired level of performance or a specified number of features is achieved. Finally, the intersection strategy has been used, which makes the use of common features from each of the given individual classifiers. Both the classifiers extracted the three important feature subsets,8 features, 16 features, and 32 features as shown in figure.5.4, figure.5.5 and figure 5.6 respectively are further used to train the final models.

5.3.3 Quantum State Preparation

Quantum state preparation is the process of creating a specific quantum state in a quantum system. A quantum state can be described by a mathematical object called a wave function or state vector, denoted as $|\Psi\rangle$. The state vector contains all the information about the quantum system, including the probabilities of measuring different

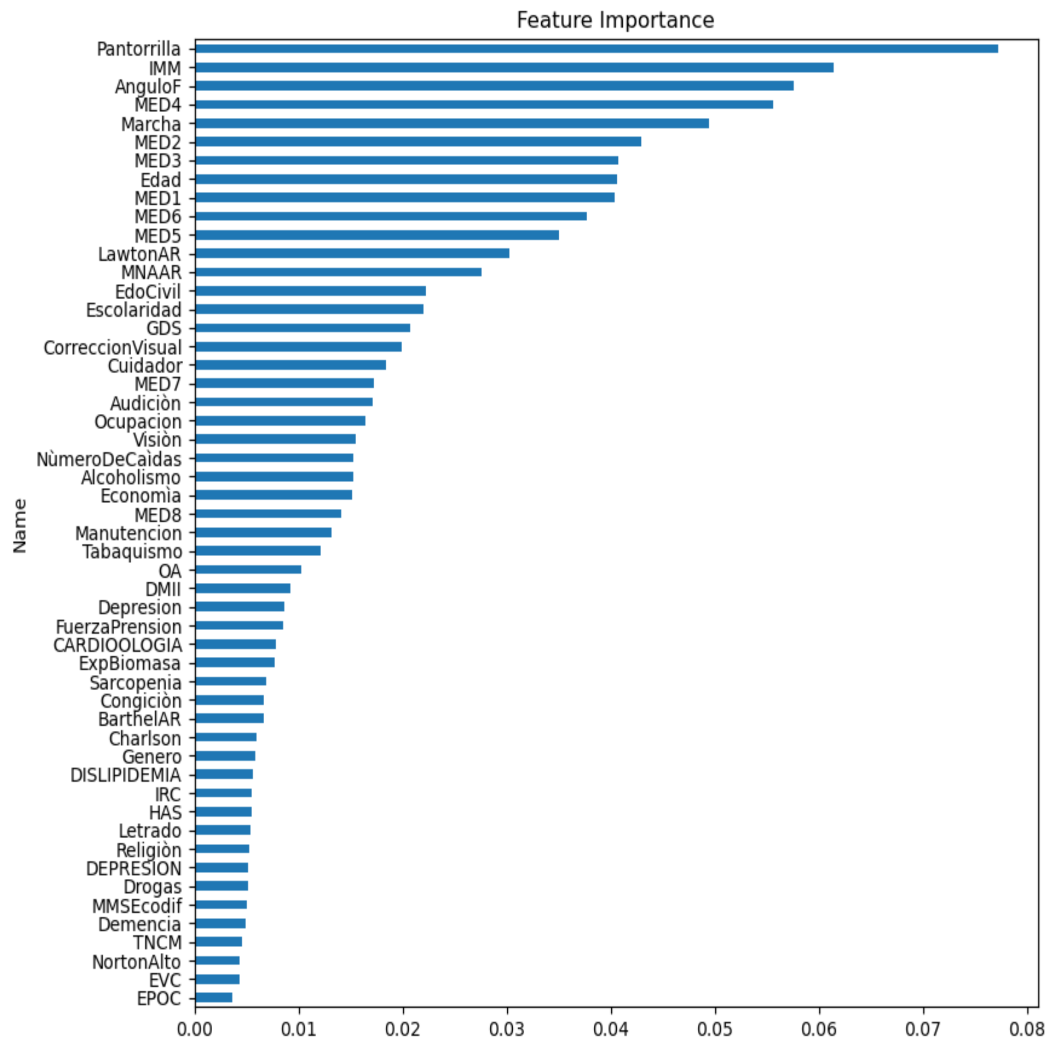

Figure 5.3: Features importance criteria for whole dataset.

outcomes from a measurement. Quantum state preparation involves choosing the appropriate basis states and coefficients to create a specific quantum state. This process can be achieved through the application of quantum gates and other quantum operations. The general equation for preparing a quantum state is give as below [274],

$$|\Psi\rangle = \sum c_i |\Psi_i\rangle \tag{5.1}$$

where c_i is the complex coefficients and $|\psi_i\rangle$ indicates the input data into the quantum system. The basis states represent the possible states that the system can be in, and they form a complete orthonormal state. To prepare a specific quantum state, I choose the appropriate basis states and coefficients. Lets suppose, to prepare a qubit in the state [274],

$$|\Psi\rangle = |0\rangle + |1\rangle \tag{5.2}$$

Name	Features	feat_importances	Technique
Pantorrilla	46	0.077199	Intersected
IMM	43	0.061371	Intersected
AnguloF	44	0.057542	Intersected
MED4	25	0.055575	Intersected
Marcha	47	0.049417	Intersected
MED2	23	0.042972	Intersected
MED3	24	0.040765	Intersected
Edad	1	0.040638	Intersected

Figure 5.4: Selected 8 important features

Name	Features	feat_importances	Technique
Pantorrilla	46	0.077199	Intersected
IMM	43	0.061371	Intersected
AnguloF	44	0.057542	Intersected
MED4	25	0.055575	Intersected
Marcha	47	0.049417	Intersected
MED2	23	0.042972	Intersected
MED3	24	0.040765	Intersected
Edad	1	0.040638	Intersected
MED1	22	0.040359	Intersected
MED6	27	0.037677	Intersected
MED5	26	0.035013	Intersected
LawtonAR	40	0.030236	Intersected
MNAAR	41	0.027541	Intersected
EdoCivil	4	0.022261	Intersected
Escolaridad	2	0.021990	Intersected
GDS	35	0.020746	Intersected

Figure 5.5: Selected 16 important features.

where $|0\rangle$ and $|1\rangle$ are the basis states of the qubit. This state is called the superposition state of $|0\rangle$ and $|1\rangle$. To prepare this state, a need to apply a quantum gate called the Hadamard gate to the initial state $|0\rangle$. The Hadamard gate is defined as, which is considered the desired state of $|\Psi\rangle$ [275],

$$\mathrm{H}|0\rangle = 1/\sqrt{2}|0\rangle + 1/\sqrt{2}|1\rangle \tag{5.3}$$

In order to analyze data for QML models, it is important to convert classical data into quantum data. This can be done by representing the classical data as training data in a quantum format as [275],

$$F_n = \{(|\psi_1\rangle, y_1), \dots (|\psi_i\rangle, y_i), \dots, (|\psi_n\rangle, y_n)\} \tag{5.4}$$

The above $|\psi_i\rangle$ is the quantum state for label y_i data stratification of $F_n, |\psi_i\rangle \in C^{2^d}$ and $y_i \in \{c_1, c_2\}$. Multiple techniques exist for converting classical data into quantum data in high dimensions, such as basis encoding and amplitude encoding. These methods allow the classical data to be embedded into the quantum data format.

5.3.3.1 Basis Encoding

The most common method for encoding classical data into a quantum state is basis encoding. It is used to convert n bit binary string x into a n qubit quantum state $|x\rangle = |i_x\rangle$, where $|i_x\rangle$ indicates the computational basis state. The approach establishes

Name	Features	feat_importances	Technique
Pantorrilla	46	0.077199	Intersected
IMM	43	0.061371	Intersected
AnguloF	44	0.057542	Intersected
MED4	25	0.055575	Intersected
Marcha	47	0.049417	Intersected
MED2	23	0.042972	Intersected
MED3	24	0.040765	Intersected
Edad	1	0.040638	Intersected
MED1	22	0.040359	Intersected
MED6	27	0.037677	Intersected
MED5	26	0.035013	Intersected
LawtonAR	40	0.030236	Intersected
MNAAR	41	0.027541	Intersected
EdoCivil	4	0.022261	Intersected
Escolaridad	2	0.021990	Intersected
GDS	35	0.020746	Intersected
CorreccionVisual	11	0.019965	Intersected
Cuidador	5	0.018424	Intersected
MED7	28	0.017224	Intersected
Audiciòn	12	0.017106	Intersected
Ocupacion	7	0.016451	Intersected
Visiòn	10	0.015453	Intersected
NùmeroDeCaìdas	38	0.015333	Intersected
Alcoholismo	31	0.015275	Intersected
Economìa	8	0.015126	Intersected
MED8	29	0.014062	Intersected
Manutencion	9	0.013133	Intersected
Tabaquismo	30	0.012155	Intersected
OA	15	0.010285	Intersected
DMII	14	0.009222	Intersected
Depresion	36	0.008693	Intersected
FuerzaPrension	45	0.008585	Intersected

Figure 5.6: Selected 32 important features.

a connection between the computational foundation of n-qubit data points and n-bit classical data points, like for the classical data (1001) the corresponding four qubit quantum state $|1001\rangle$ via following equation [290].

$$|D\rangle = \frac{1}{\sqrt{M}} \sum_{m=1}^{M} |x^m\rangle \tag{5.5}$$

Whereas, D is classical data, $D = \{x^1,\ x^2, \ldots..,\ x^M\}$ that creates binary vector, $x^m = \{b_1^m, b_2^m, \ldots b_N^m\}, b_i^m \in \{0,1\}, i \in \{1,2,\ldots.,M\}$ and M is the number of attributes.

5.3.3.2 Amplitude Encoding

The key concept of the amplitude encoding is to converts the classical data into quantum state amplitude. As seen below, classical data strings are converted to quantum amplitude variables using a normalized classical vector.

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ x_n^n \end{bmatrix} \tag{5.6}$$

The normalized $\|\mathbf{x}\|^2 = \sum_i |x_i|^2 = 1$, classical input x having length N can be converted into n-qubit quantum state amplitude, where $n = \lceil \log_2(N) \rceil$. Each input corresponds to

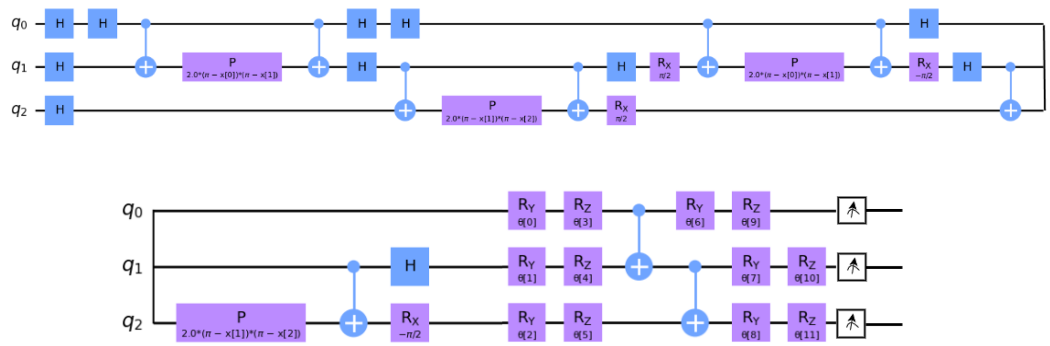

Figure 5.7: Circuital topology of basis encoding.

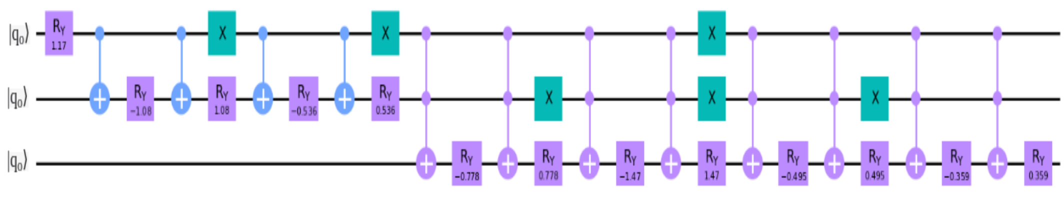

Figure 5.8: Circuital topology of amplitude encoding.

a quantum state whose computational basis's amplitudes are represented by the input vector, as shown below [282].

$$\psi_x = \sum_{i=1}^{N} x_i i \tag{5.7}$$

where $\psi \in$ Hilbert space ($\mathscr{H}$) and $\sum_i |x_i|^2 = 1$.

$$|\psi\rangle = R\left(v^i, \beta\right) |q_1 \dots q_{s-1}\rangle \, |q_s\rangle \tag{5.8}$$

The state $|\psi\rangle$ is considered a cascading circuit rotated in R_y gate rotation. The R_y gate is rotated for n number of features, where n indicates the binary power of the input feature vector. The circuital diagrams of basis encoding and amplitude encoding are indicated in figure.5.7 and figure 5.8. For simplicity, here I consider a 3 qubits circuit to distinguish the basic difference between the two schemes. The amplitude encoding has been considered for all three groups of qubits (n = 3, 4, and 5) that correspond to the three groups of feature vectors (v = 8, 16, and 32) as shown in figure 5.9. In order to construct the state $|\psi\rangle$, multiple types of circuit depth or gate rotation are utilized. In contrast to basis encoding and angle encoding, which required one qubit per dimension, and probably used a large number of qubits for higher dimension samples, this results in a loss of information or coherence in a NISQ device. Amplitude encoding, on the other hand, requires $\log(N)$ qubits with time complexity $O(N)$, where $N = 2^n$ and n is

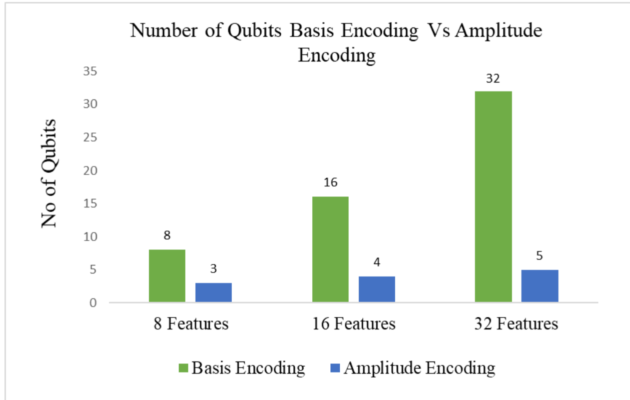

Figure 5.9: Number of qubits required for basis encoding and amplitude encoding.

the number of required qubits) for any n number of features, which assists in reducing the coherence issue in such devices.

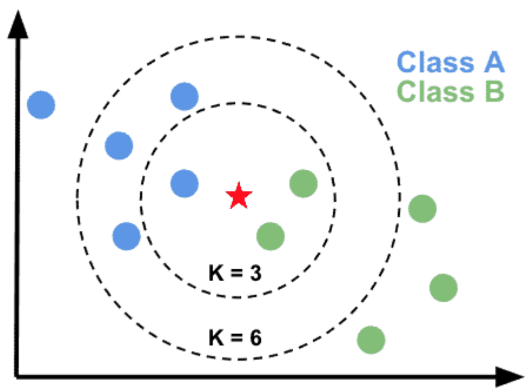

Figure 5.10: An overview of KNN algorithm.

5.3.4 Quantum K Nearest Neighbor Algorithm (QKNN)

K Nearest Neighbor (KNN) is a popular supervised machine learning technique. The concept behind its operating mechanism can be summed up as follows: Provide a testing sample, locate its closest k neighbors using some distance measure, and then classify the sample based on the information obtained from these neighbors. Generally, the majority voting approach is taken to complete the algorithm, and the test samples are given the K nearest neighbors leading category tag as a label. figure 5.10, shows how the KNN algorithm works, when K = 3, the testing sample (represented by a red star) is classified as belonging to the green dot (class B) category. However, when k = 6, the test sample is referred to as the blue dot (class A), because now that contains the majority of this category. In the KNN algorithm, the value of K is a key factor, as a change the value of K the classification outcome changes accordingly. However, this simplistic assumption contradicts actuality, particularly in the Big Data situation, because in large datasets, unexpected outliers, i.e., polluted data, may results in erroneous assessment.

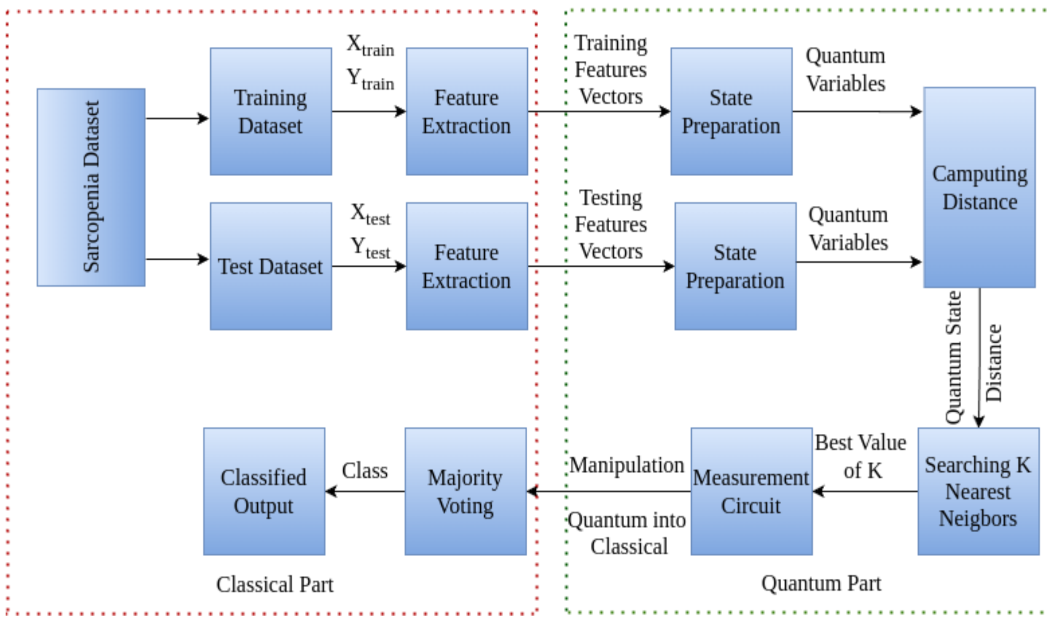

Figure 5.11: Block diagram of QKNN model.

5.3.4.1 Searching K Nearest Neighbors

One of a KNN's high time complexity features is searching K-nearest neighbors. For the unsorted search issue, a novel concept is presented with the introduction of the Grover algorithm. To determine the K-minimum values, Dürr presented a quantum algorithm [276] in 2004. In 2019, Miyamoto [277] used another concept to offer a quantum algorithm for finding K-minimum values. Finding K-minimum values in M data can be accomplished by both of their strategies with a time complexity of $O(\sqrt{kM})$. In the same way that classical KNN does, the QKNN uses the algorithm of quantum K minima finding to find the test states for K closest neighbors. First, the distance is computed in a quantum state, then a random K indices from set M are selected. A parameter t is considered which is used to find the K values which are supposed to be less than t. Utilize the binary search to identify the threshold index t minimum algorithm record that satisfies the requirement that the number less than t be near to K. To find out if the condition is satisfied, the quantum counting method is employed. Later utilizing the quantum Grover search algorithm [276], the indices are substituted in set M with new indices (that are not in set M) and continue to do so until the top k neighbors are identified as shown in figure 5.11. The most important step in this process is to set up an oracle that can perform the desired Grover operation in a quantum state. The oracle has to be able to perform calculations of the form $F_{y'} > F_y$ and verify

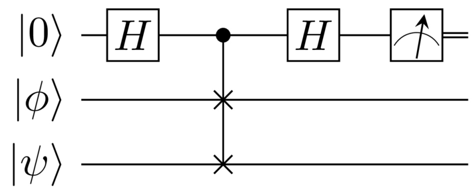

Figure 5.12: SWAP test circuit [280].

that $y' \notin A$. To create such an oracle, the fidelity is first extracted using the Swap test and encoded as the amplitudes of a quantum state. The fidelity is encoded as digital states after an analog to digital conversion of the amplitudes.

5.3.4.2 SWAP Test

The swap test is a type of quantum algorithm that can be implemented in order to produce a statistical assessment of the level of fidelity $\mathrm{F}(\psi,\phi) = |\langle\psi \mid \phi\rangle|^2$ among the two arbitrary n qubits. Figure 5.12 depicts a control swap (CSWAP), which is comprised of three registers, and its operation can be described as [279],

$$\begin{aligned} \mathrm{CSWAP}\,|0\rangle|\phi\rangle|\psi\rangle &= |0\rangle|\psi\rangle|\phi\rangle \\ \mathrm{CSWAP}\,|1\rangle|\phi\rangle|\psi\rangle &= |1\rangle|\phi\rangle|\psi\rangle \end{aligned} \tag{5.9}$$

In order to combine the three states $|0\rangle|\phi\rangle|\psi\rangle$ respectively, the implementation of the swap test between state $|\phi\rangle$ and state $|\psi\rangle$ requires three registers. Finally, at the circuit end, the measuring probability for the primary register is determined, where the difference between $\mathrm{Pr}(0) - \mathrm{Pr}(1)$ provides the desired level of fidelity [279].

$$\begin{aligned} \mathrm{Pr}(0) &= \frac{1}{2} + \frac{1}{2}|\langle\psi \mid \phi\rangle|^2 \\ \mathrm{Pr}(1) &= \frac{1}{2} - \frac{1}{2}|\langle\psi \mid \phi\rangle|^2 \end{aligned} \tag{5.10}$$

5.3.4.3 QKNN Classification

The proposed algorithm has the capability of classifying quantum states without their explicit classical description, where circuits are required to create these states. The model uses 2^n quantum states, where n is the number of qubits. The Swap test and expansions of quantum analog to digital conversion techniques were employed in the model to build an oracle, allowing us to limit the quantum KNN problem. When determining similarity, the fidelity and dot product are used, and fidelity can be applied to classify data. Let's say that $\mathscr{H}$ is the n qubit Hilbert space with a size of $N = 2^n$, and call the state represented by $|\psi\rangle \in \mathscr{H}$ the unknown test state whose label must be assigned as [280],

$$\{|\phi_j\rangle : j \in \{0, \ldots, M-1\}\} \subset \mathscr{H} \tag{5.11}$$

Where $\mathscr{H}$ is a set of M train states whose labels are known to us. For simplicity, let us suppose that $M = 2^m$ where m is considered a positive number. The goal is to identify the K nearest neighbors among the train states, and then label $|\psi\rangle$ based on the results of a majority vote. It is possible to define the fidelity as $F_j \equiv \mathrm{F}(\psi, \phi_j) = |\langle \psi \mid \phi_j \rangle|^2$ between the test state and the j_{th} train state [280].

$$F = [F_0, \ldots, F_{M-1}] \tag{5.12}$$

The subscript M is the length of the table that includes all of the fidelities together with the test state $|\psi\rangle$ and all of the train states $\{|\phi_i\rangle\}$. Additionally, the dot product $X(u,v) \equiv \langle u \mid v \rangle$ can be used to construct the QKNN algorithm in place of the fidelity method. The dot product is more helpful when dealing with real-valued vectors, which is typically the case in practical applications. In order to use fidelity with QKNN, a swap test must be performed between the test state and the train state while both are in superposition. This swap test must use analog encoded fidelity information F_i and the algorithm for a quantum analog to digital converter. The QkNN with is considering the dot products, and the swap test is replaced with the Hadamard test. By evaluating the overlap between the test state and a superposition of all training states, I perform the Hadamard test and analog-encode the dot product data X_i, and then utilize a variant similar to the real quantum analog to digital converter (QADC) technique to digitize the data.

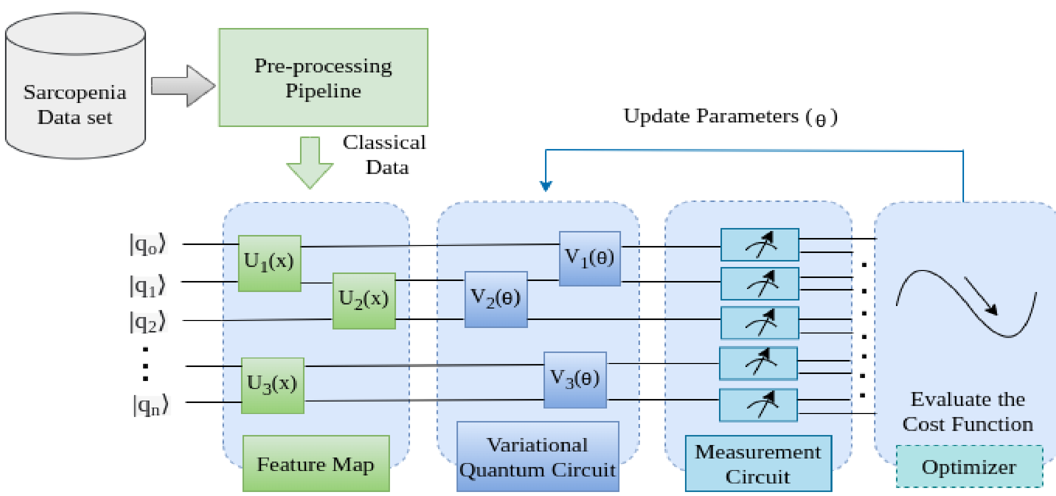

Figure 5.13: Block building of variational quantum classifier.

5.3.5 Amplitude Encoding Variational Quantum Classifier (AE-VQC)

The variational quantum classifier (VQC) is a supervised QML method that is often used to solve classification problems in the noisy intermediate scale quantum (NISQ) device. It is an important QML technique that is used to distinguish between interesting physics events and background events, allowing one to get the experimental results without using extra error-correction methods. The VQC model is based on the idea that classical information is encoded into a quantum states, processed using quantum techniques, and then a measurement circuit is used to get the classical predictions as shown in figure 5.13. The parameters are optimized and updated using a classical computer in this hybrid approach, which speeds up the optimization process without requiring longer coherence times. After the variational circuit, the qubits' state can be expressed as $|\psi(\theta)\rangle$, where θ represents the quantum gates' parameters, which are equivalent to weights in neural networks. The cost function, which is based for error mitigation by integrating noisy data in optimization computations, is computed using the device's repeated measurements. Based on quantum circuits that are challenging to replicate conventionally, the quantum approach maps classical input data into a large quantum feature space.

5.3.5.1 Quantum Feature Mapping

Quantum feature mappings' basic idea is derived from the traditional machine learning kernel technique, which involves non-linearly mapping a dataset into a higher-dimensional space to identify a hyperplane that may be used to classify non-linear data [238] [289].

$$U_{\Phi(x)} = \prod_{d} U_{\Phi(x)} H^{\otimes n} U_{\Phi(x)} H^{\otimes n} \tag{5.13}$$

A quantum feature map functions as a representation of the quantum kernel $\Phi : \mathscr{X} \rightarrow \mathscr{H}$, where $\mathscr{H}$ represents the quantum Hilbert space. I define a PauliFeatureMap for n-qubits generated by a unitary operation and considering the conventional Hadamard gates operation [281].

$$U_{\Phi(x)} = \exp\left(i \sum_{S \subseteq [n]} \phi_S(x) \prod_{k \in S} Z_i \right) \tag{5.14}$$

Where $|\phi(\vec{x})\rangle\langle\phi(\vec{x})|$, a Hilbert space vector, represents the quantum states and $\phi(\vec{x})$ represents the classical feature vector. The subscript H stands for the Hadamard gate, $U_{\phi(\mathbf{x})}$ is a diagonal gate (unitary gate) on the Paulifeature basis, and (Z_i) for the feature spacing. The requisite qubit count is directly correlated with the data dimension. By varying the angle to specific values, the data is encoded via the unitary gates $U_{\Phi(x)}$. Several feature maps, including FirstOrderExpansion $\phi S x \longmapsto x_i$, SecondOrderExpansion $\phi S : x \longmapsto (\pi - x_i)(\pi - x_i)$, and SecondOrderPauliExpanssion $\phi S : x \longmapsto \sin(\pi - x_i)\sin(\pi - x_i)$ [282] [283], are considered.

The feature map is identified by a data-encoding technique, which can be accomplished via a non-trivial or arbitrary state preparation quantum circuit and requires a time delay $\mathcal{O}(2^n)$. Before generating the state of amplitude encoding, the data is converted to corresponding angle representations using multiple controlled rotations. The input vector is used to construct the angle θ, v^i represents the i^{th} classical sample, and β and is the angle determined by the inverse sine of the classical input dimensions.

5.3.5.2 Variational Quantum Circuit

A hybrid quantum-classical method that combines the advantages of quantum and classical computing known as variational quantum circuit. A classical computer is used to optimizes that particular type of quantum circuit which consist of tunable parameters. The procedure involves setting up the states, doing the measurement, and parameterize the input x of the variational quantum circuit according to the number of parameters θ. The three fundamental elements that compose a typical quantum variational circuit are: an initial state, also referred to as the zero state or vacuum state; a quantum circuit parameterized by a set of free parameters θ; and an observable $\hat{B}$ at the output. Furthermore the observable consists of each wire local observable or set of wire in the circuit. The expectation values of one or more of these circuits, maybe with some traditional post-processing, for a particular task the scalar cost is defined as,

$$f(\theta) = \langle 0|U^{\dagger}(\theta)\hat{B}U(\theta)|0\rangle \tag{5.15}$$

Where the cost function is used to optimized the circuit free parameters $\theta = (\theta_1, \theta_2, \ldots)$. A classical optimization method probes the quantum device to train variational circuits. Typically, the optimization process is an iterative approach that looks for improved parameter θ choices at each stage. The variational form of a typical VQC model along with feature map and the encoding circuit with CNOT, RY and RZ gates are shown in figure 5.14 and figure 5.15 respectively.

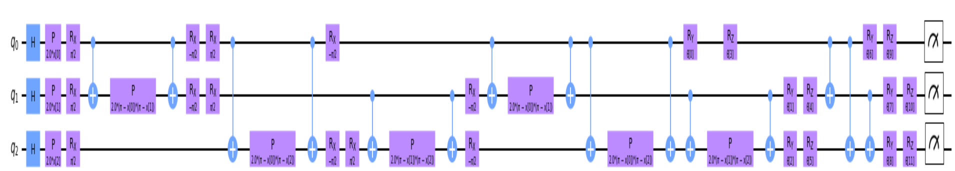

Figure 5.14: Circuital overview of typical VQC with feature map.

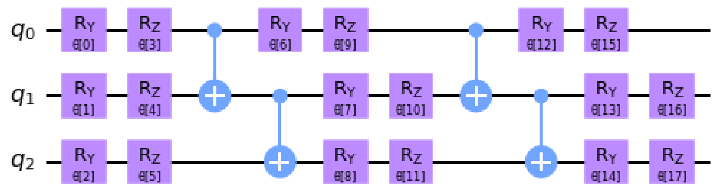

Figure 5.15: A variational circuit.

5.3.5.3 Measurement Circuit

A quantum measurement circuit is utilized to carry out a definitive measurement in order to evaluate the class possibilities. The measuring procedure is equivalent to selecting numerous samples from a distribution of possible computational base states and computing the average value. The final purpose circuit is then elaborated by using PauliFeatureMap and an EfficientSU2, where the circuit depth are consider 2. The purpose of training is to find out the parameter values that will optimize a given loss function. A quantum model can be optimized in a similar manner to that of a traditional neural network. In both cases, I calculate the gradient of the quantum circuit by running the model forward and figuring out the loss function. In order to find the difference between our predictions and the actual data, which is represented as a loss function value, the trainable parameters can be updated using gradient-based optimization techniques as a loss function during training.

5.3.5.4 Optimization

When the measurements are ready, an optimization procedure is used to update the quantum variational circuit's parameters. The cost function's value is reduced by training the circuit parameters using the classical loop. The Constrained Optimization by Linear Approximations (COBYLA) optimizer has been considered [284], which uses a n + 1 fundamental (n being the number of features) to produce successive linear assumptions of the cost function, improving these assumptions in a trusted region at each stage. COBYLA is a numerical optimization technique which is used for constrained challenges in which the objective function's derivative is unknown.

5.4 Experimental Results

In this study, the experiments are carried out on a classical device by using Python 3.9.11 with the Qiskit 0.37.0 version. StandardScaler and MinMaxScaler were imported from Sklearn 1.0.2 to standardize the input shape for the model to the range (-1, 1). For the QKNN model, the QKNeighborsClassifier is imported from qlearnkit 0.2.0 with qiskit-machine-learning library 0.3.1, qiskit-terra 0.23.0, and qiskit-aer 0.10.3.

Table 5.5: Comparison matrix of proposed models QKNN and AE-VQC with classical KNN and classical NN for all datasets.

Dataset	Classifiers	Confusion Matrix: Prediction Class → / Actual class ↓	0	1	Performance Evaluation: Precision	Recall	F1-score	Test Size	Average Values: Accuracy	F1-score
Dataset 1 (8 Features)	Classical KNN	0	11	5	65%	69%	67%	16	67%	66.9%
		1	6	11	69%	65%	67%	17		
	QKNN	0	10	7	77%	59%	67%	17	**70%**	**72.1%**
		1	3	13	65%	81%	72%	16		
	Classical NN	0	12	4	80%	75%	77%	16	79%	79.95%
		1	3	14	78%	82%	80%	17		
	AE-VQC	0	15	2	75%	88%	81%	17	**79%**	**80.9%**
		1	5	11	85%	69%	76%	16		
Dataset 2 (16 Features)	Classical KNN	0	14	3	67%	82%	74%	17	70%	73.7%
		1	7	9	75%	56%	46%	16		
	QKNN	0	9	8	90%	53%	67%	17	**73%**	**76.8%**
		1	1	15	65%	94%	77%	16		
	Classical NN	0	13	3	72%	81%	76%	16	76%	76.2%
		1	5	12	80%	71%	75%	17		
	AE-VQC	0	9	8	75%	53%	62%	17	**67%**	**70.2%**
		1	3	13	62%	81%	70%	16		
Dataset 3 (32 Features)	Classical KNN	0	12	5	80%	71%	75%	17	76%	75.2%
		1	3	13	72%	81%	76%	16		
	QKNN	0	13	4	76%	76%	76%	17	**76%**	**75.5%**
		1	4	12	75%	75%	75%	16		
	Classical NN	0	11	5	73%	69%	71%	16	72.7%	73.9%
		1	4	13	72%	76%	74%	17		
	AE-VQC	0	15	2	68%	88%	77%	17	**73%**	**76.7%**
		1	7	9	82%	56%	67%	16		

The hyper-parameters have been tuned by using different values of K (k=3, 5, 7) with different shots of quantum instances (1024, 2000, and 5000). The AE-VQC model consists of 5 layers, using COBYLA optimizer and considering a various sets of learning rates (lr = 0.001, 0.0001, and 0.0001) with different batch sizes (bs = 8, 16, and 32). Due to the lack of a publicly accessible quantum computer, the IBM quantum simulator, including the state vector and QASM simulators are used. The models have been evaluated using 80% training and 20% testing dataset. Three different sets of features (8, 16, and 32) have been considered for all the models, and then compare the competency in terms of accuracy, recall, F1-score, and precision.

5.4.1 Performance Evaluation

The performance of the proposed models has been obtained by using the mean, variance, and gradient of the classical evaluation metrics and comparing the outcome on the grounds of accuracy, precision, F1 score, recall and loss function is computed as follows,

$$loss = \frac{1}{n}\sum_{i=1}^{n}(y_a - \hat{y_a})^2 \tag{5.16}$$

In the above, the subscript y_a indicates the observed value, $\hat{y_a}$ is the predicted value, and n is the number of observations respectively. The accuracy of the model in terms of confusion metrics is obtained as follows:

$$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{5.17}$$

The subscripts T_P, T_N, F_N, F_P indicates, the values of true positive, true negative, false negative, and false positive, respectively. In order to obtain a better accuracy, the ratio of true positives and true negatives must be close to 1. Among the retrieved events, precision is a component of major cases. In the case of a good classifier, the value of precision should always be 1, when F_P is zero and $T_P = T_P + F_P$. Precision refers to a classification model's capacity to correctly recognize and isolate pertinent data points. In mathematical terms, precision is determined by dividing the number of true positives by the sum of true positives and false positives as describe bellow,

$$Precision = \frac{T_P}{T_P + F_P} \tag{5.18}$$

Where F_P indicates the values of false positive, and T_P for true positive respectively. Furthermore, another important metric is recall, which is an essential ML statistic. Recall is also known as sensitivity or true positive. A good classifier always has the recall value 1, because the equivalent values of the numerator and denominator are $T_P = T_P + F_N$ which mean the desire observation is positive and also predicted to be positive. The recall metric can be defined as,

$$Recall = \frac{T_P}{T_P + F_N} \tag{5.19}$$

A statistic called the F1 score is used to assess how well a classification model performs, especially when working with unbalanced datasets. It is the harmonic mean of recall and precision, bringing both metrics together into one number that strikes a balance between both.

$$F1score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{5.20}$$

5.4.2 Results With 8 Features

The experiment to classify the sarcopenia disease by considering 8 features using classical KNN, quantum KNN, a classical neural network, and AE-VQC model. The dataset was randomly split into 80% training and 20% test sets, and the same hyperparameters are considered to evaluate both models. Following cleaning and balancing

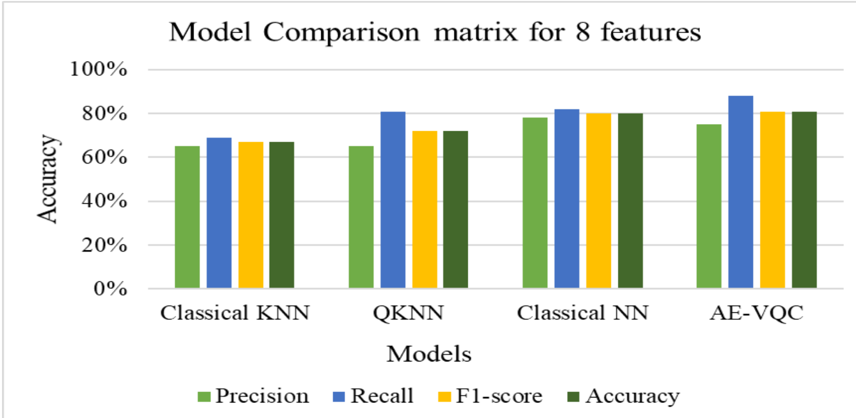

Figure 5.16: Models comparison matrix for 8 features.

the dataset, the intersection of the RF-RFE and LR-RFE was employed to select the 8 most significant features. Initially, both the classical and quantum KNN are tested for a different number of neighbors $K = 3,5,7$ and the results are recorded. However, both models perform best when $K = 5$, while $K = 3$ overfits and $K = 7$ produces the wrong prediction. Table 5.5 indicates the comparison matrix of all the models in terms of precision, recall, F1-score, and accuracy. The classical KNN model has the highest accuracy of 67% with an average F1-score of 66.9%, while the QKNN model has the highest accuracy of 70% with an F1-score of 72.1%. In the case of the QKNN model, I have used 1024 shots to get the resultant probability distribution. The amplitude encoding method used by the QKNN model reduces the dimensionality of the input data and enables the algorithm to make better predictions. Similarly, I also compared our second model, AE-VQC, to the classical NN model, and found that it performed slightly better than the traditional model, as indicated in table 5.5. Both models behave the same way since they both contain three layers, a learning rate of 0.0001 with a batch size of 16, and a loss function of binary-cross entropy. The only distinction is the input shape, because the AE-VQC model also performs the amplitude encoding method to decrease the input shape's size. The quantum models only need three qubits for the desired experiment with 8 features, which further simplifies processes. The comparison matrix of all models with 8 features is displayed in figure 5.16, which demonstrates that given the desired dataset with the intended number of features, quantum models are better performed than classical models.

5.4.3 Results With 16 Features

In this second set of experiments, I am interested in demonstrating how quantum models with amplitude encoding can assist in more accurate prediction. In particular, it could suggest that 16 attributes have been utilized to evaluate the compatibility of both the classical and quantum models. Following this, a train and test each model's

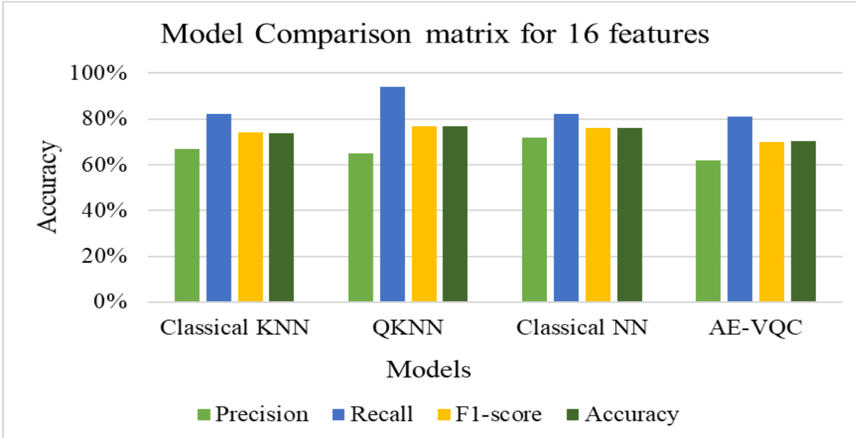

Figure 5.17: Models comparison matrix for 16 features.

performance for 16 characteristics, and subsequently evaluate the outcomes by using the testing dataset. The performance of each model for the 16-feature dataset is also displayed in the middle section of table 5.5, and it appears that the classical KNN has an average accuracy of 70% with a recall score of 73.7%, followed by the QKNN's has 73% accuracy with a recall score 76.8%, the classical NN with 76% with recall score 75.5%, and the AE-VQC model's has 67% accuracy with a recall score 70.2%. The classical KNN and QKNN models both consider the same hyper-parameters, however, in this experiment, the QKNN uses amplitude encoding with 4 qubits rather than 3 qubits. The dataset has now been encoded by the amplitude encoding technique from 16 features to 4 input vectors, which means that the data vector is converted into the quantum state amplitude. Because of the variable entanglement and data spinning, the QKNN model performs marginally better than the classical KNN in such a comparison, which allows it to get more data insight. Moreover, the input shape dimension is reduced by amplitude encoding, which reduces model complexity and improves prediction. In the comparison of the classical NN and AE-VQC models, the classical model retains better prediction, with an average accuracy of 76% and a recall score of 76.2%, while AE-VQC has 67% accuracy and a 70.2% recall score. For such an experiment, these two models yet again share the same hyper-parameters to prevent bias. The confusion matrix for each of the models with 16 features of input data is shown in figure 5.17, where the QKNN model has the highest recall score, which demonstrates how well the model detects true positives or how many of the diseased patients were identified out of all the patients.

5.4.4 Results With 32 Features

The third experiment results are shown in the final section of table 5.5, where the performance metrics comparing our proposed quantum models with classical models for dataset 3 having 32 feature dimensions are provided. The classical KNN model achieved the highest average accuracy 76% with a 75.2% recall score, and the QKNN model also retained the same accuracy of 76% with a recall score of 75.5%. In this scenario, the quantum model achieves the classical model's accuracy limit, which indicates that our proposed model is competitive with the classical model. The F1-score of the QKNN model is 0.3% higher than the classical model, which demonstrates the effectiveness of the class-wise performance over the classical model. The complexity of the models has been reduced by using an amplitude encoding scheme, which converts 32 qubits for 32 features into a 5-qubit system. The 32 important features have been selected by intersecting the output of RF-RFE and LR-RFE after cleaning the dataset. The input dataset is first normalized using the Minmax and Standard scalars, where the random state is taken to be 42 for both the classical model and the quantum model. After normalization, the classical data is transformed into a verifiable quantum state for the quantum model using the Paulifeature map, and the number of repetitions is determined to be 5. Furthermore, the suggested AE-VQC for 32 features is compared to the 32-feature traditional NN model by considering the same hyper. The average accuracy of the classical and quantum models for the intended set of features is 72.7% with an f1-score of 73.9% and 73% with an f1-score of 76.7% respectively, which seems to indicate that the quantum model performs better than the classical. The comparison results for 32 features for all the models are also shown in figure 5.18, which indicates that for the 32 features, both the quantum models outperform the classical models by predicting the sarcopenia disease more accurately.

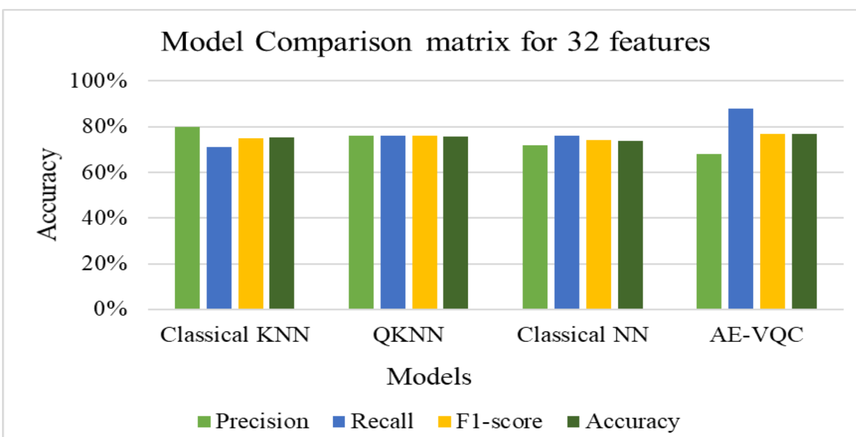

Figure 5.18: Models comparison matrix for 32 features.

5.5 Discussion

The competency of of our proposed quantum models has been obtained by comparing it to those models that has already been published. The proposed work has been evaluated to those published articles that have used the same model with different datasets

Table 5.6: Comparisons of proposed QKNN model with previous published models.

Author Name	Used Model	Dataset	Accuracy
D.Kok *et.al* [285]	Amplitude encoding	Iris	67%
Y.Dang *et.al* [286]	QKNN, Features extraction	Graz-01, Caltech-101	83.1%, 78%
N.Wiebe *et.al* [287]	Amplitude estimation	PIMA Diabetes	73%
Proposed Model	Amplitude encoding, Feature selection	Sarcopenia	70%, 73%, 76%

and methodologies. In such a scenario, [285] presents a quantum version of KNN by using Qiskit, where the distance among the two quantum states (ψ,ϕ) is measured by using fidelity $F = \|\langle\psi \mid \phi\rangle\|^2$. The SWAP test is used to evaluate the fidelity, and rather than one victor at a time, the whole data set fidelity is measured by using an Oracle $\mathscr{W}$. Before loading the data onto the quantum circuit, an analog (amplitude) encoding technique has been considered to encode the data. The model has been evaluated by using three different datasets, i.e., German-Credit, Hep, and Iris, and gets the highest accuracy of 67% for the Germen-Credit dataset. Similarly, Dang Y *et.al* in [286] improves the efficiency of classification by using the robust parallel processing capabilities of quantum computers. In order to, perform the parallel computing, the quantum version of KNN has been employed, where feature vectors from images are first extracted and then inserted into a quantum superposition state. A measurement circuit is used to obtain the classified output and get the model complexity $O(\sqrt{kM})$, which is far superior to the classical model. The two image datasets Graz-01 and Caltech-101 are analyzed and obtained with maximum accuracy of 83.1%, and 78% respectively. Furthermore, in [287] authors present a quantum K-nearest learning K means clustering models by determining the Euclidean distance through the inner product and directly. Each algorithm uses quantum superposition sequentially to simplify the structure of the cluster formation phase to a uniform cost. Distances between records and centroids can be computed simultaneously, saving time, particularly in the case of large datasets. Using quantum amplitude estimation and amplitude amplification techniques, the quantum processes speed up the distance computation. In this scenario, quantum techniques result in quadratic reductions in query complexity when compared to Monte Carlo approaches. The models have been evaluated using Heart Disease, Thyroid Disease, and Diabetes (PIMA) datasets and have obtained accuracy rates of 80%, 55%, and 73% respectively. In short, when compared to the recently published state of the art, our proposed QKNN model outperforms it in terms of accuracy for all feature sets of the sarcopenia dataset, as shown in table 5.6.

I also compared our AE-VQC model, to those models that had been published and used the same parameters for different health care applications. The variational classifier, which is being developed, is one of the methods based on a hybrid approach

Table 5.7: Comparisons of our proposed AE-VQC model with recently published models.

Author Name	Used Model	Dataset	Accuracy
D.Sierra-Sosa *et.al* [288]	VQC, state preparation	Diabetes	70%, 72%
H.Gupta *et.al* [289]	VQC, DL	PIMA Diabetes	74%
D.Maheshwari *et.al* [290]	QSVM, VQC, VQC	Diabetes	74.1%, 68.7%, 67%
Proposed Model	VQC, Feature selection	Sarcopenia	9%, 67%, 73%

integrating quantum and classical devices presented in [288]. The work improves the viability of addressing classification problems with NISQ devices by increasing the prediction rates while utilizing amplitude encoding techniques. Following the state preparation, the experiment results with two and three qubits are acquired using the amplitude encoding technique. The results of the desired scheme indicates that it is possible to classify the prevalence of acute illnesses in diabetic patients using an amplitude-encoded VQC algorithm with 70% and 72% accuracy, respectively. To improve performance when tackling classification issues, quantum computing techniques and classical neural network approaches are combined in [289] by Gupta H *et.al*. The study utilizes QML and DL frameworks to implement the diabetes prediction model. Furthermore, the importance of exploratory data analysis has been investigated and found to be essential for accurate and reliable prediction. The VQC model has employed the PauliZ operators, and the average evaluated value is used to increase statistical correctness. The model has been tested using the PIMA Indian diabetes dataset and has achieved the highest accuracy 74%. Since QML has recently advanced, many researchers are focusing on QML by applying various quantum algorithms for a variety of applications. In [290], three QML classifiers QSVM, VQC, and AE-VQC were investigated while considering three different datasets. The importance of feature extraction and pre-processing techniques, which influence model prediction, is also addressed. The total number of features in the diabetes dataset is considered to be 8, where the QSVM, basis encodes VQC using an 8-qubit quantum system while amplitude encodes VQC using a 3-qubit quantum system. By considering the diabetes dataset, the models perform well and achieve the highest accuracy of 74.1%, 68.7%, and 67% for QSVM, VQC, and amplitude-encoded VQC, respectively. Similarly, as shown in table 5.7, when comparing our second proposed AE-VQC model to those of published models, it also comparatively performs well and has the highest accuracy of 79%, 67%, and 73% for 8 features, 16 features, and 32 features, respectively.

5.6 Conclusions and Future Directions

This work involves applying and assessing QML algorithms such as QKNN and AE-VQC. These models employ an amplitude encoding method that simplifies the model by transforming the input vector into the quantum state amplitude. The first step in preparing the dataset for evaluation of the two quantum classifiers involved removing any null values, addressing missing values, and converting the data type as needed. In selecting the most effective features, the evaluation utilized a feature selection procedure that incorporated both RF with RFE and LR with RFE, and then intersected the classifier results to determine the optimal features. It should be noted that the QKNN model's use of amplitude encoding techniques was dependent on the number of features and required only a small number of qubits. It is important to understand that state preparation is merely one of among several techniques for QML algorithm enhancement. QKNN is a supervised QML algorithm that employs the quantum K maximum-finding algorithm to identify the K nearest neighbors of a test state. In this algorithm, the distance between two quantum states (ψ and ϕ) is measured using the fidelity metric, which is defined as $F(\psi,\phi) = |\langle\psi \mid \phi\rangle|^2$. To assess the fidelity between two arbitrary qubits, a swap test is utilized. The evaluation of the QKNN model is conducted using datasets with 8, 16, and 32 features, and it produces favorable results when compared to classical models. The second model, AE-VQC, consists of four components: feature mapping, a variational quantum circuit, a measurement circuit, and an optimizer that updates the θ values. The model comprises three layers and employs a system of 3, 4, and 5 qubits with amplitude encoding for three different feature sets, respectively. At the conclusion of the model, a measurement circuit is employed to convert the quantum variable back into a classical binary output. The performance of both the proposed QKNN and AE-VQC models is compared to classical KNN and NN models, with QKNN achieving the highest accuracy of 70%, 73%, and 76%, while AE-VQC attains the highest accuracy of 79%, 67%, and 73%, respectively.

QML still has several limitations, it has limited hardware, and there are only a few quantum computers available. Quantum computers are susceptible to errors due to noise in the system, which can affect the accuracy of QML algorithms. It is not yet clear if quantum computers will be able to achieve "quantum supremacy" over classical computers for all types of machine learning problems. However, besides the aforementioned challenges, QML still shows great potential.

The future of QML is quite promising. Quantum computing has the potential to revolutionize machine learning by allowing us to process and analyze data at an unprecedented scale and speed. QML algorithms could enable us to solve problems that are currently intractable with classical computing, such as simulating large molecules, optimizing complex systems, and analyzing large datasets. In short QML will play an increasingly important role in solving complex problems in the years to come.

Chapter 6

Conclusions

In this chapter, the appropriate findings obtained through the development of the thesis are presented. Following the achievement of the various objectives outlined in the introductory chapter, these results provide a comprehensive summary of the research outcomes. The conclusions highlight the significant insights and contributions derived from the study, offering a clear synthesis of its overall impact and implications. This synthesis not only emphasizes the advancements made but also suggests potential directions for future research and applications.

This thesis explored various methodologies of QML models to address critical healthcare challenges. The first case study focused on the severity classification of Covid-19 patient data using QML models, specifically the enhanced QSVM and QRF models. Both models were tested using real Covid-19 data obtained from a private hospital. The performance of these quantum models was compared with their classical counterparts, demonstrating significant improvements. These results fulfill the objective of leveraging quantum computing technology to advance healthcare solutions. The prevalence of heart disease is rapidly increasing worldwide, impacting both the global economy and public health. The second case study presents a benchmark of the FCQ-CNN for handling classification tasks on classical data. The FCQ-CNN algorithm can be customized using factors like parameterized quantum circuits, quantum filters, pooling operators, classical data preprocessing, quantum feature mapping, optimizers, and cost functions. Data from public hospitals containing patient records of heart disease was used, and the results showed better performance compared to classical models on the same data samples. The third case study significantly contributes to the primary objective of this thesis by implementing QML models for the classification of sarcopenia disease. These models utilize an amplitude encoding method, which simplifies the model by transforming the input vector into quantum state amplitudes. The evaluation of the models was conducted using datasets comprising three different subsets of features. The comparative analysis revealed that the proposed models produced favorable results, demonstrating superior performance when compared to their classical counterparts. This highlights the potential of quantum computing techniques in enhancing the accuracy and efficiency of disease classification.

The results presented in the subsequent chapters suggest that QML has the potential to improve medical decision-making as it enhances diagnostic accuracy for several kinds of health diseases. By performing accurate data analysis, QML models can deliver highly trustworthy insights, allowing medical professionals to make more informed and precise judgments. This technology development promises to improve patient care by providing reliable diagnostic results generated from complete patient data analysis.

6.1 Achievements

This dissertation has made significant contributions to medical applications by developing end-to-end frameworks encompassing data pre-processing, the creation of tailored quantum machine learning architectures and generating reliable results using standard validation metrics for each application. Achieving the objectives outlined in Section 1.2 facilitated the completion of the various stages and contributions of this thesis. All seven major objectives of this thesis were successfully met during the research process. The following sections detail the specific aims and research topics addressed in this work.

> Quantum machine learning is able to outperform classical machine learning in disease early prediction, including Covid-19, cardiovascular disease, and sarcopenia, by utilizing quantum properties, a new learning paradigms to enhance accuracy, and leading to significant advancements in personalized healthcare and treatment optimization.

Guided by the proposed hypotheses, this dissertation endeavors to unveil the potential of QML in tackling pivotal challenges within the healthcare domain. Attaining this primary objective requires meeting the following specific aims.

> **SO 1:** To establish the current state-of-the-art in quantum computing and quantum machine learning, explore their healthcare applications, and examine the diverse types of medical datasets utilized.

To achieve this objective, a comprehensive summary of the most recent advancements in the healthcare domain was compiled based on articles published recently. The methodology employed for this summary followed stringent guidelines for conducting a systematic literature review, which included formulating precise research questions and assessing the quality of the articles using established metrics. Initially, a comprehensive search across multiple databases yielded 2,038 records. To ensure relevance and quality, 468 duplicate records were identified and removed. Following this, 1,053 records that did not pertain directly to healthcare were excluded from further consideration. The remaining records were then subjected to a more detailed screening process. Based on the title evaluation, 258 records were eliminated. This was followed by an abstract review, which resulted in the exclusion of an additional 68 records. Finally, a thorough full-text analysis led to the removal of 39 more records. The rigorous screening process culminated in the selection of 49 articles, which were then thoroughly evaluated. These selected articles provided a succinct overview of the most recent literature, contributing significantly to the existing body of knowledge and understanding of QML algorithms and their applications in the healthcare sector, as referred to in chapter 2.

SO 2: To determine the pre-processing and data analysis techniques applied to the Covid-19 dataset.

SO 3: To check the performance of QML models with it's counterpart classical models for Covid-19 cases classification.

For the desired objective, the framework of the entire study has been established, by first examining the Covid-19 pandemic, which has profoundly impacted global health and socio-economic structures. To address this objective, to tackle clinical applications the QML algorithms have been employed on Covid-19 dataset, specifically focusing on electronic healthcare data and clinical features. Then, the E-QSVM and QRF models are applied to datasets related to Covid-19 and influenza, which were collected from various private hospitals. The efficacy of these models was validated by comparing their performance against classical models and other recently published quantum models. As referred to chapter 3, the outcomes indicate that the proposed QML models not only outperform their classical counterparts but also exhibit enhanced accuracy and robustness, thereby showcasing their potential to significantly advance clinical applications in the context of the Covid-19 pandemic and beyond.

SO 4: To analyze and comprehend cardiovascular disease, data types, and the significance of feature dimensionality reduction.

SO 5: To design and validate a QCNN model that outperforms its counterpart CNN in the classification of cardiovascular diseases.

This specific research question focusing on the classification problem for cardiovascular disease (CVD), a leading cause of morbidity and mortality worldwide. The study parameters include clinical and demographic factors such as age, gender, blood pressure, cholesterol levels, smoking status, and comorbidities like diabetes and obesity. For this purpose, a FCQ-CNN for IHD classification inspired by traditional CNNs has been developed and validated. The IHD dataset was processed using the MRMR filter for data cleaning and SVM-RFE for dimension reduction. The model's performance was compared with classical Optimized-CNN and FCNN models using identical optimal parameters. This quantum approach aims to enhance classification accuracy, demonstrating significant improvements over classical models as briefly described in chapter 4.

SO 6: To analyze the data pre-processing, and state preparation for QML models.

SO 7: To develop and validate amplitude-encoded QKNN and VQC models that outperform their classical NN and KNN counterparts for the prediction of sarcopenia.

In the present hypothesis, the third contribution focuses on the classification of sarcopenia using QML models. Initially, it is essential to define sarcopenia and discuss its clinical, economic, and societal impacts. The etiology of sarcopenia in the elderly population is elucidated, along with a brief history of the term. Statistical data on the prevalence of sarcopenia are presented to highlight its significant clinical implications. Data for this study were obtained from hospital electronic health records. The required sample size was calculated based on the estimated prevalence of sarcopenia, utilizing bioimpedance analysis in older adults. The VQC and QKNN models are considered, which were tested with three different set of features. These important set of features are obtained by using the output intersection of RF with RFE and logistic regression with RFE commonly using for feature selection. The models incorporated a feature map, a variational quantum circuit, a measurement circuit, and a COBYLA optimizer, as briefly explained in chapter 5.

6.2 Scientific Contributions

This dissertation has been presented as a PhD by publication, consisting of two articles published in international Q1 journals with impact factors. A third journal paper is currently in the final round of review in the Quantum Machine Intelligence journal by Springer Nature. Additionally, other significant contributions to the scientific community include a collaborative journal paper with a PhD student from the University of Deusto, published in a Q1 journal, and the presentation of a conference paper at the International Conference on Electrical, Computer, Communications, and Mechatronics Engineering (ICECCME) 2022. A summary of these scientific contributions have been shown in table 6.1.

S.No	Title	Number	Description
1	PhD Grant	1	Co-fund Marie Skłodowska-Curie
2	Journal articles	4	IF 3.4, 3.4 and 3.9, 4.1 and all journals are Q1.
3	Conference articles	1	ICECCME Maldives
4	Pipeline research	1	Currently working on a project to design and validate QML model for lupus disease prediction
5	Seminars presented	3	Related to each year PhD progress
6	Seminars Attend	9	IEEE Quantum Week 2022 (QCE22), 2024 and 2025 and 6 seminar related to PhD study
7	Study Courses	7	PhD Courses
8	Workshops	3	Workshop on Quantum Machine Learning 2022, Quantum computing programming 2023, QSciTech-Quantum BC-CMC Virtual Workshop on Quantum Machine Learning
9	Summer Schools	3	Qiskit global summer school 2022, 2023, Quantum machine learning from fundamentals to applications
10	Certificates	2	Conference participating certificate, Webinar attending certificate
11	IBM challenges	2	IBM Quantum Challenge 2022, IBM Quantum Challenge 2024
12	International Mobility	1	QCE group, TU Delft Netherlands

Table 6.1: Scientific contribution summary.

6.2.1 Published articles in international journals

The first published journal paper focuses on the implementation of Quantum Convolutional Neural Networks for classifying cardiovascular diseases, as illustrated in the accompanying table 6.2.

The second journal paper, as presented in table 6.3, concentrated on a detailed literature review of quantum computing, QML, and their applications in the healthcare domain. This study analyzed 49 relevant records exclusively based on the use of QML in healthcare.

The third journal paper, as depicted in table 6.4, is presently in the final round of review, having already undergone two rounds of revisions. This paper explores the implementation of QML models, specifically the QKNN and VQC models, for the classification of sarcopenia. The study included six experiments on three subsets for each model.

The paper shown in table 6.5, not included as a thesis chapter, contains a significant contribution through collaboration with Danyal Maheshwari, a PhD student at

Title	A Fully Connected Quantum Convolutional Neural Network for Classifying Ischemic Cardiopathy
Authors	Ubaid Ullah, Alain García Olea Jurado, Ignacio Diez Gonzalez, Begonya Garcia-Zapirain
Journal	IEEE Access
Date	26 December 2022
Impact Factor	3.4 (2023) Q1
DOI	10.1109/ACCESS.2022.3232307

Table 6.2: First publication detail [Ullah 2022].

Title	Quantum Machine Learning Revolution in Healthcare: A Systematic Review of Emerging Perspectives and Applications
Authors	Ubaid Ullah, and Begonya Garcia-Zapirain
Journal	IEEE Access
Date	12 January 2024
Impact Factor	3.4 (2023) Q1
DOI	10.1109/ACCESS.2024.3353461

Table 6.3: Second publication detail [Ullah 2024].

the University of Deusto. The research focuses on the application of quantum machine learning to electronic healthcare records for the classification of ischemic heart disease. The work introduces a novel QML model featuring an innovative pre-processing technique, which establishes an efficient methodology for ischemic heart disease classification. Remarkably, this approach achieves competitive results with numerous state-of-the-art studies that utilize considerably larger datasets.

6.2.2 International Conference

Table 6.6 provides a comprehensive overview of the contributions made through publications presented at international conferences. This body of work emphasizes the development, implementation, and meticulous fine-tuning of QML models. Specifically, it explores the QRF and QSVM models. These advanced machine-learning techniques are applied to a dataset comprised of real-world Covid-19 data. The primary objective is to enhance the performance and accuracy of these models in predicting and analyzing trends and patterns within the Covid-19 dataset, thereby demonstrating the practical applicability and potential benefits of QML in handling complex, real-time public health data.

Title	Sarcopenia Risk Prediction and Feature Selection by Using Quantum Machine Learning Algorithms
Authors	Ubaid Ullah, Danyal Maheshwari, Cristian Castillo Olea, and Begonya Garcia-Zapirain
Journal	Quantum Machine Intelligence
Date	30 October 2024
Impact Factor	4.1 (2024) Q1
DOI	Accepted

Table 6.4: Third publication detail [Ullah 2024].

Title	Quantum Machine Learning Applied to Electronic Healthcare Records for Ischemic Heart Disease Classification
Authors	Danyal Maheshwari, Ubaid Ullah, Pablo A. Osorio Marulanda, Alain García-Olea Jurado, Ignacio Diez Gonzalez, Jose M. Ormaetxe Merodio and Begonya Garcia-Zapirain
Journal	Human-centric Computing and Information Sciences (HCIS)
Date	15 February 2023
Impact Factor	3.9 (2023) Q1
DOI	10.22967/HCIS.2023.13.006

Table 6.5: Fourth publication detail [Maheshwari 2023].

Figure 6.1: Presentation certificate of conference ICECCME 2022.

Title	Severity Classification of Covid-19 Patients Data using Quantum Machine Learning Approaches
Authors	Ubaid Ullah, Danyal Maheshwari, Hanna Helene Gloyna and Begonya Garcia-Zapirain
Conference	ICECCME, IEEE 2022
Date	16, 17, 18 November 2022
Location	Maldives
DOI	10.1109/ICECCME55909.2022.9987991

Table 6.6: Fifth publication detail [Ullah 2022].

6.2.3 Quantum Computing, QML, Workshops and Summer Schools

During my PhD study, I also participated in various quantum computing and QML workshops and summer schools remotely and physically. For example Qiskit Global Summer School was an annual educational event organized by IBM, designed to introduce participants to quantum computing and quantum programming using the Qiskit framework. This summer school provides a comprehensive curriculum that includes theoretical lectures and hands-on lab sessions, making it accessible to learners at various levels. The program typically covers fundamental concepts of quantum mechanics, quantum algorithms, and their practical implementation using Qiskit. The participants are encouraged to engage with quantum computing through interactive exercises, gaining practical experience in developing quantum applications and contributing to the growing quantum computing community. I also participated in a quantum entanglement workshop for quantum computing and programming organized by Nigeria focuses on introducing participants to the fundamentals of quantum computing and programming. This workshop is part of the broader efforts to promote quantum education in Nigeria and beyond, providing a platform for learners to engage with cutting-edge quantum technologies. Moreover, the CISM Summer School on Quantum Machine Learning: From Fundamentals to Applications is an intensive educational program designed to bridge the gap between quantum computing and machine learning hosted by the International Center for Mechanical Sciences (CISM) and the University of Udine, Italy. This summer school covers foundational aspects of both quantum mechanics and machine learning before moving toward advanced topics such as quantum algorithms, quantum neural networks, and hybrid quantum-classical models. Participants engage in theoretical lectures, hands-on workshops, and discussions on the latest research and technological advancements in quantum machine learning. Figure 6.2, 6.3 and 6.5 show the certificates for the desired summer schools and workshops, respectively.

Centre International des Sciences Mécaniques
International Centre for Mechanical Sciences

Palazzo del Torso | Piazza Garibaldi, 18 | 33100 Udine Italy
cism@cism.it | www.cism.it
Phone +39 0432 248511

To whom it may concern

This is to certify that

Ubaid Ullah

has participated in the CISM-UniUD Joint Advanced School on

Quantum Machine Learning: from Fundamentals to Applications

held in Udine, Italy from September 12, 2022 to September 16, 2022 and coordinated by Professor Francesco Petruccione (University of KwaZulu-Natal, Republic of South Africa), Parofessor Carla Piazza (University of Udine, Italy) and Professor Giuseppe Serra (University of Udine, Italy).

Udine, Italy, September 16, 2022

The Secretary General of CISM
Professor Antonio De Simone

Figure 6.2: Certificate of quantum machine learning summer school 2022.

6.2.4 International Mobility

During my PhD, I had the opportunity to undertake international mobility as part of my academic development, which required a research stay of at least three months. I chose to collaborate with the Quantum and Computer Engineering Group at TU Delft Netherlands, one of the leading institutions in quantum technology research. During my time there, I engaged with a team of experts specializing in QML, focusing on investigating critical challenges and exploring advanced models in this emerging field.

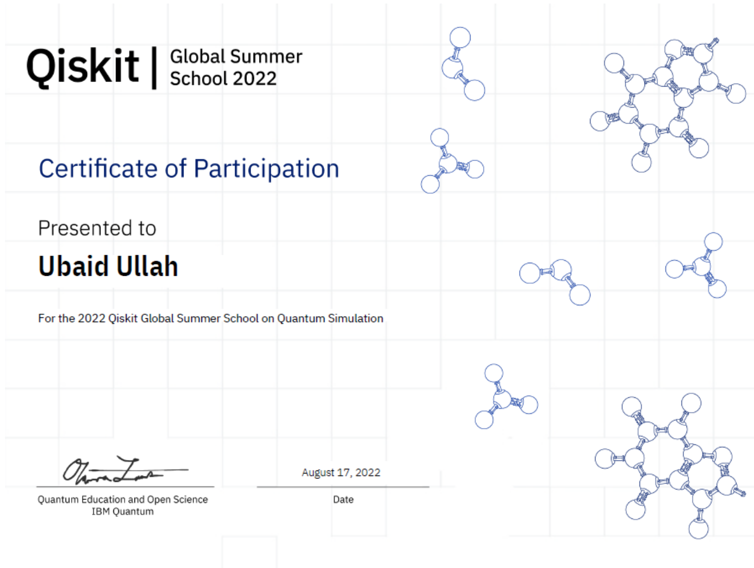

Figure 6.3: Certificate of qiskit global summer school.

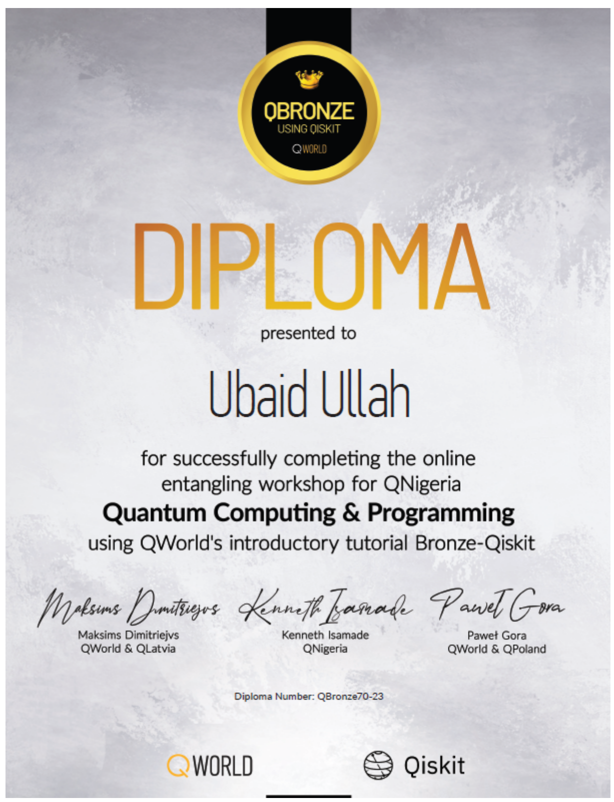

Figure 6.4: Work shop certificate of quantum computing and programming.

This experience not only expanded my knowledge of QML but also allowed me to contribute to cutting-edge research, enhancing my skills in quantum computing and its applications.

TUDelft

To Whom it May Concern

Dr. Sebastian Feld
Assistant Professor
Delft University of Technology
Faculty of Electrical Engineering,
Mathematics and Computer Science
Quantum & Computer Engineering
Mekelweg 4, 2628 CD Delft
s.feld@tudelft.nl
+31 15 278 2462

February 29, 2024

Certificate of Internship Completion

Dear Sir or Madam,

This is to certify that **Ubaid Ullah** has successfully completed an internship at **Quantum Machine Learning Group at Delft University of Technology** from **01/12/2023 to 29/02/2024**.

During the internship period, he actively participated in the group's activities.

We wish Ubaid Ullah continued success in his future endeavors, and we are confident that he will excel in his chosen field.

Yours sincerely,

Sebastian Feld

Figure 6.5: Certificate of international mobility.

6.3 Recommendation

Based on the outcomes of this research, several recommendations can be made for future work in applying QML models to healthcare aplication. As quantum computing technology continues to evolve, it is essential to explore more advanced quantum algorithms and hybrid quantum-classical approaches that can further enhance computational efficiency and predictive accuracy. Future studies should consider leveraging larger and more diverse healthcare datasets to validate the scalability and generalizability of QML models in real-world clinical scenarios. Additionally, the integration of QDL architectures could be explored to handle more complex tasks such as multi-class classification and time-series analysis in healthcare applications. Given the current limitations of quantum hardware, including qubit decoherence and noise, efforts should be directed toward developing noise-resilient quantum algorithms and optimizing quantum circuits for deployment on near-term quantum devices. Furthermore, collaborations between quantum computing researchers and healthcare professionals are recommended to ensure that the developed models address practical clinical challenges and align with healthcare industry standards. Lastly, ethical considerations, including data privacy and model interpretability, should be prioritized to ensure responsible and trustworthy use of QML in healthcare.

6.4 Thesis Conclusions, Limitation and Future Work

In summary, the objective of this research is to demonstrate the use of quantum computing with machine learning for real-world healthcare applications, which are significantly crucial to address. Beginning with the definition of quantum computing, machine learning, and the formulation of research hypotheses. The first specific objective has been full filled by conducted a detailed systematic review of the literature based on the role of QML in healthcare, involving the formulation of research questions and the evaluation of the quality of the article. The outcome of this SLR creates a foundation for future research, highlighting QML's potential to address healthcare challenges through quantum computing and machine learning integration.

The second and third specific objectives are based on the first case study, which considers the two QML algorithms E-QSVM and QRF applied to the real COVID-19 and datasets. After pre-processing and selecting the top 10 features using Chi-Square, ANOVA, and classical models, the E-QSVM model uses a parameterized quantum circuit with a ZZ feature map to capture complex data relationships. The QRF model, implemented in Pennylane with two parallel QPUs, selects the most accurate prediction. The results show, that both models outperform their classical counterparts. The second case study which is the results of the fourth and fifth specific objectives contains a quantum fully convolutional neural network (FCQ-CNN) for Ischemic Heart Disease (IHD) classification. The model combines quantum and classical convolutional layers, a pooling layer, and a fully connected layer. The experimental result shows that the proposed model performs slightly better than the classical models and the previous published quantum models. The third case study replicates the sixth and seventh objectives and explores two quantum models: QKNN and AE-VQC. Both use amplitude encoding and were tested on datasets with 8, 16, and 32 features. QKNN identifies the K nearest neighbors using a quantum K minimum-finding algorithm and swap test for fidelity assessment. AE-VQC features a variational quantum circuit with a COBYLA optimizer. Both models perform better than the classical and the previous published quantum models.

Moreover, in addition to the advantages of quantum computing, there are still various limitations and challenges associated with this field. The access of a limited number of qubits of a quantum computer restricts us to the limited number of features. Current quantum devices are relatively small and noisy, limiting their computational capabilities. Medical research frequently involves large scale datasets, which are difficult for current quantum technology and algorithms to handle. Moreover, medical data analysis requires careful consideration of important feature selection, and conventional methods may not be the best for quantum models. Currently, the access to the real hardware of a quantum computer is also quite complex and expensive.

Future studies should concentrate on creating quantum algorithms that can effectively handle and examine larger datasets, possibly using hybrid quantum-classical methodologies. This requires improving quantum error correction techniques and enhancing quantum circuits for progressively complex models. The development of quantum based feature selection methods designed specifically for high-dimensional medical data could be the main goal of research. This could involve creating novel approaches that use quantum entanglement and superposition for feature extraction or creating quantum versions of conventional methods like QPCA. To minimize noise and improve accessibility for researchers, the existing quantum computer hardware must be carefully optimized, as it is really noisy. Additionally, more research should concentrate on simplifying quantum models in order to reduce the cost of quantum computers.

6.5 Funding Declaration

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie grant agreement Number 847624. In addition, a number of institutions back and co-finance this project. Furthermore, the Basque government to the eVIDA research group IT 1536-22, University of Deusto and coordinated by Basurto Hospital in Bilbao, which provide the database. The Clinical Research Ethics Committee of Euskadi (PI202031) in Spain validated the research protocols. The article reflects only the author's view and the Agency is not responsible for any use that may be made of the information it contains.

Bibliography

[1] Hu, Wei. "Comparison of two quantum nearest neighbor classifiers on IBM's quantum simulator." Natural Science 10, no. 3 (2018): 87-98.

[2] Gibney, Elizabeth. "Quantum computer race intensifies as alternative technology gains steam." Nature 587, no. 7834 (2020): 342-344.

[3] Córcoles, Antonio D., Abhinav Kandala, Ali Javadi-Abhari, Douglas T. McClure, Andrew W. Cross, Kristan Temme, Paul D. Nation, Matthias Steffen, and Jay M. Gambetta. "Challenges and opportunities of near-term quantum computing systems." Proceedings of the IEEE 108, no. 8 (2019): 1338-1352.

[4] Ajagekar, Akshay, and Fengqi You. "Quantum computing for energy systems optimization: Challenges and opportunities." Energy 179 (2019): 76-89.

[5] Ur Rasool, Raihan, Hafiz Farooq Ahmad, Wajid Rafique, Adnan Qayyum, Junaid Qadir, and Zahid Anwar. "Quantum computing for healthcare: A review." Future Internet 15, no. 3 (2023): 94.

[6] Shara, Jollanda. "Some Notes on Quantum Machine Learning." (2023).

[7] Ray, Susmita. "A quick review of machine learning algorithms." In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon), pp. 35-39. IEEE, 2019.

[8] Solenov, Dmitry, Jay Brieler, and Jeffrey F. Scherrer. "The potential of quantum computing and machine learning to advance clinical research and change the practice of medicine." Missouri medicine 115, no. 5 (2018): 463.

[9] Shor, Peter W. "Introduction to quantum algorithms." In Proceedings of Symposia in Applied Mathematics, vol. 58, pp. 143-160. 2002.

[10] Jones, Jonathan A., Michele Mosca, and Rasmus H. Hansen. "Implementation of a quantum search algorithm on a quantum computer." Nature 393, no. 6683 (1998): 344-346.

[11] Benenti, Giuliano, Giulio Casati, Davide Rossini, and Giuliano Strini. Principles of quantum computation and information: a comprehensive textbook. 2019.

[12] Myrvold, Wayne C., Joy Christian, and Lucien Hardy. "Quantum gravity computers: On the theory of computation with indefinite causal structure." Quantum reality, relativistic causality, and closing the epistemic circle: essays in honour of Abner Shimony (2009): 379-401.

[13] Pophale, Shruti S., and Amit Gadekar. "Quantum Machine Learning Algorithms for Diagnostic Applications: A Review." In International Virtual Conference on Industry 4.0: Select Proceedings of IVCI4. 0 2021, pp. 23-32. Singapore: Springer Nature Singapore, 2023.

[14] Nguyen, Quoc Chuong, Lan Nguyen Tran, and Hung Q. Nguyen. "Qsun: an open-source platform towards practical quantum machine learning applications." Machine Learning: Science and Technology 3, no. 1 (2022): 015034.

[15] Batra, Kushal, Kimberley M. Zorn, Daniel H. Foil, Eni Minerali, Victor O. Gawriljuk, Thomas R. Lane, and Sean Ekins. "Quantum machine learning algorithms for drug discovery applications." Journal of chemical information and modeling 61, no. 6 (2021): 2641-2647.

[16] Dunjko, Vedran, and Hans J. Briegel. "Machine learning and artificial intelligence in the quantum domain: a review of recent progress." Reports on Progress in Physics 81, no. 7 (2018): 074001.

[17] Kaushik, Keshav, and Adarsh Kumar. "Demystifying quantum blockchain for healthcare." Security and Privacy 6, no. 3 (2023): e284.

[18] Ciliberto, Carlo, Mark Herbster, Alessandro Davide Ialongo, Massimiliano Pontil, Andrea Rocchetto, Simone Severini, and Leonard Wossnig. "Quantum machine learning: a classical perspective." Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 474, no. 2209 (2018): 20170551.

[19] Schuld, Maria. "Supervised quantum machine learning models are kernel methods." arXiv preprint arXiv:2101.11020 (2021).

[20] Schuld, Maria, Ryan Sweke, and Johannes Jakob Meyer. "Effect of data encoding on the expressive power of variational quantum-machine-learning models." Physical Review A 103, no. 3 (2021): 032430.

[21] Eklund, Peter W. "Comparative study of public-domain supervised machine-learning accuracy on the UCI database." Data Mining and Knowledge Discovery: Theory, Tools, and Technology 3695 (1999): 39-50.

[22] Cohen, Gregory, Saeed Afshar, Jonathan Tapson, and Andre Van Schaik. "EMNIST: Extending MNIST to handwritten letters." In 2017 international joint conference on neural networks (IJCNN), pp. 2921-2926. IEEE, 2017.

[23] Solenov, Dmitry, Jay Brieler, and Jeffrey F. Scherrer. "The potential of quantum computing and machine learning to advance clinical research and change the practice of medicine." Missouri medicine 115, no. 5 (2018): 463.

[24] Si, Yuqi, Jingcheng Du, Zhao Li, Xiaoqian Jiang, Timothy Miller, Fei Wang, W. Jim Zheng, and Kirk Roberts. "Deep representation learning of patient data from Electronic Health Records (EHR): A systematic review." Journal of biomedical informatics 115 (2021): 103671.

[25] Raghupathi, Wullianallur, and Viju Raghupathi. "Big data analytics in healthcare: promise and potential." Health information science and systems 2 (2014): 1-10.

[26] Hussain, Shah, Iqra Mubeen, Niamat Ullah, Syed Shahab Ud Din Shah, Bakhtawar Abduljalil Khan, Muhammad Zahoor, Riaz Ullah, Farhat Ali Khan, and Mujeeb A. Sultan. "Modern diagnostic imaging technique applications and risk factors in the medical field: A review." BioMed Research International 2022 (2022).

[27] Naja, Khadijeh, Susanne F. Yelin, and Xun Gao. "The development of quantum machine learning." (2022).

[28] Schuld, Maria. "Quantum machine learning for supervised pattern recognition." PhD diss., 2017.

[29] Khang, Alex, Kali Charan Rath, Karteek Madapana, Jagannadha Rao, Lakshmi Prasad Panda, and Subhasish Das. "Quantum Computing and Portfolio Optimization in Finance Services." Shaping Cutting-Edge Technologies and Applications for Digital Banking and Financial Services (2025): 27.

[30] Francq, Christian, and Jean-Michel Zakoïan. "Looking for efficient QML estimation of conditional VaRs at multiple risk levels." Annals of Economics and Statistics/Annales d'Économie et de Statistique 123/124 (2016): 9-28.

[31] Innan, Nouhaila, Muhammad Al-Zafar Khan, and Mohamed Bennai. "Financial fraud detection: a comparative study of quantum machine learning models." International Journal of Quantum Information 22, no. 02 (2024): 2350044.

[32] Seol, Jongho, Jongyeop Kim, and Abhilash Kancharla. "Integrating Blockchain-Based Security and Privacy with QML in Edge Computing." Computer Applications in Industry and Engineering (2024): 86.

[33] Hossain, Forhad, Kamrul Hasan, Al Amin, and Shakik Mahmud. "Quantum Machine Learning for Enhanced Cybersecurity: Proposing a Hypothetical Framework for Next-Generation Security Solutions." Journal of Technologies Information and Communication 4, no. 1 (2024): 32222.

[34] Naseri, Mosayeb, Sergey Gusarov, and D. R. Salahub. "Quantum Machine Learning in Materials Prediction: A Case Study on ABO3 Perovskite Structures." The Journal of Physical Chemistry Letters 14, no. 31 (2023): 6940-6947.

[35] Huang, Bing, Nadine O. Symonds, and O. Anatole von Lilienfeld. "Quantum machine learning in chemistry and materials." Handbook of Materials Modeling: Methods: Theory and Modeling (2020): 1883-1909.

[36] Akrom, Muhamad, Supriadi Rustad, Totok Sutojo, Hermawan Kresno Dipojono, Ryo Maezono, and Moses Solomon. "Quantum machine learning for corrosion resistance in stainless steel." Materials Today Quantum 3 (2024): 100013.

[37] Rahman, Syed Masiur, Omar Hamad Alkhalaf, Md Shafiul Alam, Surya Prakash Tiwari, Md Shafiullah, Sarah Mohammed Al-Judaibi, and Fahad Saleh Al-Ismail. "Climate Change Through Quantum Lens: Computing and Machine Learning." Earth Systems and Environment (2024): 1-18.

[38] Grzesiak, Marek, and Param Thakkar. "Flood prediction using classical and quantum machine learning models." arXiv preprint arXiv:2407.01001 (2024).

[39] Maheshwari, Danyal, Begonya Garcia-Zapirain, and Daniel Sierra-Sosa. "Quantum machine learning applications in the biomedical domain: A systematic review." Ieee Access (2022).

[40] García, David Peral, Juan Cruz-Benito, and Francisco José García-Peñalvo. "Systematic literature review: Quantum machine learning and its applications." arXiv preprint arXiv:2201.04093 (2022).

[41] Houssein, Essam H., et al. "Machine learning in the quantum realm: The state-of-the-art, challenges, and future vision." Expert Systems with Applications 194 (2022): 116512.

[42] Khan, Tariq M., and Antonio Robles-Kelly. "Machine learning: Quantum vs classical." IEEE Access 8 (2020): 219275-219294.

[43] Wohlin, Claes. "Guidelines for snowballing in systematic literature studies and a replication in software engineering." In Proceedings of the 18th international conference on evaluation and assessment in software engineering, pp. 1-10. 2014.

[44] Benioff, Paul. "The computer as a physical system: A microscopic quantum mechanical Hamiltonian model of computers as represented by Turing machines." Journal of statistical physics 22 (1980): 563-591.

[45] Feynman, Richard P. "Simulating physics with computers." In Feynman and computation, pp. 133-153. CRC Press, 2018.

[46] Deutsch, David. "Quantum theory, the Church–Turing principle and the universal quantum computer." Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences 400, no. 1818 (1985): 97-117.

[47] Deutsch, David, and Richard Jozsa. "Rapid solution of problems by quantum computation." Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences 439, no. 1907 (1992): 553-558.

[48] Bernstein, Ethan, and Umesh Vazirani. "Quantum complexity theory." In Proceedings of the twenty-fifth annual ACM symposium on Theory of computing, pp. 11-20. 1993.

[49] Simon, Daniel R. "On the power of quantum computation." SIAM journal on computing 26, no. 5 (1997): 1474-1483.

[50] Rieffel, Eleanor, and Wolfgang Polak. "An introduction to quantum computing for non-physicists." ACM Computing Surveys (CSUR) 32, no. 3 (2000): 300-335.

[51] Boghosian, Bruce M., and Washington Taylor IV. "Simulating quantum mechanics on a quantum computer." Physical D: Nonlinear Phenomena 120, no. 1-2 (1998): 30-42.

[52] Nielsen, Michael A., and Isaac Chuang. "Quantum computation and quantum information." (2002): 558-559.

[53] Delgado, Andrea, Kathleen E. Hamilton, Jean-Roch Vlimant, Duarte Magano, Yasser Omar, Pedrame Bargassa, Anthony Francis et al. "Quantum Computing for Data Analysis in High-Energy Physics." arXiv preprint arXiv:2203.08805 (2022).

[54] Yu, Hualong, Xibei Yang, Shang Zheng, and Changyin Sun. "Active learning from imbalanced data: A solution of online weighted extreme learning machine." IEEE transactions on neural networks and learning systems 30, no. 4 (2018): 1088-1103.

[55] Lloyd, Seth, Masoud Mohseni, and Patrick Rebentrost. "Quantum algorithms for supervised and unsupervised machine learning." arXiv preprint arXiv:1307.0411 (2013).

[56] Dunjko, Vedran, Jacob M. Taylor, and Hans J. Briegel. "Quantum-enhanced machine learning." Physical review letters 117, no. 13 (2016): 130501.

[57] Benedetti, Marcello, John Realpe-Gómez, Rupak Biswas, and Alejandro Perdomo-Ortiz. "Quantum-assisted learning of hardware-embedded probabilistic graphical models." Physical Review X 7, no. 4 (2017): 041052.

[58] Farhi, Edward, and Hartmut Neven. "Classification with quantum neural networks on near term processors." arXiv preprint arXiv:1802.06002 (2018).

[59] Schuld, Maria, Ilya Sinayskiy, and Francesco Petruccione. "An introduction to quantum machine learning." Contemporary Physics 56, no. 2 (2015): 172-185.

[60] Schuld, Maria, Francesco Petruccione, Maria Schuld, and Francesco Petruccione. "Quantum models as kernel methods." Machine Learning with Quantum Computers (2021): 217-245.

[61] Heredge, Jamie, Charles Hill, Lloyd Hollenberg, and Martin Sevior. "Quantum support vector machines for continuum suppression in B meson decays." Computing and Software for Big Science 5, no. 1 (2021): 27.

[62] Kak, Subhash C. "Quantum neural computing." Advances in imaging and electron physics 94 (1995): 259-313.

[63] Ezhov, Alexandr A., and Dan Ventura. "Quantum neural networks." Future Directions for Intelligent Systems and Information Sciences: The Future of Speech and Image Technologies, Brain Computers, WWW, and Bioinformatics (2000): 213-235.

[64] Ma, Yan-zhu, Hong-fei Song, and Jun Zhang. "Quantum algorithm for k-nearest neighbors classification based on the categorical tensor network states." International Journal of Theoretical Physics 60 (2021): 1164-1174.

[65] Chen, Samuel Yen-Chi, Shinjae Yoo, and Yao-Lung L. Fang. "Quantum long short-term memory." In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 8622-8626. IEEE, 2022.

[66] Huang, He-Liang, Yuxuan Du, Ming Gong, Youwei Zhao, Yulin Wu, Chaoyue Wang, Shaowei Li et al. "Experimental quantum generative adversarial networks for image generation." Physical Review Applied 16, no. 2 (2021): 024051.

[67] Borras, Kerstin, Su Yeon Chang, Lena Funcke, Michele Grossi, Tobias Hartung, Karl Jansen, Dirk Kruecker et al. "Impact of quantum noise on the training of quantum generative adversarial networks." In Journal of Physics: Conference Series, vol. 2438, no. 1, p. 012093. IOP Publishing, 2023.

[68] Scott, Tony C., Madhusudan Therani, and Xing M. Wang. "Data clustering with quantum mechanics." Mathematics 5, no. 1 (2017): 5.

[69] Havlíček, Vojtěch, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. "Supervised learning with quantum-enhanced feature spaces." Nature 567, no. 7747 (2019): 209-212.

[70] Schuld, Maria, and Nathan Killoran. "Quantum machine learning in feature hilbert spaces." Physical review letters 122, no. 4 (2019): 040504.

[71] Schuld, Maria, Alex Bocharov, Krysta M. Svore, and Nathan Wiebe. "Circuit-centric quantum classifiers." Physical Review A 101, no. 3 (2020): 032308.

[72] Liu, Leroy F. "DNA topoisomerase poisons as antitumor drugs." Annual review of biochemistry 58, no. 1 (1989): 351-375.

[73] Niedermaier, Tobias, Thomas Gredner, Sabine Kuznia, Ben Schöttker, Ute Mons, and Hermann Brenner. "Vitamin D supplementation to the older adult population in Germany has the cost-saving potential of preventing almost 30 000 cancer deaths per year." Molecular oncology 15, no. 8 (2021): 1986-1994.

[74] Newman-Toker, David E., Zheyu Wang, Yuxin Zhu, Najlla Nassery, Ali S. Saber Tehrani, Adam C. Schaffer, Chihwen Winnie Yu-Moe, Gwendolyn D. Clemens, Mehdi Fanai, and Dana Siegal. "Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the “Big Three”." Diagnosis 8, no. 1 (2021): 67-84.

[75] Srikanth, P., and Adarsh Kumar. "Secure Quantum Computing for Healthcare Sector: A Short Analysis." arXiv preprint arXiv:2211.10027 (2022).

[76] Azzaoui, Abir EL, Pradip Kumar Sharma, and Jong Hyuk Park. "Blockchain-based delegated Quantum Cloud architecture for medical big data security." Journal of Network and Computer Applications 198 (2022): 103304.

[77] Yuan, Xi, Chengming Yang, Qian He, Junhu Chen, Dongmei Yu, Jie Li, Shiyao Zhai et al. "Current and perspective diagnostic techniques for COVID-19." ACS infectious diseases 6, no. 8 (2020): 1998-2016.

[78] Ruffino, Andrea, Tsung-Yeh Yang, John Michniewicz, Yatao Peng, Edoardo Charbon, and Miguel Fernando Gonzalez-Zalba. "A cryo-CMOS chip that integrates silico quantum dots and multiplexed dispersive readout electronics." Nature Electronics 5, no. 1 (2022): 53-59.

[79] Munshi, Manushi, Rajesh Gupta, Nilesh Kumar Jadav, Zdzislaw Polkowski, Sudeep Tanwar, Fayez Alqahtani, and Wael Said. "Quantum machine learning-based framework to detect heart failures in Healthcare 4.0." Software: Practice and Experience 54, no. 2 (2024): 168-185.

[80] Enad, Huda Ghazi, and Mazin Abed Mohammed. "Cloud computing-based framework for heart disease classification using quantum machine learning approach." Journal of Intelligent Systems 33, no. 1 (2024): 20230261.

[81] Maheshwari, Danyal, Ubaid Ullah, Pablo A. Osorio Marulanda, Alain García-Olea Jurado, Ignacio Diez Gonzalez, Jose M. Ormaetxe Merodio, and Begonya Garcia-Zapirain. "Quantum Machine Learning Applied to Electronic Healthcare Records for Ischemic Heart Disease Classification." Hum.-Cent. Comput. Inf. Sci 13, no. 06 (2023).

[82] Prakash, B. B. N. S., B. Naveen, M. D. Akhiluzzama, and Pothuraju Rajarajeswari. "Comparative Performance Analysis of Quantum Algorithm with Machine learning Algorithms on Diabetes Mellitus." In 2023 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), pp. 1178-1183. IEEE, 2023.

[83] Abdulsalam, Ghada, Souham Meshoul, and Hadil Shaiba. "Explainable Heart Disease Prediction Using Ensemble-Quantum Machine Learning Approach." Intelligent Automation and Soft Computing 36, no. 1 (2023).

[84] Díaz-Santos, Sonia, and Daniel Escanez-Exposito. "Classical vs. Quantum Machine Learning for Breast Cancer Detection." In 2023 19th International Conference on the Design of Reliable Communication Networks (DRCN), pp. 1-5. IEEE, 2023.

[85] Shahwar, Tayyaba, Junaid Zafar, Ahmad Almogren, Haroon Zafar, Ateeq Ur Rehman, Muhammad Shafiq, and Habib Hamam. "Automated detection of Alzheimer's via hybrid classical quantum neural networks." Electronics 11, no. 5 (2022): 721.

[86] Gupta, Himanshu, Hirdesh Varshney, Tarun Kumar Sharma, Nikhil Pachauri, and Om Prakash Verma. "Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction." Complex Intelligent Systems 8, no. 4 (2022): 3073-3087.

[87] Heidari, Hanif, and Gerhard Hellstern. "Early heart disease prediction using hybrid quantum classification." arXiv preprint arXiv:2208.08882 (2022).

[88] Krunic, Zoran, Frederik F. Flöther, George Seegan, Nathan D. Earnest-Noble, and Omar Shehab. "Quantum kernels for real-world predictions based on electronic health records." IEEE Transactions on Quantum Engineering 3 (2022): 1-11.

[89] Ishwarya, M. S., and Aswani Kumar Cherukuri. "Quantum-inspired ensemble approach to multi-attributed and multi-agent decision-making." Applied Soft Computing 106 (2021): 107283.

[90] Pomarico, Domenico, Annarita Fanizzi, Nicola Amoroso, Roberto Bellotti, Albino Biafora, Samantha Bove, Vittorio Didonna et al. "A proposal of quantum-inspired machine learning for medical purposes: An application case." Mathematics 9, no. 4 (2021): 410.

[91] Saini, Shivani, P. K. Khosla, Manjit Kaur, and Gurmohan Singh. "Quantum driven machine learning." International Journal of Theoretical Physics 59, no. 12 (2020): 4013-4024.

[92] Chakraborty, Sanjay, Soharab Hossain Shaikh, Amlan Chakrabarti, and Ranjan Ghosh. "A hybrid quantum feature selection algorithm using a quantum inspired graph theoretic approach." Applied Intelligence 50, no. 6 (2020): 1775-1793.

[93] Maheshwari, Danyal, Begoña Garcia-Zapirain, and Daniel Sierra-Soso. "Machine learning applied to diabetes dataset using Quantum versus Classical computation." In 2020 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), pp. 1-6. IEEE, 2020.

[94] Sierra-Sosa, Daniel, Juan D. Arcila-Moreno, Begonya Garcia-Zapirain, and Adel Elmaghraby. "Diabetes type 2: Poincaré Data preprocessing for quantum machine learning." Comput., Mater. Continua 67, no. 2 (2021): 1849-1861.

[95] Mardirosian, Sevak. "Quantum-enhanced Supervised Learning with Variational Quantum Circuits." PhD diss., PhD thesis. Leiden University, 2019.

[96] Chen, Zhao-Yun, Cheng Xue, Si-Ming Chen, and Guo-Ping Guo. "Vqnet: Library for a quantum-classical hybrid neural network." arXiv preprint arXiv:1901.09133 (2019).

[97] Benlamine, Kaoutar, Younès Bennani, Ahmed Zaiou, Mohamed Hibti, Basarab Matei, and Nistor Grozavu. "Distance estimation for quantum prototypes based clustering." In Neural Information Processing: 26th International Conference, ICONIP 2019, Sydney, NSW, Australia, December 12–15, 2019, Proceedings, Part III 26, pp. 561-572. Springer International Publishing, 2019.

[98] Patel, Om Prakash, Neha Bharill, Aruna Tiwari, and Mukesh Prasad. "A novel quantum-inspired fuzzy based neural network for data classification." IEEE Transactions on Emerging Topics in Computing 9, no. 2 (2019): 1031-1044.

[99] Naguleswaran, Sanjeev. "Machine Learning and Quantum Intelligence for Health Data Scenarios." arXiv preprint arXiv:2410.21339 (2024).

[100] Roy, Suman Kumar, and Bhawana Rudra. "Quantum-inspired hybrid algorithm for image classification and segmentation: Q-Means++ max-cut method." International Journal of Imaging Systems and Technology 34, no. 1 (2024): e23015.

[101] Alsharabi, Naif, Tayyaba Shahwar, Ateeq Ur Rehman, and Yasser Alharbi. "Implementing Magnetic Resonance Imaging Brain Disorder Classification via AlexNet–Quantum Learning." Mathematics 11, no. 2 (2023): 376.

[102] Deshmukh, Shradha, Bikash K. Behera, Preeti Mulay, Emad A. Ahmed, Saif Al-Kuwari, Prayag Tiwari, and Ahmed Farouk. "Explainable quantum clustering method to model medical data." Knowledge-Based Systems 267 (2023): 110413.

[103] Konar, Debanjan, Siddhartha Bhattacharyya, Tapan K. Gandhi, Bijaya K. Panigrahi, and Richard Jiang. "3-D Quantum-Inspired Self-Supervised Tensor Network for Volumetric Segmentation of Medical Images." IEEE Transactions on Neural Networks and Learning Systems (2023).

[104] Mishra, Awanish Kumar, Indresh Kumar Gupta, Tarun Dhar Diwan, and Swati Srivastava. "Cervical precancerous lesion classification using quantum invasive weed optimization with deep learning on biomedical pap smear images." Expert Systems (2023): e13308.

[105] Moradi, Sasan, Christoph Brandner, Clemens Spielvogel, Denis Krajnc, Stefan Hillmich, Robert Wille, Wolfgang Drexler, and Laszlo Papp. "Clinical data classification with noisy intermediate scale quantum computers." Scientific Reports 12, no. 1 (2022): 1851.

[106] Houssein, Essam H., Zainab Abohashima, Mohamed Elhoseny, and Waleed M. Mohamed. "Hybrid quantum-classical convolutional neural network model for COVID-19 prediction using chest X-ray images." Journal of Computational Design and Engineering 9, no. 2 (2022): 343-363.

[107] Amin, Javaria, Muhammad Almas Anjum, Nadia Gul, and Muhammad Sharif. "A secure two-qubit quantum model for segmentation and classification of brain tumor using MRI images based on blockchain." Neural Computing and Applications 34, no. 20 (2022): 17315-17328.

[108] Toledo-Cortés, Santiago, Diego H. Useche, Henning Müller, and Fabio A. González. "Grading diabetic retinopathy and prostate cancer diagnostic images with deep quantum ordinal regression." Computers in biology and medicine 145 (2022): 105472.

[109] Bruno, J. Reyes, Francisco Torres-Hoyos, and Rubén Baena-Navarro. "Use of the quantum cluster algorithm and scaling dynamics in magnetic resonance imaging for prostate cancer staging." In Journal of Physics: Conference Series, vol. 2046, no. 1, p. 012007. IOP Publishing, 2021.

[110] Acar, Erdi, and Ihsan Yilmaz. "COVID-19 detection on IBM quantum computer with classical-quantum transferlearning." Turkish Journal of Electrical Engineering and Computer Sciences 29, no. 1 (2021): 46-61.

[111] Mathur, Natansh, Jonas Landman, Yun Yvonna Li, Martin Strahm, Skander Kazdaghli, Anupam Prakash, and Iordanis Kerenidis. "Medical image classification via quantum neural networks." arXiv preprint arXiv:2109.01831 (2021).

[112] Selvan, Raghavendra, Silas Ørting, and Erik B. Dam. "Locally orderless tensor networks for classifying two-and three-dimensional medical images." arXiv preprint arXiv:2009.12280 (2020).

[113] Adhikary, Soumik, Siddharth Dangwal, and Debanjan Bhowmik. "Supervised learning with a quantum classifier using multi-level systems." Quantum Information Processing 19 (2020): 1-12.

[114] Dutt, Vishal, Sriramakrishnan Chandrasekaran, and Vicente García Díaz. "Quantum neural networks for disease treatment identification." European Journal of Molecular and Clinical Medicine 7, no. 11 (2020): 57-67.

[115] Chakraborty, Sanjay, Soharab Hossain Shaikh, Amlan Chakrabarti, and Ranjan Ghosh. "A hybrid quantum feature selection algorithm using a quantum inspired graph theoretic approach." Applied Intelligence 50, no. 6 (2020): 1775-1793.

[116] Li, Yangyang, Junjie Xiao, Yanqiao Chen, and Licheng Jiao. "Evolving deep convolutional neural networks by quantum behaved particle swarm optimization with binary encoding for image classification." Neurocomputing 362 (2019): 156-165.

[117] Ding, Chen, Tian-Yi Bao, and He-Liang Huang. "Quantum-inspired support vector machine." IEEE Transactions on Neural Networks and Learning Systems 33, no. 12 (2021): 7210-7222.

[118] Benedetti, Marcello, Delfina Garcia-Pintos, Oscar Perdomo, Vicente Leyton-Ortega, Yunseong Nam, and Alejandro Perdomo-Ortiz. "A generative modeling approach for benchmarking and training shallow quantum circuits." npj Quantum Information 5, no. 1 (2019): 45.

[119] Sagheer, Alaa, Mohammed Zidan, and Mohammed M. Abdelsamea. "A novel autonomous perceptron model for pattern classification applications." Entropy 21, no. 8 (2019): 763.

[120] Mahmud, Mufti, and Stefano Vassanelli. "Processing and analysis of multichannel extracellular neuronal signals: state-of-the-art and challenges." Frontiers in neuroscience 10 (2016): 248.

[121] Major, Tyler C., and James M. Conrad. "A survey of brain computer interfaces and their applications." In IEEE SOUTHEASTCON 2014, pp. 1-8. IEEE, 2014.

[122] Zemouri, Ryad, Noureddine Zerhouni, and Daniel Racoceanu. "Deep learning in the biomedical applications: Recent and future status." Applied Sciences 9, no. 8 (2019): 1526.

[123] DL, TRUMP. "Randomized controlled trial of adjuvant therapy with lymphoblastoid interferon in resected high risk renal cell carcinoma: An EOCG study." In Proc. Am. Soc. Clin. Oncol., vol. 15, p. 253. 1996.

[124] Berkaya, Selcan Kaplan, Alper Kursat Uysal, Efnan Sora Gunal, Semih Ergin, Serkan Gunal, and M. Bilginer Gulmezoglu. "A survey on ECG analysis." Biomedical Signal Processing and Control 43 (2018): 216-235.

[125] Teplan, Michal. "Fundamentals of EEG measurement." Measurement science review 2, no. 2 (2002): 1-11.

[126] Wilcox, Teresa, and Marisa Biondi. "fNIRS in the developmental sciences." Wiley Interdisciplinary Reviews: Cognitive Science 6, no. 3 (2015): 263-283.

[127] Hefftner, Gisela, Walter Zucchini, and George G. Jaros. "The electromyogram (EMG) as a control signal for functional neuromuscular stimulation. I. Autoregressive modeling as a means of EMG signature discrimination." IEEE Transactions on Biomedical Engineering 35, no. 4 (1988): 230-237.

[128] Baygin, Nursena, Emrah Aydemir, Prabal D. Barua, Mehmet Baygin, Sengul Dogan, Turker Tuncer, Ru-San Tan, and U. Rajendra Acharya. "Automated mental arithmetic performance detection using quantum pattern-and triangle pooling techniques with EEG signals." Expert Systems with Applications 227 (2023): 120306.

[129] Sameer, Mustafa, and Bharat Gupta. "A Novel Hybrid Classical-Quantum Network to Detect Epileptic Seizures." medRxiv (2022): 2022-05.

[130] Saggi, Mandeep Kaur, and Sabre Kais. "MQML: Multi-Omic Quantum Machine Learning Based Cancer Classification, Biomarker Identification in Human Lung Adenocarcinoma." In 2024 IEEE International Conference on Quantum Computing and Engineering (QCE), vol. 1, pp. 1713-1720. IEEE, 2024.

[131] Tripathi, Sarvapriya M., Himanshu Upadhyay, and Jayesh Soni. "Advanced Cardiovascular Health in a Quantum AI-driven Healthcare Framework." Authorea Preprints (2024).

[132] Koike-Akino, Toshiaki, and Ye Wang. "quEEGNet: Quantum AI for Biosignal Processing." In 2022 IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI), pp. 01-04. IEEE, 2022.

[133] Padha, Anupama, and Anita Sahoo. "A Parametrized Quantum LSTM Model for Continuous Stress Monitoring." In 2022 9th International Conference on Computing for Sustainable Global Development (INDIACom), pp. 261-266. IEEE, 2022.

[134] Koldijk, Saskia, Maya Sappelli, Suzan Verberne, Mark A. Neerincx, and Wessel Kraaij. "The swell knowledge work dataset for stress and user modeling research." In Proceedings of the 16th international conference on multimodal interaction, pp. 291-298. 2014.

[135] Sridevi, S., T. Kanimozhi, K. Issac, and M. Sudha. "Quanvolution neural network to recognize arrhythmia from 2D scaleogram features of ECG signals." In 2022 International Conference on Innovative Trends in Information Technology (ICITIIT), pp. 1-5. IEEE, 2022.

[136] Nath, Rajdeep Kumar, Himanshu Thapliyal, and Travis S. Humble. "Quantum annealing for automated feature selection in stress detection." In 2021 IEEE Computer Society Annual Symposium on VLSI (ISVLSI), pp. 453-457. IEEE, 2021.

[137] Aishwarya, Seth, Vaishnav Abeer, Babu B. Sathish, and K. N. Subramanya. "Quantum computational techniques for prediction of cognitive state of human mind from EEG signals." Journal of Quantum Computing 2, no. 4 (2020): 157.

[138] Li, YaoChong, Ri-Gui Zhou, RuiQing Xu, Jia Luo, and She-Xiang Jiang. "A quantum mechanics-based framework for EEG signal feature extraction and classification." IEEE transactions on emerging topics in computing 10, no. 1 (2020): 211-222.

[139] Taha, Saleem MR, and Zahraa K. Taha. "EEG signals classification based on autoregressive and inherently quantum recurrent neural network." International Journal of Computer Applications in Technology 58, no. 4 (2018): 340-351.

[140] Bologna, Guido, and Yoichi Hayashi. "Qsvm: A support vector machine for rule extraction." In Advances in Computational Intelligence: 13th International Work-Conference on Artificial Neural Networks, IWANN 2015, Palma de Mallorca, Spain, June 10-12, 2015. Proceedings, Part II 13, pp. 276-289. Springer International Publishing, 2015.

[141] Khadiev, Kamil, and Liliya Safina. "The quantum version of random forest model for binary classification problem." In CEUR Workshop Proc, vol. 2842, pp. 30-35. 2021.

[142] Basheer, Afrad, A. Afham, and Sandeep K. Goyal. "Quantum k-nearest neighbors algorithm." arXiv preprint arXiv:2003.09187 (2020).

[143] Altaisky, M. V. "Quantum neural network." arXiv preprint quant-ph/0107012 (2001).

[144] Havlíček, V., Córcoles, A.D., Temme, K. et al. Supervised learning with quantum-enhanced feature spaces. Nature 567, 209–212 (2019). https://doi.org/10.1038/s41586-019-0980-2

[145] Oh, Seunghyeok, Jaeho Choi, and Joongheon Kim. "A tutorial on quantum convolutional neural networks (QCNN)." In 2020 International Conference on Information and Communication Technology Convergence (ICTC), pp. 236-239. IEEE, 2020.

[146] Wu, Zhihao, Tingting Song, and Yanbing Zhang. "Quantum k-means algorithm based on Manhattan distance." Quantum Information Processing 21, no. 1 (2022): 19.

[147] Kerenidis, Iordanis, and Jonas Landman. "Quantum spectral clustering." Physical Review A 103, no. 4 (2021): 042415.

[148] Cross, Andrew. "The IBM Q experience and QISKit open-source quantum computing software." In APS March meeting abstracts, vol. 2018, pp. L58-003. 2018.

[149] Shin, Seung Woo, Graeme Smith, John A. Smolin, and Umesh Vazirani. "How" quantum" is the D-Wave machine?." arXiv preprint arXiv:1401.7087 (2014).

[150] Efthymiou, Stavros, Sergi Ramos-Calderer, Carlos Bravo-Prieto, Adrián Pérez-Salinas, Diego García-Martín, Artur Garcia-Saez, José Ignacio Latorre, and Stefano Carrazza. "Qibo: a framework for quantum simulation with hardware acceleration." Quantum Science and Technology 7, no. 1 (2021): 015018.

[151] M. Li, S. Ma, and Z. Liu, "A novel method to detect the early warning signal of COVID-19 transmission," BMC Infect. Dis., vol. 22, no. 1, pp. 1–12, Dec. 2022.

[152] M. U. G. Kraemer et al., "The effect of human mobility and control measures on the COVID-19 epidemic in China," Science, vol. 368, no. 6490, pp. 493–497, May 2020.

[153] N. Zhu et al., "A Novel Coronavirus from Patients with Pneumonia in China, 2019," N. Engl. J. Med., vol. 382, no. 8, pp. 727–733, Feb. 2020.

[154] B. Hu, H. Guo, P. Zhou, and Z. L. Shi, "Characteristics of SARS-CoV-2 and COVID-19," Nat. Rev. Microbiol. 2020 193, vol. 19, no. 3, pp. 141–154, Oct. 2020.

[155] N. E. Tabbakha, C. P. Ooi, W. H. Tan, and Y. F. Tan, “A wearable device for machine learning based elderly’s activity tracking and indoor location system,” Bull. Electr. Eng. Informatics, vol. 10, no. 2, pp. 927–939, 2021.

[156] P. Mathur, S. Srivastava, X. Xu, and J. L. Mehta, “Artificial Intelligence, Machine Learning, and Cardiovascular Disease,” Clin. Med. Insights. Cardiol., vol. 14, 2020.

[157] B. Gallo Marin et al., “Predictors of COVID-19 severity: A literature review,” Rev. Med. Virol., vol. 31, no. 1, p. 1, Jan. 2021.

[158] C.Clarissa. "What are the differences among COVID-19, the flu, cold, and allergies?." Pharmacy Today 28.3 (2022): 24-25

[159] “Spanish Royal Decree 463/2020, of March 14th, 2020, declaring the state of alarm | Ramon y Cajal.” [Online]. Available: https://www.ramonycajalabogados.com/en/spanish-royal-decree-4632020-march-14th-2020-declaring-state-alarm. [Accessed: 17-Sep-2022].

[160] P. M. P. Fernandes and A. W. Mariani, “Life post-COVID-19: symptoms and chronic complications,” Sao Paulo Med. J., vol. 139, no. 1, pp. 01–02, Feb. 2021.

[161] Spain Government. "General Secretariat of Health and Consumption. General Directorate of Public Health, Quality and Innovation. Center for the Coordination of Health Alerts and Emergencies. Scientific-technical information. Coronavirus disease, COVID-19 Update November 12, 2020." (2020).

[162] M. Bordbar et al., “Assessment of the HScore as a predictor of disease outcome in patients with COVID-19,” BMC Pulm. Med., vol. 21, no. 1, pp. 1–9, Dec. 2021.

[163] L. Fardet et al., “Development and validation of the HScore, a score for the diagnosis of reactive hemophagocytic syndrome,” Arthritis Rheumatol. (Hoboken, N.J.), vol. 66, no. 9, pp. 2613–2620, 2014.

[164] D. Ravi et al., “Deep Learning for Health Informatics,” IEEE J. Biomed. Heal. Informatics, vol. 21, no. 1, pp. 4–21, Jan. 2017.

[165] P. Easom-Mccaldin, A. Bouridane, A. Belatreche, and R. Jiang, “Towards Building A Facial Identification System Using Quantum Machine Learning Techniques,” J. Adv. Inf. Technol., vol. 13, no. 2, pp. 198–202, Aug. 2020.

[166] X. W. Yao et al., “Quantum image processing and its application to edge detection: Theory and experiment,” Phys. Rev. X, vol. 7, no. 3, p. 031041, Sep. 2017.

[167] L. Lamata, “Quantum machine learning and quantum biomimetics: A perspective,” Mach. Learn. Sci. Technol., vol. 1, no. 3, Apr. 2020.

[168] Z. Hameed and B. Garcia-Zapirain, “Sentiment Classification Using a Single-Layered BiLSTM Model,” IEEE Access, vol. 8, pp. 73992–74001, 2020.

[169] H. Akoglu, “User’s guide to correlation coefficients,” Turkish J. Emerg. Med., vol. 18, no. 3, p. 91, Sep. 2018.

[170] A. Delilbasic, G. Cavallaro, M. Willsch, F. Melgani, M. Riedel, and K. Michielsen, “QUANTUM SUPPORT VECTOR MACHINE ALGORITHMS FOR REMOTE SENSING DATA CLASSIFICATION,” Int. Geosci. Remote Sens. Symp., pp. 2608–2611, 2021.

[171] W. Li and D. L. Deng, “Recent advances for quantum classifiers,” Sci. China Physics, Mech. Astron., vol. 65, no. 2, p. 220301, Feb. 2022.

[172] B. E. Boser, I. M. Guyon, and V. N. Vapnik, “Training algorithm for optimal margin classifiers,” Proc. Fifth Annu. ACM Work. Comput. Learn. Theory, pp. 144–152, 1992.

[173] M. Schuld, “Supervised quantum machine learning models are kernel methods,” pp. 1–25, 2021.

[174] P. Rebentrost, M. Mohseni, and S. Lloyd, “Quantum support vector machine for big data classification,” Phys. Rev. Lett., vol. 113, no. 3, p. 130503, Sep. 2014.

[175] Y. Ruan, X. Xue, H. Liu, J. Tan, and X. Li, “Quantum Algorithm for K-Nearest Neighbors Classification Based on the Metric of Hamming Distance,” Int. J. Theor. Phys., vol. 56, no. 11, pp. 3496–3507, Nov. 2017.

[176] H. Li, N. Jiang, R. Zhang, Z. Wang, and H. Wang, “Quantum Support Vector Machine Based on Gradient Descent,” Int. J. Theor. Phys., vol. 61, no. 3, Mar. 2022.

[177] A. Kariya and B. K. Behera, “Investigation of Quantum Support Vector Machine for Classification in NISQ era,” Dec. 2021.

[178] D. Maheshwari, B. Garcia-Zapirain, and D. Sierra-Soso, “Machine learning applied to diabetes dataset using Quantum versus Classical computation,” 2020 IEEE Int. Symp. Signal Process. Inf. Technol. ISSPIT 2020, 2020.

[179] N. Schetakis, (2020, October 5). Ensemble classification with Forest and Qiskit Devices : Parameters. Xanadu Discussion Forum. Retrieved September 19, 2022, from https://discuss.pennylane.ai/t/ensemble-classification-with-forest-and-qiskit-devices-parameters/608

[180] G. Renugadevi, G. Asha Priya, B. Dhivyaa Sankari, and R. Gowthamani, “Retraction: Predicting heart disease using hybrid machine learning model,” J. Phys. Conf. Ser., vol. 1916, no. 1, 2021, doi: 10.1088/1742-6596/1916/1/012208.

[181] R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles," Expert Syst. Appl., vol. 36, no. 4, pp. 7675–7680, 2009, doi: 10.1016/j.eswa.2008.09.013.

[182] S. Nyamathulla and R. Varikuti, "Predicting Heart Disease with Hybrid Machine Learning Algorithms," SSRN Electron. J., 2022, doi: 10.2139/ssrn.4121828.

[183] D. C. Yadav and S. Pal, "Prediction of heart disease using feature selection and random forest ensemble method," Int. J. Pharm. Res., vol. 12, no. 4, pp. 56–66, 2020, doi: 10.31838/ijpr/2020.12.04.013.

[184] M. S. Oh and M. H. Jeong, "Sex Differences in Cardiovascular Disease Risk Factors among Korean Adults," Korean J. Med., vol. 95, no. 4, pp. 266–275, 2020, doi: 10.3904/kjm.2020.95.4.266.

[185] P. Jousilahti, E. Vartiainen, J. Tuomilehto, and P. Puska, "Sex, Age, Cardiovascular Risk Factors, and Coronary Heart Disease," Circulation, vol. 99, no. 9, pp. 1165–1172, 1999, doi: 10.1161/01.cir.99.9.1165.

[186] "Cardiovascular diseases (cvds)," World Health Organization, 11-Jun-2021. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds). [Accessed: 17-Oct-2022].

[187] K. Uyar and A. Ilhan, "Diagnosis of heart disease using genetic algorithm based trained recurrent fuzzy neural networks," Procedia Comput. Sci., vol. 120, pp. 588–593, 2017, doi: 10.1016/j.procs.2017.11.283.

[188] A. U. Haq, J. P. Li, M. H. Memon, S. Nazir, R. Sun, and I. Garciá-Magarinõ, "A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms," Mob. Inf. Syst., vol. 2018, 2018, doi: 10.1155/2018/3860146.

[189] V. Biksham, V. Srujana, I. Meghana, B. Harshath, and G. Tarun, "Heart Disease Prediction Using Machine Learning," Ymer, vol. 21, no. 4, pp. 489–494, 2022, doi: 10.37896/YMER21.04/48.

[190] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, and P. Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning," Comput. Intell. Neurosci., vol. 2021, 2021, doi: 10.1155/2021/8387680.

[191] A. S. Y. Kong, K. S. Lai, S. H. E. Lim, S. Sivalingam, J. Y. Loh, and S. Maran, "miRNA in Ischemic Heart Disease and Its Potential asBiomarkers: A Comprehensive Review," Int. J. Mol. Sci., vol. 23, no. 16, 2022, doi: 10.3390/ijms23169001.

[192] I. Kononenko, “Machine learning for medical diagnosis: History, state of the art and perspective,” Artif. Intell. Med., vol. 23, no. 1, pp. 89–109, 2001, doi: 10.1016/S0933-3657(01)00077-X.

[193] F. M. Javed Mehedi Shamrat, M. A. Raihan, A. K. M. S. Rahman, I. Mahmud, and R. Akter, “An analysis on breast disease prediction using machine learning approaches,” Int. J. Sci. Technol. Res., vol. 9, no. 2, pp. 2450–2455, 2020.

[194] J. Mackay, G. Mensah, and K. Greenland, “The atlas of heart disease and stroke / Judith Mackay and George Mensah ; with Shanthi Mendis and Kurt Greenland,” World Health Organization, 01-Jan-2004. [Online]. Available: https://apps.who.int/iris/handle/10665/43007. [Accessed: 17-Oct-2022].

[195] M. S. Amin, Y. K. Chiam, and K. D. Varathan, “Identification of significant features and data mining techniques in predicting heart disease,” Telemat. Informatics, vol. 36, no. August 2018, pp. 82–93, 2019, doi: 10.1016/j.tele.2018.11.007.

[196] N. Kausar, S. Palaniappan, B. B. Samir, A. Abdullah, and N. Dey, “Systematic analysis of applied data mining based optimization algorithms in clinical attribute extraction and classification for diagnosis of cardiac patients,” Intell. Syst. Ref. Libr., vol. 96, pp. 217–231, 2016, doi: 10.1007/978-3-319)21212-8-9.

[197] D. Jain and V. Singh, “Feature selection and classification systems for chronic disease prediction: A review,” Egypt. Informatics J., vol. 19, no. 3, pp. 179–189, 2018, doi: 10.1016/j.eij.2018.03.002.

[198] J. Neumann, C. Schnörr, and G. Steidl, “Combined SVM-based feature selection and classification,” Mach. Learn., vol. 61, no. 1–3, pp. 129–150, 2005, doi: 10.1007/s10994-005-1505-9.

[199] M. Ashraf et al., “Prediction of Cardiovascular Disease Through Cutting-Edge Deep Learning Technologies: An Empirical Study Based on TENSORFLOW, PYTORCH and KERAS,” Adv. Intell. Syst. Comput., vol. 1165, pp. 239–255, 2021, doi: 10.1007/978-981-15-5113-0-18.

[200] X. Liu et al., “A Hybrid Classification System for Heart Disease Diagnosis Based on the RFRS Method,” Comput. Math. Methods Med., vol. 2017, 2017, doi: 10.1155/2017/8272091.

[201] N. Mathur et al., “Medical image classification via quantum neural networks,” pp. 1–14, 2021, [Online]. Available: http://arxiv.org/abs/2109.01831

[202] T. A. Shaikh and R. Ali, “Quantum computing in big data analytics: A survey,” Proc. - 2016 16th IEEE Int. Conf. Comput. Inf. Technol. CIT 2016, 2016 6th Int.

Symp. Cloud Serv. Comput. IEEE SC2 2016 2016 Int. Symp. Secur. Priv. Soc. Networks Big Data, Soc. 2016, no. 4, pp. 112–115, 2017, doi: 10.1109/CIT.2016.79.

[203] H. Gupta, H. Varshney, T. K. Sharma, N. Pachauri, and O. P. Verma, "Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction," Complex Intell. Syst., vol. 8, no. 4, pp. 3073–3087, 2022, doi: 10.1007/s40747-021-00398-7.

[204] Z. Hameed, B. Garcia-Zapirain, J. J. Aguirre, and M. A. Isaza-Ruget, "Multiclass classification of breast cancer histopathology images using multilevel features of deep convolutional neural network," Sci. Rep., vol. 12, no. 1, pp. 1–21, 2022, doi: 10.1038/s41598-022-19278-2.

[205] D. Sierra-Sosa, J. Arcila-Moreno, B. Garcia-Zapirain, C. Castillo-Olea, and A. Elmaghraby, "Dementia Prediction Applying Variational Quantum Classifier," pp. 1–12, 2020, [Online]. Available: http://arxiv.org/abs/2007.08653

[206] D. Maheshwari, B. Garcia-Zapirain, and D. Sierra-Sosa, "Quantum Machine Learning Applications in the Biomedical Domain: A Systematic Review," IEEE Access, vol. 10, no. July, pp. 80463–80484, 2022, doi: 10.1109/ACCESS.2022.3195044.

[207] J. Amin, M. Sharif, N. Gul, S. Kadry, and C. Chakraborty, "Quantum Machine Learning Architecture for COVID-19 Classification Based on Synthetic Data Generation Using Conditional Adversarial Neural Network," Cognit. Comput., pp. 1677–1688, 2021, doi: 10.1007/s12559-021-09926-6.

[208] B. K. Mishra and IGI Global, Technology road mapping for quantum computing and engineering, vol. i.

[209] A. Perdomo-Ortiz, M. Benedetti, J. Realpe-Gómez, and R. Biswas, "Opportunities and challenges for quantum-assisted machine learning in near-term quantum computers," Quantum Sci. Technol., vol. 3, no. 3, pp. 1–13, 2018, doi: 10.1088/2058-9565/aab859.

[210] G. Li, Z. Song, and X. Wang, "VSQL: Variational Shadow Quantum Learning for Classification," 2020, [Online]. Available: http://arxiv.org/abs/2012.08288

[211] J. Qi, C.-H. H. Yang, and P.-Y. Chen, "QTN-VQC: An End-to-End Learning framework for Quantum Neural Networks," no. 1, pp. 1–14, 2021, [Online]. Available: http://arxiv.org/abs/2110.03861

[212] I. Rojas, G. Joya, and A. Catala, "Advances in Computational Intelligence: 13th International work-conference on artificial neural networks, IWANN 2015 palma de mallorca, spain, june 10–12, 2015 proceedings, Part II ABC," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 9095, pp. 276–289, 2015, doi: 10.1007/978-3-319-19222-2.

[213] S. A. Alasadi and W. S. Bhaya, “Review of data preprocessing techniques in data mining,” J. Eng. Appl. Sci., vol. 12, no. 16, pp. 4102–4107, 2017, doi: 10.3923/jeasci.2017.4102.4107.

[214] P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, “New data preprocessing trends based on ensemble of multiple preprocessing techniques,” TrAC - Trends Anal. Chem., vol. 132, p. 116045, 2020, doi: 10.1016/j.trac.2020.116045.

[215] S. A. Billings and K. L. Lee, “Nonlinear Fisher discriminant analysis using a minimum squared error cost function and the orthogonal least squares algorithm,” Neural Networks, vol. 15, no. 2, pp. 263–270, 2002, doi: 10.1016/S0893-6080(01)00142-3.

[216] B. Computation, “Imperialist competitive algorithm for minimum bit error rate beamforming Arash Khabbazi * Esmaeil Atashpaz-Gargari and Caro Lucas,” Comput. Eng., vol. 1, 2009.

[217] D. D. Miller and E. W. Brown, “Artificial Intelligence in Medical Practice: The Question to the Answer?,” Am. J. Med., vol. 131, no. 2, pp. 129–133, 2018, doi: 10.1016/j.amjmed.2017.10.035.

[218] N. Mishra, A. Bisarya, S. Kumar, B. K. Behera, S. Mukhopadhyay, and P. K. Panigrahi, “Cancer Detection Using Quantum Neural Networks: A Demonstration on a Quantum Computer,” 2019, [Online]. Available: http://arxiv.org/abs/1911.00504

[219] Y. Bengio and Y. Lecun, “Scaling Learning Algorithms toward AI,” Large-Scale Kernel Mach., no. 1, pp. 1–41, 2019, doi: 10.7551/mitpress/7496.003.0016.

[220] H. Yang, L. Lu, and W. Zhou, “A novel optimization sizing model for hybrid solar-wind power generation system,” Sol. Energy, vol. 81, no. 1, pp. 76–84, 2007, doi: 10.1016/j.solener.2006.06.010.

[221] T. M. Khan and A. Robles-Kelly, “A Derivative-Free Method for Quantum Perceptron Training in Multi-layered Neural Networks,” Commun. Comput. Inf. Sci., vol. 1333, pp. 241–250, 2020, doi: 10.1007/978-3-030-63823-8-29.

[222] F. Arute et al., “Quantum supremacy using a programmable superconducting processor,” Nature, vol. 574, no. 7779, pp. 505–510, 2019, doi: 10.1038/s41586-019-1666-5.

[223] S. Boixo et al., “Characterizing quantum supremacy in near-term devices,” Nat. Phys., vol. 14, no. 6, pp. 595–600, 2018, doi: 10.1038/s41567-018-0124-x.

[224] V. Dunjko and H. J. Briegel, “Machine learning artificial intelligence in the quantum domain: A review of recent progress,” Reports Prog. Phys., vol. 81, no. 7, 2018, doi: 10.1088/1361-6633/aab406.

[225] N. Killoran, T. R. Bromley, J. M. Arrazola, M. Schuld, N. Quesada, and S. Lloyd, “Continuous-variable quantum neural networks,” Phys. Rev. Res., vol. 1, no. 3, pp. 1–22, 2019, doi: 10.1103/PhysRevResearch.1.033063.

[226] A. Garcia Olea et al., “Electronic health records features associated to coronary artery disease in patients with chest pain,” Eur. Heart J., vol. 42, no. Supplement-1, p. 2021, 2021, doi: 10.1093/eurheartj/ehab724.1151.

[227] P. A. Mundra and J. C. Rajapakse, “SVM-RFE with MRMR filter for gene selection,” IEEE Trans. Nanobioscience, vol. 9, no. 1, pp. 31–37, 2010, doi: 10.1109/TNB.2009.2035284.

[228] H. A. Le Thi, V. V. Nguyen, and S. Ouchani, “Gene selection for cancer classification using DCA,” Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 5139 LNAI, pp. 62–72, 2008, doi: 10.1007/978-3-540-88192-6-8.

[229] H. Sanz, C. Valim, E. Vegas, J. M. Oller, and F. Reverter, “SVM-RFE: Selection and visualization of the most relevant features through non-linear kernels,” BMC Bioinformatics, vol. 19, no. 1, pp. 1–18, 2018, doi: 10.1186/s12859-018-2451-4.

[230] M. Watabe, K. Shiba, C. C. Chen, M. Sogabe, K. Sakamoto, and T. Sogabe, “Quantum circuit learning with error backpropagation algorithm and experimental implementation,” Quantum Reports, vol. 3, no. 2, pp. 333–349, 2021, doi: 10.3390/quantum3020021.

[231] G. Chen, Q. Chen, S. Long, W. Zhu, Z. Yuan, and Y. Wu, “Quantum convolutional neural network for image classification,” Pattern Anal. Appl., 2022, doi: 10.1007/s10044-022-01113-z.

[232] D. Sierra-Sosa, J. D. Arcila-Moreno, B. Garcia-Zapirain, and A. Elmaghraby, “Diabetes Type 2: Poincaré data preprocessing for quantum machine learning,” Comput. Mater. Contin., vol. 67, no. 2, pp. 1849–1861, 2021, doi: 10.32604/cmc.2021.013196.

[233] Y. Wang, Y. Wang, C. Chen, R. Jiang, and W. Huang, “Development of variational quantum deep neural networks for image recognition,” Neurocomputing, vol. 501, pp. 566–582, 2022, doi: 10.1016/j.neucom.2022.06.010.

[234] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, “Evaluating analytic gradients on quantum hardware,” Phys. Rev. A, vol. 99, no. 3, pp. 1–8, 2019, doi: 10.1103/PhysRevA.99.032331.

[235] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," Phys. Rev. A, vol. 98, no. 3, pp. 1–3, 2018, doi: 10.1103/PhysRevA.98.032309.

[236] P. B. D. E. Sousa and R. V. Ramos, "The general decomposition of single-qubit, two-qubits and three-qubits quantum gates".

[237] PennyLane, dev team (2021) Ensemble classification with Forest and Qiskit devices¶, Ensemble classification with Forest and Qiskit devices-PennyLane documentation. Available at: https://pennylane.ai/qml/demos/ensemble-multi-qpu.html (Accessed: October 17, 2022).

[238] H. Yano, Y. Suzuki, R. Raymond, and N. Yamamoto, "Efficient Discrete Feature Encoding for Variational Quantum Classifier," Proc. - IEEE Int. Conf. Quantum Comput. Eng. QCE 2020, pp. 11–21, 2020, doi: 10.1109/QCE49297.2020.00012.

[239] I. Kerenidis, J. Landman, and A. Prakash, "Quantum Algorithms for Deep Convolutional Neural Networks," pp. 1–40, 2019, [Online]. Available: http://arxiv.org/abs/1911.01117

[240] Bruyère, O., Beaudart, C., Locquet, M., Buckinx, F., Petermans, J., Reginster, J. Y. (2016). Sarcopenia as a public health problem. European geriatric medicine, 7(3), 272-275.

[241] Rosenberg, I. H. (1997). Sarcopenia: origins and clinical relevance. The Journal of nutrition, 127(5), 990S-991S.

[242] Cruz-Jentoft, A. J., Bahat, G., Bauer, J., Boirie, Y., and Bruyère, O., Cederholm, T., ... and Zamboni, M. (2019). Sarcopenia: revised European consensus on definition and diagnosis. Age and ageing, 48(1), 16-31.

[243] Castillo-Olea, C., García-Zapirain Soto, B., Carballo Lozano, C., and Zuñiga, C. (2019). Automatic classification of sarcopenia level in older adults: a case study at Tijuana General Hospital. International Journal of Environmental Research and Public Health, 16(18), 3275.

[244] Cruz-Jentoft, AJ, Bahat, G., Bauer, J., Boirie, Y., Bruyère, O., Cederholm, T., ... and Zamboni, M. (2019). Sarcopenia: revised European consensus on definition and diagnosis. Age and ageing , 48 (1), 16-31.

[245] Biolo, G., Cederholm, T., and Muscaritoli, M. (2014). Muscle contractile and metabolic dysfunction is a common feature of sarcopenia of aging and chronic diseases: from sarcopenic obesity to cachexia. Clinical nutrition, 33(5), 737-748.

[246] Kim, Young-Sang, et al. "Longitudinal observation of muscle mass over 10 years according to serum calcium levels and calcium intake among Korean adults aged

50 and older: The Korean Genome and Epidemiology Study." Nutrients 12.9 (2020): 2856.

[247] Golubnitschaja, O., Liskova, A., Koklesova, L., Samec, M., Biringer, K., Büsselberg, D., ... and Kubatka, P. (2021). Caution,"normal" BMI: health risks associated with potentially masked individual underweight—EPMA Position Paper 2021. EPMA Journal, 12(3), 243-264.

[248] Tournadre, A., Vial, G., Capel, F., Soubrier, M., and Boirie, Y. (2019). Sarcopenia. Joint bone spine, 86(3), 309-314.

[249] Bian, A. L., Hu, H. Y., Rong, Y. D., Wang, J., Wang, J. X., and Zhou, X. Z. (2017). A study on relationship between elderly sarcopenia and inflammatory factors IL-6 and TNF. European journal of medical research, 22(1), 1-8.

[250] Cesari, M., Fielding, R. A., Pahor, M., Goodpaster, B., Hellerstein, M., Van Kan, G. A., ... and Vellas, B. (2012). Biomarkers of sarcopenia in clinical trials—recommendations from the International Working Group on Sarcopenia. Journal of cachexia, sarcopenia and muscle, 3(3), 181-190.

[251] Muscaritoli, M., Anker, S. D., Argilés, J., Aversa, Z., Bauer, J. M., Biolo, G. I. A. N. N. I., ... and Sieber, C. C. (2010). Consensus definition of sarcopenia, cachexia and pre-cachexia: joint document elaborated by Special Interest Groups (SIG)"cachexia-anorexia in chronic wasting diseases" and "nutrition in geriatrics". Clinical nutrition, 29(2), 154-159.

[252] Rockwood, K., and Mitnitski, A. (2006). Limits to deficit accumulation in elderly people. Mechanisms of ageing and development, 127(5), 494-496.

[253] Morley, J. E., von Haehling, S., Anker, S. D., and Vellas, B. (2014). From sarcopenia to frailty: a road less traveled. Journal of cachexia, sarcopenia and muscle, 5(1), 5-8.

[254] Sui, S. X., Holloway-Kew, K. L., Hyde, N. K., Williams, L. J., Tembo, M. C., Leach, S., and Pasco, J. A. (2021). Prevalence of sarcopenia employing population-specific cut-points: Cross-sectional data from the geelong osteoporosis study, Australia. Journal of Clinical Medicine, 10(2), 343.

[255] Ullah, Ubaid, and Begonya Garcia-Zapirain. "Quantum Machine Learning Revolution in Healthcare: A Systematic Review of Emerging Perspectives and Applications." IEEE Access (2024).

[256] Guglielmi, G., Ponti, F., Agostini, M., Amadori, M., Battista, G., and Bazzocchi, A. (2016). The role of DXA in sarcopenia. Aging clinical and experimental research, 28(6), 1047-1060.

[257] Maheshwari, Danyal, et al. "Quantum Machine Learning Applied to Electronic Healthcare Records for Ischemic Heart Disease Classification."

[258] Coorevits, P., Sundgren, M., Klein, G. O., Bahr, A., Claerhout, B., Daniel, C., ... and Kalra, D. (2013). Electronic health records: new opportunities for clinical research. Journal of internal medicine, 274(6), 547-560.

[259] Ehrenstein, V., Petersen, I., Smeeth, L., Jick, S. S., Benchimol, E. I., Ludvigsson, J. F., and Sørensen, H. T. (2016). Helping everyone do better: a call for validation studies of routinely recorded health data. Clinical epidemiology, 8, 49.

[260] Ehrenstein, V., Nielsen, H., Pedersen, A. B., Johnsen, S. P., and Pedersen, L. (2017). Clinical epidemiology in the era of big data: new opportunities, familiar challenges. Clinical epidemiology, 9, 245.

[261] Buffoni, Lorenzo, and Filippo Caruso. "New trends in quantum machine learning (a)." Europhysics Letters 132.6 (2021): 60004.

[262] Ziwei, M. E. N. G. "Review of Quantum Computing." 2020 13th International Conference on Intelligent Computation Technology and Automation (ICICTA). IEEE, 2020.

[263] Biamonte, J., Wittek, P., Pancotti, N. et al. Quantum machine learning. Nature 549, 195–202 (2017). https://doi.org/10.1038/nature23474

[264] Abohashima, Zainab, et al. "Classification with quantum machine learning: A survey." arXiv preprint arXiv:2006.12270 (2020).

[265] Ullah, Ubaid, et al. "A Fully Connected Quantum Convolutional Neural Network for Classifying Ischemic Cardiopathy." IEEE Access 10 (2022): 134592-134605.

[266] Malviya, Rishabha, and Sonali Sundram. "Exploring potential of quantum computing in creating smart healthcare." Open Biol J 9.1 (2022): 56-57.

[267] Davenport, Thomas, and Ravi Kalakota. "The potential for artificial intelligence in healthcare." Future healthcare journal 6.2 (2019): 94.

[268] Davison, Keron, and Malcolm P. Forbes. "Pre-hospital medicine: A glimpse of the future." Australasian Journal of Paramedicine 12 (2015): 1-4.

[269] Malviya, Rishabha, and Sonali Sundram. "Exploring potential of quantum computing in creating smart healthcare." Open Biol J 9.1 (2022): 56-57.

[270] Castillo-Olea, C., García-Zapirain Soto, B., Carballo Lozano, C., and Zuñiga, C. (2019). Automatic classification of sarcopenia level in older adults: a case study at Tijuana General Hospital. International Journal of Environmental Research and Public Health, 16(18), 3275.

[271] Castro, E. M. (2019). Bioestadística aplicada en investigación clínica: conceptos básicos. Revista médica clínica las Condes, 30(1), 50-65.

[272] Guyon, Isabelle, et al. "Gene selection for cancer classification using support vector machines." Machine learning 46 (2002): 389-422.

[273] Abohashima, Zainab, et al. "Classification with quantum machine learning: A survey." arXiv preprint arXiv:2006.12270 (2020).

[274] Araujo, Israel F., et al. "A divide-and-conquer algorithm for quantum state preparation." Scientific reports 11.1 (2021): 1-12.

[275] Durr, Christoph, and Peter Hoyer. "A quantum algorithm for finding the minimum." arXiv preprint quant-ph/9607014 (1996).

[276] Miyamoto, Kohei, Masakazu Iwamura, and Koichi Kise. "A Quantum Algorithm for Finding k-Minima." arXiv preprint arXiv:1907.03315 (2019).

[277] Boyer, Michel, et al. "Tight bounds on quantum searching." Fortschritte der Physik: Progress of Physics 46.4-5 (1998): 493-505.

[278] Buhrman, Harry, et al. "Quantum fingerprinting." Physical Review Letters 87.16 (2001): 167902.

[279] Basheer, Afrad, A. Afham, and Sandeep K. Goyal. "Quantum k-nearest neighbors algorithm." arXiv preprint arXiv:2003.09187 (2020).

[280] Havlíček, V., Córcoles, A.D., Temme, K. et al. Supervised learning with quantum-enhanced feature spaces. Nature 567, 209–212 (2019). https://doi.org/10.1038/s41586-019-0980-2

[281] Iten, Raban, et al. "Quantum circuits for isometries." Physical Review A 93.3 (2016): 032318.

[282] Schuld, Maria, et al. "Learning with quantum models." Supervised Learning with Quantum Computers (2018): 247-272.

[283] Powell, Michael JD. "A view of algorithms for optimization without derivatives." Mathematics Today-Bulletin of the Institute of Mathematics and its Applications 43, no. 5 (2007): 170-174.

[284] Araujo, Israel F., et al. "A divide-and-conquer algorithm for quantum state preparation." Scientific reports 11.1 (2021): 1-12. Discussion

[285] Kok, D., S. Caron, and A. Acun. "Building a quantum kNN classifier with Qiskit: theoretical gains put to practice." (2021).

[286] Dang, Y., Jiang, N., Hu, H. et al. Image classification based on quantum K-Nearest-Neighbor algorithm. Quantum Inf Process 17, 239 (2018). https://doi.org/10.1007/s11128-018-2004-9

[287] Wiebe, Nathan, Ashish Kapoor, and Krysta Svore. "Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning." arXiv preprint arXiv:1401.2142 (2014).

[288] Sierra-Sosa, Daniel, et al. "Diabetes Type 2: Poincaré data preprocessing for quantum machine learning." Comput., Mater. Continua 67.2 (2021): 1849-1861.

[289] Gupta, Himanshu, et al. "Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction." Complex and Intelligent Systems 8.4 (2022): 3073-3087.

[290] Maheshwari, Danyal, Daniel Sierra-Sosa, and Begonya Garcia-Zapirain. "Variational quantum classifier for binary classification: Real vs synthetic dataset." IEEE Access 10 (2021): 3705-3715.