UIR Ecole Supérieure d'Informatique et du Numérique
COLLEGE OF ENGINEERING & ARCHITECTURE

JAKJOUD LABS Tech for humanity &

Capstone Project

To obtain

Engineering Diploma

Degree specialty of

Computer Engineering – Big Data & Artificial Intelligence

Evaluating Classical and Quantum Hybrid Approaches for Breast Cancer Diagnosis: A Performance Comparison

---

Realized by

Ikram EL-HAJRI

Supervised by

Academic supervisor:	Prof. SBIHI Nada, Research Professor - UIR
Professional supervisor:	Dr. JAKJOUD Abdeslam, CEO – Jakjoud Labs
	Mme. BELLA Imane, Product Owner – Synergeon

Academic year 2023 - 2024

Dedication

To my beloved parents, Bouchta El-hajri & Rabiaa Karim,

Words fail to capture the depth of emotion behind this dedication. In every moment, through joys and sorrows alike, you have been unwavering pillars of support by my side. Your love knows no bounds, and I am forever indebted to the sacrifices you have made for my well-being. May the divine watch over you, granting you a life filled with health, happiness, and longevity.

To the cherished memory of my late grandfather,

Your wisdom, kindness, and strength have left an indelible mark on my heart. Though you are no longer with us, your spirit continues to guide and inspire me. I am grateful for the legacy you left behind and the lessons you imparted. May your soul rest in eternal peace, and may your memory always be a blessing.

To my beloved siblings, Imran & Hiba,

For the countless shared moments, the unbreakable bond we share, and the harmony that binds us together, I am eternally grateful. Your love and unwavering support have been the guiding light that has fueled my resilience and determination. May the blessings of the Almighty shower upon you, leading you to the best that life has to offer.

To the esteemed International University of Rabat and its esteemed faculty,

I extend my heartfelt gratitude to the institution and its resolute professionals who have provided an enriching environment for growth and learning. Your commitment to excellence, tireless dedication, and guidance have nurtured our talents and honed our skills, equipping us with the knowledge and principles to thrive in our endeavors.

Ikram El-hajri

Acknowledgements

I would like to express my gratitude to all the people who contributed to the completion of this project and played an essential role in its success.

First and foremost, I would like to warmly thank the teaching staff of the International University of Rabat, as well as the professors who imparted in-depth knowledge in our field of study. Their theoretical teachings laid the solid foundation necessary for my success during this internship.

I would like to extend my sincerest thanks to Mrs. **Nada SBIHI**, my supervisor, for her constant availability, attentive listening, and invaluable advice. Her guidance throughout this internship has been of immeasurable value, and I am extremely grateful to her for her support.

I also want to express my gratitude to Mr. **Abdeslam JAKJOUD**. For the opportunity he provided me within *JakJoud Labs* and for showing keen interest in my training, and for dedicating much of his time, attention, and energy to the tasks I had to accomplish, all within a pleasant atmosphere of camaraderie and respect.

I am extremely grateful to Mr. **Jamal Benhamou** and Mrs. **Imane Bella** for the trust they placed in me by selecting me for this internship. I am extremely grateful to have had the chance to work alongside them and to benefit from their support throughout this enriching experience. Beyond the gratitude I have for this unique opportunity, I also want to acknowledge their generosity and commitment to the professional development of young talents.

My deepest thanks go to Mr. **Karim AMOR**, the founder and CEO of the company *Synergeon*, for his warm welcome, advice, and the diversity of approaches he suggested during the project's realization.

Finally, to all those who contributed directly or indirectly to the completion of this work, please accept my warmest thanks.

Abstract

This study delves into a comparative analysis of classical machine learning models and a hybrid quantum-classical model, particularly the Variational Quantum Classifier (VQC), in the domain of breast cancer detection. The primary objective is to explore the potential enhancements that quantum integration could offer over traditional machine learning techniques. Utilizing the Breast Cancer Wisconsin (Diagnostic) Dataset, I meticulously evaluated classical models, spanning Support Vector Machine (SVM), Logistic Regression, Gradient Boosting, Random Forest, Decision Tree, KNeighbors, and XGBoost, scrutinizing their performance comprehensively.

Subsequently, I implemented the VQC model and juxtaposed its performance against these classical counterparts. Despite the empirical superiority of classical machine learning models in this specific application, it's crucial to note that quantum machine learning holds promise in numerous other domains. Indeed, quantum computing's unique capabilities may lead to breakthroughs in fields beyond medical diagnostics, emphasizing the multifaceted potential of quantum machine learning across various disciplines.

Key Words: *Breast Cancer Detection, Machine Learning (ML), quantum computing (QC), quantum algorithms, Quantum Machine Learning (QML), quantum classification, Variational Quantum Circuit (VQC), Variational Quantum Classifier (VQC), quantum encoding.*

Résumé

Cette étude explore une analyse comparative des modèles classiques d'apprentissage automatique et d'un modèle hybride quantique-classique, en particulier le Classificateur Quantique Variational (VQC), dans le domaine de la détection du cancer du sein. L'objectif principal est d'explorer les améliorations potentielles que l'intégration quantique pourrait apporter par rapport aux techniques traditionnelles d'apprentissage automatique. En utilisant l'ensemble de données du Wisconsin Breast Cancer (Diagnostic), nous avons évalué méticuleusement des modèles classiques, tels que la Machine à Vecteurs de Support (SVM), la Régression Logistique, le Boosting de Gradient, la Forêt Aléatoire, l'Arbre de Décision, KNN, et XGBoost, en scrutant leur performance de manière exhaustive.

Par la suite, nous avons implémenté le modèle VQC et avons comparé ses performances à celles de ces contreparties classiques. Malgré la supériorité empirique des modèles classiques d'apprentissage automatique dans cette application spécifique, il est crucial de noter que l'apprentissage automatique quantique présente des promesses dans de nombreux autres domaines. En effet, les capacités uniques de l'informatique quantique pourraient conduire à des percées dans des domaines autres que le diagnostic médical, mettant en avant le potentiel multifacette de l'apprentissage automatique quantique à travers différentes disciplines.

Mots-clés: *Détection du Cancer du Sein, Apprentissage Automatique (ML), Informatique Quantique (QC), Algorithmes Quantiques, Apprentissage Automatique Quantique (QML), Classification Quantique, Circuit Quantique Variational (VQC), Classificateur Quantique Variational (VQC), Encodage Quantique.*

Glossary

Term	*Definition*
Artificial Intelligence	A technique enables computers to mimic human behavior
Machine Learning (ML)	the subset of AI technique; pattern identification and analysis; machines can improve with experience from provided data sets
Deep Learning (DL)	the subset of ML technique; composed of multi-layer neural networks
Adaptive Spatial Feature Fusion	A technique in image processing that combines spatial features from different sources adaptively to improve the performance of a classification model.
Bilateral Filtering	A non-linear, edge-preserving, and noise-reducing smoothing filter used in image processing.
BIRADS Features	Breast Imaging-Reporting and Data System (BI-RADS) features are standardized descriptors used to categorize breast lesions based on their appearance in imaging studies.
Block-Based Region Segmentation Algorithm	A method in image segmentation that divides the image into smaller blocks and processes each block individually for more accurate region segmentation.
BUViTNet	A vision transformer (ViT) network specifically designed for breast ultrasound detection, leveraging pre-trained models from large datasets like ImageNet and cancer cells.
CAD System (Computer-Aided Diagnosis)	A computer system that assists radiologists in interpreting medical images, often using advanced image processing and machine learning techniques.
Coarse-to-Fine Feature Fusion	A hierarchical approach in which features are extracted in a coarse manner initially and then refined progressively to improve the accuracy of image segmentation or classification.
Contrast Limited Adaptive Histogram Equalization (CLAHE)	An image enhancement technique that improves contrast in images by adjusting the histogram within localized regions.
Convolutional Neural Network (CNN)	A type of deep neural network particularly effective for analyzing visual data, commonly used in image recognition and classification tasks.
Deep Representation Scaling (DRS) Layers	Layers in a neural network designed to reduce the number of trainable parameters while maintaining performance, often used in deep learning models to improve efficiency.
Encoder-Decoder Architecture	A neural network design commonly used in image segmentation tasks, where the encoder extracts feature from the input image and the decoder reconstructs the segmented image.
Ensemble Techniques	Methods that combine predictions from multiple models to improve overall accuracy and robustness.

Fuzzy Enhancement	A technique in image processing that enhances image features by applying fuzzy logic principles, making the image details more distinct.
GAN (Generative Adversarial Network)	A class of machine learning frameworks where two neural networks, a generator and a discriminator, are trained together to produce data that mimics a real dataset.
Image Fusion	Combining multiple images or their features to produce a single image that contains more information than any of the input images.
MLP Head (Multilayer Perceptron Head)	The final layers in a transformer model that are fully connected and used for classification tasks.
Noise Filter Network (NF-Net)	A neural network designed to handle and reduce noise in labeled training data, improving the accuracy of the trained model.
Optimal Deep Learning Feature Model	A model that selects and combines the most relevant features extracted by deep learning techniques to enhance performance in specific tasks.
RGB Channels	The red, green, and blue color channels in an image, which can be processed separately or together for various image processing tasks.
Semi-Supervised GAN Model	A type of GAN that uses both labeled and unlabeled data during training to improve performance, especially when labeled data is scarce.
Semi-Supervised Deep Learning (SSDL) Framework	A deep learning approach that uses a combination of labeled and unlabeled data to train models, improving performance with limited labeled data.
Super-Pixel Images	Images where pixels are grouped into perceptually meaningful atomic regions, which are then used for more efficient image segmentation.
Supervised Block-Based Region Segmentation Algorithm	A supervised learning technique that divides an image into blocks and segments regions based on labeled training data.
Transfer Learning	A machine learning technique where a model developed for one task is reused as the starting point for a model on a second task.
Vision Transformers (ViTs)	A type of model architecture that leverages self-attention mechanisms to process images, often outperforming traditional CNNs in certain tasks.

Abbreviations

Abbreviation	Full Form
UIR	Université Internationale de Rabat
AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
PO	Product Owner
VQC	Variational Quantum Classifier
BUS	Breast Ultrasound
WHO	World Health Organization
NF-Net	Noise Filter Network
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
BIRADS	Breast Imaging-Reporting and Data System
SSDL	Semi-Supervised Deep Learning
GAN	Generative Adversarial Network
RGB	Red, Green, Blue
DRS	Deep Representation Scaling
ViTs	Vision Transformers
CNNs	Convolutional Neural Networks
US	Ultrasound
MLP	Multilayer Perceptron
BUViTNet	Breast Ultrasound Vision Transformers Network
ViT	Vision Transformers
SPSA	Simultaneous Perturbation Stochastic Approximation
WIP	Work In Progress
ASR	age-standardized rates
AUC	Area Under the Curve
KNN	K-Nearest Neighbors
QML	Quantum Machine Learning
SVM	Support Vector Machine
QAOA	Quantum Approximate Optimization Algorithm

List of Figures

List of Tables

Table of Contents

Introduction

Breast cancer remains one of the most prevalent and deadly forms of cancer worldwide, necessitating advancements in diagnostic technologies to improve early detection and treatment outcomes. Traditional machine learning models have shown significant promise in enhancing diagnostic accuracy by analyzing complex datasets to identify patterns indicative of cancer. These classical models, including Support Vector Machines (SVM), Logistic Regression, Gradient Boosting, Random Forest, Decision Tree, KNeighbors, and XGBoost, have been widely adopted due to their robustness, ease of implementation, and effectiveness.

In recent years, the advent of quantum computing has opened new avenues for computational innovation. Quantum computing, leveraging the principles of quantum mechanics, promises to revolutionize various fields by solving problems more efficiently than classical computers. One such promising application is in the realm of machine learning, where quantum-enhanced models, such as the Variational Quantum Classifier (VQC), aim to surpass the capabilities of classical algorithms by processing vast amounts of data and identifying intricate patterns more effectively.

This report aims to conduct a comparative analysis between classical machine learning models and a hybrid quantum-classical model in the context of breast cancer detection. The study focuses on evaluating the performance of these models using the Breast Cancer Wisconsin (Diagnostic) Dataset, examining metrics such as accuracy, precision, recall, and F1-score to determine the efficacy of each approach.

The primary objective is to investigate whether the integration of quantum computing can enhance the performance of classical machine learning models in medical diagnostics. While quantum computing holds theoretical advantages, such as improved computational efficiency and superior handling of high-dimensional data, it remains crucial to empirically validate these benefits. This research seeks to contribute to the ongoing discourse on the potential and current limitations of quantum computing in practical applications, particularly in the critical field of breast cancer detection.

By providing a detailed comparison, this report aims to highlight the strengths and weaknesses of both classical and quantum-enhanced models, offering insights into their current capabilities and future potential.

Chapter 1 : General Project Framework

Introduction:

This chapter provides an insightful overview of the hosting organizations involved in the project, shedding light on their missions, visions, strategic issues, and organizational structures. The chapter begins with an introduction to Synergeon, a startup founded by Mr. Karim Amor, specializing in disruptive solutions in robotics, software, and the Internet of Things, with a keen focus on artificial intelligence and blockchain technologies. Following this, the contracting laboratory, Jakjoud Labs, is introduced, along with its parent company, Terradoxa SAS. The chapter delves into the core values, missions, and technological segments addressed by both organizations, providing a comprehensive understanding of their strategic objectives and areas of expertise.

I. Presentation of the hosting organization

1. Presentation of Synergeon

Synergeon is a startup founded in 2022 by Mr. **Karim Amor** specialized in the design and development of disruptive solutions in the fields of robotics, software, and the Internet of Things, with a particular penchant for the technological segments of artificial intelligence and blockchain (NFT).

1 Logo Synergeon

2. Missions, Visions and Values

Its mission is mainly focused on the development and realization of innovative projects referred to as "**Moonshot projects**". These are exploratory projects initially intended to have a massive impact on society, both economically and socially, and capable of addressing a meta-problem through a radical creative approach, using cutting-edge technologies.

3. Strategic Issues:

- Recruitment and retention of talents.
- Investment in R&D.
- Adaptation to technological developments.
- Strategic partnerships.

II. The contracting laboratory Jakjoud Labs:

1. Presentation of the Company Jakjoud Labs:

2 Logo Jakjoud Labs

Technological company founded in 2022 by four individuals:

- **Abdeslam Jakjoud**: Doctor in artificial intelligence
- **Widad Jakjoud**: Doctor in computer science
- **Hicham Jakjoud**: Doctor in physics
- **Fatima Zohra Jakjoud**: Doctor in IoT and embedded systems

3 Logo Terradoxa

Based in Marrakech, specialized IoT and AI subsidiary of the company **Terradoxa**.

2. Presentation of Terradoxa SAS:

Terradoxa SAS is a technological company active since 2019, based in Paris, specialized in digital identity solutions, software engineering consulting services, and cloud solutions.

3. Visions, values, and missions:

a. Visions:

- Development of customized solutions and integration of software packages for companies of all sizes.
- Scientific research in the fields of quantum physics, materials physics, computer science, artificial intelligence, and embedded systems.
- Implementation of customized solutions combining IoT, computer science, and artificial intelligence for industrial companies.
- Consulting services (in systems engineering, software engineering, cloud, logistics, etc.) offered to public and private sectors.

b. Core Values:

The environment at Jakjoud Labs is based on a collective spirit: a spirit of family and team, united around strong values.

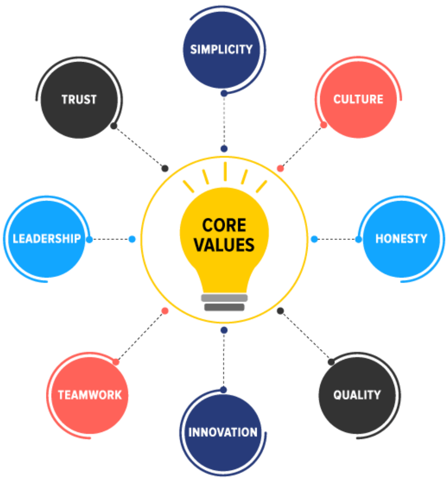


Figure 4 Core Values

c. Missions:

Provide technological solutions adapted to the needs of Moroccan companies through the combination of research and engineering in computer science, mathematics, electronics, and physics.

"We aim to create/integrate a self-sufficient Moroccan technological ecosystem first and then African. With the objective of being able to design sovereign and non-coercive technological solutions."

4. Addressed Technological Segments:

- Engineering of computer solutions.
- Cloud solution architectures.
- Engineering of embedded systems.
- Artificial intelligence and data sciences.
- Web3.0 solutions and blockchain.

- **Quantum computing**.

5. Special Features:
   - **Stakeholder in the Moroccan quantum computer project (first in Africa).**
   - Patent pending in the blockchain domain.
   - They offer workshops, particularly for high schools, for:
     - ➔ Video game creation.
     - ➔ Robotics.

6. Strategic Issues:
   - Recruitment and retention of talents.
   - Investment in R&D.
   - Adaptation to technological developments.
   - Strategic partnerships.

7. Organizational Chart:

During my internship, I was at the level of Mr. *Jihane Abdeslam JAKJOUD*'s department (applied mathematics and AI).

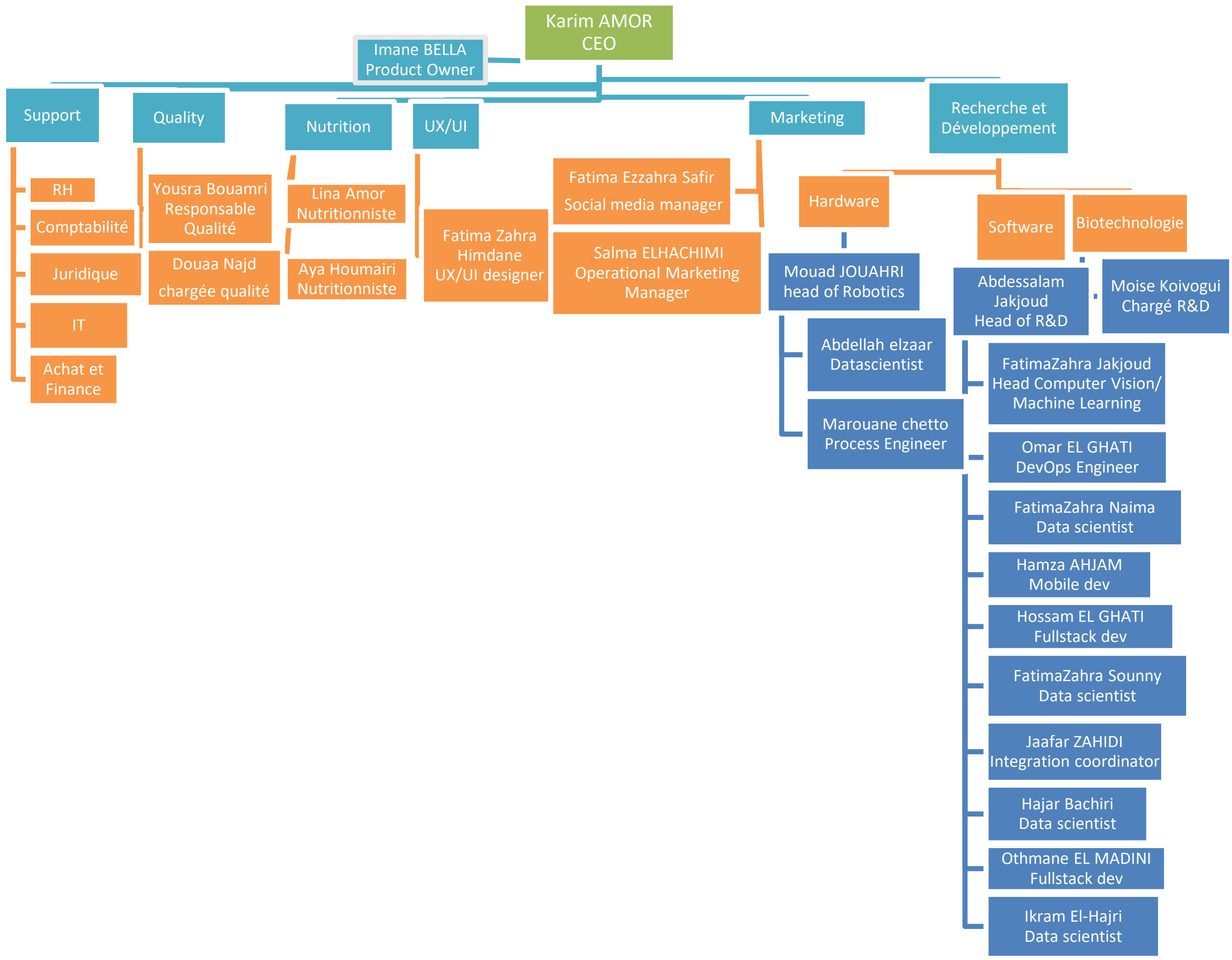


Figure 5 Organizational Chart

III. Presentation of the project

1. Project Presentation:

This project focuses on enhancing breast cancer detection models by leveraging quantum nodes. The objective is to explore the integration of quantum computing techniques to improve the accuracy and efficiency of breast cancer classification.

The project involves a comprehensive literature review to investigate existing methodologies and approaches in the field of quantum computing for healthcare applications, specifically for breast cancer detection. Through this review, insights are gained into how quantum nodes can be incorporated into machine learning pipelines to enhance classification performance.

Two distinct approaches are pursued in this project. Firstly, a classical approach utilizing traditional machine learning algorithms is implemented to establish a baseline for comparison. Secondly, a hybrid approach employing Variational Quantum Circuits (VQC) is explored to leverage the capabilities of quantum computing for breast cancer classification.

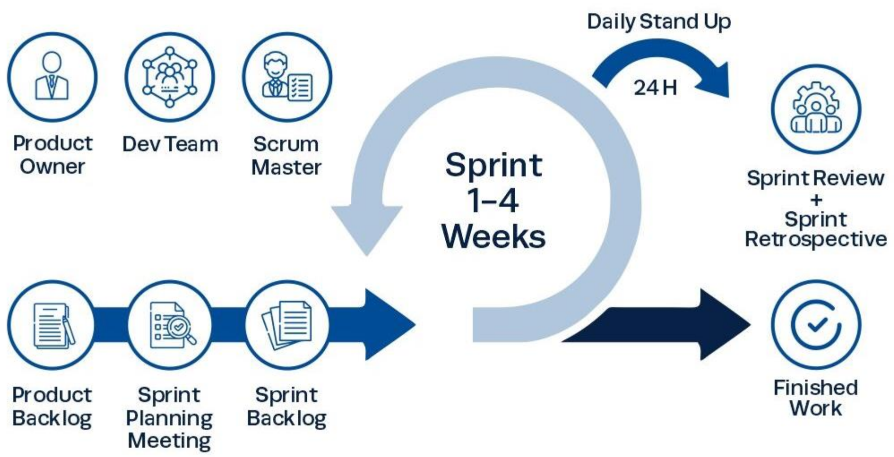
Overall, the project seeks to contribute to the advancement of breast cancer detection techniques by harnessing the potential of quantum computing, with the ultimate goal of improving diagnostic accuracy and patient outcomes.

2. Project Management

Developing IT projects requires good organization both technically and functionally to manage the increasing complexity of information systems and meet client needs. It's important to adopt a project management method that enhances system understanding, ensures a good level of quality, and guarantees effective maintenance by facilitating communication and common understanding among stakeholders. In this regard, I chose an agile method called "**Kanban**," as it best meets the needs of my project while adhering to the rules expressed at the project's outset.

a. Agile Method:

The project methodology aims to minimize risks and validate the development of an agile process. Agile methods focus on designing software by maximizing customer involvement and actively responding to their requests. They are based on three core values:



6 Agile Methodology

- **Team:** Focus on individuals and their interactions rather than processes and tools.
- **Collaboration**: Involving the client in development and direct communication with them.

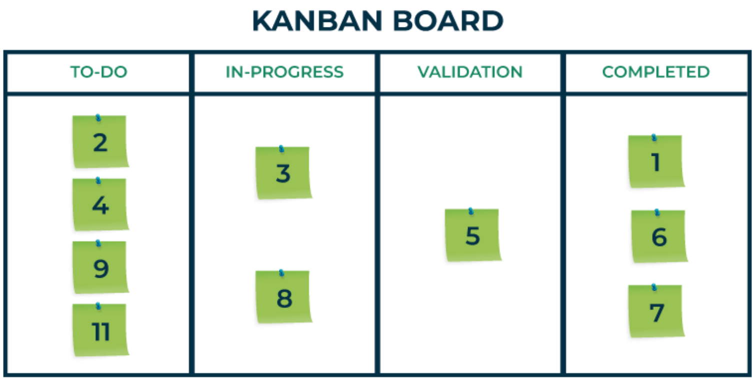
- **Embracing Change**: Initial planning and software structure should be flexible and adaptable. It's crucial to react to change rather than rigidly adhering to a plan.
- **Sprints**: During each sprint, various meetings are scheduled:
- **Daily Meeting**: A quick 15-minute daily synchronization meeting to review task progress.
- **Delivery**: After each sprint, completed tasks are delivered in the form of a report.
- **Sprint Review**: The team presents their achievements to the client's representative, the "PO," for sprint validation. Detected issues are either addressed in the next sprint or grouped into a separate one.
- **Sprint Retrospective**: After each sprint, the team holds a meeting to discuss the previous sprint and propose improvements for the next one.
- **Product Owner or "PO":** Represents the client and describes the tasks to be done. In our case, it represents "Synergeon."

b. Kanban:

Kanban is a visual project management approach focusing on transparency, collaboration, and workflow management. It's widely used in project management, including in Data Science projects.

Here's how Kanban can be applied to such projects:

Kanban Board:



7 **Kanban Board**

Visualizes the project's workflow, typically divided into columns representing different stages such as "To Do," "In Progress," and "Done." For my project, columns can be customized for specific stages like "review literature," "Preprocessing," "Model Training," " Benchmark Against Classical Models," "," etc.

Kanban Cards: Each task or work item is represented by a Kanban card containing information like description, dependencies, required resources, and deadlines. In the context of translation and paraphrasing, Kanban cards can represent specific tasks like literature review, study the fundamentals of quantum computing and quantum nodes, Gather datasets for training models, result evaluation, etc.

Workflow: Kanban cards move from left to right as tasks progress through the workflow. For instance, a card starts in the "To Do" column and moves to the "In Progress" column when someone starts working on it. Upon completion, it moves to the "Done" column. This allows the entire team to see the project's progress and identify potential bottlenecks or issues.

Work In Progress (WIP) Limits: A key practice of Kanban is limiting the number of tasks in progress at each workflow stage. This helps maintain focus and high efficiency by avoiding overloading team members.

Continuous Improvement: Kanban encourages continuous process improvement. By regularly analyzing workflow, cycle times, and encountered issues, the team can identify optimization and improvement opportunities. This could include identifying bottlenecks, finding ways to speed up the translation and paraphrasing process, or adjusting priorities based on outcomes.

Conclusion

In conclusion, this chapter serves as a foundational exploration of the hosting organizations, Synergeon and Jakjoud Labs, and their respective parent company, Terradoxa SAS. By elucidating their missions, visions, values, and strategic issues, the chapter sets the stage for the subsequent phases of the project. The insights garnered from this exploration will inform the project's direction, ensuring alignment with the organizations' objectives and fostering collaboration towards the successful realization of the project goals. Additionally, the chapter highlights the importance of effective project management methodologies such as Agile and Kanban, along with tools like Gantt charts, in ensuring efficient coordination and execution of project activities.

Chapter 2: Project Background

Introduction

This chapter aims to provide the motivation behind enhancing breast cancer detection models, focusing on the significance of breast cancer as a global health issue and its impact in Morocco. The chapter will present key statistics and highlight the urgency for improved detection methods.

I. Motivation

Breast cancer is the most prevalent cancer among women worldwide, with an estimated **2.3 million new cases** diagnosed each year according to the World Health Organization (WHO). It is also the leading cause of cancer death in women, claiming over **685,000 lives** annually.

In Morocco, the situation is concerning as well. While breast cancer rates are generally lower than in developed countries, they are on the rise. Studies suggest an age-standardized incidence rate of **around 36.4 per 100,000 women**, with a significant increase observed between 2004 and 2008. This rise is attributed to factors like changes in reproductive behavior and lifestyle choices.

Early detection is crucial for improving breast cancer prognosis. However, current detection methods face limitations. Mammography, the gold standard screening tool, can miss some cancers and lead to unnecessary biopsies.

This capstone project is driven by the urgent need to develop more accurate and reliable breast cancer detection models. By integrating quantum computing with classical machine learning techniques, we aim to create a powerful tool for early and effective diagnosis.

II. Understanding Breast Cancer

1. Overview of Breast Cancer

Breast cancer is a disease in which abnormal breast cells grow out of control and form tumors. If left unchecked, the tumors can spread throughout the body and become fatal.

Breast cancer cells begin inside the milk ducts and/or the milk-producing lobules of the breast. The earliest form (in situ) is not life-threatening and can be detected in the preliminary stages. Cancer cells can spread into nearby breast tissue (invasion). This creates tumors that cause lumps or thickening.

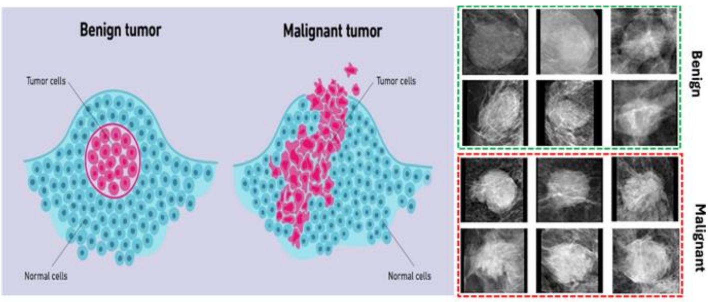


Figure 8 Visualization of breast cancer.

Invasive cancers can spread to nearby lymph nodes or other organs (metastasize). Metastasis can be life-threatening and fatal.

Treatment is based on the person, the type of cancer and its spread. Treatment combines surgery, radiation therapy and medications.

2. Causes of Breast Cancer

The exact cause of most breast cancers is unknown. Researchers believe it arises from a complex interplay between genetics and environmental factors. Here are some of the established risk factors that can increase a person's chances of developing breast cancer:

Age: Breast cancer risk increases as women age. Most breast cancers are diagnosed in women over 50.

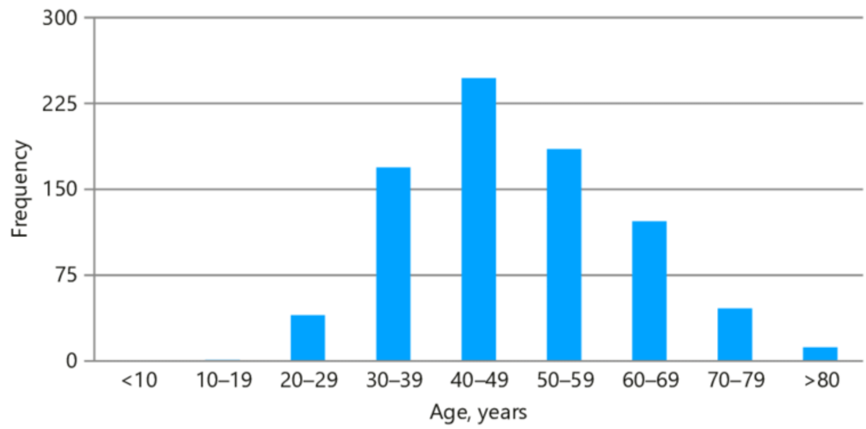


Figure 9 Age Distribution of Breast Cancer

Family history: Having a close relative (mother, sister, daughter) with breast cancer increases your risk. The risk is even higher if the relative was diagnosed at a young age or has multiple breast cancer diagnoses in the family.

Risk Factor	Systematic Review: Study Type or Number of Studies (Reference)	Systematic Review: Breast Cancer Risk Ratio (95% CI)	BCSC Breast Cancer Risk Ratio (95% CI)*
First-degree relatives with breast cancer			
0	Meta-analysis (26)	Reference group	Reference group
1	Meta-analysis (26)	2.14 (1.92–2.38)	
2	Meta-analysis (26)	3.84 (2.37–6.22)	None vs. any, 1.86 (1.69–2.06)
≥3	Meta-analysis (26)	12.05 (1.70–85.16)	
Age at diagnosis of first-degree relatives with breast cancer			
None	Meta-analysis (26)	Reference group	Reference group
<40 y	Meta-analysis (26)	3.0 (1.8–4.9)	None vs. <50 y, 2.17 (1.86–2.53)
40–49 y	Meta-analysis (26)	2.0 (1.5–2.8)	
50–59 y	Meta-analysis (26)	2.3 (1.7–3.2)	None vs. ≥50 y, 1.68 (1.44–1.96)
≥60 y	Meta-analysis (26)	1.7 (1.3–2.1)	
Second-degree relatives with breast cancer			
0	Meta-analysis (27)	Reference group	NA
≥1	Meta-analysis (27)	1.7 (1.4–2.0)	NA
Breast density (BI-RADS category)†			
1	1 (75)	0.46 (0.37–0.58)	0.41 (0.31–0.55)
2	1 (75)	Reference group	Reference group
3	1 (75)	1.62 (1.51–1.75)	1.75 (1.59–1.93)
4	1 (75)	2.04 (1.84–2.26)	2.33 (2.04–2.66)
Prior breast procedure‡			
None	1 (76)	Reference group	Reference group
Any	1 (76)	1.87 (1.64–2.13)	1.51 (1.36–1.67)

BCSC = Breast Cancer Surveillance Consortium; BI-RADS = Breast Imaging Reporting and Data System; NA = not available.
* Model included age, race, body mass index, family history of breast cancer, and site. Numbers of women included in estimates varied by risk factor because of missing data.
† 1 = almost entirely fat; 2 = fibroglandular densities; 3 = heterogeneously dense; 4 = extremely dense.
‡ Surgical and needle biopsies.

Table 1 Breast Cancer Risk Associated with Family History, Breast Density, and Breast Procedures

- **Genes:** Inherited gene mutations, particularly in the BRCA1 and BRCA2 genes, significantly elevate breast cancer risk. These genes play a role in DNA repair, and mutations can hinder this process, increasing the risk of errors that can lead to cancer.

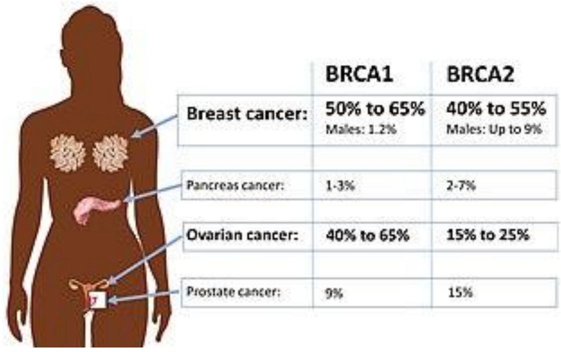


Figure 10 BRCA Genes and Breast Cancer Risk

- **Dense breast tissue:** Women with denser breast tissue have a higher risk of breast cancer. Dense breast tissue appears white on mammograms, which can make it harder to detect tumors.

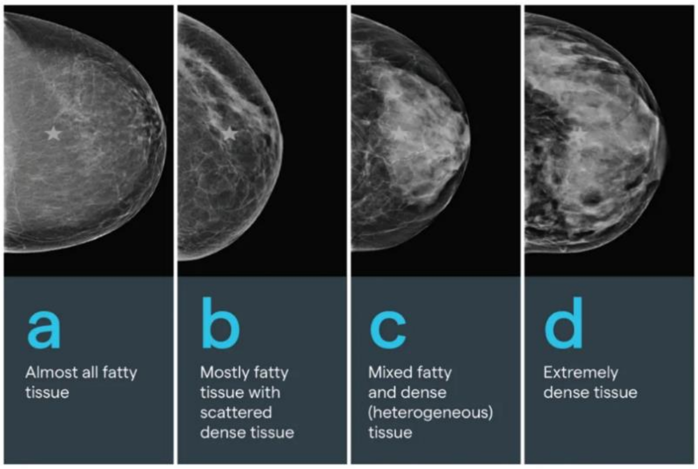


Figure 11 Dense Breast Tissue

- **Hormonal factors:** Exposure to estrogen over a longer period increases breast cancer risk. This includes factors like starting menstruation at a young age, having a late menopause, taking hormone replacement therapy (HRT) for an extended period, and not having children or having your first child after age 30.
- **Lifestyle factors:** Obesity, lack of physical activity, and excessive alcohol consumption can contribute to an increased risk of breast cancer.

3. Types of Breast Cancer

Breast cancer can be broadly categorized into two main types: invasive and non-invasive.

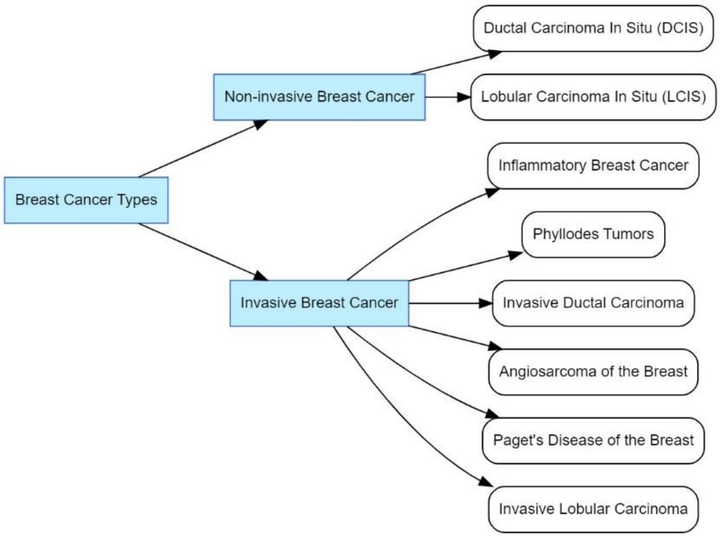


Figure 12 Types of Breast Cancer

Non-invasive breast cancer:

- **Ductal carcinoma in situ (DCIS):** This is abnormal cell growth within the milk ducts of the breast. While not cancerous itself, it has the potential to become invasive if left untreated.
- **Lobular carcinoma in situ (LCIS):** Similar to DCIS, this involves abnormal cell growth, but within the milk-producing lobules of the breast. It also carries an increased risk of developing invasive cancer in the future.

Invasive breast cancer:

- **Invasive ductal carcinoma (IDC):** This is the most common type of invasive breast cancer, accounting for roughly 70-80% of all cases. It originates in the milk ducts and has the potential to spread to surrounding tissues.
- **Invasive lobular carcinoma (ILC):** The second most common invasive type, ILC starts in the lobules and can also spread beyond the breast.

4. Existing Breast Cancer Detection Techniques & Technologies

Breast cancer can be detected using various methods, including:

1. **Ultrasound** uses a probe impregnated with a gel that facilitates movement.

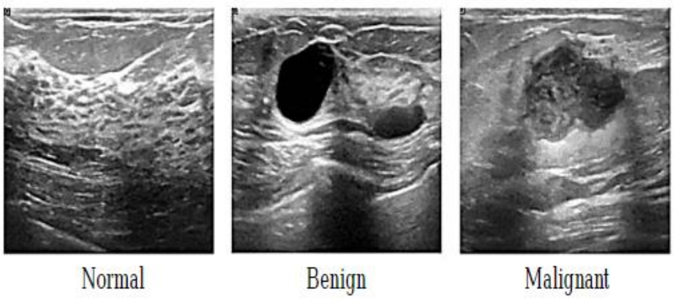


Figure 13 Sample of ultrasound breast image.

➔ ultrasound has limitations, such as **low specificity**, which can lead to **false-positive** results, and its **effectiveness depends on the skill of the sonographer**.

2. **Mammogram.** X-ray of the breast performed standing up with a device called a mammograph which requires moderate compression of the breast.

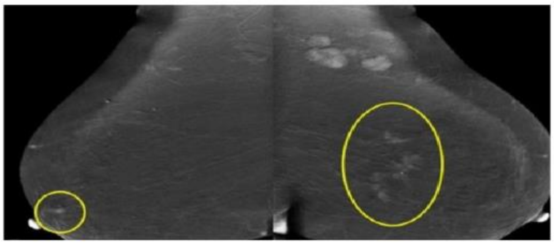


Figure 14 Sample of mammogram breast image.

➔ mammography has limitations, such **as low sensitivity for women with dense breasts**, **exposure to radiation,** and the **potential for false-positive results**.

3. **Breast magnetic resonance imaging (MRI).** A kind of body scan that uses a magnet linked to a computer. The MRI scan will make detailed pictures of areas inside the breast.

Figure 15 Typical finding of MRI-based breast cancer(arrow)

➔ MRI has limitations, such as **high cost and limited availability.**

4. **Biopsy.** This is a test that removes tissue or fluid from the breast to be looked at under a microscope and does more testing. There are different kinds of biopsies (for example, fine-needle aspiration, core biopsy, or open biopsy).

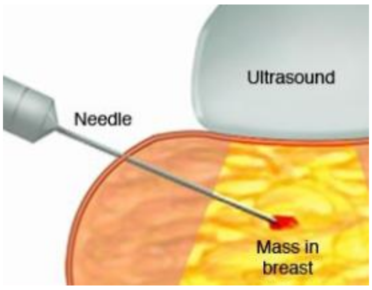


Figure 16 Biopsy.
Figure 17 Microwave Tomography (MWT):

5. **Radar-based Imaging** employs Ultra-Wideband (UWB) microwave radar imaging, utilizing reflected waves from illuminated breast tissues for reconstruction. This technique, also known as beamformers, offers faster detection compared to microwave tomography, requiring lower computational power and simpler signal processing.

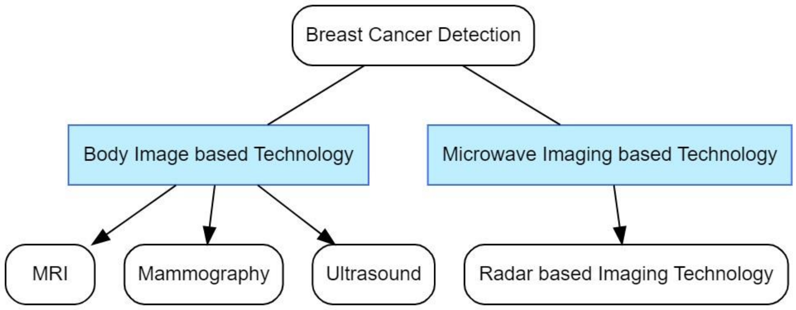


Figure 18 Block diagram showing the different modalities in breast cancer detection.

5. Existing diagnostic method, advantages, and limitations.

The table below gives a fast reference to comprehend the state of breast cancer diagnosis today and the need for an enhanced diagnostic strategy.

Diagnostic Method	Advantages	Limitations
Mammography	Well-established Widely accessible Detects structural changes and calcifications	Limited sensitivity in dense breast tissue False positives/negatives
Ultrasound	No radiation Useful for dense breasts Differentiates cysts from solid masses	Limited specificity Operator-dependent Limited detection in deep tissues
MRI (Magnetic Resonance Imaging)	High sensitivity No radiation Detailed soft tissue visualization	High cost Longer exam duration Requires specialized expertise to detect benign lesions
Biopsy	Provides tissue samples for definitive diagnosis. High diagnostic accuracy	Invasive and uncomfortable Small risk of complications Requires skilled medical staff. Sample may not be representative

Table 2 Existing diagnostic method, advantages, and limitations.

III. Global Statistics on Breast Cancer

Breast cancer casts a long shadow across the globe, posing a significant threat to women's health. This section delves into the sobering statistics that paint a clear picture of the immense challenge I face.

A Staggering Prevalence: Breast cancer reigns as the most diagnosed cancer among women worldwide. According to the World Health Organization (WHO), an estimated 2.3 million women received a breast cancer diagnosis in 2022. This translates to a staggering statistic: roughly one in four new cancer cases diagnosed in women is breast cancer. This sheer number of cases highlights the widespread impact of this disease, affecting millions of women and their families each year.

1. Leading Cause of Cancer Death:

The picture becomes even more concerning when considering mortality rates. Sadly, breast cancer isn't just the most common cancer; it's also the leading cause of cancer death in women globally. WHO estimates that in 2022, an estimated 670,000 women succumbed to breast cancer. This translates to roughly one in six cancer deaths among women being attributed to breast cancer.

These numbers underscore the urgency of improving early detection and treatment strategies to save lives.

This table shows global breast cancer mortality in women in 2020. Barbados had the highest rate of breast cancer mortality in women in 2020.

Rank	Country	Number	ASR/ 100,000
	World	*684,996*	*13.6*
1	Barbados	111	42.2
2	Fiji	184	41.0
3	Jamaica	637	34.1
4	Bahamas	80	31.0
5	Papua New Guinea	847	27.7
6	Somalia	1,189	27.2
7	Mali	1,425	26.6
8	Dominican Republic	1,577	26.4
9	Syria	1,946	26.2
10	Samoa	21	25.6

Table 3 global breast cancer mortality in women

2. Unequal Burden Across Regions:

The global landscape of breast cancer showcases significant disparities. Developed countries tend to have higher detection rates due to more readily available screening programs. Women in these regions are more likely to be diagnosed at an earlier stage, when treatment success rates are higher. However, advancements in treatment options also play a crucial role. Developed nations often have access to cutting-edge therapies, contributing to lower mortality rates. Conversely, lower- and middle-income countries often grapple with higher death rates due to limited access to early detection and effective treatment modalities. This creates a situation where the burden of breast cancer falls disproportionately on those with fewer resources to fight it.

This table shows global breast cancer incidence in women in 2020. Belgium had the highest rate of breast cancer in women in 2020, followed by the Netherlands.

Rank	Country	Number	ASR/100,000
	World	*2,261,419*	*47.8*
1	Belgium	11,734	113.2
2	The Netherlands	15,725	100.9
3	Luxembourg	497	99.8
4	France	58,083	99.1

5	France, New Caledonia	185	99.0
6	Denmark	5,083	98.4
7	Australia	19,617	96.0
8	New Zealand	3,660	93.0
9	Finland	5,228	92.4
10	US	253,465	90.3

Table 4 global breast cancer incidence in women in 2020

While statistics paint a stark picture, it's essential to consider the human cost of this disease. Breast cancer not only affects the physical health of women but also disrupts families and communities. It can lead to emotional distress, financial hardship, and a loss of productivity. The impact of breast cancer transcends mere numbers, highlighting the need for a multi-faceted approach to address this global health challenge.

IV. Breast Cancer in Morocco

Breast cancer casts a long shadow over the health of women in Morocco. This section delves into the concerning statistics, explores potential contributing factors, and highlights the need for improvement.

Breast cancer holds the dubious distinction of being the most prevalent cancer among Moroccan women. Studies indicate a concerning **33-36%** of all female cancers are breast cancer. This translates to a massive portion of the female population potentially facing this disease. This high prevalence necessitates a critical evaluation of the situation and a call to action for preventative measures.

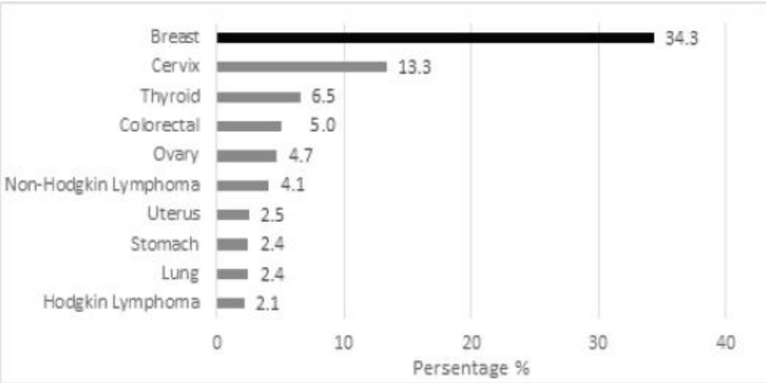


Figure 19 Ten Most Common Cancers among Moroccan Women, according to the Cancer Registry of Casablanca 2005-2007

In Morocco, two population-based registries are the principal sources of reliable cancer morbidity data:

- **Cancer Registry of Casablanca**: Covers around 12% of the Moroccan population and provides incidence data from 2004 onwards.
- **Cancer Registry of Rabat**: Covers around 2.1% of the Moroccan population and reports all new cancer cases from 2005 onwards.

1. The Challenge of Late Diagnosis:

Several factors might contribute to the high prevalence of late-stage diagnoses and the overall burden of breast cancer in Morocco:

- **Limited access to screening programs:** Early detection programs like mammograms are crucial for catching breast cancer early. While initiatives exist in Morocco, ensuring **wider accessibility** throughout the country remains a challenge, especially in remote areas.
- **Cultural and social barriers:** Stigmas surrounding breast cancer, or a lack of awareness can deter women from seeking preventive screenings or consultations. Open communication and educational campaigns can help address these barriers and empower women to take charge of their health.
- **Socioeconomic factors:** Poverty and limited access to healthcare can also hinder timely diagnosis and treatment. Addressing these factors is crucial for ensuring equitable access to healthcare services for all women in Morocco.

2. Comparison with Other Countries:

The age-standardized incidence rate of breast cancer in Morocco is higher than some neighboring countries like Algeria and Tunisia, but lower than some Western nations[1]. This highlights the need for tailored strategies considering regional factors.

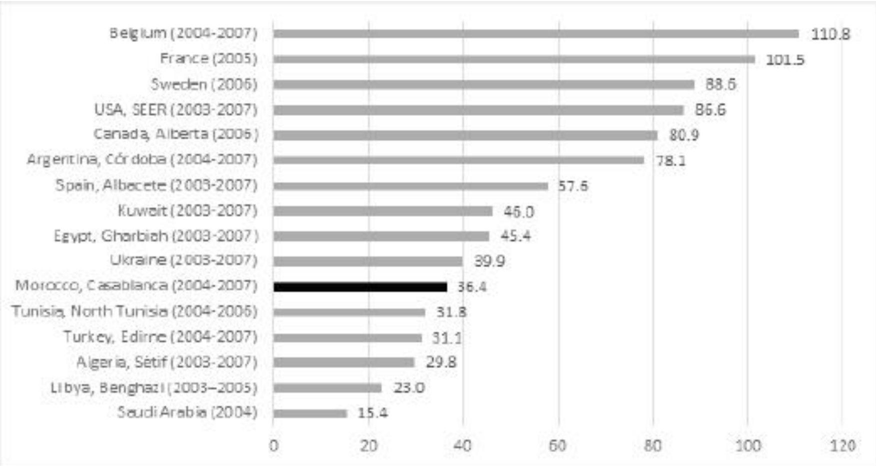


Figure 20 Age-Standardized Incidence (World Population) Rate for Female Breast Cancer in Morocco Compared to Other Countries.

V. Potential of Quantum Computing in Breast Cancer Detection

1. Advantages of Quantum Integration

Integrating quantum computing into breast cancer detection models offers several significant advantages over classical approaches. The unique properties of quantum computing can address the limitations of current detection methods, enhancing their effectiveness and efficiency. Here are the key benefits:

a. Enhanced Data Analysis Capabilities

Quantum computers excel at handling complex, high-dimensional data patterns. Breast cancer datasets often include intricate relationships and subtle patterns that classical algorithms may miss. Quantum computing can analyze these patterns more comprehensively, leading to more accurate detection models.

b. Improved Prediction Accuracy

The ability of quantum algorithms to process and analyze large datasets with greater precision can lead to improved prediction accuracy in breast cancer detection. Quantum models can potentially identify early-stage cancerous changes that classical models might overlook, thereby improving early diagnosis and patient outcomes.

c. Reduced Computational Time

Quantum computing can significantly reduce the time required to train and test machine learning models. Classical algorithms, especially those dealing with large and complex datasets, can be computationally intensive and time-consuming. Quantum algorithms can perform these tasks more efficiently, speeding up the overall process from data processing to model deployment.

d. Handling Data Quality and Quantity Issues

Breast cancer detection models often face challenges related to data quality, such as class imbalance and missing values. Quantum algorithms, particularly those designed for optimization problems, can handle these issues more effectively. They can optimize model parameters and manage imbalanced datasets to ensure robust and reliable detection outcomes.

Conclusion

This chapter has provided a comprehensive overview of the motivation behind enhancing breast cancer detection models with quantum integration. By addressing the global and regional context of breast cancer, the causes of the disease, the limitations of current detection methods, and the potential advantages of quantum computing, it sets the stage for the subsequent chapters that will delve into the technical details and implementation of the proposed approach.

Chapter 3: Literature Review and Theoretical Framework

Introduction:

Breast cancer remains a significant global health concern, necessitating effective and efficient detection methods to improve patient outcomes. In recent years, there has been a surge in research focusing on the utilization of advanced technologies, particularly machine learning and artificial intelligence, for breast cancer diagnosis. This literature review examines the current landscape of breast cancer detection methodologies, with a specific emphasis on ultrasound imaging and machine learning-based approaches. By synthesizing findings from selected studies published between 2018 and 2024, this review highlights the advancements, challenges, and future directions in the field of breast cancer detection. Additionally, the theoretical framework section provides an overview of quantum mechanics principles and their relevance to quantum computing, laying the foundation for exploring quantum algorithms and their potential applications in breast cancer detection.

I. Literature Review:

1. Research Questions:

- RQ1: To what extent is classical computing combined with quantum computing?
- RQ2: Does quantum machine learning provide speed increases compared to classical machine learning?
- RQ3: Does combining classical and quantum approaches lead to increased accuracy, or are there cases where classical machine learning performs better?
- RQ4: What are the advantages and limitations of quantum machine learning?

2. Selection Criteria:

Inclusion	**Exclusion**
Studies involving experimental results only	Studies without experimental results were not considered
Studies published from 2018 to 2024	Studies published before 2018 were excluded
Studies involving breast cancer detection only	Studies involving other cancer detections
Papers focusing on machine learning-based breast cancer detection	Papers focusing on other techniques used for breast cancer diagnosis
Studies written in English only	Studies written in other languages
Only journals and conferences are used	Other sources such as books, theses and magazines were excluded

Table 5 Selection Criteria

3. Previous work

In reference [2], a combination of deep learning technology and ultrasound imaging was employed. Tumor regions were segmented from breast ultrasound (BUS) images using a supervised block-based region segmentation algorithm. The best diagnostic results were achieved by creating a combination feature model based on the depth features of ultrasonic imaging and strain elastography. [3] addresses noisy labels in training breast tumor classification models with the noise filter network (NF-Net). Reference [4] describes a CAD system for tumor diagnosis using an image fusion method that combines various image content representations and employs ensemble techniques with different convolutional neural network (CNN) architectures on ultrasound images.

Reference [4] introduces a new BIRADS-SSDL network that incorporates clinically approved breast lesion characteristics (BIRADS features) into a task-oriented semi-supervised deep learning (SSDL) framework, aiming for precise diagnosis with limited training data. Reference [5] discusses a CAD system using an image fusion method and ensemble CNN architectures, noting that different image content representations enhance prediction accuracy, and including tumor shape features improves diagnostic effectiveness. Reference [6] focuses on pre-processing BUS images by resizing and enhancing them with contrast limited adaptive histogram equalization before generating a segmentation mask through concatenated convolutions.

Reference [7] presents a semi-supervised GAN model to enhance breast ultrasound images, which are then used for breast mass classification via a CNN, evaluated with 5-fold cross-validation. Reference [8] applies fuzzy enhancement and bilateral filtering algorithms to enhance original images, fusing decomposition images through RGB channels and using an optimal deep learning feature model with adaptive spatial feature fusion for classification. Reference [9] describes a deep learning-based method for classifying breast masses in ultrasound images, using deep representation scaling (DRS) layers to reduce trainable parameters and improve performance over conventional transfer learning methods.

Reference [10] explores transfer learning methods for classifying and detecting breast images in ultrasound, focusing on pre-processing techniques, pre-training models, and CNNs. Reference [11] proposes a novel CNN with a coarse-to-fine feature fusion approach for breast image segmentation, using an encoder-decoder architecture and super-pixel images with a weighted-balanced loss function to handle lesion size variations. Reference [12] introduces BUViTNet, which uses vision transformers (ViTs) instead of CNNs for breast ultrasound detection, leveraging datasets from ImageNet and cancer cells to achieve superior performance compared to traditional ViT and CNN-based transfer learning methods.

Reference [13] presents standards for breast ultrasound image segmentation evaluation, proposing standardized procedures for accurate annotations and introducing a loss-based approach to assess the impact of user interactions on segmentation sensitivity. Reference [14] proposes a BI-RADS classifier model for categorizing US breast lesions using a multi-class US image, bilinear interpolation, and neighborhood component analysis to generate informative features for automated classification. Reference [15] introduces an enhanced ViT architecture with a shared MLP head to balance feature learning between class and patch tokens, utilizing these tokens to distinguish between malignant and benign images and determine tumor area overlap.

4. Quantum Computing: Literature Analysis

This report provides a systematic literature review exploring the current state of Quantum Machine Learning (QML) from historical, theoretical, and technical viewpoints. A thorough analysis was ensured by applying specific criteria based on relevant keywords in publication titles.

Quantum computing, a rapidly evolving research field, aims to solve complex problems quickly, efficiently, and accurately using quantum systems. Key milestones in quantum computing history include Feynman's 1980s observation about classical computers' limitations in simulating quantum phenomena, Deutsch's 1985 demonstration of quantum computing's universality, Shor's 1994 prime factorization algorithm, and Grover's 1996 quantum cryptography algorithm. Achieving faster computation speeds remains a major challenge, with significant research efforts dedicated to this goal. In 2019, Google announced Quantum Supremacy, showcasing the superior computational power of quantum systems over classical computers.

Since 2011, publications on quantum computing have increased significantly, peaking in 2022, though there was a slight dip in 2016. The first decade (2000-2010) saw fewer publications, but interest grew notably from 2011 onward.

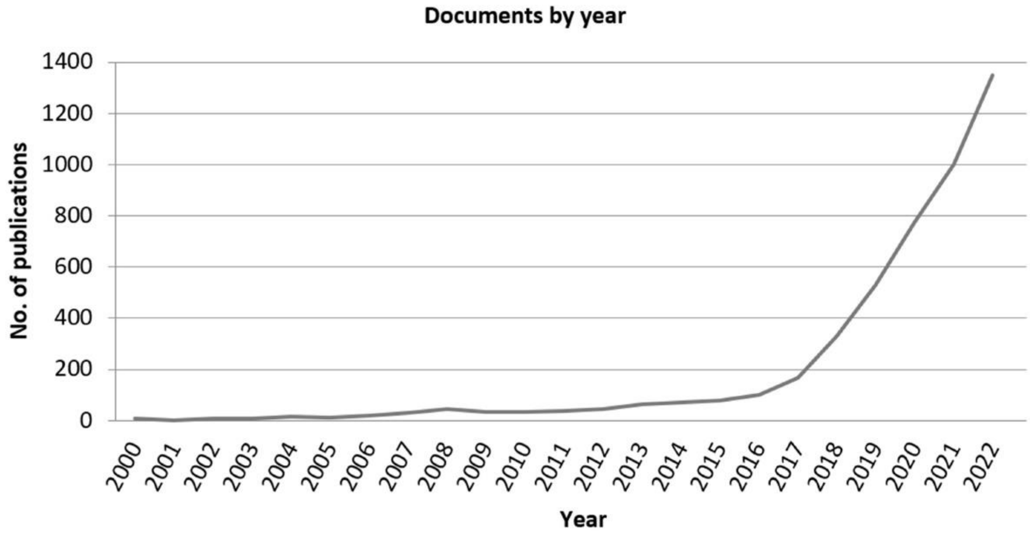


Figure 21 Number of relevant publications per year about QML (Statistics March 2023).

The top ten countries publishing QML articles include the United States and China, contributing almost 50% of the total. These countries' dominance is supported by their prominent quantum hardware providers.

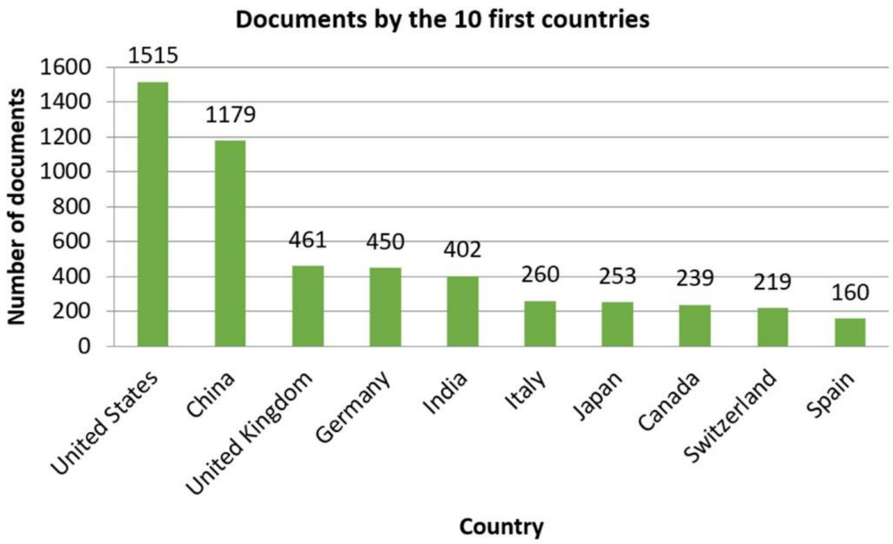


Figure 22 Countries that have published (Scopus source) papers related to QML.

Figure 23 Countries and collaboration network that have published researchers in Scopus related to QML.

Quantum machine learning has developed almost in parallel with classical machine learning, beginning before the advent of quantum computers. Important milestones include Feynman's 1981 lecture on quantum computation benefits, Deutsch's 1985 universal quantum computer concept, Shor's 1994 algorithm for prime factorization, and Grover's 1996 search optimization algorithm. In 1998, Oxford University's Jones, Mosca, and Hansen executed Grover's algorithm on a 2-qubit quantum computer.

In 2001, IBM and Stanford University implemented Shor's algorithm on a 7-qubit processor. In 2012, physicist John Preskill described the concept of quantum supremacy. In 2019, Google achieved quantum supremacy by performing a task beyond classical computers' capabilities.

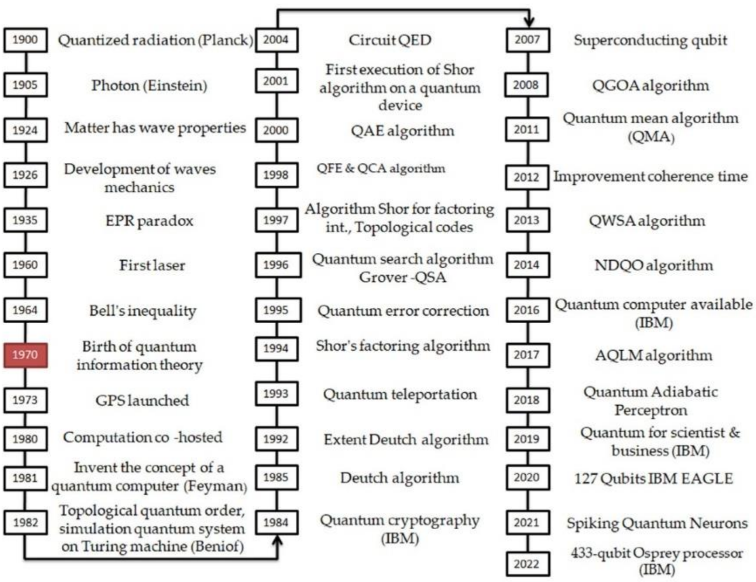


Figure 24 Timeline progression of QML milestones. Diagram highlighting major advancements in quantum machine learning.

Today, QML has evolved into quantum brain-inspired machine learning, which seeks to develop algorithms using the interactions and dynamics of quantum systems as learning resources, mimicking brain computation.

5. Current challenges in breast cancer detection Approaches

Breast cancer detection faces several challenges in the current scenario. One challenge is the variability and heterogeneity of cancer cells, as well as individual responses to treatment, which makes the treatment of cancer difficult. Another challenge is the discomfort, inaccuracy, cost, harmful radiation, and inconvenience associated with traditional detection methods like physical examination, X-rays, echography, and microwave imaging. Additionally, there is a need to improve the efficiency and outcomes of breast imaging through the adoption of artificial intelligence (AI)-based applications. Underserved populations in low- and middle-income countries (LMIC) face significant disparities in breast cancer detection and access to healthcare, leading to late diagnoses and suboptimal treatment results. Finally, the current clinical detection techniques have reliability issues, long waiting times, and painful procedures, which hinder early diagnosis and proper treatment.

6. Limitations of Classical Models

Machine learning (ML) in breast cancer diagnosis faces challenges and limitations despite its potential. One major issue is the lack of diversity in training datasets, often not representing all populations[16] .Additionally, the scarcity of highly annotated medical data poses a challenge for effective algorithm training[17], [18]. Deep learning models, requiring substantial data, may not perform optimally due to limited medical image datasets[18]. Moreover, the need for standardized processes for data collection and analysis remains crucial for enhancing ML accuracy in breast cancer diagnosis. Addressing these challenges through improved dataset diversity, increased annotation, and standardized procedures can enhance the efficacy of machine learning applications in breast cancer diagnosis.

II. Quantum Mechanics Principles

$$i\hbar\frac{\partial}{\partial t}|\Psi\rangle = \hat{H}|\Psi\rangle$$

Figure 25 Schrödinger equation.

Quantum mechanics is the branch of physics that describes the behavior of particles at the smallest scales, such as atoms and subatomic particles. It revolutionized our understanding of nature by introducing concepts that defy classical intuition.

1. Superposition:

In quantum mechanics, particles can exist in multiple states simultaneously, known as superposition. This means that until a measurement is made, a particle can be in a combination of different states with different probabilities.

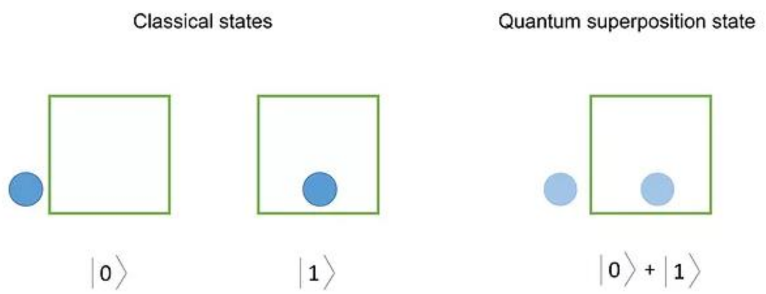


Figure 26 Superposition.

One of the most famous examples of superposition in quantum mechanics is the double-slit experiment. In this experiment, a beam of electrons is fired at a screen with two slits in it. On the other side of the screen, an observation screen is placed to detect the electrons. When electrons are fired one at a time, a diffraction pattern of bright and dark bands is observed on the observation screen. This pattern is the result of the electrons passing through both slits simultaneously and interfering with each other.

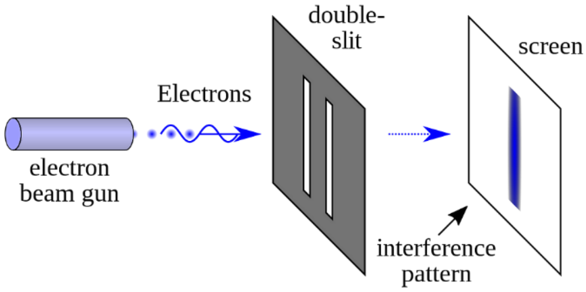


Figure 27 Double-slit experiment.

Another example of superposition in quantum mechanics is the superposition of energy states in atoms. Atoms can exist in multiple energy states at the same time, known as superposition of energy levels. This is known as the Schrödinger's cat paradox, where a cat can be both alive and dead at the same time, until the state is observed.

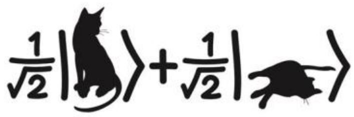


Figure 28 Schrödinger's cat paradox.

Superposition is a key principle in the field of quantum computing, which aims to harness the unique properties of quantum systems for computational purposes. Quantum computers use qubits (quantum bits) instead of classical bits, and the qubits can exist in superposition of states, which allows them to perform certain calculations much faster than classical computers. Quantum algorithms such as Shor's algorithm and Grover's algorithm, which can factor large numbers and search large databases much faster than classical algorithms, rely on superposition.

2. Entanglement:

Entanglement occurs when the quantum states of two or more particles become correlated in such a way that the state of one particle cannot be described independently of the state of the others, no matter the distance between them.

One of the most striking examples of entanglement is the Einstein-Podolsky-Rosen (EPR) paradox, proposed in 1935. In this thought experiment, two particles are prepared in such a way that their properties are correlated. According to quantum mechanics, measuring the state of one particle instantaneously determines the state of the other, even if they are separated by vast distances.

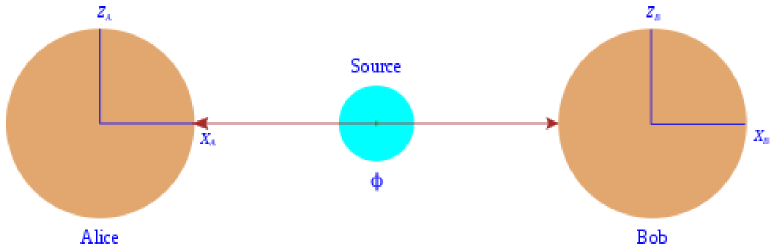


Figure 29 Einstein-Podolsky-Rosen (EPR) paradox.

Another example of entanglement is demonstrated in Bell tests, which are experiments designed to test the violation of Bell inequalities. These tests confirm that entanglement is a real physical phenomenon that cannot be explained by classical theories.

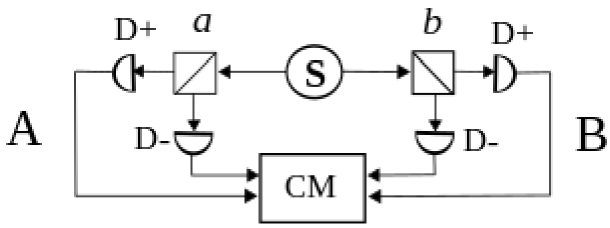


Figure 30 Bell tests.

Entanglement plays a crucial role in various quantum technologies, including quantum cryptography and quantum teleportation. In quantum computing, entanglement enables the creation of quantum circuits with exponentially more computational power than classical circuits. Quantum algorithms such as teleportation and superdense coding exploit entanglement to achieve tasks that are impossible with classical resources alone.

3. Quantum Measurement:

Measurement in quantum mechanics is a process that causes the quantum state of a system to 'collapse' to one of the possible eigenstates of the observable being measured. The outcome of a measurement is probabilistic and depends on the state of the system.

One of the key aspects of quantum measurement is that the outcome is probabilistic. Unlike classical physics, where the outcome of a measurement can be predicted with certainty given sufficient information, quantum mechanics introduces inherent uncertainty. The outcome of a measurement is determined by the probabilities associated with each possible state of the system.

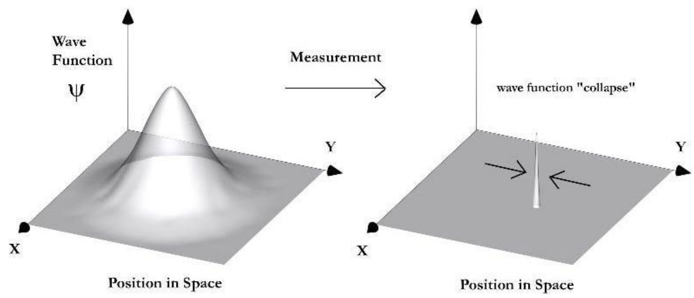


Figure 31 Quantum Measurement

Let's consider a quantum system described by a wavefunction $|\psi\rangle$in a Hilbert space. Suppose we want to measure an observable A, represented by a Hermitian operator $\hat{A}$, with eigenvalues a_i and corresponding eigenvectors $|a_i\rangle$.

The state $|\psi\rangle$ can be expanded in terms of the eigenstates of A as:

$$|\psi\rangle=\sum_i c_i|a_i\rangle$$

where c_i are the probability amplitudes of the system being in the eigenstates $|a_i\rangle$.

Upon measurement of A, the outcome a_i is obtained with probability $|c_i|^2$. After measurement, the state of the system collapses to the corresponding eigenstate $|a_i\rangle$.

Mathematically, the collapse of the wavefunction can be represented as:

$$|\psi\rangle \rightarrow \text{Measurement of } A \rightarrow |ai\rangle \text{ with } |c_i|^2$$

This process illustrates the probabilistic nature of quantum measurement, where the outcome is not deterministic but governed by the probabilities associated with each possible state of the system.

Despite its crucial role in quantum mechanics, the nature of quantum measurement remains a subject of debate and interpretation among physicists. Various interpretations, such as the Copenhagen interpretation and the Many-Worlds interpretation, offer different perspectives on the underlying mechanisms of measurement and its implications for our understanding of reality.

4. Interpretations of Quantum Mechanics

a. Copenhagen Interpretation:

Proposed by Niels Bohr and Werner Heisenberg in the 1920s, the Copenhagen interpretation is one of the most widely known and accepted interpretations of quantum mechanics.

According to this interpretation, quantum systems exist in superposition of states until they are measured, at which point the wavefunction collapses, and the system assumes a definite state.

The role of the observer is central in the Copenhagen interpretation, as it emphasizes the act of measurement as the means by which reality is defined.

However, the Copenhagen interpretation does not provide a clear explanation of what constitutes a measurement or why collapse occurs.

b. Many-Worlds Interpretation:

Quantum Bayesianism is a subjective interpretation of quantum mechanics that views quantum probabilities as Bayesian probabilities, representing an agent's beliefs or degrees of certainty.

According to QBism, quantum states and observables are not objective properties of the world but rather expressions of an observer's personal beliefs or experiences.

Below we have a Bloch ball, each point in the Bloch ball is a possible quantum state for a qubit. In QBism, all quantum states are representations of personal probabilities.

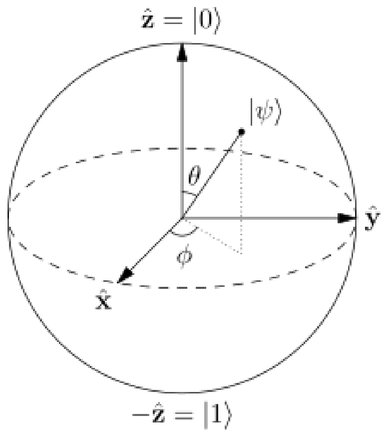


Figure 32 Bloch ball.

QBism emphasizes the role of the observer in constructing reality and understanding quantum phenomena.

III. Quantum Computing Principles

Quantum computing represents a paradigm shift in computation, harnessing the principles of quantum mechanics to perform computations that are infeasible for classical computers. This section explores the foundational concepts underlying quantum computing and its potential applications.

1. Quantum Computer

Quantum computers utilize the principles of quantum physics to perform certain tasks more efficiently than classical supercomputers. They act as co-processors alongside traditional computers, similar to how GPUs assist in video gaming and deep learning.

Control and Operation

Traditional computers precisely control quantum operations by triggering qubit operations via quantum gates, considering the execution time and coherence time of the qubits. This hybrid system ensures accurate control and synchronization.

Key Components

1. **Quantum Registers:**
   - Collections of qubits that store information.
   - Exploit superposition to hold and manipulate multiple values simultaneously.
   - Record benchmark: 433 qubits (IBM, November 2022).
2. **Quantum Gates:**
   - Physical systems act on qubits to initialize and perform computations.
   - Applied iteratively based on the algorithms in use.
3. **Measurement Interface:**
   - Retrieves computational results post quantum gate operations.
   - Involves multiple setup, computation, and measurement cycles to ensure accuracy.
   - Results are averaged and translated into digital values for interpretation by the traditional computer.
   - Typical quantum computers, such as those from D-Wave and IBM, repeat calculations at least 1024 times.
4. **Quantum Chipset:**

- Includes quantum registers, gates, and measurement devices, especially in superconducting qubits.
- Current chipsets are relatively small, with the 433-qubit Osprey being about the size of a quarter.

5. **Refrigerated Enclosure:**
   - Maintains near absolute zero temperatures to minimize disturbances.
   - Houses part of the control electronics and the quantum chipset to ensure qubit coherence and reduce operational noise.
6. **Electronic Writing and Reading:**
   - Controls initialization, updating, and reading of qubit states within the refrigerated enclosure.

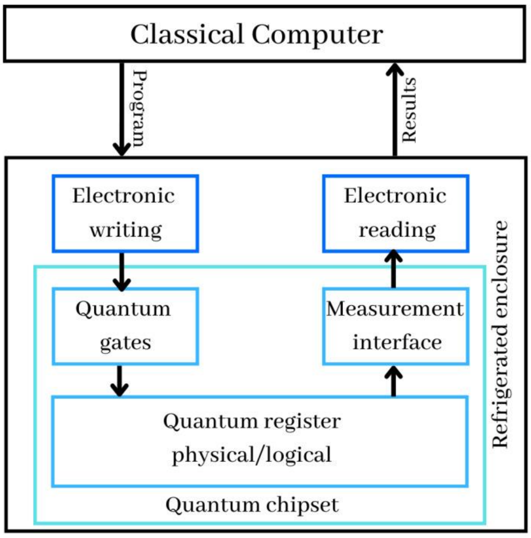


Figure 33 The quantum computer's architecture.

Quantum computers today are built based on atoms (e.g., cold atoms, trapped ions and nuclear magnetic resonance), electrons (e.g., superconducting, silicon and topological), or photon (e.g., linear optics). D-Wave, IBM and Google have all shown significant quantum computer development progress. The D-Wave model is built using the quantum adiabatic technique, which handles issues involving optimization or probabilistic sampling, whereas the IBM model is mostly dealing with non-adiabatic models. Most of those companies provide a quantum computer simulator that runs on a classical computer to test code before executing it on a quantum device. It is possible to use this simulator in either a local environment or the cloud. It cannot handle true quantum states because it is operating on a classical computer, but it is useful for testing code syntax and flow.

1. Dirac (Bra-Ket) Notation

States and operators in quantum mechanics are represented as vectors and matrices, respectively. Instead of utilizing standard linear algebra symbols, Dirac notation is used to represent the vectors. Let a and b be in C^2:

Ket:

$$|a\rangle = \begin{matrix} a_1 \\ a_2 \end{matrix}$$

Bra:

$$\langle b| = |b\rangle^* = \begin{matrix} b_1 \\ b_2 \end{matrix}^* = \begin{matrix} b_1^* & b_2^* \end{matrix}$$

The complex conjugate of any complex number can be generated by inverting the sign of its imaginary component—for example, the complex conjugate of b=a+i×d is b*=a−i×d.

Bra-Ket: Inner product

$$\langle b \mid a\rangle = a_1 b_1^* + a_2 b_2^* = \langle a \mid b\rangle^*$$

Ket-Bra: Outer product

$$|a\rangle\langle b| = \begin{matrix} a_1 b_1^* & a_1 b_2^* \\ a_2 b_1^* & a_2 b_2^* \end{matrix}$$

2. Quantum Bits (Qubits)

Quantum computing operates on quantum bits or qubits, which are the fundamental units of information in quantum systems. Unlike classical bits, which can only exist in states of 0 or 1, qubits can exist in superposition, representing both 0 and 1 simultaneously. This superposition enables quantum computers to perform multiple calculations in parallel, exponentially increasing computational efficiency.

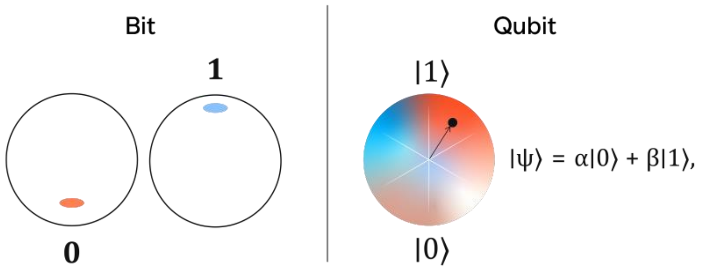


Figure 34 Quantum Bits (Qubits)

3. Quantum Gates

Quantum gates are fundamental components in quantum computing, analogous to classical logic gates in traditional computing. They operate on qubits, the basic units of quantum information, and manipulate their quantum states to perform computations. Quantum gates are essential for implementing quantum algorithms and executing quantum operations.

a. Types of Quantum Gates:

1. **Pauli Gates:**

These gates, named after physicist Wolfgang Pauli, are fundamental quantum gates that perform basic operations on qubits. The Pauli-X gate (σx), Pauli-Y gate (σy), and Pauli-Z gate (σz) correspond to rotations around the x, y, and z axes of the Bloch sphere, respectively. They are equivalent to classical NOT gates but operate in the quantum domain.

2. **Hadamard Gate (H):**

The Hadamard gate is a key quantum gate that creates superposition states.

$$H: = \frac{1}{\sqrt{2}} \begin{matrix} 1 & 1 \\ 1 & -1 \end{matrix}$$

When Hadamard gate is applied to state $|0\rangle$ we obtain a qubit in an equal superposition of $|0\rangle$ and $|1\rangle$ states

$$\frac{1}{\sqrt{2}} \begin{matrix} 1 & 1 \\ 1 & -1 \end{matrix} \begin{matrix} 1 \\ 0 \end{matrix} = \frac{1}{\sqrt{2}} \begin{matrix} 1 \\ 1 \end{matrix} = \frac{|0\rangle + |1\rangle}{\sqrt{2}},$$

and to a $|1\rangle$ we obtain

$$\frac{1}{\sqrt{2}} \begin{matrix} 1 & 1 \\ 1 & -1 \end{matrix} \begin{matrix} 0 \\ 1 \end{matrix} = \frac{1}{\sqrt{2}} \begin{matrix} 1 \\ -1 \end{matrix} = \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

The Hadamard gate is crucial for generating and manipulating quantum superposition, a fundamental property exploited in many quantum algorithms.

3. **Phase Gates:** Phase gates introduce phase shifts to qubits, modifying the relative phase between different quantum states. The most common phase gate is the Phase Shift gate (S), which introduces a $\pi/2$ phase shift to the $|1\rangle$ state. The Controlled Phase gate (CPHASE) applies a phase shift to the target qubit based on the state of the control qubit, enabling the creation of entangled states.

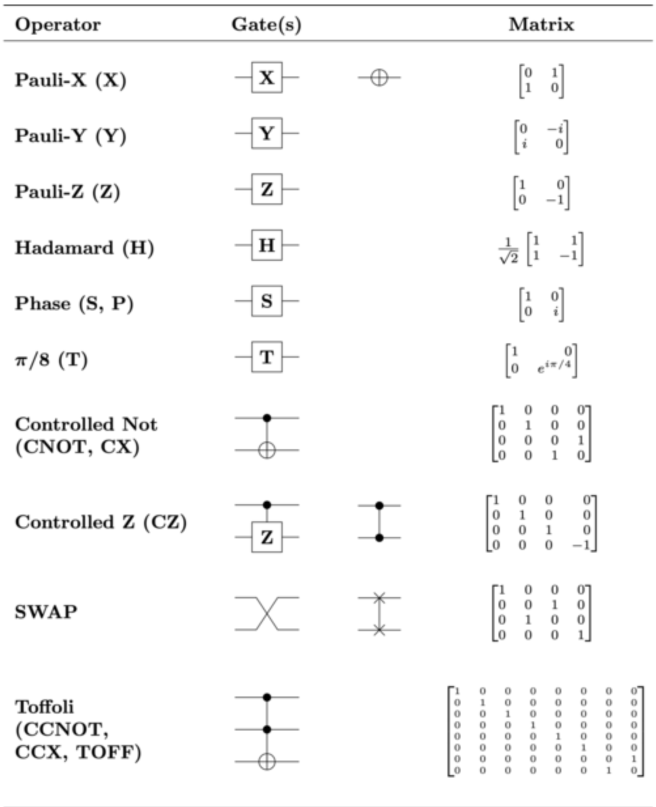
4. **Controlled Gates:** Controlled gates operate on multiple qubits, with one qubit serving as the control and another as the target.

The most widely used controlled gate is the Controlled-NOT gate (CNOT), which flips the state of the target qubit if and only if the control qubit is in the $|1\rangle$ state.

Before Applying CNOT		**After Applying CNOT**	
Controlled	**Targeted**	**Controlled**	**Targeted**
$|0\rangle$	$|0\rangle$	$|0\rangle$	$|0\rangle$
$|0\rangle$	$|1\rangle$	$|0\rangle$	$|1\rangle$
$|1\rangle$	$|0\rangle$	$|1\rangle$	$|1\rangle$
$|1\rangle$	$|1\rangle$	$|1\rangle$	$|0\rangle$

Table 6 CNOT gate truth table.

Controlled gates are essential for implementing quantum logic and creating entangled states, which are central to many quantum algorithms.



Operator	Gate(s)	Matrix
Pauli-X (X)	X	$\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix}$
Pauli-Y (Y)	Y	$\begin{bmatrix}0 & -i\\i & 0\end{bmatrix}$
Pauli-Z (Z)	Z	$\begin{bmatrix}1 & 0\\0 & -1\end{bmatrix}$
Hadamard (H)	H	$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\1 & -1\end{bmatrix}$
Phase (S, P)	S	$\begin{bmatrix}1 & 0\\0 & i\end{bmatrix}$
$\pi/8$ (T)	T	$\begin{bmatrix}1 & 0\\0 & e^{i\pi/4}\end{bmatrix}$
Controlled Not (CNOT, CX)		$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&0&1\\0&0&1&0\end{bmatrix}$
Controlled Z (CZ)	Z	$\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&-1\end{bmatrix}$
SWAP		$\begin{bmatrix}1&0&0&0\\0&0&1&0\\0&1&0&0\\0&0&0&1\end{bmatrix}$
Toffoli (CCNOT, CCX, TOFF)		$\begin{bmatrix}1&0&0&0&0&0&0&0\\0&1&0&0&0&0&0&0\\0&0&1&0&0&0&0&0\\0&0&0&1&0&0&0&0\\0&0&0&0&1&0&0&0\\0&0&0&0&0&1&0&0\\0&0&0&0&0&0&0&1\\0&0&0&0&0&0&1&0\end{bmatrix}$

Figure 35 Quantum Gates

b. Quantum Gates Implementation:

Quantum gates are realized physically through various quantum systems, including trapped ions, superconducting circuits, and photonics. Each quantum system has its advantages and challenges, but all rely on manipulating the quantum states of qubits to execute quantum operations.

1. **Trapped Ions:** In trapped ion quantum computing, ions are trapped using electromagnetic fields and manipulated using laser pulses. Quantum gates are implemented by applying laser beams to induce transitions between different energy levels of the trapped ions, effectively executing quantum operations.

2. **Superconducting Circuits:** Superconducting quantum computing utilizes superconducting circuits composed of Josephson junctions to create qubits. Quantum gates are implemented by applying microwave pulses to the superconducting qubits, which induce transitions between different energy levels, allowing for the execution of quantum operations.

3. **Photonics:** Photonic quantum computing relies on manipulating photons to encode and process quantum information. Quantum gates are implemented using optical elements such as beam splitters, phase shifters, and detectors, which manipulate the quantum states of photons to perform quantum operations.

In all these implementations, precise control and manipulation of qubits are crucial for executing quantum gates accurately and reliably. Quantum error correction techniques and fault-tolerant protocols are employed to mitigate errors and decoherence, ensuring the integrity of quantum computations.

4. Representation of Qubit States

The state of a qubit may be represented in different ways. Dirac notation allows us to express this state in a readable form. A qubit in state $|0\rangle$, for instance, will transfer to state $|1\rangle$ after the application of the X operator.

$$X|0\rangle \rightarrow |1\rangle$$

In the figure below, a state of a single qubit is represented by the Bloch sphere. Quantum states are represented by vectors that extend from the origin to a certain point on the surface of the Bloch sphere. The top and bottom antipodes of the sphere are $|0\rangle$ and $|1\rangle$, respectively.

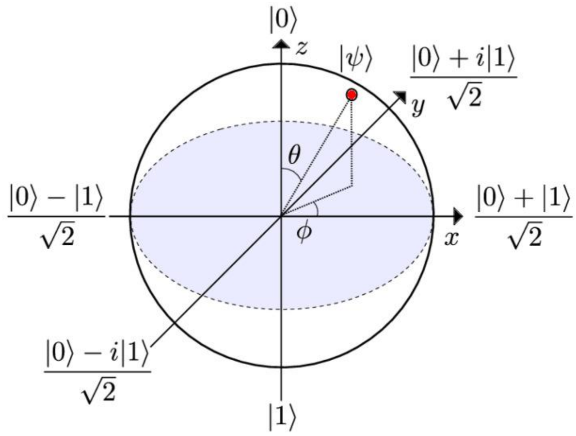


Figure 36 Bloch sphere representation of the state of qubit.

On the Bloch sphere, we may write any *pure state* (i.e., a qubit state specified by a vector of norm 1. *Mixed state*, in contrast, is a state that combines numerous pure quantum states or qubits) as seen below:

$$|\psi\rangle = \cos\theta / 2|0\rangle + e^{i\phi}\sin\theta / 2|1\rangle$$

with $\theta \in [0, \pi]$, determines the probability to measure the state $|0\rangle$ as $P(0)=\cos^2\theta/2$ and the state $|1\rangle$ as $P(1)=\sin^2\theta/2$, and $\phi \in [0,2\pi]$ describes the relative phase. On the surface of a Bloch sphere, all of these pure states may be represented with a radius of $|\vec{r}|=1$. The Bloch vector generates such a state's coordinates:

$$\vec{r} = \begin{matrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\theta) \end{matrix}$$

as result,

$$\begin{aligned}
&|0\rangle: \theta = 0, \phi \text{ arbitrary } \rightarrow \vec{r} = (0,0,1); \\
&|1\rangle: \theta = \pi, \phi \text{ arbitrary } \rightarrow \vec{r} = (0,0,-1); \\
&|+\rangle: \theta = \pi / 2, \phi = 0 \rightarrow \vec{r} = (1,0,0); \\
&|-\rangle: \theta = \pi / 2, \phi = \pi \rightarrow \vec{r} = (-1,0,0); \\
&|+i\rangle: \theta = \pi / 2, \phi = \pi / 2 \rightarrow \vec{r} = (0,1,0); \\
&|-i\rangle: \theta = \pi / 2, \phi = 3\pi / 2 \rightarrow \vec{r} = (0,-1,0).
\end{aligned}$$

Bloch sphere is only capable of representing the state of a particular qubit. Therefore, the Q−$sphere$ is used for multi-qubits (and single qubits as well).

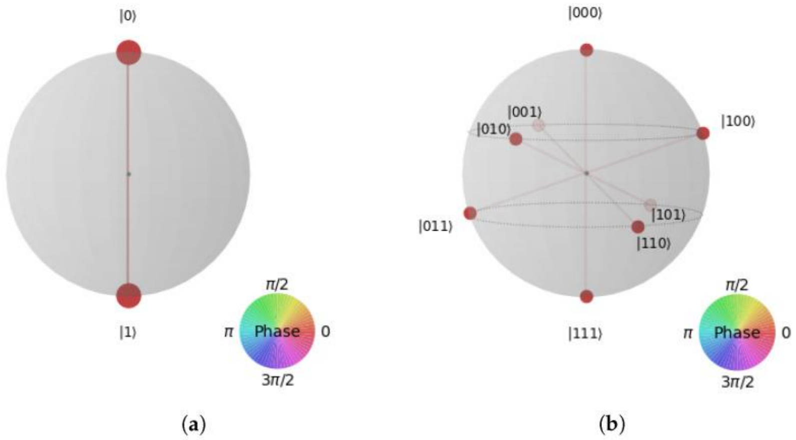


Figure 37 Representations of single qubit and multi-qubit in Q-sphere. (a) Representation of a superposition state; (b) Representation of three-qubit states.

- The south pole represents the state $|1\rangle$.
- The north pole represents the state $|0\rangle$.
- The size of the blobs is related to the likelihood that the relevant state will be measured.
- The color indicates the relative phase compared to the state $|0\rangle$.

5. Quantum Algorithms

Quantum algorithms represent a significant departure from classical algorithms, leveraging the principles of quantum mechanics to solve computational problems with unprecedented efficiency. This section explores some of the most notable quantum algorithms and their applications across various domains.

a. Shor's Algorithm:

Shor's algorithm, proposed by mathematician *Peter Shor* in 1994, revolutionized cryptography by demonstrating the potential to efficiently factor large integers, a task believed to be intractable for classical computers.

The algorithm exploits quantum parallelism and the periodicity properties of modular exponentiation to factorize integers into their prime factors exponentially faster than classical algorithms such as the General Number Field Sieve.

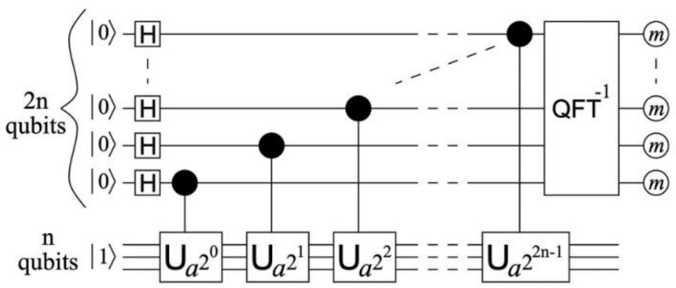


Figure 38 Shor's Algorithm

Shor's algorithm has profound implications for cryptographic protocols such as RSA, threatening their security by enabling the rapid factorization of large composite numbers used in encryption keys.

b. Grover's Algorithm:

Grover's algorithm, developed by Lov Grover in 1996, addresses the problem of unstructured search, where the goal is to find a specific item in an unsorted database.

Unlike classical search algorithms that require O(N) time complexity, where N is the number of items in the database, Grover's algorithm achieves a quadratic speedup, requiring only O(√N) iterations to find the target item.

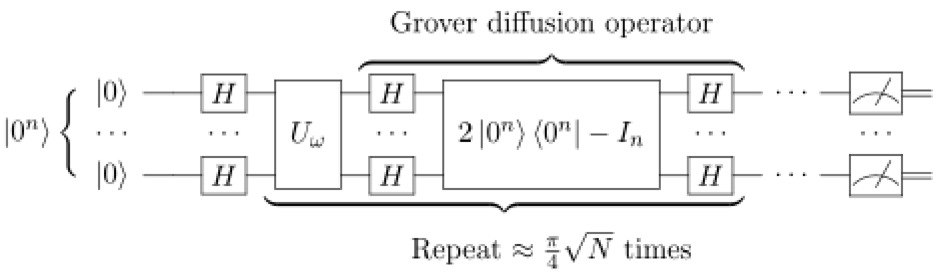


Figure 39 Grover's Algorithm

The algorithm achieves this speedup by exploiting quantum amplitude amplification, which amplifies the amplitude of the target state while suppressing the amplitudes of non-target states, leading to a probabilistic speedup in search.

c. Quantum Simulation Algorithms:

Quantum simulation algorithms aim to simulate the behavior of quantum systems, such as molecular dynamics, condensed matter physics, and quantum chemistry, which are inherently difficult to simulate on classical computers.

These algorithms, including the Variational Quantum Eigensolver (VQE), Quantum Phase Estimation (QPE), and Quantum Approximate Optimization Algorithm (QAOA), leverage the quantum nature of matter to accurately model complex quantum phenomena.

Quantum simulation algorithms have diverse applications, ranging from materials science and drug discovery to optimization and machine learning, enabling the exploration of novel materials, the design of efficient chemical processes, and the discovery of new pharmaceutical compounds.

d. Quantum Machine Learning:

Quantum machine learning algorithms represent a fusion of quantum computing principles with classical machine learning techniques, aimed at solving computational tasks more efficiently and addressing challenges that arise in quantum domains.

There are four strategies of how to combine machine learning and quantum computing, based on if one considers the data was created by a classical (C) or quantum (Q) system, and whether the computer that processes data are classical (C) or quantum (Q), (as shown in the figure below).

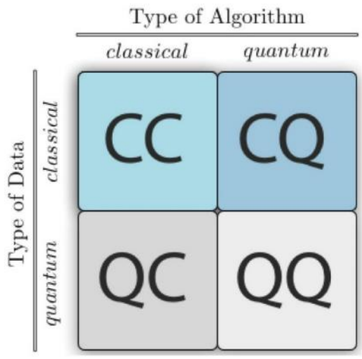


Figure 40 Four separate ways to combine quantum computing with machine learning.

In exploring the intersection of classical and quantum computing, various methodologies emerge, each with distinct focuses and techniques.

The CC scenario employs traditional data treatment methods, leveraging machine learning techniques inspired by quantum information research, such as tensor networks and quantum-inspired algorithms.

Conversely, the QC scenario investigates how machine learning can enhance quantum computing, aiding in the analysis of measurement data and distinguishing quantum states.

CQ methodology applies quantum computing to analyze classical datasets, aiming to develop quantum algorithms for data mining.

Meanwhile, the QQ scenario centers on processing quantum data with quantum devices, either analyzing experimental data or simulating quantum systems' behavior.

In this study, emphasis is placed on the CQ scenario, exploring two main approaches: running traditional machine learning algorithms on quantum computers for speedups, or developing quantum machine learning algorithms based on quantum subroutines.

The figure below illustrates the processing strategies used in conventional machine learning and QML. In traditional machine learning, data is a direct input to the algorithm, which then analyses the data and generates an output. QML, on the other hand, demands the initial encoding of the data into quantum data. QML receives quantum data as input, processes it, and generates quantum data as output. The quantum data is then converted to conventional data.

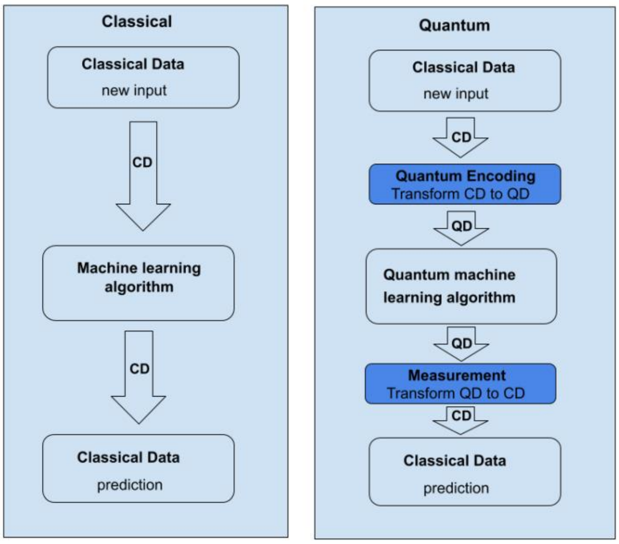


Figure 41 Processing techniques of conventional machine learning and quantum machine learning.

e. Quantum Machine Learning Algorithms:

QMLs leverage quantum properties such as superposition and entanglement to enhance learning and classification tasks, offering potential advantages over classical approaches. Here are some prominent examples:

1. Quantum Support Vector Machine (QSVM):

QSVM is a quantum-enhanced version of the classical support vector machine (SVM), a widely used supervised learning algorithm for classification tasks.

It utilizes quantum algorithms to efficiently compute the inner product between feature vectors, which forms the basis of SVM's decision boundary calculations.

QSVM exploits quantum parallelism and the quantum kernel trick to speed up the classification process and improve the accuracy of SVM models.

2. Quantum Neural Networks (QNN):

QNNs are quantum-analogue counterparts of classical artificial neural networks (ANNs), designed to perform machine learning tasks using quantum processing units.

QNNs leverage quantum entanglement and superposition to process information and learn complex patterns more efficiently than classical ANNs.

They offer potential advantages in tasks such as pattern recognition, feature learning, and optimization, particularly in scenarios where the input data or the learning process exhibits quantum behavior.

3. Quantum k-Means Clustering:

Quantum k-Means Clustering is a quantum algorithm inspired by the classical k-means clustering algorithm, which partitions data points into k clusters based on their similarities.

It uses quantum circuits to represent data points and quantum operations to perform clustering in a quantum state space.

Quantum k-Means Clustering has applications in data analysis, pattern recognition, and unsupervised learning, offering potential speedups over classical clustering algorithms.

4. Variational Quantum Circuits (VQC):

The Variational Quantum Classifier (VQC) is based on the use of a function $f(x, \theta)=y$ which can be implemented in a quantum computer using a circuit denoted as Sx in order to encode the input data *x* toward a quantum state and especially as amplitudes of the state, then a quantum circuit denoted as $U\theta$, finally a one simple qubit measurement. This latter measurement yields the likelihood of the VQC predicting '0' or '1', which may be used to determine the prediction of a binary classification. The circuit parameters (θ) of this classification are also trainable, and a variational technique may be employed to train these parameters. The four steps of VQC are shown in the figure and explained below.

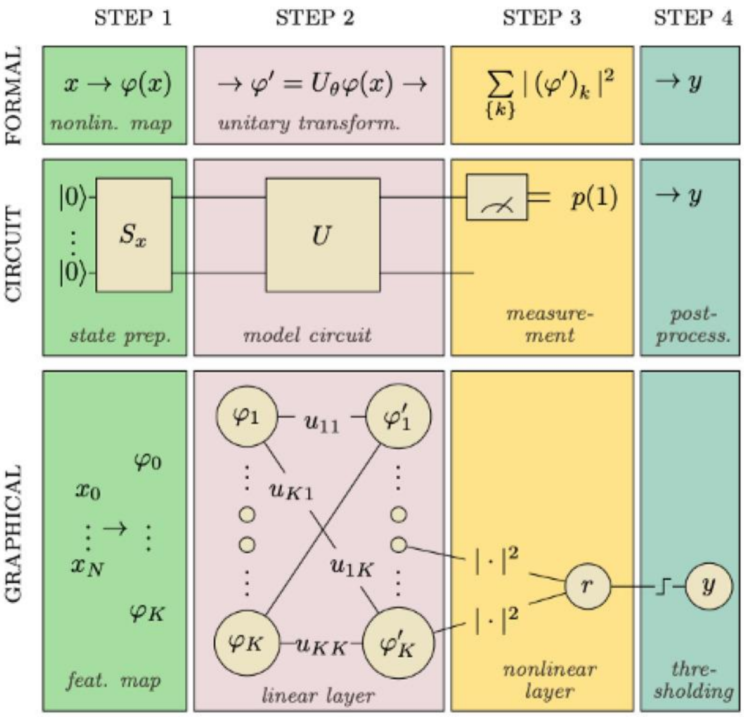


Figure 42 The four steps of a VQC.

- **State preparation:** To be able to encode classical data into quantum states, we use operations to help us work with data in a quantum circuit. As mentioned earlier, quantum encoding is one these methods that consists of representing classical data in the form of a quantum state in Hilbert space employing a quantum feature map. Recall that a feature map is a mathematical mapping that allows us to integrate our data into higher dimensional spaces, such as quantum states in our case. It is similar to a variational circuit in which the parameters are determined by the input data. It is essential to emphasize that a variational circuit depends on parameters that can be optimized using classical methods.
- **The model circuit:** The next step is the model circuit, or the classifier in precise terms. φ' is generated using a parameterized unitary operator $U\theta$ applied to the feature vector noted as $\varphi(x)$ which became a vector of a quantum state in an n-qubit system (in the Hilbert space). The model uses a circuit that is composed of gates which change the state of the input and are built on unitary processes, and they depend on external factors that can be

adjusted. U_θ translates $\varphi(x)$ into another vector φ' with a prepared state $|\varphi(x)\rangle$ in the model circuit. U_θ is comprised of a series of unitary gates.

- **Measurement:** We take measurements to obtain information from a quantum system. Although a quantum system has an infinite number of potential states, we can only recover a limited amount of information from a quantum measurement. Notice that the number of qubits is equal to the number of results.
- **Post-process**: Finally, the results were post-processed including a learnable bias parameter and a step function to translate the result to the outcome 0 or 1.

Conclusion

In conclusion, this chapter provides a comprehensive overview of the literature on breast cancer detection methodologies and introduces the fundamental principles of quantum mechanics and quantum computing. The reviewed studies underscore the growing interest in leveraging machine learning techniques to enhance the accuracy and efficiency of breast cancer diagnosis, particularly in the realm of ultrasound imaging analysis. However, several challenges persist, including dataset diversity, annotation scarcity, and the need for standardized processes. Looking ahead, the integration of quantum algorithms holds promise for revolutionizing computational approaches to breast cancer detection, offering unprecedented speed and efficiency. By addressing existing limitations and embracing emerging technologies, the field is poised to make significant strides in early detection and treatment of breast cancer, ultimately improving patient outcomes and reducing the global burden of this disease.

Chapter 4: Design and Methodology

Introduction

This chapter details the methodology for comparing classical and quantum models in breast cancer detection. It covers data acquisition, exploratory data analysis, preprocessing steps, and the setup for both classical machine learning models and the Variational Quantum Classifier (VQC).

I. Data Acquisition

The dataset utilized in this analysis is the Breast Cancer (Wisconsin) Diagnosis dataset, sourced from the UCI Machine Learning Repository. This dataset comprises diagnostic information and a comprehensive set of 30 features characterizing cell nuclei in digitized images obtained from fine needle aspirates (FNAs) of breast masses.

1. Data Source

The dataset originates from research conducted by the University of Wisconsin Hospitals, Madison, and Dr. William H. Wolberg, W. Nick Street, and Olvi L. Mangasarian. It was generously made available through the University of California, Irvine (UCI) Machine Learning Repository, a widely recognized and respected source for datasets used in research and education.

2. Data Description

1. **ID number**: Unique identifier assigned to each sample.
2. **Diagnosis**: Indicates the diagnosis of the breast mass, with 'M' representing malignant and 'B' representing benign.
3. **Ten real-valued features**: Computed for each cell nucleus, including radius, texture, perimeter, area, smoothness, compactness, concavity, concave points, symmetry, and fractal dimension.

   These features are measured as the mean, standard error (SE), and “worst” or largest (mean of the three largest values) across the cell nuclei, resulting in 30 features in total.

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_mean	compactness_mean	concavity_mean	points_mean	...
0	87139402	B	12.32	12.39	78.85	464.1	0.10280	0.06981	0.03987	0.03700	...
1	8910251	B	10.60	18.95	69.28	346.4	0.09688	0.11470	0.06387	0.02642	...
2	905520	B	11.04	16.83	70.92	373.2	0.10770	0.07804	0.03046	0.02480	...
3	868871	B	11.28	13.39	73.00	384.8	0.11640	0.11360	0.04635	0.04796	...
4	9012568	B	15.19	13.21	97.65	711.8	0.07963	0.06934	0.03393	0.02657	...
...	...	...	...	...	...	...	...	...	...	...	...
564	911320502	B	13.17	18.22	84.28	537.3	0.07466	0.05994	0.04859	0.02870	...
565	898677	B	10.26	14.71	66.20	321.6	0.09882	0.09159	0.03581	0.02037	...
566	873885	M	15.28	22.41	98.92	710.6	0.09057	0.10520	0.05375	0.03263	...
567	911201	B	14.53	13.98	93.86	644.2	0.10990	0.09242	0.06895	0.06495	...
568	9012795	M	21.37	15.10	141.30	1386.0	0.10010	0.15150	0.19320	0.12550	...

Table 7 Breast Cancer Wisconsin (Diagnostic) Dataset

II. Exploratory Data Analysis

EDA is primarily a strategy to see what the data can express, away from the formal modelling or hypothesis testing task, and assists in the analysis of data sets in order to describe their statistical properties, which include measures of central tendency (the mean, mode, and median), measures of spread (standard deviation and variance), and the shape of the distribution, and the presence of outliers Following the data collecting phase [19]. The Figure below represents the essential steps in the exploratory data analysis phase, which are explored in greater detail below:

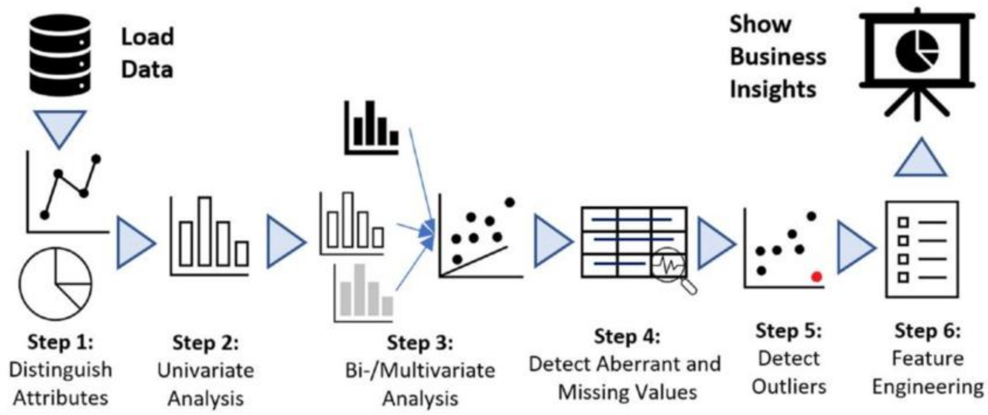


Figure 43 Exploratory Data Analysis Steps

1. Descriptive Statistics

In this section, I provide a summary of the fundamental statistical properties of the dataset, such as mean, median, mode, standard deviation, and variance for each feature. This gives a quick overview of the data's central tendency and variability.

Attribute	count	mean	std	min	25%	50%	75%	max
id	569	30371831.432	25020585.612	8670	869218	906024	8813129	911320502
radius_mean	569	14.1273	3.524	6.981	11.7	13.37	15.78	28.11
texture_mean	569	19.2896	4.301	9.71	16.17	18.84	21.8	39.28
perimeter_mean	569	91.969	24.299	43.79	75.17	86.24	104.1	188.5
area_mean	569	654.8891	351.9141	143.5	420.3	551.1	782.7	2501
smoothness_mean	569	0.0964	0.0141	0.0526	0.0864	0.0959	0.1053	0.1634
compactness_mean	569	0.1043	0.0528	0.0194	0.0649	0.0926	0.1304	0.3454
concavity_mean	569	0.0888	0.0797	0	0.0296	0.0615	0.1307	0.4268
points_mean	569	0.0489	0.0388	0	0.0203	0.0335	0.074	0.2012
symmetry_mean	569	0.1812	0.0274	0.106	0.1619	0.1792	0.1957	0.304
dimension_mean	569	0.0628	0.0071	0.05	0.0577	0.0615	0.0661	0.0974
radius_se	569	0.4052	0.2773	0.1115	0.2324	0.3242	0.4789	2.873
texture_se	569	1.2169	0.5516	0.3602	0.8339	1.108	1.474	4.885
perimeter_se	569	2.8661	2.0219	0.757	1.606	2.287	3.357	21.98
area_se	569	40.3371	45.491	6.802	17.85	24.53	45.19	542.2
smoothness_se	569	0.007	0.003	0.0017	0.0052	0.0064	0.0081	0.0311

Table 8 Basic Statistical Details

2. Visualizations

Visualizing data is an effective way to explore and understand the distribution and relationships among variables.

The figure below shows the distribution of key numerical features in the dataset. It helps in identifying the shape of the distribution (normal, skewed, etc.) and any potential outliers.

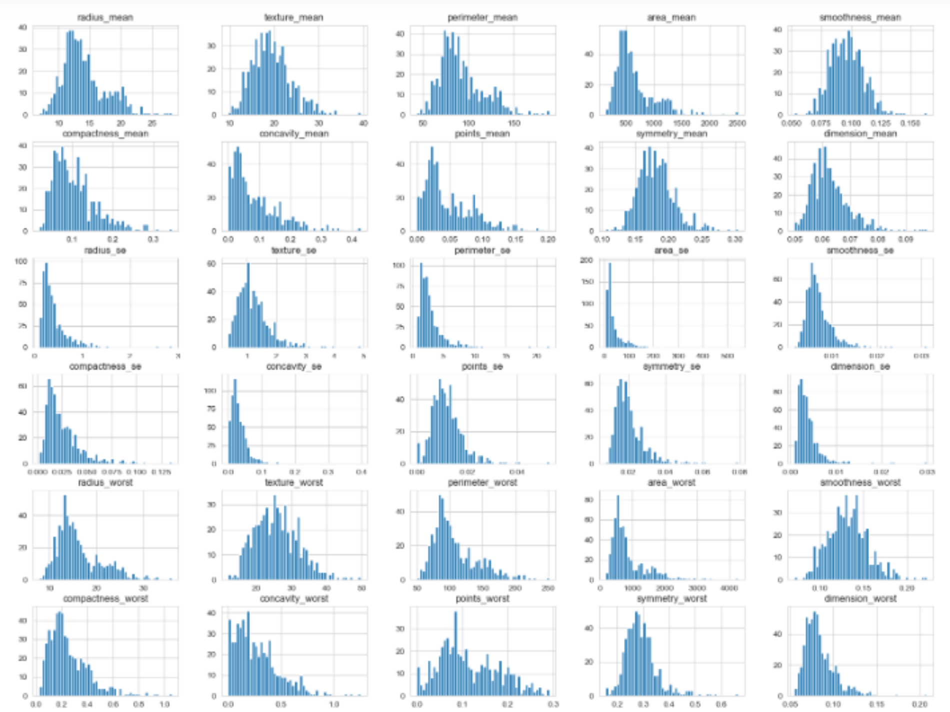


Figure 44 Distributions of Numerical Data

a. Diagnosis Distribution

This section analyzes the distribution of diagnoses within the Wisconsin Breast Cancer Dataset.

The bar chart below shows that malignant diagnoses are more frequent than benign diagnoses in the dataset.

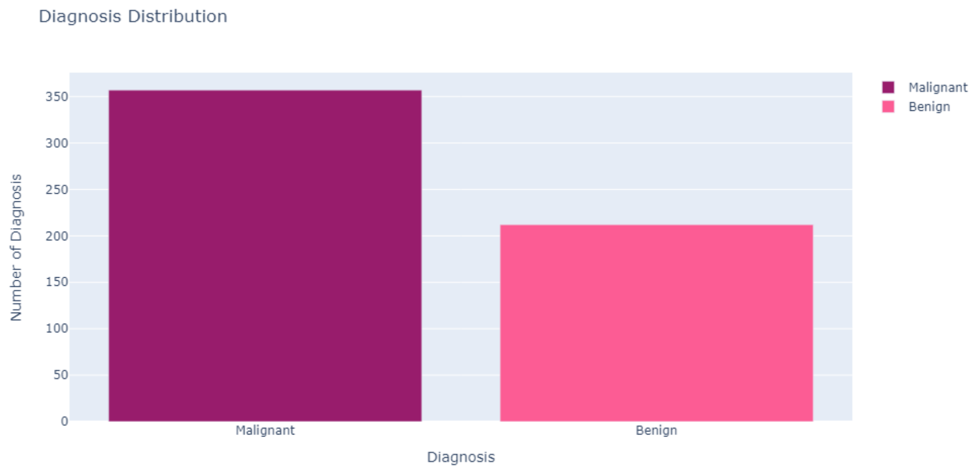


Figure 45 Diagnosis Distribution

The dataset consists of 37.3% benign and 62.7% malignant diagnoses according to the pie chart below. This indicates a slightly imbalanced distribution between the two classes.

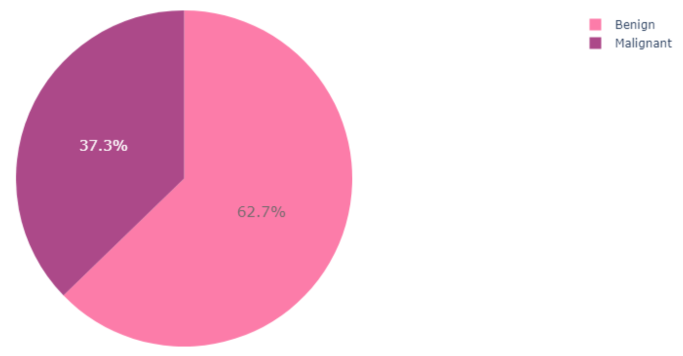


Figure 46 Pie Chart

A slightly imbalanced class distribution can be unfavorable for machine learning tasks. In this case, the model might be biased towards malignant cases due to their higher representation in the data. To address this bias, specific techniques such as data augmentation during model training might be necessary.

b. Violin Plot of Features by Diagnosis

Figure below shows the violin plot of features by diagnosis. In the texture_mean feature, the median of the Malignant and Benign classes is separated, which can be a good indicator for classification. However, in the dimension_mean feature, the median of the Malignant and Benign classes does not appear to be separated, which may not provide useful information for classification.

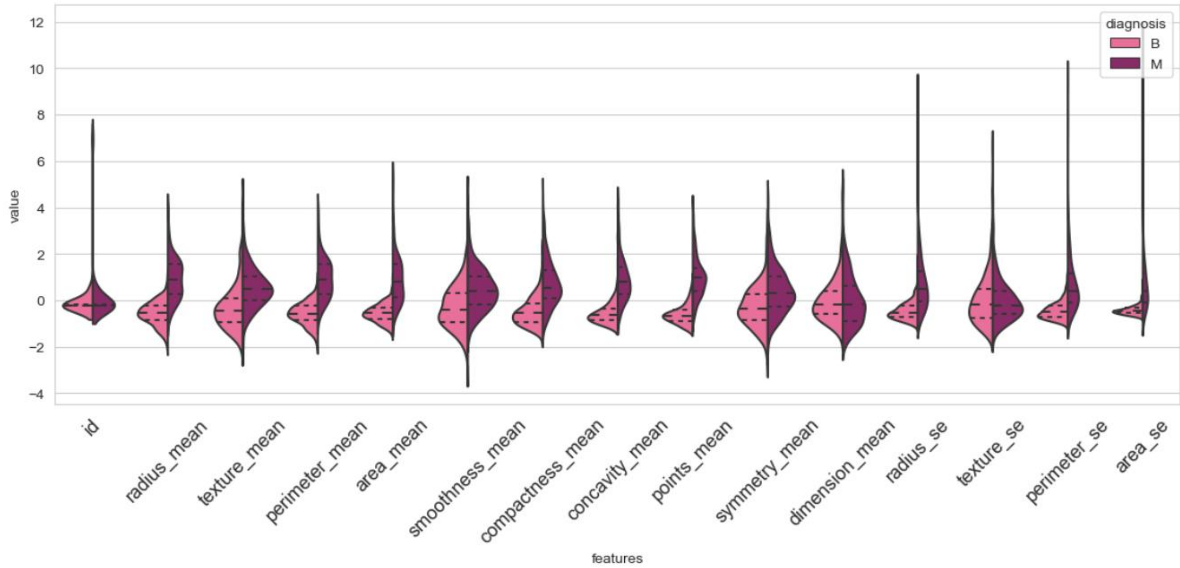


Figure 47 Violin Plot of Features by Diagnosis

c. Correlation Heatmap

The correlation heatmap, shown in Figure below, is used to understand the relationships between distinctive features and the target variable in the dataset. This heatmap is crucial for feature selection in machine learning models. The heatmap is organized in a grid format, where both rows and columns are labeled with feature names. The color intensity and numerical value within each cell represent the strength and direction of the correlation. Darker blue indicates stronger positive correlations, while darker orange indicates stronger negative correlations.

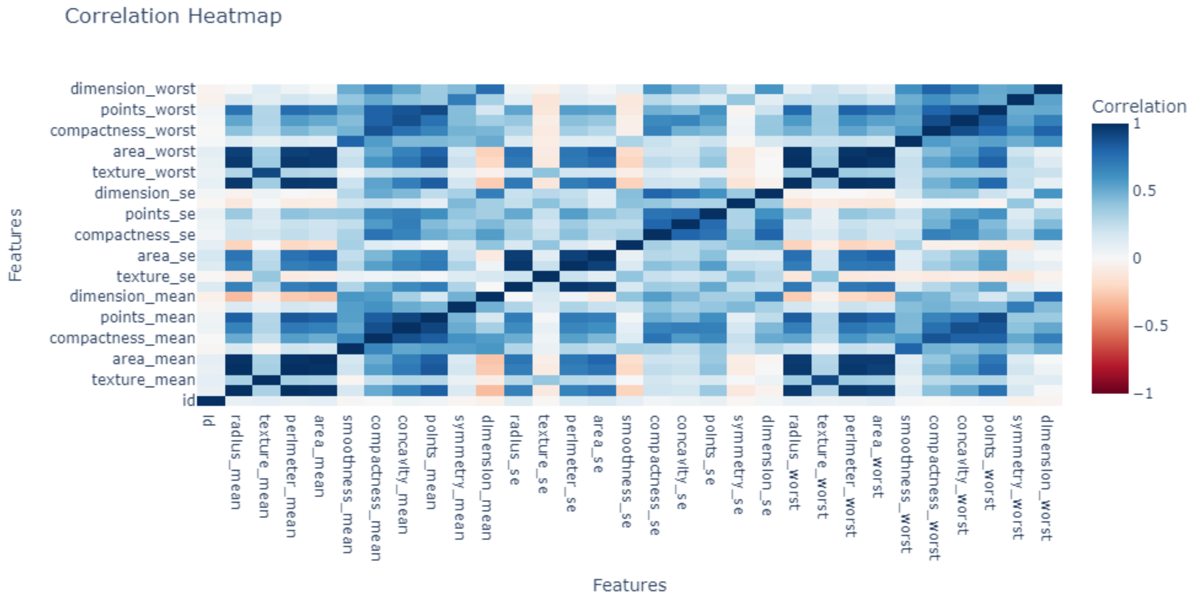


Figure 48 Correlation Heatmap

From the correlation heatmap, we observe the following:

- Radius, Area, and Perimeter are highly correlated (corr > 0.9), which is expected since Area and Perimeter are calculated using the Radius values.
- Texture_mean and texture_worst are highly correlated (corr = 0.98), with texture_worst being the largest value of all textures.
- Compactness_mean, concavity_mean, and concave_points_mean are highly correlated (corr = 0.7 to 0.9).
- Symmetry_mean and symmetry_worst are correlated (corr = 0.7).
- Dimension_mean and dimension_worst are correlated (corr = 0.77).

Based on these observations, we conclude that:

- The radius, parameter, and area are highly correlated, so we will use only one of them.
- Compactness_mean, concavity_mean, and concavepoint_mean are highly correlated, so we will use compactness_mean.
- The selected parameters for use are perimeter_mean, texture_mean, compactness_mean, and symmetry_mean.

d. Feature Pair

This section explores the correlation between features within the Wisconsin breast cancer dataset. Analyzing feature correlations can help identify redundant information and improve machine learning model performance.

The figures below show the feature pair, positively correlated features, and negatively correlated features, respectively. These plots help visualize the relationships between distinctive features.

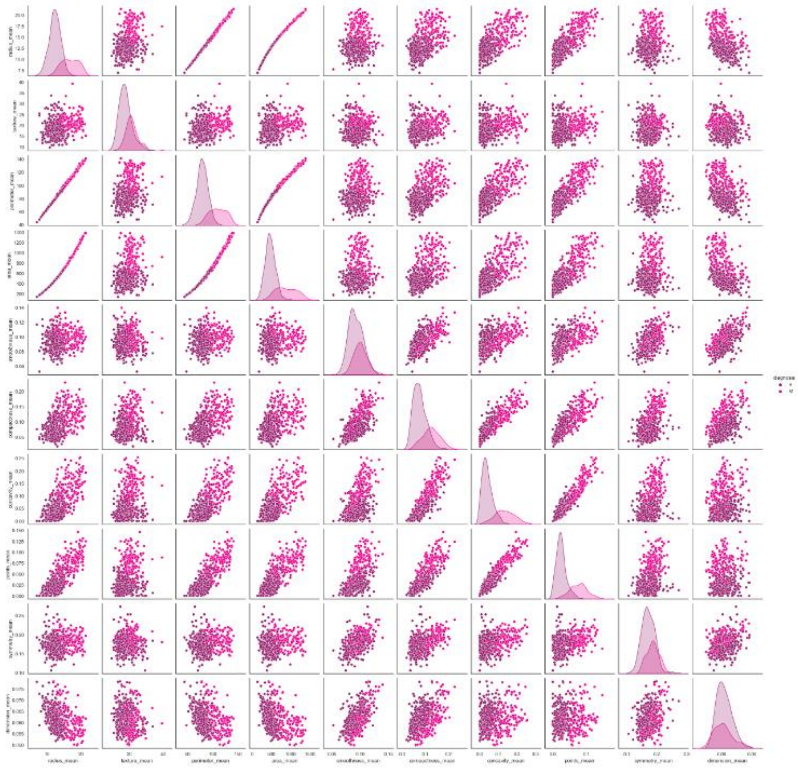


Figure 49 Feature Pair

e. Highly Correlated Features

In a scatter plot with a positive correlation, as the value of one feature increases, the value of the other feature also tends to increase. The data points will show a general upward trend.

Including highly correlated features in a machine learning model can be detrimental. It can lead to overfitting, where the model performs well on the training data but poorly on unseen data. Additionally, it can cause multicollinearity, which creates difficulties in interpreting the model's coefficients and reduces its overall performance.

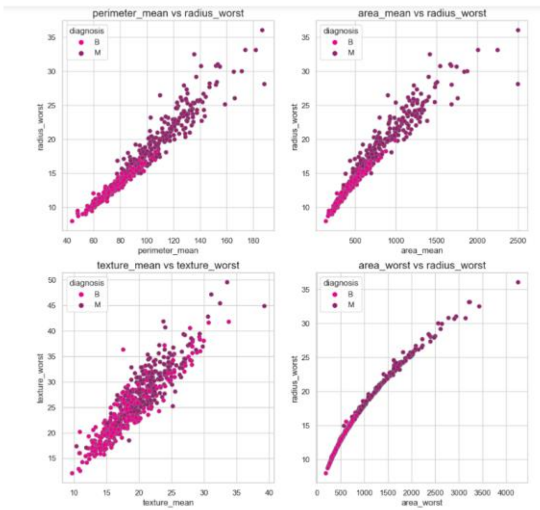


Figure 50 Positively Correlated Features

f. Low Correlated Features

In a scatter plot with a negative correlation, as the value of one feature increases, the value of the other feature tends to decrease. The data points will show a general downward trend.

In machine learning models, features with low correlation can be beneficial because they provide diverse information for the model to learn from. This can potentially improve the model's ability to distinguish between benign and malignant cases.

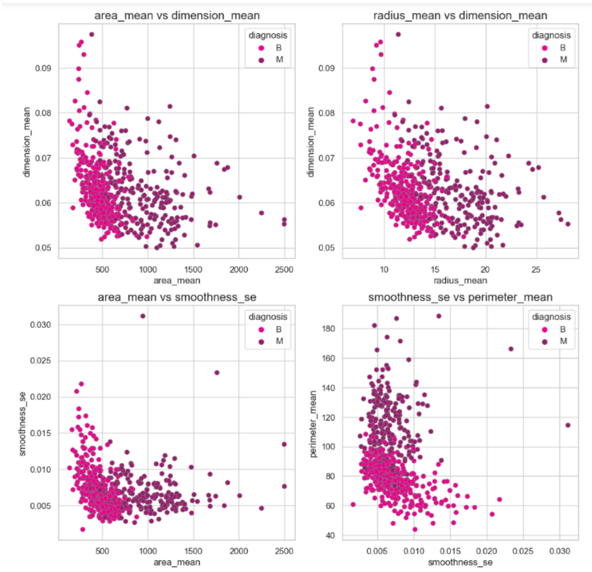


Figure 51 Negatively Correlated Features

g. Selected Features based on correlation.

The table below presents selected features from the Wisconsin breast cancer dataset chosen based on their correlation analysis. This analysis aims to identify features with informative relationships that can be beneficial for building machine learning models.

The table highlights several groups of correlated features. For example:

- **Concavity measures:** Features like compactness_mean, concavity_mean, and concave points_mean all show high correlation, suggesting they capture similar information. From this group, concavity_mean is likely chosen as a representative.
- **Shape and size features:** Features like area_worst, area_mean, and radius_se might capture related properties of cell size and shape. The table selects area_se, potentially due to its lower correlation with other features in the group.
- **Texture properties:** Texture_mean and texture_worst exhibit correlation, and the table selects texture_mean, possibly because it represents an average texture property.

The selection process aims to avoid redundancy and choose features that provide diverse information for machine learning models. This can help improve model performance and reduce the risk of multicollinearity.

Correlated attributes	Selected attribute
compactness_mean, concavity_mean, concave points_mean	concavity_mean
radius_se, perimeter_se, area_worst	area_se
compactness_worst, concavity_worst, concave points_worst	concavity_worst
compactness_se, concavity_se, concave points_se	concavity_se
texture_mean, texture_worst	texture_mean
area_worst, area_mean	area_mean

Table 9 Selected attributes based on correlation.

This process reduced the number of features from 30 to 16.

h. Outliers Detection

This section explores the presence of outliers in the Wisconsin breast cancer dataset. Outliers are data points that fall outside the typical range of the data and can potentially influence statistical analysis.

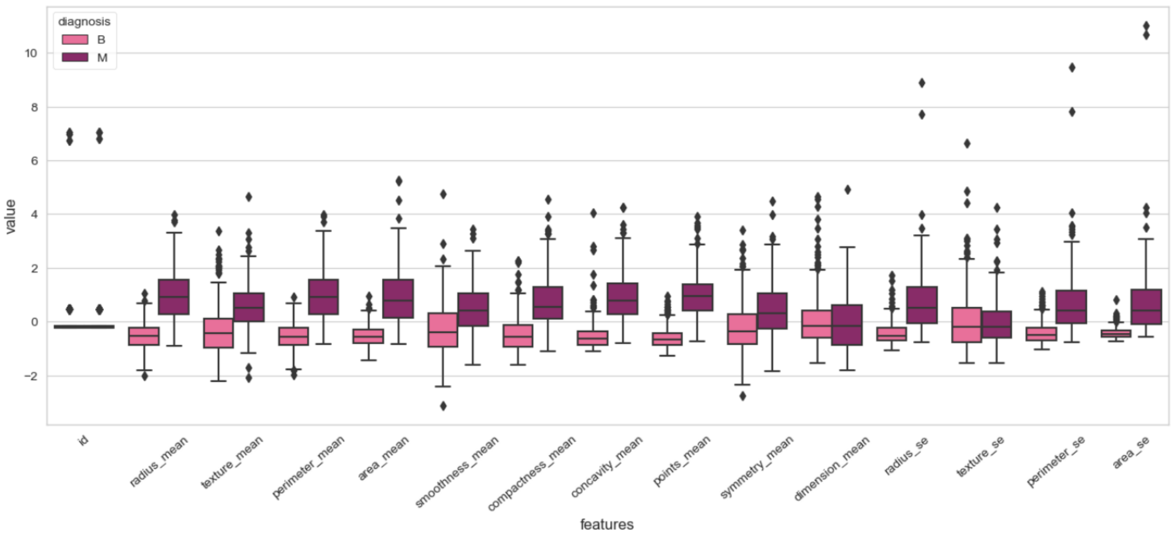


Figure 52 Box Plot

The box plot is a common tool for visualizing outliers. The box in the plot represents the interquartile range (IQR) of the data, which contains the middle 50% of the data points. The horizontal line within the box represents the median value. The whiskers extend from the box to show the remaining data points that fall within 1.5 times the IQR from the quartiles. Data points beyond the whiskers are potential outliers.

i. Dimensionality Reduction

Exploring high dimensional data by mapping it to a new two-dimensional plane is another important visualization technique used commonly in machine learning. The dimensionality reduction can be made in two separate ways:

1. Feature Selection: By only keeping the most relevant variables from the original dataset.
2. Dimension Reduction: By exploiting the redundancy of the input data and by finding a smaller set of new variables, each being a combination of the input variables, containing the

same information as the input variables. Principal Component Analysis and tSNE are two popular methods used based on this principle.

Principal Component Analysis (PCA)

Principal Component Analysis (PCA) technique was introduced by the mathematician Karl Pearson in 1901. It works on the condition that while the data in a higher dimensional space is mapped to data in a lower dimension space, the variance of the data in the lower dimensional space should be maximum.

- **Principal Component Analysis (PCA)** is a statistical procedure that uses an orthogonal transformation that converts a set of correlated variables to a set of uncorrelated variables. CA is the most widely used tool in exploratory data analysis and in machine learning for predictive models. Moreover,
- Principal Component Analysis (PCA) is an unsupervised learning algorithm technique used to examine the interrelations among a set of variables. It is also known as a general factor analysis where regression determines a line of best fit.
- The main goal of Principal Component Analysis (PCA) is to reduce the dimensionality of a dataset while preserving the most important patterns or relationships between the variables without any prior knowledge of the target variables.
- Principal Component Analysis (PCA) is used to reduce the dimensionality of a data set by finding a new set of variables, smaller than the original set of variables, retaining most of the sample's information, and useful for the regression and classification of data.

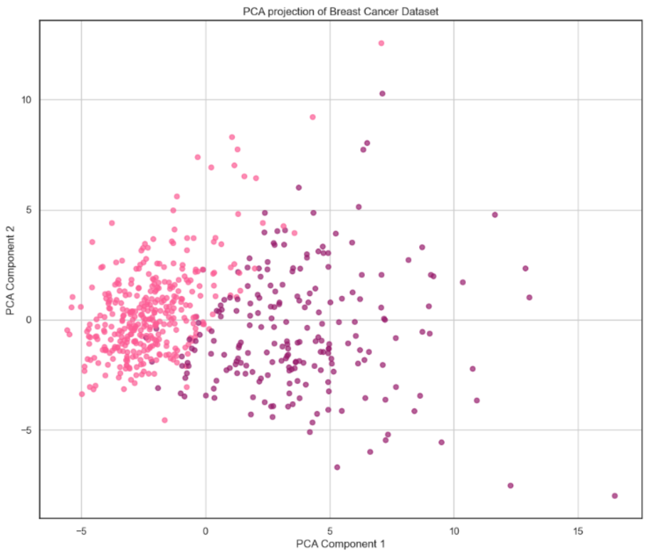


Figure 53 Principal Component Analysis (PCA)

The PCA projection shows some separation between the benign and malignant data points. This suggests that the principal components capture relevant information that can help differentiate between the two classes.

t-SNE (t-distributed Stochastic Neighbor Embedding)

T-distributed Stochastic Neighbor Embedding (t-SNE) is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions.

- Dimensionality Reduction represents n-dimensions data (multidimensional data with many features) in 2 or 3 dimensions. An example of dimensionality reduction can be discussed as a classification problem i.e. student will play football or not that relies on both temperature and humidity can be collapsed into just one underlying feature since both features are correlated to a high degree. Hence, we can reduce the number of features in such problems. A 3-D classification problem can be hard to visualize, whereas a 2-D one can be mapped to simple 2-dimensional space and a 1-D problem to a simple line.
- Even though PCA and t-SNE both are unsupervised algorithms that are used to reduce the dimensionality of the dataset. PCA is a deterministic algorithm to reduce the dimensionality of the algorithm and the t-SNE algorithm a randomized non-linear method to map the high dimensional data to the lower dimensional. The data that is obtained after reducing the dimensionality via the t-SNE algorithm is generally used for visualization purposes only.

Figure 54 t-SNE (t-distributed Stochastic Neighbor Embedding)

The t-SNE projection shows some visual separation between the benign and malignant data points. This suggests that the t-SNE embedding captures relevant relationships between the features that can help differentiate between the two classes.

III. Classical Approach

In this section, we will delve into the classical approach of machine learning classification, which forms the foundation for my comparative analysis with quantum hybrid approaches. The classical methodology encompasses several key steps, including feature set construction, model selection, evaluation metrics, and performance analysis. Let us explore each facet in detail.

1. Methodology

The primary objective of this classical approach is to identify the most effective and predictive algorithm for breast cancer detection. This will serve as a benchmark for comparison with the results obtained from the quantum hybrid VQC (Variational Quantum Circuit) models. To achieve this, I applied several machine learning classifiers to the Breast Cancer Wisconsin Diagnostic dataset. The classifiers used include Support Vector Machine (SVM), Random Forests, Logistic Regression, and K-Nearest Neighbors (KNN). The proposed architecture is detailed in the figure below.

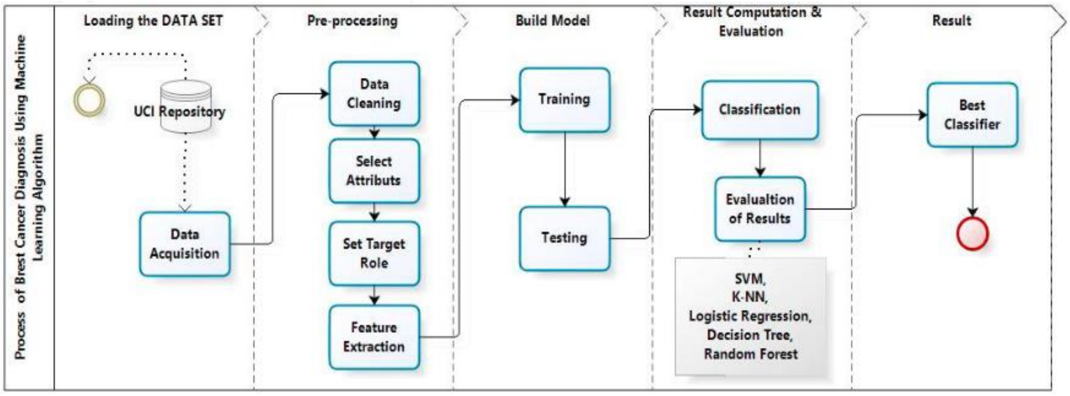


Figure 55 Process Flow Diagram

Preprocessing Steps

Preprocessing serves as a critical initial phase in any data analysis or modeling task. It involves cleaning and transforming raw data into a format suitable for analysis, thereby laying the groundwork for accurate insights and robust model performance. By addressing issues such as missing values, duplicates, and redundant features, preprocessing enhances the reliability and effectiveness of subsequent analytical processes.

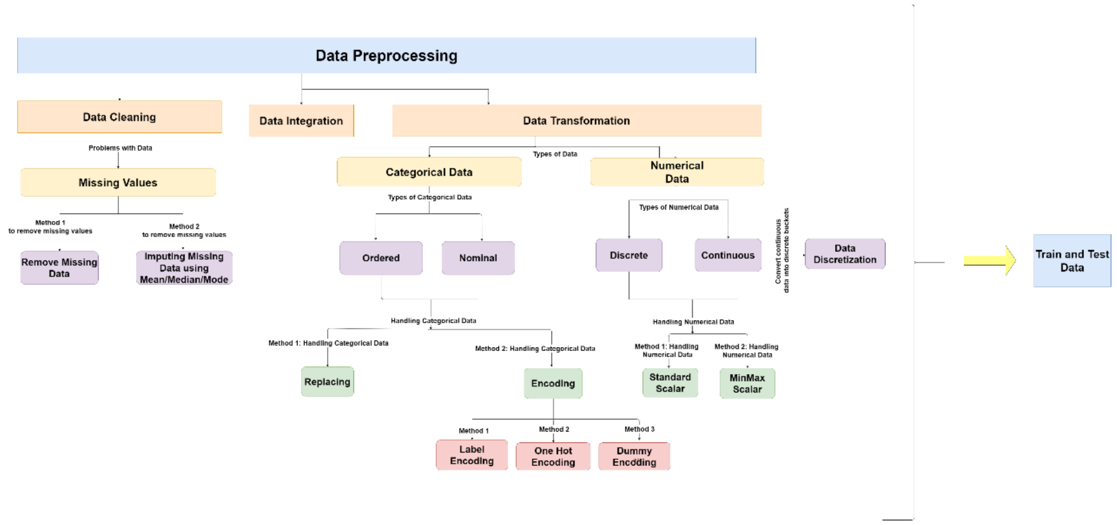


Figure 56 Pre-processing Steps

Data Cleaning

The first step of **Data Preprocessing** is **Data Cleaning.** Most of the data that we work today are not clean and requires substantial amount of **Data Cleaning**. Some have missing values, and some have junk data in them. If these missing values and inconsistencies are not managed properly then our model wouldn't give accurate results.

Missing Values

The initial step in preparing the dataset involved checking for missing or null values. Using the Pandas library in Python, the dataset was scanned for any missing values. Fortunately, upon inspection, it was found that the dataset contained no missing values. This ensured the integrity of the dataset, eliminating the need for imputation or removal of incomplete data points.

Check for null and missing values

```
null_values = df.isnull().values.any()
if null_values == True:
    print("There are some missing values in data")
else:
    print("There are no missing values in the dataset")
```

```
There are no missing values in the dataset
```

Figure 57 Handling Missing Values

Handling Duplicate Entries

To ensure data accuracy and avoid redundancy, the dataset was examined for any duplicate entries. By leveraging Pandas duplicated () function, it was determined that the dataset contained no duplicate elements. This step confirmed the uniqueness of each data point, preventing any potential biases that could arise from duplicate observations.

Check for duplicate elements

```
sum(df.duplicated())
```

```
0
```

Figure 58 Handling Duplicate Entries

Feature Scaling:

Feature scaling is applied to standardize the features to have a mean of 0 and a standard deviation of 1.

$$z = \frac{x - \mu}{\sigma}$$

Figure 59 Standardization.

This step is crucial for algorithms sensitive to feature scales, such as SVM, k-NN, and neural networks. Python sklearn library offers us with StandardScaler () function to standardize the data values into a standard format.

```
]:
X = df.drop('diagnosis',axis=1).values
y = df['diagnosis'].values

sc = StandardScaler()
X = sc.fit_transform(X)
```

Figure 60 Feature Scaling

Data Splitting

Splitting the dataset into training and testing sets is a crucial step in machine learning model development. This process ensures that the model's performance can be evaluated on unseen data, providing an estimate of its generalization ability.

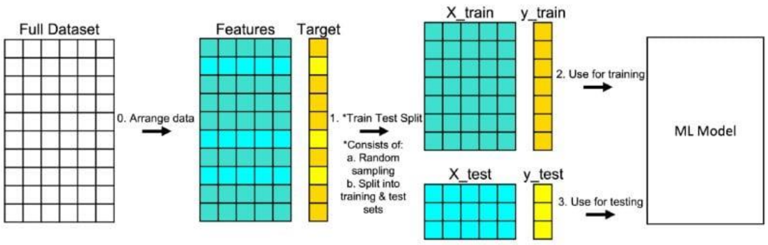


Figure 61 Data Splitting

In this experiment, a 60-20-20 split ratio is chosen, allocating 60% of the data to the training set and 20% to the validation set and 20% testing set. This ratio strikes a balance between having sufficient data for training the model and retaining an adequate amount for evaluating its performance.

Figure 62 Data splitting 60-20-20.

Setting Up the Environment for the Experiment:

To ensure a consistent and reproducible experimental environment, I have set up a virtual environment tailored for this project. This allows us to isolate the dependencies and configurations specific to the classical approach from the global environment. Below are the steps taken to set up the environment:

I have created a virtual environment using tools like virtualenv. This environment serves as a self-contained workspace for our project, preventing conflicts with other projects' dependencies.

```
PS C:\Users\HP\Downloads\PFE\Enhancing Breast Cancer Detection M
odels with Quantum Nodes in Neural Networks using Pennylane and
Python> python -m venv venv
```

Figure 63 Creating a virtual environment.

Once created, I activated the virtual environment. Activation ensures that any subsequent installations or executions are confined within this isolated environment.

```
PS C:\Users\HP\Downloads\PFE\Enhancing Breast Cancer Detection M
odels with Quantum Nodes in Neural Networks using Pennylane and
Python> venv\Scripts\activate
```

Figure 64 Activating the virtual environment.

Within the virtual environment, I installed the Jupyter kernel using the ipykernel package. This allows us to use Jupyter notebooks within the context of our project's environment.

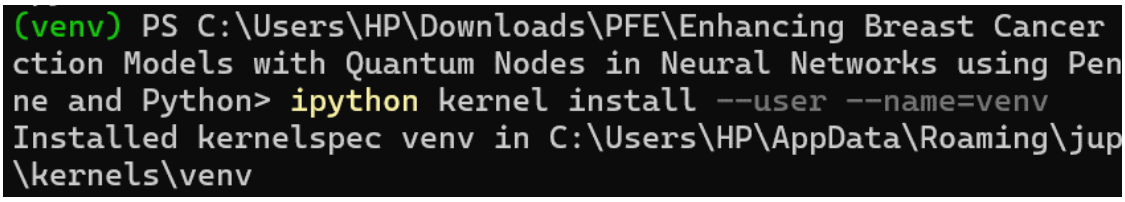


Figure 65 Installing Jupiter kernel for the virtual environment.

After installing the kernel, I selected it within the Jupyter notebook interface. This ensures that when running notebooks, they utilize the packages and configurations specific to the project's environment.

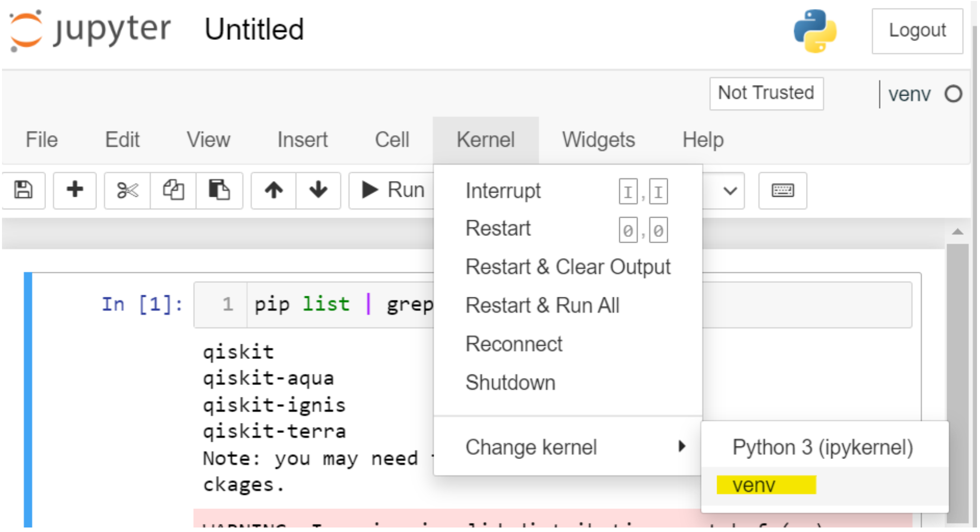


Figure 66 Selecting the installed kernel in Jupiter notebook in this virtual environment.

The table below provides an overview of the important packages and libraries utilized throughout the classical approach. These packages play a crucial role in data preprocessing, model training, evaluation, and visualization.

Package/Library	**Description**
pandas	Data manipulation and analysis
numpy	Numerical computing library
matplotlib	Data visualization library
seaborn	Statistical data visualization
scikit-learn	Machine learning library for classification, regression, clustering, etc.
plotly	Interactive visualization library for creating rich, interactive visualizations

Table 10 Important Packages/Libraries Used for the Classical Approach

1. Model Training and Selection

Machine learning is the most practical way of predicting breast cancer sickness. Reading the literature review, it becomes clear that the bulk of the work has been accomplished using machine learning and deep learning techniques. It is often understood that deep learning falls within the umbrella of machine learning. Five separate machine learning methods were used to this new dataset to find the most accurate method. Decision Tree, XGBoost, Logistic regression, Naive Bayes, and Random Forest are the categories used to organize these methods. In this section, I will get a brief overview of a few of these designs.

Random forest (RF)

Random Forest is a popular machine-learning technique used for both classification and regression tasks. It creates multiple decision trees from different parts of the training data. Each tree classifies the data separately and the final prediction is the sum of all the individual forecasts. This approach reduces the risk of overfitting, leading to more accurate and reliable predictions. Additionally, as the number of trees increases, the method becomes more robust to noise and outliers in the data. However, there is a trade-off between accuracy and computational efficiency, as training more trees requires more time and resources. These algorithms divide the data recursively. The Random Forest algorithm is provided below.

Node t, randomly selection v of the p independent variables.

∀of the k=1, …, vis the variable that was sampled; for each possible split of kth, determine the optimal split sk.

In s*, gain the optimal splitting k from k=1 to k=m; Determine the optimal split sk for splitting node t; jth variable is defined cut point cs* that is used for splitting node t.

The data is separated here, with the i=1, …, n; observation with xij<cs* going to the left descending node and all other observations going to the right descending node.

To grow a tree of maximum size Tb, simply repeat steps 1–4 for each node in the tree's descendent set.

Figure 67 Random Forest (RF) Algorithm

XGBoost (XGB)

XGBoost (Extreme Gradient Boosting) is a widely used method for building robust predictive models, especially in machine learning. It uses aggregated decision trees to make accurate predictions from complex datasets. Although decision trees are easy to interpret, XGBoost can be difficult to understand at first glance. However, data scientists and machine learning experts prefer XGBoost because it efficiently processes large datasets and quickly builds accurate models. This powerful and versatile tool strikes a balance between model complexity and prediction accuracy using gradient descent and regularization techniques. It is highly adaptable, making it valuable for extracting insights from complex datasets. However, careful optimization of the hyperparameters is necessary to achieve the best results. With the right approach, XGBoost enables data scientists to build accurate and reliable models that offer valuable insights into complex datasets. XGBoost algorithm can be written as

$$Obj = \sum_{i=1}^{m} l\left(y_i, \hat{y}_i\right) + \sum_{k=1}^{m} \Omega(f_k)$$

where f_k is the leaf node's regular term of the k[th] classification tree, $l(y_i, \hat{y}_i)$ is the training error of sample x_i, and Obj is the objective function.

Naive Bayes (NB)

Naive Bayes is a classification algorithm that assumes that features are conditionally independent, given class labels. Although this assumption is often violated in real databases, it is still useful in practical applications. Although not independent, classification can derive information from features. For high-dimensional data, is fast and efficient with less training data than complicated and complex models because it estimates the probability of each feature separately. It can handle both discrete and continuous data, making it versatile for different databases. In natural language processing, it is essential for text categorization and spam filtering. Overall, Naive Bayes is a useful machine learning tool for solving classification problems. Naive Bayes classifier can be written as –

$$f_c^{NB}(x) = \prod_{j=1}^{n} P(X_j = x_j | C = c) P(C = c)$$

Naive Bayes(x) is the probability that observation x is in class c, f_c is the class-c observations, and n is the number of observations. $P(X_j{=}x_j|C{=}c)P(C{=}c)$ is the conditional probability of seeing feature j in class c. Divide class c's observations by feature j's xj usage. Training data can be used to calculate P($C{=}c$). Multiplying feature conditional probabilities and prior probabilities classifies new data into the most likely class.

Logistic regression (LR)

Using labeled data, I train our model in supervised learning. Logistic regression is used for categorization problems in supervised learning. Logistic regression's discrete output variable (y) is usually 0 or 1. A sigmoid function simulates X's effect on the output variable. This function provides a probability between 0 and 1 indicating the input's likelihood of being positive (1). Finance, marketing, and healthcare use logistic regression. Based on medical history and demographic data, logistic regression can estimate patients' cancer risk. Logistic regression handles nonlinear input–output relationships. For massive datasets, it requires fewer computing resources. Logistic regression's simplicity and effectiveness make it a common classification problem solution. Logistic regression can also provide relative relevance information for feature selection and model interpretation. Logistic regression classifier can be written as

$$g(x) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \beta_0 x_1 + \cdots + \beta_m x_m$$

where $\pi(x)$ denotes the probability of a binary outcome (such as success or failure) given the values in the predictor vector x. The log odds are predicted as a linear combination of the predictor variables, and the coefficients β_0, β_1, …, β_m show the impacts of each predictor on the log chances. Exponentiating the equation allows one to determine the odds of a successful outcome given specific values of the predictors.

K-Nearest Neighbors (KNN)

K-Nearest Neighbors is a simple, yet powerful machine learning algorithm used for both classification and regression tasks. It operates on the principle that similar instances exist in proximity. KNN works by finding the K closest training examples in the feature space to the instance that needs to be classified or predicted.

In the classification context, the class label for a new instance is determined by the majority vote of its K nearest neighbors. For regression, the prediction is the average of the values of its K nearest neighbors. KNN is non-parametric, meaning it does not make any assumptions about the underlying data distribution. This flexibility makes KNN highly versatile and easy to implement. However, KNN can be computationally expensive, especially as the dataset grows since it requires calculating the distance between the instance and all training examples.

The algorithm is sensitive to the choice of K, the number of neighbors. A small K can lead to overfitting, while a large K might smooth out the predictions too much, leading to underfitting. Additionally, the performance of KNN can be affected by the choice of distance metric, with Euclidean distance being the most common.

Nearest-neighbor algorithm

a) A pseudo code for the nearest neighbor algorithm is

ALGORITHM *Nearest-neighbor*$(D[1..n,1..n], s)$

//Input: A $n \times n$ distance matrix $D[1..n,1..n]$ and an index s of the starting city.

//Output: A list Path of the vertices containing the tour is obtained.

for $i \leftarrow 1$ to n do Visited $[i] \leftarrow$ false

Initialize the list Path with s

Visited $[s] \leftarrow$ true

Current $\leftarrow s$

for $i \leftarrow 2$ to n do

  Find the lowest element in row current and unmarked column j containing the element.

  Current $\leftarrow j$

  Visited$[j] \leftarrow$ true

  Add j to the end of list Path

Add s to the end of list Path

return Path

Figure 68 KNN Algorithm

Support Vector Machine (SVM)

Support Vector Machine (SVM) is a supervised learning algorithm used for classification and regression tasks. It is particularly effective in high-dimensional spaces and for problems where the number of dimensions exceeds the number of samples. SVM works by finding the hyperplane that best separates the classes in the feature space. The optimal hyperplane is defined as the one that maximizes the margin, which is the distance between the hyperplane and the nearest data points from each class. These points are called support vectors.

For classification, SVM can handle both linear and non-linear data by using kernel functions. A linear SVM uses a straight hyperplane, while non-linear SVMs use kernel functions (such as polynomial, radial basis function, and sigmoid) to transform the data into a higher-dimensional space where it can be linearly separable. This flexibility allows SVMs to model complex relationships.

In the case of regression (known as Support Vector Regression, SVR), the algorithm tries to fit the best possible line within a predefined margin of tolerance. SVMs are robust to overfitting, especially in high-dimensional space, due to the regularization parameter (C), which controls the trade-off between achieving a low error on the training data and minimizing the model complexity.

The SVM algorithm can be described as follows:

1. Choose the type of SVM (linear or non-linear) and the kernel function (if non-linear).
2. Define the optimization problem to maximize the margin between the classes while minimizing classification error.
3. Solve the optimization problem to find the support vectors and the coefficients defining the hyperplane.
4. Use the hyperplane to classify new data points.

By balancing complexity and accuracy, SVMs provide a robust approach for various classification and regression problems, particularly when dealing with high-dimensional and non-linear data.

IV. Hybrid Approach: Quantum Variational Classifier (VQC)

1. Introduction to Quantum Computing Tools:

a. Qiskit Overview

Qiskit is an open-source software development kit (SDK) for working with quantum computers at the level of circuits, pulses, and algorithms. It provides tools for creating and manipulating quantum programs and running them on prototype quantum devices on IBM Quantum Platform or on simulators on a local computer. It follows the circuit model for universal quantum computation and can be used for any quantum hardware (currently supports superconducting qubits and trapped ions) that follows this model.

Qiskit was founded by *IBM Research* to allow software development for their cloud quantum computing service, IBM Quantum Experience. External supporters also make contributions, typically from academic institutions.

The primary version of Qiskit uses the Python programming language. Versions for Swift and JavaScript were initially explored, though the development for these versions has halted. Instead, a minimal re-implementation of basic features is available as MicroQiskit, which is made to be easy to port to alternative platforms.

69 Qiskit Logo

Developer(s)	**IBM Research, Qiskit community**
Initial release	March 7, 2017
Written in	Python
Operating system	Cross-platform
Type	SDK for Quantum Computing

Table 11 Qiskit details

In this project, the following versions of Qiskit components were used:

```
!pip list | grep qiskit
```

```
qiskit                        0.33.1
qiskit-aer                    0.10.3
qiskit-algorithms             0.3.0
qiskit-aqua                   0.9.5
qiskit-ibmq-provider          0.18.2
qiskit-ignis                  0.7.0
qiskit-machine-learning       0.3.1
qiskit-terra                  0.19.1
```

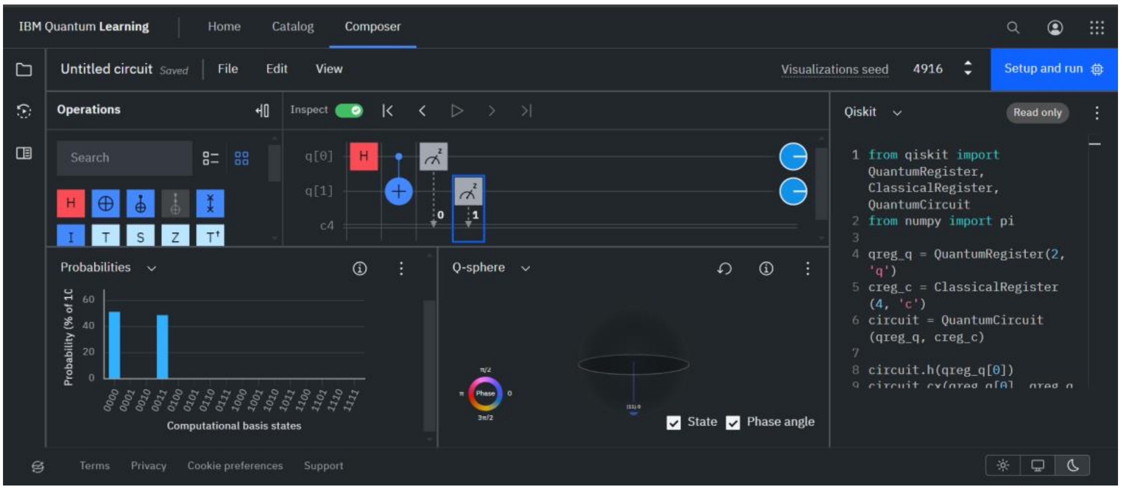
70 Qiskit Versions

Component	Explanation
qiskit	The main Qiskit package which serves as the umbrella package that includes the core functionalities.
qiskit-aer	Provides high-performance simulators for quantum circuits.
qiskit-algorithms	Contains algorithms for quantum computing including optimization, chemistry, and machine learning algorithms.
qiskit-aqua	Provides algorithms for quantum computing applications in chemistry, AI, and optimization. (Now mostly integrated into qiskit-algorithms.)
qiskit-ibmq-provider	Allows access to IBM Quantum devices and simulators.
qiskit-ignis	Provides tools for quantum error correction and mitigation.
qiskit-machine-learning	Contains quantum machine learning algorithms and tools.
qiskit-terra	The foundation upon which the rest of Qiskit is built. It includes the necessary tools for creating and manipulating quantum circuits, compiling, and running them on quantum hardware or simulators.

Table 12 Qiskit Components

b. IBM Quantum Composer

The Quantum Composer is a graphic user interface (GUI) designed by IBM to allow users to construct various quantum algorithms or run other quantum experiments. Users may see the results of their quantum algorithms by either running it on a real quantum processor or by using a simulator. Algorithms developed in the Quantum Composer are referred to as a "quantum score", in reference to the Quantum Composer resembling a musical sheet.

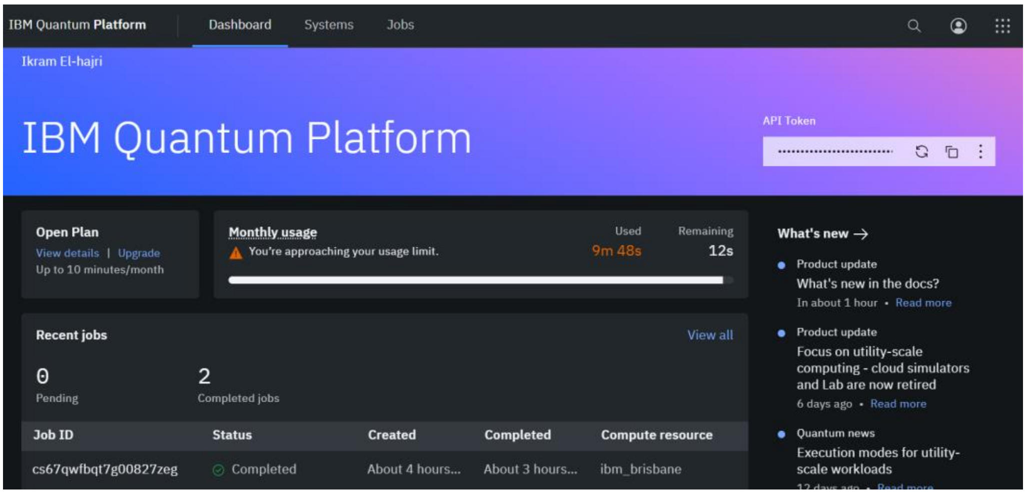


71 IBM Quantum Composer

c. IBM Quantum API

The IBM Quantum API is a fundamental component of the Qiskit ecosystem, providing access to IBM's quantum devices and simulators through a cloud-based platform. Through the IBM Quantum API, users can interact with quantum hardware and execute quantum circuits remotely, without the need for specialized quantum hardware on their local machines.

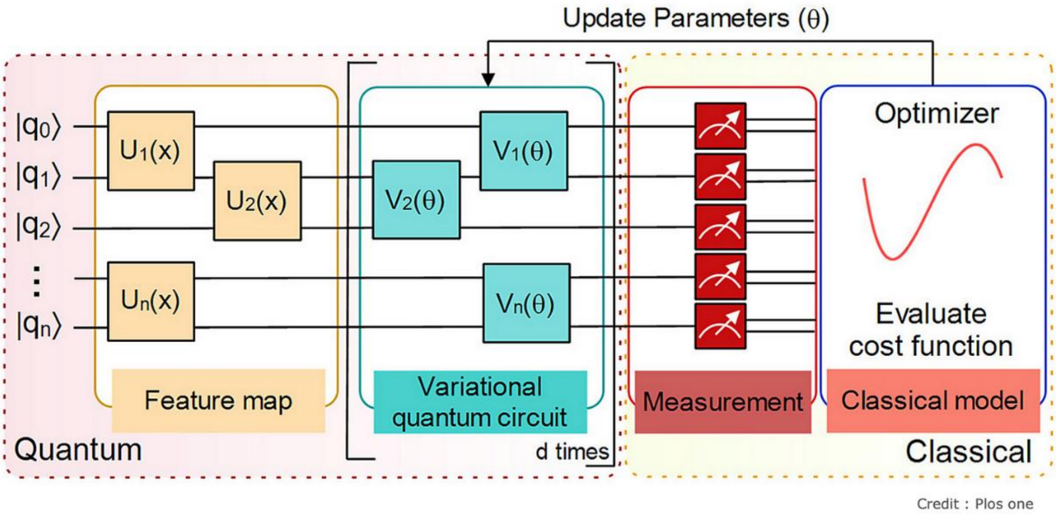
In this project, the IBM Quantum API was utilized to access IBM's quantum devices for executing quantum circuits, as well as simulators for testing and debugging purposes. Authentication credentials were obtained from the IBM Quantum platform.



72 IBM Quantum Platform

2. Variational Quantum Classifier (VQC)

Variational quantum classifiers (VQCs) are a type of hybrid quantum machine learning algorithm that can be used to solve a wide variety of classification problems. VQCs combine the power of quantum computing with the flexibility of classical machine learning algorithms to achieve state-of-the-art performance on many tasks.



73 Variational Quantum Classifier (VQC)

3. VQC Modeling

In general, the VQC architecture consists of three basic components, namely *quantum feature mapping*, *ansatz*, and *measurement* [20],as shown in the figure below.

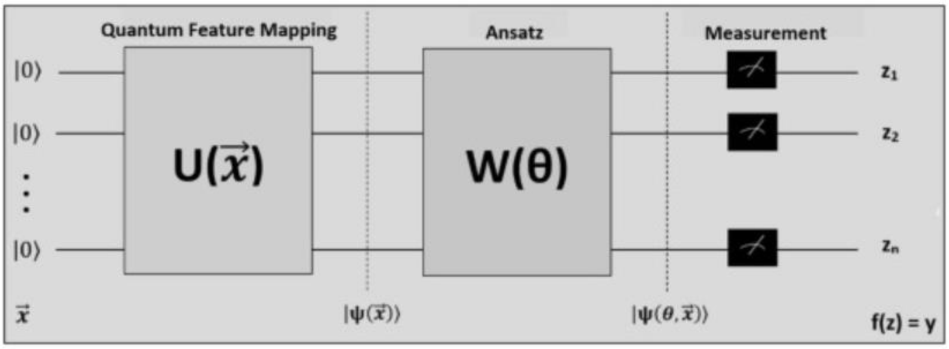


Figure 74 The VQC architecture is composed of quantum feature mapping, ansatz, and measurement.

Quantum feature mapping serves to transform classical input attributes into quantum states, while ansatz extends these quantum states utilizing adaptable parameters. Measurements are then conducted to acquire results, followed by post-processing through decoding to generate the ultimate output. In the initial stage of VQC, the process commences with quantum feature mapping $U(\vec{x})$, which encodes the classical feature vector $\vec{x}$ into a quantum state $|\psi(\vec{x}) >$ within the Hilbert space [21], [22]. According to the equation below.

$$\mathrm{U}\left(\vec{x}\right)|0\rangle = \left|\psi\left(\vec{x}\right)\right\rangle$$

this quantum feature mapping is specifically applied to the ground state **|0>** for every qubit. The second component, denoted as the ansatz $\boldsymbol{W(\theta)}$, dynamically evolves the quantum states $\psi\theta$,x→ within the system, as elucidated in the equation below.

$$\left|\psi\left(\theta, \vec{x}\right)\right\rangle = \mathrm{W}(\theta)\mathrm{U}\left(\vec{x}\right)|0\rangle$$

This form of parameterized circuit construction is notably suitable for the particular problem at hand due to its circuit depth scaling linearly with the number of qubits.[23], [24], [25] In the third stage, the process of measuring the quantum state $\boldsymbol{z}$ is executed, marking the conclusive phase in the quantum sequence. Subsequently, a classical computer engages in further post-processing on this outcome by decoding and retrieving the output $\boldsymbol{y}$, which is a function of $\boldsymbol{f(z)}$. Serving as the ultimate step in the VQC framework, this decoding procedure concludes the regression task and generates the final output of the VQC.

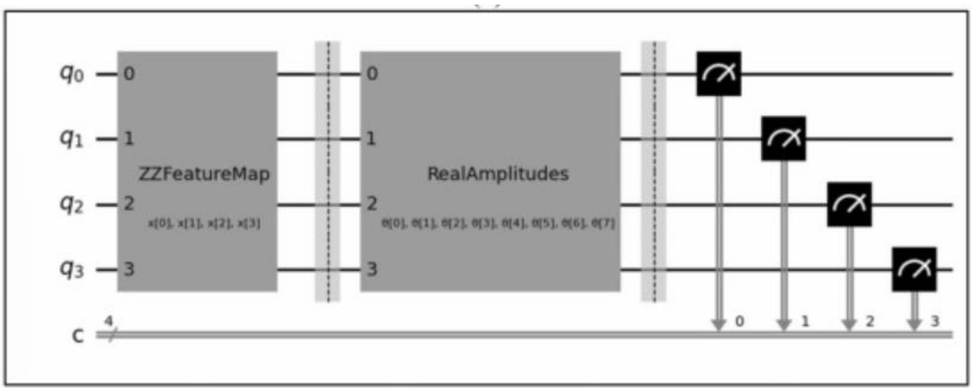


Figure 75 The proposed VQC component consists of ZZFeatureMap as quantum feature mapping, RealAmplitudes as ansatz, and measurements.

4. Data Encoding for Quantum Circuits:

a. Feature Map Encoding

Feature map encoding is a crucial step in preparing classical data for processing by quantum circuits in the VQC. It involves mapping classical feature vectors to quantum states, often through techniques such as amplitude encoding or angle encoding. This process transforms classical data into quantum states, enabling quantum circuits to operate on the encoded data.

In this implementation, the ZZFeatureMap is utilized, which consists of layers of Hadamard gates followed by parameterized Pauli-Z rotations.

```
from qiskit.circuit.library import ZZFeatureMap
feature_map = ZZFeatureMap(feature_dimension=num_features, reps=1)
```

Figure 76 ZZFeatureMap.

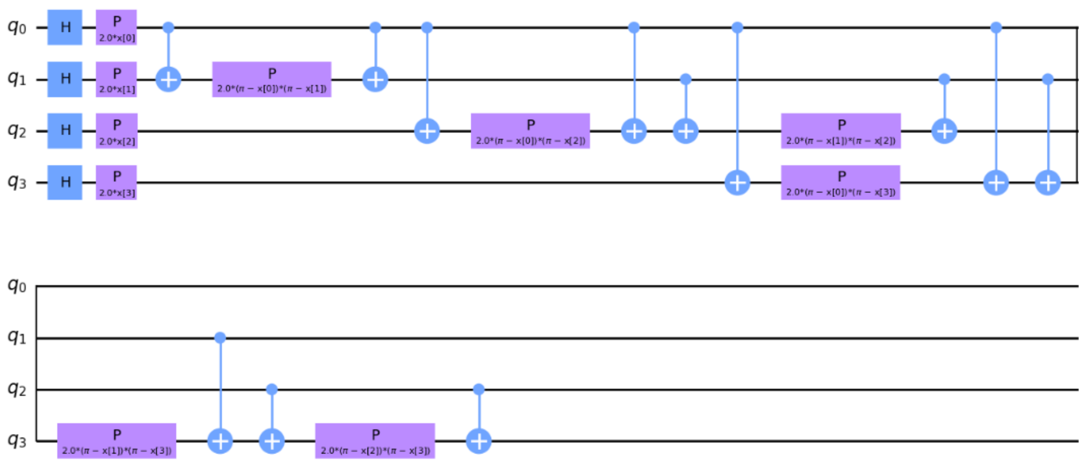


Figure 77 ZZ Feature Map Circuit

Each qubit in the circuit represents a feature from the input data, and the rotations introduce entanglement between qubits based on pairwise interactions among features.

b. Variational Form/ Ansatz Encoding

Variational form encoding / Ansatz defines the structure of the quantum circuit used in the VQC. It represents a parametrized quantum circuit with tunable parameters, often referred to as variational parameters. These parameters are optimized during the training process to minimize classification errors. Variational form encoding allows for flexibility in the representation of quantum circuits, enabling adaptation to different classification tasks and data distributions.

```python
from qiskit.circuit.library import RealAmplitudes
```

```python
ansatz = RealAmplitudes(num_qubits=num_features, reps=3)
```

Figure 78 RealAmplitudes.

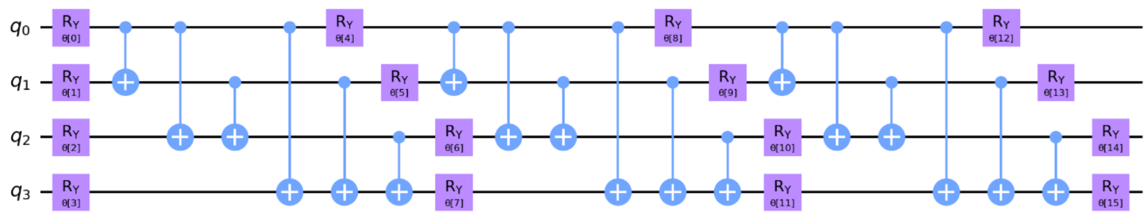


Figure 79 RealAmplitudes Circuit

In this implementation, the RealAmplitudes circuit is employed, which consists of layers of single-qubit rotations and entangling gates (typically CX gates).

5. Training Process:

a. Quantum Circuit Construction

In the training process of the VQC, quantum circuits are constructed based on the encoded data and the variational form. The feature map encoding transforms classical input data into quantum states, which are then processed by the variational form. The structure of the quantum circuit, including the arrangement of qubits and gates, is determined by the chosen variational form.

b. Quantum Circuit Execution:

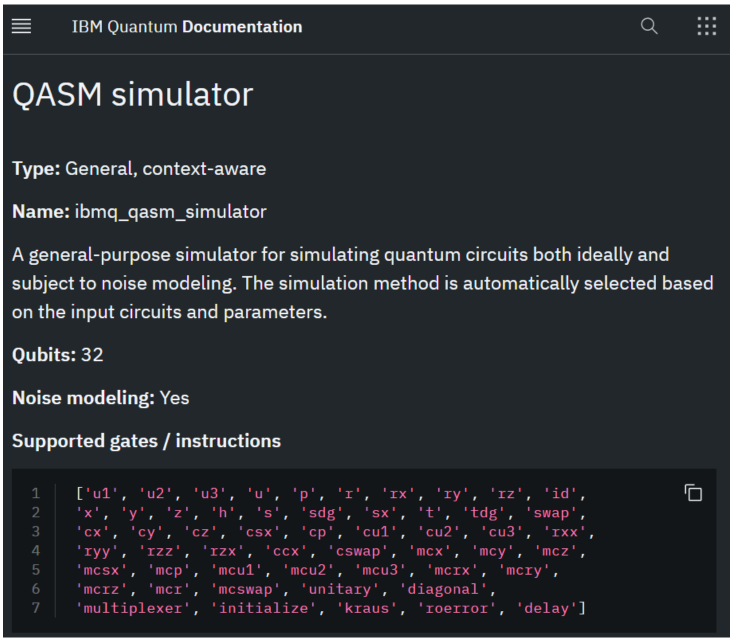
Once the quantum circuits are constructed, they are executed on quantum hardware or simulators to obtain measurement outcomes. These outcomes are then used to compute the cost function, typically based on the discrepancy between the predicted and actual labels. Quantum circuit execution can be performed iteratively as part of an optimization process, where variational parameters are adjusted to minimize the cost function using classical optimization algorithms.

6. Training Approach:

In this study, the training of the Variational Quantum Classifier (VQC) was conducted using two distinct approaches: one utilizing the QASM Simulator and the other leveraging IBM Quantum Devices.

a. Training on QASM Simulator:

The QASM Simulator is a cloud-based quantum simulator provided by IBM Quantum, which emulates the behavior of quantum computers.



IBM Quantum **Documentation**

QASM simulator

Type: General, context-aware

Name: ibmq_qasm_simulator

A general-purpose simulator for simulating quantum circuits both ideally and subject to noise modeling. The simulation method is automatically selected based on the input circuits and parameters.

Qubits: 32

Noise modeling: Yes

Supported gates / instructions

```
['u1', 'u2', 'u3', 'u', 'p', 'r', 'rx', 'ry', 'rz', 'id',
'x', 'y', 'z', 'h', 's', 'sdg', 'sx', 't', 'tdg', 'swap',
'cx', 'cy', 'cz', 'csx', 'cp', 'cu1', 'cu2', 'cu3', 'rxx',
'ryy', 'rzz', 'rzx', 'ccx', 'cswap', 'mcx', 'mcy', 'mcz',
'mcsx', 'mcp', 'mcu1', 'mcu2', 'mcu3', 'mcrx', 'mcry',
'mcrz', 'mcr', 'mcswap', 'unitary', 'diagonal',
'multiplexer', 'initialize', 'kraus', 'roerror', 'delay']
```

Figure 80 QASM Simulator

```python
from qiskit import BasicAer, execute
from qiskit.algorithms.optimizers import SPSA
from qiskit_machine_learning.algorithms.classifiers import VQC
# Patch the OneHotEncoder to use sparse_output
original_init = OneHotEncoder.__init__

def patched_init(self, *args, **kwargs):
    if 'sparse' in kwargs:
        kwargs['sparse_output'] = kwargs.pop('sparse')
    original_init(self, *args, **kwargs)

OneHotEncoder.__init__ = patched_init
vqc = VQC( feature_map=feature_map,
          ansatz=ansatz,
          loss='cross_entropy',
          optimizer=SPSA(),
          initial_point=initial_point,
          quantum_instance=BasicAer.get_backend( 'qasm_simulator')
```

Figure 81 Creating a VQC object for the QASM Simulator

b. Training on Quantum Computer (IBM Quantum Devices):

The second approach involved using actual IBM Quantum Devices. IBM Quantum Devices represent physical quantum processors hosted by IBM Quantum, providing access to real quantum hardware for experimentation.

To use the actual IBM quantum computer as a backend, I had to set up the [IBM Quantum Experience] account. After that, I could copy the authentication token I need for the initial connection. The info will be saved on the computer the first time I run this code.

```
from qiskit import IBMQ
IBMQ.save_account('91d1f02dc4cb016a2a65d9c3736fbc77803910675f57980

IBMQ.load_account()
```

Figure 82 Setting up the IBM Credentials

```
# Get the provider
provider = IBMQ.get_provider(hub='ibm-q', group='open', project='m

# List all available backends
available_backends = provider.backends()
print("Available backends:", available_backends)

```

```
Available backends: [<IBMQBackend('ibm_sherbrooke') from IBMQ(hub='ib
m-q', group='open', project='main')>, <IBMQBackend('ibm_brisbane') fr
om IBMQ(hub='ibm-q', group='open', project='main')>, <IBMQBackend('ib
m_kyoto') from IBMQ(hub='ibm-q', group='open', project='main')>, <IBM
QBackend('ibm_osaka') from IBMQ(hub='ibm-q', group='open', project='m
ain')>]
```

Figure 83 Getting the provider.

For this experience, the *IBM_Brisbane* backend was used.

```
provider = IBMQ.get_provider(hub='ibm-q', group='open', project= 'main' )
q_computer = provider.get_backend( 'ibm_brisbane')
```

Figure 84 settigng up the IBM quantum computer as a backend for our instance.

The *ibm_brisbane* backend is a quantum processor provided by IBM Quantum as part of their cloud-based quantum computing services.

It is one of the actual quantum devices available through the IBM Quantum Experience platform. Users can submit quantum circuits to be executed on this hardware, allowing them to perform real quantum computations and experiments. This backend is characterized by its specific qubit count, connectivity, and noise properties, which are crucial for running and optimizing quantum algorithms effectively.

```
from qiskit import BasicAer, execute
from qiskit.algorithms.optimizers import SPSA
from qiskit_machine_learning.algorithms.classifiers import VQC
vqc17 = VQC( feature_map=feature_map,
          ansatz=ansatz,
          loss= 'cross_entropy',
          optimizer=SPSA(),
          initial_point=initial_point,
          quantum_instance=q_computer)
```

Figure 85 Creating a VQC object for a quantum computer.

7. Optimization Process:

In this study, the optimization process for training the Variational Quantum Classifier (VQC) involved the use of **Simultaneous Perturbation Stochastic Approximation (SPSA)** algorithm. SPSA is a gradient-free optimization technique that is particularly well-suited for optimizing black-box functions, where the gradients of the objective function are unknown or expensive to compute. It is commonly used in scenarios where traditional gradient-based optimization methods are impractical or infeasible.

Note

SPSA can be used in the presence of noise, and it is therefore indicated in situations involving measurement uncertainty on a quantum computation when finding a minimum. If you are executing a variational algorithm using a Quantum ASseMbly Language (QASM) simulator or a real device, SPSA would be the most recommended choice among the optimizers provided here.

Figure 86 SPSA note from IBM Documentation

How SPSA Works:

1. **Initialization:** SPSA begins by initializing the parameter vector to an initial guess.
2. **Iteration:** At each iteration, SPSA evaluates the objective function at two perturbed points along each parameter dimension. These perturbations are chosen randomly according to a specific distribution, often based on a Rademacher distribution.
3. **Gradient Estimation:** Using the function evaluations at the perturbed points, SPSA estimates the gradient of the objective function with respect to each parameter dimension.
4. **Parameter Update:** SPSA updates the parameter vector based on the estimated gradient and a step size determined by the optimization schedule. The step size may decrease over time to improve convergence.
5. **Convergence:** The optimization process continues iteratively until a termination criterion is met, such as reaching a maximum number of iterations or achieving a satisfactory level of convergence.

Conclusion

Chapter 4 tackles the design and methodology employed in this study, which includes the use of classical machine learning techniques and a hybrid quantum-classical method using Variational Quantum Classifiers. The Breast Cancer (Wisconsin) Diagnosis dataset is utilized for the analysis, and exploratory data analysis, feature selection, dimensionality reduction, and model training are conducted. The classical approach involves Support Vector Machine (SVM), Random Forests, Logistic Regression, and K-Nearest Neighbors (KNN) classifiers, while the hybrid approach utilizes quantum computing tools like Qiskit and the IBM Quantum platform. The training process for the Variational Quantum Classifier (VQC) involves constructing quantum circuits, executing them on quantum hardware or simulators, and optimizing variational parameters using classical optimization algorithms like Simultaneous Perturbation Stochastic Approximation (SPSA). The goal is to classify benign and malignant breast masses using these methods.

Chapter 5: Results and Analysis

Introduction

This chapter presents the results obtained from the implementation of classical machine learning algorithms followed by the integration of Quantum Variational Classifier (VQC). This section delves into the performance evaluation, comparative analysis, and interpretation of results, aiming to assess the efficacy of quantum integration in enhancing breast cancer detection models.

1. Evaluation Metrics

Before delving into the results, it is imperative to introduce the evaluation metrics employed to assess the efficacy and performance of the models. These metrics serve as quantitative measures to gauge the models' accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC). A brief overview of each metric's significance and interpretation is provided to facilitate comprehension of the subsequent analyses.

1. Accuracy:

Accuracy is the ratio of correctly predicted instances to the total instances. It is a measure of the overall effectiveness of the model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

- **TP (True Positive):** Correctly predicted positive instances.
- **TN (True Negative):** Correctly predicted negative instances.
- **FP (False Positive):** Incorrectly predicted as positive instances.
- **FN (False Negative):** Incorrectly predicted as negative instances.

Accuracy indicates how often the model makes the correct prediction. It is a useful metric when the classes are balanced but can be misleading if the classes are imbalanced.

2. Precision:

Precision, also known as Positive Predictive Value, is the ratio of true positive predictions to the total predicted positives.

$$\text{Precision} = \frac{TP}{TP + FP}$$

Precision measures the accuracy of positive predictions. It is particularly important in scenarios where the cost of false positives is high.

3. Recall (Sensitivity or True Positive Rate):

Recall is the ratio of true positive predictions to the total actual positives. It is also known as Sensitivity or True Positive Rate.

$$\text{Recall} = \frac{TP}{TP + FN}$$

Recall measures the model's ability to identify all relevant instances. It is crucial in situations where the cost of false negatives is high.

4. F1 Score:

The F1 Score is the harmonic means of precision and recall. It provides a single metric that balances both concerns.

$$F_1 = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$$

The F1 Score is useful when you need a balance between precision and recall, especially in the presence of imbalanced datasets.

5. Confusion Matrix:

A confusion matrix is a table used to describe the performance of a classification model. It shows the actual versus predicted classifications and includes true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

		Predicted	
		Negative (**N**) **-**	Positive (**P**) **+**
Actual	Negative **-**	True Negatives (**TN**)	False Positives (**FP**) **Type I error**
	Positive **+**	False Negatives (**FN**) **Type II error**	True Positives (**TP**)

Figure 87 Confusion Matrix

The confusion matrix provides a comprehensive breakdown of model performance and helps in understanding the types of errors the model is making.

6. AUC-ROC (Area Under the ROC Curve):

The AUC-ROC is the area under the ROC curve. It provides a single value that summarizes the performance of the classifier across all thresholds.

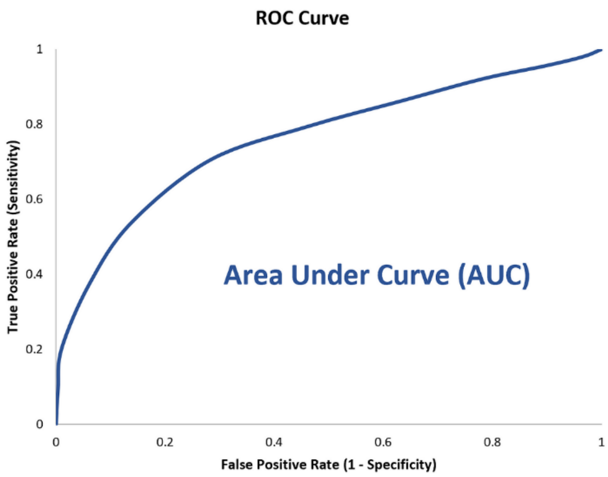


Figure 88 AUC-ROC Curve

Axes :

- **X-axis:** False Positive Rate (FPR)
- **Y-axis:** True Positive Rate (TPR)

AUC-ROC ranges from 0 to 1, with 1 representing a perfect classifier and 0.5 indicating a performance no better than random chance. It is a robust metric that considers both the sensitivity and specificity of the model.

2. Performance of Classical Machine Learning Classifiers

1. Support Vector Machine (SVM):

a. Linear Kernel:

- Accuracy: Achieved an accuracy of 95% on the test data.
- Precision: Achieved a precision of 97% for malignant (1) class and 94% for benign (0) class.
- Recall: Malignant class had a recall of 81%, indicating the model correctly identified 81% of actual malignant cases.
- F1-score: Obtained an F1-score of 88% for malignant class and 96% for benign class.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.94      0.99      0.96       110
           1       0.97      0.81      0.88        36

    accuracy                           0.95       146
   macro avg       0.95      0.90      0.92       146
weighted avg       0.95      0.95      0.94       146
```

Figure 89 SVM On Linear Kernel Classification Report

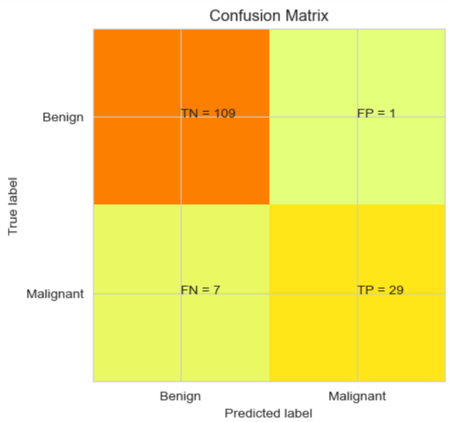


Figure 90 SVM On Linear Kernel Confusion Matrix

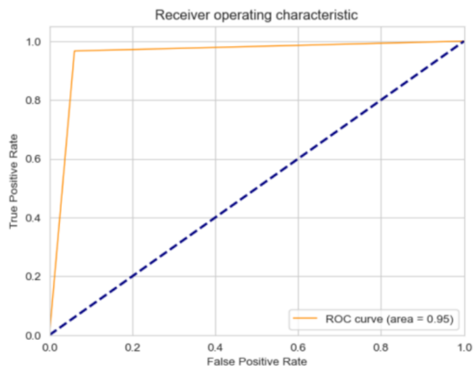


Figure 91 SVM on Linear Kernel Receiver Operating Characteristics

b. Radial Basis Function (RBF) Kernel:

- Accuracy: Maintained a similar accuracy of 95% on the test data.
- Precision: Achieved 100% precision for malignant class and 94% precision for benign class.
- Recall: Malignant class recall was 81%, consistent with the linear kernel.
- F1-score: Maintained an F1-score of 89% for malignant class and 97% for benign class.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.94      1.00      0.97       110
           1       1.00      0.81      0.89        36

    accuracy                           0.95       146
   macro avg       0.97      0.90      0.93       146
weighted avg       0.95      0.95      0.95       146
```

Figure 92 SVM on RBF Kernel Classification Report

- Confusion Matrix: Similar to the linear kernel, it predicted 110 TN, 0 FP, 7 FN, and 29 TP.

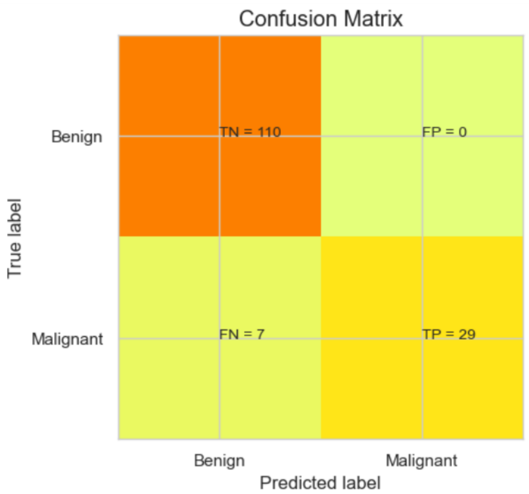


Figure 93 SVM on RBF Kernel Confusion Matrix

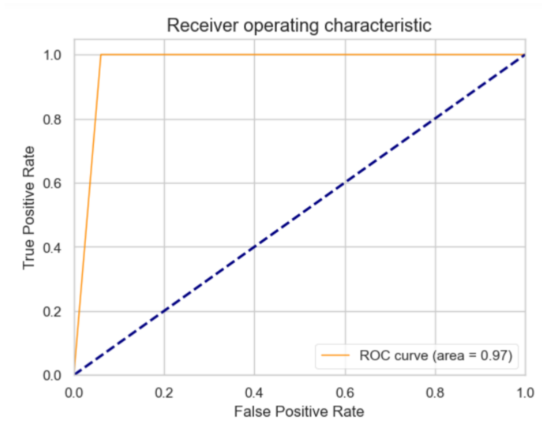


Figure 94 SVM on RBF Kernel Receiver Operating Characteristic

c. Kernel Selection Using Learning Curve

In machine learning, understanding the bias-variance trade-off is crucial. High bias indicates underfitting, where the model is too simple to capture the complexities in the data. On the other hand, high variance suggests overfitting, where the model fits the training data too closely but fails to generalize well to unseen data.

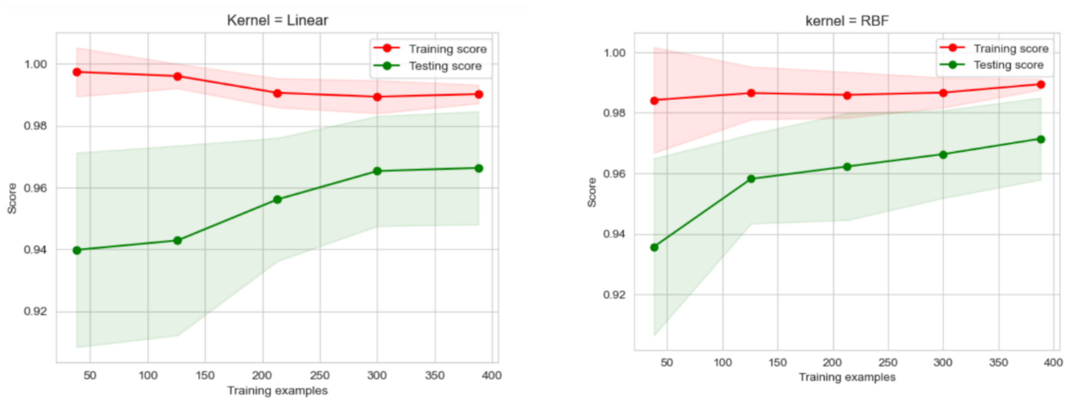


Figure 95 SVM on Linear and RBF kernel Learning Curve

On the RBF kernel, I observed high variance in the Support Vector Machine (SVM) model. This is evident from the significant gap between the training and testing scores. While the training accuracy is close to 1, indicating a near-perfect fit to the training data, the testing accuracy is comparatively lower, around 0.7. This discrepancy signals overfitting—the model is memorizing the training data rather than learning underlying patterns, leading to poor generalization.

d. Selection of Regularization Parameter (C)

The regularization parameter C in SVM controls the trade-off between minimizing training error and maintaining simplicity to generalize well on unseen data.

I plotted learning curves for SVM with both linear and RBF kernels, varying the value of C. The objective was to find an optimal C value that balances bias and variance, achieving a good trade-off between training and testing accuracy.

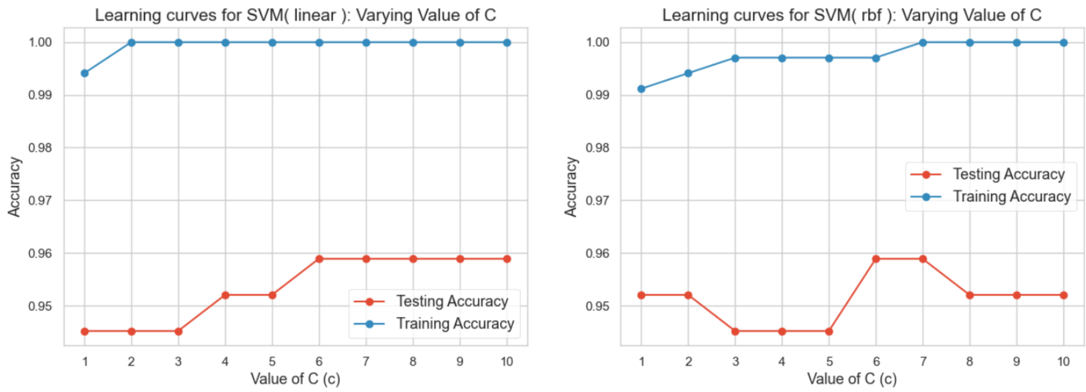


Figure 96 Learning Curve for SVM (linear & RBF): varying value of C

e. Optimal Model

After evaluating various models and hyperparameters, I identified an optimal SVM model with an RBF kernel and C=1. This model demonstrated a high level of performance, achieving an accuracy of approximately 95%.

2. Logistic Regression:

- Accuracy: The logistic regression classifier achieved a high accuracy of 96%, indicating that it correctly classified 96% of the test instances.
- Precision: The precision values were high for both classes, with 95% precision for the benign class (0) and 100% precision for the malignant class (1). This indicates that when the classifier predicted a class, it was correct 95% of the time for the benign class and 100% of the time for the malignant class.
- Recall: The recall value for the malignant class was 83%, suggesting that the classifier correctly identified 83% of the actual malignant cases.
- F1-score: The F1-score, which is the harmonic mean of precision and recall, was 91% for the malignant class and 97% for the benign class.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.95      1.00      0.97       110
           1       1.00      0.83      0.91        36

    accuracy                           0.96       146
   macro avg       0.97      0.92      0.94       146
weighted avg       0.96      0.96      0.96       146
```

Figure 97 Logistic regression Classification Report

- Confusion Matrix: The confusion matrix showed 110 true negatives (TN), 0 false positives (FP), 6 false negatives (FN), and 30 true positives (TP).

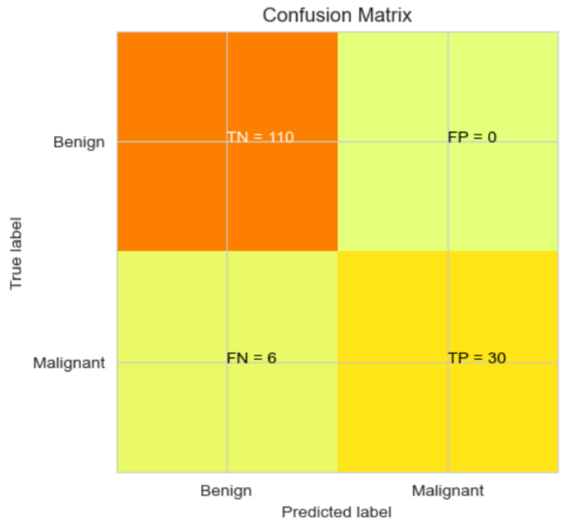


Figure 98 Logistic regression Confusion Matrix

- AUC-ROC Score: The AUC-ROC score, which measures the classifier's ability to distinguish between classes, was approximately 0.983, indicating an important level of performance.

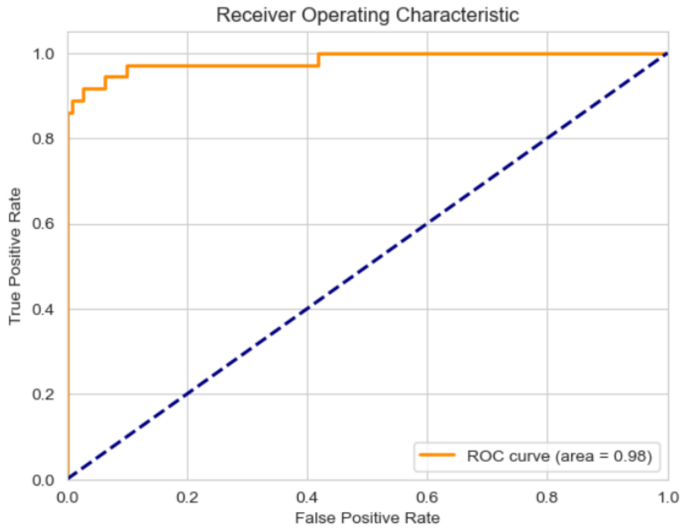


Figure 99 Logistic regression Receiver operating characteristic curve.

3. Gradient Boosting:

- Accuracy: Maintained a high accuracy of 95%, similar to logistic regression.
- Precision: Achieved high precision values for both classes, with 95% precision for the benign class and 94% for the malignant class.
- Recall: The recall for the malignant class was 83%, consistent with logistic regression.
- F1-score: Obtained F1-scores of 88% for the malignant class and 96% for the benign class.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.98      0.96       110
           1       0.94      0.83      0.88        36

    accuracy                           0.95       146
   macro avg       0.94      0.91      0.92       146
weighted avg       0.94      0.95      0.94       146
```

Figure 100 Gradient Boosting Classification Report

Confusion Matrix: Predicted 108 TN, 2 FP, 6 FN, and 30 TP.

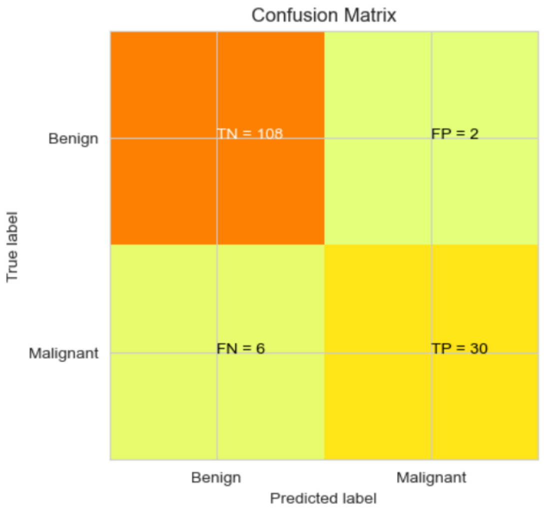


Figure 101 Gradient Boosting Confusion Matrix

AUC-ROC Score: Attained an AUC-ROC score of approximately 0.982, indicating strong discriminative ability.

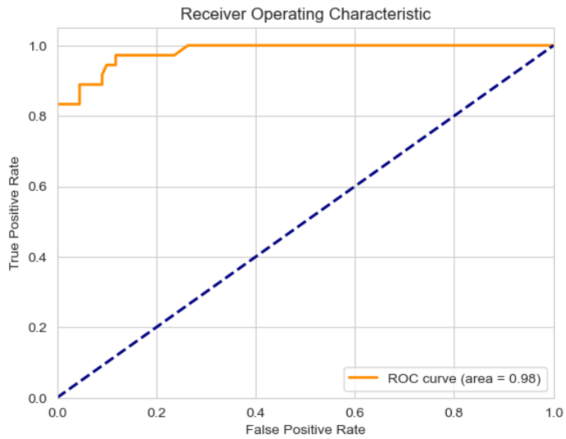


Figure 102 Gradient Boosting Receiver operating characteristic curve.

4. Random Forest:

- Accuracy: Achieved a slightly lower accuracy of 94% compared to logistic regression and gradient boosting.
- Precision: Obtained high precision values, with 95% precision for the benign class and 91% for the malignant class.
- Recall: Showed a recall of 83% for the malignant class, consistent with previous classifiers.
- F1-score: Maintained F1-scores of 87% for the malignant class and 96% for the benign class.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.97      0.96       110
           1       0.91      0.83      0.87        36

    accuracy                           0.94       146
   macro avg       0.93      0.90      0.91       146
weighted avg       0.94      0.94      0.94       146
```

Figure 103 Random Forest Classification Report

Confusion Matrix: Resulted in 107 TN, 3 FP, 6 FN, and 30 TP.

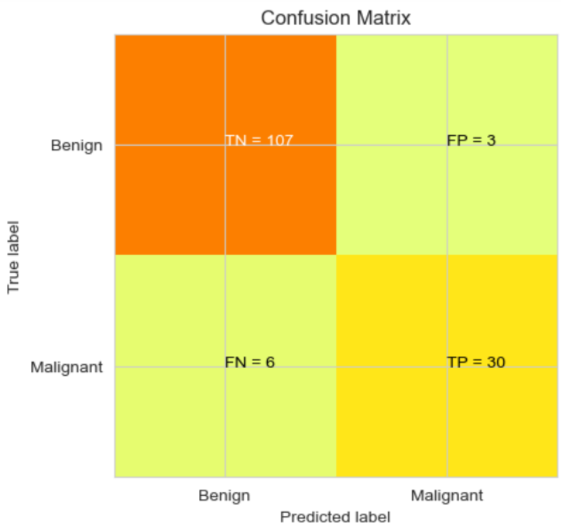


Figure 104 Random Forest Confusion Matrix

AUC-ROC Score: Achieved an AUC-ROC score of approximately 0.976, indicating good discriminative ability.

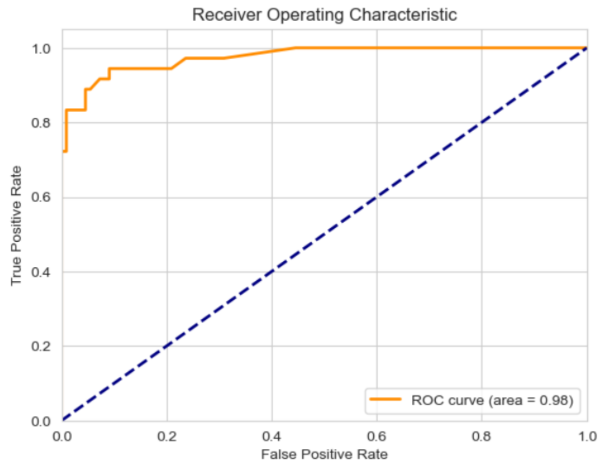


Figure 105 Random Forest Receiver operating characteristic curve.

5. Decision Tree:

- Accuracy: Achieved a lower accuracy of 91% compared to previous classifiers.
- Precision: Attained high precision values, with 94% precision for the benign class and 83% for the malignant class.
- Recall: Demonstrated a recall of 81% for the malignant class.
- F1-score: Obtained F1-scores of 82% for the malignant class and 94% for the benign class.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.94      0.95      0.94       110
           1       0.83      0.81      0.82        36

    accuracy                           0.91       146
   macro avg       0.88      0.88      0.88       146
weighted avg       0.91      0.91      0.91       146
```

Figure 106 Decision Tree Classification report

Confusion Matrix: Displayed 104 TN, 6 FP, 7 FN, and 29 TP.

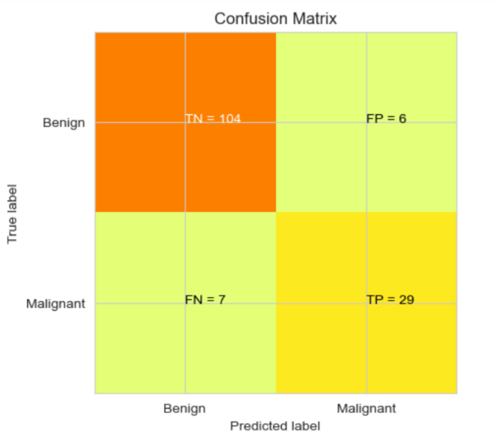


Figure 107 Decision Tree Confusion Matrix

AUC-ROC Score: Achieved an AUC-ROC score of approximately 0.876, indicating moderate discriminative ability.

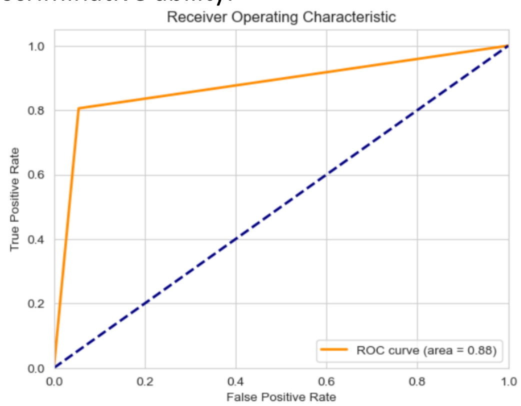


Figure 108 Random Forest Receiver operating characteristic curve.

6. KNeighbors:

- Accuracy: Maintained a high accuracy of 94%.
- Precision: Achieved high precision values, with 93% precision for the benign class and 97% for the malignant class.
- Recall: Showed a recall of 78% for the malignant class, indicating some misclassification of actual malignant cases.
- F1-score: Obtained F1-scores of 86% for the malignant class and 96% for the benign class.

Classification Report:

	precision	recall	f1-score	support
0	0.93	0.99	0.96	110
1	0.97	0.78	0.86	36
accuracy			0.94	146
macro avg	0.95	0.88	0.91	146
weighted avg	0.94	0.94	0.94	146

Figure 109 KNN Classification Report

Confusion Matrix: Predicted 109 TN, 1 FP, 8 FN, and 28 TP.

AUC-ROC Score: Attained an AUC-ROC score of approximately 0.949, indicating good discriminative ability.

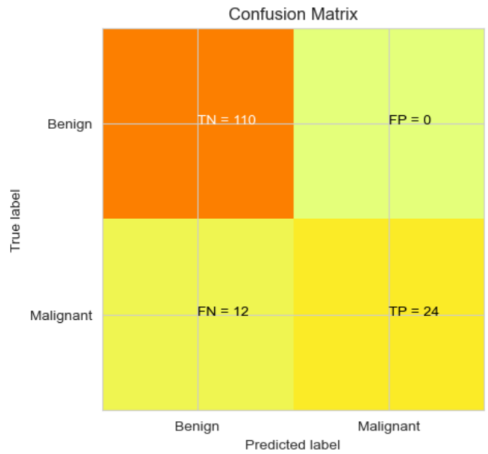


Figure 110 KNN Confusion Matrix

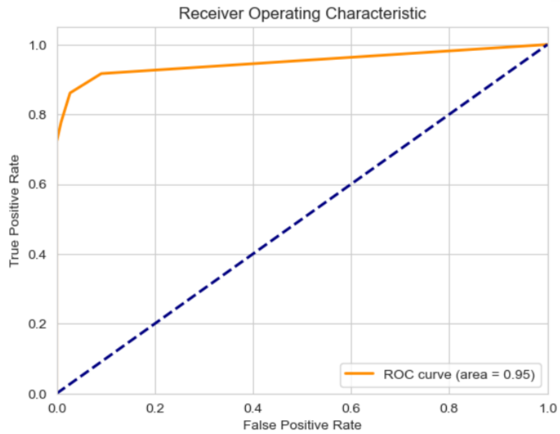


Figure 111 KNN Receiver operating characteristic curve.

a. Best Value of K:

- Utilizing mean cross-validation scores, the optimal value of kk was determined to be 7, enhancing the classifier's performance and contributing to its accuracy and effectiveness in classifying cancerous and non-cancerous cells.

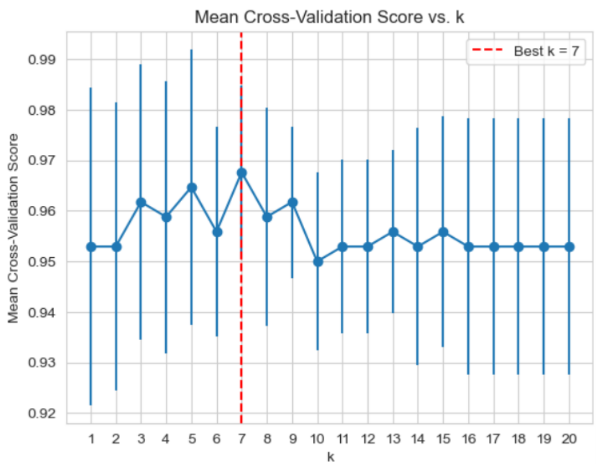


Figure 112 Mean Cross-Validation Score vs. k.

7. XGB:

- Accuracy: Achieved a slightly lower accuracy of 93% compared to some previous classifiers.
- Precision: Obtained high precision values, with 95% precision for the benign class and 88% for the malignant class.
- Recall: Demonstrated a recall of 83% for the malignant class.
- F1-score: Maintained F1-scores of 86% for the malignant class and 95% for the benign class.

```
Classification Report:
              precision    recall  f1-score   support

           0       0.95      0.96      0.95       110
           1       0.88      0.83      0.86        36

    accuracy                           0.93       146
   macro avg       0.91      0.90      0.91       146
weighted avg       0.93      0.93      0.93       146
```

Figure 113 XGB Classification report.

Confusion Matrix: Resulted in 106 TN, 4 FP, 6 FN, and 30 TP.

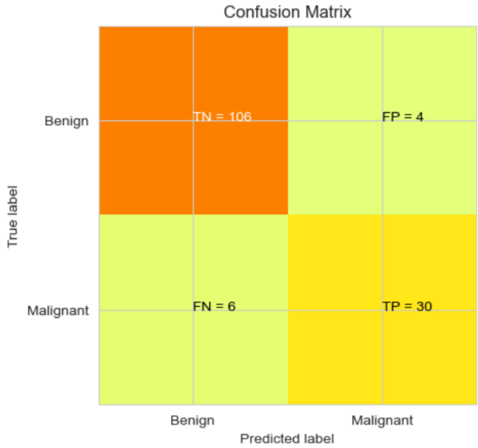


Figure 114 XGB Confusion matrix.

AUC-ROC Score: Obtained an AUC-ROC score of approximately 0.969, indicating strong discriminative ability.

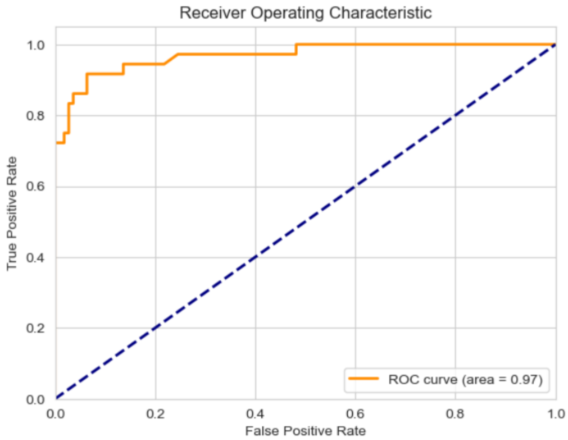


Figure 115 XGB Receiver operating characteristic curve.

3. Performance of the Variational Quantum Classifier (VQC)

1. QASM Simulator As a Backend

```
Classification Report:
              precision    recall  f1-score   support

           0       0.78      0.44      0.56        41
           1       0.75      0.93      0.83        73

    accuracy                           0.75       114
   macro avg       0.76      0.69      0.70       114
weighted avg       0.76      0.75      0.73       114
```

Figure 116 QASM Simulator as a Backend Classification report.

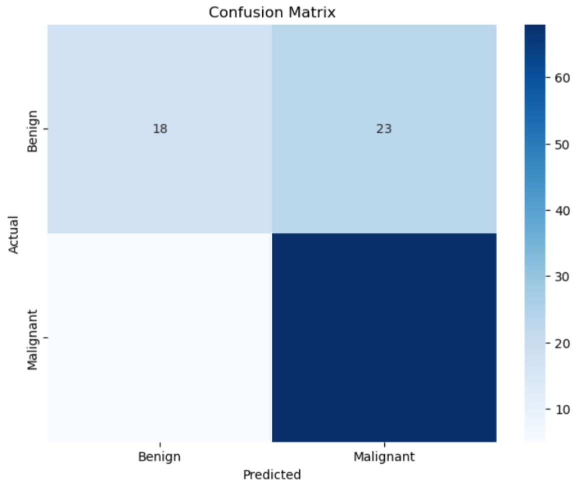


Figure 117 QASM Simulator as a Backend Confusion Matrix

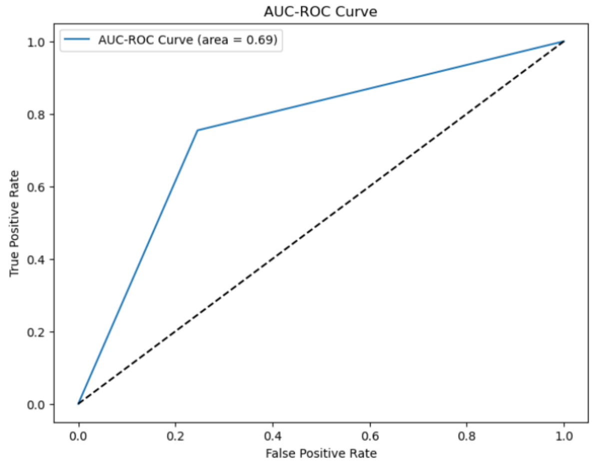


Figure 118 QASM Simulator as a Backend Receiver operating characteristic curve.

2. Quantum Computer as A Backend

```
Classification Report:
              precision    recall  f1-score   support

         0.0       0.78      0.70      0.74        57
         1.0       0.73      0.80      0.76        57

    accuracy                           0.75       114
   macro avg       0.75      0.75      0.75       114
weighted avg       0.75      0.75      0.75       114
```

Figure 119 Quantum Computer as A Backend Classification Report

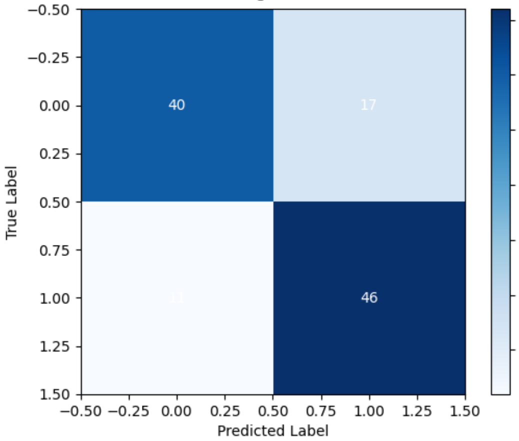


Figure 120 Quantum Computer as A Backend Confusion Matrix

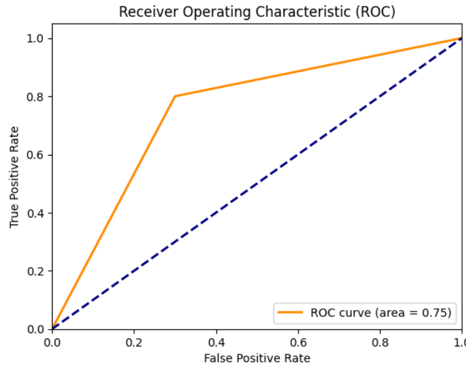


Figure 121 Quantum Computer as A Backend Receiver operating characteristic curve

4. Comparative Analysis: Classical Machine Learning vs. Variational Quantum Classifier (VQC)

This section aims to compare the performance of classical machine learning models with the Quantum Variational Classifier (VQC). The analysis will highlight why classical models exhibit superior performance compared to the VQC in the context of breast cancer detection.

1. Performance Comparison

We begin by analyzing the performance of classical machine learning algorithms, which have been extensively developed and optimized over many years. These algorithms leverage sophisticated techniques for handling high-dimensional data and capturing complex patterns. Among the classical models evaluated, logistic regression emerges as the top performer, achieving an accuracy of 95.89%. Support Vector Machine (SVM) closely follows with an accuracy of 95.21%, demonstrating robust performance in distinguishing between malignant and benign cases. Gradient Boosting, K-Nearest Neighbors (KNN), and Random Forest classifiers also exhibit impressive performance, with accuracies ranging from 93.15% to 93.84%. However, the Decision Tree classifier lags with an accuracy of 89.04%.

In contrast to classical machine learning, the VQC represents a novel approach to classification leveraging the principles of quantum computing. However, the current state of quantum hardware imposes significant limitations on the performance of the VQC. In our evaluation, the VQC achieves an accuracy of 75% when utilizing both QASM simulator and real quantum device backends. This lower accuracy can be attributed to the nascent stage of quantum computing technology, as well as the challenges posed by hardware constraints such as limited qubit counts, noise, and decoherence.

Accuracy	Algorithm
0.952055	SVC
0.890411	DecisionTreeClassifier
0.958904	LogisticRegression
0.938356	KNeighborsClassifier
0.931507	RandomForestClassifier
0.938356	GradientBoostingClassifier
0.945205	XGB

Table 13 Table of Summary of model's performance

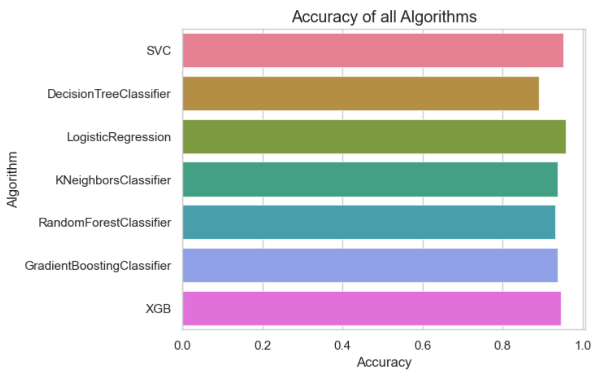


Figure 122 Accuracy of all algorithms.

2. Interpretation

Breast cancer detection was evaluated using both conventional machine learning (ML) algorithms and quantum computing (QC) models, with the performance assessed on the Wisconsin Breast Cancer dataset using an 80-20 dataset split. For the classical machine learning algorithms, all thirty features were utilized, whereas for the quantum computing models, only four features were mapped onto four qubits. Among the classical ML algorithms, logistic regression achieved the highest accuracy at 95.89%, followed closely by the Support Vector Machine (SVC) with an accuracy of 95.21%. Gradient Boosting and K-Nearest Neighbors (KNN) both attained accuracies of 93.84%, while the Random Forest classifier lagged slightly with an accuracy of 93.15%. The Decision Tree classifier had the lowest accuracy among the classical models at 89.04%. In contrast, the Quantum Variational Classifier (VQC) method, using both QASM simulator and real quantum device backends, achieved an accuracy of 75%. While the SVC with RBF and polynomial kernels performed the best among all models evaluated, the quantum methods demonstrated lower accuracy but still showed promising potential for future improvements in quantum computing technology and algorithmic development.

The significant performance disparity between classical and quantum approaches can be attributed to several factors:

- **Maturity and Development:** Classical algorithms have benefited from years of research, optimization, and practical applications. This maturity translates into robust techniques for handling high-dimensional data and extracting complex patterns.
- **Data Representation and Feature Space:** Classical models excel at utilizing all 30 features in the Wisconsin Breast Cancer dataset. Established techniques for feature selection, engineering, and scaling empower them to efficiently capture intricate relationships within the data. In contrast, VQC, limited by the available qubits (four in this case), can only represent a smaller subset of features, hindering its ability to capture the full complexity of the data.
- **Computational Infrastructure:** Classical machine learning leverages robust and mature computational infrastructure, enabling efficient training and deployment on large datasets. This translates to faster processing times and more scalable solutions compared to the nascent quantum computing infrastructure.

3. Future Work:

1. Investigate the impact of feature engineering techniques on the performance of VQC, potentially leading to improved accuracy and robustness.
2. Explore optimization techniques for quantum circuits to reduce errors and improve the performance of VQC.
3. Apply VQC to larger breast cancer datasets to evaluate its scalability and performance on more complex data.
4. Investigate the application of quantum error correction techniques to mitigate the impact of noise and errors on VQC performance.
5. Experiment with alternative quantum models like Quantum Support Vector Machines (QSVM).
6. Monitor advancements in quantum computing hardware, including increases in qubit count, reduction of noise and errors, and improvements in gate fidelities and coherence times. As quantum hardware matures, the scalability and performance of quantum machine learning models are expected to improve significantly, potentially narrowing the performance gap between classical and quantum approaches.

Conclusion

This chapter compares classical machine learning algorithms (e.g. logistic regression, SVM) with the Quantum Variational Classifier (VQC) for breast cancer detection. Classical algorithms outperform VQC in terms of accuracy and robustness, due to their maturity, data representation, computational infrastructure, and stability. Despite this, VQC shows promise and encourages further research in quantum computing for breast cancer detection, with future directions including feature engineering, quantum circuit optimization, large-scale datasets, and quantum error correction.

Conclusion

The comparative analysis conducted in this study provides valuable insights into the current capabilities and limitations of classical machine learning models versus hybrid quantum-classical models, specifically the Variational Quantum Classifier (VQC), in the context of breast cancer detection. The findings indicate that classical machine learning models, such as Support Vector Machine (SVM), Logistic Regression, Gradient Boosting, Random Forest and XGBoost, generally exhibit superior performance in terms of accuracy and other evaluation metrics when compared to the VQC model.

The empirical results underscore the effectiveness of classical models in handling the Breast Cancer Wisconsin (Diagnostic) Dataset, achieving higher accuracy and more reliable diagnostic performance. These models have benefited from decades of optimization and refinement, leading to robust algorithms that excel in various data-driven applications, including medical diagnostics.

On the other hand, the implementation of the VQC model, while promising in theory due to the unique capabilities of quantum computing did not surpass the performance of its classical counterparts in this specific application. This outcome highlights the current developmental stage of quantum computing technologies. Quantum machine learning is still in its infancy, and there are significant challenges to be addressed, including the development of more sophisticated quantum algorithms.

Despite these challenges, the potential of quantum computing in revolutionizing various fields remains substantial. Quantum machine learning could lead to breakthroughs in areas where classical algorithms struggle, especially as quantum technology matures. For instance, quantum algorithms may offer significant advantages in solving optimization problems, modeling complex systems, and analyzing large-scale data sets more efficiently than classical methods.

This study's findings contribute to the broader understanding of the interplay between classical and quantum machine learning approaches. It emphasizes the need for continued research and experimentation to enhance the performance and applicability of quantum models. Future research directions could include exploring different quantum algorithms, optimizing quantum circuit designs, and improving error correction methods to enhance the reliability and accuracy of quantum computations.

In conclusion, while classical machine learning models currently outperform the VQC in breast cancer detection, the integration of quantum computing presents an exciting frontier with the potential to transform numerous scientific and technological domains. Continued advancements in quantum technology, combined with innovative algorithmic developments, are expected to unlock new possibilities, and significantly impact the future of machine learning and data analysis.

Bibliography and Webography

[1] K. Benallel, R. El Kilali, R. Benjelloun, and M. Kadiri, "Depression and Breast Cancer in Morocco: Prevalence and Associated Factors," *Int J Breast Cancer*, vol. 2023, pp. 1–10, Nov. 2023, doi: 10.1155/2023/3277929.

[2] W.-X. Liao *et al.*, "Automatic Identification of Breast Ultrasound Image Based on Supervised Block-Based Region Segmentation Algorithm and Features Combination Migration Deep Learning Model," *IEEE J Biomed Health Inform*, vol. 24, no. 4, pp. 984–993, Apr. 2020, doi: 10.1109/JBHI.2019.2960821.

[3] Z. Cao, G. Yang, Q. Chen, X. Chen, and F. Lv, "Breast tumor classification through learning from noisy labeled ultrasound images," *Med Phys*, vol. 47, no. 3, pp. 1048–1057, Mar. 2020, doi: 10.1002/mp.13966.

[4] E. Zhang, S. Seiler, M. Chen, W. Lu, and X. Gu, "BIRADS Features-Oriented Semi-supervised Deep Learning for Breast Ultrasound Computer-Aided Diagnosis," Apr. 2019, doi: 10.1088/1361-6560/ab7e7d.

[5] W. K. Moon, Y.-W. Lee, H.-H. Ke, S. H. Lee, C.-S. Huang, and R.-F. Chang, "Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks," *Comput Methods Programs Biomed*, vol. 190, p. 105361, Jul. 2020, doi: 10.1016/j.cmpb.2020.105361.

[6] A. E. Ilesanmi, U. Chaumrattanakul, and S. S. Makhanov, "A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning," *Biocybern Biomed Eng*, vol. 41, no. 2, pp. 802–818, Apr. 2021, doi: 10.1016/j.bbe.2021.05.007.

[7] T. Pang, J. H. D. Wong, W. L. Ng, and C. S. Chan, "Semi-supervised GAN-based Radiomics Model for Data Augmentation in Breast Ultrasound Mass Classification," *Comput Methods Programs Biomed*, vol. 203, p. 106018, May 2021, doi: 10.1016/j.cmpb.2021.106018.

[8] Z. Zhuang, Z. Yang, A. N. J. Raj, C. Wei, P. Jin, and S. Zhuang, "Breast ultrasound tumor image classification using image decomposition and fusion based on adaptive multi-model spatial feature fusion," *Comput Methods Programs Biomed*, vol. 208, p. 106221, Sep. 2021, doi: 10.1016/j.cmpb.2021.106221.

[9] M. Byra, "Breast mass classification with transfer learning based on scaling of deep representations," *Biomed Signal Process Control*, vol. 69, p. 102828, Aug. 2021, doi: 10.1016/j.bspc.2021.102828.

[10] G. Ayana, K. Dese, and S.-W. Choe, "Transfer Learning in Breast Cancer Diagnoses via Ultrasound Imaging.," *Cancers (Basel)*, vol. 13, no. 4, Feb. 2021, doi: 10.3390/cancers13040738.

[11] K. Wang, S. Liang, S. Zhong, Q. Feng, Z. Ning, and Y. Zhang, "Breast ultrasound image segmentation: A coarse-to-fine fusion convolutional neural network," *Med Phys*, vol. 48, no. 8, pp. 4262–4278, Aug. 2021, doi: 10.1002/mp.15006.

[12] G. Ayana and S. Choe, "BUViTNet: Breast Ultrasound Detection via Vision Transformers," *Diagnostics*, vol. 12, no. 11, p. 2654, Nov. 2022, doi: 10.3390/diagnostics12112654.

[13] Y. Zhang *et al.*, "BUSIS: A Benchmark for Breast Ultrasound Image Segmentation," *Healthcare*, vol. 10, no. 4, p. 729, Apr. 2022, doi: 10.3390/healthcare10040729.

[14] E. Kaplan *et al.*, "Automated BI-RADS classification of lesions using pyramid triple deep feature generator technique on breast ultrasound images," *Med Eng Phys*, vol. 108, p. 103895, Oct. 2022, doi: 10.1016/j.medengphy.2022.103895.

[15] H. Feng *et al.*, "Identifying Malignant Breast Ultrasound Images Using ViT-Patch," *Applied Sciences*, vol. 13, no. 6, p. 3489, Mar. 2023, doi: 10.3390/app13063489.

[16] N. Manjunathan, N. Gomathi, and S. Muthulingam, "Early Detection of Breast Cancer using Machine Learning," in *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)*, IEEE, Jun. 2023, pp. 165–169. doi: 10.1109/ICSCSS57650.2023.10169777.

[17] S. Zakareya, H. Izadkhah, and J. Karimpour, "A New Deep-Learning-Based Model for Breast Cancer Diagnosis from Medical Images," *Diagnostics*, vol. 13, no. 11, p. 1944, Jun. 2023, doi: 10.3390/diagnostics13111944.

[18] K. Meghana, N. Nandal, R. Tanwar, L. Goel, and G. Chhabra, "Breast Cancer Detection with Machine Learning-A Review," in *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)*, IEEE, Mar. 2023, pp. 168–172. doi: 10.1109/ICSCDS56580.2023.10104644.

[19] K. Purohit, "Separation of Data Cleansing Concept from EDA," *International Journal of Data Science and Analysis*, vol. 7, no. 3, p. 89, 2021, doi: 10.11648/j.ijdsa.20210703.16.

[20] M. Akrom, S. Rustad, and H. K. Dipojono, "Development of quantum machine learning to evaluate the corrosion inhibition capability of pyrimidine compounds," *Mater Today Commun*, vol. 39, p. 108758, Jun. 2024, doi: 10.1016/j.mtcomm.2024.108758.

[21] E. I. Elsedimy, S. M. M. AboHashish, and F. Algarni, "New cardiovascular disease prediction approach using support vector machine and quantum-behaved particle swarm optimization," *Multimed Tools Appl*, vol. 83, no. 8, pp. 23901–23928, Aug. 2023, doi: 10.1007/s11042-023-16194-z.

[22] G. Abdulsalam, S. Meshoul, and H. Shaiba, "Explainable Heart Disease Prediction Using Ensemble-Quantum Machine Learning Approach," *Intelligent Automation & Soft Computing*, vol. 36, no. 1, pp. 761–779, 2023, doi: 10.32604/iasc.2023.032262.

[23] S. Alsubai, A. Alqahtani, A. Binbusayyis, M. Sha, A. Gumaei, and S. Wang, "Heart Failure Detection Using Instance Quantum Circuit Approach and Traditional Predictive Analysis," *Mathematics*, vol. 11, no. 6, p. 1467, Mar. 2023, doi: 10.3390/math11061467.

[24] R. Narain, S. Saxena, and A. Goyal, "Cardiovascular risk prediction: a comparative study of Framingham and quantum neural network based approach," *Patient Prefer Adherence*, vol. Volume 10, pp. 1259–1270, Jul. 2016, doi: 10.2147/PPA.S108203.

[25] S. S. Kavitha and N. Kaulgud, "Quantum K-means clustering method for detecting heart disease using quantum circuit approach," *Soft comput*, vol. 27, no. 18, pp. 13255–13268, Sep. 2023, doi: 10.1007/s00500-022-07200-x.

Annex

QBronze Certification, QWorld, 2024: Acquired through an immersive workshop series on quantum computing and programming, with special thanks to Prof. TARIK EL BARAKY for the invitation to this enlightening experience.