Informatics Institute of Technology
In Collaboration with
University Of Westminster

Fibro-QuanNet

Pulmonary Fibrosis Prognosis Prediction using Quantum Machine Learning

A Dissertation by
Mr Nuwin Hansitha Godakanda Arachchi

Supervised by
Mr Guhanathan Poravi

May 2023

Submitted in partial fulfilment of the requirements for the
BEng (Hons) in Software Engineering degree at the University of Westminster, UK

ABSTRACT

Pulmonary fibrosis is a progressive lung condition caused by damaged or scarred lung tissue obstructing the exchange of carbon dioxide and oxygen gasses in the alveoli Thereby, leaving the body deprived of the oxygen required for blood oxygenation and less lung volume. As per state-of-the-art medical practice, the deterioration/ scarring of the lung tissue is not entirely reversible or correctable, merely leaving patients with symptom management using therapy and clinical drug trials. An accurate judgment of the lung function decline is crucial for the management and trial treatment of the patient.

This research project endeavors to automate the process of prognosis prediction of pulmonary fibrosis using a hybrid-classical quantile regression hybrid model built using a variational quantum circuit. Although quantum computing is still in its formative years, research activities done in similar domains have proved to have immaculate success in both the correctness and speed of the results. The project explores the advantages one might gain by utilizing the developing quantum computing over the use of classical computational approaches, which will in return facilitate and encourage more optimization of machine learning using quantum computing.

The model has shown promising results so far, with a Laplace Log Likelihood matrix of -7.13, and a mean absolute error of just 212.31. For a regression model trained with a small dataset such as the OSIC dataset with just 700+ DICOM images with its metadata, the evaluations are noticeable and promising.

Keywords: Pulmonary Fibrosis, Prognosis Prediction, Quantum Computing, Quantum Machine Learning,

Subject Descriptors:

Computing methodologies → Machine learning → Machine learning approaches → Neural networks

Theory of computation → Models of computation → Quantum computation theory → Quantum complexity theory

Applied computing → Life and medical sciences → Consumer health

Human-centered computing → Human computer interaction (HCI) → Interaction paradigms → Graphical user interfaces

DECLARATION

I hereby certify to that this dissertation and all its associated artifacts are the product of my own research work, and that none of these components have ever been nor in the process of being submitted at research programs or other academic qualification programs to any other universities or academic institutions. All information and particulars were procured from trusted, reliable sources and have been cited appropriately.

Students' Full Name: Nuwin Hansitha Godakanda Arachchi
Registration Number: W1761350 | 2019443

Signature: *Date*: 9th of May 2023

ACKNOWLEDGEMENT

This research journey wasn't an easy one. It had its ups and downs along the way and took hours of dedication and commitment to successfully complete the project.

Firstly, I would like to thank Mr Guhanathan Poravi, my supervisor, for his constant motivation and push to go above and beyond with the research project and showing us the best way possible in every steep hill of the research process. Your constant support and encouragement is to thank for the successful completion of this project. I would also like to thank and remind the quantum computing researchers at Microsoft India especially Mr Kinshuk Sengupta, for their guidance throughout the development process of the project deliverables. The quantum variational circuits would not have been a success if not for them. I would also like to extend my thank Mr Azhar Aneeze, who also helped me throughout the process amid their busy schedules. This project would also not be a success without the doctors and medical students that helped me throughout the process, especially to Dr Yumandhee Ayanajee, who was also available when a clafification raised, and so was all the staff at Durdans hospital, Colombo, and the medical students at Grodno State Medical University.

Finally, but never the least, I would like to thank my dear parents, for their constant loving boost of energy right when I need it, to keep me driven and focused throughout this tough year. I would also like to thank my sister, for her lame jokes that kept me smiling and driven through this immense pressure.

Once again, thank you very much.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Acronym	Description
AI	Artificial Intelligence
DICOM	Digital Imaging and Communications in Medicine
FVC	Forced Volume Capacity
GUI	Graphical User Interface
HRCT	High-Resolution Computed Tomography
IDL	Interstitial Lung Disease
IPF	Idiopathic Pulmonary Fibrosis
MONAI	Medical Open Network for Artificial Intelligence
MRI	Magnetic Resonance Imaging
PF	Pulmonary Fibrosis
QC	Quantum Computing
QML	Quantum Machine Learning
ROCc	Receiver Operating Characteristic curve
UIP	Usual Interstitial Pneumonia

CHAPTER 1: INTRODUCTION

1.1 Prolegomena

This dissertation illustrates the authors attempt at automating the prognosis prediction of pulmonary fibrosis using a novel quantum machine learning approach, while clearly demonstrating the methodology adhered to while performing the research along with a thorough and critical evaluation of the existing work in the literature. Finally, the document indicates the evaluation of the novel ensemble models developed as product of the research and future work avenues identified as by-product of the research.

1.2 Problem Domain

1.2.1 Pulmonary Fibrosis (PF)

PF is a progressive lung disease caused by damaged or scarring in the lung tissue or cavity, occasionally prefixed as idiopathic PF (IPF), when of unknown causality (Devaraj, 2014). The scarred/ damaged area will fibrous the pulmonic tissue, obstructing the exchange of carbon dioxide and oxygen gasses in the alveoli (tiny air-sacks at the end of the of the airflow branches (U.S. Department of Health and Human Services, 2011)), thereby leaving the body deprived of the oxygen required for blood oxygenation and less lung volume (American Lung Association, 2022).

Consequent to the significant loss of lung real estate and capacity, the amount of air the patient may respire declines considerably, thus, encountering persistent dyspnea (shortness of breath) (Pulmonary Fibrosis Foundation, 2022). As per state-of-the-art medical practice, the deterioration/ scarring of the lung tissue is not entirely reversible or correctable, merely leaving patients with symptom management using therapy and clinical drug trials (Mayo Foundation for Medical Education and Research, 2021). Therefore, PF has received awareness and concern as one of the most common and lethal forms of idiopathic interstitial lung disease, with an associated survival median of just three years (Das and Chakraborty, 2015). According to the Pulmonary Fibrosis Foundation, one in every two-hundred adults' over 70 years of age may be affected by IPF in the united states alone, with nearly 250,000 cases diagnosed and over 50,000 reported yearly (Pulmonary Fibrosis Foundation, 2022). Approximately 40,000 people lose their lives to PF/IPF in the United States alone (Schwartz, 2018).

Early diagnosis of PF plays a crucial part in early treatments to maintain the patient's lung capacity and function within an acceptable and healthy range while attempting to prevent further deterioration (Choi, 2021). The survival medians of mild, moderate, and severe PF categorized by their FVC percentages were 55.6, 38.7, and 27.4 months (about 2 and a half years) respectively (Nathan et al., 2011), indicating the necessity for early diagnosis of the disease.

1.2.2 Prognosis Prediction of Pulmonary Fibrosis

The deterioration of the lung capacity due to PF and IPF is impossible to be determined and may range from minimal or no degeneration over multiple years to immediate decline within a short time, often even weeks (Das and Chakraborty, 2015). Therefore, an accurate judgment of the lung function decline is crucial for the management and trial treatment of the patient. Prognosis also helps medical professionals determine the best course of treatment for the patient or where treatment may not be possible results will back the professionals' decision to suggest lung transplantation as the ultimate alternative (Kistler et al., 2014).

For conducting manual detection and progress prediction of PF, multiple institutions have published guidelines, including invasive techniques such as pulmonary tissue biopsy; as well as less invasive techniques such as video-assisted thoracoscopy (Richeldi et al., 2014). Similarly, several techniques are also used to access the lung function decline such as spirometry testing, a classical derivative of lung function and FVC. However, spirometry tests do not take into attention the underlying mechanisms nor the progression techniques (Wuyts et al., 2016).

High-Resolution Computed Tomography (HRCT) has thus come to wide-spread use as the most effective method of assessing lung function and had now become monotonous to conduct a HRCT/CT imaging to understand the underlying mechanism of the lung (Rachel, 2020). In order to identify PF, radiologists and other medical professionals may look for radiologic features such as “honeycombing” or “lung architectural distortion” which indicates the presence of PF/IPF, then required to be verified using a biopsy sample of the lung wall tissue (Niknejad, 2022).

1.3 Problem Definition

Given the severity of PF, it is eminent that prognosis prediction of the diagnosis is crucial in both the patients’ well-being as well as providing medical professionals with authoritative and reliable information to perform informed and objective decisions. However, for a manual prognosis prediction, the patient will be subjected to constant, expensive, and rigorous testing such as spirometry testing/ HRCT imaging (Mayo Foundation for Medical Education and Research, 2021). This delays the prognosis prediction at a whole, taking away the competitive advantage of the prediction, thus missing out on valuable time medical professionals may have to make profound and life-saving decisions regarding the patient (Choi, 2021).

As the National Health Service (2022) reports, IPF is as difficult to detect and prognose due to its quite similarity to other lung conditions such as Chronic Obstructive Pulmonary Disease (COPD), thereby supporting the dire need for further and more thorough research into the automation of prognosis of PF, along with computation intervention to quicken the process.

1.3.1 Problem Statement

Pulmonary fibrosis is a progressive pulmonary condition at which state-of-the-art medical practices are defenseless leaving experts to yield critical decisions based on the prognosis of lung functionality manually, which is time-consuming and prone to error.

1.4 Research Motivation

Pulmonary fibrosis is one of the lethal pulmonary interstitial lung conditions known, claiming thousands of lives, which ultimately leaves patients at an extremely un-comfortable situation having thickening, stiffening and scarring the lung tissue (Schwartz, 2018). Accurate and efficient prediction of the likely prognosis of this disease will be able to give patients more time for clinical treatment trials (Kumaran and Preethi, 2022), which might one day be able to find a potential cure for PF/ IPF.

The manual diagnosis process by itself costs a good fortune for rigorous and constant follow up testing and consumes a longer time span to derive an accurate prognosis using accumulated test FVC results and HRCT imagery (Rachel, 2020). Getting to perform an accurate prognosis prediction using a single HRCT testing would be able to save up millions being spent on follow up, overwhelming testing and unnecessary hospital visits. Pulmonary fibrosis is also close to the author personally, with PF diagnosis cases in close home.

1.5 Research Gap

Based on previous studies concerning pulmonary fibrosis prognosis prediction, the literature implies minimal research activity utilizing quantum computing technologies as a scalable option to drive prognosis prediction for PF using machine learning to far more efficient and accurate outcomes (Rachel, 2020).

- A quantum prediction model using neural networks have so far not been attempted.
- Utilize patient metadata as input fields, allowing external biomarkers to interfere.

This project proposes to address this gap by introducing a novel approach to pulmonary fibrosis prognosis prediction using quantum machine learning utilizing data imagery from medical apparatuses.

1.6 Contribution to the Body of Knowledge

The author's contribution to the problem and research domain bodies of knowledge can be summarized as follows:

- **Automated Prognosis Prediction**: Data Engineering + Quantum Machine Learning (QML) + Ensemble models
- **Pulmonary Fibrosis Prognosis**: Prediction automation + Artificial Intelligence (AI) + Data analysis

1.6.1 Contribution to the Research Domain

This project proposes a novel quantum computing approach to PF prognosis prediction and explores QML algorithms, which have not been attempted or explored before, to facilitate the prognosis prediction of pulmonary fibrosis. Using the models produced through the processing of biomarkers, the project also explores the advantage one might gain by utilizing the developing quantum computing over the use of classical computational approaches, which in return will facilitate and encourage more optimization of machine learning using quantum computing.

1.6.2 Contribution to the Problem Domain

Identifies the existing limitations in PF prognosis prediction using medical imaging and biomarkers, often prognosed after manual critical referencing and analysis of biomarkers of the pulmonary cavity of the patient and explores the possibility of automating the process in a more scalable path, in terms of accuracy, rather than using classical computation or manual prognosis.

1.7 Research Challenge

Considering the review of existing work done above, the following can be identified as the most challenging in creating a pulmonary fibrosis prognosis prediction QML model:

- Prediction model selection – Reviewing and selecting suitable quantum machine learning algorithms for predictions and classification.
- Feature identification – Reviewing methodologies available to extract input data parameters from HRCT data.
- Quantum computing access – Exploring available quantum computers and computational tools allowing optimal quantum performance.

Usage of quantum-based data – Exploring the usage of quantum-based feature data extracted from HRCT imagery data and fine-tuning them towards PF prognosis prediction.

1.8 Research Questions

RQ1: What are the existing technologies used in classical and quantum computing methodologies to predict the prognosis of PF?

RQ2: How much further can prognosis prediction of PF be developed and enhanced using classical computing methodologies till it reaches a physical research periphery?

RQ3: How can prognosis prediction of PF using HRCT imagery data be optimized and enhanced to achieve the quantum advantage?

1.9 Research Aim & Objectives

1.9.1 Research Aim

This research aims to design, develop, and evaluate a novel prediction model which is capable of providing accurate and efficient prognosis predictions of pulmonary fibrosis utilizing High-Resolution Computer Tomography data through quantum machine learning.

In elaboration of the aim identified above, this project will be focused and produce a system that can be used to automate the process of predicting the prospective, imminent deterioration in the pulmonary forced volume capacity of a patient diagnosed with PF/IPF with the use of High-Resolution Computed Tomography (HRCT) pulmonary cavity imagery data. To achieve this, patient baseline pulmonary HRCT imagery data and pulmonary FVC data will be established along with further FVC results acquired from either pre- or post-baseline testing. With this research, the review and usage of present pre-processing techniques for HRCT imagery data in quantum computing, the model construct will be capable of handling raw HRCT imagery feed input and analyzing the characteristics to produce a prognosis prediction output in terms of FVC as per a defined period.

Prioritizing the validation or invalidation of the hypothesis, extensive study, and research along with the development of said components and evaluating its performance will be performed. The system is expected to be able to run locally or in hosted environments for personal or public use respectively. The machine learning models, and their codes will be made available as open source to facilitate future research. Finally, following the outcomes and the findings of the research project, a research-paper will be published along with a literature-survey of the existing literature in the domain.

1.9.2 Research Objectives

In order to ensure all the research questions and aims are addressed and achieved lucratively, the following objectives and milestones were identified and defined to set course for the successfully completion of the project.

Table 1.1: Research Objectives

Objective	**Description**	**Learning Outcomes**	**Research Questions**
Problem Identification	Carry out the following tasks to identify the problem. • **RO1**: Research interested domains and identify a potential problem which may be feasible to solve within the limited time constraints and technologies.	L01	
Literature Survey	Carry out an in-depth review of the following areas, • **RO2**: Analyze and understand fundamental concepts of quantum	L01, L04, L05	RQ1, RQ2

	machine learning and understanding training models and working with them. • **RO3**: Conduct preliminary studies into existing prognosis prediction systems. • **RO4**: Study on Computer Tomography to understand the need and technique of using HRCT in PF prognosis. • **RO5**: Analyzing existing models and identifying their limitations. • **RO6**: Critical review of the literature and elaborate on the research gap and methodologies.		
Requirement Analysis	Carry out in-depth user requirement gathering in the following areas, • **RO7**: Understand and gather requirements users may expect from a prognosis prediction system for PF. • **RO8**: Get insight and opinion from Pulmonologists, Radiologists, and quantum data scientists to build the system and mitigate any legal/ social/ ethical issues. • **RO9**: Identifying the tools and techniques (software requirements) and expected behavior for the system through questionnaires.	L01, L02, L03, L06	RQ1, RQ3
Design	Develop the design architecture of the proposed system, capable of solving the gap.	L01	RQ3

	• **RO10**: Design a prognosis prediction system to demonstrate the FVC and HRCT data. • **RO11**: Design a data-preprocessing pipeline for HRCT imagery data feed. • **RO12**: Design the QML prediction model which can produce the prognosis prediction.		
Development	Implement a system that's capable of addressing the gap aimed to solve. • **RO13**: Develop a prognosis prediction system that can predict the prognosis of PF efficiently and quickly. • **RO14**: Develop the QML model that can use quantum super-states to produce quick and efficient predictions. • **RO15**: Develop a pre-processing pipeline for HRCT imagery feed data. • **RO16**: Develop the hyperparameter tuning component that improves the prediction system.	L07, L05	RQ3
Testing & Evaluation	Testing and evaluating the prototype. • **RO17**: Create a test plan and perform unit, integration, and functional testing of the prediction system. • **RO18**: Evaluate how efficient quantum models maybe be in comparison to classical models to perform predictions.	L04	RQ3

	• **RO19**: Perform requirement validation against all requirements and evaluate accuracy measures.		
Publish Findings	Produce and publish well-structured papers that will critically evaluate and review the research area. • **RO20**: Publishing evaluation and testing results of the project system. • **RO21**: Making code and models created during the project open-source and publicly available for future work.	L06, L08	

1.10 Chapter Summary

This chapter delivered a synopsis of the problem the author attempts to solve through this research project along with why PF is such a pressing matter on a global scale being evidently proven along with the gap the author has identified. The aims and objectives of this project were also discussed parallel to the learning outcomes each of those objectives are targeted to achieve. The contributions this project would make to the have also been discusses in detail, along with the novelty of the project, in both solution and the problem. The challenges faced when attempting this research were covered finally.

CHAPTER 2: LITERATURE REVIEW

2.1 Chapter Overview

This chapter aims to review and critically evaluate the existing literature in the domain of quantum computing and pulmonary fibrosis prognosis prediction. With a thorough and deep understanding of PF and quantum computing and machine learning, a literature survey has been written regarding PF and all the existing technologies and algorithms for building state of the art machine learning models using quantum computing, and all the steps that lead up to it including pre-processing techniques, all the way to testing and evaluation. The concept map is also included, which covers a bird-eye view of the research domain, demonstrating how the research gap was derived and all other possibilities to help for future work in similar domains.

2.2 Concept Map

Upon conducting thorough literature surveying across the domain, a mind-map graph was derived of the literature and visualized as a concept map. This graph covers all the existing literature in the prediction domain and shows how the problem has been narrowed down into the gap chosen for this project, including review of the problem domain, technologies, existing similar work, and testing techniques, etc. The map is attached in the appendix section A.

2.3 Problem Domain

2.3.1 Pulmonary Fibrosis & Idiopathic Pulmonary Fibrosis

Pulmonary fibrosis, occasionally prefixed as idiopathic pulmonary fibrosis can be defined as a rare chronic lung fibrosing interstitial pneumonia which may have occurred of an unknown cause, and have radiological and histological features very similar to that of a usual interstitial pneumonia (UIP) (Raghu et al., 2022). Soft lung tissue in the pulmonary cavity is given into the interstice fibrosing, and is left with a fiber like structure, which immensely reduces the lung capacity and function, leaving patients in high-risk of a lot more physiological jeopardies (Huang et al., 2015). Since PF/ IPF is identified as a progressive interstitial pulmonary disorder and giving to its non-specific, chronic interstitial inflammation supported by the extensive collagen deposition, PF is branded as one of the most lethal and rapidly progressing pulmonary disorders related to tissue fibrosis (Antoniu, 2008).

Predominantly present in males (López-Muñiz Ballesteros et al., 2021), pulmonary fibrosis is known to effect older adults the most in the ages 50-70 and above, where most cases have been un-detected with asymptomatic fibrosis, or due to unexplained symptomatic cases, where prefixed idiopathic will be diagnosed (Díaz-Piña et al., 2018). Potential risk factors include bad lifestyles such as smoking (Taskar, 2006), inhalation of wood, metal or silicone particles in work environments (Steele et al., 2005), where mostly men are associated at work and even due to genetic conditions in some cases (Pardo and Selman, 2012). With these numbers, patients are mostly un-fit or will be declared too far-along for any surgical intervention (if-any) such as the only possibility of cure, lung transplant (Lee et al., 2010). The pulmonary tissues completely alter structure, inclusive of the phenotype of alveolar, epithelial cells, etc., and presents only a mean lifespan of 2.5 to 3.5 years after diagnosis (Epstein Shochet et al., 2021). Further, PF is also considered quite novel to its name and age, by being designated its name only in 2003, and the number of cases rising every year, with the most amount of cases recorded yet in 2021 with the discovery of Covid-19, which forced people to go into treatments, where upon inspection and testing was diagnosed with IPF (Perera et al., 2021).

2.3.2 Impact of Pulmonary Fibrosis

According to Díaz-Piña (2018), PF has been categorized as the most lethal and aggressive pulmonary diseases among all other diffusive interstitial pulmonary pneumopathies. With no direct medication or treatment course designated in the state of art practices, patients are left with drug-trials and trial drug courses, which can be brutal, and cause long-lasting, sometimes lethal side-effects which eventually might eradicate the patient set for cure (Lacedonia et al., 2021). In surgical intervention, the only treatment option set of 100% recovery so far is lung-transplantation which is also a very dangerous procedure, and a preposterously long wait-list for the organ transplant (Raghu et al., 2022).

As convened above, PF does not affect anyone and everyone the same way. The symptoms, where applicable, resembles that of very common and curable diseases such as the common cold, or respiratory infection (Raghu et al., 2011). Symptoms for PF may vary tremendously from case to case (Epstein Shochet et al., 2021), and in most cases, the disease begins to develop itself and worsen over time with the development of more scarring in the pulmonary cavity, all while being asymptomatic (Andersson-Sjöland et al., 2008). In the lucky cases where symptoms may show at an early stage, the patient may experience symptoms such as spurts of shallow breathing and shortness of breath (García-Sancho et al., 2011), dry-coughing, constant extreme fatigue even after a good amount of sleep/ rest (Taskar and Coultas, 2008) and even sudden involuntary weight-loss reported in some cases (Patel et al., 2007). On the contrary, however, lung-transplant might well be the only option available. However, as the disease worsens over time, due to either non-treatment or asymptomatic mis-diagnosis, more complex and serious symptoms maybe experienced by patients such as finger-clubbing, where fingertips and/ or toes may look more wider and wider than usual (Morena et al., 2023), and cyanosis, where fair-skinned people may experience bluing of the skin and darker-skinned people may experience graying or whitening of the skin around the mouth and eyes where the skin is the thinnest and high-oxygenated blood-streams pass by, making it clearly visible that the body received too little oxygen from the bloody, which indicates the alveoli's may be obstructed, a key symptom of sever fibrosis in the lung cavity (Pergolizzi et al., 2023).

2.3.3 Prognosis Prediction of Pulmonary Fibrosis

2.3.3.1 Importance of Prognosis Prediction

As portrayed above, PF is a fatal pulmonic condition which deteriorates the quality of life of its victims, rapidly and drastically. The effects of fibrosis, even though never fully reversible, can be managed and treated and managed, so that patients experience some ease instead of constant dis-orientation and discomfort. The following treatment options are the deemed the safest and most effective when providing treatments for PF.

- Oxygen Therapy – Prescribed in sever PF, where the blood oxygen levels may fall below 88%, and to prevent further downfall (Kim et al., 2021).
- Pulmonary Rehabilitation – Allows the lungs to restore the ability to function using healthy lung segments more effectively and allow patients to feel less breathless and allows them to function with more ease and less fatigue (Alamri et al., 2023).
- Drug Therapy – Catered for each specific type of PF, drug therapy trials different accepted drugs on patients to counterattack further fibrosing in the pulmonary cavity. Few of the most common therapies include anti-inflammatory medication, immune system suppressing drugs, which also comes with a list of side effects (Selman, 2006).
- Lung Transplantation – Organ transplant is a viable, where over 2,714 people have undergone lung transplant of either a single lobe, or both lobes in 2019 in the United States alone, of which 35% were for IPF related cases. Though this is considered a viable option, due to its severe risks and unavailability of donations, most patients of IPF/ PF will never receive a fresh-set of lungs through-out their life (Kim et al., 2021).
- Palliative Care – Considered the viable option for final stage IPF, patients are treated for each symptom as they arise, with no direct treatment for IPF (Pergolizzi et al., 2023).

The treatment management listed above, clearly demonstrates the time sensitivity of the prognosis prediction of IPF. Without a proper and valid prognosis prediction, it is impossible to determine how to start treating patients of IPF, accurate to their symptoms are the stage in which they are (King et al., 2014). Prognosis prediction will be able to roughly predict the time-span a patient has left before they drive into lung-failure (respiratory-failure), which is life-threatening and can be fatal in most cases (Zarogoulidis et al., 2023). Prognosis predicted might entirely vary and depend on other crucial factors of the patient such as their age, gender which has a tendency to worsen in males rapidly, smoking status where smoking is an added drive to the progression of the scarring in lung tissue, as well as genetic conditions, discussed in detail above (Astor et al., 2023), however it is possible to determine the likely outcome of the patient's condition, based on past cases and considering notable and distinguishable information such as their gender, smoking status, etc. (Hunninghake, 2014).

2.3.3.2 Medical Apparatus & Biomarkers for Diagnosis & Prognosis Prediction

As discussed above, when suspected of Interstitial Lung Disease, the super domain of Pulmonary Fibrosis, medical professionals will collect as much information as possible of the daily activities of the patient, such as their daily work, personal and genetical medical history, home environment, which may help them identify any potential exposures they may include risk factors for PF (Lee et al., 2023). Doctors may utilize different testing, imaging, and blood work to accurately diagnose PF through other ILDs, which includes the following apparatus.

Pulmonary Function Testing

Lung function testing/ pulmonary function testing is a measurement into the volume of air ones lungs are capable of holding in a single breath (Jagpal, Pistun and Mikhail, 2014). Tissue scarring causes the pulmonary tissue to wither, and stiff, which makes it harder and ultimately unable to make a full expansion, which in return makes it unable hold air and effectively transfer oxygen thought alveoli to the bloodstream (Chuliá-Peris et al., 2022).

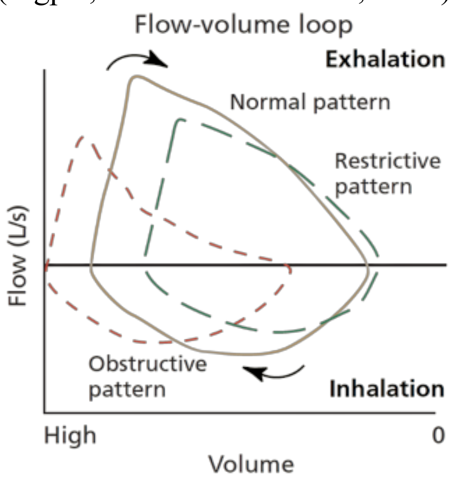


Figure 2.1: Spirometry Graph showing Lung Function.

Figure 2.1, clearly demonstrates the reduction in lung function in air-volume (the expiration and inspiration), where the normal lung volume capacity, indicated in blue, clearly demonstrates a very high-volume capacity, with an expiration flow of over 7.5 – 8.5 L/sec topping at 6 – 7 L of lung volume; and an inspiration flow of over 4 – 5.5 L/sec, whereas obstructed lung cavities, which may indicate chronic obstruction in the pulmonary cavity, may indicate IPF of fibrosing in the lung, while showing a severe reduction in the flow and volume capacity, with just barely 2.8 – 4.5 L/sec of expiration flow and 3.5-5 L/sec of inspiration flow with a maxing out lung-volume capacity of just 5 L, which can be life-threatening (M Liu et al., 2019).

Serology (Blood Work)

Serology is one of the most trusted forms of testing in the world (Hunninghake, 2014) and provides information into the patients' medical history to detect and understand why the fibrosis/ lung scarring may have taken place. This includes identification of diseases such as auto-immune disease which is a known and a highly influential cause of PF, traces of nicotine or drug usage, environmental exposure, which increases the chances of developing scarring in the lung tissue, and makes it possible to give PF a cause rather than define it IPF (Lake, 2014).

High-Resolution Computed Tomography (HRCT)

HRCT, is a specialized version of a X-Ray imagery, which will be focused towards the chest cavity in order to get a clear and unobstructed vision into a patients pulmonary cavity (Qi et al., 2023). Being a more concentrated version of the usual CT scan (also known as the CAT scan), HRCTs are able to give a more close up and clear view of the lung cavity, by showing healthy lung tissue in deep black, and scarred or inflamed lung will be shown in a more lighter scale of grey or white (Lawrence and Nho, 2018). The use of HRCT is preferred in the process of detection of PF, due to attention to simple details such as the whitening explained above, which can go undetected and unnoticed to an untrained eye in the use of a regular CT imaging where not prominent (Raghu, 2006).

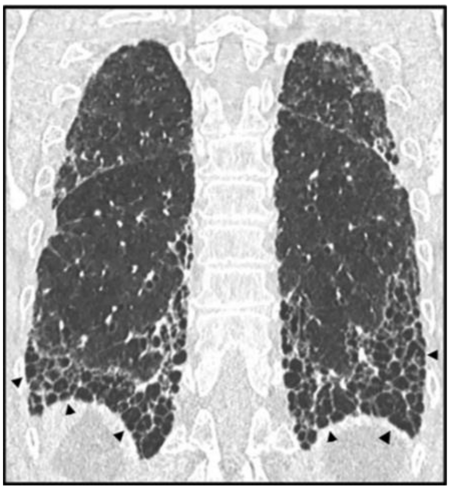


Figure 2.2: Lung HRCT imagery clearly showing Honeycombing.

As figure 2.2 noticeably demonstrates, the lung cavity is clearly shown in black, with a honey-combing effect taking place around the sides of the pulmonary tissue, indicating fibrosing of the lungs, which suggests the presence and diagnosis of IPF, else if diagnosed, this can be given a name. However in most cases, this might have to be classified as cause unknown scarring of the pulmonary cavity, and thus the patient will be diagnosed as Idiopathic PF, and treated according to symptoms and progression (Gui et al., 2018).

2.3.4 Quantum Computing & Machine Learning

2.3.4.1 Quantum Computing

With the rapid emergence in our understanding in quantum physics and mechanics, scientists actively thrive to harness the physical capabilities and aspects of quantum mechanics to enhance and further develop our boundaries of communication and information processing (Biamonte et al., 2017). This has led to the development of quantum computing, which has demonstrated the ability to perform complex logical calculations and problem-solving that would have taken classical computers years to complete (Huang et al., 2021). Quantum computation has been utilized and successfully implemented in numerous real-world applications so far, with outstanding performance being reported with just a few qubits available for use in the state-of-art so far such as simulating exact chemical and matter reactions, optimizing logistics and transport, and even being able to conceptually break-through the classical encryption conventions, that were deemed unbreachable (Astor et al., 2023).

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

Figure 2.3: Qubit State Representation

The fundamental difference between quantum and classical machine learning is the computational unit they use. Classical machine learning utilizes the well-established bits to withhold data in the form on zeros and ones (Simeone, 2022). On the contrary, quantum computers utilize a specialized type of bits, namely qubits to represent data retained. The difference between classical bits and qubits is that while classical bits hold information in either a zero or one at a single given moment, qubits are able to exist in multiple states simultaneously as shown in figure 2.3 (Li, Xiao and Li, 2016). Qubits are controlled, stabilized and synthesized by implementing various physical schemes by using objective phenomena defined and established in quantum mechanics (Cerezo et al., 2022), thus allowing them to achieve a quantum phenomenon called superposition, which allows quantum computers to perform calculations on many different possibilities at a single moment, concurrently, until it collapses into a definite position, which unleashes an immense amount of computational power (Astor et al., 2023). Figure 2.4

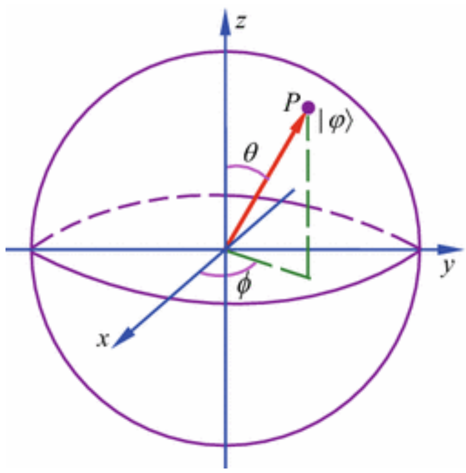


Figure 2.4: Qubit Representation on the Bloch Sphere

represents the conceptual qubit on the Bloch sphere, which provides a graphical representation of the qubit. The mapping can be produced in a sphere by embedding the Cartesian product in a three-dimensional space. Quantum computers also utilize another quantum phenomenon named quantum entanglement. Quantum entanglement is the process in which two or more quantum systems becomes so tightly correlated to each other, which makes it impossible to describe the state of one without the other (Gisin et al., 2002). This plays a vital role in the computational power of quantum computing as well as in quantum mechanics, such as enabling phenomenon such as quantum-teleportation. Entanglement is also vital in making the quantum computer reliable, by enabling the implementation of quantum error correction (Preskill, 2018).

2.3.4.2 Quantum Machine Learning

With the advance in quantum computing, applying machine learning to these powerful computers have always been a topic of discussion and dialogue. Applying machine learning techniques to solve complex problems by developing native algorithms to optimize the use of quantum computers leverages the quantum computational capabilities such as quantum-entanglement and superposition to achieve faster training speeds and inference in machine learning models by simply processing vast amounts of data at the same time (Dunjko, Taylor and Briegel, 2016). The use of quantum computation with machine learning has been modernizing the field of machine learning by providing faster and more accurate predictions and detections, encouraging further research work into the domain (Cao et al., 2019). According to Bernhardt (2019) from MIT, quantum machine learning has been able to, and has been outperforming state-of-the-art classical machine learning models and algorithms, where quantum algorithms used for data clustering has been performed rampantly faster in comparison to its classical counterpart. While quantum computing is still at its very infancy, with these promising and outstanding results, QML would definitely play a huge role in the advancement in ML techniques and development (Lloyd, Mohseni and Rebentrost, 2013).

2.3.5 Achieving the Quantum Advantage for Prognosis Prediction

As demonstrated above, quantum machine learning is more than capable of showing great results, even at its very young age. Quantum phenomenon such as quantum entanglement and superposition of qubits have proven that the use of QML will improve the efficacy of prediction models, thus improving the possibilities of exponential speed-ups over classical ML algorithms in the medical diagnosis and prediction fields, allowing achieving the quantum advantage over classical machine learning approaches (Azevedo, Silva and Dutra, 2022).

2.4 Existing Work

2.4.1 Classical Machine Learning Systems

2.4.1.1 Fibro-CoSANet using Convolutional Self Attention Network

The Fibro-CoSANet model's performance should be compared to previous approaches. The paper's authors claim that their model beat previous approaches like CNNs and LSTMs. However, recent research by Furukawa et al., (2022) demonstrated that deep learning methods such as CNNs and transfer learning approaches may produce good results in the detection and segmentation of pulmonary fibrosis. As a result, before forming judgments regarding the superiority of a certain methodology, it is critical to analyze the strengths and shortcomings of several methodologies. The dataset used to train and verify the Fibro-CoSANet model, however, should be thoroughly scrutinized (Inomata et al., 2021). The researchers utilized the publicly accessible Labeled Imaging Data Repository (LIDC-IDRI) dataset, which comprises CT images of lung nodules. However, this dataset might not be typical of the pulmonary fibrosis patient group. A recent study by Zhou et al., (2022) shown that using a broad and balanced dataset is critical for attaining good results in medical image analysis.

2.4.1.2 Prediction Analysis of IPF Progression from OSIC Dataset

Using the OSIC dataset, Mandal et al., (2020) used machine learning methods to predict the evolution of idiopathic pulmonary fibrosis (IPF). However, they did not apply convolutional neural networks or attention processes, which have been proven to be beneficial in previous experiments. Furukawa et al., (2022) conducted a review of the usage of deep learning-based techniques in pulmonary fibrosis detection and segmentation. CNNs have been successful in detecting and segmenting lung lesions from computed tomography images, and they have also shown potential in predicting disease progression, according to scientists.

Zhang et al. (2021) employed an ensemble of deep CNNs to classify medical images, demonstrating great accuracy in distinguishing between healthy and sick lungs. CNNs have showed potential in a variety of medical image classification tasks, including lung disease detection, according to the scientists.Liu et al. (2020) did a meta-analysis comparing the performance of deep learning algorithms in diagnosing illnesses from medical imaging to that of healthcare experts. Deep learning algorithms performed similarly to healthcare professionals and may give advantages in some circumstances, according to the researchers.

2.4.1.3 Fibrosis-Net: Deep Convolutional Neural Network

Wong et al., (2021) Fibrosis-Net is a customized deep convolutional neural network built for predicting pulmonary fibrosis development from chest computed tomography (CT) data. Guan et al. (2019) employed a deep learning strategy for feature extraction and classification, as opposed to Mandal et al. (2020), who used typical machine learning algorithms. Furukawa et al.,)2022) found that deep learning-based techniques, such as CNNs, were effective in pulmonary fibrosis detection and segmentation from CT images. They also said that 3D CNNs, which were employed in Fibrosis-Net, performed better in medical picture processing than 2D CNNs.

Another research, (G Liu et al., 2019a), created a deep learning-based algorithm for predicting disease progression in IPF patients using CT images. The authors used a 3D CNN architecture similar to Fibrosis-Net to predict IPF development and discovered that their technique beat existing machine learning algorithms. Using a pre-trained 3D CNN, Li et al. (2020) suggested a transfer learning technique for pulmonary fibrosis detection from CT images. They observed great accuracy in their technique, which may give benefits in data efficiency and training time.

2.4.2 Quantum Machine Learning Systems

2.4.2.1 Quantum Particle Swarm Optimization

Rathore et al. (2020) used computed tomography images at baseline to predict progression in idiopathic pulmonary fibrosis. While Rathore et al. (2020) employed a machine learning technique similar to Mandal et al. (2020), their work used QPSO, a metaheuristic optimization algorithm commonly used in machine learning applications. Kaur et al. (2021) employed a similar strategy to predict IPF development using CT scans, but instead of QPSO, they optimized the hyperparameters of a gradient boosting model using an Artificial Bee Colony (ABC) algorithm. The authors claimed that their method was very accurate, which might benefit in the early detection of IPF. Yang et al. (2020) developed a deep learning-based strategy for predicting IPF development from CT images using a 3D CNN architecture in their study. The authors found that their technique had good accuracy and specificity, which might provide advantages over typical machine learning algorithms. In contrast, Flaherty et al. (2020) used radiometric characteristics taken from CT images to predict the course of IPF. According to scientists, radiometric characteristics might help predict IPF development and identify individuals who would benefit from early management.

2.4.2.2 Quanvolutional Neural Network

QCNNs are a sort of neural network that use quantum mechanics concepts to boost performance in a variety of machine learning applications. Although the notion of quantum convolutional neural networks is still in its early stages, various research have highlighted its potential advantages over traditional convolutional neural networks (CNNs) in image classification tasks (Houssein et al., 2022) and quantum chemical applications. However, one of the major obstacles of QCNNs is the scarcity of quantum hardware. Because current state-of-the-art quantum computers have a restricted number of qubits, training and deploying QCNNs in real-world applications is difficult (Romero, Olson and Aspuru-Guzik, 2017). Furthermore, QCNNs need substantial expertise in both quantum physics and machine learning, both of which are currently in short supply. Future research should concentrate on producing scalable quantum hardware, enhancing interpretability, and comparing the performance of QCNNs in diverse applications to that of classical neural networks.

2.5 Technological Review

2.5.1 Proposed Architecture of the Prognosis Prediction Model

After thorough investigation into the existing work, the author had come up with the following proposed architecture for Fibro-QuanNet. The steps have been discussed in further detail and precision in the technological review section. Starting from the pre-processing techniques, all the technologies existing and utilized during the last 5 years have been brough into attention and been discussed in detail. Finally, the benchmarking and evaluation techniques usable in the project has been discussed thereafter.

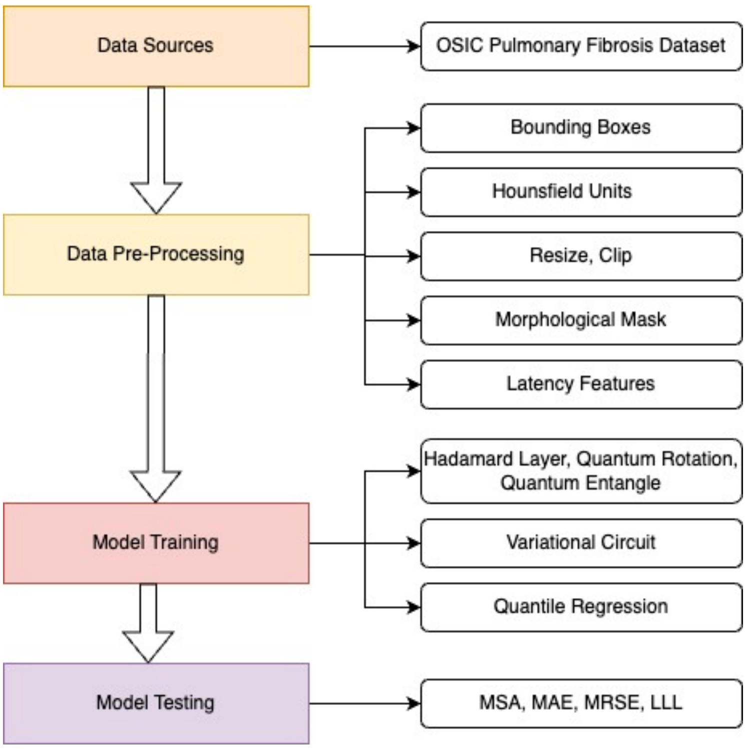


Figure 2.5: Proposed Architecture of Fibro-QuanNet

2.5.2 Pre-Processing Techniques

The first step of the prediction model, as depicted above in figure 2.5 is to pre-process the input DICOM images and the tabular data in order to pass them as computer-comprehensible data to the prediction algorithm, discussed in further details over chapter section 2.5.3. Fibro-QuanNet would feature an array of pre-processing stages, to successfully convert these input imagery and tabular information into the input torches as projected.

2.5.2.1 Data Annotation

Data annotation is crucial in developing machine-learning models for object detection tasks. The computer vision community widely uses bounding box annotation among the various annotation techniques. In this technique, a rectangular box is drawn around the object of interest in an image, and the box coordinates are recorded as labels for the training data. However, the accuracy of the bounding box annotation depends on the annotator's expertise, leading to the problem of annotation inconsistency. For instance, in research by Cai et al., (2019), the authors discovered that bounding box annotations can vary greatly amongst annotators. This can lead to inconsistencies in the quality of the training data and have an impact on the performance of the trained models. To address this issue, several researchers have advocated using crowd-sourcing strategies for data annotation (Welinder et al., 2010). Crowdsourcing, on the other hand, might pose additional issues, such as the requirement to assure the quality and trustworthiness of annotations contributed by non-expert annotators.

2.5.3 Hounsfield Units

Hounsfield units are an important metric in DICOM picture pre-processing, notably in medical imaging. Hounsfield units are a unit of radiodensity, which is a material's capacity to absorb x-ray radiation. Hounsfield units are used in DICOM pictures to measure the attenuation of x-rays as they travel through different tissues and structures in the body, allowing for the differentiation of diverse anatomical aspects. The segmentation of medical pictures, which is the act of distinguishing various sections of an image depending on their attributes, is one major use of Hounsfield units. Guo, Schwartz and Zhao, (2019) employed Hounsfield units to segment CT images of the liver to detect liver cancers in research. The scientists discovered that utilizing Hounsfield units in combination with other image processing approaches enhanced tumor identification accuracy when compared to using Hounsfield units alone. Hounsfield units are also employed to repair picture artefacts produced by non-uniformity in the x-ray beam or the imaging equipment itself. (Wang, Xu and Shi, 2018) employed Hounsfield units to adjust for beam-hardening artefacts in CT images in their investigation.

2.5.4 Morphological Masking

A morphological mask, also known as a binary mask or morphological filter, is a method extensively used in image processing for picture segmentation. It entails producing a binary picture that emphasizes certain regions of interest within an original image depending on their form, size, and other morphological properties. The binary mask that results can then be used to isolate sections of interest for additional examination or alteration. Morphological masks have a medical imaging use, specifically in the processing of MRI and CT data. Tiwari and Singhai, (2017) employed morphological masks to segment brain pictures for the goal of producing 3D models of the brain in research. The scientists discovered that morphological masks were successful in segregating distinct anatomical components, such as gray matter, white matter, and cerebrospinal fluid, allowing for precise brain modeling and visualization. Thus, morphological masks are a popular image processing approach for segmenting and isolating certain regions of interest within an image. They have several uses in sectors like as medical imaging and computer vision and have been demonstrated to improve the accuracy of different image analysis techniques.

2.5.5 Algorithm Selection

2.5.5.1 Quantum Particle Swarm Optimization

Quantum Particle Swarm Optimization (QPSO) is an optimization technique that blends quantum mechanics with swarm intelligence principles. Initially proposed in 2006, QPSO as an extension of classical Particle Swarm Optimization (PSO) for application in quantum computing (Li, Zhan and Zhang, 2022). Due to its encouraging results in several optimization situations, QPSO, a relatively new optimization approach, has drawn attention recently. To find the best answer in a high-dimensional search space, QPSO combines the ideas of quantum mechanics with swarm intelligence. One of QPSO's key benefits is its capacity for global search, which enables it to look for a problem's global optimum rather than becoming bogged down in local optima(Qi et al., 2020).

The optimization of artificial neural networks is one area where QPSO has been used. For the categorization of medical pictures in research by Mohan and Nair (2021), the authors utilized QPSO to optimize the weights of a neural network. In comparison to other optimization methods like the Genetic Algorithm and PSO, the authors discovered that QPSO was able to obtain a greater classification accuracy. The optimization of quantum circuits for quantum computing is another use for QPSO. A quantum circuit's characteristics were optimized for the goal of preparing a quantum state in research by Goldin et al., (2019) using QPSO. Comparing QPSO to other optimization techniques, the authors discovered that it was more effective and fidelity oriented. The principles of quantum mechanics are employed to update the velocity and location of particles in the search space. The location of a particle indicates a possible solution, while its velocity denotes its direction of travel. The quantum gate sequence is used in controlling the velocity and position modifications of the particles. The gates operate as operators, modifying the particle states, and their sequence is adjusted to produce the highest performance in the search space (G Liu et al., 2019b). In the search space, each particle in QPSO is represented by a location vector and a velocity vector. The location vector represents a possible solution to the problem, whereas the velocity vector represents the direction and amplitude of the particle's travel. At each iteration of the method, the velocity and position of each particle are updated depending on the optimal position discovered by the particle and its neighbors.

$$v_i = Wv_i + c_1r_1\left(P_{best,i} - x_i\right) + c_2r_2\left(g_{best} - x_i\right)$$
$$x_i = x_i + v_i$$

Figure 2.6: Particle Swarm Optimization Equation

The equation depicted in figure 2.5, where v_i(t) is the particle's velocity at time t, w is its inertia weight, c_1 and c_2 are its cognitive and social learning factors, respectively, r_1 and r2 are random numbers between 0 and 1, p_i(t) is its best position so far, g_i(t) is its best position according to its neighbors, and x_i(t) is its current position.

2.5.5.2 Quanvolutional Neural Networks

Convolutional neural networks (CNNs) and quantum mechanics are two different types of neural networks that work together to create Quanvolutional neural networks (QCNNs). They use the characteristics of quantum mechanics to improve CNN performance in tasks like image identification and classification. They are designed to function on quantum data, such as quantum pictures or quantum states (Acampora and Schiattarella, 2021). A high-level description of how QCNNs work is that they execute a number of quantum operations on the input quantum state or picture, then use classical convolutional operations to extract features from the output quantum state. Quantum operations are made to take advantage of entanglement and superposition, two quantum features, to encrypt more data into the input state than conventional processes can (Al-jumaili et al., 2022).

The three primary parts of the QCNN architecture are the quantum layer, the classical layer, and the pooling layer. The Hadamard and phase gates, which enable the encoding of quantum information into the state, are examples of quantum operations the quantum layer conducts on the input quantum state or picture. The resultant quantum state is then processed by the classical layer using traditional convolutional techniques to extract features. In order to decrease the dimensionality of the feature maps, the pooling layer down samples the data. The efficiency with which QCNNs can analyze and categorize quantum data is one of its main features. Due to their enormous dimensionality and requirement for a significant number of conventional processing units, traditional CNNs have difficulty handling quantum input effectively. However, QCNNs may effectively handle quantum data with fewer classical processing units by taking use of quantum mechanics' features like entanglement and superposition.

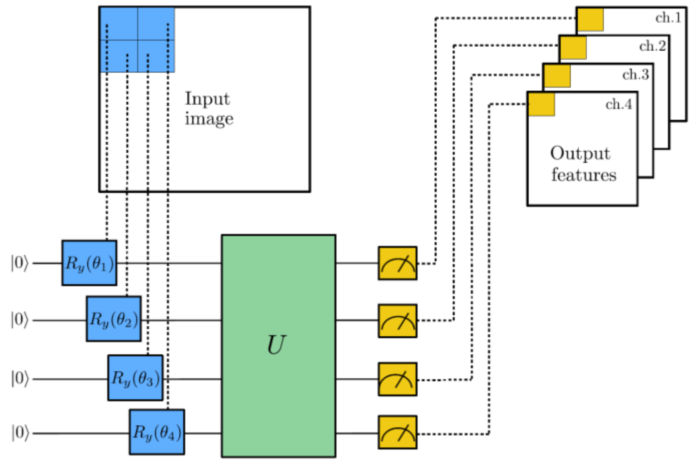


Figure 2.7: Typical Architecture of a Quanvolutional Neural Network

2.5.5.3 Quantum Convolutional Neural Networks

In order to carry out machine learning tasks, quantum neural networks, a particular sort of neural network, utilize quantum computing concepts. QNNs process information using the principles of quantum physics, in contrast to traditional neural networks, which depend on classical computing concepts. One of the main distinctions between quantum neural networks and quantum neural networks is that the former is more general-purpose, whilst the latter are especially made to handle quantum pictures or quantum data. In general, QNNs work by encoding the input data into a quantum state, then manipulating and extracting information from the quantum state via a sequence of quantum operations. The quantum operations employed in QNNs are created to take use of the superposition and entanglement aspects of quantum mechanics to perform calculations that would be challenging or impossible for classical neural networks. The output is then acquired after measuring the ensuing quantum state (Amin et al., 2022).

Quantum Boltzmann machines, quantum Hopfield networks, and quantum variational autoencoders are a few examples of QNNs (Cong, Choi and Lukin, 2019). Different machine learning problems, such classification, clustering, and feature extraction, are addressed by each form of QNN. Although QNN research is still in its infancy, several encouraging papers have recently been published. For instance, Beer et al., (2020) showed that a straightforward QNN may be trained for classification tasks using a quantum computer. Wan et al. (2020) suggested a QNN-based framework for unsupervised learning tasks in another study. In conclusion, quantum neural networks are a promising field of study that apply the ideas of quantum mechanics to machine learning challenges. QNNs have the potential to dramatically enhance the performance of machine learning algorithms in a variety of applications, even though research is still in its early phases.

2.5.5.4 Quantum Generative Adversarial Networks

QGANs are a sort of quantum machine learning technique that combines the ideas of generative adversarial networks with quantum computing principles. QGANs are intended to create new quantum data that is comparable to a training set of quantum data and may be used for a range of tasks including quantum data reduction and synthesis. QGANs are made up of two neural networks: a generator network and a discriminator network. The generator network is taught to generate quantum data that is comparable to the training data, whilst the discriminator network is trained to differentiate between the generated quantum data and the actual training data (Beer et al., 2020b).

QGANs employ quantum mechanics concepts to conduct some calculations more efficiently than conventional GANs. QGANs, for example, may use quantum entanglement features to produce vast volumes of data in parallel, whereas traditional GANs must generate data sequentially. Furthermore, QGANs may use quantum algorithms for optimization and matrix operations, which can accelerate training. QGANs employ quantum mechanics principles to conduct specific calculations more efficiently than classical GANs. For example, QGANs may use the features of quantum entanglement to produce enormous volumes of data in parallel, whereas traditional GANs must generate data sequentially. Furthermore, QGANs may use quantum algorithms for optimization and matrix operations, which can accelerate the training process (Kulkarni, Pawale and Kharat, 2022).

2.5.5.5 Deep Convolutional Neural Networks

Deep Quanvolutional Neural Networks are a sort of quantum machine learning technique that combines deep neural network principles with Quanvolutional neural network principles. DQNNs are intended to categorize quantum data and extract quantum features. DQNNs are made up of many layers of Quanvolutional filters that extract quantum characteristics from incoming quantum data. These characteristics are subsequently processed by numerous layers of fully connected neural networks that perform classification tasks. DQNNs, like QNNs, use quantum physics concepts such as superposition and entanglement to conduct certain computations more effectively than classical DNNs.

Several intriguing research on DQNNs have been published in recent years. For quantum state classification problems, Cong et al. (2020) presented a DQNN architecture. They proved that DQNNs can categorize quantum states more accurately than classical machine learning systems. (Karthikeyan and Priyakumar, 2022) suggested a DQNN technique for quantum chemistry applications in another paper. They demonstrated that DQNNs can forecast the energy of quantum chemical systems accurately. DQNNs have the potential to vastly increase the efficiency and accuracy of quantum machine learning algorithms in a variety of applications. However, research in this subject is still in its infancy, and more study is required to fully explore the possibilities of DQNNs.

2.5.5.6 Hough Algorithm

The Hough transform is a computer vision algorithm that detects fundamental forms in digital pictures such as lines, circles, and ellipses. The method operates by mapping the picture space into a parameter space and recognizing points or areas in that space that match to the forms being sought. Paul Hough invented the Hough transform in 1962 as a way for automatically recognizing tracks in particle physics experiments, and it has since become widely used in computer vision and image processing applications (Zhang, Wang and Sun, 2017). The Hough algorithm has been intensively investigated in the computer vision field, with several modifications and extensions offered over the years. Gisin et al., (2002) for example, presented the Randomized Hough Transform (RHT) as a more efficient variant of the Hough transform that randomly picks points from the image space to minimize computing complexity. Matas, Galambos and Kittler, (2000) presented the Progressive Probabilistic Hough Transform (PPHT) as a technique for identifying lines in real-time applications that gradually collects evidence for line segments in the picture space.

The Hough transform is utilized in a variety of computer vision applications, including object detection, picture segmentation, and feature extraction. (Potempa and Porebski, 2022), for example, employed the Hough transform to locate and track football players in video footage and obtained great accuracy in player localization and tracking. (Mittal and Rajam, 2020) employed the Hough transform to extract the outlines of blood vessels in retinal pictures and established the method's efficiency in identifying diabetic vascular abnormalities.

2.5.5.7 Classical-Quantum Quantile Regression Hybrid Model

Classical-quantum quantile regression models are a form of machine learning method that predicts the quantiles of a target variable by combining classical and quantum computing approaches. These models are based on the quantum machine learning concept, which employs quantum computing principles to improve the performance of conventional machine learning models (Azevedo, Silva and Dutra, 2022). The classical-quantum quantile regression model works by training a classical regression model on the target variable's low quantiles and a quantum regression model on the high quantiles. Combining the two models yields the final prediction. The model is trained using a hybrid classical-quantum computer, which is made up of both a classical and a quantum computer.

Several recent research have investigated the efficacy of classical-quantum quantile regression models. In one research, Ding et al. (2021) created a classical-quantum quantile regression model for predicting stock market volatility. In terms of prediction accuracy and resilience, the model beat standard regression models. However, classical-quantum quantile regression models have certain disadvantages. One of the most challenging tasks is implementing and optimizing the hybrid classical-quantum algorithm. Another issue is the requirement for specialized gear, such as quantum computers, which may be scarce.

2.5.6 Hyper Parameter Tuning

In order to improve the performance of a machine learning model on a particular task, hyperparameter tweaking is a crucial stage in the training process. For hyperparameter tweaking, several methods have been put forth, each having pros and cons.

2.5.6.1 Grid Search

Grid search is a straightforward and obvious approach for hyperparameter tuning that includes setting a range of values for each hyperparameter and then assessing the model's performance for all combinations of hyperparameters in the grid. The benefit of grid search is that it is exhaustive and can locate the best hyperparameters given enough computer resources. The drawback of grid search is that it may be computationally expensive, especially when the number of hyperparameters and their ranges is enormous.

2.5.6.2 Random Search

Random search is a more efficient alternative to grid search, which includes randomly picking values for each hyperparameter from their corresponding ranges and assessing the model's performance for a given number of iterations. The benefit of random search is that it is less computationally costly than grid search and can frequently identify suitable hyperparameters with less iterations. However, the downside of it is that it may overlook crucial parts of the hyperparameter space, especially when the ranges are not well-defined.

2.5.6.3 Bayesian Optimization

Using a probabilistic model, such as Gaussian processes or tree structured Parzen estimators, Bayesian optimization is a probabilistic method of hyperparameter tuning. In comparison to grid search or random search, Bayesian optimization has the benefit of efficiently exploring the hyperparameter space and converging to the ideal hyperparameters with fewer evaluations. The drawback of Bayesian optimization is that it uses more computer resources to train the probabilistic model, and the performance of the model may be affected by the selection of hyperparameters for the probabilistic model.

2.5.7 Evaluation Review

2.5.7.1 Mean Squared Error

The average squared difference between the predicted and actual values is calculated using the MSE measure, which is frequently used to assess the effectiveness of regression models. The benefit of MSE is that it penalizes large errors more severely than minor errors and offers a reliable indicator of the model's overall correctness. MSE's sensitivity to outliers and potential inability to offer a clear understanding of the model's performance are its drawbacks.

2.5.7.2 Root Mean Squared Error

Another often employed statistic, RMSE, determines the average squared difference between the expected and actual values. As RMSE is in the same unit as the response variable, it has the advantage of being simpler to comprehend than MSE. However, RMSE has the drawback of being sensitive to outliers and maybe not giving a clear view of the model's performance.

2.5.7.3 Laplace Log Likelihood

Laplace log-likelihood is a machine learning regularization approach intended to prevent model overfitting. It augments the log-likelihood function with a penalty term proportional to the sum of the absolute values of the model's parameters. This penalty promotes the model's parameter values to be sparse, i.e., many of the parameters are set to zero, resulting in a simpler model that is less likely to overfit the training data. One of the key advantages of Laplace log-likelihood is that it can handle high-dimensional data with numerous characteristics, which in standard models can lead to overfitting. The L1 penalty in Laplace log-likelihood efficiently performs feature selection and reduces model complexity by shrinking the coefficients of less significant features to zero.

2.6 Chapter Summary

This chapter reviewed and critically evaluated the existing literature in the domain of quantum computing and pulmonary fibrosis prognosis prediction. With a thorough and deep understanding of PF and quantum computing and machine learning, a literature survey has been written regarding PF and all the existing technologies and algorithms for building state of the art machine learning models using quantum computing, and all the steps that lead up to it.

CHAPTER 3: METHODOLOGY

3.1 Chapter Overview

This chapter aims to bring the methodologies the author adhered to through-out the research project. The research methodologies, development methodologies, data-elicitation methodologies, project management methodologies used in the project has been discussed with supporting justifications for each selection made for each methodology. Finally, the project deliverables, Gantt chart, software, hardware, data, and technical requirements have also been reviewed, along with the potential risks associated and corresponding mitigation plans in place.

3.2 Research Methodology

In order to efficiently manage cost, time and scope, the key factors governing the quality of any project, the Saunders Research Onion Model (Saunders, Lewis and Thornhill, 2007) has been used to construct the appropriate methodologies for this research project.

Table 3.1: Research Methodologies

Research Philosophy	The research philosophy affects the nature in which data is collected and analyzed due to its relation to the investigation of reality. Among the different philosophical paradigms available and discussed, **Pragmatism** has been selected and used in this research, giving mostly to its involvement of both quantitative and qualitative data.
Research Approach	The research approach discusses the approach taken by the researcher when performing the research. For this research, the **Deductive** approach has been followed over the alternative Inductive approach since the research aims to test and prove the hypothesis and expected to be quantitative.
Research Strategy	The research strategy discusses the data collection strategies adhered to when collecting data to answer the research questions. Among the different strategies available, **Surveys** has been chosen as the primary focus since it shed light to a more quantitative

	approach. The target audience being in the medical profession, it would be quite beneficial to conduct interviews with them as well.
Research Choice	The research choice focuses on the choice of the methodologies selected and identifies the qualitative and quantitative characteristics of the research. The **Mixed method** has been selected as the research choice for this research due to its qualitative and quantitative character. Both surveys as well as interviews conducted with the target audience will be used to defend the hypothesis.
Time Horizons	The time horizon selected for this research will be **Longitudinal** since required data will be gathered and used prior and after the development of the system.

3.3 Development Methodology

3.3.1 Life-Cycle Model

The software development life cycle governs the continuous progression of the project from its initiation to the very end. Off the available life cycle methodologies, the **Agile** Software Development life cycle has been chosen for this research as the research development method giving to its adaptability to rapid changes in the requirements of the project.

3.3.2 Requirement Elicitation Methodology

Requirements of the project will be gathered through medical practitioners and PF patients using various methods of data collection such as **questionnaires and interviews** and will be analyzed to identify the requirements for this project as well as for the betterment of the domain application.

3.3.3 Design Methodology

Object-Oriented Analysis & Design (OOAD) has been selected as the design paradigm for this research due to its alignment and focus on data structures and tendency to relate to real-world objects, this object orientation, which may additionally also enable and help in the future scaling and re-usability of the models and applications built through this project.

3.3.4 Development Methodology

This project uses of **Object-Oriented Programming (OOP)** over its counterparts such as structured and functional programming due to its tendency to orient the application towards object orientation and thus be easier to handle and scale in the future.

3.3.5 Evaluation Methodology

Classification matrix evaluation techniques such as Confusion Matrix, ROCc (Receiver Operating Characteristic curve) and the Laplace Log will be used in this project. The ROC curve and the Laplace Log will play a key-role in producing a confidence score which is a crucial factor in medical predictions.

3.3.6 Solution Methodology

Similar to classical machine learning, quantum machine learning also requires several steps be taken in order for a successful model production. The architecture of the QML prediction model has been planned as follows.

For the data sources, the author will be utilizing the OSCI public dataset which includes a large sum of HRCT imagery for the training and the testing of the model. The application further aims to use the MONAI library to pre-process the HRCT imagery data, following which a QCNN will be utilized as the quantum net after a series of quantum gates, while version controlling using version controlling software discussed below.

Once the machine learning models are functional, CM and ROCcs will be utilized evaluate the performance of the models and an API will be developed to enable communication between the GUI application and the back-end model. The application will be deployed in a CI/CD pipeline to continuously integrate changes to the systems and keep up-to date.

3.4 Project Management Methodology

Adhering to a project timeline, scoping and costs plays a crucial part in any project. Therefore, selecting a suitable management methodology plays a key-role in the project.

Accounting for all opportunities and take-aways from methodologies available, **Prince2** has been selected to make sure the project is product oriented and will be flexible to any fluctuations that may occur and will need to be adjusted accordingly as well as to compartmentalize the work.

3.4.1 Schedule

3.4.1.1 Gantt Chart

The Gantt chart has been attached in the appendix section B.

3.4.1.2 Project Deliverables and Dates

Table 3.2: Project Deliverables & Dates

Deliverable	Date
Project Proposal Document The initial proposal of the project.	3rd November 2022
Literature Review Document The Critical review of existing work and solutions.	14th November 2022
Software Requirement Specification Document The document specifying requirements to be satisfied and developed as the final prototype and means of collecting data.	24th November 2022
System Design Document The document specifies the design developed for the Prognosis Prediction system and overviews of the algorithms to be developed.	1st December 2022
Prototype The prototype with main core features functional.	5th December 2022
Thesis The final report documenting the project and research process.	23rd March 2023
Review Research Paper A research paper reviewing existing systems in the prognosis prediction domain published at a conference.	30th March 2023
Final Research Paper A research paper introducing the prognosis prediction system developed at the end of this project.	27th April 2023

Public Package / Library Public library to access the prognosis prediction system developed.	30th April 2023

3.5 Resource Requirements

The following are the software, hardware and data resources required in order to successfully navigate through the research project, as well as delivering expected solutions based on the objectives.

3.5.1 Hardware Resources

- **Intel Core i7 (10th Gen)/ Apple Silicone M1 Pro or higher** – To perform heavy and intense QML processing and computational tasks.
- **16 GB RAM or higher** – To multi-task smoothly with large DICOM datasets and QML integrated models running.
- **Disk space of 40 GB or higher –** To store project related documentation, trained models, source code, dataset, and other data.

3.5.2 Software Resources

- **Operating System (Windows/ macOS/ Linux) –** An operating system that is able to handle the computational demand of Quantum Computing and run/ support language and framework needed in the development process. macOS will be used by default.
- **IBM Quantum/ Google Quantum –** A publicly available remote quantum device is crucial to process the quantum machine learning project. IBM Quantum will be used by default for this project.
- **Python** – An all-purpose language widely used and compatible with many quantum machine learning projects. Python will be used to develop QML models.
- **Qiskit/ Cirq Frameworks** – Open-source framework working for each IBM and Google quantum devices respectively. Qiskit will be used here by default.
- **PennyLane Python Library** – Python library used to support the development, training, and testing of the QML models.
- **NodeJS –** The API that will be needed to communicate with QML backend and the frontend.
- **JavaScript (ReactJS)** – Front-end application where the prognosis prediction will be initiated, and results are to be shown.

- **JB DataSpell/ Visual-Studio Code/ Jupyter Notebook** – Integrated Development Environments used to support the development.
- **Google-Colabotory/ IBM Quantum Lab** – Cloud IDEs to support the training and testing of the QML models.
- **Zotero/ Mendeley** – Reference management tool used to keep track of research artifacts and manage references. Zotero is used in this project by default.
- **MS Office/ MS Sharepoint/ Google Docs/ Canva/ OmniPlan –** Services to store, design, develop and document the research process.
- **Google Drive/ GitHub –** File back-up for documents and source code.

3.5.3 Technical Skills

- Understand the fundamentals of quantum computing and quantum machine learning.
- Ability to successfully build, train and test QML models relevant to prognosis prediction.
- Ability to document technical and scientific writing required for the research.

3.5.4 Data requirements

- Prognosed pulmonary HRCT imagery with their FVC capacity data and other supporting patient metadata such as gender, smoking status, age, etc.

3.6 Risks & Mitigation

Table 3.3: Risk Mitigation Plan

Risk Item	**Severity**	**Frequency**	**Mitigation Plan**
Corruption/ access restriction to development source code	5	5	Keeping a back-up of all source code on Version Controlling (GitHub) and external physical drives.
Deletion/ corruption of documentation	5	4	Using the cloud-first approach through tools such as Share-point and regular local back-ups.

Inability to meet deliverable deadlines within the expected time span	5	4	Making sure personal deadlines are set and met as per the Gantt structure to make sure smooth and complete coverage of all deliverables.
Lack of access to quantum hardware	5	5	Making sure multiple cloud quantum computational solutions are explored and ready to use when needed and essential.
Keeping in-par with the latest technological improvements in QML	4	5	Making sure constant literature survey is performed to ensure constant refresh of the latest advancement of the technologies applicable and hardware capacities.

3.7 Chapter Summary

This chapter discussed in detail the methodologies the author had adopted through-out the research process. From the research methodologies adhered, development methodologies, data-elicitation methodologies, evaluation, and solution methodologies to the project management methodologies followed in this project have been discussed, justifying each decision made throughout. The Gantt chart of the project, along with the project deliverables have also been shown, followed by the hardware, software, data, and technical resources requirements needed to successfully complete the project. Finally, potential risks have been identified and relevant mitigation plans were laid to moderate such complications that may arise through the process.

CHAPTER 4: SOFTWARE REQUIREMENTS SPECIFICATION

4.1 Chapter Overview

This chapter focuses on identifying and analyzing all the potential stakeholders of the project and the system developed through it by identifying potential beneficiaries through a rich picture diagram. Further, mechanisms to collect and analyze perspective to produce the context diagram, use cases and requirements have also been discussed.

4.2 Rich Picture

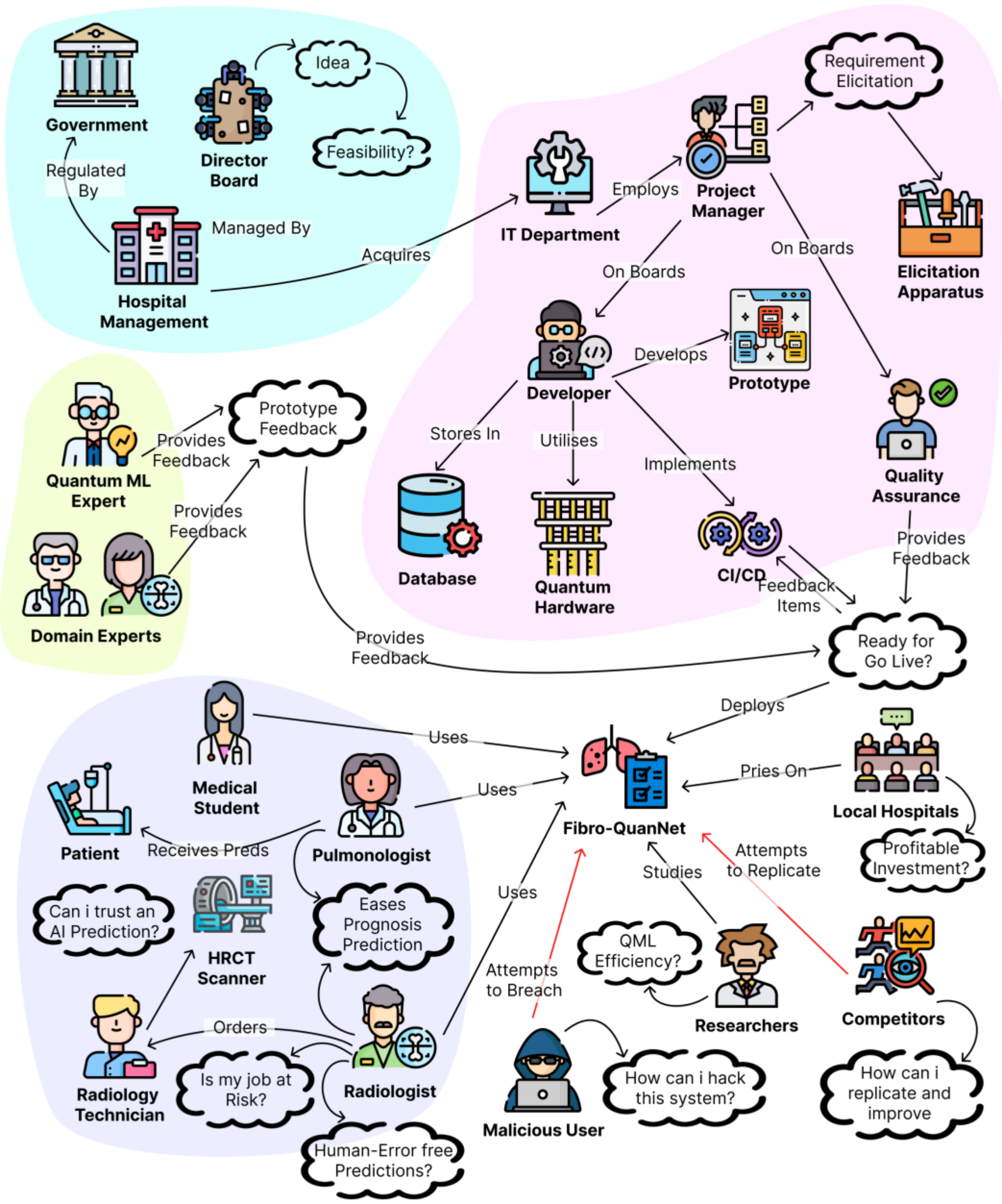


Figure 4.1: Rich Picture Diagram

The rich picture diagram above (Figure 4.1) provides a bird view of all the potential stakeholders/ beneficiaries of Fibro-QuanNet. The diagram outlines most possible implications and how the expected behavior of the system would be of benefit.

4.3 Stakeholder Analysis

The stakeholder (SH) analysis has been illustrated using the stakeholder onion model in Figure 4.2, which represents the stakeholders in association with the system in co-response to their respective environments followed by each stakeholder's unique and valuable viewpoint being discussed under the Stakeholder Viewpoints (Chapter 4.3.2).

4.3.1 Stakeholder Onion Model

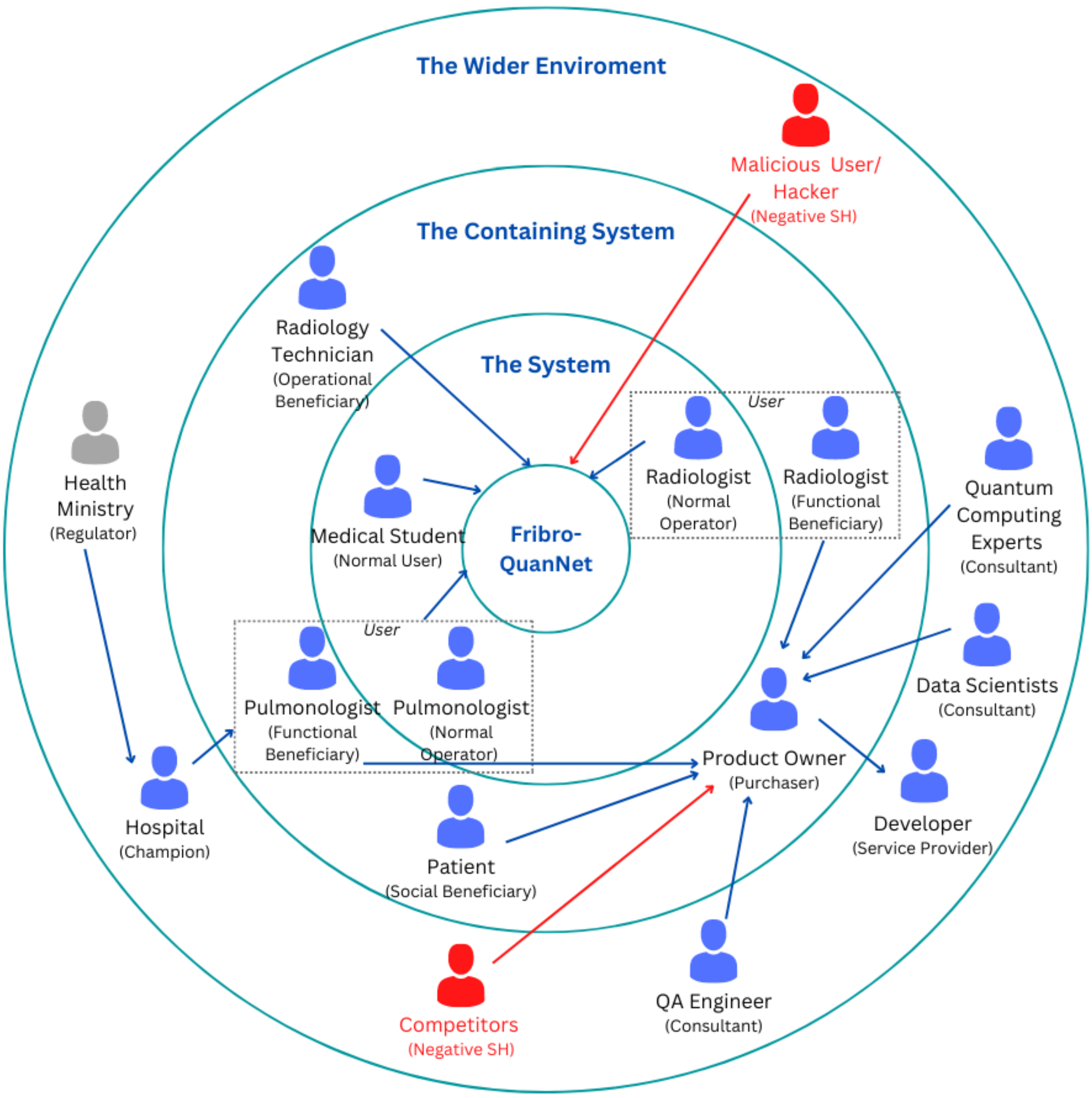


Figure 4.2: Stakeholder Onion Model

4.3.2 Stakeholder Viewpoints

Table 4.1: Stakeholder Viewpoints

Stake Holder	**Role**	**Benefits/ Role Description**
Hospital Admin/ Technology Officer (Product Owner)	Purchaser	Makes sure the application is functional and organizes while gathering and communicating requirements from the functional beneficiaries to the financial beneficiaries and the developer.
Pulmonologist	Functional Beneficiary & Normal User (Hybrid User)	Can use the system to produce more accurate prediction of the progression/ prognosis of pulmonary fibrosis in the patients the pulmonologist gets.
Radiologist		Can use the system to produce accurate reports of the prognosis when referred to by pulmonologists after performing the initial HRCT.
Patient	Social Beneficiary	Will receive accurate prognosis predictions much faster and accurately than usually would manually, which may provide time for proper clinical drug trials.
Medical Students	Normal User	Can use the application to produce accurate prognosis predictions of PF as well as communicate requirements and changes directly to the normal operators received from higher functional beneficiaries.
Radiology Technician	Operational Beneficiary	Will provide the application with necessary data for training models and other requirements of HRCT imagery

Quantum Machine Learning Expert	Consultant	Aids and guides to improve performance of the quantum machine learning models and algorithms.
Pulmonary Disease Expert		Provides advice and guides the developers into the domain to improve the system accuracy and performance.
Dev-ops Engineer		Advice & aids web-app accessibility.
QA Engineer		Advice on quality aspect of the application.
Developer	Hybrid Service Provider	Develops and maintains the system.
Hospital	Product Champion	Reduced opportunity cost for prognosis prediction manually. May attract potential competitive customers due to accurate automated predictions.
Competitor	Negative Stakeholder	May observe and attempt to produce similar systems to manipulate and turn over existing and potential purchasers and product champions.
Hacker	Hostile Agent & Negative Stakeholder	May observe and attempt to breach system security and be hazardous to the system and data integrity of the application laws and conduct.
Health Ministry	Regulator	Will be politically influencing the system to comply all laws and control any loopholes.

4.4 Selection of Requirement Elicitation Methodologies

Gathering requirements from related and significant stakeholders have been using a variety of requirement elicitation tools in this project. This section discusses the tools utilized and the reasoning for usage. Questions and other material used in the preparation of the tools have been attached to the appendix section D.

Table 4.2: Requirement Elicitation Apparatus & Reasoning

SH Role	Target SH	Gathering Apparatus
Negative Stakeholder	Competitors	***Apparatus 1:*** **Literature Review** The author has reviewed and critically evaluated existing applications with capabilities to successfully preform prognosis predictions of pulmonary fibrosis and other similar domains (pulmonary and HRCT based) grounded on the last five years to identify a potential gap in the field. Therefore, using the features studied from the existing system, the author can gather key requirements for the system.
Functional Beneficiary	Pulmonologists, Radiologists & Medical Students	***Apparatus 2:*** **Structured Interviews** As means of collecting requirements from the users of the system, pulmonologists, radiologists, and medical students, structured interviews were performed. Due to the lack of availability of many medical professionals, interviews were deemed the most effective path to collect requirements, additionally giving to its ability to gather qualitative feedback than through text feedback.
Regulator & Consultant	QML Experts, Pulmonology Experts	***Apparatus 3:*** **Structured Interviews** As means of collecting expert-opinion of the system and its models, interviews were conducted with QML and Pulmonology experts. Due to the restricted amount of availability of medical experts and the quantum computing domain being new, conducting interviews were deemed efficient.
Developer	Developer	***Apparatus 4:*** **Experimenting/ Prototyping** Prototyping allows the developer the ability to attempt different methodologies progressively and repetitively for

		the implementation of the system, giving to the agile development life-cycle method the project follow.

4.5 Discussion of Results

The result-data acquired through the requirement elicitation apparatus versed above have been presented and discussed below.

4.5.1 Findings from Literature Review

Table 4.3: Findings from Literature Review

Citation	**Findings**
(Goldin et al., 2019)	Firstly, the authors explore the use of a wrapper method for feature selection over the use of the Filter methods, more popularly used, due to their tendency to misalign the feature selection with the classification steps, resulting in a lesser performance in the classification model. Demonstrates how the use of Evolutionary Computation (EC) to improve the overall performance of wrapper methods. Combines the feature selection prediction algorithms using a wrapper method which combines a quantum particle swarm optimization algorithm to select a few features with random forest for classification.
(Amin et al., 2022)	The authors have identified and utilized a two-stage DP paradigm which includes the use of a CGAN for synthetic data generation, as well as the utilization of both classical and quantum machine learning for the classification stage, using classical convolutional neural networks and quantum Quanvolutional neural networks for the other. The CGAN takes use of three different layers, such as ReLu, for the generative and Leaky Elu for the discriminative networks. For the CNN, ReLu and SoftMax activation techniques are used and based on the selected hyper-parameters such as the batch-size, Epoch, the model is trained. The QNN uses 3 layers including quantum-layers, dense layers of specified activation and drop-out layers. The experiments performed on both

	classical and quantum algorithms, distinguishes the increased performance and accuracy of the CNN and the QNN, and a Confusion Matrix demonstrates 0.96 accuracy in the QML models where the classical algorithm only reached an accuracy of 0.86.
(Sengupta and Srivastava, 2021)	Uses the OpenCV library to de-noise the HRCT imagery as pre-processing techniques. The study also discusses the use of other QML algorithms such as QCNN and hybrid CNN, which utilizes both classical and quantum machine learning to produce a hybrid model in both environments. The authors also use hinge loss and adaptive learning rate optimization (ADAM) over the use of the typical gradient descent-based optimizer giving to its lesser cost computationally and implementation.
(Mandal et al., 2020)	Attempts an ensemble of classical machine learning models as listed under techniques used. Laplace Log technique has been used here for evaluation requirements, which is convenient to measure the accuracy which in return is essential for medical practices to derive a confidence score.
(Wong et al., 2021)	Attempts to utilize GSInquire deep neural networks to validate the performance of the deep neural network using explainability driven performance validation. This ensures transparency in the decision-making process using the clinical imagery data. This also helps identify gaps that exist in classical ML due to the models making "right decisions for wrong reasons". Further, this builds a sense of trust on the system, with the decisions made clearly viewable.

4.5.2 Findings from Functional Beneficiary Interviews

Pulmonologists, radiologists, and medical students were identified as the core functional beneficiary of Fibro-QuanNet as depicted in the onion model above. Structured interviews were carried out with experts in both the pulmonology and radiology domains to collection opinions and establish the requirements of the system. A consultant pulmonologist from the Sri Lankan Council of Pulmonologists (SLCP), two radiologists from a private hospital, a senior house officer from the Karapitiya National Hospital of SL, and four medical students both pre-intern and MD/ MBBS readers were interviewed and produced below is the thematic analysis of the results and opinions gathered during the interviews. The complete set of questions directed and the expected outcomes of is are attached in the appendix section D.

Table 4.4: Thematic analysis of findings from interviews with functional beneficiaries

Code	**Theme**	**Conclusion**
Biomarkers for Prognosis Prediction	Identification and analysis of available and widely used biomarkers for the prediction of the prognosis in PF.	As imagery data inputs from the user plays a key role in Fibro-QuanNet, this question attempts to recognize biomarkers are widely accepted in the industry. Answers received includes HRCT imagery, MRI imagery and X-Ray imagery, of which all medical professionals agreed, HRCT would be take prominence.
Automating prognosis prediction	Identifies the need for an automated system to predict the prognosis of PF.	As this project deals with a very time sensitive lung condition, which, by the most subtle of detail can make a mass difference in the well-being of the patient, and given the errors a manual system tends to pose, all medical professionals agreed to the need and use of an automated system for PP.
Detection of Fibrosis	Identifies the indications and potential symptoms of PF in the lung cavity through imagery.	Training the model to identify the presence of PF is one of the first steps in the process of PP. Thus, information such as honeycombing in the lung cavity, sectional darkening, abnormality in cavities, etc., was collected from pulmonologists of ways to identify and properly the analyze the presence of PF in the lung cavity using HRCT imagery.
Performing prognosis	Identifies the methodologies adhered to when	Fibro-QuanNet attempts at automating the prognosis prediction of PF, where learning and knowing how to effectively prognose PF obviously plays a crucial role. The medical professionals directed the author with resources to refer, as well

	performing prognosis.	as giving a high-level idea and a step-by-step walkthrough of the process.

4.5.3 Findings from Consultant Interviews

Data scientists and quantum computing experts have been identified as the key technology consultants of Fibro-QuanNet system. Structured interviews were carried with classical and quantum machine learning experts in-order to collect their opinions and suggestions regarding the development of the system. A Ph.D. quantum machine learning expert at the Microsoft Quantum Research Center, India; a masters (reading) machine learning expert; and four other classical ML experts have been interviewed to establish a solid foundation to the understanding of the system development and produced below is the thematic analysis of the results and opinions gathered during the interviews. The complete set of questions directed and the expected outcomes of is are attached in the appendix D.

Table 4.5: Thematic analysis of findings from interviews with the Consultants

Code	**Theme**	**Conclusion**
Public Quantum Computers	Identifies publicly available quantum computers usable during the project.	Since this project deals with utilizing quantum computational capabilities, publicly available quantum computers had to be identified and initialized to get maximum computational capabilities for the development process, as well as for risk mitigation. Systems such as IBM Quantum, Microsoft QML Stimulator, Google Quantum, etc. was suggested by the experts.
Quantum Algorithms implemented	Identifies quantum machine learning algorithms the technical experts have implemented before to get pointers	Since this project revolves around being able to utilize quantum computing capabilities to run quantum machine learning algorithms, the technical experts were asked what algorithms they've implemented to identify and get potential pointers for the implementation of Fibro-QuanNet.

Medical pre-processing libraries	Identified potential libraries that can be used for the processing of biomarkers effectively.	The first step of the development process as suggested involves pre-processing the input HRCT imagery. Thus, most experts suggested the use of third-party libraries to ensure valid pre-processing and to effectively manage time within the short time constraints. Libraries such as OpenCV, MONAI, etc. were suggested.
Feature extraction/ selection techniques	Identifies possible feature extraction and feature selection techniques for the implementation of the system.	Experts suggested, to keep the application as efficient as possible and to make the predictions as accurate as possible, feature extraction and selection techniques would be helpful for fine-tuning. Techniques such as sampling, noise reduction, contour sampling were suggested.
Evaluation Techniques	Identifies possible evaluation techniques for evaluating Fibro-QuanNet.	Since evaluation of the model plays a crucial role in the project, possible and the best evaluation techniques were queried from the technical experts for smooth and acceptable evaluation of the models built in the project. Techniques such as AUC, PR curve, error rate mappings, etc., were suggested by the experts.

4.5.4 Findings from Prototyping

Findings through reiterated prototyping, many challenges and problems were faced, listed under criterion below.

Table 4.6: Criterion of the findings from prototyping

Criteria	**Conclusion**
Quantum model development	During the repetitive building of the quantum models involved in Fibro-QuanNet, problems such as the allocation of qubits and dataset

	integrations were faced. This had to be addressed by using multiple quantum computers simultaneously with intervals of utilizations.
Quantum computational resource allocation	One of the key issues faced during the prototyping was the availability of quantum resources for the building and run-time of the models. Quantum resources were re-located amid or after idle time of the resources. To address this issue, again, multiple quantum computers were made available during the development process.
Dataset availability	Finding a dataset for this project was a challenging process due to the information the system requires. Even though HRCT images could be collected and analyzed, patient metadata such as the age, genetic history, etc. which are also key factors deciding the prognosis of the patient, had to also be identified and evaluated. To overcome this challenge, the author reached out to other researchers that have worked on classical approaches of the PF prognosis prediction for data collection and try to get datasets they may have created or utilized, where he was able to identify the OSIC pulmonary fibrosis dataset created for a ML competition, which includes all the information required for the project development.

4.6 Summary of Findings

Table 4.7: Summary of Findings

Id	**Finding**	**Literature Re.**	**Interview (FB)**	**Interview (consultant)**	**Prototyping**
1	PF prognosis prediction would benefit and save lives of PF patients who require urgent predictions for treatments.	✓	✓		
2	Prediction systems may reach a physical boundary for further research in classical ML, making way for alternative pathways.			✓	

3	Achieving the quantum advantage would be of benefit for prediction systems of the prognosis for PF by ramping its speed and efficiency.	✓		✓	✓
4	The gap identified will contribute to both the Medical/ health domain and the quantum computing domain alike.	✓	✓	✓	
5	Building and maintaining custom and use-case specific models for Prediction systems are beneficial and preferred over the use of typical libraries	✓		✓	✓
6	Having a method to effectively and swiftly prediction the prognosis of PF will be a benefit to the patients and users.		✓	✓	
7	Using Quantum Neural Networks/ Quanvolutional Neural Networks over other quantum algorithms will be effective and easier to work within the set time frame	✓		✓	✓
8	Using HRCT imagery data as the input data set for PF prognosis would be deemed superior to using other biomarkers such as MRI etc.	✓	✓		
9	The application should be usable and maintainable due to the fast-improving quantum machine learning field and to incorporate its changes	✓		✓	
10	Having an adaptable and variable prediction model would be preferable to allow for future use of the same models and datasets.	✓		✓	✓
11	Opinions of quantum machine learning experts, pulmonologists and radiologists would be a major influence in the application usage.		✓	✓	

12	Having a sufficient set of well-cleaned & pre-processed data would be vital for the performance of the system.	✓		✓	

4.7 Context Diagram

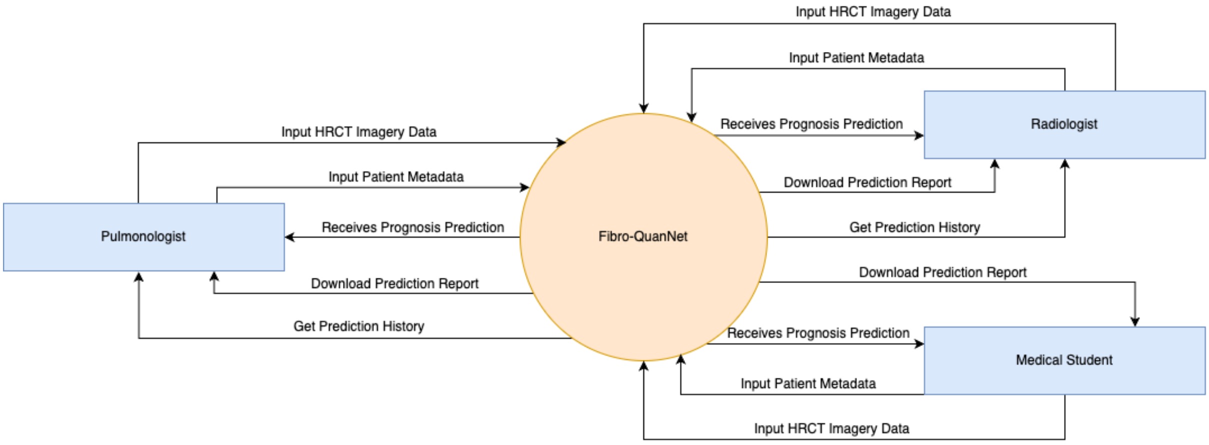


Figure 4.3: Context Diagram

4.8 Use Cases

4.8.1 Use Case Diagram

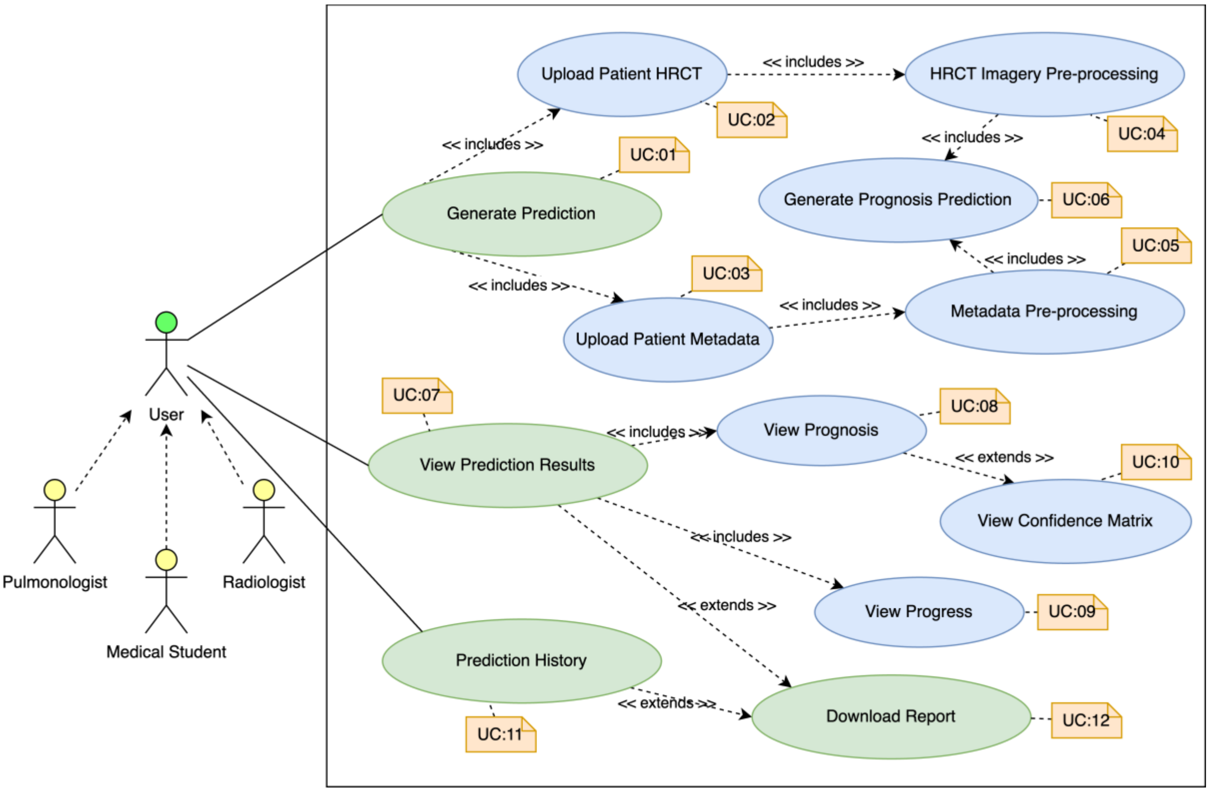


Figure 4.4: Use Case Diagram

4.8.2 Use Case Descriptions

Table 4.8: Use Case Description UC:01

<table>
<tr><td>Use Case</td><td colspan="2">Generate Prediction</td></tr>
<tr><td>Id</td><td colspan="2">UC:01</td></tr>
<tr><td>Description</td><td colspan="2">This functionality allows the use to enter the HRCT imagery and metadata to the system which triggers the pre-processing and PM.</td></tr>
<tr><td>Primary Actor</td><td colspan="2">User</td></tr>
<tr><td>Supporting Actors</td><td colspan="2">None</td></tr>
<tr><td>Stakeholders</td><td colspan="2">Pulmonologist, Radiologist, Medical Student</td></tr>
<tr><td>Pre-condition</td><td colspan="2">The user needs to have uploaded HRCT (UC:01) and the metadata (UC:03) data to the system</td></tr>
<tr><td>Post Condition</td><td colspan="2">Success End: Results produced with the FVC as expected with the confidence matrix.</td></tr>
<tr><td>Trigger</td><td colspan="2">Success end of the UC:05</td></tr>
<tr><td rowspan="2">Main Flow</td><td>Actor</td><td>System</td></tr>
<tr><td>1. User opens the upload wizard
3. 30-100 slides of “. dicom” files selected.
6. Inputs and uploads Metadata</td><td>2. Prompts to upload HRCT
4. dicom type files uploaded
5. Prompts to upload metadata
7. Starts the pre-processing
8. Quantum Module runs</td></tr>
<tr><td>Alternative Flows</td><td colspan="2">-</td></tr>
<tr><td rowspan="2">Exception Flow</td><td>Actor</td><td>System</td></tr>
<tr><td>3a. Select less than 30 dicom images.
3b. Selects over 100 dicom images</td><td>4a./ 5a. uploading failed (network).
7a. models starting failed</td></tr>
</table>

Included Use Cases	UC:02, UC:03, UC:04, UC:05, UC:06

Table 4.9: Use Case Description UC:07

Use Case	View Prediction Results	
Id	UC:07	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

The rest of the use-case descriptions have been moved to the appendix section D.

4.9 Requirements

4.9.1 Functional Requirements

The MoSCoW principle has been followed to regulate the priority levels given to each requirement in par with their significance. Priority keymap has been attached in appendix C.

Table 4.10: Functional Requirements

FR-ID	Requirement	Use Case	Priority Level
FR01	User must be able to upload HRCT imagery to the system	UC:01	M
FR02	The user must be able to enter prevailing FVC to the system	UC:03	M
FR03	The system must be able to pre-process the HRCT	UC:02	M
FR04	The system should choose and borrow most efficient quantum-processing time from the fleet of available quantum machines	UC:05	S
FR05	The system must be able to detect and predict the prognosis of PF in the input HRCT	UC:05	M
FR06	The system should output the confidence matrix of the prognosis prediction	UC:09	S
FR07	The user should be able to view past prognosis predictions	UC:10	S
FR08	The user should be able to download prognosis prediction report	UC:11	C
FR09	The system should have a user-friendly graphical user interface	UC:06	S
FR10	The system should be able to view the current progression of PF	UC:08	C
FR11	The system could produce, and print reporting based on the prognosis prediction	UC:11	C
FR12	The user should be able to elect to contribute HRCT data	UC:07	C
FR13	The system will not show the reasoning behind the prognosis prediction	NA	W

FR14	The system will not take in any other biomarkers as input	NA	W

4.9.2 Non-Functional Requirements

Table 4.11: Non-Functional Requirements

NFR-ID	Requirement	Description	Priority
NFR01	Quality of Output	The accuracy of the prognosis predictions must be of highest quality possible due to its key-role in patient care. The system must be able to produce a quality attribute along with the PP describing the accuracy and quality of the prediction.	Crucial
NFR02	Performance	The system must be able to produce prediction in a much faster and with higher precision using and achieving the quantum advantage for PF PP.	Crucial
NFR03	Usability	Since the system caters for pulmonologists and radiologists who can be of all ages and abilities, the system must incorporate an easy-to-use user interface, which will abide by all design guidelines for accessibility.	Crucial
NFR04	Scalability	Since QC is still at its early ages, and PF still being researched on heavily for treatments, much more functionalities can be added to the system in the future.	Desirable

4.10 Chapter Summary

Through this chapter, the author has brought forth illustrations and descriptions of all the stakeholders that may interact directly, will be benefited, or will take per-role in any external viewpoints using a rich-picture diagram and the stakeholder onion model. Additionally, the author also discussed the tools used the in gathering the requirements of the system through various stakeholders identified and described. Lastly, the systems use cases, functional and non-functional requirements were also per tailed in detail.

CHAPTER 5: SOCIAL, LEGAL, ETHICAL & PROFESSIONAL ISSUES

5.1 Chapter Overview

This chapter reviews the BSC code of conduct, and discusses how the relevant SLEP issues were identified, and plans were made for the mitigation of those issues.

5.2 SLEP Issues & Mitigations

5.2.1 Social Issues

- Verbal consent was appropriated from all interviewees before interviews to record the proceedings of the interviews and to disclose names & styles in the documentation.
- All interviews sent out as questionnaires clearly noted that data collected will be treated confidentially and findings were brought to the thesis only under a thematic analysis.

5.2.2 Legal Issues

- All data inputs in the application developed will collect nor store no unique identification of individuals nor collect any personal data other than what's required for functioning.
- All programming languages, frameworks, tools, and datasets utilized were under valid open-source licensing and had followed adequate data collection privacy (OSIC).
- All source code produced in this project will be licensed under GPL3 license.

5.2.3 Ethical Issues

- All information in this thesis were obtained from trusted sources along with citations and referencing, and holds integrity with no fabrications, falsification, or plagiarism.

5.2.4 Professional Issues

- The research's outcomes and limitations are exact and are documented with no alteration.
- All software and tools utilized were used under valid open-source or student licenses only.
- Best coding practices and research standards were followed throughout the project.
- Prototype was developed in password-protected devices with the latest security patches.

5.3 Chapter Summary

This chapter identified potential SLEP of the BSC code of conduct & how they were mitigated.

CHAPTER 6: DESIGN

6.1 Chapter Overview

This chapter aims to discuss the design decisions taken to decide on the architecture for Fibro-QuanNet based on the requirements gathered through discussed in chapter 4. These justifications include the selection of the design paradigm, high-level architecture diagram, data flow diagrams, system process activity diagrams, and low-fidelity user interface designs.

6.2 Design Goals

The design quality attributes have been attached in the appendix section E.

6.3 System Architecture Design

The system follows the tiered architecture approach, with three layers, the presentation, logic, and data tiers giving to its ability to establish a logical separation within components.

6.3.1 Architecture Diagram

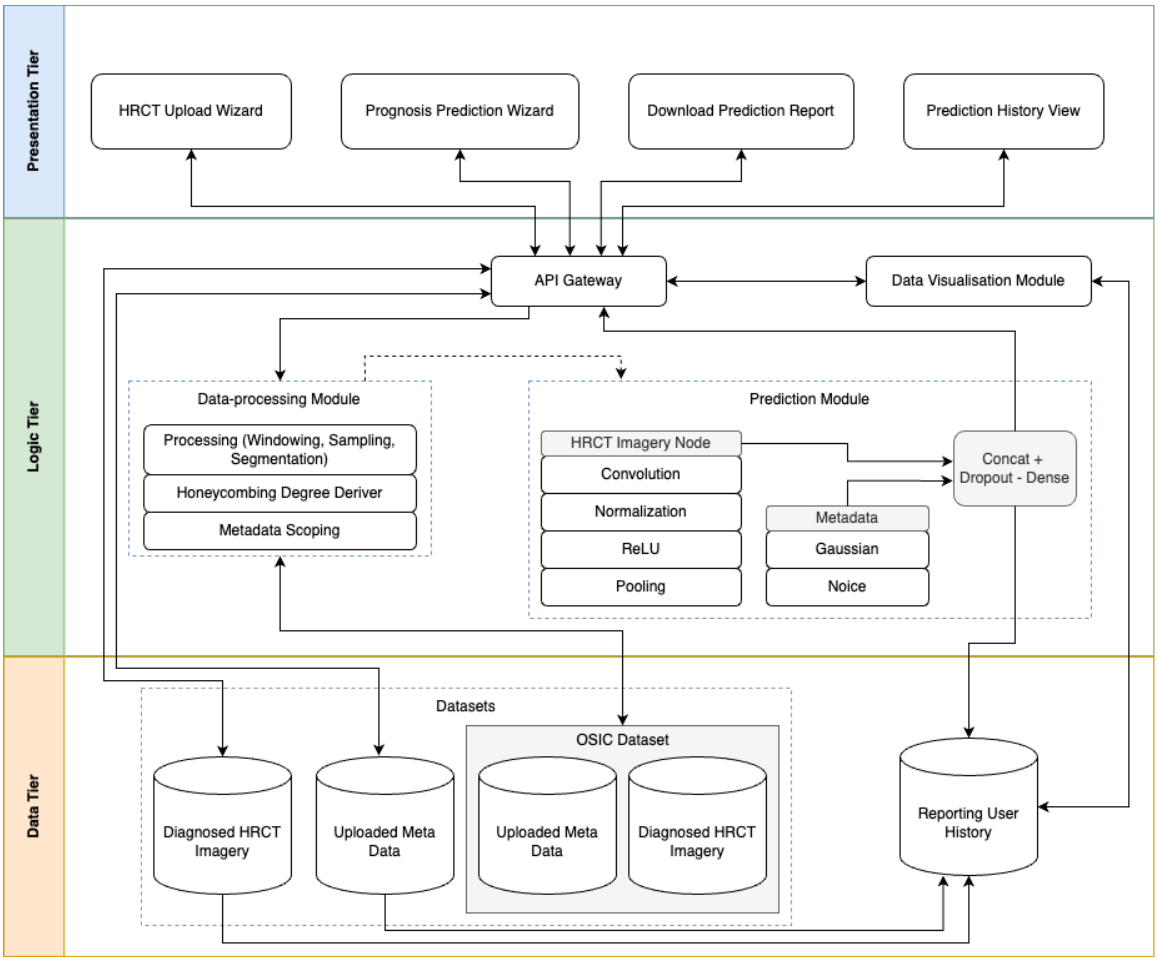


Figure 6.1: High-Level Tiered Architecture Design

6.3.2 Layers of the Architecture

As illustrated above in figure 6.1, Fibro-QuanNet follows a three-tier architecture which consists of the presentation layer, the business logic layer, and the data layer. The three layers play a vital role in keeping the architecture modular, and allows further plugins of microservices to the application, facilitating the design quality attributes.

6.3.2.1 Presentation Tier

- *HRCT Upload Wizard* – This UI allows the user to upload the essential input data such as the HRCT inputs and the metadata of the patient such as their age, smoking status, etc.
- *Prognosis Prediction Wizard* – This UI wizard allows the user to view the prognosis prediction output after the uploading data functionality is complete. This will give the user the exact prediction for a set period in FVC, along with a confidence score of the prediction performed.
- *Download Prediction Report* – This UI allows the user to download/ print the predicted prognosis as a medical report, where all the necessary attributes will be included to a reporting structure.
- *Prediction History View* – This UI allows the user to view the past prognosis predictions he/ she may have done using the system. The application will allow a user sign-in functionality, where they will be able to track all the cases, they've worked on using Fibro-QuanNet.

6.3.2.2 Logic Tier (Business Logic)

- *Data Pre-processing Module* – The pre-processing model required to pre-process the user inputs before sending them into the prediction module. The preprocessing model is broken down into three main steps:
    - *Processing (Windowing, Sampling, Segmentation)* – This segment is responsible for the initial processing of the data by performing actions such as windowing, sampling and finally segmenting the HRCT sections.
    - *Honeycombing Degree Deriver* – The processed data is scanned for attributes of honeycombing, which is a key indicator for pulmonary fibrosis.
    - *Metadata Scoping* – Input metadata is taken into the scope here and added to prepared for the input to the prediction model.

- *Prediction Module* – The core model of the system which is has been specifically trained to predicting the prognosis of pulmonary fibrosis. The core model also includes several nodes and a concatenation node.
    - *HRCT Imagery Node* – The HRCT node is the node which deals with the pre-processed imagery data. This node involved four other sub-processes:
        - Convolution Filtering
        - Batch Normalization
        - ReLu (rectified linear unit) Activation Function
        - 2D Average Pooling

      Each of these layers are repeatedly performed until the desired size is achieved.
    - *Metadata Node* – The metadata node includes of a Gaussian Noise normalization model which aims to mimic the randomization of the processes in nature.
    - *Concatenation and Drop-out* – Finally, both these branches are concatenated, followed by the final dense and drop-out layers.
- *Data Visualization Module* – This module is responsible for creating visual components of the application such as the report generating and prediction result visualization.
- *API Gateway* – The API gateway acts like the router which exposes the endpoints to the public. API calls will be routed towards each controller to perform the necessary modeling.

6.3.2.3 Data Tier

- *Datasets Database* – This database is responsible of storing and serving the datasets used for the system. The initial dataset of the system as well as HRCT imagery that the user had opted to save in the systems will also be saved in these NoSQL databases. The datasets database can be broken down into two further parts.
    - *OSIC Dataset* – This database holds the initial database used for the training of the system.
    - *HRCT & Metadata Input Database* – This database keeps track of all the HRCT and metadata that had been uploaded to the system (given the user had opted for the contributing to the system development)
- *User History* – This database will store the information and the reports generated by users of the system that has opted for keeping their history tracked on the application.

6.4 System Design

6.4.1 Choice of Design Paradigm

After thorough consideration of the requirements for Fibro-QuanNet, the author had seen **Structured System Analysis and Design Method (SSADM)** design paradigm as the preferred design paradigm that aligns with all the requirements of this system over the use of its rival, Object Oriented Analysis and Design (OOAD). Several key decisions that lead to this choice are explored and justified below.

- The research project aims to develop a machine learning based solution which does not require development concerned to object orientation.
- Due to the tight time constraints of the research, it was deemed convenient to develop the application without complex objectification such as object orientation.
- Most quantum programming languages are not object-oriented languages.
- It is also proven that following SSADM assists in future additions to the system.

6.4.2 Design Diagrams

6.4.2.1 Level One Data-Flow Diagram

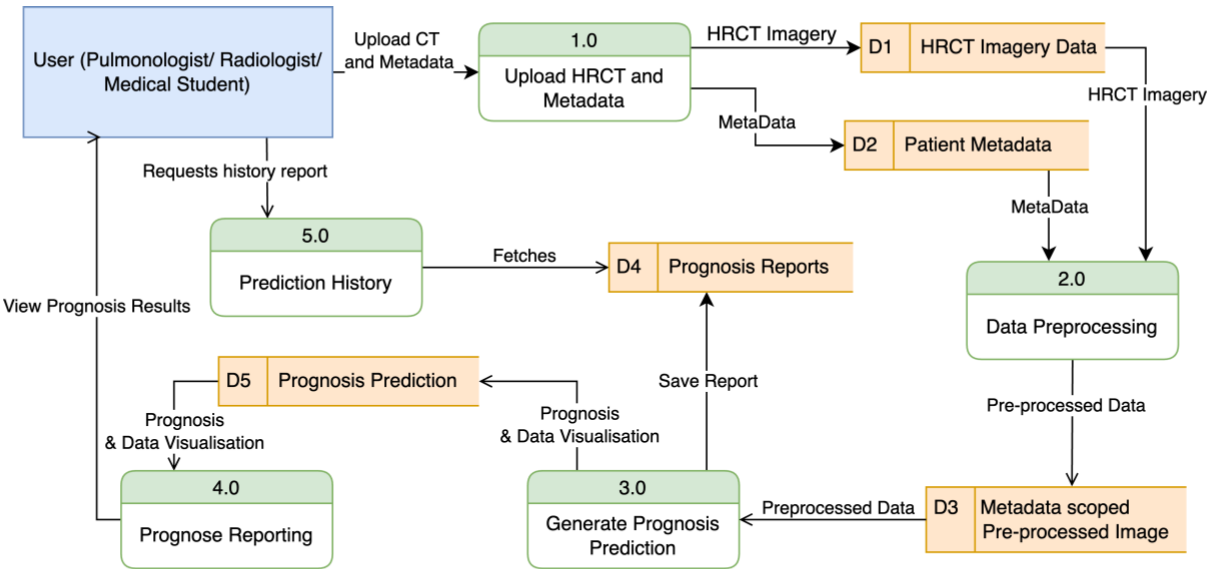


Figure 6.2: Level One Data-Flow Diagram

The Level 1 data flow diagram, presented as Figure 6.2, illustrates the sequential flow of each action segment that is processed at the system's server end. The diagram is composed of several components that work together to meet the functional requirements of the system. Additionally, the diagram portrays the essential outputs that are to be provided to the user after certain processes have been completed.

6.4.2.2 Level Two Data-Flow Diagrams

Data Preprocessing Level Two Data-Flow Diagram

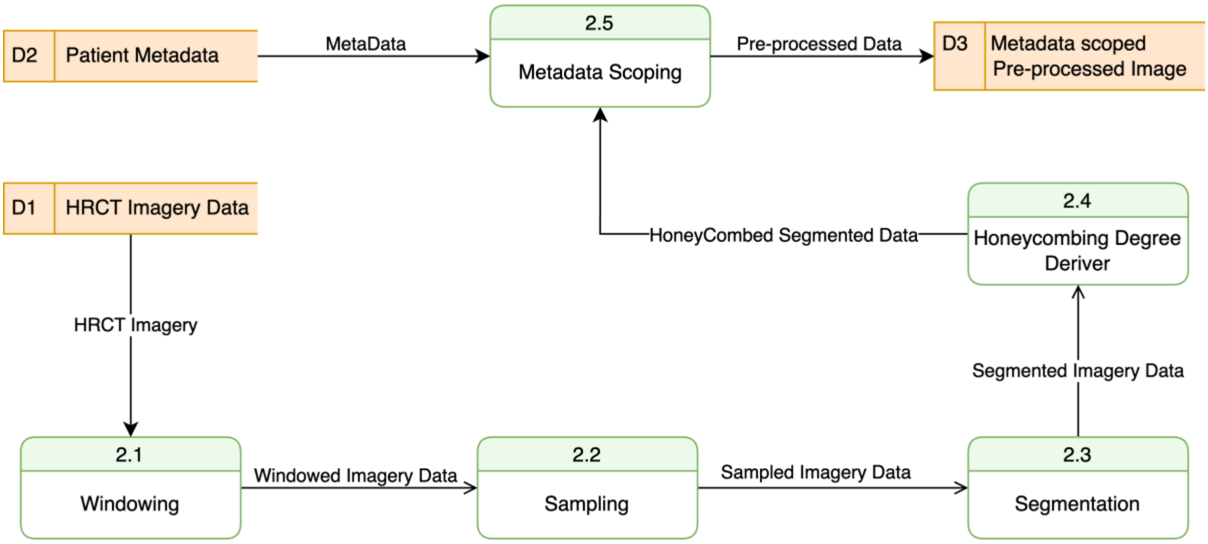


Figure 6.3: Data Preprocessing Level Two Data-Flow Diagram

Generate Prognosis Prediction Level Two Data-Flow Diagram

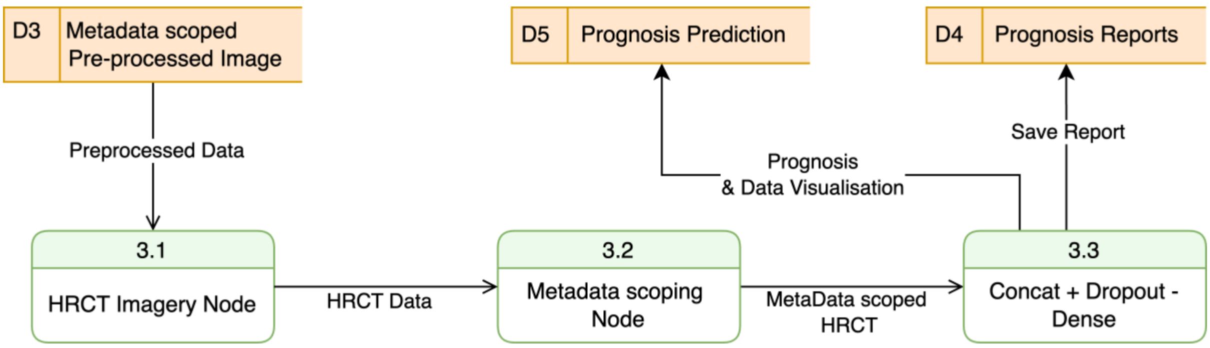


Figure 6.4: Generate Prognosis Prediction Level Two Data-Flow Diagram

6.4.3 System Process Flow Diagram

This diagram is a depiction of the entire system process flow of Fibro-QuanNet with all the system components and validations where necessary.

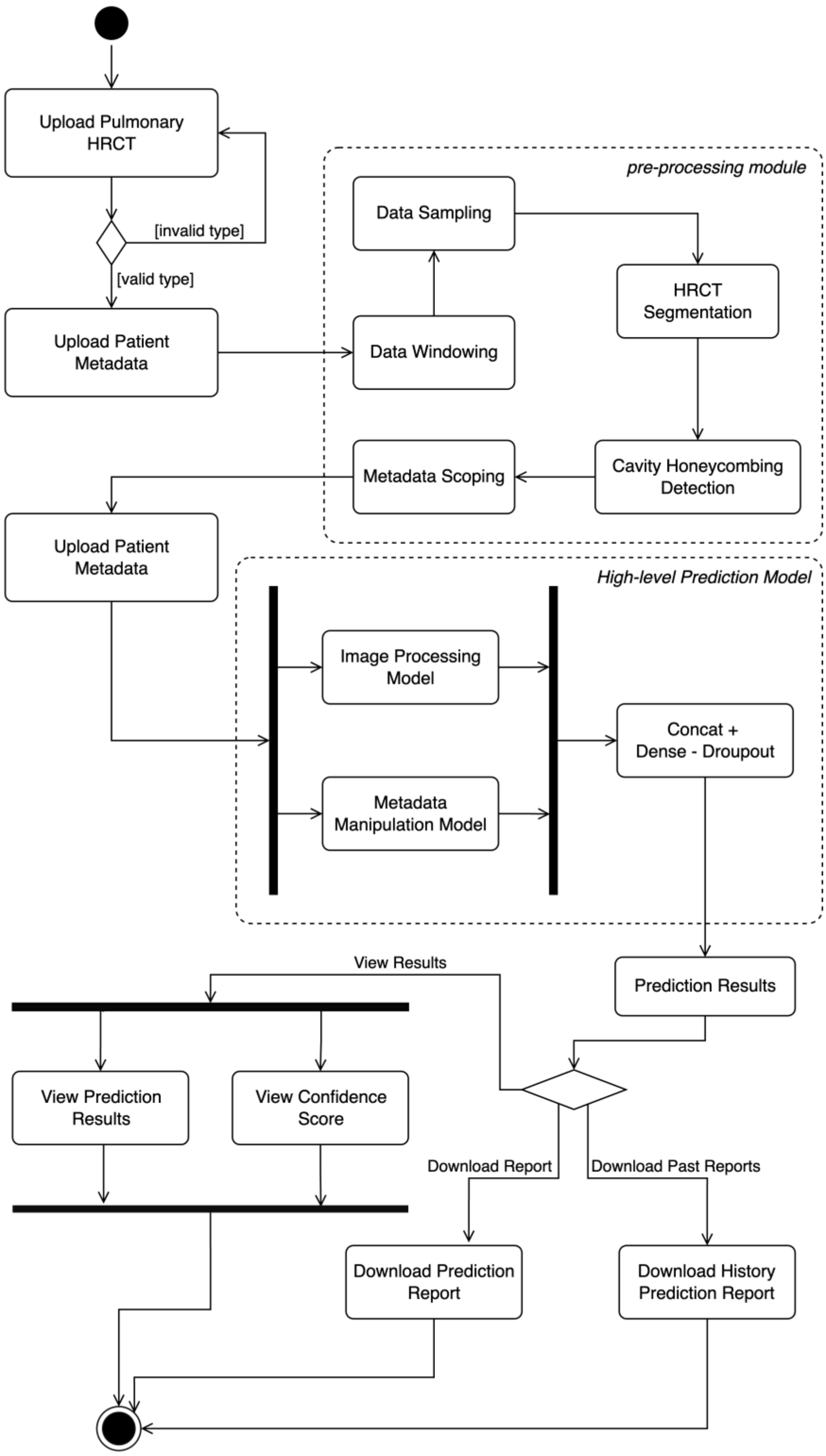
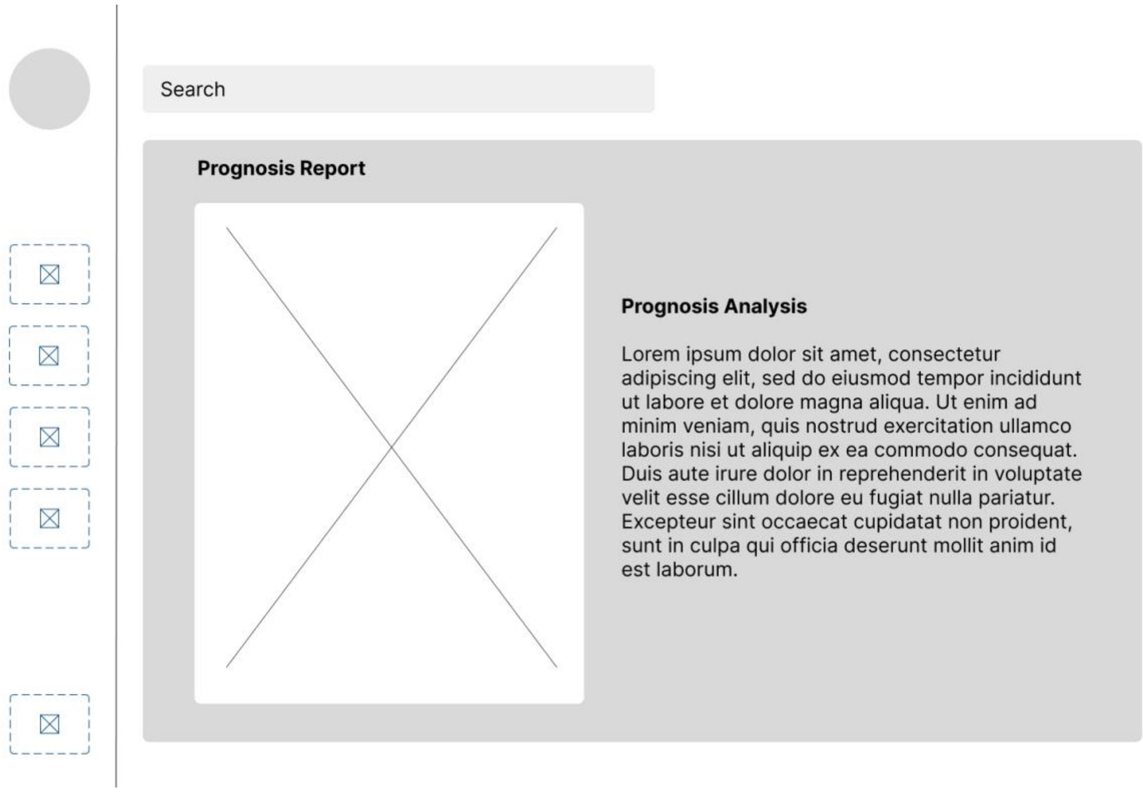


Figure 6.5: System Process Flow Diagram

6.4.4 Low-Fidelity User Interface Design

As a representation of how the interface will look like for a regular user interacting with the system, a low fidelity prototype is presented below. The complete collection of wireframes can be found in the appendix section D.



6.5 Chapter Summary

In this chapter, the design decisions for the system were derived based on the requirements collected during the requirement gathering and elicitation stage. To represent the system, a 3-Tier architecture was chosen, and each of its components was extensively examined. The design paradigm used was explained and supported by reasons, and diagrams were created to illustrate the system flow. The user interface component was then designed according to the stated design goals. Finally, a user experience process was presented to represent the entire system.

CHAPTER 7: IMPLEMENTATION

7.1 Chapter Overview

This chapter provides and overview of the initial implementation of the core of Fibro-QuanNet. Decisions made regarding the preferred dataset used, programming language, libraries and other tools used have be justified here.

7.2 Technology Selection

7.2.1 Technology Stack

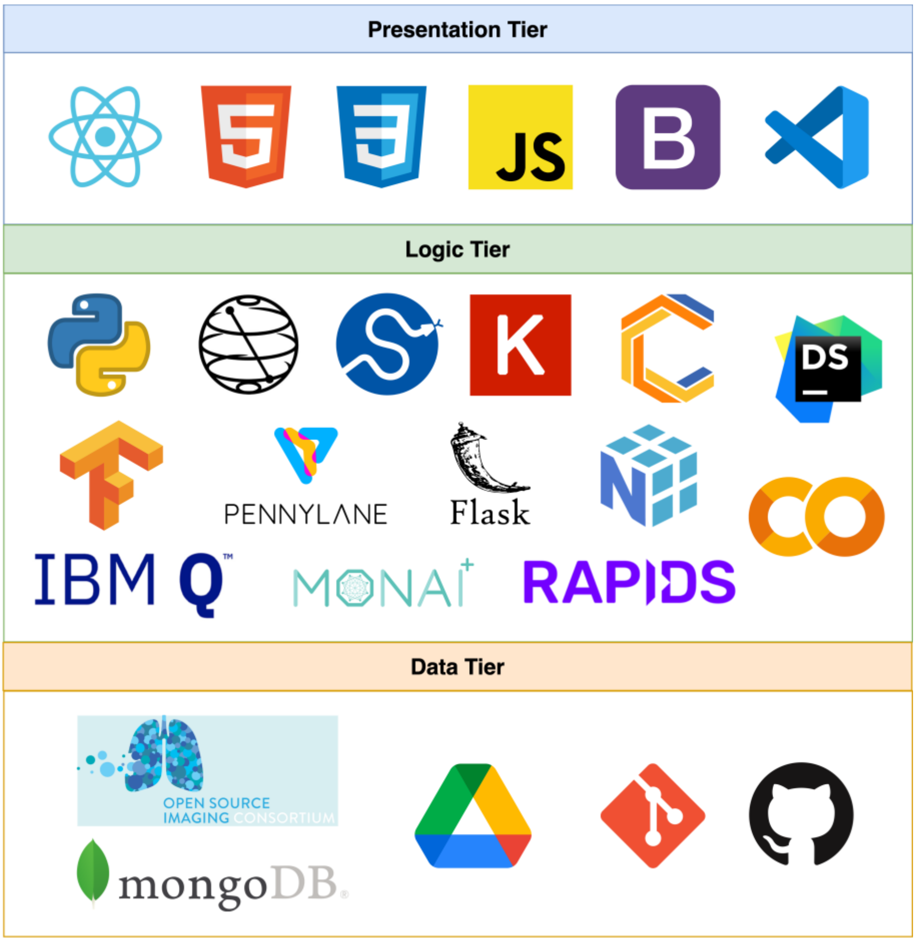


Figure 7.1: Technology Stack

7.2.2 Dataset

The core of Fibro-QuanNet is a data science project, thus requiring a dataset with pulmonary fibrosis prognosis predictions, to train and test the models. Additionally, since this project focuses on medical data, the data must be coming from a valid data point for accuracy. The data required would be:

- Pulmonary cavity HRCT imagery data
- Patient Metadata – Smoking status, age, FVC, etc.

The data requirement was fulfilled using a dataset from the **open-source imaging consortium** which is has made available a pulmonology prognosis prediction dataset which includes data such as the HRCT imagery on baseline, and the patient meta data including the smoking status, age, FVC and lung deterioration percentage.

7.2.3 Programming Languages

Since Fibro-QuanNet is a machine learning application as discussed above, **Python** has been selected as the preferred programming language for the development of the models. Python is a general-purpose programming language and can be used for a wide range of tasks beyond machine learning, including web development, scientific computing, and automation.

It is a popular choice for machine learning projects giving to its ease of use, rich libraries such as NumPy, Pandas, Matplotlib, and TensorFlow which make it easier to preprocess and analyze data, build models, and visualize results; and frameworks, versatility, and support for parallel and distributed computing make it a popular choice for machine learning projects.

7.2.4 Libraries

Table 7.1: Libraries and Justifications

Library	**Justification**
PennyLane	PennyLane is a robust library capable of which consists of the capacity to cater a wide range of quantum computation engines. PennyLane allows the development and training of quantum machine learning models, and then easily run them on different quantum platforms, and compare results across different quantum systems.
Cirq	Cirq is also a robust QML library which on the other hand allows the use and practicing of the use of inbuilt quantum machines for creating and running quantum models in googles hardware itself. One of the key reasons to use Cirq is its ease of use, which allows you to define quantum circuits and algorithms using a simple and intuitive syntax.

Qiskit	Qiskit provides a range of tools and libraries for quantum computing tasks, including support for quantum algorithms, quantum circuit design, and quantum machine learning.
MONAI	MONAI (Medical Open Network for AI) is a deep learning framework for medical imaging and related tasks to develop, train, and evaluate deep learning models for a range of tasks, such as image segmentation, classification, and generation. MONAI provides a comprehensive and user-friendly framework for developing and training deep learning models for medical imaging tasks.
RAPIDS	RAPIDS Acceleration of Diagnostics Deployment is a software library that provides GPU acceleration for data science tasks. It is designed to provide high-performance computing capabilities for data processing and analysis, allowing data scientists to work more efficiently and quickly. By leveraging the power of GPUs, RAPID can help to significantly speed up data processing and analysis, allowing data scientists to focus on creating meaningful insights from their data.

7.2.5 Development Frameworks

Giving to the application being a python-based application, the author has chosen a python-based development framework to implement the connection to its backend applications. Thus, **Flask** has been selected as the preferred development framework giving to its lightweight and ease in the deployment of the application. Flask also supports the use of different libraries and applications used in developing the system, thus making it ideal for the use of Fibro-QuanNet.

7.2.6 Integrated Development Environments

Table 7.2: Integrated Development Environments Selected

IDE	**Justification**
DataSpell	DataSpell provides a collaborative workspace and user-friendly interface, making it easy for teams to work together on projects and share insights.

	This can help facilitate better communication and collaboration within your team, leading to more effective and efficient development processes.
Google Collaboratory	Collaboratory provides a valuable platform for people who want to run data science and machine learning workflows without having to worry about managing their own computing resources.
VSCode	VSCode provides a clean and easy-to-use interface that makes it easy for programmers to navigate their code and find the tools needed.
Jupyter	Jupyter is an easy-to-use application which provides the entire use of a notebook locally and eases the use of quantum libraries.

7.2.7 Summary of Technology Selection

Table 7.3: Summary of Technology Selection

Component	**Technologies Selected**
Programming Languages	Python, JavaScript
Development Framework	Flask
Quantum Computing Library	PennyLane with IBM Quantum
UI Framework	React Bootstrap
Libraries	Cirq, Qiskit, MONAI, RAPIDS, NumPy, PyTorch
IDE – Research	Google Collaboratory, IBM Quantum Labs
IDE – Product	VSCode, DataSpell
Version Controlling	GitHub, Git

7.3 Implementation of Core Functionalities

7.3.1.1 Quantum Gates Declaration

The use of quantum gates is a crucial factor when constructing the quantum circuit used further into the application. The quantum gates produce the core simulations of the quantum phenomena, allowing the achievement of the quantum advantage.

```python
# Single qubit Hadamard Gates layer
def hadamard_layer(nqubits):
    for idx in range(nqubits):
        qml.Hadamard(wires=idx)

# Rotates the parameterized qubits against the y-axis
def rotation_y_layer(w):
    for idx, element in enumerate(w):
        qml.RY(element, wires=idx)

def entangle_layer(nqubits):
    for i in range(0, nqubits - 1, 2):
        qml.CNOT(wires=[i, i + 1])
    for i in range(1, nqubits - 1, 2):
        qml.CNOT(wires=[i, i + 1])
```

Figure 7.2: Quantum Gates Code Snippet

7.3.2 Quantum-Variational Circuit

The quantum variational circuit is responsible of connecting the quantum gates shown above, allowing the sequential flow of parameters mimicking quantum simulations. Variational circuits are used here due the small dataset size and prominent accuracy shown against other circuits available, as discussed in the literature review over in chapter 2.

```python
@qml.qnode(dev, interface="torch")
def variational_circuit(q_input_features, q_weights_flat):
    # Reshape weights
    q_weights = q_weights_flat.reshape(q_depth, n_qubits)

    # Start from state |+> , unbiased w.r.t. |0> and |1>
    hadamard_layer(n_qubits)

    # Embed features in the quantum node
    rotation_y_layer(q_input_features)

    # Sequence of trainable variational layers
    for k in range(q_depth):
        entangle_layer(n_qubits)
        rotation_y_layer(q_weights[k])

    # Expectation values in the Z basis
    exp_vals = [qml.expval(qml.PauliZ(position)) for position in range(n_qubits)]
    return tuple(exp_vals)
```

Figure 7.3: Quantum Variational Circuit Code Snippet

7.3.3 Quantum-Classical Quantile Regression Hybrid Model

Quantile regression has been chosen over the use of neural networks due to several factors discussed over in the literature review, including the small data size the author had to work with. This, instead of being a limiting factor, was actually a turn to the better with the use of quantum variational circuits, which demonstrated a lot more potential, given a bigger dataset and more qubits.

```python
class RegressionModel(nn.Module):
    def __init__(self, in_tabular_features=9, in_ctscan_features=76800, out_quantiles=3):
        super(RegressionModel, self).__init__()
        self.in_ctscan_features = in_ctscan_features

        self.fc1 = nn.Linear(in_tabular_features, 512)
        self.fc2 = nn.Linear(in_ctscan_features, 512)
        self.fc3 = nn.Linear(1024, 512)
        self.fc4 = nn.Linear(512, out_quantiles)

    def forward(self, x1, x2):
        # Now the quant model has 2 inputs: x1 (the tabular features)
        # and x2 (the pre-computed latent features)
        x1 = F.relu(self.fc1(x1))

        # obtain the input features for the quantum circuit
        # by reducing the feature dimension from 512 to 4
        pre_out = self.pre_net(x2)
        q_in = torch.tanh(pre_out) * np.pi / 2.0

        # Flattens the latent features and concatenate with tabular features
        x2 = x2.view(-1, self.in_ctscan_features)
        x2 = F.relu(self.fc2(x2))
        x = torch.cat([x1, x2], dim=1)

        # Apply the quantum circuit to each element of the batch and append to q_out
        q_out = torch.Tensor(0, n_qubits)
        q_out = q_out.to(device)
        for elem in q_in:
            q_out_elem = variational_circuit(elem, self.q_params).float().unsqueeze(0)
            q_out = torch.cat((q_out, q_out_elem))

        x = self.post_net(q_out)
        x = F.relu(self.fc3(x))
        x = self.fc4(x)
        return x
```

Figure 7.4: Quantile Regression Model Code Snippet

More of the implementation code snippets can be viewed over in the appendix section F.

7.4 User Interface

Due to space constrains, the user interface of the implementation has been moved over to the appendix section F.

7.5 Chapter Summary

The implementation chapter outlined the tools and technologies that are used by Fibro-QuanNet along with its libraries and frameworks utilized with reasoning for each of these selections. Further, a step-by-step implementation of the codebase has also been attached with its explanations to match each code snippet. The use of both quantum and classical implementations have been notified and clearly defined where necessary and required.

CHAPTER 8: TESTING

8.1 Chapter Overview

This chapter aims to discuss the testing done towards ensuring the project application functions and produces valid and accurate responses as expected. The expectations of the testing along with the criteria adhered to have been clearly defined in this chapter. Different testing done towards making sure the quantum machine learning model functions as expected have also been demonstrated, along with the functional & non-functional requirements testing. Finally, the limitations identified in the testing process have been acknowledged.

8.2 Objectives and Goals of Testing

The core objective of the testing phase is to make sure that the system can perform as expected in the requirements gathered and produces a valid response to a given input. The following expectations were set down as requirements:

- The machine learning models perform as expected which has been tested to produce the most accurate prognosis.
- The product satisfies the functional requirements from chapter 4 and meets all the “Must have” and “Should have” functionality according to the MoSCoW technique.
- The product satisfies the non-functional requirements from chapter 4.
- The product must be thoroughly tested to ensure a minimum of bugs and improvements made where possible along with bug-fixes.
- The product must be capable of handling any unforeseen, un-expected run-time errors.

8.3 Testing Criteria

In order to demonstrate clear waypoints of between the conceptuality of the system and the implementation of the system, the following methodologies of testing were chosen:

- Functional Quality – Determines the quality and the achievement of the functional and non-functional requirements, and how well the requirements have been met.
- Structural Quality – Ensures the code meets structural code quality, security and the how the non-functional requirements have been met without harming functionality.

8.4 Model Testing

8.4.1 Training and Validation Loss Curve

The training and validation loss curve is a helpful model testing statistic because it shows how effectively the model learns from data during training and how well it generalizes to new data during validation. The training loss gauges how well the model matches the training data after the model has been fitted to it during training. The training loss goes down as the model gets better at fitting the training set of data. The model may, however, overfit the training data, in which case it memorizes the training data as opposed to discovering the underlying patterns. A decrease in training loss with an increase in validation loss serve indicators of this.

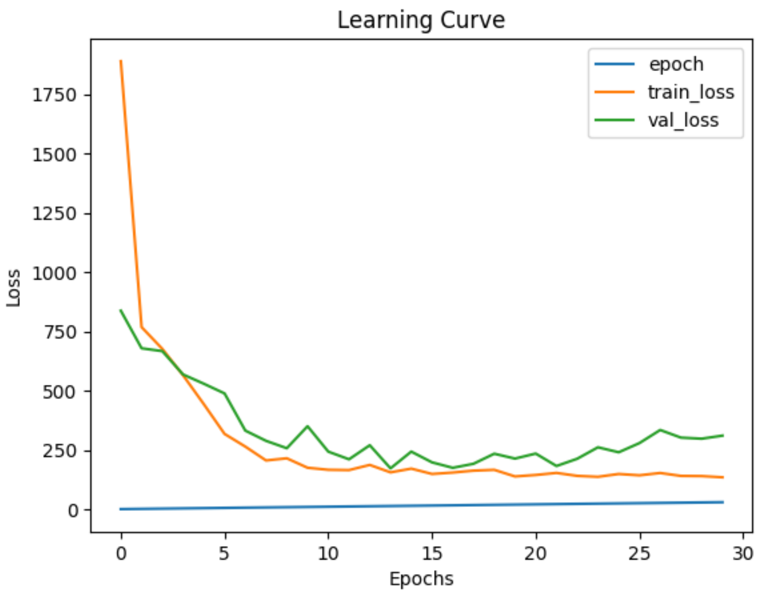


Figure 8.1: Training & Validation Loss Curve

8.4.2 Root Mean Squared Error

The RMSE calculates the average deviation in response variable units between expected and actual values. The square root of the average of the squared discrepancies between the anticipated and actual values is used to compute RMSE. The lower the RMSE, the better the prediction ability of the model. RMSE is a suitable model test for regression models since it is simple to grasp. Because the matric is given in the same units as the response variable, it gives an easy-to-understand measure of how accurate the model's predictions are. This is especially critical when explaining the model's conclusions to stakeholders who may not be familiar with statistics or data science. Another reason RMSE is a valuable statistic is that it accounts for all faults in the model's predictions. They reflect both the bias and the variability of the errors, unlike some other metrics, like as R-squared, which simply assess the fraction of the variation in the response variable that is explained by the model. As a result, it provides a more comprehensive statistic for assessing model performance.

```
mse = sklearn.metrics.mean_squared_error(actual, predicted)

rmse = math.sqrt(mse)

print("Mean Squared Error", rmse)
```

```
Mean Squared Error 179.3242
```

Figure 8.2: Root Mean Squared Error Snippet

As depicted above, the RMSE graph for Fibro-QuanNet, with over 700 predictions, shows a promising value of just 179.32, which is a very promising value for a regression model.

8.4.3 Mean Absolute Error

The mean absolute error (MAE) is a popular statistic for assessing the effectiveness of regression models. It calculates the average difference in response variable units between expected and actual values. The MAE assesses the average size of the mistakes and is more resilient to outliers than the root mean squared error (RMSE), which squares the errors and gives greater weight to huge errors. It is an excellent model test for regression models because of its resilience to outliers. Outliers are exceptional values in data that differ significantly from the other values. They can have a huge influence on the RMSE since it squares the errors, resulting in a large penalty for outliers. The MAE, on the other hand, does not square the errors, making it less susceptible to outliers. This indicates that if your data contains outliers, the MAE may be a better choice.

```
# for loop for iteration
for i in range(n):
    summation += abs(actual[i] - predicted[i])

error = summation/n

print("Mean absolute error : " + str(error))
```

```
Mean absolute error : 212.329
```

Figure 8.3: Mean Absolute Error Snippet

As demonstrated above, the MAE shows a performance of 212.33 (rounded), which is a satisfactory performance for a prediction model.

8.4.4 Laplace Log-Likelihood Graph

The Laplace log-likelihood is a useful metric for assessing the performance of statistical models, especially those that include regularization techniques such as L1 regularization (also known as the Lasso). Regularization approaches such as L1 regularization promote sparsity in the model coefficients, which means that many of them are exactly zero. The Laplace log-likelihood can be used to assess the model's sparsity by punishing non-zero coefficients more severely than zero coefficients. This is significant because sparsity can lead to better interpretable and understandable models. By comparing the log-likelihood values of different models, the Laplace log-likelihood may be utilized for model selection. Models with higher log-likelihood values are more likely to fit the data, whereas models with lower log-likelihood values are less likely to fit. It is feasible to choose the best model for a given problem by comparing the log-likelihood values of many models.

```
Epoch 21 train      : 100%|██████████| 38/38 [00:32<00:00,  1.18it/s, loss=153.6, metric=-6.5520]
Epoch 21 val        : 100%|██████████| 10/10 [00:04<00:00,  2.03it/s, loss=524.7, metric=-8.4796]
Epoch 22 train      : 100%|██████████| 38/38 [01:04<00:00,  1.70s/it, loss=141.0, metric=-6.4523]
Epoch 22 val        : 100%|██████████| 10/10 [00:05<00:00,  1.85it/s, loss=512.3, metric=-8.1952]
Epoch 23 train      : 100%|██████████| 38/38 [00:31<00:00,  1.20it/s, loss=137.1, metric=-6.4219]
Epoch 23 val        : 100%|██████████| 10/10 [00:03<00:00,  3.32it/s, loss=518.6, metric=-8.1900]
Epoch 24 train      : 100%|██████████| 38/38 [00:31<00:00,  1.22it/s, loss=149.0, metric=-6.5137]
Epoch 24 val        : 100%|██████████| 10/10 [00:02<00:00,  3.50it/s, loss=518.6, metric=-7.7485]
Epoch 25 train      : 100%|██████████| 38/38 [00:30<00:00,  1.23it/s, loss=143.5, metric=-6.4738]
```

Figure 8.4: Laplace Log-Likelihood Scores

8.4.5 Quantum Performance Benchmarking

Quantum benchmarking is an essential method for measuring the performance of a quantum computer, including hybrid quantum circuits. It entails comparing the quantum circuit's output to a known or predicted output in order to measure how well the quantum circuit is operating. Dealing with mistakes is one of the most difficult aspects of constructing a quantum computer. Quantum systems are extremely susceptible to noise and other types of interference, which can cause calculation mistakes. Quantum benchmarking allows you to quantify the mistakes in a quantum circuit, which is important for understanding the system's constraints and devising techniques for error correction and mitigation. Another important aim in developing a quantum computer is to optimize its performance. Quantum benchmarking may be used to detect underperforming circuit elements and design methods for enhancing performance. For example, if a specific gate or qubit routinely produces mistakes, it may be able to enhance its performance by adjusting the gate's settings or recalibrating the qubit.

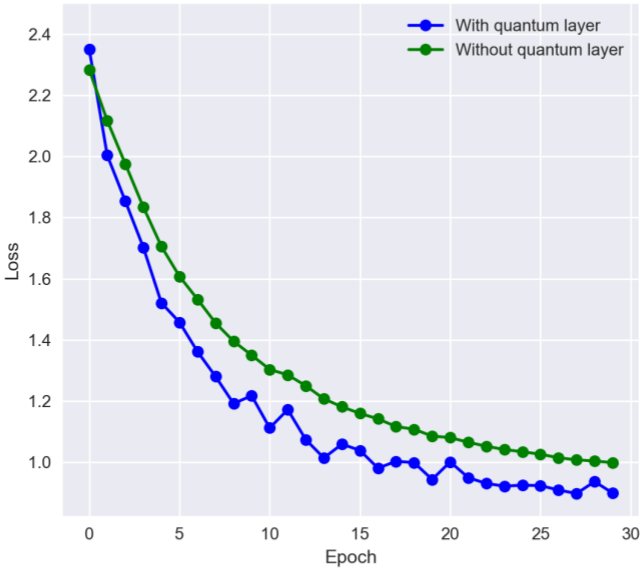


Figure 8.5: Quantum Benchmarking- Training and Validation Graph

As shown in the figure above, quantile regression has performed much better when working with the quantum layer, rather than without the quantum layer. However, the results above have been taken through direct training through IBM Quantum Labs, thus reducing its reproducibility and the quantum optimization being higher than actual performance of simulations demonstrated in figure 8.1. Nevertheless, the model shows significant improvement when trained on quantum processors and parameters as demonstrated.

8.5 Benchmarking

Table 8.1: Benchmarking results on OSIC dataset

Research	**Dataset**	**Regression Mode**	**LLL**	**RMSE**
Proposed Fibro-QuanNet	OSIC	Hybrid Quantile Regression	**-7.128**	212.33
(Nazi et al., 2021)	OSIC	ResNet, EfficentNet-b1, EfficentNet-b2, EfficentNet-b3,	-6.68	**201.5**
(Mandal et al., 2020)	OSIC	Quantile Regression, Ridge, ElasticNet	-6.92	-

(Wong et al., 2021)	OSIC	ElasticNet	-6.28	-

Due to the limited usage of usage of the dataset of published papers, the author decided to resort to benchmarking to different competition applications that has taken place under the image consortium dataset.

Table 8.2: Competition Leaderboard Benchmarking

Competition Ranking	**Models Used**	**LLL**	**RMSE**
1st Place	Quantile Regression, EfficientNet b5	-6.14	-
4th Place	Linear Decay, Quantile Regression	-5.38	-
6th Place	EfficentNet-b0, Ridge, ElasticNet	-6.53815	-
9th Place	Concatenate Tile Pooling	-5.98	-

8.6 Functional Testing

Through the functional testing, each functional requirement identified in the requirement elicitation stage, and brough into scope will be evaluated.

Table 8.3: Functional Testing

Test Case	**FR-ID**	**User Action**	**Expected Result**	**Actual Result**	**Status**
1	FR01	User must be able to upload HRCT imagery to the system	HRCT imagery is accepted to the system	A zip file input of the HRCT DICOM images is accepted into the system	Pass
2	FR02	The user must be able to enter prevailing FVC to the system	FVC is accepted to the system	The FVC entered to the text field is accepted	Pass

3	FR03	The system must be able to pre-process the HRCT	Pre-processing of the system must happen after uploading.	Pre-processing takes place after uploading HRCT images	Pass
4	FR04	The system should choose and borrow most efficient quantum-processing time from the fleet of available quantum machines	IBM Quantum and Google Cirq is repeatedly followed up on to produce outputs	IBM quantum is the only freely available functionality. Thus, physical boundaries to reach	Fail
5	FR05	The system must be able to detect and predict the prognosis of PF in the input HRCT	Detect and predict the prognosis of PF	Successfully detects and predict the prognosis of PF	Pass
6	FR06	The system should output the confidence matrix of the prognosis prediction	Outputs confidence matric	Confidence matrix is displayed in the outputs table	Pass
7	FR08	The user should be able to download prognosis prediction report	Should be able to download prediction	Downloadable CSV with 133 weeks of prediction	Pass
8	FR09	The system should have a user-friendly graphical user interface	User-friendly GUI with documentation	Easy-to-use UI with a documentation has been implemented	Pass

9	FR10	The system should be able to view the current progression of PF	The current progression must be shown	The current level of fibrosis can be viewed at week 0	Pass
10	FR11	The system could produce, and print reporting based on the prognosis prediction	Downloadable report	Downloadable CSV available	Pass

8.7 Module and Integration Testing

Table 8.4: Module and Integration Testing

Module	**Input**	**Expected Outcome**	**Actual Outcome**	**Status**
Upload HRCT	DICOM Zip File	System accepts the zip file upload, and upload into server	System accepts the zip file upload, and upload into server	Pass
HRCT Pre-processing	HRCT Images, Metadata	Returns latency features of CTs	Returns latency features of CTs	Pass
Prognosis Prediction	Latency Features	Returns a CSV with predictions and confidence	Returns a CSV with predictions and confidence	Pass

8.8 Non-Functional Testing

The system's accuracy and performance would be the most important non-functional needs. To evaluate these system sectors, multiple testing criteria and metrics were used, and the results were examined to find any system performance bottlenecks. All the testing of the non-functional requirements has been evidently proven to a completion percentage of 100% available in the appendix section G.

8.9 Limitations of the Testing Process

One of the main issues the author faced during the testing phase was the loss of actual data of predictions in the dataset. The only available information in the dataset were for specific weeks of prognosis, which the author manipulated to check for on a cellular level. This made the evaluation of the system much trickier, and as a result, was restricted to a set amount of evaluation matrices at hand. This issue has also been encountered by other authors, discussed in the benchmarking section above, where they have resorted to the evaluation techniques the author too had utilized for the testing phase. Have the actual values be given, the testing phase could have been increased drastically to in-corporate more matrices.

Another major limitation to the testing of the system was the limited availability of quantum space and allocation, where the author did not resort to quantum simulators such as PennyLane or StrawberryFields, which would have drastically hindered the performance of the system. Thus, as with time quantum computers increases their capabilities, the models could be more enhanced and trained with more qubits, which now is scaled to just 4.

8.10 Chapter Summary

This chapter discussed the testing of the system developed through the research project. Initially, the model testing has been covered, alongside the different evaluation techniques used for evaluating the model’s performance. The author has then covered these matrices being benchmarked against other research that has been conducted with the same datasets. Finally, the application has been tested against all the functional and non-functional requirements.

CHAPTER 9: EVALUATION

9.1 Chapter Overview

Following the successful implementation of the minimum viable product for Fibro-QuanNet, and thorough testing to ensure accurate and efficient predictions discussed in the above chapter, the product was evaluated in perspective of the requirements identified in chapter 4. This chapter includes self-evaluation performed by the author as well as expert opinion in both domain and technical, on the outcomes of the project.

9.2 Evaluation Methodology and Approach

The research project includes both qualitative and quantitative aspects as crucial needs considering the weightage a prognosis prediction holds. Through chapter 8, the author has identified and rectified any potential bugs in the system including verifying the accuracy of the system, however, expert feedback on the qualitative aspect of the system will be discussed in detail over this chapter, as well as self-evaluating the system co-responding to the functional and non-functional requirements identified in chapter 4 through requirement elicitation. A thematic analysis has been performed to present the criticism received from evaluators.

The document shared with the evaluators for explanation of the research project and its outcomes have been attached in the appendix section.

9.3 Evaluation Criteria

In order to perform a focused thematic analysis of the evaluation performed on the research project with conspicuous attention to the achievement of the requirements, the following criteria has been identified and followed through.

Table 9.1: Evaluation Criteria

Criteria	**Purpose of Evaluation**
Choice of Research Concept	To evaluate the significance of this research gap, background, and the complexity undertaken through the project.
Research contributions	To evaluate the significance of the contributions made to the quantum machine learning technical field, as well the pulmonary fibrosis problem domain.

Development approach	To validate if approach adhered to when developing the system and its functionality follows appropriate algorithmic and theoretical concepts.
Usability and UX	To evaluate the usability and to verify that the MVP is user friendly and convenient for use.
Documentation quality and literature	To verify the entire research process has been documented to a substantial quality along with a decent review of the literature.
Qualitative Outcomes	To validate and verify that appropriate testing has been performed with significant performance in the models developed.
Improvements and Limitations	To identify any improvements that can be made to the system or be openings into future work.

9.4 Self-Evaluation

A self-evaluation was performed by the author based on the criteria identified above considering the entire research journey.

Table 9.2: Self-Evaluation as per the Evaluation Criteria

Criteria	**Self-Evaluation**
Choice of Research Concept	The problem domain chosen for the research project is a pressing matter in the health domain, with very little facts known about the disease and treatments, making prognosing the only viable option. The technical domain chosen for the project is a very novel technology, still in its infancy, with a great potential to be the next immense advancement in computing technology.
Research contributions	The core research contributions made primarily lie within the technical domain of the research. A hybrid model for the prediction of prognosis in the med-tech industry has never been attempted using quantum-convolutional neural networks, neither has the tabular information such as the age, gender, etc. be considered before for the prognosis prediction of IPF even in the field of classical prognosis prediction models built for IPF.

Development approach	A tremendous amount of work has gone into the development of the pre-processing of the system to in-corporate the tabular information available through the dataset, as well as to incorporate a hybrid-quantum machine learning layer into the py-torch models which are classically made to run on classical machine learning layers, and further to incorporate the necessary tensors to process the information to predict prognosis for IPF.
Usability and UX	The UI/UX developed for the MVP is quite straight-forward and includes a documentation too to guide the user.
Documentation quality and literature	The highest quality standards have been adhered to when documenting the research project, using the latest versions of Zotero citations and MS word. All the documents referred to were cited appropriately and sourced from DOI where possible.
Quantitative Outcomes	Although difficult to perform measure a prognosis prediction using quantitative values, measures such as lap-lace loglikelihood graphs has been taken to produce quantitative values for the evaluation of the developed MVP.
Improvements and Limitations	After getting feedback from experts regarding improvements that can be made to the system, some were addressed where time-permits such as the lap-lace loglikelihood graph.

9.5 Selection of the Evaluators

A few remarkable individuals were chosen and reached-out for in order to perform expert evaluation on the research developments, the MVP, and the research documentation. The individuals were chosen through the following break-down of stakeholders.

Table 9.3: Evaluators Selection Criteria

EC ID	**Category**
EC1	Research domain experts including quantum machine learning experts, machine learning experts, software technical leads, technology officers, etc.

EC2	Problem domain experts including pulmonologists, radiologists, general physicians, general surgeons, medical officers.
EC3	Potential users of the system, including pulmonologists, radiologists, pre-intern doctors, medical-students, medico-legal officers

9.6 Evaluation Results

Table 9.4: Thematic Analysis of Feedback from Evaluators

Criteria	EC ID	Theme	Feedback
Choice of Research Concept	EC1	Quantum Prognosis Prediction Gap	The study of quantum computing along with machine learning is still an emerging technology and would be a research area with very much potential and impact.
	EC2	PF problem domain	PF is one of the most lethal lung diseases that even state-of-art medicine is defenseless, and accurate prognosis prediction is crucial for treatments.
	EC3	Domain applicability	No research similar to this has happened in Sri Lanka and has not been applied in any hospital systems to speedup nor measure accuracy of PP. Locally sourced products can be considered into being used at hospitals too.
Research contributions	EC1	Technical Contribution	Quantum hybrid machine learning models are very new and have shown promising results so far, utilizing this in a prognosis prediction system has not been seen before, and is a great contribution.
	EC2	Domain Contribution	Not many people have attempted to take essential medical information of the patient

			like their age, gender into consideration when predicting, making this a solid addressing to the gap in prognosis prediction of IPF.
	EC3	Domain Contribution	Accurate predictions with contingencies towards medical information is a great move and not many applications exist solving this.
Development approach	EC1	Use of DICOM and metadata	The use of both DICOM and external meta-data that are deciding factors is fabulous.
		Use of hybrid ML models	The use of a hybrid model over a full-quantum ML model is a smart move here, giving to the state of quantum computers yet.
	EC2	Use of quantum computing,	The use of quantum computing brings an edge into the accuracy of predictions.
	EC3	Use of quantum computing,	Prognosis predictions is such a crucial factor in IPF and is great to have solutions with such high accuracy and processing.
Usability and UX	EC1, EC3	Presents prognosis predictions	The UI is smooth and is easy to use along with the documentation, especially the reports.
Documentation quality and literature	EC1	Quality and practicality of LR and content	The documentation of the research journey and the literature and its critical review was commended for its accuracy and deep review into technologies and existing work.
	EC2, EC3	Literature of the domain and approach taken to the solution	The deep review into the problem domain and the thorough discussion into PF was commended and pointed out as accurate. The solution approach was praised as the steps identified are familiar to logical decisions made during a physical prognosis prediction.

Quantitative Outcomes	EC1	Analysis of the prognosis prediction system	The current testing pathways taken are satisfactory, and the use of non-value-oriented testing matrices such as lap-lace loglikelihood as an added advantage giving the nature of the results produced, which is not necessarily positive or negative values.
	EC1, EC2, EC3	Analysis of Fibro-QuanNet application	The application works without errors in a smooth run and produces the results within a minute, which is efficient for such a process that would usually take weeks to predict. Error handling has also been added.
Improvements and Limitations	EC1	Usage of hybrid-quantum models	The author has considered using entirely quantum neural networks as well, which shows the amount of work put into the research, however, future researchers could utilize full quantum algorithms.
	EC2, EC3	Analysis of Fibro-QuanNet application	Could further utilize more patient data such as genetic factors which can be hard to process, however would open a great potential gap.

9.7 Limitations of Evaluation

Quantum computing is a fairly new subject of research, which leaves only a few researchers and experts in the field to contact and work with, and almost no experts with the knowledge of quantum computing in Sri Lanka, where the research is based in. Thus, it was very difficult to find and contact experts in the field and get their expert feedback and opinions. Therefore, the author mostly has done evaluation with classical machine learning experts, who do not possess expert knowledge regarding quantum machine learning, which can be identified as a hindering in the evaluations, and the only contact with actual quantum ML experts were online and in different time-zones which made it further difficult to connect. Further, connecting with medical officers and consultants in the medical field is extremely difficult with their busy time-schedules, results in only calls to gather feedback about the application.

9.8 Evaluation on Functional Requirements

The functional requirements identified as the most crucial requirements (M) has been cut out with the evaluations for each. The full set of FRs can be found in the appendix section E.

Table 9.5: Evaluation of the Must Have Functional Requirements

FR-ID	Use Case	Priority Level	Evaluation
FR01	UC:01	M	Implemented
FR02	UC:03	M	Implemented
FR03	UC:02	M	Implemented
FR04	UC:05	M	Implemented
Function Requirements completion percentage: 10/12 (ignoring W requirements) = 83.3*%			

9.9 Evaluation of Non-Functional Requirements

The full set of NFRs can be found in the appendix section E.

Table 9.6: Evaluation of the Crucial Non-Functional Requirements

NFR-ID	Requirement	Priority	Evaluation
NFR01	Quality of Output	Crucial	Implemented
NFR02	Performance	Crucial	Implemented
NFR03	Usability	Crucial	Implemented
Non-Functional Requirements completion percentage: 4/4 = 100%			

9.10 Chapter Summary

This chapter discussed the self-evaluation and the thematic analysis of the feedback received from selected evaluators, who have also been carefully selected through discussed consideration, and the criterion chosen for the evaluation of the project. The evaluation of the functional and non-functional requirements has also been covered through this chapter. The full set of requirements can be found in the appendix section E.

CHAPTER 10: CONCLUSION

10.1 Chapter Overview

This chapter brings the dissertation of this research to a conclusion and includes the concluding remarks of the project. The contributions made to the research domain has been given in a concise, along with the achievement of the research aims and objectives. The challenges the author face, and use of existing skills as well as the development of new and important skills and been discussed followed by a confab into future work that can be based off this project.

10.2 Achievement of Research Aims and Objectives

10.2.1 Achievement of the Aim

This research aims to design, develop, and evaluate a novel prediction model which is capable of providing accurate and efficient prognosis predictions of pulmonary fibrosis utilizing High-Resolution Computer Tomography data through quantum machine learning.

The aim of the research was achieved successfully by designing, developing, and evaluating a quantum-hybrid Quanvolutional neural network, capable of accurately performing prognosis predictions of pulmonary fibrosis using HRCT imagery input as well as other patient meta-data, while utilizing quantum machine learning techniques.

10.2.2 Achievement of the Objectives

The full list of objectives of the research and their corresponding achievement status can be found in the appendix section.

Table 10.1: Research Objectives Achievements Summary

Objective	**RO ID**	**Learning Outcomes**	**Status**
Problem Identification	RO1	L01	Completed
Literature Survey	RO2, RO3, RO4, RO5, RO6	L01, L04, L05	Completed
Requirement Analysis	RO7, RO8, RO9	L01, L02, L03, L06	Completed

Design	RO10, RO11, RO12	L01	Completed
Development	RO13, RO14, RO15, RO16	L07, L05	Completed
Testing & Evaluation	RO17, RO18, RO19	L04	Completed
Publish Findings	RO20, RO21	L06, L08	Completed

10.3 Utilization of Knowledge from the Course

Table 10.2: Knowledge Utilized from the Course Content

Module(s)	**Knowledge Utilized**
Software Development Group Project	The initial experience of performing a research, documentation, the research process was comprehended during the undertaking of the SDGP module. Thanks to its research experience, performing research project individually was made leisurelier and armed.
Object Oriented Programming, Programming Principle I & II	These modules gave us the understanding needed to develop, document, and test a complete software, while also demonstrating the programming principles and standards followed and are deemed required. The module also pushed us into utilizing different frameworks and databases, which helped a lot during the research.
Algorithms: Theory, Design, and Implementation	The module opened us into the world of logical and algorithmic thinking, which helped understand and solve complex logical equations and algorithms, and helped working with intricate, complex algorithms such as neural networks.
Web Designing and Development	Introduced the aspect and the foundation required for developing web-applications using HTML, CSS, and JavaScript and the necessary UI/UX principles needed to adhere to when developing the UI of the application.

10.4 Use of Existing Skills

- *Full-Stack Development* – While working in his internship during the industrial placement year at iTelaSoft, the author was exposed to full-stack development with different technology stacks. These development skills came in handy when developing the minimum viable product for the project.
- *Machine Learning* – Through different courses and seminars held at university as well as through online-learning platforms made available through the university contributed massively to the development of the machine learning models in the project. Software development group project also enabled to lay the foundation in image processing neural networks, of which the knowledge has been utilized when building and coming up with the models used in this research project.
- *Algorithmic Understanding and Knowledge* – The review and understanding of algorithms really came in convenient when finding, understanding, and utilizing different algorithms needed to achieve the most accuracy and performance in the models and the applications developed through the research project.

10.5 Use and Development of New Skills

- *Quantum Computing* – The author had to review, research, and acquire knowledge about quantum computing in a short-time span, for which he followed two quantum-computing short-courses on LinkedIn-Learning, provided through the university.
- *Quantum Machine Learning* – The author was expected acquire knowledge on utilizing quantum exclusive machine learning algorithms and techniques, which was a challenging task by itself with the limited number of resources available to do so. However, upon thorough investigation and suggestions from industry experts, a course was followed on edX, offered by the University of Toronto, which incredibly helped in developing the machine learning models of the research project.
- *High-Resolution Computed Tomography Application and Reconstruction* – The author had to understand and learn how HRCT imager data works, including how DICOM image formats work and can be reconstructed and worked with to utilize this high-resolution imagery to be pre-processed and employed in the quantum machine learning model, which has not been attempted before. Exploring the literature of HRCT images usage allowed the author to gather the necessary knowledge for this purpose and utilize this knowledge in the research project MVP development.

10.6 Achievement of Learning Outcomes

Table 10.3: Achievement of Learning Outcomes

Learnings	**LOs**
The problem identified was broken down into two main components, the model building, training, testing and application development; as well as the documentation and literature review. Thus, the core gap was attempted in a methodical manner utilizing necessary tools and technologies required. Sub-tasks were assigned timespans and systematized to complete within timescale.	LO1, LO2
The research project uses quantum computing alongside quantum machine learning and image processing techniques such as neural networking, in addition to HRCT image reconstruction and pre-processing. These tasks were all carried out using the appropriate tools and technologies, after thorough review of existing tools.	LO3, LO5
Over 40 papers were read, analyzed, and critiqued in order to review literature.	LO4
The supervisor was regularly met with expected and required deliverables.	LO5
All social, legal, ethical, and professional issues were taken into consideration, and mitigation plans have been laid, discussed in chapter 5.	LO6
The project features the development of the full-stack application with a front-end application built using ReactJS, a back-end application built using Python with a flask middle tier, and the python-Jupyter notebook to train models.	L07
This dissertation documents the entire research process, including critical analysis and reviewing of existing applications, project management and planning methodologies, as well as new skills utilized in the project, to the best standards of academic documentation possible.	LO8

10.7 Problems and Challenges Faced

Table 10.4: Mitigations to Problems and Challenges Faced

Problem/ Challenge	**Mitigation**
The selected domain area consisted of very complex mathematical	Constant communication and help from technical experts at the Microsoft Quantum Research

phenomenon in both computational and mechanical.	Center in India, as well as guidance from the supervisor helped immensely to overcome these challenged, and to look for appropriate resources.
Since quantum computing and quantum machine learning along with pulmonary fibrosis prognosis prediction are both fairly new research topics and complex ones at that, a steep learning curve had to be covered in the very short time span.	Short courses were found on platforms like LinkedIn-learning and edX, which enabled a quick grasp of knowledge and understanding of the technical domain, along with some practicing in the IBM quantum-labs. For problem understanding, the author being a victim of the disease, and information already being known came in very handy and helpful.
Since quantum computers are still in its infancy, getting hold of a physical quantum-computer is practically impossible at date, making it harder to access quantum computational hardware through clouds computing, which required a lot of network-capacity.	The author was lucky to have been working in a time-zone with not much cloud-traffic to the quantum computer, since not much research activity was being channeled towards the quantum-computer from the +0530-time zone. However, the time was restricted when it came to rigorous utilization at times like training of the model, which required constant connection and communication to the quantum hardware.
Lack of quantum computing and quantum machine learning experts, due to the research domain being fairly new.	The author reached out to quantum computing experts in well-known organizations like the Microsoft Research center and the Strawberry-Fields quantum research center during the initial proof-of-concept stage, which helped to always have contact to quantum computing experts for any doubts or issues during development.
Extremely busy time-schedules of medical professionals and consultants for the problem domain, which entirely depends on the medical field.	The author was able to contact two consultant pulmonologists, who he refers to personally, as well as medical officers and medical students he personally has contact with, and gather necessary clarifications, requirements, and evaluations.

10.8 Deviations

The following changes had to be made to the scope of the project as constrained by time and performance of the models:

- The initial plan of the project was to develop an entirely quantum-machine learning centric quantum neural network. However, after development and testing of different algorithms applicable to quantum-neural networks, it was deemed, according to the state-of-art quantum computational power available free (7 qubits per session on IBM Quantum), the hybrid quantum-classical neural network utilized in the implementation shows better accuracy and performance speed than a purely quantum neural network.
- The initial project plan included "could have" features to develop a sign-in user portal, where a user's prediction histories could be accessed and contribute data for the development of the models. However, this poses a legal and ethical issue where it is deemed illegal in some countries and states to store medical information with user-identification enabled in any form, especially when contributing the data.

10.9 Limitations of the Research

- As mentioned in the deviations, a purely quantum neural network was attempted during the research, however scrapped due to non-performance. This is a physical limitation due to the limited availability of qubits in the freely available quantum hardware access.
- This research focuses on using the pinnacle of prognosis prediction imagery, HRCT - DICOM images, while not considering the possibility of predictions using typical CT imagery or x-ray imagery, which are both, though not popular, viable options to use for prognosis predictions.
- The system only takes meta-data input of a patients age, gender, and smoking status, limited to which are made available by the dataset used during the training. However, genetical information, which are also a key dependency has not been accounted for anywhere in the meta-data inputs.
- The dataset considered for the training and the testing of the models were really small for its class, which is a common problem in the medical research domain, resulting of its time-consuming data labeling and reconstructions, thus making deep-learning approaches very difficult to approach and achieve the quantum advantage.

10.10 Future Enhancements

Based on the limitations identified of the system above as well as input from the evaluators during the evaluation of the project, different suggestions and inputs have been recognized as work that may be able to further improve the application of PF prognosis predication.

- As discussed in the limitations above, pure quantum neural networks could be a great avenue of research in the near future. As clearly stated, with an higher amount of qubits made available, it would be possible to utilize these qubits and build and train a potential quantum neural networking model. With IBMs goal towards reaching 300 qubits, the capacity available for free tier plans could be made higher, which may allow and support the use and development of pure quantum models.
- Recognizing patterns in which genetical information can be stored and processed to a tensor to add more biases to the system could also potentially be a future work.
- Introducing explainable AI into the domain of prognosis prediction systems, would also be a great piece of enhancement to the domain. Viewing the chronological steps taken by a machine learning model in the health industry would bring more trustworthiness into the health care-AI industry and MedTech.
- Collecting data of existing pulmonary fibrosis diagnosed lung imagery, and putting them together while creating a much larger dataset would also be a added value future work. The limited about of data available makes it harder to train deep learning models and other neural networks.

10.11 Achievement of the Contribution to the Body of Knowledge

With the conclusion of the project, the author has been able to address the following contributions to the two domains of PF prognosis prediction using quantum ML.

10.11.1 Research Domain

- The research contributes a novel hybrid quantum-classical neural networking model, which is able to successfully predict the prognosis of PF considering all the possible metadata which may deeply impact the performance and accuracy of the system.
- DICOM imagery and tabular data pre-processing model, capable of receiving the DICOM image file and a CSV file with patient metadata, and be able to convert into tensor outputs, capable of feeding into a prognosis prediction model.

According to the latest literature, performed until the very last day of dissertation writing, no quantum-classical hybrid applications, nor any quantum machine learning models exist to perform any prognosis prediction of PF in a faster, accurate, and convenient with outstanding accuracy and lap-lace similarity for any classical ML implementation of a prediction model.

10.11.2 Problem Domain

The recognition and utilization of patient meta-data in the prognosis prediction of PF has never been attempted before and has been successfully achieved through this project. However, there are more potential data that may be involved, but were not made available on the dataset.

10.12 Concluding Remarks

This marks the conclusion of the research project which attempted to address one of the world's most lethal, uncurable pulmonary diseases, where the state-of-art medical practices are still deemed helpless. The hybrid model implemented through this paper will make the prognosis prediction of PF for patients much more accurate and simpler. The quantum-classical hybrid variational circuit will also be utilizable and improvable in the future with the daily advancement in quantum computation.

REFERENCES

Acampora, G. and Schiattarella, R. (2021). Deep neural networks for quantum circuit mapping. *Neural Computing and Applications*, 33 (20), 13723–13743. Available from https://doi.org/10.1007/s00521-021-06009-3.

Alamri, A.K. et al. (2023). Pulmonary Hypertension Association's 2022 International Conference Scientific Sessions Overview. *Pulmonary Circulation*, 13 (1), e12182. Available from https://doi.org/10.1002/pul2.12182.

Al-jumaili, S. et al. (2022). Recent Advances on Convolutional Architectures in Medical Applications: Classical or Quantum? *2022 International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT)*. 20 October 2022. Ankara, Turkey: IEEE, 800–805. Available from https://doi.org/10.1109/ISMSIT56059.2022.9932857 [Accessed 20 December 2022].

American Lung Association, N.H.C. (2022). Each Breath: Facts about Pulmonary Fibrosis. *7 Things Everyone Should Know about Pulmonary Fibrosis*. Available from https://www.lung.org/blog/7-things-know-pulmonary-fibrosis [Accessed 6 September 2022].

Amin, J. et al. (2022). Quantum Machine Learning Architecture for COVID-19 Classification Based on Synthetic Data Generation Using Conditional Adversarial Neural Network. *Cognitive Computation*, 14 (5), 1677–1688. Available from https://doi.org/10.1007/s12559-021-09926-6.

Andersson-Sjöland, A. et al. (2008). Fibrocytes are a potential source of lung fibroblasts in idiopathic pulmonary fibrosis. *The International Journal of Biochemistry & Cell Biology*, 40 (10), 2129–2140. Available from https://doi.org/10.1016/j.biocel.2008.02.012.

Antoniu, S.A. (2008). Targeting the endothelin pathway in the idiopathic pulmonary fibrosis: the role of bosentan. *Expert Opinion on Therapeutic Targets*, 12 (9), 1077–1084. Available from https://doi.org/10.1517/14728222.12.9.1077.

Astor, T.L. et al. (2023). Anti-fibrotic therapy and lung transplant outcomes in patients with idiopathic pulmonary fibrosis. *Therapeutic Advances in Respiratory Disease*, 17, 175346662311659. Available from https://doi.org/10.1177/17534666231165912.

Azevedo, V., Silva, C. and Dutra, I. (2022). Quantum transfer learning for breast cancer detection. *Quantum Machine Intelligence*, 4 (1), 5. Available from https://doi.org/10.1007/s42484-022-00062-4.

Beer, K. et al. (2020a). Training deep quantum neural networks. *Nature Communications*, 11 (1), 808. Available from https://doi.org/10.1038/s41467-020-14454-2.

Beer, K. et al. (2020b). Training deep quantum neural networks. *Nature Communications*, 11 (1), 808. Available from https://doi.org/10.1038/s41467-020-14454-2.

Bernhardt, C. (2019). *Quantum Computing for Everyone*. MIT Press.

Biamonte, J. et al. (2017). Quantum Machine Learning. *Nature*, 549 (7671), 195–202. Available from https://doi.org/10.1038/nature23474.

Cai, Q. et al. (2019). Exploring Object Relation in Mean Teacher for Cross-Domain Detection. Available from https://doi.org/10.48550/ARXIV.1904.11245 [Accessed 10 May 2023].

Cao, Y. et al. (2019). Quantum Chemistry in the Age of Quantum Computing. *Chemical Reviews*, 119 (19), 10856–10915. Available from https://doi.org/10.1021/acs.chemrev.8b00803.

Cerezo, M. et al. (2022). Challenges and opportunities in quantum machine learning. *Nature Computational Science*, 2 (9), 567–576. Available from https://doi.org/10.1038/s43588-022-00311-3.

Choi, W.-I. (2021). Current and future treatment for idiopathic pulmonary fibrosis. *Journal of the Korean Medical Association*, 64, 256–263. Available from https://doi.org/10.5124/jkma.2021.64.4.256.

Chuliá-Peris, L. et al. (2022). Matrix Metalloproteinases and Their Inhibitors in Pulmonary Fibrosis: EMMPRIN/CD147 Comes into Play. *International Journal of Molecular Sciences*, 23 (13), 6894. Available from https://doi.org/10.3390/ijms23136894.

Cong, I., Choi, S. and Lukin, M.D. (2019). Quantum convolutional neural networks. *Nature Physics*, 15 (12), 1273–1278. Available from https://doi.org/10.1038/s41567-019-0648-8.

Das, S. and Chakraborty, K. (2015). Review paper of Idiopathic Pulmonary Fibrosis. *International Journal of Advanced Research*, 3, 1565–1570.

Devaraj, A. (2014). Imaging: How to recognise Idiopathic Pulmonary Fibrosis. *European Respiratory Review*, 23 (132), 215–219. Available from https://doi.org/10.1183/09059180.00001514.

Díaz-Piña, G. et al. (2018). The Role of ADAR1 and ADAR2 in the Regulation of miRNA-21 in Idiopathic Pulmonary Fibrosis. *Lung*, 196 (4), 393–400. Available from https://doi.org/10.1007/s00408-018-0115-9.

Dunjko, V., Taylor, J.M. and Briegel, H.J. (2016). Quantum-Enhanced Machine Learning. *Physical Review Letters*, 117 (13), 130501. Available from https://doi.org/10.1103/PhysRevLett.117.130501.

Epstein Shochet, G. et al. (2021). MiR-608 overexpression in idiopathic pulmonary fibrosis (IPF). *BMC Pulmonary Medicine*, 21 (1), 1. Available from https://doi.org/10.1186/s12890-020-01377-3.

Furukawa, T. et al. (2022). A comprehensible machine learning tool to differentially diagnose idiopathic pulmonary fibrosis from other chronic interstitial lung diseases. *Respirology*, 27 (9), 739–746. Available from https://doi.org/10.1111/resp.14310.

García-Sancho, C. et al. (2011). Familial pulmonary fibrosis is the strongest risk factor for idiopathic pulmonary fibrosis. *Respiratory Medicine*, 105 (12), 1902–1907. Available from https://doi.org/10.1016/j.rmed.2011.08.022.

Gisin, N. et al. (2002). Quantum cryptography. *Reviews of Modern Physics*, 74 (1), 145–195. Available from https://doi.org/10.1103/RevModPhys.74.145.

Goldin, J.G. et al. (2019). Prediction of progression in idiopathic pulmonary fibrosis using CT scans at baseline: A quantum particle swarm optimization - Random forest approach.

Artificial Intelligence in Medicine, 100, 101709. Available from https://doi.org/10.1016/j.artmed.2019.101709.

Gui, X. et al. (2018). Leptin promotes pulmonary fibrosis development by inhibiting autophagy via PI3K/Akt/mTOR pathway. *Biochemical and Biophysical Research Communications*, 498 (3), 660–666. Available from https://doi.org/10.1016/j.bbrc.2018.03.039.

Guo, X., Schwartz, L.H. and Zhao, B. (2019). Automatic liver segmentation by integrating fully convolutional networks into active contour models. *Medical Physics*, 46 (10), 4455–4469. Available from https://doi.org/10.1002/mp.13735.

Houssein, E.H. et al. (2022). Hybrid quantum-classical convolutional neural network model for COVID-19 prediction using chest X-ray images. *Journal of Computational Design and Engineering*, 9 (2), 343–363. Available from https://doi.org/10.1093/jcde/qwac003.

Huang, H.-Y. et al. (2021). Power of data in quantum machine learning. *Nature Communications*, 12 (1), 2631. Available from https://doi.org/10.1038/s41467-021-22539-9.

Huang, Y. et al. (2015). A functional genomic model for predicting prognosis in idiopathic pulmonary fibrosis. *BMC Pulmonary Medicine*, 15 (1), 147. Available from https://doi.org/10.1186/s12890-015-0142-8.

Hunninghake, G.M. (2014). A New Hope for Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine*, 370 (22), 2142–2143. Available from https://doi.org/10.1056/NEJMe1403448.

Inomata, M. et al. (2021). Clinical impact of the radiological indeterminate for usual interstitial pneumonia pattern on the diagnosis of idiopathic pulmonary fibrosis. *Respiratory Investigation*, 59 (1), 81–89. Available from https://doi.org/10.1016/j.resinv.2020.07.001.

Jagpal, S., Pistun, O. and Mikhail, J. (2014). Pirfenidone for Idiopathic Pulmonary Fibrosis, Thrombocytosis in Chronic Obstructive Pulmonary Disease Exacerbations, and a Longitudinal Study on E-Cigarettes. *American Journal of Respiratory and Critical*

Care Medicine, 190 (6), 699–700. Available from https://doi.org/10.1164/rccm.201406-1188RR.

Karthikeyan, A. and Priyakumar, U.D. (2022). Artificial intelligence: machine learning for chemical sciences. *Journal of Chemical Sciences*, 134 (1), 2. Available from https://doi.org/10.1007/s12039-021-01995-2.

Kim, G.H.J. et al. (2021). A study design for statistical learning technique to predict radiological progression with an application of idiopathic pulmonary fibrosis using chest CT images. *Contemporary Clinical Trials*, 104, 106333. Available from https://doi.org/10.1016/j.cct.2021.106333.

King, T.E. et al. (2014). A Phase 3 Trial of Pirfenidone in Patients with Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine*, 370 (22), 2083–2092. Available from https://doi.org/10.1056/NEJMoa1402582.

Kistler, K.D. et al. (2014). Lung transplantation in idiopathic pulmonary fibrosis: a systematic review of the literature. *BMC Pulmonary Medicine*, 14 (1), 139. Available from https://doi.org/10.1186/1471-2466-14-139.

Kulkarni, V., Pawale, S. and Kharat, A. (2022). A Classical-Quantum Convolutional Neural Network for Detecting Pneumonia from Chest Radiographs. Available from http://arxiv.org/abs/2202.10452 [Accessed 20 December 2022].

Kumaran, N.V. and Preethi, D.M.D. (2022). Impact of Chronic Lung Disease Using Deep Learning: A Survey. In: Neuhold, E.J. Fernando, X. Lu, J. et al. (eds.). *Computer, Communication, and Signal Processing*. IFIP Advances in Information and Communication Technology. Cham: Springer International Publishing, 46–59. Available from https://doi.org/10.1007/978-3-031-11633-9_5 [Accessed 30 September 2022].

Lacedonia, D. et al. (2021). Downregulation of exosomal let-7d and miR-16 in idiopathic pulmonary fibrosis. *BMC Pulmonary Medicine*, 21 (1), 188. Available from https://doi.org/10.1186/s12890-021-01550-2.

Lake, F. (2014). Pirfenidone reduced disease progression in idiopathic pulmonary fibrosis. *Annals of Internal Medicine*, 161 (8), JC4. Available from https://doi.org/10.7326/0003-4819-161-8-201410210-02004.

Lawrence, J. and Nho, R. (2018). The Role of the Mammalian Target of Rapamycin (mTOR) in Pulmonary Fibrosis. *International Journal of Molecular Sciences*, 19 (3), 778. Available from https://doi.org/10.3390/ijms19030778.

Lee, J.K. et al. (2023). Clinical impact of weight loss on mortality in patients with idiopathic pulmonary fibrosis: a retrospective cohort study. *Scientific Reports*, 13 (1), 5774. Available from https://doi.org/10.1038/s41598-023-32843-7.

Lee, J.S. et al. (2010). Does Chronic Microaspiration Cause Idiopathic Pulmonary Fibrosis? *The American Journal of Medicine*, 123 (4), 304–311. Available from https://doi.org/10.1016/j.amjmed.2009.07.033.

Li, J.-Y., Zhan, Z.-H. and Zhang, J. (2022). Evolutionary Computation for Expensive Optimization: A Survey. *Machine Intelligence Research*, 19 (1), 3–23. Available from https://doi.org/10.1007/s11633-022-1317-4.

Li, P., Xiao, H. and Li, B. (2016). Quantum representation and watermark strategy for color images based on the controlled rotation of qubits. *Quantum Information Processing*, 15 (11), 4415–4440. Available from https://doi.org/10.1007/s11128-016-1413-x.

Liu, G. et al. (2019a). A Quantum Particle Swarm Optimization Algorithm with Teamwork Evolutionary Strategy. *Mathematical Problems in Engineering*, 2019, e1805198. Available from https://doi.org/10.1155/2019/1805198.

Liu, G. et al. (2019b). A Quantum Particle Swarm Optimization Algorithm with Teamwork Evolutionary Strategy. *Mathematical Problems in Engineering*, 2019, e1805198. Available from https://doi.org/10.1155/2019/1805198.

Liu, M. et al. (2019). Ligustrazin increases lung cell autophagy and ameliorates paraquat-induced pulmonary fibrosis by inhibiting PI3K/Akt/mTOR and hedgehog signalling via increasing miR-193a expression. *BMC Pulmonary Medicine*, 19 (1), 35. Available from https://doi.org/10.1186/s12890-019-0799-5.

Lloyd, S., Mohseni, M. and Rebentrost, P. (2013). Quantum algorithms for supervised and unsupervised machine learning. Available from https://doi.org/10.48550/ARXIV.1307.0411 [Accessed 2 May 2023].

López-Muñiz Ballesteros, B. et al. (2021). Sex Differences in the Incidence and Outcomes of Patients Hospitalized by Idiopathic Pulmonary Fibrosis (IPF) in Spain from 2016 to 2019. *Journal of Clinical Medicine*, 10 (16), 3474. Available from https://doi.org/10.3390/jcm10163474.

Mandal, S. et al. (2020). Prediction Analysis of Idiopathic Pulmonary Fibrosis Progression from OSIC Dataset. *2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON)*. October 2020. 861–865. Available from https://doi.org/10.1109/GUCON48875.2020.9231239.

Matas, J., Galambos, C. and Kittler, J. (2000). Robust Detection of Lines Using the Progressive Probabilistic Hough Transform. *Computer Vision and Image Understanding*, 78 (1), 119–137. Available from https://doi.org/10.1006/cviu.1999.0831.

Mayo Foundation for Medical Education and Research. (2021). Pulmonary fibrosis - Symptoms and causes. *Mayo Clinic*. Available from https://www.mayoclinic.org/diseases-conditions/pulmonary-fibrosis/symptoms-causes/syc-20353690 [Accessed 6 September 2022].

Mittal, K. and Rajam, V.M.A. (2020). Computerized retinal image analysis - a survey. *Multimedia Tools and Applications*, 79 (31–32), 22389–22421. Available from https://doi.org/10.1007/s11042-020-09041-y.

Morena, D. et al. (2023). Clinical Profile of Patients with Idiopathic Pulmonary Fibrosis in Real Life. *Journal of Clinical Medicine*, 12 (4), 1669. Available from https://doi.org/10.3390/jcm12041669.

Nathan, S.D. et al. (2011). Long-term Course and Prognosis of Idiopathic Pulmonary Fibrosis in the New Millennium. *Chest*, 140 (1), 221–229. Available from https://doi.org/10.1378/chest.10-2572.

National Health Service. (2022). Idiopathic Pulmonary Fibrosis - Diagnosis. *nhs.uk*. Available from https://www.nhs.uk/conditions/idiopathic-pulmonary-fibrosis/diagnosis/ [Accessed 18 October 2022].

Nazi, Z.A. et al. (2021). Fibro-CoSANet: pulmonary fibrosis prognosis prediction using a convolutional self attention network. *Physics in Medicine &amp$\mathsemicolon$ Biology*, 66 (22), 225013. Available from https://doi.org/10.1088/1361-6560/ac36a2.

Niknejad, M.T. (2022). Pulmonary fibrosis | Radiology Reference Article. *Radiopedia*. Available from https://radiopaedia.org/articles/pulmonary-fibrosis [Accessed 18 October 2022].

Pardo, A. and Selman, M. (2012). Role of matrix metaloproteases in idiopathic pulmonary fibrosis. *Fibrogenesis & Tissue Repair*, 5 (S1), S9. Available from https://doi.org/10.1186/1755-1536-5-S1-S9.

Patel, N.M. et al. (2007). Pulmonary Hypertension in Idiopathic Pulmonary Fibrosis. *Chest*, 132 (3), 998–1006. Available from https://doi.org/10.1378/chest.06-3087.

Perera, U.E. et al. (2021). Evaluation of microRNA expression in a sheep model for lung fibrosis. *BMC Genomics*, 22 (1), 827. Available from https://doi.org/10.1186/s12864-021-08073-4.

Pergolizzi, J.V. et al. (2023). What Do We Need to Know About Rising Rates of Idiopathic Pulmonary Fibrosis? A Narrative Review and Update. *Advances in Therapy*, 40 (4), 1334–1346. Available from https://doi.org/10.1007/s12325-022-02395-9.

Potempa, R. and Porebski, S. (2022). Comparing concepts of quantum and classical neural network models for image classification task. 61–71. Available from https://doi.org/10.1007/978-3-030-81523-3_6 [Accessed 20 December 2022].

Preskill, J. (2018). Quantum Computing in the NISQ era and beyond. *Quantum*, 2, 79. Available from https://doi.org/10.22331/q-2018-08-06-79.

Pulmonary Fibrosis Foundation. (2022). What Is Pulmonary Fibrosis? *Default*. Available from https://www.pulmonaryfibrosis.org/understanding-pff/about-pulmonary-fibrosis/what-is-pulmonary-fibrosis [Accessed 31 August 2022].

Qi, F. et al. (2020). Hsa_circ_0044226 knockdown attenuates progression of pulmonary fibrosis by inhibiting CDC27. *Aging*, 12 (14), 14808–14818. Available from https://doi.org/10.18632/aging.103543.

Qi, F. et al. (2023). LncRNA TUG1 promotes pulmonary fibrosis progression via up-regulating CDC27 and activating PI3K/Akt/mTOR pathway. *Epigenetics*, 18 (1), 2195305. Available from https://doi.org/10.1080/15592294.2023.2195305.

Rachel, M. (2020). Detection of Cystic Fibrosis Symptoms Based on X-Ray Images Using Machine Learning- Pilot Study. *Biomedical Journal of Scientific & Technical Research*, 24 (2). Available from https://doi.org/10.26717/BJSTR.2020.24.004019 [Accessed 6 August 2022].

Raghu, G. (2006). High prevalence of abnormal acid gastro-oesophageal reflux in idiopathic pulmonary fibrosis. *European Respiratory Journal*, 27 (1), 136–142. Available from https://doi.org/10.1183/09031936.06.00037005.

Raghu, G. et al. (2011). An Official ATS/ERS/JRS/ALAT Statement: Idiopathic Pulmonary Fibrosis: Evidence-based Guidelines for Diagnosis and Management. *American Journal of Respiratory and Critical Care Medicine*, 183 (6), 788–824. Available from https://doi.org/10.1164/rccm.2009-040GL.

Raghu, G. et al. (2022). Idiopathic Pulmonary Fibrosis (an Update) and Progressive Pulmonary Fibrosis in Adults: An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *American Journal of Respiratory and Critical Care Medicine*, 205 (9), e18–e47. Available from https://doi.org/10.1164/rccm.202202-0399ST.

Richeldi, L. et al. (2014). Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. *New England Journal of Medicine*, 370 (22), 2071–2082. Available from https://doi.org/10.1056/NEJMoa1402584.

Romero, J., Olson, J.P. and Aspuru-Guzik, A. (2017). Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2 (4), 045001. Available from https://doi.org/10.1088/2058-9565/aa8072.

Saunders, M.N.K., Lewis, P. and Thornhill, A. (2007). *Research methods for business students*, 4th ed. Harlow, England ; New York: Financial Times/Prentice Hall.

Schwartz, D.A. (2018). Idiopathic Pulmonary Fibrosis. *National Organization for Rare Disorders (NORD)*. Available from https://rarediseases.org/rare-diseases/idiopathic-pulmonary-fibrosis/ [Accessed 18 October 2022].

Selman, M. (2006). Role of Epithelial Cells in Idiopathic Pulmonary Fibrosis: From Innocent Targets to Serial Killers. *Proceedings of the American Thoracic Society*, 3 (4), 364–372. Available from https://doi.org/10.1513/pats.200601-003TK.

Sengupta, K. and Srivastava, P.R. (2021). Quantum algorithm for quicker clinical prognostic analysis: an application and experimental study using CT scan images of COVID-19 patients. *BMC Medical Informatics and Decision Making*, 21, 227. Available from https://doi.org/10.1186/s12911-021-01588-6.

Simeone, O. (2022). An Introduction to Quantum Machine Learning for Engineers. Available from https://doi.org/10.48550/ARXIV.2205.09510 [Accessed 2 May 2023].

Steele, M.P. et al. (2005). Clinical and Pathologic Features of Familial Interstitial Pneumonia. *American Journal of Respiratory and Critical Care Medicine*, 172 (9), 1146–1152. Available from https://doi.org/10.1164/rccm.200408-1104OC.

Taskar, V. and Coultas, D. (2008). Exposures and Idiopathic Lung Disease. *Seminars in Respiratory and Critical Care Medicine*, 29 (06), 670–679. Available from https://doi.org/10.1055/s-0028-1101277.

Taskar, V.S. (2006). Is Idiopathic Pulmonary Fibrosis an Environmental Disease? *Proceedings of the American Thoracic Society*, 3 (4), 293–298. Available from https://doi.org/10.1513/pats.200512-131TK.

Tiwari, M. and Singhai, Rakesh. (2017). A Review of Detection and Tracking of Object from Image and Video Sequences.

U.S. Department of Health and Human Services. (2011). Definition of alveoli - NCI Dictionary of Cancer Terms - NCI. Available from https://www.cancer.gov/publications/dictionaries/cancer-terms/def/alveoli [Accessed 6 September 2022].

Wang, H., Xu, Y. and Shi, H. (2018). A new approach for reducing beam hardening artifacts in polychromatic X-ray computed tomography using more accurate prior image. *Journal of X-Ray Science and Technology*, 26 (4), 593–602. Available from https://doi.org/10.3233/XST-17325.

Welinder, P. et al. (2010). The Multidimensional Wisdom of Crowds. *Advances in Neural Information Processing Systems*. 2010. Curran Associates, Inc. Available from https://papers.nips.cc/paper_files/paper/2010/hash/0f9cafd014db7a619ddb4276af0d692c-Abstract.html [Accessed 10 May 2023].

Wong, A. et al. (2021). Fibrosis-Net: A Tailored Deep Convolutional Neural Network Design for Prediction of Pulmonary Fibrosis Progression From Chest CT Images. *Frontiers in Artificial Intelligence*, 4, 764047. Available from https://doi.org/10.3389/frai.2021.764047.

Wuyts, W.A. et al. (2016). Daily Home Spirometry: A New Milestone in the Field of Pulmonary Fibrosis. *American Journal of Respiratory and Critical Care Medicine*, 194 (8), 1033–1034. Available from https://doi.org/10.1164/rccm.201605-1090LE.

Zarogoulidis, P. et al. (2023). Positron Emission-Computed Tomography, Cryobiopsy versus Bronchoalveolar Lavage and Computed Tomography Findings for Interstitial Lung Disease: A Long-Term Follow-Up. *Medicina*, 59 (4), 787. Available from https://doi.org/10.3390/medicina59040787.

Zhang, W., Wang, B. and Sun, F. (2017). Recognition Method Based on Wigner-Hough Transform for Poly-Phase Code Radar Signal. *International Journal of Communications, Network and System Sciences*, 10 (08), 128–137. Available from https://doi.org/10.4236/ijcns.2017.108B014.

Zhou, T. et al. (2022). Dense Convolutional Network and Its Application in Medical Image Analysis. *BioMed Research International*, 2022, 1–22. Available from https://doi.org/10.1155/2022/2384830.

APPENDIX A: CONCEPT MAP

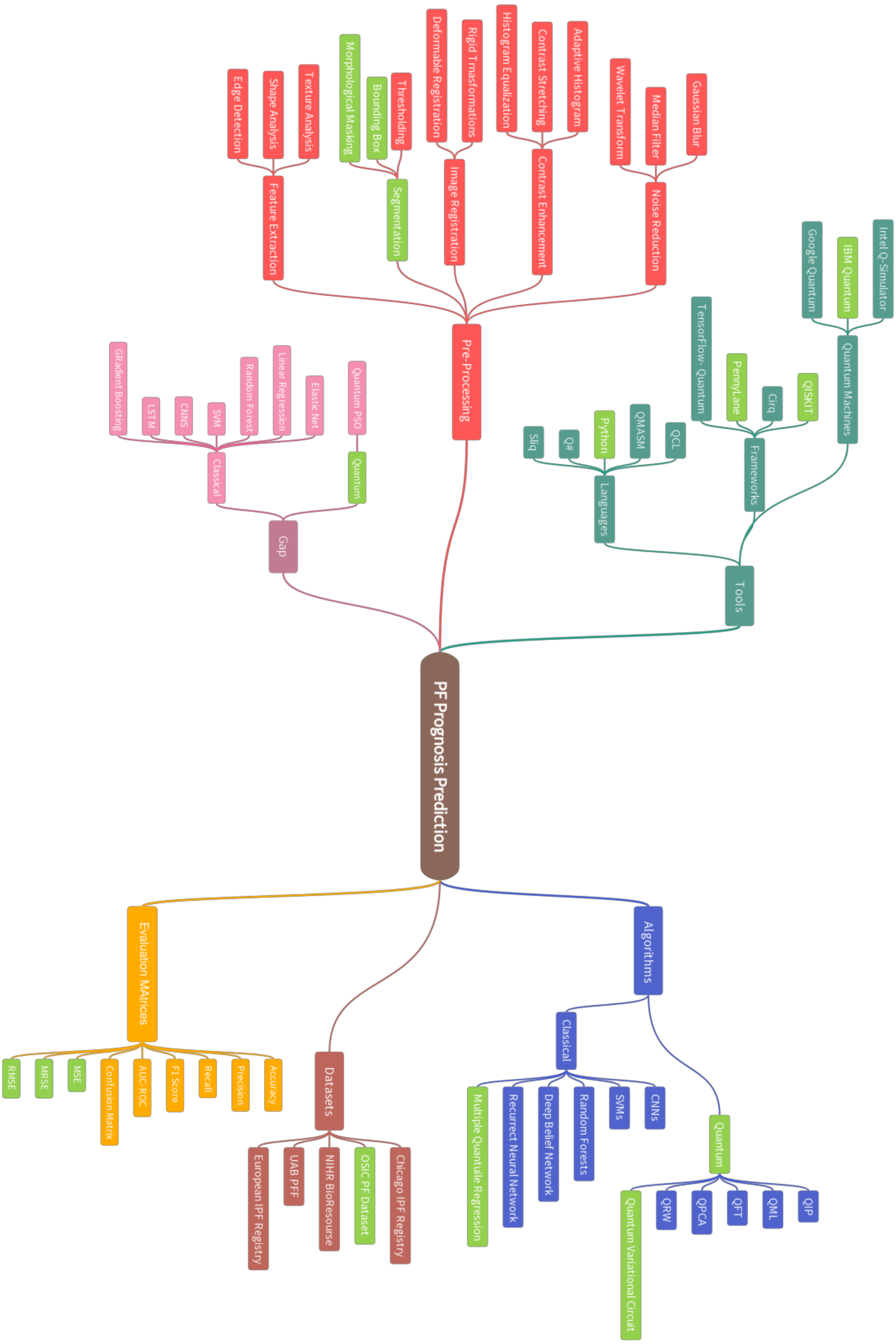


Figure 0.1: Concept Map

APPENDIX B: GANTT CHART

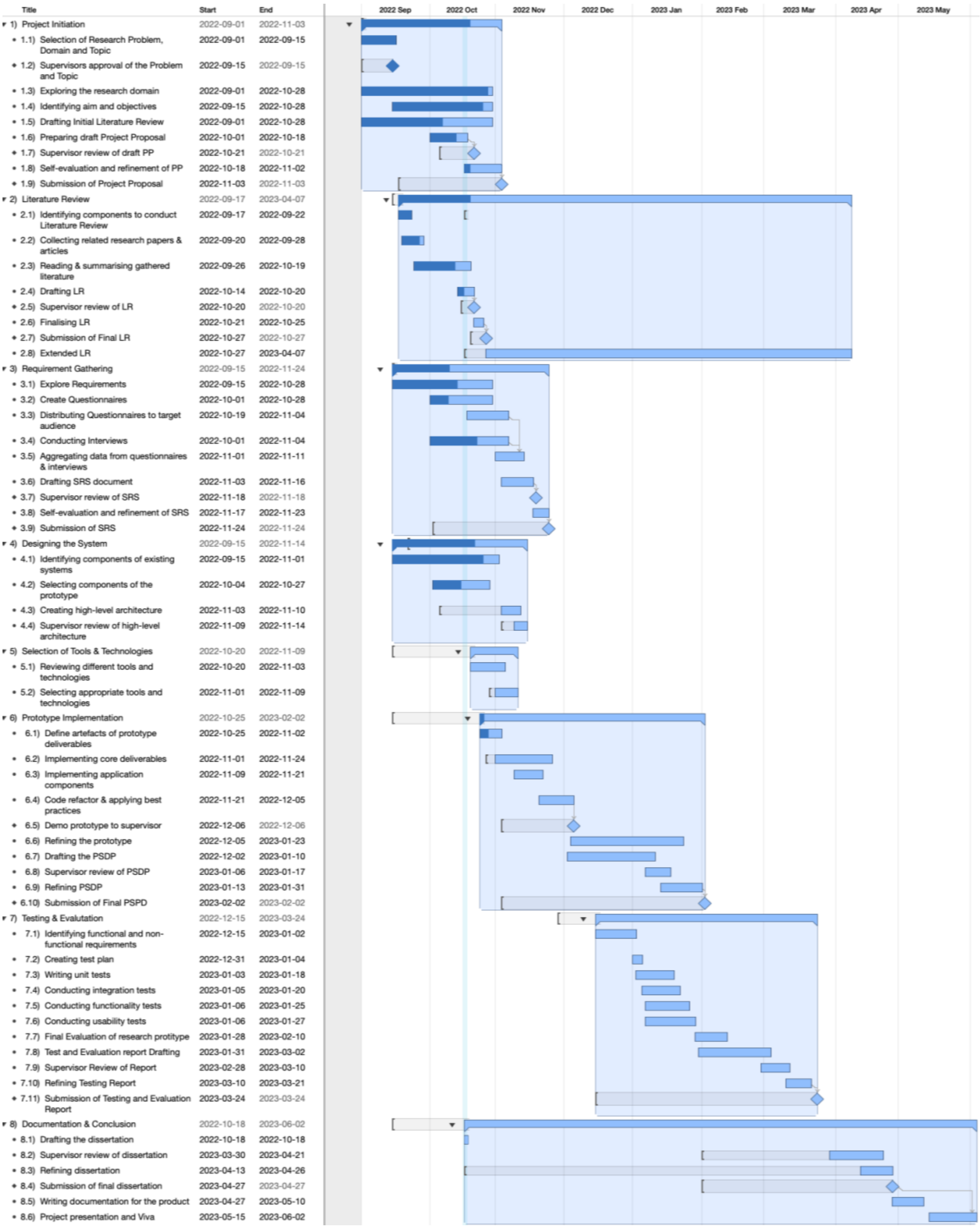


Figure 0.2: Gantt Chart

APPENDIX C: REQUIREMENTS PRIORITY KEY MAPPING

Table 0.1: Requirements Priority Key Mappings

Priority	Description
Must have **(M)**	The core functionality of Fibro-QuanNet and must be implemented.
Should have **(S)**	These features are not core requirements of the application, however, would add value if produced.
Could have **(C)**	These features are desirable requirements that will be considered luxury features and are not essential for the product.
Will not have **(W)**	These features are excluded from the product implementation.

APPENDIX D: SOFTWARE REQUIREMENTS SPECIFICATIONS

D.1 Questionnaires to Experts

Table 0.2: Questionnaire to Technical Experts

Question	Relevance	Research Question
Have you worked with cases regarding pulmonary fibrosis?	To identify whether the target audience has worked with the domain disease.	RQ2
What medical biomarkers have you used to predict PF?	To identify the potential biomarkers to predict PF	RQ1
Which of those biomarkers would you suggest being the best to predict the prognosis of PF?	To identify the requirement for a reliable system	RQ4
Have you performed a prognosis prediction of a PF case manually?	To identify whether the target audience has performed the manual prediction of the automated system.	RQ6
What features do you look for in the lung cavity for effective PF diagnosis?	To identify the technique used to identify PF and recognize effective ways to do so.	RQ3
What features in the lung cavity do you look for when predicting the prognosis of a PF patient?	The identify the technique used to predict the prognosis of patient.	RQ3
How crucial do you think the prediction of PF prognosis is to the medical outcome of a patient?	To identify the crucialness of the research project, emphasizing the significance.	RQ7

How would you rate the difficulty of performing a prognosis prediction?	To identify the difficulty in predicting manually the prognosis.	RQ3
How would you rate the likeliness of a manual prognosis prediction being incorrect?	To identify the requirement for a reliable system	RQ9
Do you think automating the process of prognosis prediction of PF would be worthwhile and benefit patients while saving time for medical professionals?	To identify the requirement for a reliable system	RQ10
Would you be interested in using an automated tool to predict the prognosis of PF automatically?	Validating the need for the system.	RQ8
Have you used or heard about any other applications that can predict the prognosis of PF?	To identify potential competitors	RQ2
What features did you like about while using those existing applications?	To identify the difficulty in predicting manually the prognosis.	RQ4
What downsides did you face when using these existing applications?	Validating the need for the system.	RQ6
What other features would you like to see in an application that can predict the prognosis of pulmonary fibrosis? This may not always be the best option for pulmonary fribrosis preeiction however.	To identify the difficulty in predicting manually the prognosis.	RQ5

Table 0.3: Questionnaire to Domain Experts

Question	**Relevance**	**Research Question**
Have you worked with quantum computing?	To identify the requirement for a reliable system	RQ3
What quantum machines have you used which are open for public use?	Validating the need for the system.	RQ9
Have you worked with any quantum machine learning algorithms?	To identify potential competitors	RQ10
Have you worked with any quantum image recognition-based machine learning?	To identify the difficulty in predicting manually the prognosis.	RQ8
Have you worked with any transfer learning for quantum machine learning?	Validating the need for the system.	RQ2
Have you worked on QML for the MedTech industry?	To identify the difficulty in predicting manually the prognosis.	RQ4
What quantum ML algorithms/ models for image recognition do you think would help PF prognosis prediction achieve the quantum advantage?	To identify the requirement for a reliable system	RQ6
Have you worked with pre-processing models for HRCT medical imagery?	To identify the requirement for a reliable system	RQ3
What pre-processing library would you say is best for HRCT/ any medical biomarking?	Validating the need for the system.	RQ9

What do you think is the best approach is for developing a prediction model using image recognition?	To identify potential competitors	RQ10
What evaluation methodologies have you used in your ML projects?	To identify the difficulty in predicting manually the prognosis.	RQ8
Which evaluation method do you think will be best for a prediction model?	To identify the requirement for a reliable system	RQ3
Have you developed/ come across any quantum ML models capable of predicting the prognosis of pulmonary fibrosis?	Validating the need for the system.	RQ9

D.2 Use Cases

Table 0.4: Use Case Description U0:02

Use Case	View Prediction Results	
Id	UC:02	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**

	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.5: Use Case Description U0:03

Use Case	View Prediction Results	
Id	UC:03	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization

Alternate Flows	Actor	System
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.6: Use Case Description U0:04

Use Case	View Prediction Results	
Id	UC:04	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations

Included Use Cases	UC:08, UC:09, UC:10

Table 0.7: Use Case Description U0:05

Use Case	View Prediction Results	
Id	UC:05	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.8: Use Case Description U0:06

Use Case	View Prediction Results	
Id	UC:06	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.9: Use Case Description U0:08

Use Case	View Prediction Results
Id	UC:08
Description	View the prognosis prediction results

Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.10: Use Case Description U0:09

Use Case	View Prediction Results
Id	UC:09
Description	View the prognosis prediction results
Primary Actor	User
Supporting Actors	None
Stakeholders	Pulmonologist, Radiologist, Medical Student

Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.11: Use Case Description U0:10

Use Case	View Prediction Results
Id	UC:10
Description	View the prognosis prediction results
Primary Actor	User
Supporting Actors	None
Stakeholders	Pulmonologist, Radiologist, Medical Student
Pre-condition	Successful UC:01 and attaches
Post Condition	Success End: The prediction model successfully concludes the prediction
Trigger	User Action

Main Flow	Actor	System
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.12: Use Case Description U0:11

Use Case	View Prediction Results	
Id	UC:11	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result

		4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations
Included Use Cases	UC:08, UC:09, UC:10	

Table 0.13: Use Case Description U0:13

Use Case	View Prediction Results	
Id	UC:13	
Description	View the prognosis prediction results	
Primary Actor	User	
Supporting Actors	None	
Stakeholders	Pulmonologist, Radiologist, Medical Student	
Pre-condition	Successful UC:01 and attaches	
Post Condition	Success End: The prediction model successfully concludes the prediction	
Trigger	User Action	
Main Flow	**Actor**	**System**
	1. User selects results view 5. Selects to download the report	2. Checks for percentage in quantum model 3. Output the model result 4. Output the data visualization
Alternate Flows	**Actor**	**System**
	1. User selects past prediction 4. Can download the older report	2. Extracts data from the database 3. Produces the visualizations

Included Use Cases	UC:08, UC:09, UC:10

APPENDIX E: DESIGN

E.1 Design Goals

Table 0.14: Design Quality Attributes

Quality Attribute	Description
Correctness	The correctness of the predictions made by Fibro-QuanNet must be of highest accuracy possible by utilizing the latest and all available data. This has been addressed in the architecture by specially allocating a presentation tier for a confidence matrix of the prediction performed.
Performance Efficiency	The architecture of Fibro-QuanNet is built such that not all the models utilize the scarce availability of quantum allocation, reserving it for the more crucial models ensuring resource availability.
Scalability	Since quantum computing is still in its early development stages, scalability will be crucial. The architecture allows scalability at its core by segmenting the model components for separate quantum allocations.
Flexibility	Fibro-QuanNet can be identified as one of the most flexible applications due to its operability on any device or OS smoothly. This is achieved in the architecture by utilizing remote computational devices for the foreseeable future with QC still in its early ages.
Reusability	The architecture of the system must provide compartmentalized, reusable components which will enable the system to be updated and facilitate the addition of new functionality and use cases, as required through the adaptability goals which will be in-return benefit both patients and the product champion.

APPENDIX F: IMPLEMENTATION

F.1 Implementation Code Snippets

```python
class CropBoundingBox:
    @staticmethod
    def bounding_box(img3d: np.array):
        mid_img = img3d[int(img3d.shape[0] / 2)]
        same_first_row = (mid_img[0, :] == mid_img[0, 0]).all()
        same_first_col = (mid_img[:, 0] == mid_img[0, 0]).all()
        if same_first_col and same_first_row:
            return True
        else:
            return False

    def __call__(self, sample):
        image = sample['image']
        if not self.bounding_box(image):
            return sample

        mid_img = image[int(image.shape[0] / 2)]
        r_min, r_max = None, None
        c_min, c_max = None, None
        for row in range(mid_img.shape[0]):
            if not (mid_img[row, :] == mid_img[0, 0]).all() and r_min is None:
                r_min = row
            if (mid_img[row, :] == mid_img[0, 0]).all() and r_max is None \
                    and r_min is not None:
                r_max = row
                break

        for col in range(mid_img.shape[1]):
            if not (mid_img[:, col] == mid_img[0, 0]).all() and c_min is None:
                c_min = col
            if (mid_img[:, col] == mid_img[0, 0]).all() and c_max is None \
                    and c_min is not None:
                c_max = col
                break

        image = image[:, r_min:r_max, c_min:c_max]
        return {
            'features': sample['features'],
            'image': image,
            'metadata': sample['metadata'],
            'target': sample['target']
        }
```

Figure 0.3: Bounding Box Implementation Code Snippet

```python
class ConvertToHU:
    def __call__(self, sample):
        image, data = sample['image'], sample['metadata']

        img_type = data.ImageType
        is_hu = img_type[0] == 'ORIGINAL' and not (img_type[2] == 'LOCALIZER')
        if not is_hu:
            warnings.warn(f'Patient {data.PatientID} CT Scan not cannot be'
                          f'converted to Hounsfield Units (HU).')

        intercept = data.RescaleIntercept
        slope = data.RescaleSlope
        image = (image * slope + intercept).astype(np.int16)
        return {
            'features': sample['features'],
            'image': image,
            'metadata': data,
            'target': sample['target']
        }
```

Figure 0.4: Hounsfield Units Code Snippet

```python
class Resize:
    def __init__(self, output_size):
        assert isinstance(output_size, tuple)
        self.output_size = output_size

    def __call__(self, sample):
        image = sample['image']
        resize_factor = np.array(self.output_size) / np.array(image.shape)
        image = zoom(image, resize_factor, mode='nearest')
        return {
            'features': sample['features'],
            'image': image,
            'metadata': sample['metadata'],
            'target': sample['target']
        }
```

Figure 0.5: Resizing Code Snippet

```python
class Mask:
    def __init__(self, threshold=-400,
                 root='../data/test'):
        self.threshold = threshold
        self.root = root

    def __call__(self, sample):
        image = sample['image']
        for slice_id in range(image.shape[0]):
            m = self.get_morphological_mask(image[slice_id])
            image[slice_id][m == False] = image[slice_id].min()

        return {
            'features': sample['features'],
            'image': image,
            'metadata': sample['metadata'],
            'target': sample['target']
        }

    def get_morphological_mask(self, image):
        m = image < self.threshold
        m = clear_border(m)
        m = label(m)
        areas = [r.area for r in regionprops(m)]
        areas.sort()
        if len(areas) > 2:
            for region in regionprops(m):
                if region.area < areas[-2]:
                    for coordinates in region.coords:
                        m[coordinates[0], coordinates[1]] = 0
        return m > 0
```

Figure 0.6: Morphological Masking Code Snippet

```python
class Normalize:
    def __init__(self, bounds=(-1000, 500)):
        self.min = min(bounds)
        self.max = max(bounds)

    def __call__(self, sample):
        image = sample['image'].astype(np.float)
        image = (image - self.min) / (self.max - self.min)
        return {
            'features': sample['features'],
            'image': image,
            'metadata': sample['metadata'],
            'target': sample['target']
        }

class ToTensor:
    def __init__(self, add_channel=True):
        self.add_channel = add_channel

    def __call__(self, sample):
        image = sample['image']
        if self.add_channel:
            image = np.expand_dims(image, axis=0)

        return {
            'features': sample['features'],
            'image': torch.from_numpy(image),
            'metadata': sample['metadata'],
            'target': sample['target']
        }

class ZeroCenter:
    def __init__(self, pre_calculated_mean):
        self.pre_calculated_mean = pre_calculated_mean

    def __call__(self, sample):
        return {
            'features': sample['features'],
            'image': sample['image'] - self.pre_calculated_mean,
            'metadata': sample['metadata'],
            'target': sample['target']
        }
```

Figure 0.7: Tensor Normalization Code Snippet

```python
class AutoEncoder(nn.Module):
    def __init__(self):
        super(AutoEncoder, self).__init__()
        # Encoder
        self.conv1 = nn.Conv3d(1, 16, 3)
        self.conv2 = nn.Conv3d(16, 32, 3)
        self.conv3 = nn.Conv3d(32, 96, 2)
        self.pool1 = nn.MaxPool3d(kernel_size=2, stride=2, return_indices=True)
        self.pool2 = nn.MaxPool3d(kernel_size=3, stride=3, return_indices=True)
        self.pool3 = nn.MaxPool3d(kernel_size=2, stride=2, return_indices=True)
        # Decoder
        self.deconv1 = nn.ConvTranspose3d(96, 32, 2)
        self.deconv2 = nn.ConvTranspose3d(32, 16, 3)
        self.deconv3 = nn.ConvTranspose3d(16, 1, 3)
        self.unpool1 = nn.MaxUnpool3d(kernel_size=2, stride=2)
        self.unpool2 = nn.MaxUnpool3d(kernel_size=3, stride=3)
        self.unpool3 = nn.MaxUnpool3d(kernel_size=2, stride=2)

    def encode(self, x, return_partials=True):
```

Figure 0.8: Auto Encoder Pre-processor Code Snippet

```python
# Helper function that generates all latent features
class GenerateLatentFeatures:
    def __init__(self, autoencoder, latent_dir):
        self.autoencoder = autoencoder
        self.latent_dir = Path(latent_dir)
        self.cache_dir = Path(cache_dataset)

    def __call__(self, sample):
        patient_id = sample['metadata'].PatientID
        cached_latent_file = self.latent_dir/f'{patient_id}_lat.pt'

        if cached_latent_file.is_file():
            latent_features = torch.load(cached_latent_file)
        else:
            with torch.no_grad():
                img = sample['image'].float().unsqueeze(0)
                latent_features = self.autoencoder.encode(
                    img, return_partials=False).squeeze(0)
            torch.save(latent_features, cached_latent_file)

        return {
            'tabular_features': sample['features'],
            'latent_features': latent_features,
            'target': sample['target']
        }
```

Figure 0.9: Latent Feature Extraction Code Snippet

```python
# Prognosis prediction route
@app.route('/predict', methods=['POST'])
def prognosis():
    # check if the post request has the file part
    if 'files[]' not in request.files:
        resp = jsonify({
            "message": 'No file part in the request',
            "status": 'failed'
        })
        resp.status_code = 400
        return resp

    files = request.files.getlist('files[]')

    errors = {}
    success = False

```

Figure 0.10: End-point Implementation Code Snippet

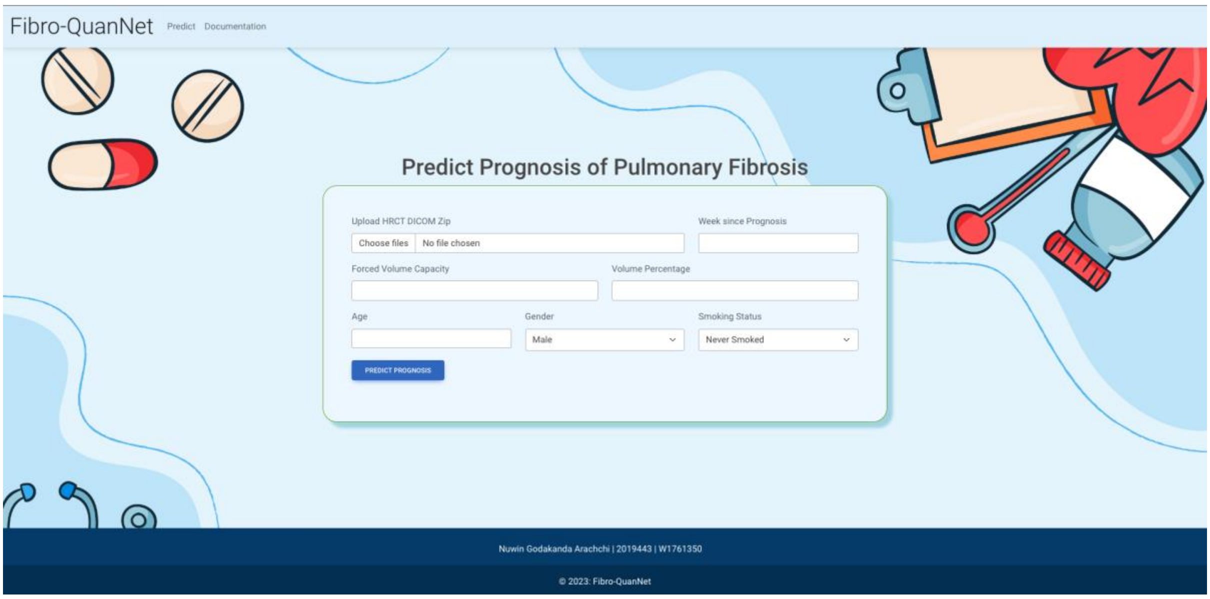


Figure 0.11: Upload HRCT Predict UI Screen

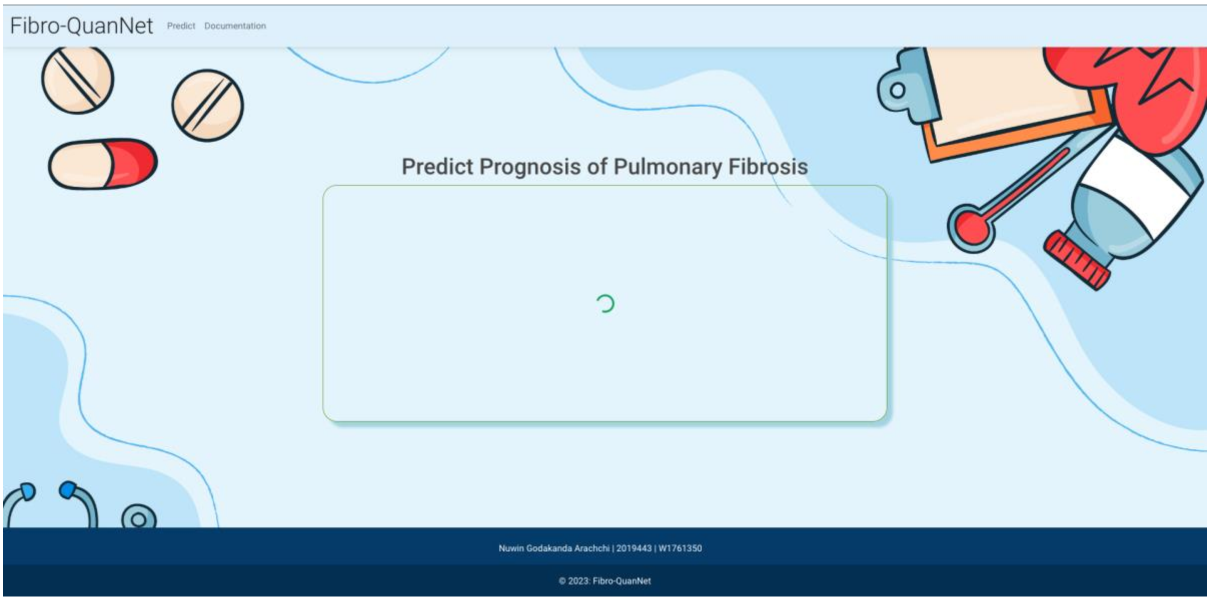


Figure 0.12: Processing Prognosis Prediction UI

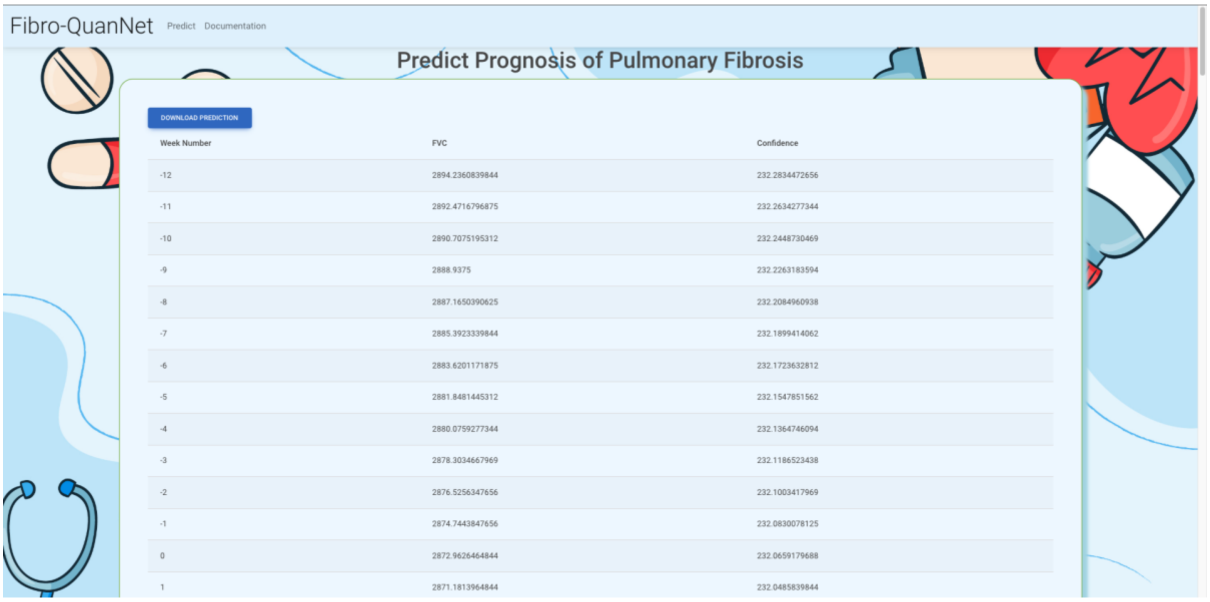


Figure 0.13: View Predictions and Download CSV UI

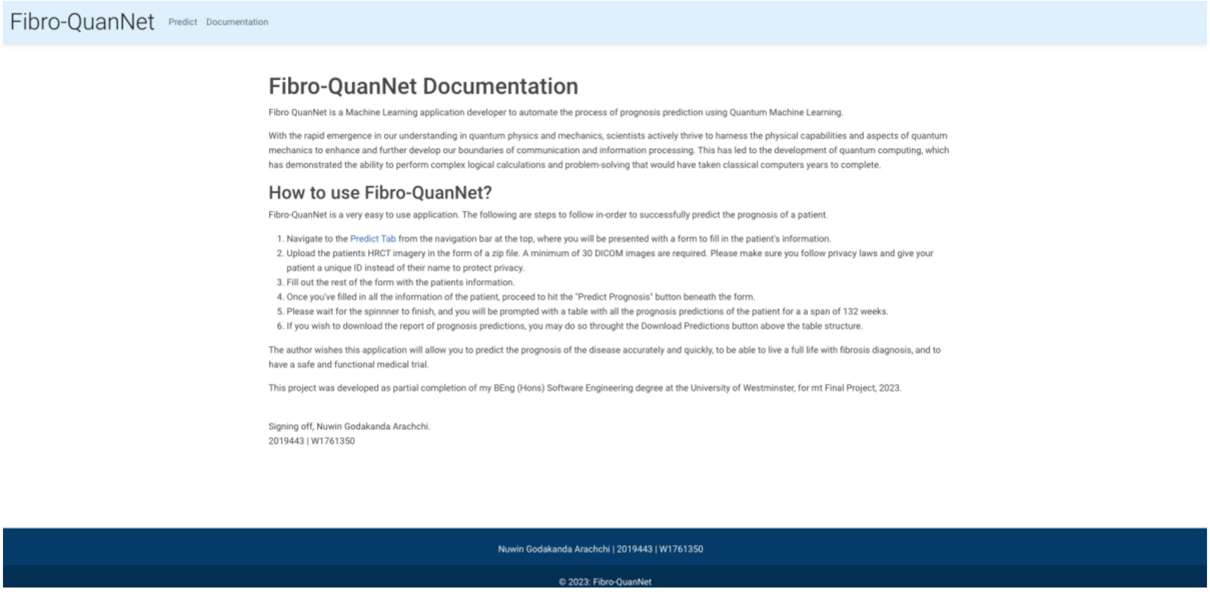


Figure 0.14: Documentation UI

APPENDIX G: EVALUATIONS

G.1 Functional Requirements Evaluation

Table 0.15: Full List of the Functional Requirements Evaluation

FR-ID	Requirement	Use Case	Priority Level	Evaluation
FR01	User must be able to upload HRCT imagery to the system	UC:01	M	Implemented
FR02	The user must be able to enter prevailing FVC and metadata to the system	UC:03	M	Implemented
FR03	The system must be able to pre-process the HRCT	UC:02	M	Implemented
FR04	The system must be able to detect and predict the prognosis of PF in the input HRCT	UC:05	M	Implemented
FR05	The system should output the confidence matrix of the prognosis prediction	UC:09	S	Implemented
FR06	The user should be able to view past prognosis predictions	UC:10	C	Not Considered
FR07	The user should be able to download prognosis prediction report	UC:11	S	Implemented
FR08	The system should choose and borrow most efficient quantum-processing time from the fleet of available quantum machines	UC:05	C	Accessed Successfully
FR09	The system should have a user-friendly graphical user interface	UC:06	S	Implemented
FR10	The system should be able to view the current progression of PF	UC:08	C	Implemented

FR11	The system could produce, and print reporting based on the prognosis prediction	UC:11	C	Implemented
FR12	The user should be able to elect to contribute HRCT data	UC:07	C	Not Considered
FR13	The system will not show the reasoning behind the prognosis prediction	NA	W	Not Considered
FR14	The system will not take in any other biomarkers as input	NA	W	Not Considered
Function requirements completion percentage: 10/12 = 83.3*%				

G.2 Non-Functional Requirements Evaluation

Table 0.16: Full List of Non-Functional Requirements Evaluation

NFR-ID	Requirement	Description	Priority	Evaluation
NFR01	Quality of Output	The accuracy of the prognosis predictions must be of highest quality possible due to its key-role in patient care. The system must be able to produce a quality attribute along with the PP describing the accuracy and quality of the prediction.	Crucial	Implemented
NFR02	Performance	The system must be able to produce prediction in a much faster and with higher precision using and achieving the quantum advantage for PF PP.	Crucial	Implemented
NFR03	Usability	Since the system caters for pulmonologists and radiologists who can be of all ages and abilities, the	Crucial	Implemented

		system must incorporate an easy-to-use user interface, which will abide by all design guidelines for accessibility.		
NFR04	Scalability	Since QC is still at its early ages, and PF still being researched on heavily for treatments, much more functionalities can be added to the system in the future.	Desirable	Implemented
Non-Functional Requirements completion percentage: 4/4 = 100%				

APPENDIX H: CONCLUSION

H.1 Achievement status of the Research Objectives

Table 0.17: Achievement Status of the Research Objectives

Objective	Description	Learning Outcomes	Status
Problem Identification	Carry out the following tasks to identify the problem. • **RO1**: Research interested domains and identify a potential problem which may be feasible to solve within the limited time constraints and technologies.	L01	Completed
Literature Survey	Carry out an in-depth review of the following areas, • **RO2**: Analyze and understand fundamental concepts of quantum machine learning and understanding training models and working with them. • **RO3**: Conduct preliminary studies into existing prognosis prediction systems. • **RO4**: Study on Computer Tomography to understand the need and technique of using HRCT in PF prognosis. • **RO5**: Analyzing existing models and identifying their limitations. • **RO6**: Critical review of the literature and elaborate on the research gap and methodologies.	L01, L04, L05	Completed

Requirement Analysis	Carry out in-depth user requirement gathering in the following areas, • **RO7**: Understand and gather requirements users may expect from a prognosis prediction system for PF. • **RO8**: Get insight and opinion from Pulmonologists, Radiologists, and quantum data scientists to build the system and mitigate any legal/ social/ ethical issues. • **RO9**: Identifying the tools and techniques (software requirements) and expected behavior for the system through questionnaires.	L01, L02, L03, L06	Completed
Design	Develop the design architecture of the proposed system, capable of solving the gap. • **RO10**: Design a prognosis prediction system to demonstrate the FVC and HRCT data. • **RO11**: Design a data-preprocessing pipeline for HRCT imagery data feed. • **RO12**: Design the QML prediction model which can produce the prognosis prediction.	L01	Completed
Development	Implement a system that's capable of addressing the gap aimed to solve. • **RO13**: Develop a prognosis prediction system that can predict the prognosis of PF efficiently and quickly.	L07, L05	Completed

	• **RO14**: Develop the QML model that can use quantum super-states to produce quick and efficient predictions. • **RO15**: Develop a pre-processing pipeline for HRCT imagery feed data. • **RO16**: Develop the hyperparameter tuning component that improves the prediction system.		
Testing & Evaluation	Testing and evaluating the prototype. • **RO17**: Create a test plan and perform unit, integration, and functional testing of the prediction system. • **RO18**: Evaluate how efficient quantum models maybe be in comparison to classical models to perform predictions. • **RO19**: Perform requirement validation against all requirements and evaluate accuracy measures.	L04	Completed
Publish Findings	Produce and publish well-structured papers that will critically evaluate and review the research area. • **RO20**: Publishing evaluation and testing results of the project system. • **RO21**: Making code and models created during the project open-source and publicly available for future work.	L06, L08	Completed