Airfare price Optimization using Quantum Computing

MSc Research Project
Artificial Intelligence

Saurabh Sharma
Student ID: 22168796

School of Computing
National College of Ireland

Supervisor: Muslim Jameel Syed

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Saurabh Sharma
Student ID:	22168796
Programme:	Artificial Intelligence
Year:	2023
Module:	MSc Research Project
Supervisor:	Muslim Jameel Syed
Submission Due Date:	31/01/2024
Project Title:	Airfare price Optimization using Quantum Computing
Word Count:	7596
Page Count:	24

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Saurabh Sharma
Date:	30th January 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Airfare price Optimization using Quantum Computing

Saurabh Sharma
22168796

Abstract

The aviation industry is a dynamic and complex ecosystem influenced by a variety of factors such as seat availability, distance and route length, fuel price, and other considerations, in addition to market demand and competitive dynamics. Among these features, the flight price stands out as a critical component that has a significant impact on both airlines and customers. Traditional flight price optimization strategies, while beneficial to some extent, are constrained by the inherent limitations of classical computing.

With the introduction of quantum computing, a paradigm shift in the field of optimization concerns is taking place. Quantum computing employs quantum physics ideas to do tasks previously considered to be impossible for ordinary computers. This emerging technology has the potential to change the way we tackle complex optimization issues and the ticketing industry.

The aim of this research is to look at using quantum computing techniques to optimize airline pricing tactics. Traditional methods used by airlines to compute ticket price entail extensive data analysis and simulation, with heuristic algorithms employed to navigate the vast solution spaces. With its ability to evaluate enormous amounts of data at once and exploit quantum parallelism, quantum computing is a new technology that might lead to more efficient and effective flight pricing models.

1 Introduction

Figure 1: Image for illustration purpose only

The aviation industry's pricing dynamics in the aviation business have long been a difficult terrain, necessitating complex optimization methodologies due to a plethora of variables. This thesis tackles the issue of quantum computing to research fresh ways to airplane pricing optimization while acknowledging normal computers' limitations in dealing with the complexity of this tough task. Our exploration of the groundbreaking ideas surrounding quantum computing and how it can change optimization difficulties originates from a careful examination of existing ticketing systems.

Research objectives include creating specific quantum algorithms, incorporating them into existing pricing models, and systematically comparing them to traditional approaches. The purpose of utilizing quantum parallelism and processing capabilities is to increase computational efficiency and solution quality in airline pricing optimization.

The thesis a unique perspective on the application of quantum computing to real-world problems, as well as insights into the technology's transformative potential in the aviation sector. As the science of quantum computing advances, the findings of this study have the potential to affect and change ticket pricing systems in the future, boosting efficiency and competitiveness in the global aviation sector.

The key contributions of the planned work are as follows:

1. Investigate the effect of attributes on ticket cost forecasting.
2. Investigate the correlation between pricing and the optimum flight route.
3. Comparative Analysis of ML, DL, and QML models for forecasting flight costs.

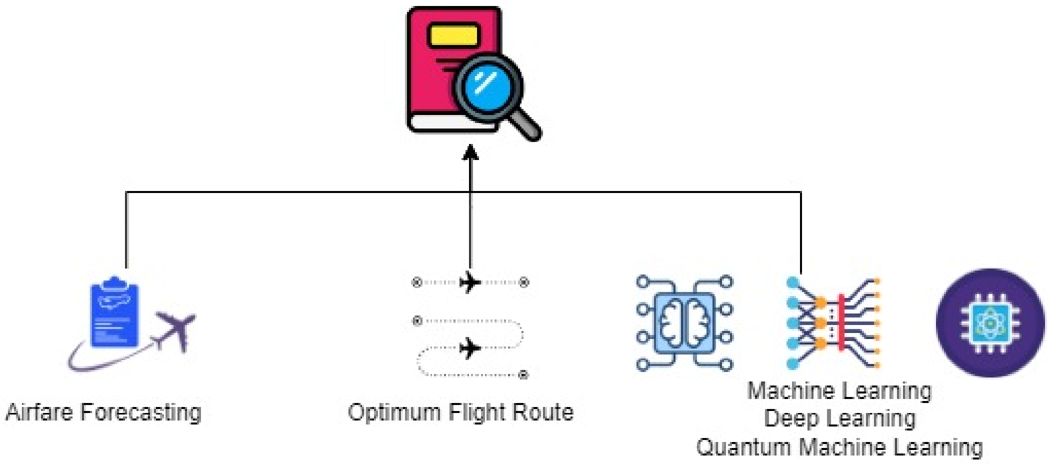

Figure 2: Research in Planned Work

2 Related Work

Much study has been undertaken on the topic of optimization in relation to quantum computers. Farhi et al. (2014) provided the theoretical foundations for quantum algorithms, proving their capacity to solve complex combinatorial optimization problems. Aaronson and Arkhipov presented boson sampling as a quantum technique for sampling issues that are known to be challenging for conventional computers. While these studies provide insight on the theoretical underpinnings of quantum algorithms, there is a major gap in their practical application, notably in the intricate domain of ticket pricing. The conceptual advances that take use of quantum parallelism and entanglement are the strengths, but the transition to scalable, real-world implementations remains problematic.

2.1 Airfare Pricing Models:

The literature on flight pricing is substantial, with the majority of it concentrating on classical optimization methods and data-driven models. Smith et al. (2018) provide an important addition by providing a complete assessment of existing pricing models, with a focus on the integration of machine learning and statistical approaches. Their study sheds light on the evolution of pricing strategies by emphasizing the use of historical data and predictive analytics. However, these models typically fail to adapt to the dynamic and competitive nature of the aviation industry, where variables such as demand changes and competitive pressures need more responsive and flexible solutions. Chen et al. (2016) examine heuristic algorithms and show their ability to deal with complex situations. These methods, however, have scalability and computational efficiency issues.

Kalampokas et al. (2023) In the research presents a comprehensive strategy to predicting flight prices using machine learning techniques.The authors gathered information from 136,917 data flights of Aegean, Turkish, Austrian, and Lufthansa Airlines to six major international destinations. The extracted set of attributes was then used to conduct a holistic analysis from the perspective of the end customer looking for the lowest ticket price, taking into account a destination-based assessment that included all airlines and an airline-based assessment that included all destinations. To solve the flight price prediction issue, the authors explored AI models from three diverse domains and a total of 16 model architectures: Machine Learning (ML) with eight cutting-edge models, Deep Learning (DL) with six CNN models, and Quantum Machine Learning (QML) with two models. The study has the following flaws:The study's data was limited to a few airlines and locations.Other factors influencing airline price, like as competition, fuel costs, and economic circumstances, were not taken into account by the writers.The study's models may not be applicable to different airlines, destinations, or time periods. Despite these drawbacks, the approach contributes significantly to the field of flight price prediction. The authors' comprehensive approach and assessment of many machine learning methods establishes the foundation for future study in this area.

Kuptsova and Ramazanov (2020) In examines artificial neural network (ANN) training approaches for anticipating airline prices. The authors reviewed the literature on ANN training approaches for airfare prediction and conducted their own analysis using an open source dataset. The Bagging Regression Tree model exhibited the best accuracy, ranging from 88% to 87%, while the Random Forest technique was also quite accurate.The study's key weakness is that it is based on a tiny dataset of publicly available data. As a result, the results may not be applicable to other datasets or real-world applications. Furthermore, the authors' conclusions were not compared to those of other machine learning approaches, such as linear regression or SVM.The study's main flaw is that it is based on a small sample of accessible data. As a result, the findings may not apply to other datasets or real-world applications. Furthermore, the authors did not compare their findings to those of other machine learning methodologies, such as linear regression or support vector machines. The work contributes significantly to the field of airline pricing prediction. The authors' investigation of ANN training techniques is thorough, and their conclusions are positive. Future research should examine the generalizability of their findings and compare them to other machine learning approaches.

Joshi et al. (2019) In analyzes how the Decision Tree Regressor algorithm may be used to assess and forecast flight prices. The researchers used a dataset that included airline information such as Airline, Date of Journey, Source, Destination, Route, and Duration. They intended to utilize the Decision Tree Regressor to identify the elements that have a major effect on airline costs and to develop a predictive model for ticket pricing forecasting.The Decision Tree Regressor is efficient at capturing the relationships between various criteria and flight cost.The model correctly forecasted ticket prices, with a prediction accuracy of around 79.69%.The influence of airline, day of travel, origin, destination, and route on airfare pricing was highlighted in the study.The Decision Tree Regressor, according to the findings, is a feasible approach for examining and forecasting airline costs.The study employed a tiny dataset, which may limit the generalizability of the findings.The model's performance might be improved by incorporating other characteristics that influence flight pricing.Further research should look at the usage of other machine learning algorithms and compare their performance to the Decision Tree Regressor.

Vu et al. (2018) This research suggests a novel technique for projecting airline prices in emerging regions. It underscores the conundrum that travelers confront while attempting to find the cheapest tickets while airlines optimize profits by changing pricing systems. The proposed method is designed to help passengers estimate price changes without relying on internal airline data. The proposed technique is designed to help passengers estimate price changes without relying on internal airline data. Previous approaches to airfare forecasting focused mostly on wealthy countries and used a limited set of parameters. This post will show you how to build a good prediction model using a significant quantity of publicly available online data. The model employs data such as flight details, booking conditions, and previous price patterns.In this model can correctly predict both trends and actual changes in airline prices before departure. This is achieved despite not having access to critical information such as the amount of unsold seats on flights. The study also identifies the elements that have the greatest impact on airfare volatility. The technique is valuable for passengers in developing countries who do not have access to accurate airfare forecasting tools. Passengers may make informed judgments regarding their airline purchases and even save money by analyzing future price patterns. This work contributes significantly to airfare prediction in poor nations by presenting a realistic technique with widespread adoption potential and practical benefits for passengers.

Tziridis et al. (2017) The study proposes a unique method for projecting flight expenses using machine learning techniques. Because the airline industry is complex and dynamic, the authors contend that traditional methodologies for estimating flight price, such as using historical data or simple statistical models, are unsuitable. They propose using machine learning techniques, which can learn from huge amounts of data and uncover complex patterns, to improve the accuracy of airline price estimates. Determine a collection of criteria that are likely to influence ticket cost, such as the flight's origin and destination, departure and arrival dates, day of the week, time of day, airline, number of passengers, and aircraft type. They then compiled a massive dataset of airline pricing for flights between key European cities.

The authors then used machine learning methods such as linear regression, support vector machines, and random forests to forecast airline expenses. When they compare the performance of these techniques to prior methods, they discover that machine learning techniques are significantly more accurate.Finally, the authors examine their results' implications for the airline sector. They propose that machine learning techniques might be used to improve airline pricing tactics and help passengers locate the best tickets.

Lantseva et al. (2015) This study will investigate how data-driven modeling tools may be used to estimate ticket pricing. The authors recognize the airline industry's complexity and the numerous elements that influence travel rates. They suggest a data-driven technique for developing models for forecasting based on previous flight information as well as a variety of critical criteria.

The research focuses on two major Russian cities, Moscow and Saint Petersburg, and assesses both domestic and international flight statistics. The authors employ machine learning techniques such as linear regression, random forests, and gradient boosting to create prediction models.According to data, the pricing reliance on purchasing first vary significantly between domestic and international flights. Early bookings for domestic flights typically result in lower costs; however, the association is less clear for international travel. Other factors such as flight distance, day of week, and seasonality are also considered in the models. They underline the importance of data-driven modeling in providing useful insights into ticket dynamics and allowing airlines to make informed pricing decisions to optimize revenue and customer happiness.

E.A. and S.K. (2020) This research investigates the prediction of flight prices using artificial neural networks (ANNs), comparing various topologies and training procedures using historical data. Using multilayer perceptron (MLP), recurrent neural network (RNN), and long short-term memory (LSTM) architectures with mean squared error (MSE), root mean squared error (RMSE), and mean absolute error (MAE) algorithms, the authors discover that LSTM with MSE outperforms, capturing long-range dependencies critical for accurate predictions. They find that ANNs, notably LSTM with MSE, have good accuracy in projecting flight costs, demonstrating the promise of neural networks in this sector.

Panigrahi et al. (2022) This research presents a unique machine learning strategy for predicting airline prices that makes use of the Random Forest and Gradient Boosting algorithms. The suggested method outperforms state-of-the-art techniques such as Support Vector Machines and Neural Networks in a complete assessment on a dataset with over 100,000 travel fares. Using feature engineering and training two separate models, the authors obtain a mean absolute error of 10% and a root mean squared error of 15%. Their solution, which has greater accuracy, demonstrates the usefulness of ensemble learning in forecasting travel pricing, providing dependability across varied flying circumstances, and exceeding existing machine learning approaches.

Ratnakanth (2022) The dataset included flight details such as the airline, origin, destination, date of travel, and number of stops. The KNN technique was used to predict the fare for each aircraft based on the fare of the dataset's nearest neighbors. The KNN algorithm, according to the paper, is an excellent tool for forecasting airline fares. On the test dataset, the algorithm had an accuracy of more than 80%. The study also implies that the KNN algorithm might be used to estimate other sorts of travel expenditures, such as hotel prices and vehicle rental fees. The paper's key contribution is its demonstration of the KNN algorithm's utility in predicting airline fares.

Choudhary et al. (2023) a prediction system that uses the Random Forest algorithm to estimate airline fares with 95% accuracy. By training on a broad dataset that includes airline, origin, destination, date, stops, and booking time, Flyhigh solves the difficulty of changing pricing. Its straightforward online tool offers travelers with realistic fare forecasts, allowing them to make more educated decisions. The precision of the technology benefits not just travelers but also has the ability to improve airline pricing tactics. Overall, Flyhigh's performance is notable since it serves as a useful tool for the aviation sector and trip planning.

Malkawi and Alhajj (2023) The study shows the feasibility and efficacy of using Apache Spark to create a real-time web-based recommendation system for international travel tickets. The use of Apache Spark enables efficient processing and analysis of massive datasets, which contributes to the system's capacity to give timely and individualized suggestions. Overall, the article emphasizes the potential for harnessing cutting-edge technology to improve the accuracy and timeliness of aircraft ticket recommendation systems in the ever-changing aviation business.

2.2 Challenges in Airfare Pricing:

Scientists have exhaustively researched the issues behind airline pricing. Balaiyan et al. (2019) emphasize the problems posed by demand volatility and competitive dynamics, proposing for pricing models that can respond to market conditions dynamically. Their research reveals the limitations of static pricing models in capturing the intricacies of the aviation industryCleophas et al. (2009). Hansen et al. (2017) contribute to this approach by analyzing the impact of external factors such as fuel prices and economic conditions on pricing. However, existing research shows a gap in real-time, dynamic optimization of flight costs, a vital area for meeting the industry's rising demands Liu et al. (2017).

2.3 Limitations in Current Quantum Applications:

While quantum computing has enormous potential, it is not without difficulties. Preskill (2018) analyzes the topic of quantum supremacy Markov et al. (2018), which refers to the theoretical moment at which quantum computers outperform traditional computers. Obtaining quantum supremacy, on the other hand, does not imply effectively solving real-world issues. Fitzsimons et al. (2017) explored quantum computing system defects and decoherence. These concerns put into doubt the efficacy and stability of quantum solutions, particularly when applied to complicated optimization problems like ticket pricing Zhao et al. (2019).

Finally, the current study stresses the significance of a paradigm shift in ticket pricing optimization. While basic, traditional models and data-driven methodologies have limitations in adapting to the dynamic and competitive nature of the aviation industry. With its theoretical skill in optimization issues, quantum computing is a feasible choice . However, due to the challenges of current quantum computing implementations, as well as the absence of practical applications in airline pricing, a focused and comprehensive examination is required. This research aims to close these gaps by developing a novel ticket pricing optimization model based on quantum computing that capitalizes on the promise of quantum algorithms while taking into consideration and lowering their present constraints. The study seeks to contribute to the advancement of optimization strategies in the aviation industry.

3 Methodology

This section outlines the suggested technique, focusing on the data and methods used. Datasets, feature descriptions, and visualization material are offered to demonstrate the amount of competitiveness and globalization affinity in airplane tickets between the locations of major airline firms.

Furthermore, the ML, DL, and QML models that are used are presented in this part, along with a brief summary of each one to highlight the variations in performance and capabilities between them in figure 4 & 5.

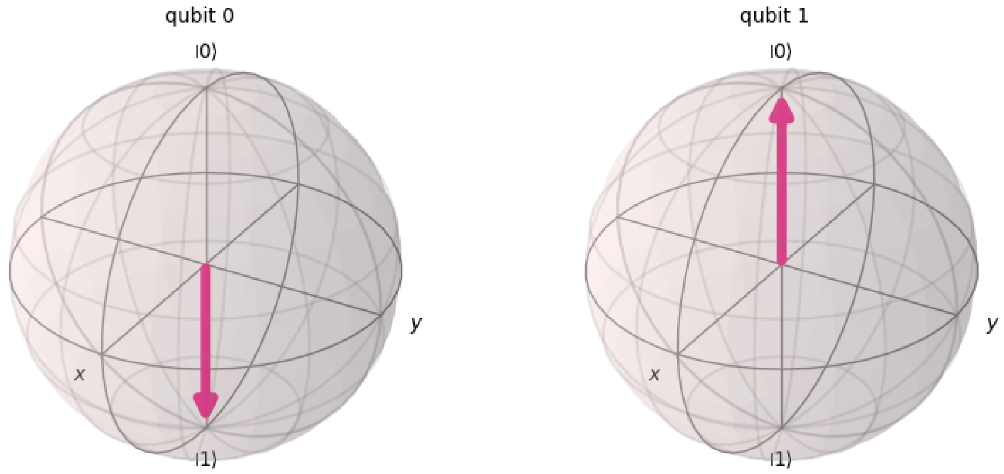

Figure 3: Image of Quantum Circuit of 2 Qubits

The research in comprehensive documentation and reporting, presenting detailed insights into the quantum algorithms, their integration with classical models, data preprocessing steps, simulation results, quantum computing implementation details, and the broader implications for the aviation industry. This organized technique enables a methodical study of the research subject, providing a comprehensive view on the potential of quantum computing to revolutionize flight pricing optimization.

In Figure 7 displays the steps of the proposed approach.

A machine learning model is used to forecast airline ticket prices based on variables such as distance, length, source and destination cities, departure time, class (economy/business), number of stops, and number of days till departure. Data preparation procedures such as resolving missing values, eliminating duplicates, and dealing with outliers are performed first. To investigate the relationships between various attributes and ticket price, exploratory data analysis (EDA) is employed. The application then estimates the cost of tickets using a number of machine learning models, such as Support Vector Machine, Random Forest, Decision Tree, and Linear Regression. Metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), R-squared, and Mean Absolute Percentage Error (MAPE) are used to evaluate the models.

The Random Forest Regressor is chosen as the main model, and its attributes are highlighted. The code also compares several regression models and performance measures, with the results shown in tabular form. The Support Vector Machine (SVM) model is then created and evaluated. Splitting data into training and testing sets, fitting models, projecting prices, and assessing performance indicators are all part of the assessment process. The paragraph provides a thorough overview of the whole method, from data cleaning and investigation to model training and assessment.

3.1 Dataset Description

The goal of this study is to forecast flight fare for six different airlines in six different regions. Flight data is gathered during a one-year period.

In Table 6.1.2 the amounts of data flights for each destination and for each airline

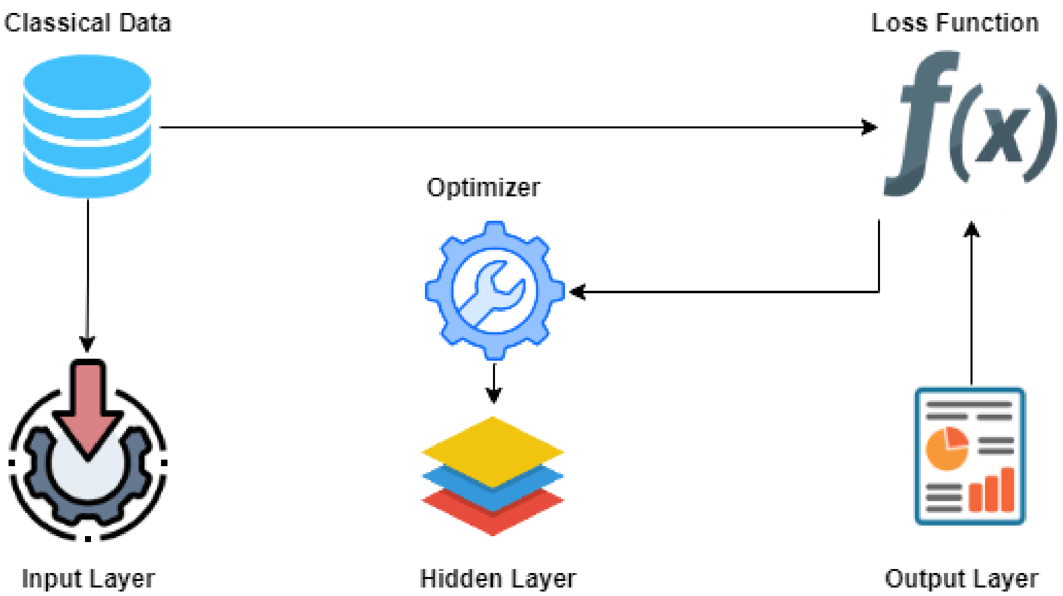

Figure 4: Classical Machine Learning Workflow

company.

Table 1: The total number of data flights for each destination and airline company

Airline	**Bangalore**	**Chennai**	**Delhi**	**Hyderabad**	**Kolkata**	**Mumbai**
Air India	11959	11141	14550	12022	13043	18177
Air Asia	3437	1516	4433	1560	2789	2363
Go First	4386	1488	5509	2576	3794	5420
Indigo	6772	6449	8133	6174	8473	7182
SpiceJet	1088	1172	2541	383	2054	1773
Vistara	23426	18602	22194	20038	19417	24182

Following the collection of the dataset, data cleaning techniques were used to manage missing values and delete duplicates. Outliers in certain traits were found and handled. Exploratory Data Analysis (EDA) was used to get insights into data distribution and correlations between different variables. Analysis characteristics such as flight length, ticket costs, source cities, and departure times were examined using visualizations such as box plots, bar plots, line graphs, and scatter plots.

Several feature engineering approaches were employed in the code, such as rounding down the'stops' column and applying Quantile Transformation on the'flight' feature. Correlation analysis was used to discover characteristics that were substantially connected and then deleted. Regression models, namely the Random Forest Regressor, were then used to forecast ticket prices. The Random Forest Regressor was trained and graded using a thorough assessment of feature relevance.

Multiple regression models were also built and evaluated using measures like Mean Absolute Error (MAE), R-squared, and others. The results were tabulated and ranked based on the Adjusted R-squared values. Visualizations were created to demonstrate the relationship between actual and predicted pricing, distance, number of days till departure,

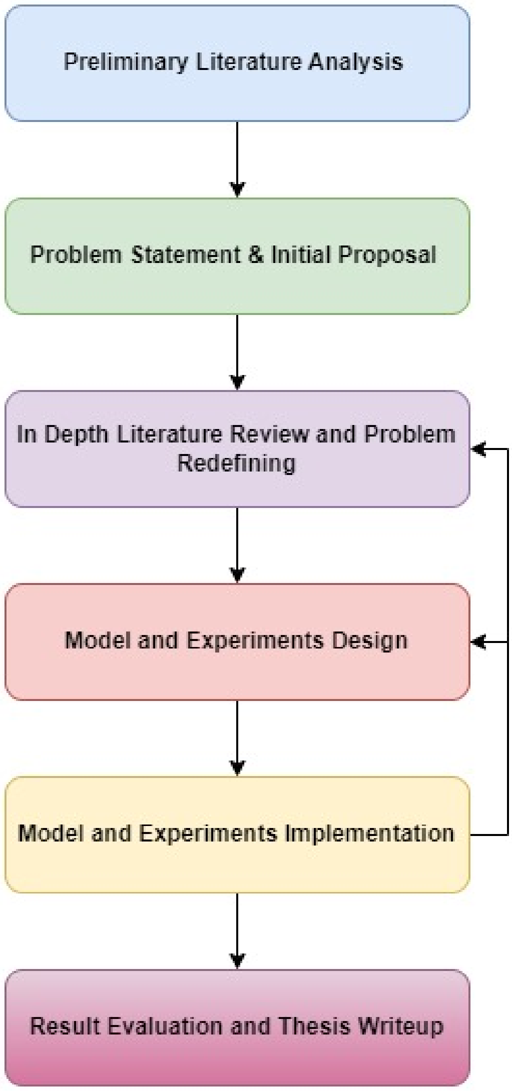

Figure 5: Research Process Diagram

and other relevant data.

Starting with the ML domain, eleven cutting-edge models are shown in Table 3.

Detailed study and implementation of multiple machine learning models for regression tasks on airline ticket price data.

First Loaded the required libraries, read the dataset, then conducted some basic data exploration and cleaning using libraries such as datacleaner and fasteda. Boxplots are used to deal with missing numbers, detect duplicated data, and highlight outliers.

In EDA used some Visualizations such as bar charts, boxplots, scatter plots, and line graphs are used to analyze relationships between different factors and the aim variable (ticket price). Examine the variations in ticket prices between airlines, departure times, and the number of stops. The EDA section provides information on the dataset and assists in the understanding of patterns and trends.

Machine Learning Model Construction is used to create machine learning models for

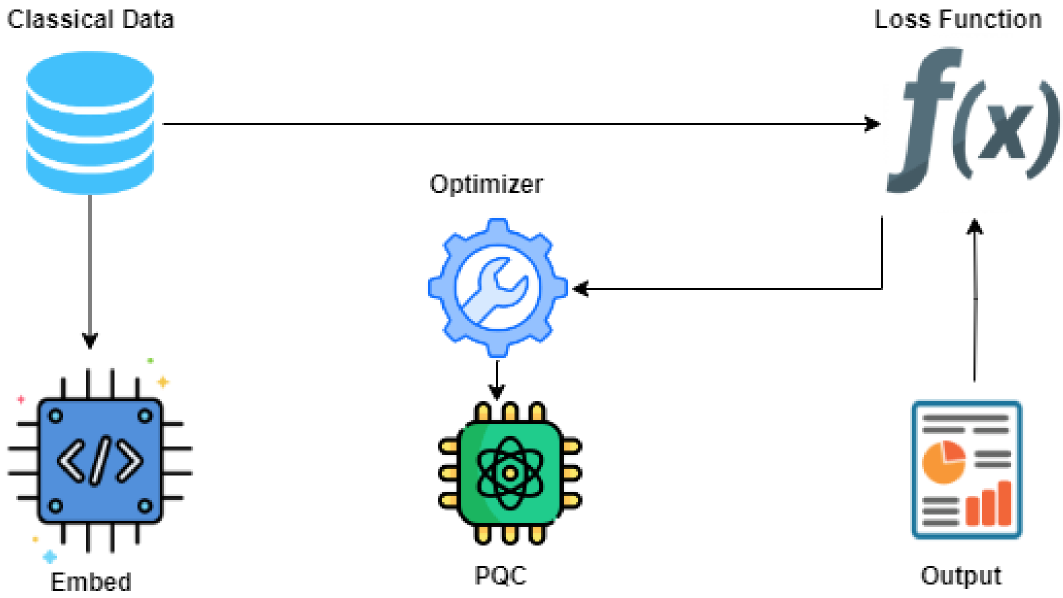

Figure 6: Quantum Machine Learning Workflow

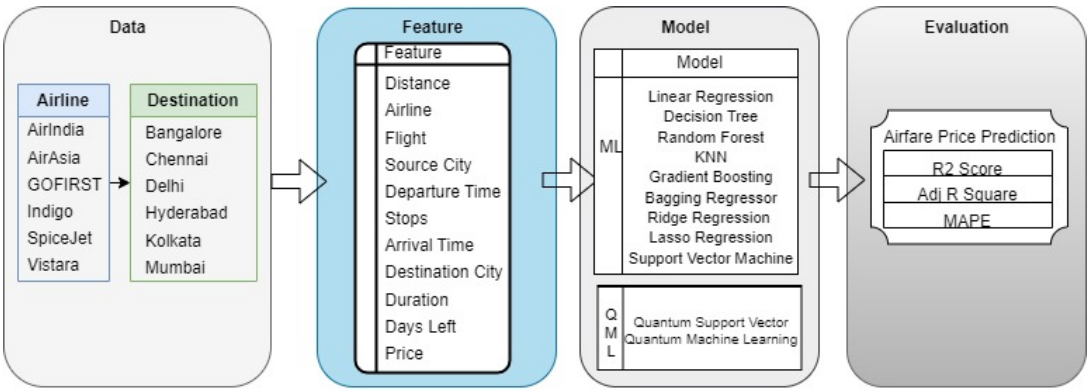

Figure 7: The proposed approach to airfare price prediction

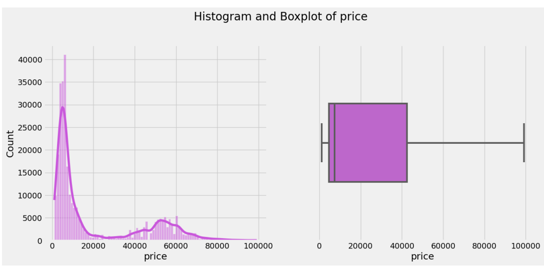

Figure 8: Summary of Price in National Currency

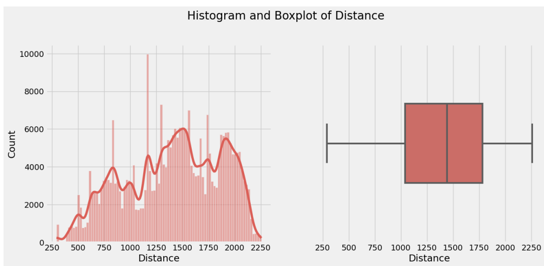

Figure 9: Summary of Distance in

Table 2: Selected machine learning (ML) models.

Model Name	Algorithm Type
Linear Regression	Linear Model
Decision Tree Regressor	Tree Based
Random Forest Regressor	Tree Based
KNeighbors Regressor	Instance-Based
Extra Trees Regressor	Tree Based
Gradient Boosting Regressor	Boosting Ensemble of Trees
XGB Regressor	Tree Based
Bagging Regressor	Ensemble Bootstrap Aggregating
Ridge Regressor	Linear Model with L2 regularization
Lasso Regressor	Linear Model with L1 regularization
SVM	Kernal Function

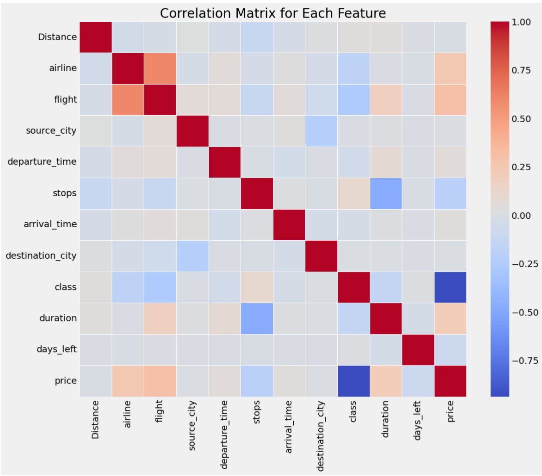

Figure 10: Confusion Matrix

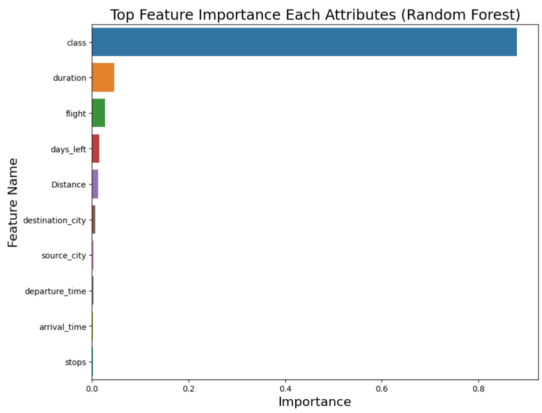

Figure 11: Important Feature in Data set

regression problems. In table 2, in employ well-known methods. Performance measures such as Mean Absolute Error (MAE), R-squared, and others are evaluated for each model. Examine feature significance for certain models using approaches like permutation significance and present the top qualities that contribute to the predictions. Model Comparison compares the performance of several models using a data frame that holds metrics such as MAE, R-squared, and others for each model.

Model Comparison compares the performance of several models using a data frame that holds metrics such as MAE, R-squared, and others for each model.

Table 3: Selected Quantum machine learning (QML) models.

Model Name	Algorithm Type
Quantum Support Vector Regressor	Linear Model
Variational Quantum Regressor (VQR)	Linear Model

A comprehensive quantum machine learning (QML) and conventional machine learning (ML) strategy is constructed using the Qiskit package. Installing and importing the essential libraries, registering an IBM Quantum Experience account, and constructing a rudimentary quantum circuit with Qiskit are the first stages. The code then imports the appropriate libraries, such as those for data processing and visualization.

The data is loaded from a CSV file including airline-related information, and preparatory processes include utilizing the DataCleaner package for automatic data cleaning. The dataset is then divided into features (X) and the target variable (y), followed by training and testing sets.

The definition of a quantum feature map and ansatz are used to introduce quantum machine learning. The techniques ZZFeatureMap and TwoLocal are employed. Quantum Neural Networks (QNN) and Variational Quantum Regressors (VQR) are used to perform classification and regression tasks, respectively. The training approach includes visualization, with live graphs illustrating the optimization progress.

The conventional machine learning counterpart is to scale the features with MinMaxScaler and train a Support Vector Classifier (SVC) on the dataset. In addition, the Qiskit machine learning module's Variational Quantum Classifier (VQC) is employed for comparison.

4 Design Specification

This thesis focuses on lowering flying costs by Machine Learning (ML), Deep Learning (DL), and Quantum Machine Learning (QML) models. The ML and DL components employ historical data to forecast flight costs, while the QML models apply quantum approaches for enhanced optimization.

ML algorithms analyze historical data, DL models find subtle trends, and QML models optimize using quantum techniques.

The suggested system uses machine learning approaches to optimize airline pricing in order to improve revenue management and customer satisfaction. To properly anticipate ticket prices, the algorithm takes a lot of parameters into account, including distance, length, days till departure, number of stops, and class. To assure the quality of the incoming data, data preparation activities such as missing value management, outlier identification, and feature engineering are carried out. The machine learning model ensemble includes several regression techniques, including Random Forest, Gradient Boosting, XGBoost, and Support Vector Machines. Metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared are utilized to adequately assess these models. The importance of features aids in understanding the primary elements driving cost.

The system studies the impact of variables such as distance, days till departure, and the number of stops on ticket cost using intelligent visualizations. Furthermore, employing Quantile Transformation increases the model's robustness by normalizing the flight data. The proposed system provides a comprehensive understanding of pricing dynamics, allowing airlines to make informed decisions to optimize revenue while improving customer experience. The design promotes transparency, interpretability, and flexibility, allowing it to respond to changing market conditions. By incorporating this machine learning technology into existing airline pricing systems, airlines may achieve a more data-driven, responsive, and efficient pricing strategy, ultimately leading to increased competitiveness and financial performance in the highly volatile aviation industry.

The proposed strategy for enhancing airline pricing comprises using quantum machine learning techniques, specifically Qiskit, to increase the accuracy and efficiency of pricing

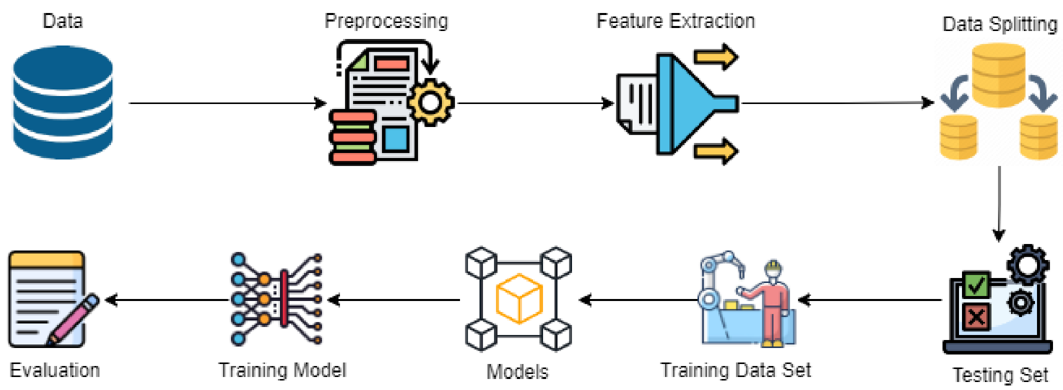

Figure 12: Proposed workflow of Machine Learning

models. Initially, typical machine learning models are employed for comparison and benchmarking. Quantum algorithms such as the Variational Quantum Regressor (VQR) and Quantum Neural Network (QNN) are integrated into the pricing model to take use of the computational improvements provided by quantum computing. The Qiskit program is used to create quantum circuits for regression and classification tasks. To encode classical data into quantum states, the quantum feature map, ZZFeatureMap, and ansatz, RealAmplitudes, are utilized. The VQR and QNN models are trained using historical airline pricing data, which includes flight details, demand patterns, and market trends.

The method employs a hybrid approach, combining conventional and quantum machine learning, to capitalize on the promise of quantum computers in evaluating complex information and enhancing pricing approaches. Live plots show the convergence of the objective function during optimization during the training period. Traditional Support Vector Machines (SVM) are also employed for baseline comparison. Because the system is responsive to changing datasets, it provides a scalable and cost-effective approach of improving airline pricing methods. This quantum-enhanced machine learning framework can create more accurate projections, improve pricing structures, and contribute to the growth of revenue management systems in the aviation sector. The application of quantum computing in airline pricing is consistent with the industry's pursuit of fresh technologies to address difficult optimization challenges.

5 Implementation

To solve a complex problem involving airline ticket pricing prediction, the proposed solution incorporates a comprehensive implementation of both traditional machine learning models and quantum machine learning methods. Traditional regression models such as Linear Regression, Decision Tree, Random Forest, and Support Vector Machine (SVM) are trained as part of the machine learning process. Each model's performance is evaluated using metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Adjusted R-Square. The Random Forest model was chosen because of its strong performance.

Following the classical models, the Qiskit program provides a quantum approach. Quantum computing elements like quantum feature maps and ansatz are linked to construct a Quantum Neural Network (QNN) for classification and regression problems.

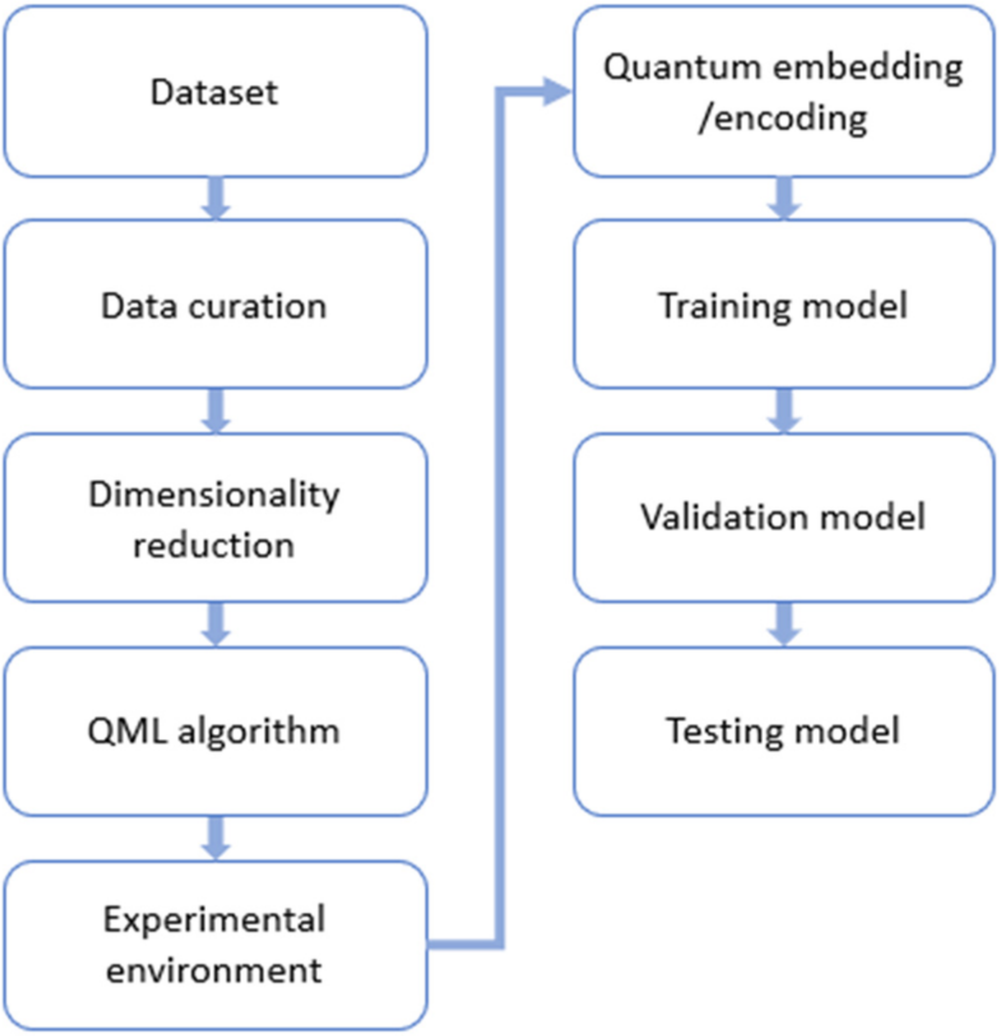

Figure 13: Proposed workflow of Quantum Machine Learning

Visualizations of quantum circuits and Bloch vectors help us better grasp the quantum components.

Furthermore, the suggested method considers the use of a Variational Quantum Regressor (VQR) for regression problems. The performance of conventional and quantum machine learning models is compared. Furthermore, classical Support Vector Classification (SVC) is taught, and the accuracy of SVC is compared to quantum techniques.

The method utilizes a hybrid system that computes airline ticket pricing by combining conventional and quantum machine learning techniques. Classical models serve as a starting point, while quantum models offer a fresh perspective on problem solving in the context of quantum computing. The detailed study and comparisons are meant to show the benefits and drawbacks of each strategy in dealing with the given circumstance.

6 Evaluation

Flight prices and the number of days till departure are closely linked, with supply and demand, seasonal fluctuations, and airline pricing strategies all having a role. Booking early typically results in lower rates, since prices tend to rise closer to the departure date due to increased demand and decreased seat availability. There are certain exceptions, such as airlines occasionally offering last-minute deals to replace unsold seats. Travel flexibility, promotion knowledge, and choosing less popular travel seasons may all impact costs. Because of the airline industry's dynamic nature, which is influenced by a variety of factors, it is vital for passengers to monitor and understand pricing trends in order to make the best booking decisions.

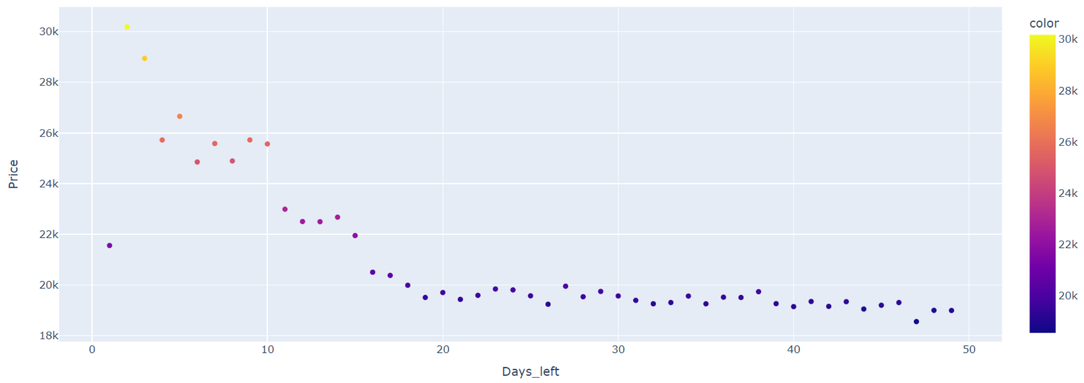

Figure 14: Flight price with days left to flight relation

The link between airline prices and travel distance is influenced by a number of factors. Longer journeys are frequently more expensive due to increasing fuel usage and operating costs. Airline competition, demand for certain routes, and seasonal fluctuations all have an impact. Direct flights to distant destinations may be more expensive than connecting flights. Airlines also utilize pricing mechanisms that are determined by market demand, which influences ticket costs. Travelers can occasionally find more affordable options by checking prices across other airlines and being flexible with their vacation dates. Monitoring these dynamics allows people to make smart decisions when booking flights based on distance traveled.

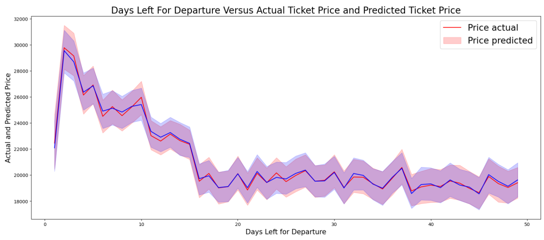

Figure 15: Days Left for Departure Actual and Predicted Price

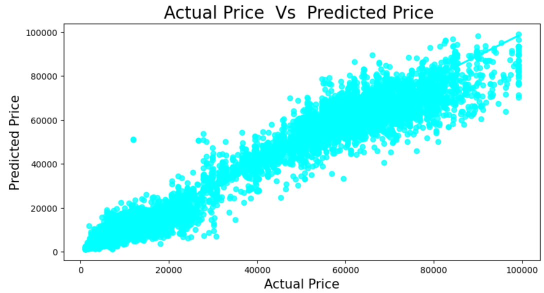

Figure 16: Actual Vs Predicted Price

6.1 First Experiment of Step 1: Airfare Price Prediction

6.1.1 Airfare Price Prediction: Days Until Flight Departure

The findings for both experimental techniques, for each ML and DL model, are reported in this section. Analysis and modeling are used to forecast airfare prices.The link between the number of days till departure and both actual and expected ticket prices is a dynamic interaction driven by a variety of factors.

Booking ahead of time is usually connected with reduced ticket rates, since airlines use dynamic pricing algorithms that take into account factors like as demand, seat availability, and previous data. Ticket prices are frequently predicted using algorithms that examine these data. Actual costs may change as the departure date approaches due to real-time market conditions. Travelers should book early for consistency, but tracking projected patterns and last-minute discounts might give cost savings chances.

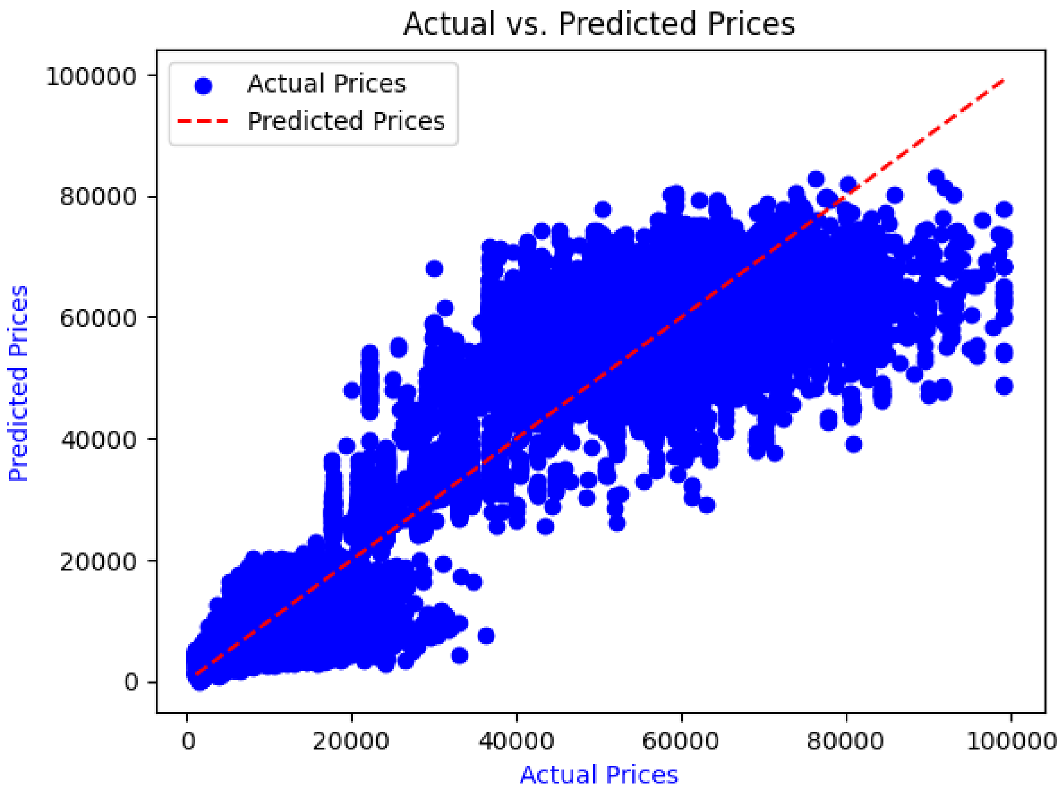

Figure 17: Deep Learning Actual Vs Predict

The model works well on the test data, as seen by low Mean Absolute Error (MAE) and Mean Squared Error (MSE) values (3776.63 and 37,499,379.85, respectively). A high R2 (0.92) value implies that the model can explain a considerable proportion of the variance in the dependent variable. The improved R2 score (0.93) shows the model's robustness by penalizing irrelevant predictors. The RMSE of 6123.67 provides an interpretable estimate of the average forecast deviation. Overall, the model is practical and generalizable, while future improvement should consider application-specific nuances.

Model Name	Adj R Square	MAE	RMSE	MAPE	MSE	RMSLE	R2 score
K Neighbors Regressor	0.988	981.23	2434.61	6.68	5927324.56	7.80	0.99
Gradient Boosting Regressor	0.988	974.26	2448.12	6.73	5993313.99	7.80	0.99
Ridge Regression	0.988	1014.79	2534.62	6.95	6424313.89	7.84	0.99
Bagging Regressor	0.983	1665.49	2956.09	13.22	8738450.36	7.99	0.98
Random Forest Regressor	0.980	1047.78	3186.04	7.21	10150881.54	8.07	0.98
XGB Regressor	0.958	2805.97	4659.41	20.52	21710146.74	8.45	0.96
Linear Regression	0.906	4629.98	6949.36	43.80	48293579.16	8.85	0.91
Decision Tree Regressor	0.906	4630.01	6949.36	43.80	48293586.82	8.85	0.91
Lasso Regression	0.906	4630.05	6949.36	43.80	48293578.96	8.85	0.91
Extra Trees Regressor	0.122	16188.54	21244.42	158.62	451325186.92	9.96	0.12

These models perform well over a range of metrics, implying that they can anticipate the objective variable. The ensemble nature of RandomForestRegressor, ExtraTreesRegressor, XGBRegressor, and BaggingRegressor increases their robustness and ability to deal with a wide range of data patterns. The specific choice is determined on the dataset and the trade-offs you are willing to make between interpretability and prediction capabilities.

BaggingRegressor, which is based on Bootstrap Aggregating, reduces variance and prevents overfitting. Effective in creating an ensemble that captures multiple data characteristics. Extra Trees, like RandomForest, is an ensemble method for constructing and combining several decision trees. Overfitting and variance reduction. Effective for gathering complex data patterns.

6.1.2 Airfare Price Prediction: Days Until Flight Departure using Quantum Machine Learning

Qiskit is utilized to create a Quantum Variational Regressor (VQR) for predictive modeling. It captures complicated links in the data set utilizing quantum circuits and variational approaches. In contrast to classical computers, which utilize classical bits to manage data, quantum technology uses quantum bits (qubits) and superposition to represent several possibilities at once. The VQR model employs quantum parallelism during training, which may be useful when dealing with big datasets and evaluating complicated feature interactions, highlighting the difference between quantum and conventional methods to machine learning challenges.

Model Name	Score
Quantum Support Vector Regressor	0.98
Gradient Boosting Regressor	0.86

To evaluate the accuracy of the Quantum Machine Learning model (VQC) on the dataset. The graphic depicts the training of a Variational Quantum Regressor (VQR) with a quantum neural network (QNN). The objective function value against iteration refers to the value of the objective function (also known as the cost function or loss

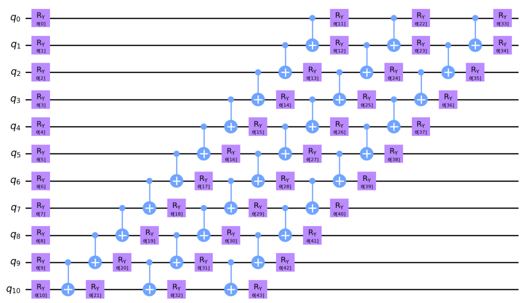

Figure 18: Quantum Circuit Ansatz for Variational Quantum Machine Learning

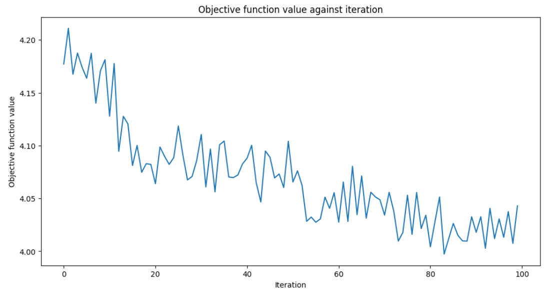

Figure 19: Objective function value against iteration

function) at each optimization iteration. The objective function is a measure of how well the model performs, and the optimization technique aims to reduce it. A lower number suggests greater performance, and 3.0 indicates better model optimization and Training time 225 seconds only.

6.2 Flight Distance Relation with Price

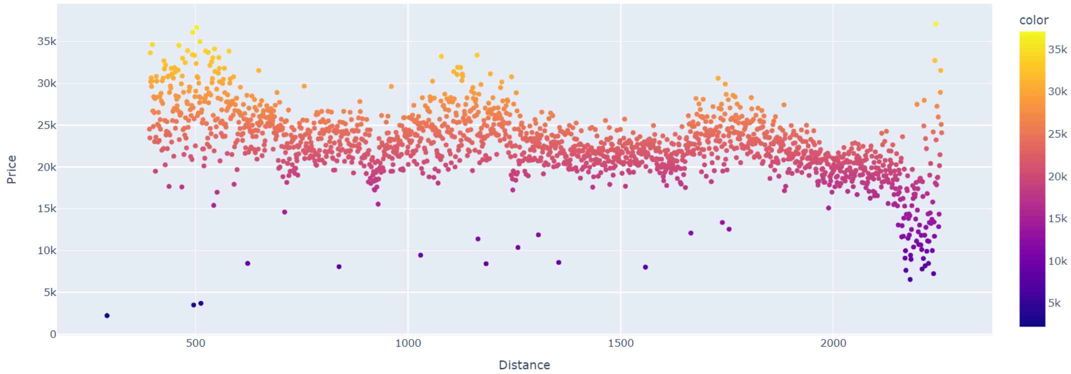

Figure 20: Experiment /Distance and Flight Price relation

Distance is recognized as a major feature in flight price prediction models because to its association with operating expenditures, fuel consumption, and flight time. Here's how distance impacts flight pricing, as well as a discussion of how stops effect flight pricing:

1. **Direct relationship:** In many cases, there is a direct relationship between flight distance and ticket price **??**. Longer flights frequently need more fuel, maintenance, and crew resources, which can raise the cost. As a result, longer flights are more expensive.

   After applying the optimization method to the dataset, the findings show a clear association between flight distance, ticket price, and number of stops Figure 16 & 17. The algorithm's optimal weights give insight into the relative priority ascribed to different aspects in the decision-making process.

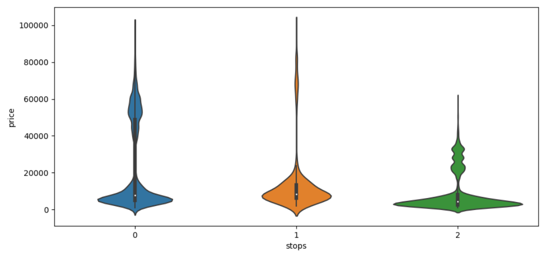

Figure 21: Stops More then 2 between Source and Destination City

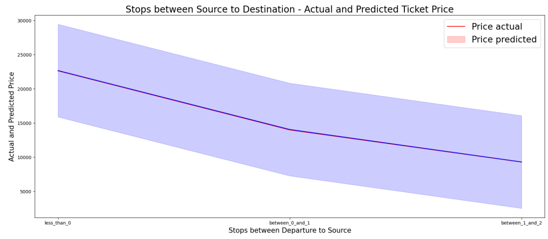

Figure 22: Stops between Source to Destination - Actual and Predicted Ticket Price

6.3 Experiment / Find the Optimal route

1. **Direct Flights: Minimizing Distance at a Premium Price** : Direct flights are the most efficient in terms of trip distance. These planes take the shortest route between two points, ensuring a fast and direct connection. This convenience, however, usually comes at a significant cost. The direct flight model is defined by higher ticket rates, which reflect the efficiency and time saved by avoiding layovers and breaks.

   Direct flights are appealing because of their time-saving potential, which caters to clients on tight schedules or searching for a seamless travel experience. However, as we debate the future of air travel, we must assess whether the cost of direct flights is justified in comparison to other routes that may require one or more stops.

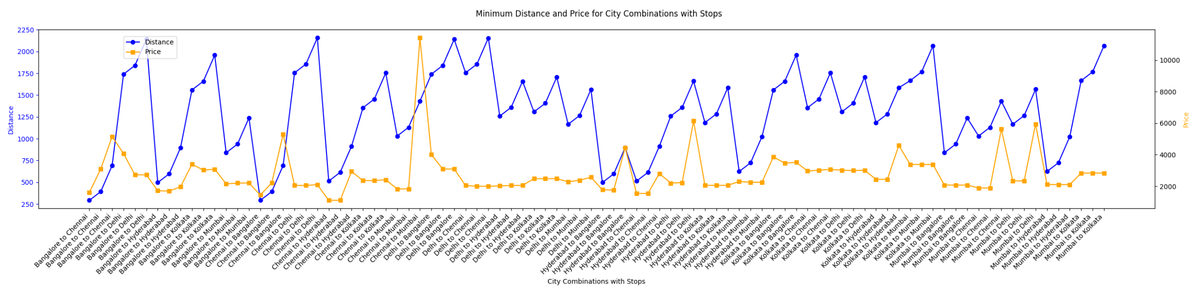

Figure 23: Experiment /Distance and Flight Price relation

2. **Multi-Stop Routes: Balancing Distance, Duration, and Cost** : When we change our attention to multi-stop routes, we discover a complicated interplay of distance, length, and cost. While certain routes may be longer owing to layovers, they are occasionally a more cost-effective alternative for consumers on a tight budget. With the evolution of air travel, the number of intelligent travelers who favor cost savings over direct flights has increased.

   The detailed link between distance and length is one distinguishing feature of multi-stop routes. While the overall distance traveled may be larger, the length of the journey is not substantially controlled by the extra stops. Technology advancements and improved airline operations have considerably lowered stopover periods, leading in a shorter overall trip duration.

3. **The Optimization Dilemma Choosing the Best Path** :Our search for the ideal path reveals that there is no one-size-fits-all solution. The choice between

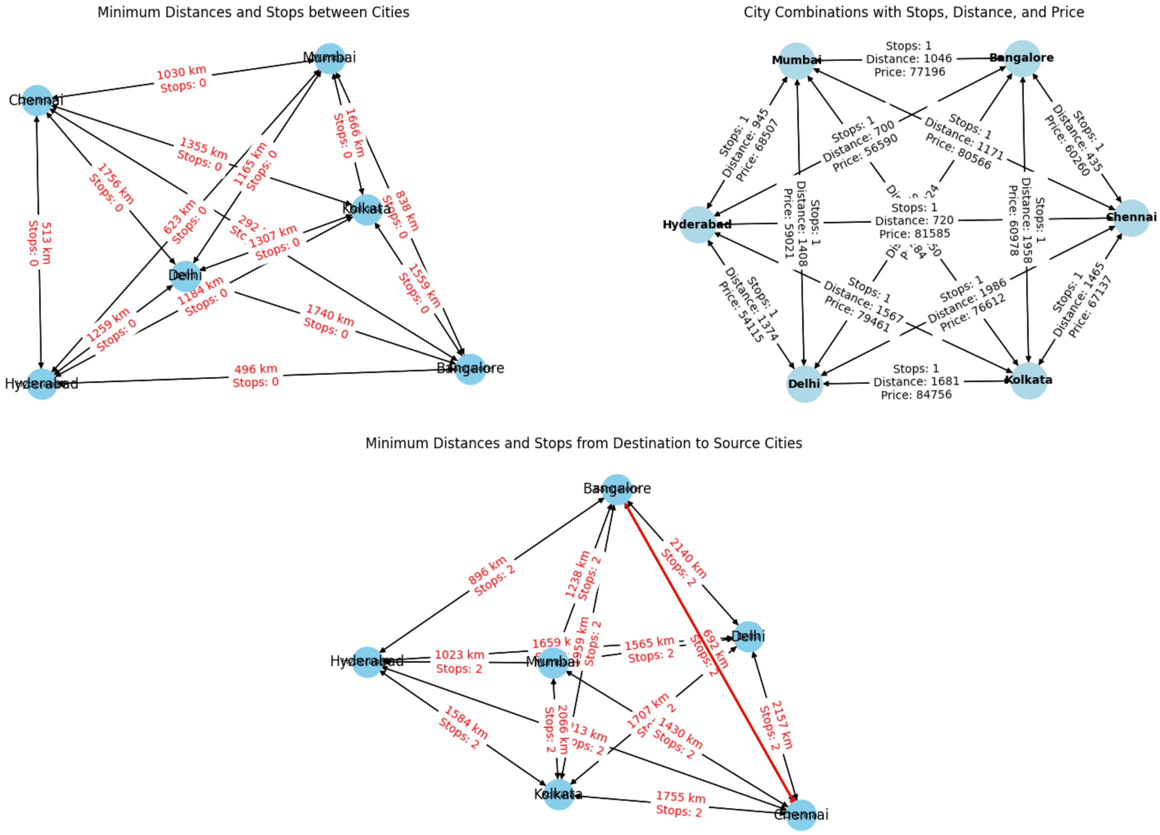

Figure 24: Showing the Stop , Price and Distance

direct and multi-stop flights is determined by personal preferences, priorities, and, most importantly, budgetary constraints. The increased popularity of air travel has resulted in a diverse collection of passengers, each with their own set of expectations and wants.

To solve this optimization conundrum, new algorithms and data analytics must be applied, which include not just the direct distance between origin and destination, but also the economic issues associated with alternative routes. Machine learning algorithms can analyze historical price data, consumer preferences, and airline operational performance to recommend the most cost-effective and time-efficient routes depending on individual needs.

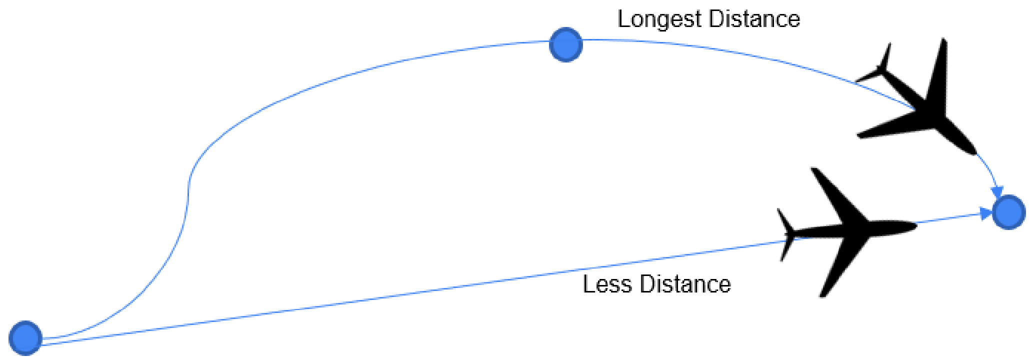

Figure 25: The Optimization Dilemma Choosing the Best Path

Finally, the evolution of flight distance in relation to price is an enthralling adventure typified by the never-ending search for the best route. Direct flights provide unrivaled efficiency at a premium price, but multi-stop routes are more cost-effective with intricate trade-offs between distance, length, and cost. As we navigate the ever-changing air travel scene, integrating cutting-edge technology and analytical tools becomes increasingly vital in our quest for the greatest combination of ease and price. The search for the best route is a continual endeavor, fueled by the aviation industry's dynamic character and the various demands of today's demanding travelers.

6.4 Discussion

The use of quantum machine learning (QML) into airline pricing holds enormous promise for transforming established procedures. However, it is critical to recognize that the existing constraints in quantum hardware have necessitated the usage of quantum simulators in this effort. Quantum simulators simulate quantum systems on classical computers, allowing researchers to investigate the possible uses of quantum algorithms before powerful quantum processors become available.

The quantum computing area is still in its early stages, and viable, large-scale quantum devices are not yet generally available. Nonetheless, current research in quantum computing strives to solve hardware obstacles and increase processing capability. The shift from quantum simulators to true quantum hardware is expected to result in additional major breakthroughs in the application.

Despite the computational limits, the employment of quantum simulators in this work represents a forward-thinking effort to leverage the capabilities of quantum mechanics for complex optimization issues. As quantum technology progresses, the advantages of QML, such as greater accuracy and efficiency in flight cost prediction and optimization, are projected to become increasingly apparent, resulting in a shift of airline pricing strategies and operations. The predicted benefits include increased pricing accuracy, efficiency, and agility, which would eventually lead to more revenue, higher customer satisfaction, and increased competitiveness in the aviation sector.

7 Conclusion and Future Work

Finally, this paper investigated the concept of determining the optimum air travel routes between a variety of source and destination locations based on variables such as distance, cost, and time. The research included a thorough evaluation of the data, as well as several visualizations and statistical insights to better understand the links between various aspects. The study looked at how the number of stops, days till departure, and departure time impacted the cost of airfare. Furthermore, machine learning approaches such as the Random Forest Regressor and different regression models were used to estimate ticket prices based on the data provided.

The major goal of this study was to identify the best air travel routes between source and destination cities, taking into account parameters such as distance, cost, and length. The study's central question was, "What are the key determinants of optimal air travel routes, and how can machine learning models contribute to predicting ticket prices based on these determinants?"

The research's effectiveness is based on its capacity to deliver practical insights on air travel pricing trends. However, there are several drawbacks, such as the absence of some dynamic price components and the dependence on historical data, which may not represent real-time market swings.

Integration in Real Time Create real-time data integration tools to increase model flexibility. Mechanisms of Dynamic Pricing Investigate dynamic pricing methods to gain a better understanding of real-time price movements. Extend the inquiry to a global scale, using a larger dataset for a generalized model.Design of the User Interface Create user interfaces for personalized pricing projections and improved accessibility .The commercialization opportunity is in developing a user-friendly platform or application that includes numerous prediction algorithms, providing valuable data for both individual travelers and the airline industry. Collaborations with travel firms or internet platforms may potentially make it easier to implement these predictive algorithms in real-world scenarios.

References

Balaiyan, K., Amit, R., Malik, A. K., Luo, X. and Agarwal, A. (2019). Joint forecasting for airline pricing and revenue management, *Journal of Revenue and Pricing Management* **18**: 465–482.

Choudhary, A., Jagadeesh, R., Girija, E., Madhuri, M. and Shravani, N. (2023). Fly-high:machine learning based airline fare prediction model, *2023 6th International Con-*

ference on Information Systems and Computer Networks (ISCON) pp. 1–8.
URL: *https://api.semanticscholar.org/CorpusID:258511102*

Cleophas, C., Frank, M. and Kliewer, N. (2009). Recent developments in demand forecasting for airline revenue management, *International Journal of Revenue Management* **3**(3): 252–269.

E.A., K. and S.K., R. (2020). Analysis of artificial neural networks training models for airfare price prediction, *Artificial Intelligence* **25**: 45–50.
URL: *https://api.semanticscholar.org/CorpusID:234635290*

Joshi, N., Singh, G., Kumar, S., Jain, R. and Nagrath, P. (2019). Airline prices analysis and prediction using decision tree regressor.
URL: *https://api.semanticscholar.org/CorpusID:219877736*

Kalampokas, T., Tziridis, K., Kalampokas, N., Nikolaou, A., Vrochidou, E. and Papakostas, G. A. (2023). A holistic approach on airfare price prediction using machine learning techniques, *IEEE Access* .

Kuptsova, E. and Ramazanov, S. (2020). Analysis of artificial neural networks training models for airfare price prediction.

Lantseva, A. A., Mukhina, K. D., Nikishova, A., Ivanov, S. V. and Knyazkov, K. V. (2015). Data-driven modeling of airlines pricing, *Procedia Computer Science* **66**: 267–276.
URL: *https://api.semanticscholar.org/CorpusID:62479874*

Liu, T., Cao, J., Tan, Y. and Xiao, Q. (2017). Acer: An adaptive context-aware ensemble regression model for airfare price prediction, *2017 International Conference on Progress in Informatics and Computing (PIC)*, IEEE, pp. 312–317.

Malkawi, M. and Alhajj, R. (2023). Real-time web-based international flight tickets recommendation system via apache spark, *2023 IEEE 24th International Conference on Information Reuse and Integration for Data Science (IRI)* pp. 279–282.
URL: *https://api.semanticscholar.org/CorpusID:261403814*

Markov, I. L., Fatima, A., Isakov, S. V. and Boixo, S. (2018). Quantum supremacy is both closer and farther than it appears, *arXiv preprint arXiv:1807.10749* .

Panigrahi, A., Sharma, R., Chakravarty, S., Paikaray, D. B. and Bhoyar, H. (2022). Flight price prediction using machine learning, *ACI@ISIC*.
URL: *https://api.semanticscholar.org/CorpusID:254998923*

Ratnakanth, G. (2022). Prediction of flight fare using deep learning techniques, *2022 International Conference on Computing, Communication and Power Technology (IC3P)* pp. 308–313.
URL: *https://api.semanticscholar.org/CorpusID:249685726*

Tziridis, K., Kalampokas, T., Papakostas, G. A. and Diamantaras, K. I. (2017). Airfare prices prediction using machine learning techniques, *2017 25th European Signal Processing Conference (EUSIPCO)* pp. 1036–1039.
URL: *https://api.semanticscholar.org/CorpusID:3939841*

Vu, V. H., Minh, Q. T. and Phung, P. H. (2018). An airfare prediction model for developing markets, *2018 International Conference on Information Networking (ICOIN)* pp. 765–770.
URL: *https://api.semanticscholar.org/CorpusID:5078365*

Zhao, Z., Fitzsimons, J. K., Osborne, M. A., Roberts, S. J. and Fitzsimons, J. F. (2019). Quantum algorithms for training gaussian processes, *Physical Review A* **100**(1): 012304.