Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7628863

PERFORMANCE OPTIMIZATION USING QUANTUM MACHINE LEARNING TECHNIQUE FOR PREDICTING FLIGHT DELAYS

Shruti S Pophale
School of Computer Sciences and Engineering, Sandip University Nashik, India
shruti.pophale@gmail.com

Purushottam R. Patil
School of Computer Sciences and Engineering, Sandip University Nashik, India

Amol D. Potgantwar
Department of Computer Engineering, Sandip Institute of Technology and Research Center Nashik, India

Pawan R. Bhaladhare
School of Computer Sciences and Engineering, Sandip University Nashik, India

Abstract:
Numerous companies rely on a number of different airlines to connect them with other regions of the world as a result of the increasingly important position that the aviation industry plays in today's global transportation sector. On the other hand, severe weather can have a direct impact on airline services, most notably in the form of flight delays. The solution to this problem is to accurately predict these flight delays, which enables passengers to be well prepared for the deterrent that will be caused to their journey and enables airlines to respond to the potential causes of the flight delays in advance in order to diminish the negative impact. The goal of this work, which makes use of quantum machine learning, is to investigate the methodologies that are applied in the construction of models for forecasting flight delays that are brought on by inclement weather. In the initial phase of the project, we investigate the feasibility of employing Python-based linear Regression in conjunction with Support Vector Machine. After that, we feed the dataset into our classifier to obtain the results. In the second half of the project, our primary focus is on acquiring data, and we investigate the possibility of combining quantum machine learning with the process of gathering data by first dissecting the dataset and then determining which features are important. After looking over the results, we compared them to the outcomes of other models, such as the machine learning classifier with quantum linear regression and the quantum SVM-SVR, in order to determine which classifier would be the most effective in resolving the issue.
Keywords: Airline services, delays, mean delay times, quality mean delays and performance matrices.

1. INTRODUCTION

The management of air travel in the United States has a significant obstacle in the form of flight delays. From 2010 to 2018, an average of approximately 20% of domestic flights inside the

United States were delayed. According to estimates provided by the Federal Aviation Administration (FAA), the yearly expenditures associated with delays in 2018 totaled $28 billion [1]. Flight delays will not only result in an increase in economic losses, but they will also have a detrimental influence on the operations of the air transport system as well as passengers, airlines, and the airline industry. To begin, flights that are delayed will have an effect on the on-time performance of airports and will cause passengers to experience inconvenience [2]. According to a research by the United States Department of Transportation [3,] one of the primary causes of customer dissatisfaction was long delays on their flights. Second, a high delay rate will cause damage to the market reputation of the airline and the loyalty of its passengers; it will also increase the cost of operations and incur additional costs [4]. Flight delays will throw early flight plans into chaos, which will enhance the risk to the safety of the air transportation system [5]. These kinds of adverse effects serve as a driving force behind the research and development of improved delay management systems [6].

It is common knowledge that high precision prediction models are useful tools that can cut down on flight delays, hence lowering the economic costs incurred by airlines and improving the level of passenger satisfaction [7]. The purpose of this study is to assist aviation authorities in gaining a deeper understanding of the possible behaviours and causes that contribute to flight delays, as well as to construct a prediction model with a high level of accuracy that can monitor aircraft delays in real-time. It is essential for the prediction model to have a high level of accuracy because the results of predictions with lower accuracy will mislead airport operators when it comes to understanding the future delay status of flights, which will, in turn, lead to ineffective alleviation of flight delays [8, 9].

The report from the Bureau of Transportation Statistics (BTS) states that an initial flight delay can be attributed to a variety of factors, including severe weather, a breach in security, air carrier, etc. [10]. However, the high number of connected resources in the National Airspace System (NAS), such as aeroplanes, crews, and passengers, results in the propagation of flight delays [11]. In order for airlines to get the most out of their fleets, they typically decrease the amount of time that elapses between arrival and departure flights, which also makes the likelihood of delay propagation higher [12].

There are two significant obstacles that need to be taken into consideration from the point of view of data analytics. First, the majority of the parameters that were included in the earlier research for predicting flight delays were on a macro-level. These factors include meteorological conditions, air traffic management, temporal variables, and seasonal effects [13–18]. However, factors at the macro-level characterise the general trend of flight delays, which may or may not be directly connected with delays in actual flight times [19]. Second, because of the interrelationships between airports, it is possible to divide flight delays that occur in a single airport into "local" and "network" categories [20]. The term "local" refers to the associated delays that are experienced in turnaround operations at each unique airport. The term "Network" is used to measure the delays that have been noticed at particular airports, however these delays cannot be attributed to local operations because they occurred earlier in the day at upstream airports. However, only a small number of works have been dedicated to the assessment and analysis of the cause behind flight delays, and more especially to the local and network effects on delays for leaving flights. This is a major limitation of the field. It is essential to do research that can quantify and analyse the local and network effects on delays [20]. On the one hand, it is anticipated that network effects in the NAS will be effective indications for delay prediction [17]. Consequently, there is a possibility that the accuracy of the predictions will improve if the network effects can be captured and incorporated into a model. However, only a small number of studies looked at the influence that the air traffic network has on flight delays [17], and the study that was important looked at the network impacts as different groups based on a clustering algorithm. On the other hand, the modern aviation information system is composed of a sizable number of different components. When all of the potential elements are included in a model, the complexity of the model rises, and it also has a negative impact on the results of the prediction, primarily as a result of factors that are irrelevant or collinearity [19]. Therefore, in order to make a prediction with a high level of accuracy, it is essential to investigate the reason behind the formation of delays and isolate the components that are directly relevant to flight delays.

- The number of people and goods that need to be transported on time has led to a considerable growth in the demand for air transportation. Flight delays have been identified as one of the most difficult challenges faced by the aviation industry.
- The ability to accurately estimate when a flight may be delayed is essential to the development of a more productive airline industry.
- This work provides aviation authorities with the ability to understand the underlying behaviour and mechanism of flight delays.

Additionally, it identifies a novel set of influential elements that can be used to construct a high accuracy prediction model to predict flight delays. The findings indicate that local effects are primarily responsible for short-term delays, whereas network effects are primarily responsible for long-term delays. Furthermore, the findings indicate that large airports are affected by a large number of upstream airports, whereas small airports are primarily affected by local operations. If it is used correctly, the proposed prediction model has the potential to assist operators of aviation systems in performing more efficient planning and budgeting, as well as reducing the adverse effects of undesirable congestion.

In situations that take place in real life, information is obtained from a variety of sources. At any given time, a dynamically updated version of the information is accessible. The following is a list of research objectives:

1.To address a variety of concerns for the purpose of developing a QML framework that enables effective decision making.

2.Using quantum machine learning to improve the overall performance of the flight delay model.

3.The gathering, processing, and synergistic combining of information obtained from a variety of information sources in order to achieve a deeper comprehension of the phenomenon that is the focus of the investigation.

4.Intelligence and the ability to regulate a system can be improved by using data gathered from a variety of sources.

5. Real-time system that makes optimal use of available information and carefully selects appropriate procedures

The remaining portions of the paper are structured as follows: In the following section, we will discuss the relevant research on flight delay prediction. In Section 3, the specific details regarding the data set as well as the definitions of aircraft delays are provided in further depth. In the fourth section, both the prediction model and the accompanying complex network technique are presented. The prediction performance is proposed and discussed in Section 5, which then presents the results. In the final section, we will analyze the findings and what comes next.

2. LITERATURE REVIEW

In this section, we will offer a literature analysis to show the uses of machine learning models as well as traditional techniques for assessing the probability of flight delays [15, 18, 21]. The characteristics of aircraft arrival and departure delays were investigated by Muelleret al. [21], who came to the conclusion that departure delays were better modeled by a Poisson distribution, whilst arrival delays were more accurately represented by a normal distribution. Pfeil et al. [15] established a probabilistic forecast of whether or not a terminal area route will be blocked based on raw convective weather forecasts. This forecast was used to determine whether or not a terminal area route will be closed. An asymmetric logit probability model was presented by Perez-Rodriguez et al. [22] to predict the daily probabilities of arrival delays based on a variety of influential factors. These factors include the departure delay, the size of the airline, the size of the airport, and the day of the flight. The model was able to make this prediction by taking into account all of these factors. Researchers started to predict flight delays by applying machine learning models [30, 31] as a result of the rapid development of data science. Neural networks, k-NN, and random forests are examples of models that are frequently utilized for the purpose of flight delay prediction. When compared to the use of probabilistic models in previously conducted research, the application of machine learning algorithms typically results in better prediction performance [19]. As shown in Table 1, macro-level characteristics are frequently utilized as inputs for delay prediction models. These include temporal variables [13, 14, 17,23], seasonal impacts [23, 24], airlines and airport [13, 14, 25,27]. An adaptive network model was developed by Khan Mohammadi et al. [13] to predict flight delays. The model was based on macro-level factors such as the origin of the departure flight, departure time, and scheduled arrival time. The predictions were used as inputs for a fuzzy decision-making system to schedule aircraft landings. Lu et al. [32] developed a recommendation system by generating a k-Nearest Neighbor algorithm to simulate the delays in historical records and predict similar delay situations in the future. These attributes included airports, airlines, departure time, and the number of passengers, among other things. Other macro-level parameters, such as the time of day, the day of the week, and the month of the year, are also utilized rather frequently [14, 28, 29]. To anticipate flight delays, Daniel et al., [26] constructed a neural network with 4 hidden layers and further incorporated the route type. They attained the highest accuracy of 87% of the aforementioned strategies by using this method. However, the climatological aspects are not taken into account in any of their models, thus there is still an opportunity for development in this area. Therefore, In their Bayesian Network model for arrival delay prediction, Rodrguez-Sanz et al. [27] took into account airport infrastructure, airline, type of aircraft, operational times, and weather factors (wind, clouds) in airports. A two-stage model for predicting flight delays was proposed by Young et al. [24], who incorporated recurrent neural networks and neural networks in their study. The initial step consisted of utilizing deep RNN to make a prediction about the daily delay status.

Table 01: Overview of relevant studies

Reference	Methodology	Considered factors
Belcastro et al. [25]	Random forest	The airports of origin and destination, the departure and arrival times that are scheduled, and the weather conditions at both the origin and the destination airports (temperature, humidity, wind direction, wind speed, barometric pressure, sky condition, visibility, weather phenomena descriptor)
Daniel et al. [26]	Neural Network	The airline, the route type code, the airport of departure, the airport of arrival, the scheduled day and time of departure, and the scheduled day and time of arrival
Rodríguez-Sanz et al. [27]	Bayesian Networks	Infrastructure of the airport, airline, kind of aircraft, operational hours, passenger throughput, and current weather conditions
Lambelho et al. [28]	Machine learning	the airline, the type of aircraft, the distance, the airport of origin and airport of destination, the number of seats, the time of day, the day of the week, the month of the year, and the later arrival delay.
Guan et al. [29]	Random forest and LSTM	The starting point and the finish line airport, flight number, airline identification number, departure and arrival planned time, day of the month, month, season, day of the week, weather condition, wind direction, wind speed, and traffic flow of air route are all included.

2.1 Research Motivation

The typical amount of time that an airplane is delayed is frequently used as a measurement of airport capacity. In this world, there is almost always some sort of difficulty with flights being delayed. It can be quite challenging to articulate the rationale behind a delay. Bad weather appears to be the most common reason of aircraft delays, in contrast to a few other variables responsible for the delays, such as runway work and high traffic. Some planes experience delays as a result of reactive delays, which are caused by the late arrival of the flight that came before them. Because airlines and airports rely on consumer loyalty to maintain their frequent flyer programs, this is detrimental to both industries. It also has an impact on the marketing methods used by businesses.

3. Methodology

In this section, we will attempt to depict an approach that we have utilized in the past with the goal of resolving issues about enormous amounts of data, complicated processing, inadequate computational space, overfitting, and existing noise in data. The development of the proposed model is illustrated by the following thing. As can be seen from the accompanying image, the proposed method is broken up into a number of stages.

- **Classical Machine Learning**
- **Quantum computing**
- **Quantum machine learning**
- **Datasets**

3.1 Classical Machine Learning

The following diagram illustrates the process of doing research using the proposed machine learning and QNN approach.

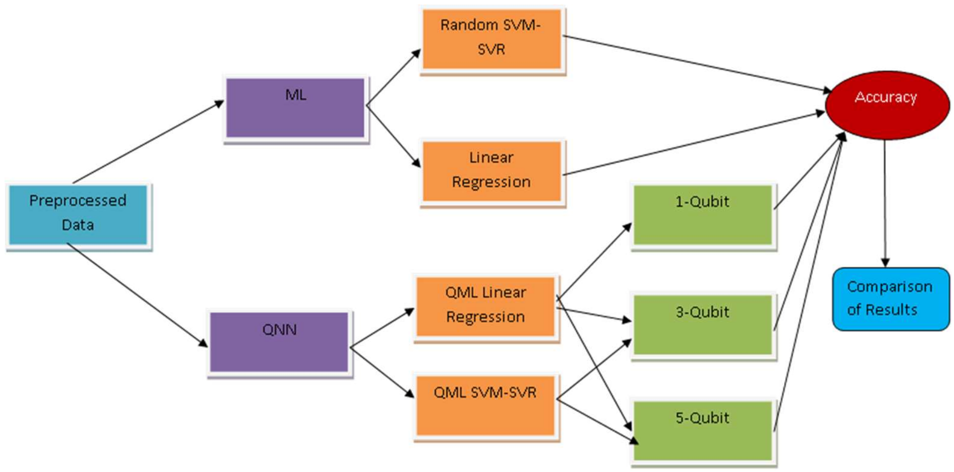

Figure 01: Flow of Research work

3.2 Load Dataset: In order to import the dataset from the local directory, we made use of one of the libraries in deep learning known as Pandas. Pandas provide an expressive, powerful, and array that makes it simple to manipulate the data on many different platforms. The dataset used in data science is in the format of a csv file. Pandas can import the dataset from the local directory if it is in the comma-separated values format and it uses CSV, which stands for comma-separated values.

Dataset:

- **US flight dataset: Jan 2015 - Dec 2015**
- **Dataset: flights.csv, airports.csv, airlines.csv**

3.3. Data Preprocessing: The most important purpose of this model is to provide accurate and precise predictions, and the algorithms should be able to quickly interpret the characteristics of the data. As a result of their varied points of origin, the vast majority of datasets derived from the actual world have data that is either missing, inconsistent, or noisy. It is impossible to get results from data mining algorithms applied to noisy data; therefore, it is essential to treat the data in order to increase the quality of the datasets. Four steps make up the preprocessing of the data, and they are as follows:

1. Data cleaning

2. Data integration
3. The Transformation of Data
4. the decrease of data

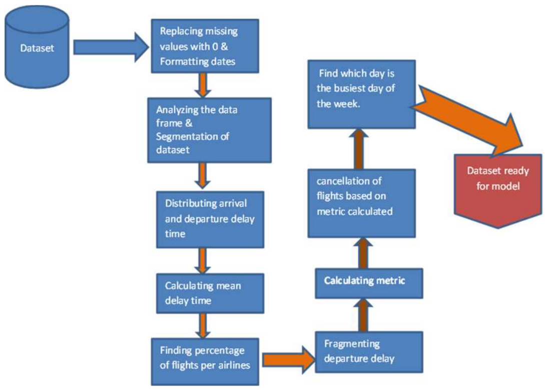

Figure 02: Data Preprocessing approach using ML

Feature Extraction: Feature extraction is a method or technique that is used to reduce the enormous amount of input data into a useful feature. It does this by focusing on the features that are most important. The massive amounts of data that are supplied are shrunk down into more manageable and understandable groups through the use of dimensionality reduction.

When training a machine learning model, feature extraction can be beneficial since it leads to the following outcomes:

•An increase in training speed
•An improvement in model accuracy
•A reduction in the risk of overfitting
•An improvement in the quality of data visualization

4. Proposed approach:

The term "quantum machine learning" refers to the practice of incorporating quantum algorithms into existing machine learning software. The most popular application of the phrase refers to machine learning algorithms for the analysis of classical data that are carried out on a quantum computer; this type of machine learning is often referred to as quantum-enhanced machine learning. Quantum machine learning is a form of machine learning that makes use of qubits, quantum operations, or other types of specialized quantum systems to improve the speed with which algorithms in a program can compute and store data. Machine learning algorithms are used to compute enormous amounts of data. This covers hybrid approaches that utilize both classical and quantum processing, such as outsourcing computationally complex subroutines to a quantum device. Hybrid methods encompass both of these types of processing. The steps that can be taken while working with the proposed strategy are as follows,

1. Dataset Loading
2. Parameter Initialization

X=dept_hr_min
Y= mean
Learning –rate α=0.01
Epochs=26

Batch size=30
Qubit=Qn
qnode =€=(X,Y,Θ(wt)-> Random no.)
Ψ=Data_Dimension
3 . for Each Qn
4. qlayer=(€, Θ,Qn)
5. clayer1=(Qn, activation="relu")
6. clayer2=(Ψ, activation="linear")
7. model=(q,c1,c2)
8. modelfit(X,Y,epochs,batchsize)
9. QMLalgo()
10.end for

The phase that was just completed provides an overview of the proposed method, which can make use of QML in order to enhance the performance of the system and cut down on the amount of delay in flight.

5. Result analysis & Discussion:

The performance of the prediction models is evaluated using domestic flights inside the United States from January 2015 through December 2015. Classification and QML are both produced using the same data set. Benchmark models employing existing factors from previous research are also developed in order to assess the effectiveness of our suggested factors using QML. The figure that follows demonstrates the two distinct kinds of delays, such as arrival and departure, on an airline-by-airline basis. It draws the conclusion that certain airlines do not experience delays upon arrival, but rather only experience delays upon departure.

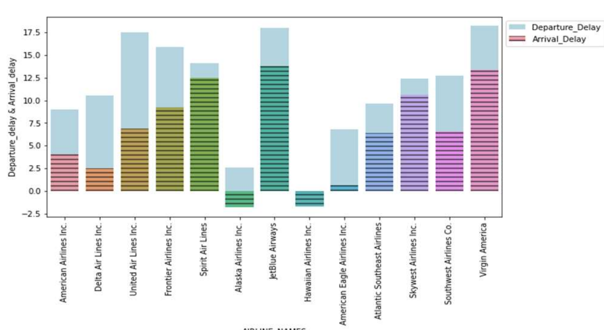

Figure 03: Distribution of Arrival delay and Departure delay per airline

5.1 Performance Measure:

Following are some performance matrix use for evaluating the proposed approach with the exiting one

1) Mean squared error: The Mean Squared Error is a statistical method that determines how closely a regression line corresponds to a given group of data points. It is a risk function with a value that corresponds to the expected squared error loss and its expected value. The calculation for mean square error involves finding the average, or

more specifically the mean, of the squared mistakes that result from analyzing data in relation to a function.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2 \quad (1)$$

2) Mean absolute error: Mean Absolute Error is a statistic that is used for the evaluation of models, specifically regression models. The mean absolute error of a model with regard to a test set is the mean of the absolute values of the individual prediction errors averaged out across all of the occurrences that make up the test set.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \quad (2)$$

3) Root mean square error : It does this by calculating the average deviation from the values predicted by a model and the values that actually occur. It provides a rough estimate of how accurately the model can anticipate the value that is being sought after (accuracy). The better the model is, as measured by the RMSE, the lower the value of the root mean square error.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (3)$$

4) R-squared (R^2): The use case determines whether or not a score is considered to be good; nevertheless, in general, an R-Squared value of 0.5 is considered to be acceptable. This would imply that the model's independent variables are responsible for explaining fifty percent of the variance in the dependent variable.

$$\textbf{R2 Squared} = 1 - \frac{SSr}{SSm} \quad (4)$$

5.2 Exiting approaches:

The paragraph that follows demonstrates the existing method that is used for observing the flight delay by using linear Regression and Support Vector Regression with a few fundamental parameters,

5.2.1 Linear Regression

```
X = depart_hour_min'
Y = mean
X_train,X_test,Y_train,Y_test =  train_test_split(X,Y,test_size = 0.25)
regr1 = LinearRegression()
poly = PolynomialFeatures(degree = 3)
Xpoly = poly.fit_transform(X_train)
regr1.fit(Xpoly, Y_train)
```

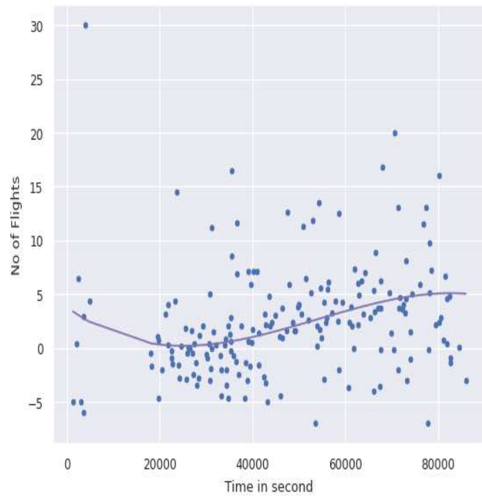

Figure 04: flights delays per second with ref to avg delay line of all the airlines using LR

The number of delayed flights displayed above is updated every second. in reference to the overall average delay line for all airlines, which is computed with the assistance of linear regulation.

5.2.2 SVM-SVR: Support Vector Regression

```
X = depart_hour_min
Y = mean
X_train,X_test,Y_train,Y_test = train_test_split(X,Y,test_size = 0.25)
regr1 = SVR(kernel ='poly', degree = 3)
poly = PolynomialFeatures(degree = 3)
Xpoly = poly.fit_transform(X_train)
regr1.fit(Xpoly, Y_train)
Xpolytest = poly.fit_transform(X_test)
result_poly = regr1.predict(Xpolytest)
```

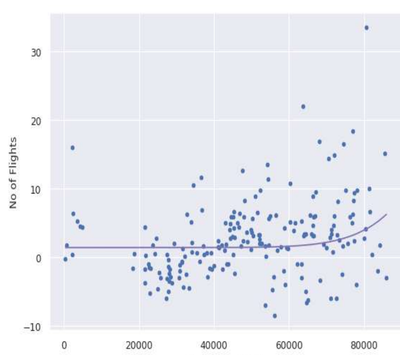

Figure 05: no of flights delays per second With ref to avg delay line of all the airlines using SVM

The number of delayed flights displayed above is updated every second. In reference to the overall average delay line for all airlines, which is computed with the assistance of the SVM-SVR regulation.

The next paragraph provides an overview of the Quantum Machin learning approach, along with linear regression and SVM while maintaining the same parameters.

5.3 QML

Following are Quantum machine learning approach using Linear Regression & SVM by considering upto 5 qubit,

1-Qubit

Quantum Linear Regression

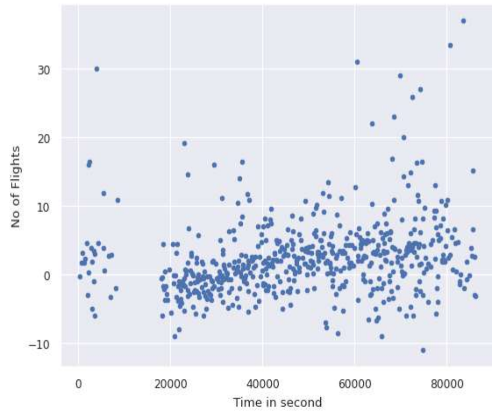

Figure 06: population of no of flights delays per second with help of quantum linear regression.

The population of the number of flights that are delayed per second is shown in the figure that was just above with the help of quantum linear regression. Some of these flights are located in the negative quadrant since that region does not have any depth. However, they are affected by the delay in arrival.

1-Qubit

Quantum SVM-SVR

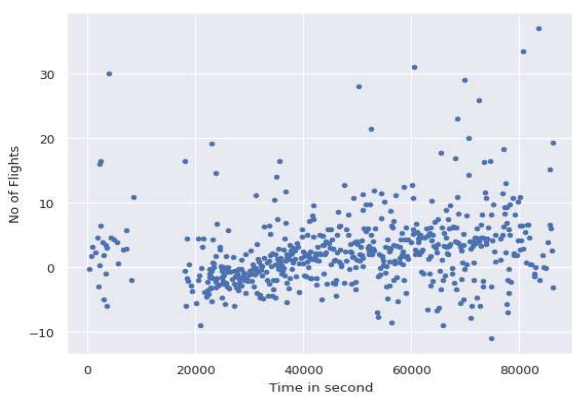

Figure 07: population of no of flights delays per second with help of quantum svm

The population of the number of flights that are delayed per second is shown in the figure that was just above with the help of quantum linear regression. Some of these flights are located in the negative quadrant since that region does not have any depth. However, they are affected by the delay in arrival.

3-Qubit

Quantum Linear Regression

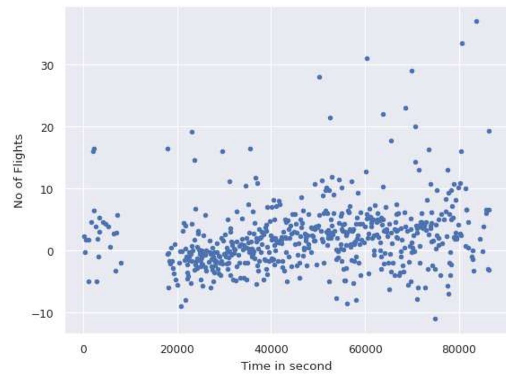

Figure 08: population of no of flights delays per second with help of quantum linear regression.

The population of the number of flights that are delayed per second is shown in the figure that was just above with the help of quantum linear regression. Some of these flights are located in the negative quadrant since that region does not have any depth. Delay, but also those who are affected by the delayed arrival,

3-Qubit

Quantum SVM-SVR

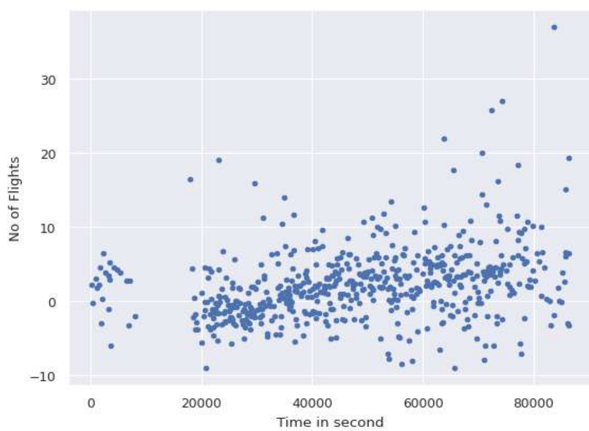

Figure 09: population of no of flights delays per second with help of quantum SVM-SVR regression.

The population of the number of flights that are delayed per second is displayed in the figure that was just above with the assistance of quantum SVM-SVR regression. Some of these flights are located in the negative quadrant since that region does not have any depth. However, they are affected by the delay in arrival.

5-Qubit

Quantum Linear Regression

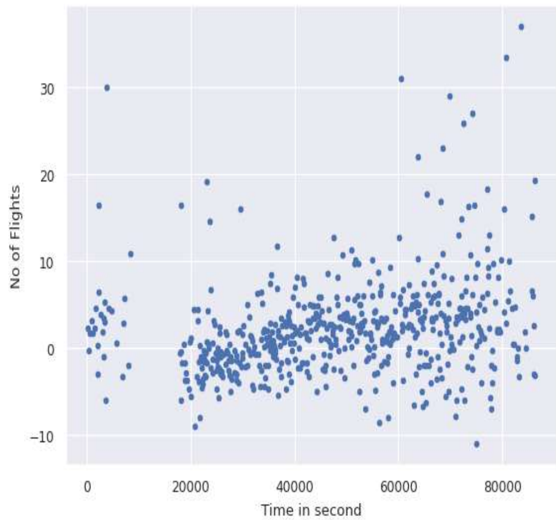

Figure 10: population of no of flights delays per second with help of quantum linear regression.

The population of the number of flights that are delayed per second is shown in the figure that was just above with the help of quantum linear regression. Some of these flights are located in the negative quadrant since that region does not have any depth. Delay, but also those who are affected by the delayed arrival

5-Qubit

Quantum SVM-SVR

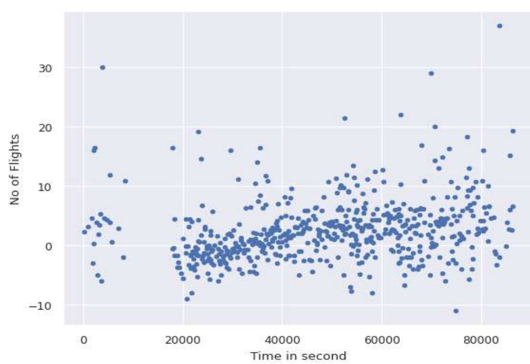

Figure 11: population of no of flight delays per second with help of quantum SVM-SVR regression.

The population of the number of flights that are delayed per second is displayed in the figure that was just above with the assistance of quantum SVM-SVR regression. Some of these flights are located in the negative quadrant since that region does not have any depth. Delay, but also those who are affected by the delayed arrival

Table 02: Comparative Table

Algorithms	**Parameters**		**MSE**	**RMSE**	**MAE**	**R2_Score**
Machine Learning		**Linear Regression**	24.8554	4.9855	3.3640	0.05
		SVM-SVR	30.4346	3.9488	5.5167	0.03

QML	1-Qubit	Q- Linear Regression	15.5555	3.9438	4.0339	0.035
		Q- SVM-SVR	10.2262	3.1978	3.0209	1.0053
	3-Qubit	Q- Linear Regression	15.5532	3.9437	3.9166	0.0393
		Q- SVM-SVR	8.9080	2.9846	3.7877	0.8798
	5-Qubit	Q- Linear Regression	15.5120	3.93853	3.982536	0.054918
		Q- SVM-SVR	14.7466	3.8401	4.203	0.9152

After examining a number of algorithms, as shown in the table above, we came to the conclusion that the strategy that we proposed, which utilized quantum machine learning, was the most effective of all of them and made use of a novel approach.

6. CONCLUSION

Forecasting the length of flight delays is an intriguing topic that has received increased attention in recent years. The majority of research efforts have been focused on developing and expanding models in an effort to improve the predictability of flight delays in terms of both their precision and accuracy. This model is dependent on a dataset containing information about aircraft delays, such as origin, destination, arrival delay, carrier type, etc. Flight delay prediction models need to have a high level of precision and accuracy in order to be useful. Keeping planes on schedule is a very significant concern. According to the analysis of the findings, it is clear that the combination of the integration of multidimensional heterogeneous data with the use of various techniques for feature selection and regression might produce potentially useful tools. In this work, we have used a variety of key Performance Metrics, such as mean squared error, root mean square error, mean absolute error, and R-squared (R2) We have also used popular modern classification algorithms, such as 3qbit, as compared to conventional machine learning after comparing these algorithms with a variety of performance metrics. We have arrived at the conclusion that our system demonstrates that quantum linear regression and quantum support vector machine support vector regression are very ideal for predicting flight delays.

References:

1. Eurocontrol Network Manager Annual Report. 2019. Available online: https://www.eurocontrol.int/publication/network-manager-annual-report-2019 (accessed on 24 February 2021).
2. Eurocontrol Annual Network Operations Report. 2019. Available online: https://www.eurocontrol.int/publication/annual-network-operations-report-2019 (accessed on 24 February 2021).

3. Eurocontrol Five-Year Forecast 2020–2024. Available online: https://www.eurocontrol.int/publication/eurocontrol-five-year-forecast-2020-2024 (accessed on 24 February 2021).
4. Kim, Y.J.; Choi, S.; Briceno, S.; Mavris, D. A deep learning approach to flight delay prediction. In Proceedings of the AIAA/IEEE Digital Avionics Systems Conference, Sacramento, CA, USA, 25–29 September 2016; pp. 1–6. [Google Scholar] [CrossRef]
5. Lambelho, M.; Mitici, M.; Pickup, S.; Marsden, A. Assessing strategic flight schedules at an airport using machine learning-based flight delay and cancellation predictions. J. Air Transp. Manag. 2020, 82, 101737. [Google Scholar] [CrossRef]
6. Choi, S.; Kim, Y.J.; Briceno, S.; Mavris, D. Prediction of Weather-Induced Airline Delays Based on Machine Learning Algorithms. In Proceedings of the IEEE/AIAA 35th Digital Avionics Systems Conference (DASC), Sacramento, CA, USA, 25–29 September 2016. [Google Scholar] [CrossRef]
7. Alonso, H.; Loureiro, A. Predicting flight departure delay at Porto Airport: A preliminary study. In Proceedings of the 2015 7th International Joint Conference on Computational Intelligence (IJCCI), Lisbon, Portugal, 12–14 November 2015; IEEE: New York, NY, USA, 2015; Volume 3, pp. 93–98. [Google Scholar]
8. Chen, J.; Li, M. Chained predictions of flight delay using machine learning. In Proceedings of the AIAA Scitech 2019 Forum, San Diego, CA, USA, 7–11 January 2019; pp. 1–25. [Google Scholar] [CrossRef]
9. Kalliguddi, A.M.; Leboulluec, A.K. Predictive Modeling of Aircraft Flight Delay. Univers. J. Manag. 2017, 5, 485–491. [Google Scholar] [CrossRef][Green Version]
10. Manna, S.; Biswas, S.; Kundu, R.; Rakshit, S.; Gupta, P.; Barman, S. A statistical approach to predict flight delay using gradient boosted decision tree. In Proceedings of the ICCIDS 2017—International Conference on Computational Intelligence in Data Science, Chennai, India, 2–3 June 2018; pp. 1–5. [Google Scholar] [CrossRef]
11. Yu, B.; Guo, Z.; Asian, S.; Wang, H.; Chen, G. Flight delay prediction for commercial air transport: A deep learning approach. Transp. Res. Part E Logist. Transp. Rev. 2019, 125, 203–221. [Google Scholar] [CrossRef]
12. Thiagarajan, B.; Srinivasan, L.; Sharma, A.V.; Sreekanthan, D.; Vijayaraghavan, V. A machine learning approach for prediction of on-time performance of flights. In Proceedings of the AIAA/IEEE Digital Avionics Systems Conference, St. Petersburg, FL, USA, 17–21 September 2017; pp. 5–10. [Google Scholar] [CrossRef]
13. Ayhan, S.; Costas, P.; Samet, H. Predicting estimated time of arrival for commercial flights. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Data Mining, London, UK, 19–23 August 2018; pp. 33–42. [Google Scholar]
14. Shao, W.; Prabowo, A.; Zhao, S.; Tan, S.; Koniusz, P.; Chan, J.; Hei, X.; Feest, B.; Salim, F.D. Flight delay prediction using airport situational awareness map. In Proceedings of the GIS: Proceedings of the ACM International Symposium on Advances in Geofigureic Information Systems, Chicago, IL, USA, 5–8 November 2019; pp. 432–435. [Google Scholar] [CrossRef][Green Version]

15. Mueller, E.; Chatterji, G. Analysis of aircraft arrival and departure delay characteristics. In Proceedings of the AIAA's Aircraft Technology, Integration, and Operations (ATIO) 2002 Technical Forum, Los Angeles, CA, USA, 1–2 October 2002; p. 5866. [Google Scholar]
16. Novianingsih, K.; Hadianti, R. Modeling flight departure delay distributions. In Proceedings of the 2014 International Conference on Computer, Control, Informatics and Its Applications (IC3INA), Bandung, Indonesia, 21–23 October 2014; IEEE: New York, NY, USA, 2014; pp. 30–34. [Google Scholar]
17. Itoh, E.; Mitici, M. Queue-based modeling of the aircraft arrival process at a single airport. Aerospace 2019, 6, 103. [Google Scholar] [CrossRef][Green Version]
18. Kleinbekman, I.C.; Mitici, M.; Wei, P. Rolling-Horizon Electric Vertical Takeoff and Landing Arrival Scheduling for On-Demand Urban Air Mobility. J. Aerosp. Inf. Syst. 2020, 17, 150–159. [Google Scholar] [CrossRef]
19. Tu, Y.; Ball, M.O.; Jank, W.S. Estimating flight departure delay distributions—A statistical approach with long-term trend and short-term pattern. J. Am. Stat. Assoc. 2008, 103, 112–125. [Google Scholar] [CrossRef]
20. Şeker, M.; Noyan, N. Stochastic optimization models for the airport gate assignment problem. Transp. Res. Part E Logist. Transp. Rev. 2012, 48, 438–459. [Google Scholar] [CrossRef]
21. Van Schaijk, O.R.; Visser, H.G. Robust flight-to-gate assignment using flight presence probabilities. Transp. Plan. Technol. 2017, 40, 928–945. [Google Scholar] [CrossRef][Green Version]
22. L'Ortye, J.; Mitici, M.; Visser, H.G. Robust flight-to-gate assignment with landside capacity constraints. Transp. Plan. Technol. 2021, 44, 1–22. [Google Scholar] [CrossRef]
23. Iowa State University. ASOS-AWOS-METAR Data Download. 2020. Available online: https://mesonet.agron.iastate.edu/request/download.phtml (accessed on 1 March 2020).
24. Young, J.K., et al.: A deep learning approach to flight delay prediction.In: Proceedings of Digital Avionics Systems Conference, pp. 1–6. IEEE,Piscataway (2016)
25. Belcastro, L., et al.: Using scalable data mining for predicting flight delays.ACM Trans. Intell. Syst. Technol. 8(1), 1–20 (2016)
26. Daniel, A.P., et al.: Supervised neural network with multilevel input layersfor predicting of air traffic delays. In: Proceedings of International JointConference on Neural Networks, pp. 1–6. IEEE, Piscataway (2018)
27. Rodriguez-Sanza, A., et al.: Assessment of airport arrival congestion anddelay: Prediction and reliability. Transport. Res. Part C: Emerg. 98, 255–283 (2019)
28. Lambelho, M., et al.: Assessing strategic flight schedules at an airport usingmachine learning-based flight delay and cancellation predictions. J. AirTransport Manage. 82, 101737.1–101737.11 (2020)
29. Guan, G., et al.: Flight delay prediction based on aviation big data andmachine learning. IEEE Trans. Veh. Technol. 69(1), 140–150 (2020).https://doi.org/10.1109/TVT.2019.2954094.

30. McCarthy, N., et al.: Amsterdam to Dublin eventually delayed? LSTM andtransfer learning for predicting delays of low cost airlines. In: Proceedingsof the AAAI Conference on Artificial Intelligence, vol. 33, pp. 9541–9546.AAAI Press, Palo Alto (2019)31. Chen, J., Li, M., Chained Predictions of Flight Delay Using Machine Learning. AIAA Scitech 2019 Forum. p. 1661 (2019)
31. Joint Economic Committee report: Your flight has been delayedagain: Flight delays cost passengers, airlines, and the U.S. economybillions.https://www.jec.senate.gov/public /index.cfm/democrats/2008/5/your-flight-has-been-delayed-again_1539. Accessed 15 July 2020