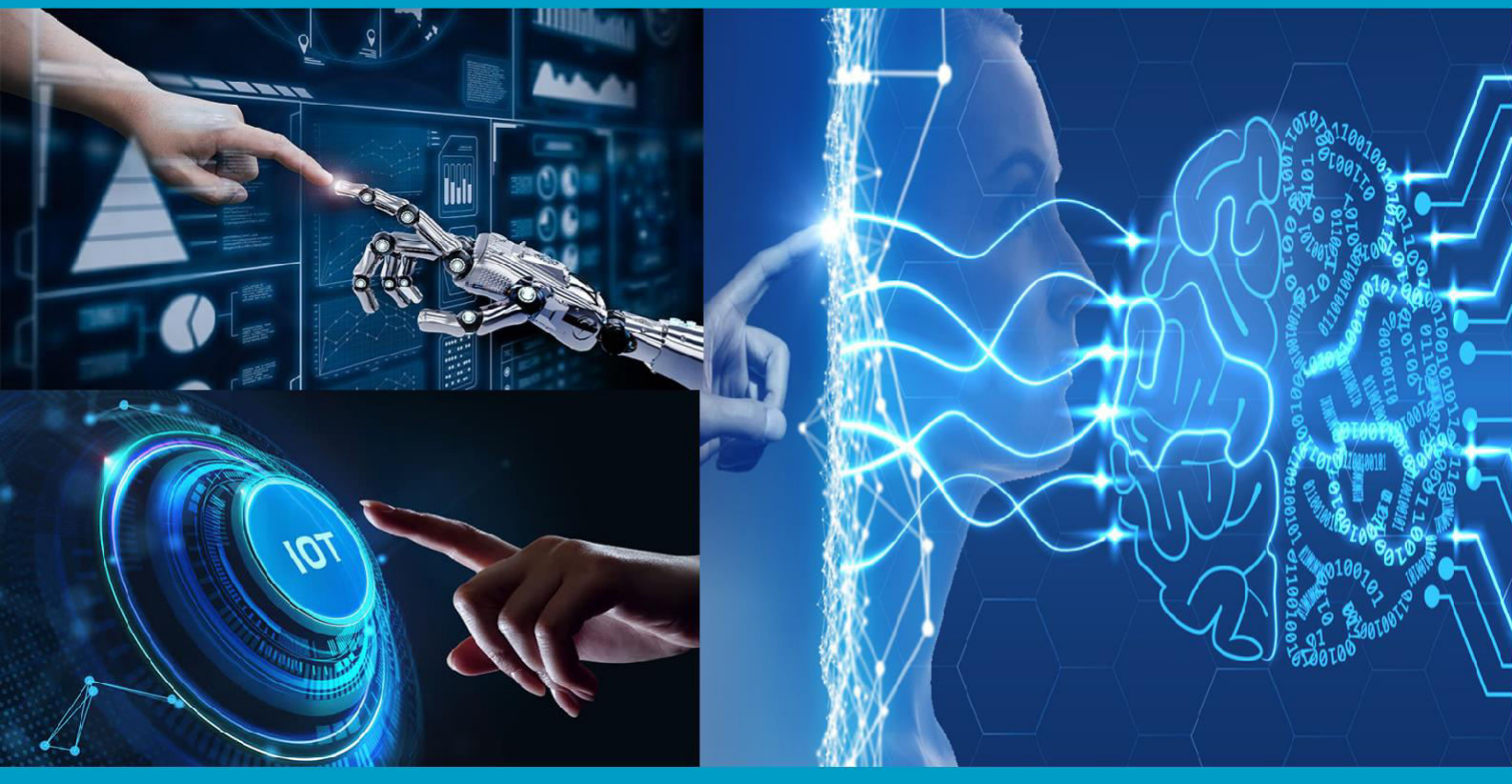
IJMRSETM

e-ISSN: 2395 - 7639



INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

Volume 9, Issue 12, December 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.580

International Journal of Multidisciplinary Research in Science, Engineering, Technology & Management (IJMRSETM)

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

Flight Delay Prediction with Machine Learning

[1]T Prasanna Krishna, [2]Subhani Shaik

[1]PG Scholar, Dept of AIML, St Marys Group of Institutions Guntur, AP, India

[2]Assistant Professor, Dept of CSE, St Marys Group of Institutions Guntur, AP, India

ABSTRACT: Flight is an important mode of transport. In fact, it is the only time-feasible mode of transport for travelling abroad. Every airport across the globe witnesses frequent flight delays for many unforeseen reasons. In the contemporary era, there is Artificial Intelligence (AI) usage is prevalent in every domain. Transportation domain is no exception. Due to phenomenal growth of aviation sector across the globe, it became very important to ascertain possible flight delays using AI based methods. In this paper, we proposed a Machine Learning (ML) based framework for automatic prediction of flight delays. It is meant for taking multiple datasets as inputs and have mechanisms for feature selection and learning based approach. The training data is used for learning from the data and gain knowledge. Then the resultant knowledge or prediction model is used to predict possible flight delays. An algorithm known as Predicting Flight Delays with Error Calculation (PFD-EC) is proposed to realize the framework. The experimental results revealed that the proposed system has significant accuracy in prediction of flight delays.

KEYWORDS: Flight delay prediction, machine learning, feature selection, knowledge models

I. INTRODUCTION

In the present world, the major components of any transportation system include passenger airline, cargo airline, and air traffic control system. With the passage of time, nations around the world have tried to evolve numerous techniques of improving the airline transportation system. This has brought drastic change in the airline operations. Flight delays occasionally cause inconvenience to the modern passengers [1]. Every year approximately 20% of airline flights are cancelled or delayed, costing passengers more than 20 billion dollars in money and their time. Average aircraft delay is regularly referred to as an indication of airport capacity. Flight delay is a prevailing problem in this world. It's very tough to explain the reason for a delay. A few factors responsible for the flight delays like runway construction to excessive traffic are rare, but bad weather seems to be a common cause. Some flights are delayed because of the reactionary delays, due to the late arrival of the previous flight. It hurts airports, airlines, and affects a company's marketing strategies as companies rely on customer loyalty to support their frequent flying programs.

From the literature, many existing ML models are found for flight delay prediction. Gopalakrishnan and Balakrishnan [4] investigated on different methods of predicting delays in flight arrivals and air traffic control using AI based phenomena. Alharbi and Prince [5] proposed a hybrid approach considering AI enabled models towards making a system that predicts flight delays. Musaddi *et al.* [6] proposed a model based on binary classification using ML approach for flight delay prediction process. Al-Tabbakh *et al.* [7] has explored Egyptian flight delays by exploiting ML techniques. More *et al.* [8] used back propagation and radial bus function combination for flight delay prediction. Besides, they made comparative analysis with different models in ML. From the review of literature, it is understood that there is need for considering more appropriate factors in ML models for improving prediction performance. Our contributions in this paper are as follows.

1. We proposed a ML framework for automatic prediction of flight delays based on different ML techniques.
2. We proposed an algorithm named Predicting Flight Delays with Error Calculation (PFD-EC) for prediction of flight delays.
3. We proposed a web based application to realize the framework and evaluate its performance.

The remainder of the paper is structured as follows. Section 2 reviews literature relevant to flight delay prediction. Section 3 presents proposed framework and algorithm. Section 4 presents results and discussion. Section 5 concludes the work and provides future work.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

II. RELTED WORK

This section reviews literature on existing methods meant for prediction of flight delay. Meel *et al.* [1] explored ML models and error computation approach towards detecting delays in flight arrivals. Yazdi *et al.* [2] explored a deep learning based prediction model for forecasting flight delay probabilities. Levenberg-Marquart algorithm and deep learning are combined to achieve this. Kuhn and Jamadagni [3] discussed different ML models and their mechanisms suitable for prediction of flight delays. Gopalakrishnan and Balakrishnan [4] investigated on different methods of predicting delays in flight arrivals and air traffic control using AI based phenomena. Alharbi and Prince [5] proposed a hybrid approach considering AI enbled models towards making a system that predicts flight delays. Musaddi *et al.* [6] proposed a model based on binary classification using ML approach for flight delay prediction process. Al-Tabbakh *et al.* [7] has explored Egyptian flight delays by exploiting ML techniques. More *et al.* [8] used back propagation and radial bus function combination for flight delay prediction. Besides, they made comparative analysis with different models in ML. Stefanovič *et al.* [9] considered flights of Lithuanian airport for examination. They proposed a prediction model based on Support Vector Machine (SVM) in order to predict possible deviations in flight arrivals. Elangovan *et al.* [10] proposed system based on ML for predictive analysis of flight departures. Their system was capable of finding flight delays more accurately.

Ugwu *et al.* [11] invented a decision support system that dynamically considers different kinds of inputs to know the possible deviations in flight arrivals. Their system exploits Decision Tree (DT) for prediction. Achenbach and Spinler [12] investigated on airline operations with regard to arrival of flights and their prediction and optimization of indices. It could provide details of predictions for short-haul flights. Chen *et al.* [13] made chained approach for predictions of flight delays using ML models. This approach collaborates knowledge of different models. Vandal *et al.* [14] used historical flight data in order to make predictions and quantify flight delays on daily basis. It could help in having certain clarity on making important decisions. Xu *et al.* [15] considered several factors such as historical data, weather, forecast of rains etc. in order to predict possible flight delays. Guo *et al.* [16] studied ML models and applied them in order to identify possible occurrences in which passengers are to be transferred to different airport or platforms based on real time data analysis. Belcastro *et al.* [17] proposed a scalable approach in prediction of flight delays. It exploits computing resources associated with distributed computing in order to have more scalable and efficient prediction of flight delays. Other contributions include predictive modelling [18], ensemble learning [19], air traffic flow management system [20] and trajectory prediction approach [21]. From the review of literature, it is understood that there is need for considering more appropriate factors in ML models for improving prediction performance.

III. PROPOSED METHODOLOGY

We proposed a web based interface along with ML models in order to exploit historical data to forecast flight delays. It is the system that is based AI which injects knowledge to a ML algorithm. It is a supervised learning approach mixed with traditional web based interface that considers different factors. It makes use of multiple datasets covering flight dataset, weather dataset and airport events dataset. It considers several factors as illustrated in Figure 1. The final flight information dataset is a hybrid with required features that are used to train AI based classifier. The supervised learning approach enables learning based phenomenon that gains knowhow of flight arrival and departure dynamics. Based on the learning process, the proposed system gains substantial knowledge and provide highly accurate predictions.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

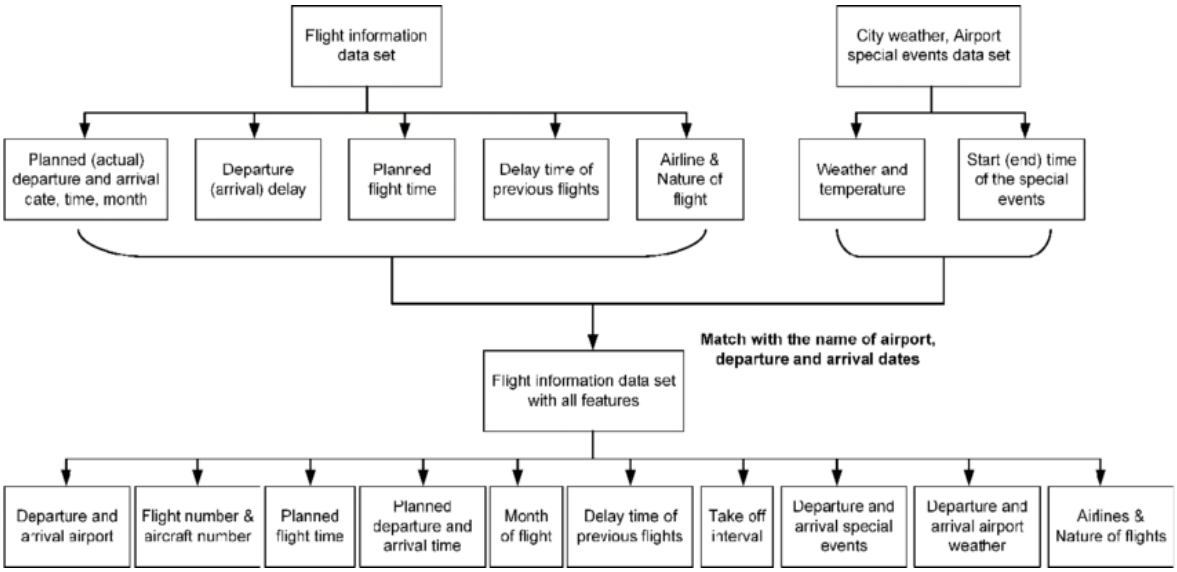


Fig 1: Different factors considered in the ML process

As presented in Fig. 1, it is understood that the data considered for flight arrival and departure prediction has many features considered. It has actual arrival time and departure time, planned flight time, delay time of previous trips, weather, temperature, events at airport and their temporal dimensions. After considering the historical data, the proposed system is based on the learning based approach illustrated in Fig. 2.

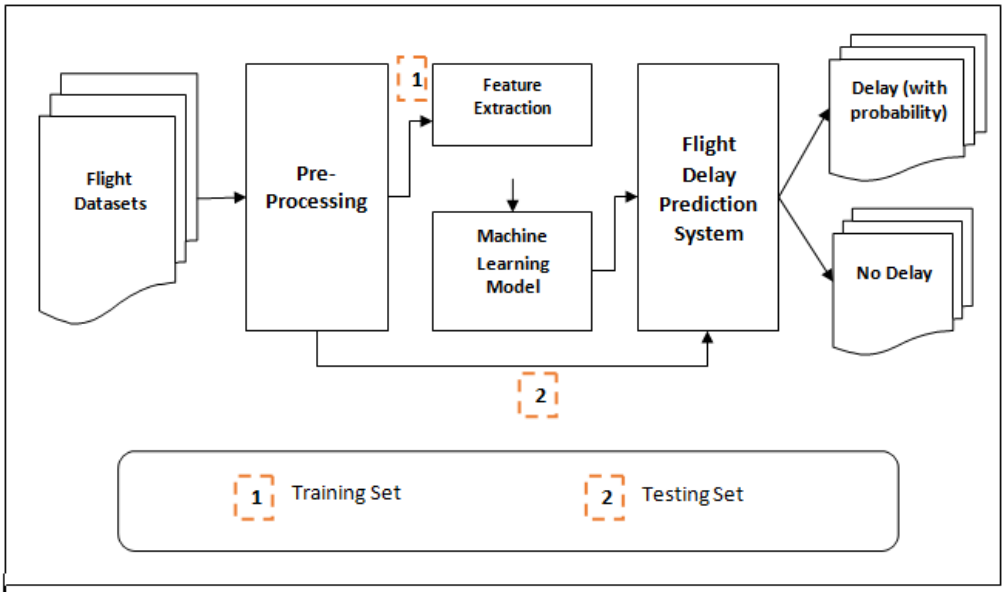


Fig. 2: ML based framework for flight delays prediction

As presented in Fig. 2, the datasets are combined in order to have a sophisticated dataset. Then it is subjected to pre-processing to get training and test datasets. The training data is given to feature selection mechanism which selects contributing features. Then the selected features are given to ML models in order to train classifiers. After gaining knowledge through AI based approach, it will result in a knowledge model that can take test data and perform predictions.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

3.1Algorithm

Algorithm: Predicting Flight Delays with Error Calculation (PFD-EC)
Inputs: airport.csv,carrier.csv,flight_data.csv, prediction models M
Output: predicting flight delay results,

1. Start
2. Input feedback dataset
3. Pre-processing
4. Extract features from training set()
5. For each model m in M
6. Train the model m
7. End For
8. For each model m in M
9. Use model for testing
10. Evaluate
11. Display results
12. End For
13. End

Algorithm 1 Proposed algorithm

As presented in Algorithm 1, it takes input datasets and ML models in order to have a prediction system to know flight delays. It has an iterative approach for training each model and then perform prediction process.

3.2 Performance Evaluation Metrics

Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are the two metrics used to evaluate the proposed ML models for flight delay prediction.

IV. RESULTS AND DISCUSSION

The proposed system using AI based approach is linked to a web based application. It has provisions for exploiting historical data and perform flight delay predictions.

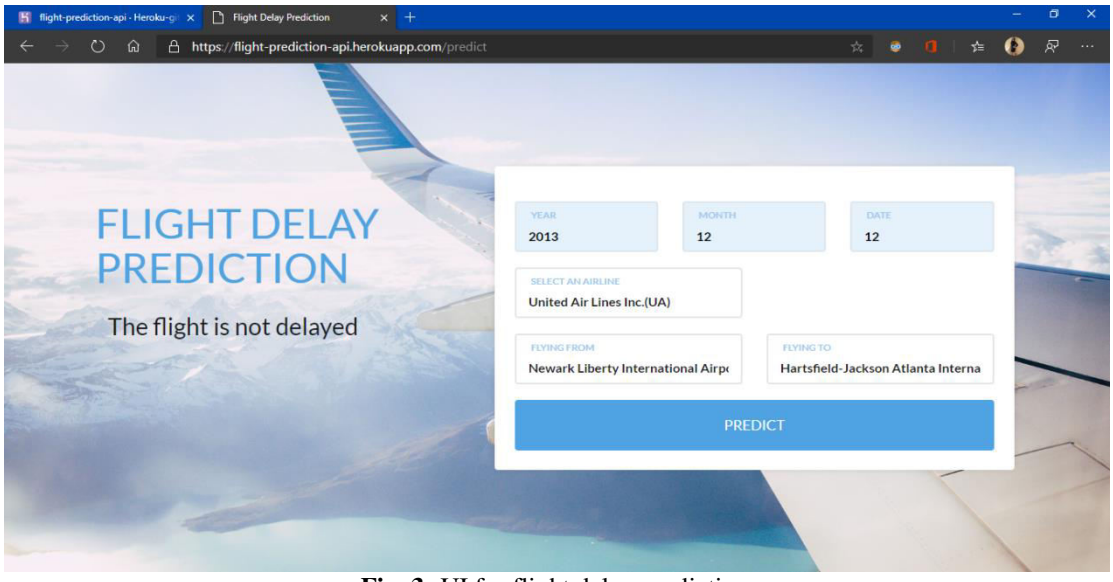


Fig. 3: UI for flight delay prediction

As presented in Fig. 3, flight delay prediction is provided and this is possible due to underlying supervised learning model that gains knowledge from training data.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

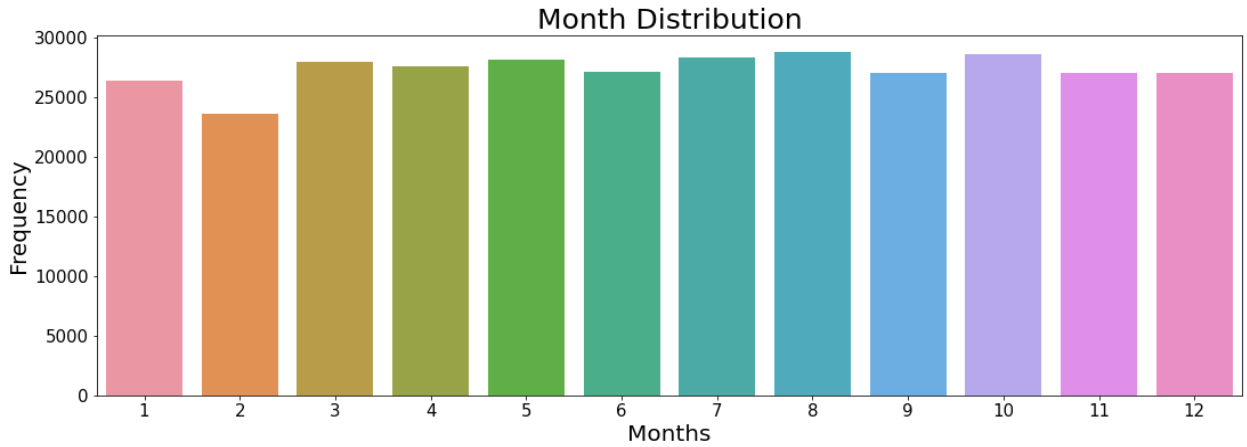


Fig. 4: Flight frequency distribution month wise

As presented in Figure 4, flight frequency is provided for a given year. It visualized flight frequency for 12 months of the given year.

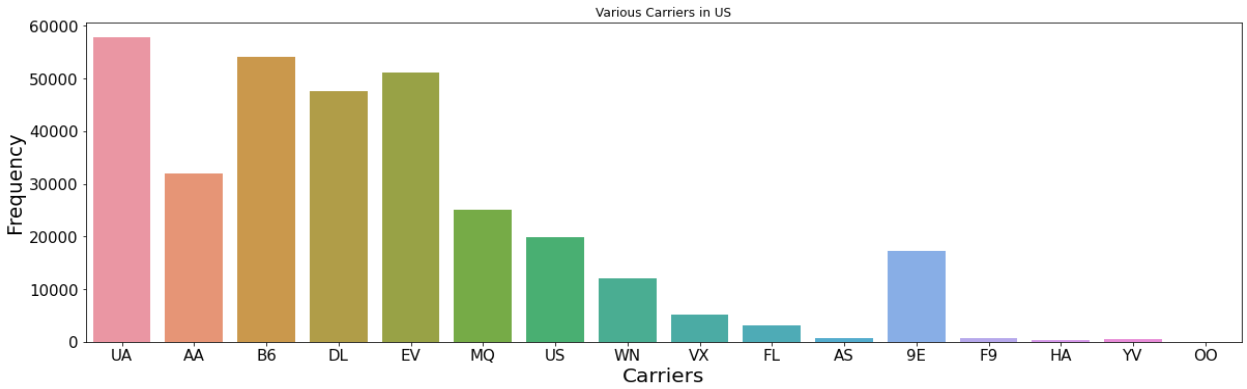


Fig. 5: Various flight carriers and their frequency

As presented in Fig. 5, flight carriers to different destinations are considered. Each carrier has different frequency as indicated in visualization.

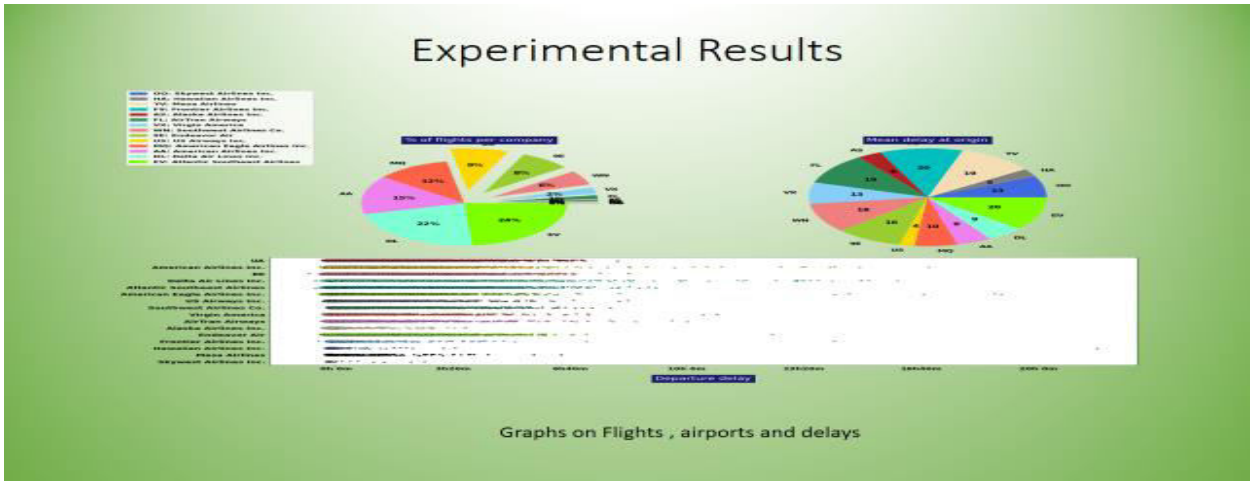


Fig. 6: Results of experiments

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

As presented in Fig. 66, flight delays are estimated as per the proposed algorithm and results are provided for different flights in different airports.

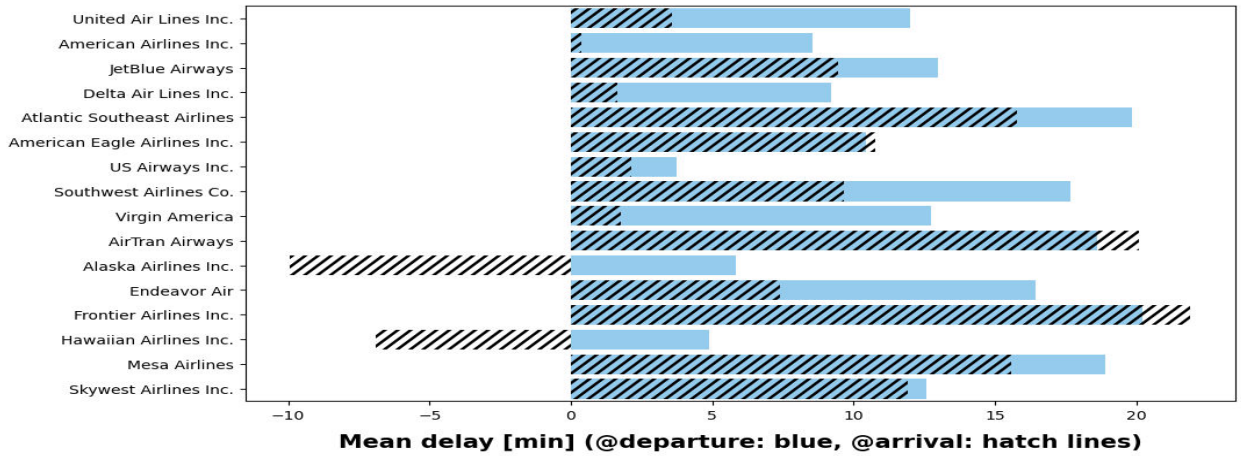


Fig. 7: The mean delay time of different airlines

As Shown in the Fig.7, the mean delay time for each airline are provided in terms of arrival time and departure time.

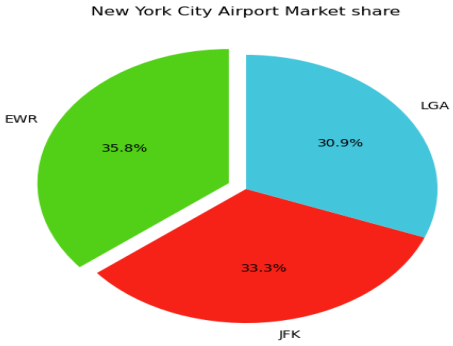


Fig. 8: Market share of different airports in Ney York city

As presented in Figure 8, it is evident that each airport has its market share in the given city. Highest market share is by EWR with 35.8%.

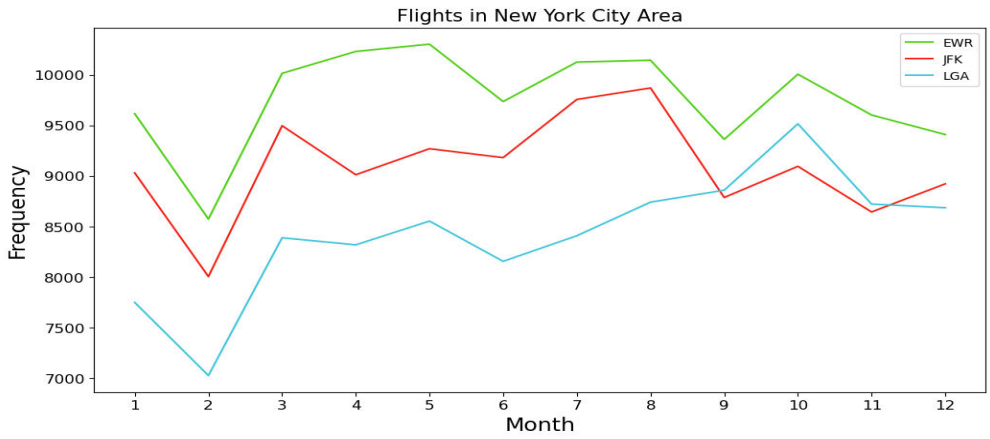


Fig. 9: Market share of different airports in Ney York city

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

As presented in Figure 9, it is evident that each airport has its frequency in the given city. Highest frequency is by EWR with more than 10000.

Model	MAE	RMSE
KNear Neighbours	0.276584695	0.525913201
Logistic Regression	0.279776997	0.528939502
Naïve Bayes	0.280355179	0.529485768
Random Forest	0.298068111	0.545956144

Table 1: Shows performance comparison among ML models

As presented in Table 1, MAE and RMSE values of different ML models are provided for performance comparison.

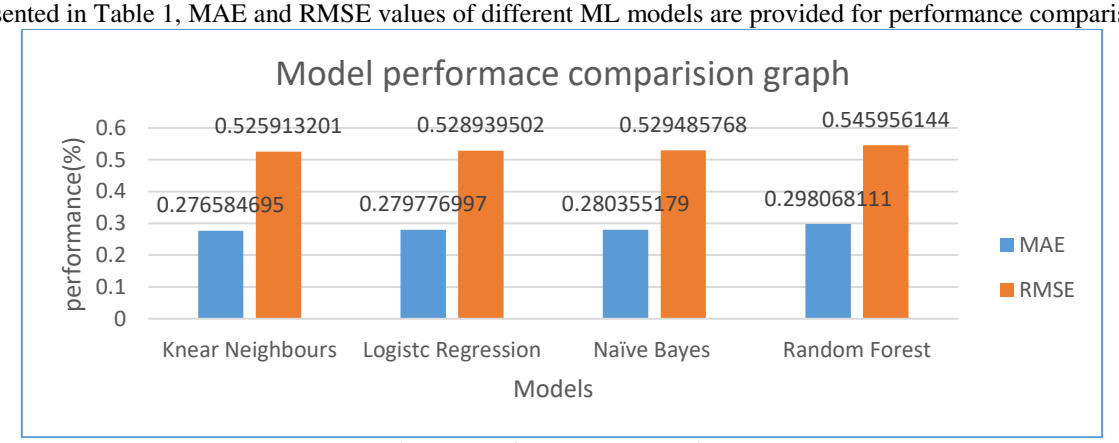


Fig. 10: Performance comparison

As presented in Figure 10, performance of different prediction models is provided. More in MAE or RMSE indicates less performance. KNN model showed highest performance in terms of MAE and RMSE.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a Machine Learning (ML) based framework for automatic prediction of flight delays. It is meant for taking multiple datasets as inputs and have mechanisms for feature selection and learning based approach. The training data is used for learning from the data and gain knowledge. Then the resultant knowledge or prediction model is used to predict possible flight delays. An algorithm known as Predicting Flight Delays with Error Calculation (PFD-EC) is proposed to realize the framework. The experimental results revealed that the proposed system has significant accuracy in prediction of flight delays. In future we intend to exploit deep learning models for improving prediction performance in flight delays.

REFERENCES

[1]Meel, P., Singhal, M., Tanwar, M., &Saini, N. (2020). Predicting Flight Delays with Error Calculation using Machine Learned Classifiers.2020 7th International Conference on Signal Processing and Integrated Networks (SPIN).
[2]Yazdi, M. F., Kamel, S. R., Chabok, S. J. M., &Kheirabadi, M. (2020). Flight delay prediction based on deep learning and Levenberg-Marquart algorithm.Journal of Big Data, 7(1).
[3] Nathalie Kuhn∗ and NavaneethJamadagni†.(2017). Application of Machine Learning Algorithms to Predict Flight Arrival Delays, p1-6.
[4] KarthikGopalakrishnan and HamsaBalakrishnan. (2017). A Comparative Analysis of Models for Predicting Delays in Air Traffic Networks. Twelfth USA/Europe Air Traffic Management Research and Development Seminar, p1-10.
[5] Bashayer Alharbi1 and Master Prince2,*.(2020). A Hybrid Artificial Intelligence Approach to Predict Flight Delay. International Journal of Engineering Research and Technology.13 (4), p814-822.
[6] RoshniMusaddi1 , Anny Jaiswal 2 , Pooja J 3 , MansviGirdonia 4 , Minu M.S 5. (2018). Flight Delay Prediction using Binary Classification. International Journal of Emerging Technologies in Engineering Research (IJETER).6 (10), p34-38.

| ISSN: 2395-7639 | www.ijmrsetm.com | Impact Factor: 7.580|

| Volume 9, Issue 12, December 2022 |

| DOI: 10.15680/IJMRSETM.2022.0912014 |

[7] M. Al-Tabbakh, S., M. Mohamed, H., & H. El, Z. (2018). Machine Learning Techniques for Analysis of Egyptian Flight Delay. International Journal of Data Mining & Knowledge Management Process, 8(3), 01–14.
[8]Nachiket More1, Karan Pagare2, Meet Bhanushali3, Dr. Kishor T. Patil4. (2020). Comparative Study of Flight Delay Prediction using Back Propagation and Radial Basis Function. International Research Journal of Engineering and Technology (IRJET).7 (3), p5092-5097.
[9] Stefanovič, P., Štrimaitis, R., &Kurasova, O. (2020). Prediction of Flight Time Deviation for Lithuanian Airports Using Supervised Machine Learning Model. Computational Intelligence and Neuroscience, 2020, 1–10.
[10]ThavanisaElangovan, 2Mafas Raheem, 3Amad Arshad. (2020). PREDICTIVE ANALYTICS FOR AIRLINE DEPARTURE DELAYS. JOURNAL OF CRITICAL REVIEWS.7 (11), p4034-4039.
[11] C. Ugwu1 ,Ntuk, Ekaete2. (2016). DYNAMIC DECISION TREE BASED ENSEMBLED LEARNING MODEL TO FORECAST FLIGHT STATUS. European Journal of Computer Science and Information Technology.4 (6), p15-24.
[12] Achenbach, A., &Spinler, S. (2018). Prescriptive analytics in airline operations: Arrival time prediction and cost index optimization for short-haul flights. Operations Research Perspectives, 5, 265–279.
[13]Chen, J., & Li, M. (2019). Chained Predictions of Flight Delay Using Machine Learning.AIAA Scitech 2019 Forum.
[14] Thomas Vandal, Max Livingston, CamenPiho and Sam Zimmerman.(2017). Prediction and Uncertainty Quantification of Daily Airport Flight Delays. Proceedings of Machine Learning Research.82, p45-51.
[15] YiyangXu1 ,Luyao Liu2 , Xichen Gao2 and Fanyu Frank Zeng3,*. (2019). Analysis of Factors in Flight Delay . 2nd International Conference on Mathematics, Modeling and Simulation Technologies and Applications .93, p170-176.
[16] Guo, X., Grushka-Cockayne, Y., & De Reyck, B. (2018). Forecasting Airport Transfer Passenger Flow Using Real-Time Data and Machine Learning.SSRN Electronic Journal.
[17]Belcastro, L., Marozzo, F., Talia, D., &Trunfio, P. (2016). Using Scalable Data Mining for Predicting Flight Delays. ACM Transactions on Intelligent Systems and Technology, 8(1), 1–20.
[18]Anish M. Kalliguddi* ,Aera K. Leboulluec. (2017). Predictive Modeling of Aircraft Flight Delay. Universal Journal of Management.5 (10), p485-491.
[19] ZahabiyaMhowwala1 , A. Razia Sulthana2 , Sujala D. Shetty3. (2020). Movie Rating Prediction using Ensemble Learning Algorithms. (IJACSA) International Journal of Advanced Computer Science and Applications.11 (8), p383-388.
[20] DeepudevS.1 ,Palanisamy P.2 , Varun P. Gopi3 , Manjunath K. Nelli4 and Asokkumar K.5. (2020). Performance Improvement of Air Traffic Flow Management Ground Delay Program using Machine-learning and Mixed Integer Linear Programming based Algorithm. International Journal on Emerging Technologies .11 (2), p1071-1081.
[21] Esther CalvoFernández, José Manuel Cordero, George Vouros, Nikos Pelekis, TheocharisKravaris, Georg Fuchs, Natalya Andrienko, Gennady Andrienko and Enrique Casado, David Scarlatti, Pablo Costas and . (2020). DART: A Machine-Learning Approach to Trajectory Prediction and Demand-Capacity Balancing. Seventh SESAR Innovation Days, p1-8.

Impact Factor: 7.580

INTERNATIONAL JOURNAL
OF MULTIDISCIPLINARY RESEARCH
IN SCIENCE, ENGINEERING, TECHNOLOGY AND MANAGEMENT

+91 99405 72462 +91 63819 07438 ijmrsetm@gmail.com