ISSN: 2454-9940

# INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

# SYSTEM FOR PREDICTING FLIGHT DELAYS BASED ON MACHINE LEARNING

**[1]B. HARI BABU, ,[2] G. HARIPRIYA,[3] K. KAVITHA, [4] L. SWATHI**

[1] Associate Professor, Department of Computer Science Engineering,
**AUDISANKARA COLLEGE OF ENGINEERING & TECHNOLOGY**,
NH-16, By-Pass Road, Gudur, Tirupati Dist, Andhra Pradesh, -524101, India.
[234] Associate Professor, Department of Computer Science Engineering,
**AUDISANKARA COLLEGE OF ENGINEERING & TECHNOLOGY**,
NH-16, By-Pass Road, Gudur, Tirupati Dist, Andhra Pradesh, -524101, India.

## ABSTRACT

Flight delay is a major problem in the aviation sector. During the last two decades, the growth of the aviation sector has caused air traffic congestion, which has caused flight delays. Flight delays result not only in the loss of fortune also negatively impact the environment. Flight delays also cause significant losses for airlines operating commercial flights. Therefore, they do everything possible in the prevention or avoidance of delays and cancellations of flights by taking some measures. In this paper, using machine learning models such as Logistic Regression, Decision Tree Regression, Bayesian Ridge, Random Forest Regression and Gradient Boosting Regression we predict whether the arrival of a particular flight will be delayed or not.

**Index Terms**— Flight Prediction, Machine Learning, Error Calculation, Logistic Regression, Decision Tree, Bayesian Ridge, Rando.

## I. INTRODUCTION

Flight delay is studied vigorously in various research in recent years. The growing demand for air travel has led to an increase in flight delays. According to the Federal Aviation Administration (FAA), the aviation industry loses more than $3 billion in a year due to flight delays [1] and, as per BTS [2], in 2016 there were 860,646 arrival delays. The reasons for the delay of commercial scheduled flights are air traffic congestion, passengers increasing per year, maintenance and safety problems, adverse weather conditions, the late arrival of plane to be used for next flight [3] [4]. In the United States, the FAA believes that a flight is delayed when the scheduled and actual arrival times differs by more than 15 minutes. Since it becomes a serious problem in the United States, analysis and prediction of flight delays are being studied to reduce large costs.

## II. LITERATURE SURVEY

Much research has been done on studying flight delays. The prediction, analysis and cause of flight delays have been a major problem for air traffic control, decision-making by airlines and ground delay response programs. Studies are conducted on the delay propagation of the sequence. Also, studying the predictive model of arrival delay and departure delay with meteorological features is encouraged. In the past, researchers have tried to predict flight delays with Machine Learning. Chakrabarty et al. [5] used supervised automatic learning algorithms (random forest, Gradient Boosting Classifier, Support Vector Machine and

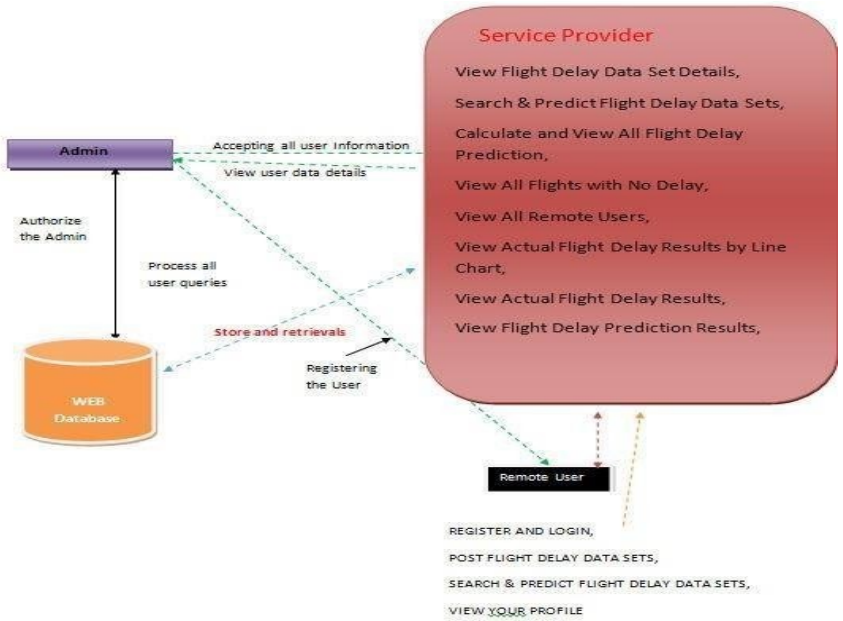
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

the k nearest neighbour algorithm) to predict delays in the arrival of operated flights including the five busiest US airports. The maximum precision achieved was 79.7% with gradient booster as a classifier with a limited data set. Choi et al. [6] applied machine learning algorithms like decision tree, random forest, AdaBoost and k Nearest Neighbors to predict delays on individual flights. Flight schedule data and weather forecasts have been incorporated into the model. Sampling techniques were used to balance the data and it was observed that the accuracy of the classifier trained without sampling was more that of the trained classifier with sampling techniques.

Cao et al. [7] used a Bayesian Network model to analyse the turnaround time of a flight and delay prediction. Juan José Rebollo and Hamsa Balakrishnan [8] used a hundred pairs of origin and destination to summaries the result of various regression and classification models. The find outs reveal that among all the methods used, random forest has the highest performance. However, predictability may additionally range because of factors such as the number of origin destination pairs and the forecast horizon. Sruti Oza, Somya Sharma [9] used multiple linear regression to predict weather induced flight delays in flight data, as well as climatic factors and probabilities due to weather delays. The forecasts were based on some key attributes, such as carrier, departure time, arrival time, origin and destination. Anish M. Kalliguddi and Aera K. Leboulluec [10] predicted both departure and arrival delays using regression models such as Decision Tree Regressor, Multiple Linear Regression and Random Forest Regressor in flight-data. It has been observed that the longer forecast horizon is useful for increasing the accuracy with a minimum forecast error for random forests. Etani J Big Data [11] A supervised model of on-schedule arrival fight is used using weather data and flight data.

## III. PROPOSED SYSTEM

The overview of our proposed system is shown in the below figure.



INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

**Fig-1: System Overview**

### Implementation Modules

Service Provider

In this module, the sp has to login by using valid user name and password. After login successful he can do some operations such as View Flight Delay Data Set Details, Search & Predict Flight Delay Data Sets, Calculate and View All Flight Delay Prediction, View All Flights with No Delay, View All Remote Users, View Actual Flight Delay Results by Line Chart, View Actual Flight Delay Results, View Flight Delay Prediction Results. User in this module, there are n numbers of users are present. User should register before doing some operations. After registration successful he has to login by using authorized user name and password. Login successful he will do some operations like POST FLIGHT DELAY DATA SETS, SEARCH & PREDICT FLIGHT DELAY DATA SETS, VIEW YOURPROFILE.

## IV. RESULTS


**Fig-2: Home Page**


**Fig. 3: View All Flight Details**

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

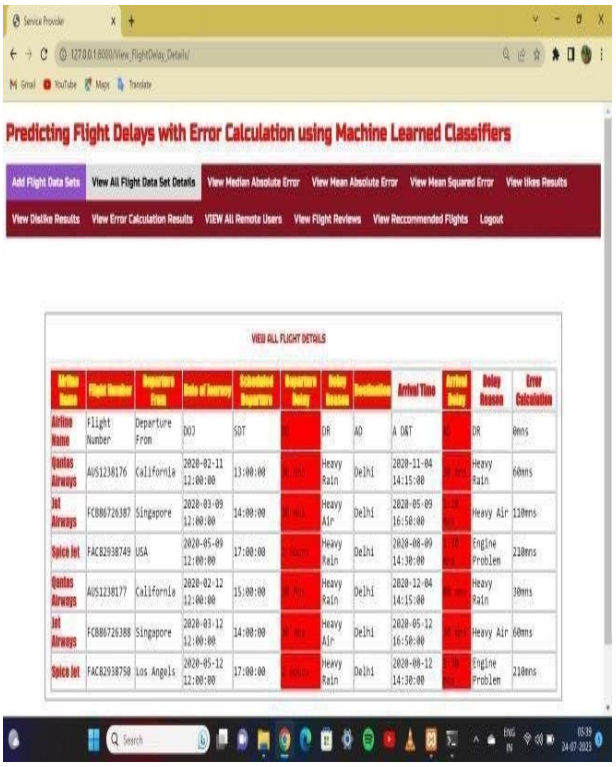



Fig-4: User Registration Page


Fig-5: Search Delay Flight Details

## V. CONCLUSION

Machine learning algorithms were applied progressively and successively to predict flight arrival & delay. We built five models out of this. We saw for each evaluation metric considered the values of the models and compared them. In Departure Delay, Random Forest Regressor was observed as the best model with Mean Squared Error 2261.8 and Mean Absolute Error 24.1, which are the

minimum value found in these respective metrics. In Arrival Delay, Random Forest Regressor was the best model observed with Mean Squared Error 3019.3 and Mean Absolute Error 30.8, which are the minimum value found in these respective metrics.

In the rest of the metrics, the value of the error of Random Forest Regressor although is not minimum but still gives a low value comparatively. In maximum metrics, we found out that Random Forest Regressor gives us the best value and thus should be the model selected.

The future scope of this paper can include the application of more advanced, modern and innovative preprocessing techniques, automated hybrid learning and sampling algorithms, and deep learning models adjusted to achieve better performance.

To evolve a predictive model, additional variables can be introduced. e.g., a model where meteorological statistics are utilized in developing error-free models for flight delays. In this paper we used data from the

ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

US only, therefore in future, the model can be trained with data from other countries as well. With the use of models that are complex and hybrid of many other models provided with appropriate processing power and with the use of larger detailed datasets, more accurate predictive models can be developed.

Additionally, the model can be configured for other airports to airports would be required to incorporate into this research.

**REFERENCE**

1. N. G. Rupp, "Further Investigation into the Causes of Flight Delays," in Department of Economics, East Carolina University, 2007.
2. "Bureau of Transportation Statistics (BTS) Databases and Statistics," [Online]. Available: http://www.transtats.bts.gov.
3. Mr. Ch. Dayakar, V. Chinnabbigari Srinivasulu, “Stock Portfolio Management System”, Journal of Engineering Sciences, Vol 15, Issue 08, ISSN:0377-9254, 2024.
4. Penchalaiah, N. and Seshadri, R. “Effective Comparison and Evaluation of DES and Rijndael Algorithm (AES)”, International Journal of Computer Science and Engineering, Vol. 02, No. 05, 2010, 1641-1645.
5. Navoneel, et al., Chakrabarty, "Flight Arrival Delay Prediction Using Gradient Boosting Classifier," in Emerging Technologies in Data Mining and Information Security, Singapore, 2019.
6. M Rajaiah, A Sudhakaraih, S V K Varma and P Venkatalakshmi (2015): Chemical and Soret effect on MHD free convective flow past an accelerated vertical plate in presence of inclined magnetic field through porous medium. i-manager’s Journal on Mathematics, Vol. 4(1), pp. 32-39.
7. Shyja, V. & Ranganathan, Govindasamy & Bindhu, V.. (2023). Link quality and energy efficient optimal simplified cluster based routing scheme to enhance lifetime for wireless body area networks. Nano Communication Networks. 37.
8. S. Sharma, H. Sangoi, R. Raut, V. C. Kotak, S. Oza, "Flight Delay Prediction System Using Weighted Multiple Linear Regression," International Journal of Engineering and Computer Science, vol. 4, no. 4, pp. 11668 - 11677, April 2015.
9. A. M. Kalliguddi, Area K., Leboulluec, "Predictive Modelling of Aircraft Flight Delay," Universal Journal of Management, pp. 485 -491, 2017.