IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 5, Issue 2, May 2021
Impact Factor: 4.819

Flight Delay Prediction Using Machine Learning

Prajwal Dhone[1], Uday Kirange[2], Rushabh Satarkar[3] and Prof. Shashant Jaykar[4]
Students, Department of Electronics and Telecommunication Engineering[1,2,3]
Professor, Department of Electronics and Telecommunication Engineering[4]
Rajiv Gandhi College of Engineering and Research, Nagpur, Maharashtra, India
Affiliation to Rastrasant Tukdoji Maharaj Nagpur University

Abstract: *In this fast growing world as airplanes continue flying, flight delays are the part of the experience. According to the Bureau Of Statistics(BOS), about 20% of all flights are delayed by 15 minutes or more. Flight delays causes a negative impact, mainly economical for airport authorities, commuters and airline industries as well. Furthermore, in the domain of sustainability, it can even cause environmental harm by the rise in fuel consumption and gas emissions and also some of the important factors including adverse weather conditions, preparing the aircraft, fixing of mechanical issue, getting security clearance, etc. Hence, these are the factors which indicates the necessity it has become to predict the delays of airline problems. To carry out the predictive analysis, which includes a range of statistical techniques from machine learning, this studies historical and current data to make predictions about the future delays, taking help of Regression Analysis using regularization technique used in Python.*

Keywords: Flight delay, predictive analysis, machine leaning, regression analysis.

I. INTRODUCTION

The primary goal of this paper is to predict the airline delays caused by various factors such as technical issue with aircraft, bad weather, air traffic control problems, bird strikes. Major problem arrises in the current aviation system due to flight delay. Some methods are required to introduce how delay propagates in the airport networks. Flight delays becomes an important subject and problem for air transportation system all over the world. The aviation industry is continuing to suffer from economic losses associated with the flight delays all the time. Many popular data driven methods have been used to predict flight delay, including the random forest algorithm, artificial neural network ,logic probability and deep learning. The critical aim of these methods is to extract appropriate influential factors from real system to build reliable and high-efficient prediction model.

II. METHODOLOGY

In this project the machine learning flow can be explained in three stages:

1. Gathering of data.
2. Data Pre-processing.
3. Training & Testing of Model.

2.1 Gathering of Data

The three databases are selected for the best combination based features as per the previous experience of the data analysed. The BTS provides the database on the real time performance data. To explain the arrival delay of the flights, causes of the delay are reported in five categories: air carrier, extreme weather, national aviation system, late arriving aircraft and security. In our model, only the hourly departures and hourly arrivals are applied in corresponding airports.

2.2 Data Pre-processing

Instead of the precise delay time, we use the delay group in our research for random forest classification. The application of the delay group is suitable for the uncertainty of the actual air traffic control. All the non-positive delay

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 5, Issue 2, May 2021
Impact Factor: 4.819

groups are classified as on-time flights. The maximum of the delay group is 11 (delay time greater than 170 minutes) per day . The Following details shown in Table 1. Reference table of delay groups Delay groups Delayed time (min) -2 (-inf, -15) -1 [-15,0) 0 [0,15) 1 [15,30) 2 [30,45) 3 [45,60) 4 [60,75) 5 [75,90) 6 [90,105) 7 [105,120) 8 [120,135) 9 [135,150) 10 [150,165) 11 [165,180) 12 [180,+inf) The following delay-related data are extracted from the BTS database. • Day of month • Day of week • Scheduled departure time • Scheduled arrival time • Scheduled elapsed time

2.3 Training & Testing the Model:

- DATE_MONTH_YEAR: date, month, year by user.
- FLYING_AIRLINE : flight selection by user.
- FLYING_FROM & FLYING_TO: departure destination & arrival destination.

2.4 Data Source

We have used the following data sets:

1. Statistical Computing & Statistical Graphics .This data is taken from the Research and Technology Administration (RITA) database and structured for our use.
2. For weather data we have: From NOAA the weather observation data has been fetched at hourly land based weather format. Daily & hourly data of various airports is included in the source.

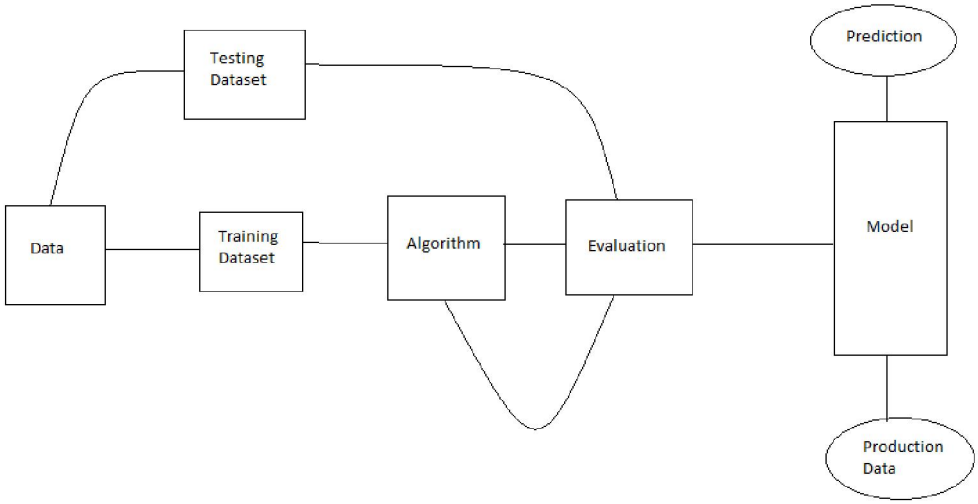
2.5 Software & Packages Used

A. Pycharm

Python Packages

1. Pandas : It is an open source python package that is most widely used for data analysis.
2. Numpy & Scipy : These are multidimensional arrays and general mathematical algorithms, etc.
3. Sci-kit Learn : It is a free machine learning library for python. It has features like K-neighbours and also scientific libraries like Numpy & Scipy (which we have used here).
4. Matplotlib : It is a cross - platform , data visualization and graphical plotting library for python and its numerical extension NumPy
5. CSV : "Comma-Separated Values" is a simple file format used to store tabular data like spreadsheets or databases. It is generally imported and exported from programs.

III. MODELLING AND ANALYSIS



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 5, Issue 2, May 2021
Impact Factor: 4.819

3.1 Model of Workflow

The workflow model explains the working flow of the project in a sequence manner.

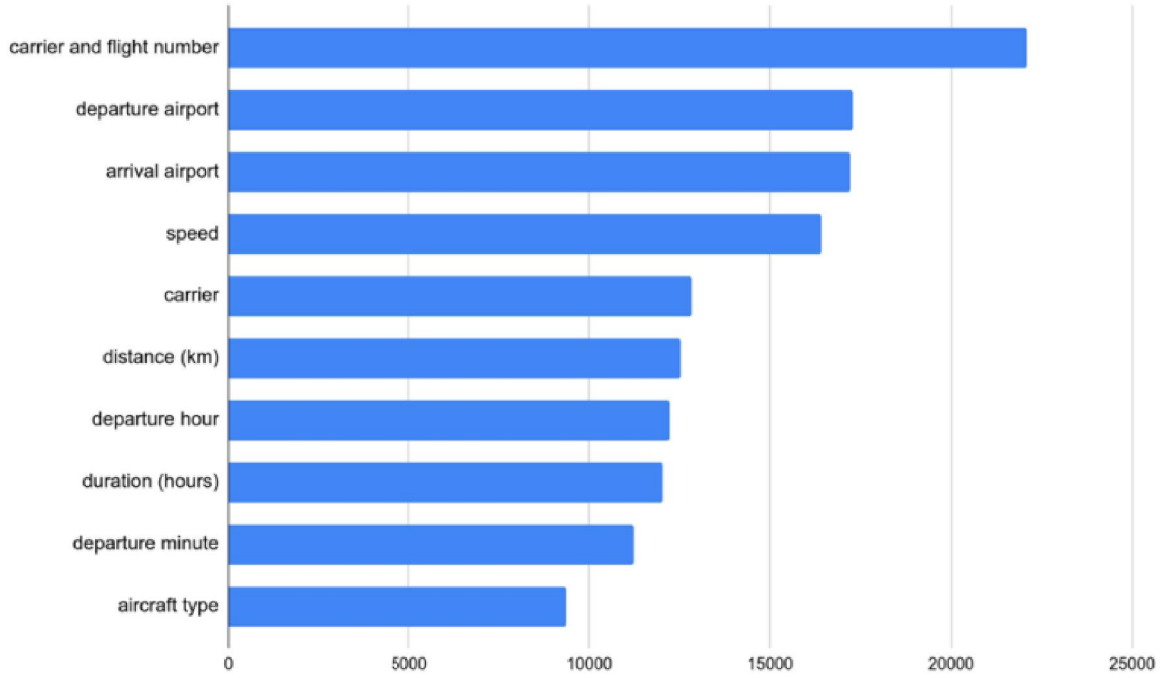
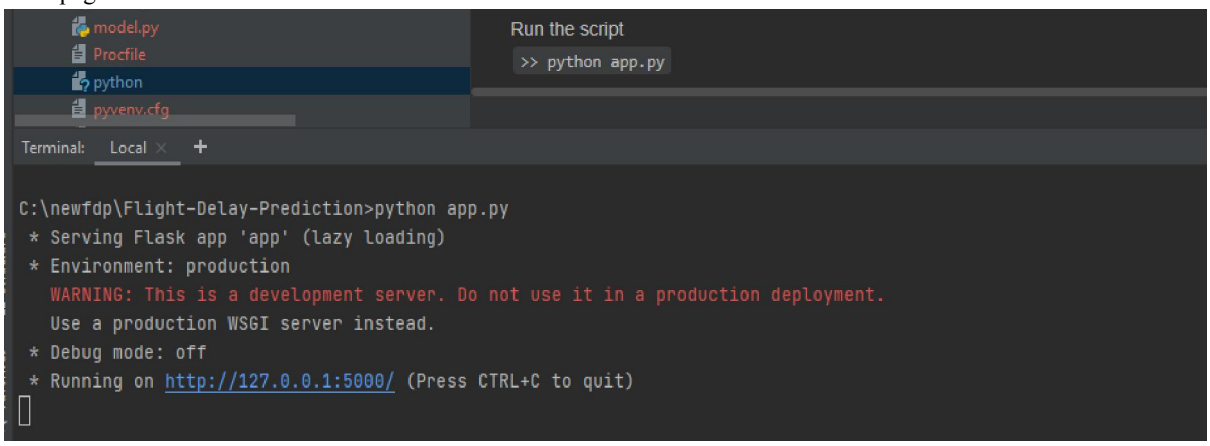


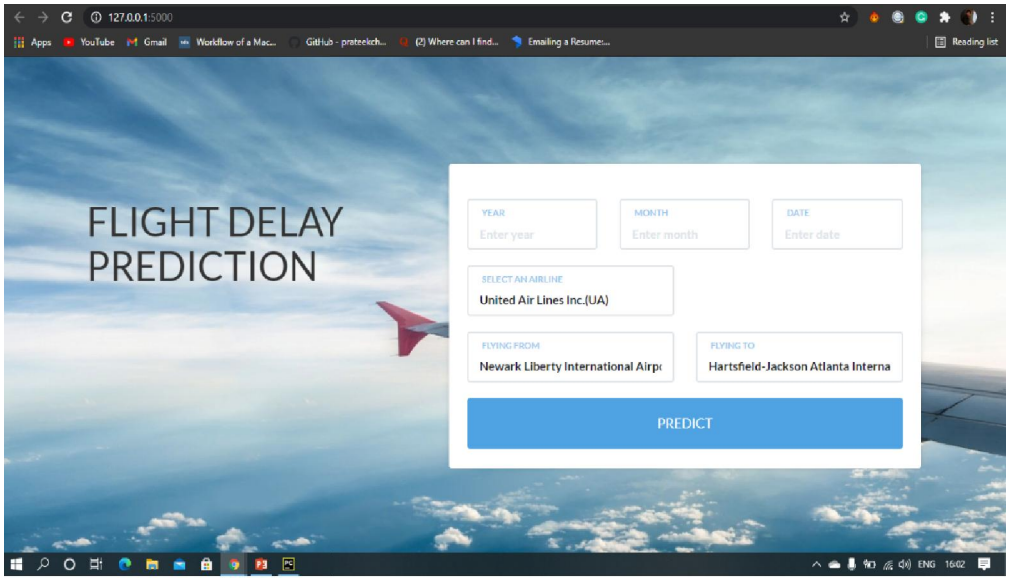
Figure: Features Of Flight Delay

IV. RESULTS

After successfully preparing data, filtering of data and training data which is to be predicted or shown by the ATC(Air Traffic Control) and successfully run this script using **>>python app.py** command.then we generate the sever page link.



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 5, Issue 2, May 2021
Impact Factor: 4.819



Thus, the Flight Delay Prediction API was created!

V. CONCLUSION

This prediction not only helps the aviation industry but also the passengers taking a flight. The flight delay also shows the negative reputation of the airline and also thus decreases the reliability of the airlines.The analysis carried here not only predicts delays based on the pervious available data, but also give statistical description of airlines, their rankings based on their on-time performance, the peak hours delays.

ACKNOWLEDGEMENTS

We would like to thank our project guide professor Shashant Jaykar of our Department of Electronics and Telecommunications Engineering for his guidance and support..

REFERENCES

[1] "Airline Delay Predictions using Supervised Machine Learning" by Pranali Chandarana and Prabakaran.N and Kannadasan.R, VIT University, Vellore.
[2] Chained Predictions of Flight Delay Using Machine Learning Jun Chen*, San Diego State University, San Diego, California, 92182 Meng Li† Purdue University, West Lafayette, Indiana, 47906.
[3] 1. Rebollo, J.J.; Balakrishnan, H. Characterization and prediction of air traffic delays. Transp. Res. Part C 2014, 44, 231–241. [CrossRef]
[4] MIT, Lexington, Massachusetts, Allan, S.S., S.G. Gaddy, and J.E. Evans, (2001) Delay Causality and Reduction at the New York City A.
[5] 1. E. Balaban, I. Roychoudhury, L. Spirkovska, S. Sankararaman, C. Kulkarni, and
[6] T. Arnon. Dynamic routing of aircraft in the presence of adverse weather using a POMDP framework. In 17th AIAA Aviation Technology, Integration, and Operations Conference, 2017.
[7] https://www.claimcompass.eu/blog/why-is-my-flight-delayed/
[8] http://www.github.com
[9] http://scikit-learn.org/
[10] https://www.scipy.org/
[11] https://developers.amadeus.com/blog/flight-delay-prediction-machine-learning
[12] Towards Data Science