Aviation Delay Prediction

Yash Tijil Department of Computer Science & Engineering Galgotias University, Greater Noida, India Yash.21SCSE1010898@galgotiasuniversity.edu.in	Shubhankar Bagchi Department of Computer Science & Engineering Galgotias University, Greater Noida, India shubhankar.21SCSE1011395@galgotiasuniversity.edu.in	Mr Rahul Anjana-Assistant Professor, Department of CSE Galgotias University, Greater Noida, India rahul.anjana@galgotiasuniversity. edu.in

***Abstract*— This project is dedicated to predicting aviation delays through the application of deep learning techniques, tackling a pressing issue that significantly impacts passengers, airlines, and regional economies. The goal is to develop a highly advanced prediction system that utilizes detailed flight data to forecast arrival delays with high accuracy. To achieve this, we will investigate a range of deep learning methods, applying them to an extensive dataset that includes various aspects of flight-related information. This project will involve a thorough exploration and comparative analysis of these techniques to assess their effectiveness in predicting flight delays. By rigorously evaluating the performance of each method, we aim to determine which techniques provide the most reliable and precise delay predictions. The ultimate objective is to equip airlines with accurate delay forecasts, which will enhance their flight planning processes and contribute to minimizing overall delays, thereby improving the efficiency of air travel.**

Keywords—Deep learning, Aviation Delay Prediction, Logistic Regression, Neural Networks, Machine Learning

I. Introduction

Aviation delays have become a significant concern for airlines, passengers, and airport operators. Despite the wealth of historical flight data available, many airlines and airport operators struggle to leverage this information to forecast delays accurately. Current predictive models are often insufficient, resulting in inadequate responses to operational disruptions and diminished customer trust. There is a clear need for advanced data-driven approaches to improve the accuracy of arrival delay predictions, ensuring that airlines can enhance operational efficiency and passenger satisfaction.

II. Literature Review

Many machine learning-based techniques in data mining like clustering, classification rules, and regression have been proposed to create and extract model predictions from historical data.

The configuration of the forecast model may indicate the lack of influence on the initiation of airport ground delay measures. For the purpose of weather-dependent analysis of dates and evaluation of performance, unsupervised data modelling approaches such as clustering are applied.

Akpinar and Karabacak used data mining categorization criteria, taking into account critical aspects such as airports, airlines, cargo, passengers, efficiency, and safety [1]. This study provides a comprehensive overview of data mining applications in civil aviation. Data mining may help with fuel price optimization, cargo optimization, passenger tracking, airport conditions, weather forecast, revenue per flight, cost per seat, catering and handling cost per seat etc.

Zhang and Nazeri used data mining to study the influence of weather on the performance of the National Airspace System (NAS) [2]. They use C5.0 decision tree learning technology and K-means clustering algorithm. The study found that weather patterns/conditions have an impact on NAS performance. They concluded that the revealed criteria pertain to flights that were blocked that can be used to predict performance on specific days based on weather.

Ha et al. used the CRISP-DM (CRoss Industry Standard Process for Data Mining) method to develop an experimental model and applied it to the exploration of big data [3]. They classified airports based on arrival delays and deduced that they could make recommendations on the best airports for arrival delays.

Sridhar and Mukherjee utilized two models: decision tree and logistic regression to predict when a Ground Delay Program will take place relying on traffic demand and climatic circumstances [4]. These algorithms are used to estimate the GDP of two major US airports. Evaluation of the models is done using data from weather variables such as wind, convection, precipitation, cloud height and visibility, as well as arrival requirements based on flight time. The logistic regression method calculates the probability of GDP occurring in that period, while the decision tree classifies hours as GDP or non-GDP.

Natarajan et al. utilized logistic regression and decision trees (random forest) using the same method to estimate delays [5]. They also evaluated the projected arrival time and delays for both models and determined that the decision tree method was more successful.

Tu et al. used probabilistic models to demonstrate that when delays are smaller than 2 hours, the chance of delay may be predicted [6]. Mueller et al. evaluated Normal and Poisson distributions and concluded that although arrival delays can be fitted to a normal distribution, the Poisson distribution is appropriate for measuring departure delays [7].

The outcomes of neural networks, decision trees, and Naïve Bayes were also assessed in a study by Lu et al. where decision trees performed the best, with a prediction confidence of 70% [8].

As per Chen et al., fuzzy SVM with a weighted margin is more accurate for predicting flight delays than a standalone SVM [9].

Delay predictions have also been made using ST-Random Forest and CNN-LSTM deep learning frameworks with 92.39% accuracy [10, 11].

III. Methodology

The BTS only classifies a flight as delayed when it arrives fifteen minutes or more later than expected. According to this study, a flight is considered "Delayed" if it is more than fifteen minutes behind schedule. Entries for cancelled or diverted flights have been excluded to enhance accuracy and results. The information from the "Departure Delay" column has been used to create a new column named "Delayed." The 'Delayed' column contains two values: 0 and 1, which represent the flight status. Flights that departed more than 15 minutes after the scheduled departure time are denoted by the numeral 1, while those that took off on time are represented by 0. Next, we used the 'Delayed' column to determine the f1-score, accuracy, recall, and support.

In this section, we provide an overview of the suggested framework. The first and most crucial stage in creating a model is gathering data. When obtaining records, it is important to consider important factors including the dataset's authenticity, correctness, and legality. Data gathering was followed by data pre-processing, which included formatting and cleaning the gathered data. Data cleaning, which included eliminating null values and unnecessary information from tuples, was also done at this phase.

Logistic regression was selected as the baseline model because of its interpretability and effectiveness in binary classification. However, it doesn't perform well with complex relationships. Unlike logistic regression, a neural network was selected as it is better suited for capturing intricate, nonlinear patterns. After the dataset was analysed, pertinent features were extracted to prepare for the testing and training stages. A thorough training and testing process was applied to each algorithm. Next, the two algorithms' f-score, recall, accuracy, and precision scores were contrasted.

IV. Methods

A. Logistic Regression

It is a statistical method used to calculate the likelihood of a binary result. Despite its name, logistic regression is employed for classification tasks rather than regression tasks. By estimating probabilities, it illustrates the relationship between one or more independent variables and a binary dependent variable and falls within the general linear model (GLM) group.

B. Neural Networks

The human brain serves as the inspiration for neural networks, which are machine learning models. They can comprehend intricate patterns because they are made up of interconnected nodes arranged in layers. Deep learning uses multi-layered neural networks to extract complex properties from input, allowing for applications such as autonomous systems, natural language processing, and picture and audio recognition.

V. Data Collection and Preprocessing

A. Dataset

The Bureau of Transportation Statistics (BTS) was used as the data source for acquiring the following comprehensive flight information data for the year 2015:

'Month,' 'Day,' 'Day of Week,' 'Year,' 'Flight Number,' 'Airline,' 'Tail Number,' 'Origin Airport,' 'Destination Airport,' 'Scheduled Departure,' 'Departure Time,' 'Departure Delay,' 'Taxi Out,' 'Wheels Off,' 'Scheduled Time,' 'Elapsed Time,' 'Air Time,' 'Distance,' 'Wheels On,' 'Taxi In,' 'Scheduled Arrival,' 'Arrival Time,' 'Arrival Delay,' 'Diverted,' 'Cancelled,' 'Cancellation Reason,' 'Air System Delay,' 'Airline Delay,' 'Aircraft Delay,' 'Aircraft Delay,' and 'Weather Delay.'

B. Data Preprocessing

Prior to model training, it is essential to preprocess the data to prevent potential errors later on. In this study, various Python programming techniques and libraries were used to preprocess the data:

1. Columns with reasons for delay had some null values. They were replaced with 0.

2. Handling missing values: The 'Arrival Delay' and 'Departure Delay' columns had some missing values. These rows were dropped.

3. Eliminating superfluous traits: While the majority of the traits are pertinent, some were discarded since they weren't necessary.

For instance, as cancelled flights are not regarded as delayed in this work, the "Cancelled" column and its corresponding "Cancellation Reason" column were eliminated from our dataset. For the same reason, the "Diverted Flights" column was removed. Since the models used do not support categorical variables, categorical columns like "Airline," "Tail Number," "Origin Airport," and "Destination Airport" were eliminated

VI. Results

A. Confusion Matrix

After training, the model was tested and obtained an approximate 95% Accuracy Score. Within the confusion matrix, the diagonal components match the total number of correctly classified tuples [1]. Here is a summary of the Confusion Matrix:

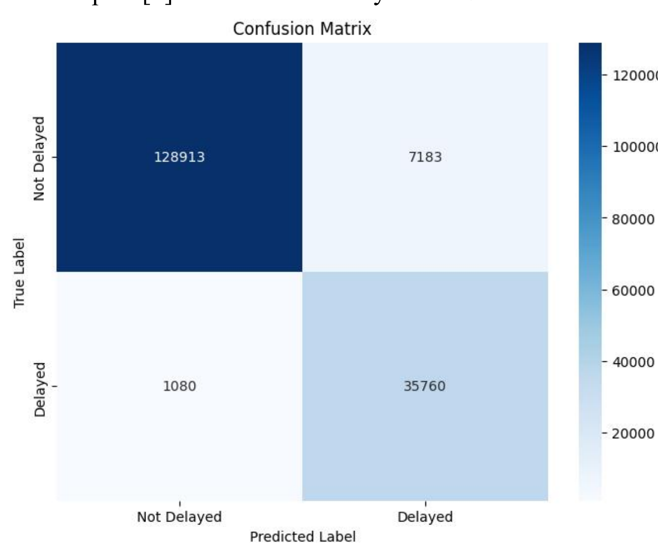

Fig. 1. Confusion Matrix generated for Neural Network model

B. Comparative Analysis of Algorithms

The models were applied to the dataset with the main purpose of identifying and classifying flights with delays of more than 15 minutes. The 'Delayed' attribute within the dataset served as the target variable for assessing and classifying flight delays.

The model's functionality was assessed using the following criteria:

- Validation accuracy characterizes the model's ability to accurately predict samples within a set of given values.
- Recall is determined by the ratio of the correct return of critical events to all related events.

  True positive / (true positive + false negative) equals recall.
- Precision is determined by the ratio of accurately identified positive instances to the total positive observations.

  True positive / (true positive + false positive) equals precision.
- F1-Score or weighted average of recall and precision, is a statistically defined accuracy metric.

TABLE I. RESULTS OF ACCURACY, F1-SCORE, RECALL, AND PRECISION

Algorithm	Precision		Recall		F1-Score		Accuracy
	0	*1*	*0*	*1*	*0*	*1*	
Logistic Regression	0.95	0.66	0.89	0.82	0.92	0.73	0.87
Neural Network	0.99	0.83	0.95	0.97	0.97	0.90	0.95

- Therefore, achieving 95% accuracy with the Neural Network learning model when provided with any set of attributes (feature values) distinctly demonstrates its efficiency for predicting aircraft delays.

VII. CONCLUSION

Our study demonstrated that machine learning techniques can be used to anticipate airplane delays with a high degree of accuracy. The aforementioned classification and analysis attempts to extensively investigate factors that affect delays, such as "weather delay," "airline delay," and "security delay," in addition to assessing delays for various human requirements. With 95% accuracy, the neural network model appropriately divides delays into two categories. Airports, airlines, and passengers will all benefit from its ability to accurately forecast aviation delays.

Therefore, the study of flight delays presented in this paper is grounded entirely in scientific parameters, underscoring its crucial significance in the aviation industry.

A. Future Scope

While this study looked at a number of factors that can affect flight delays, future studies can concentrate on examining the effects of additional factors like seasonality and temporal patterns. Moreover, sequential patterns in the data can be captured by using deep learning models, like recurrent neural networks. Additionally, the scope of this study was limited to a specific dataset; future research should include international flight data to generalize the model.

REFERENCES

[1] Akpinar, M.T. and Karabacak, M.E. (2017). Data mining applications in civil aviation sector: State-of-art review. In *CEUR Workshop Proc* (Vol. 1852, pp. 18-25).

[2] Nazeri, Z. and Zhang, J. (2017). Mining Aviation Data to Understand Impacts of Severe Weather. In *Proceedings of the International Conference on Information Technology: Coding and Computing* (ITCC.02) pp. 518-523.

[3] Ha. ,. J. N. a. H. P. S. Man. (2015) "Analysis of Air-Moving on Schedule Big Data based on CrispDm Methodology," ARPN Journal of Engineering and Applied Sciences, pp. 2088-2091.

[4] Mukherjee, A., Grabbe, S. R., and Sridhar, B. (2014). Predicting Ground Delay Program at an airport based on meteorological conditions. In 14th AIAA Aviation Technology, Integration, and Operations Conference (pp. 2713-2718).

[5] Natarajan, V., Meenakshisundaram, S., Balasubramanian, G. and Sinha, S. (2018) "A Novel Approach: Airline Delay Prediction Using Machine Learning," 2018 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA pp. 1081-1086, doi: 10.1109/CSCI46756.2018.00210

[6] Tu, Y., Ball, M.O, and Jank, W.S. (2008) Estimating flight departure delay distributions—a statistical approach with long-term trend and short-term pattern", Journal of the American Statistical Association, 103, pp. 112-125.

[7] Mueller, E.R. and Chatterji, G.B. (2002) Analysis of aircraft arrival and departure delay characteristics", In AIAA's Aircraft Technology, Integration, and Operations (ATIO) 2002 Technical Forum, Los Angeles, California, USA

[8] Lu, Z., Wang, J., and Zheng, G. (2008) A new method to alarm large scale of flights delay based on machine learning, in knowledge acquisition and modelling", KAM '08. International Symposium on, pp. 589-592.

[9] Chen, H., Wang, J., and Yan, X. (2008) A fuzzy support vector machine with weighted margin for flight delay early warning", In *2008 Fifth International Conference on Fuzzy Systems and Knowledge Discovery* (Vol. 3, pp. 331-335). IEEE

[10] Li, Q. and Jing, R. (2022) Flight delay prediction from spatial and temporal perspective, Expert Systems with Applications, Volume 205, 117662.

[11] Li, Q., Guan, X., and Liu, J. (2023) A CNN-LSTM framework for flight delay prediction, Expert Systems with Applications. Volume 227, 120287.