Journal of Advances in Developmental Research (IJAIDR)
E-ISSN: 0976-4844 ● Website: www.ijaidr.com ● Email: editor@ijaidr.com

A Novel Approach for Predicting Flight Delay Based on Historical Data Using Machine Learning Techniques

Ms. P. Anuja[1], M. Divya Bharathy[2]

[1,2]B.SC Computer Science
School of Science and Allied Health Science
Jeppiaar University, Tamil Nadu, Chennai-600119
[1]anujamail2508@gmail.com
[2]divyabharathym635@gmail.com

Abstract
One of the biggest problems facing the aviation sector is flight delays, which result in both passenger discomfort and airline financial losses. Precise forecasting of flight delays can enhance operational efficiency and assist passengers and airlines make better decisions. Conventional delay prediction techniques rely on statistical analysis and rule-based algorithms, which frequently fall short of capturing intricate interactions between affecting elements including weather, airline timetables, and traffic congestion. Using flight data from the past, this research suggests a machine learning-based method for forecasting flight delays. The suggested methods use machine learning algorithms and feature analysis to find patterns in delays and forecast potential delays. Data gathering and machine learning-based prediction are two of the system's parts. According to experimental investigation, machine learning models can efficiently prediction accuracy and support improved management of aviation operations.

Keywords: Aviation analytics, machine learning, historical flight data, data preprocessing, feature extraction, and flight delay prediction

INTRODUCTION

Current Studies on Predicting Flight Delays Statistical Models & Regression Techniques
Machine Learning-Based Methods (Random Forest, SVM, Decision Tree, etc.)
Restrictions in Current Systems
Inadequate Prediction Precision
• Insufficient Real-Time Updates
• Inadequate Analysis of Features
The need for an enhanced machine learning-based prediction system has been identified as a research gap.
The suggested flight delay prediction approach

Benefits
1. Improved Predictive Precision

2. Instantaneous Updates
3. A Better Experience for Passengers
4. Improved Airline Management
5. Adaptable and Scalable
6. Making Proactive Decisions
7. Insight Driven by Data

LITERATURE SURVEY

Current System:

At the moment, airline-reported delays, historical averages, and predetermined rules are used to forecast aircraft delays. Numerous airports and airlines rely on statistical analysis and basic regression models, which are unable to capture complex interactions between several variables. These systems are not flexible in real time and frequently offer

erroneous forecasts as a result of changing weather and flight traffic. Furthermore, travelers have little access to accurate pre-flight delay forecasts, which causes confusion and annoyance.

the suggested method. Additionally, the system demonstrated adaptability to various operational scenarios and flight conditions.

The findings show that combining machine learning methods with historical data analysis can greatly increase the precision of flight delay forecasts. By identifying the main causes of delays, the suggested system offers airline management helpful insights. Airlines can use this data to improve operational efficiency, minimize passenger discomfort, and optimize travel schedules.

METHODOLOGY

Data Preprocessing

Before training the machine learning models, the dataset undergoes several preprocessing steps to improve prediction accuracy. Missing values are identified and handled using appropriate techniques such as mean substitution and removal of incomplete records. Categorical attributes including airline name, source airport, and destination airport are converted into numerical representations using label encoding.

Feature scaling is also applied to normalize numerical values such as departure delay, arrival delay, and flight duration. Duplicate entries and inconsistent records are removed to maintain data quality. These preprocessing techniques help improve model efficiency and reduce prediction errors.

Advantages of Proposed System

The proposed flight delay prediction system provides several advantages for both passengers and airline companies. The system helps passengers plan their travel schedules effectively by predicting possible delays in advance. Airlines can use the prediction results to improve operational efficiency and reduce passenger dissatisfaction.

Machine learning algorithms improve prediction accuracy by analyzing historical flight data and identifying hidden patterns. The system also reduces manual analysis and supports real-time decision making. In addition, accurate delay prediction helps airport authorities manage air traffic more efficiently.

The proposed system can also be integrated with real-time data sources to improve prediction performance. Live weather updates, airport traffic information, and aircraft maintenance reports can be

continuously analyzed to provide more accurate delay predictions. This helps airlines and passengers receive updated information instantly and take necessary actions in advance. The integration of real-time analytics makes the system more reliable and practical for modern air transportation management.

RESULT AND DISCUSSION

A historical flight dataset that included details including airline schedules, departure and arrival times, weather, and prior delay records was used to assess the suggested flight delay prediction system. To assess how well the machine learning models performed, the dataset was split into training and testing sets. To examine trends associated with aircraft delays, an incorporating real-time changes in between number of machine learning methods were used, including Random Forest, Support Vector Machine (SVM), Gradient Boosting, and CatBoost.

The outcomes of the experiment show that the suggested system can recognize intricate linkages interplay several elements that impact flight delays. Because they can successfully handle huge datasets and categorical features, boosting-based models like CatBoost and Gradient Boosting demonstrated superior performance among the studied algorithms. These models provide higher prediction accuracy than the conventional regression-based methods seen in current systems.

Important influencing factors like weather, airline operating performance, airport congestion, and past delay trends were all analyzed by the prediction algorithm. The analysis showed that the likelihood of flight delays is significantly influenced by weather and past delay trends. Based on these characteristics, the program was able to accurately categorize flights into delayed and non-delayed groups.

Overall, the experimental assessment demonstrates that the suggested machine learning-based architecture offers a trustworthy method for forecasting flight delays. The technology can assist travelers airport officials, and airlines in making wise choices and increasing the effectiveness of the aviation system as a whole

RESULTS

By utilizing machine learning to provide precise and up-to-date delay estimates, the suggested flight delay prediction system transforms air travel planning. The system gives airlines, airports, and travelers useful insights by evaluating past flight data and

airport officials, and airlines in making wise choices and increasing the effectiveness of the aviation system as a whole

incorporating real-time changes. This strategy saves operational costs, decreases passenger discomfort, and increases scheduling efficiency.

During the training phase, cross-validation techniques were used to assess the system's dependability. The resilience of the machine learning models was proved by their consistent performance across several dataset subsets

REFERENCES

1. E,Simic and M. Begovic, “Airport delay prediction using machine learning regression models as a tool for decision making process”,Proc.45th Jubliee Int.
   Conf. Inf. Reuse Integr . Data Sci. (IRI),pp. 49-54, Aug. 2022.
   Conv. Inf. Commun. Electron. Technol. (MIPRO), pp. 841-846, May 2022.

2. S.S.B.T. Lincy, H. Al Ali, A.A.A. M. Majid, O .A.A.A. Alhammadi, A.M.Y.M. Aljassmy and Z.Mukandavire,"Analysis of flight delay data using different machine learning algorithims",Proc.New Trends Civil Aviation (NTCA),pp. 57-62, Oct.2022
3. D. Jadav, D . Patel, S. Thacker, A. Nair, R. Gupta, N.K.Jadav, et al.,"EmReSys:AI based efficient employee ranking and recommendation system for organizations", Proc. Int. Conf. Comput . Commun . Intellu . Syst. (ICCCIS), pp. 440-445, Nov. 2022
4. R. Vane, "Flight delay analysis and possible enhancement with big data", Int.
5. A. Dand, "Airline delay prediction using machine learning algorithms",2020.
6. T. Wang and S.-C. Chen, "Multi-task local-global graph network for flight delay prediction", Proc. IEEE 23rd Int
7. Y . Wang,M. Z. Li, K. Gopalakrishnan and T . Liu, "Timescale of delay propagation in
8. A. Dand, "Airline delay prediction using machine learning algorithms",2020.
9. T. Wang and S.-C. Chen, "Multi-task local-global graph network for flight delay prediction", Proc. IEEE 23rd Int
10. Y . Wang,M. Z. Li, K. Gopalakrishnan and T . Liu, "Timescale of delay propagation in airport networks", Transp. Res. E Logistics Transp. Rev., vol. 161, May 2022.