

Deciphering Airline Performance Data

Shraddha Ramprakash Gupta
Indiana University

Radhika Ganesh
Indiana University

Subhadra Mishra
Indiana University

project-gupta37-rganesh-sumishra

Abstract

Reliability and timeliness are crucial to the aviation sector. In addition to causing annoyance to travelers, flight delays and cancellations cause airlines to suffer significant financial losses. In order to determine causes, develop predictive models, optimize scheduling, evaluate airline performance, and enhance the overall customer experience, this research analyzes a dataset of airline delays and cancellations. Airlines may improve their operating efficiency, guarantee safety and compliance, and help travelers make better decisions by utilizing this dataset. The report also discusses the potential future uses of such data for emerging technologies and industry advancements.

Keywords

Machine learning, data science, prediction, classification, exploratory data analysis, data visualization, airline delays, flight cancellations, operational efficiency, compliance, future applications.

1 Introduction

The aviation industry's relentless pursuit of efficiency and reliability is continuously challenged by flight delays and cancellations. These disruptions, caused by an array of factors such as weather, air traffic, and technical issues, not only inconvenience passengers but also impact airlines' operational and financial performance. Analyzing a comprehensive dataset of airline delays and cancellations is pivotal to understanding the root causes and consequences of these disruptions. This report focuses on the methods and approaches for analyzing such a dataset and explores the potential insights and applications that can arise from this analysis. By identifying the primary reasons behind delays and cancellations, airlines can refine their operations and optimize passenger satisfaction. Predictive modeling can enable proactive resource management and improved passenger notifications. Evaluating airline performance facilitates informed passenger choices, and the dataset can play a critical role in enhancing safety and regulatory compliance within the industry.

Previous work

The research papers that are cited below, collectively provide a comprehensive overview of flight delay prediction and management within the aviation industry. They explore various methodologies, including data mining, machine learning, and predictive modeling, to shed light

on the factors influencing aviation disruptions. The papers emphasize the significance of using advanced technologies to enhance flight delay prediction and offer insights into the practical implementation of these methodologies for improving aviation data analysis. Together, they contribute valuable perspectives and approaches to address the challenges of flight delays in the aviation sector.

2 Methods

To extract meaningful insights from the airline delay and cancellation dataset, a variety of analytical methods can be employed. We are planning to carry out the following methods and approaches:

Before any analysis, we will pre-process the dataset to ensure data quality and consistency. This includes handling missing values, data cleaning, encoding the categorical attributes into numerical values and transforming data into a suitable format for analysis.

Then Exploratory Data Analysis (EDA) will be performed to gain a holistic understanding of the dataset. Various visualization techniques can be applied to identify trends, patterns, and outliers in the data. It will help us understand the distribution of delays and cancellations, their frequency, and their relationship to different variables such as airline, airport, and time of day.

Further, we can analyze factors like weather-related delays, air traffic, technical issues in the data to find the primary reasons for flight delays and cancellations through descriptive statistics and visualizations.

Classification models like Naïve bayes Classifier, SVM, etc. will be employed to build predictive models that forecast potential delays or cancellations based on historical data. These models can help airlines proactively manage resources and notify passengers in advance, resulting in improved customer satisfaction. Further, performance of airlines can be assessed in terms of their punctuality and reliability. Passengers can use this information to make more informed choices when booking flights.

References

1. Sternberg, Alice Soares, Jorge Carvalho, Diego Ogasawara, Eduardo. (2017). *A Review on Flight Delay Prediction*.
2. Rebollo JJ, Balakrishnan H. *Characterization and prediction of air traffic delays*. Transportation Res Part C Emerg Technol. 2014;44:231–41.
3. C. Ariyawansa and A. Aponso. *Review on state of art data mining and machine learning techniques for intelligent Airport systems*. In Proceedings of 2016 International Conference on Information Management, ICIM2016, pages 134–138, 2016.
4. Kalliguddi, Anish Leboulluec, Aera. (2017). *Predictive Modeling of Aircraft Flight Delay*. Universal Journal of Management. 5. 485-491. 10.13189/ujm.2017.051003.
5. Somani, Saksham. (2021). *An Approach of Applying Machine Learning Model in Flight Delay Prediction- A Comparative Analysis*. Revista Gestão Inovação e Tecnologias. 11. 1233-1244. 10.47059/revistageintec.v11i3.2006.
6. Sahu, Prof Desale, Kunal Patil, Ashish Laishetty, Prithvi Patil, Bhuvaneshwar. (2023). *Flight Delay Prediction System in Machine Learning using Support Vector Machine Algorithm*. IJARCCCE. 12. 10.17148/IJARCCCE.2023.125188.

7. Yiu, Cho Yin Ng, Kam K.H. Kwok, Kin Lee, Wing Mo, Ho. (2021). *Flight delay predictions and the study of its causal factors using machine learning algorithms*. 10.1109/ICCASIT53235.2021.9633571.
8. <https://www.kaggle.com/datasets/yuanyuwendymu/airline-delay-and-cancellation-data-2009-2018/data>
9. <https://www.transtats.bts.gov/Homepage.asp>