Literature Review

Flight Delay Analysis

Literature Review

1. Introduction:

Flight delays are a significant problem within the aviation industry, affecting passenger satisfaction, operating efficiency, and airline profitability. Delays can be brought on by a host of factors, including weather, air traffic density, technical faults, and inefficiencies in scheduling by airlines. As air travel demand increases, minimizing flight delays has become imperative for airlines and regulators. This review of literature talks about earlier studies on flight delay analysis, focusing on conventional methods, machine learning techniques, key factors affecting them, and challenges in predictive modelling.

2. Theoretical Background:

There are various theories and frameworks that inform flight delay analysis. Queuing theory has been used widely to model air traffic congestion and schedule optimization (Bazargan, 2016). Stochastic process theory helps in predicting the likelihood of delays based on historical data (Hansen, 2002). Operations research methods have also been applied to optimize resource use at airports and reduce turnaround time (Liu et al., 2021).

3. Traditional Approaches to Flight Delay Analysis:

Early methods for analysing flight delays relied on statistical models and time-series analysis.

- **Regression Models:** Linear regression and logistic regression have been widely used to estimate delay probabilities based on historical data (Ball et al., 2010). These models identify significant factors influencing delays, such as peak-hour congestion and weather conditions.
- **Time-Series Analysis:** ARIMA models have been applied to forecast delay trends over time. Studies have shown that ARIMA models can effectively capture seasonal variations in flight delays (Garrow, 2012).
- **Survival Analysis:** Kaplan-Meier estimators and Cox proportional hazards models have been used to analyze the duration of delays and the likelihood of extended disruptions (Tu et al., 2008).

4. **Machine Learning-Based Flight Delay Prediction:**
   With the advent of big data, machine learning models have enhanced the ability to predict flight delays with greater accuracy.
   - **Decision Trees & Random Forest:** Ensemble models such as Random Forest have been employed to classify flight delays based on multiple attributes, including departure time, airport congestion, and weather conditions (Choi et al., 2020).
   - **Gradient Boosting Techniques (XGBoost):** Recent studies show that XGBoost performs well in handling large and imbalanced flight delay datasets, providing superior predictive accuracy compared to traditional models (Chen & Guestrin, 2016).
   - **Deep Learning Models:** Neural networks, particularly LSTMs, have been explored to analyze sequential flight delay data. These models capture temporal dependencies, making them effective in delay forecasting (Zhang et al., 2021).
5. **Key Factors Influencing Flight Delays:**

A number of studies have pinpointed main causes of flight delays:

• **Weather Conditions:** Inclement weather conditions such as storms, fog, and heavy snowfall significantly impact flight schedules (Klein et al., 2019).

• **Air Traffic Congestion:** High-density airports experience frequent delays due to runway congestion and air traffic control limitations (Pyrgiotis et al., 2013).

• **Aircraft Turnaround Time:** Delays in baggage handling, security clearance, and refuelling are responsible for departure delays (Wu & Caves, 2004).

• **Carrier-Specific Factors:** Certain airlines experience more delays due to inefficiencies in their scheduling and operations (Fleurquin et al., 2013).

6. **Challenges in Flight Delay Prediction:**
   Despite advancements in predictive modelling, several challenges remain:
   - **Data Quality & Availability**: Incomplete or biased datasets hinder accurate delay prediction (Hansen & Huang, 2019).
   - **Generalization Across Airports:** Models trained on specific airports may not generalize well to other locations due to varying operational conditions (Balakrishna et al., 2010).

- **External Unpredictable Factors**: Unexpected events such as security threats, strikes, and global pandemics disrupt flight schedules beyond model predictions (Sun et al., 2022).

7. Contributions of This Project:

This capstone project aims to build a robust flight delay analysis model that integrates predictive accuracy with practical usability:

- **Explainable AI:** Unlike black-box models, the project employs SHAP values to provide interpretable insights into delay causes.
- **Data-Driven Decision Making:** By leveraging historical flight data, the model helps airlines and regulators make informed scheduling improvements.
- **Operational Recommendations:** The study provides actionable insights for airports and carriers to mitigate delays and optimize operations.

8. Conclusion:

Recent research focuses on the performance of machine learning models, particularly ensemble techniques like XGBoost and Random Forest, in predicting flight delays. Even then, there are concerns related to data quality, generalization, and external interference. This project aims at bridging these gaps by combining explainable AI techniques and real-world decision-making tools in order to improve flight scheduling and passenger experience.

9. References:

Ball, M., et al. (2010). "Statistical Methods for Predicting Flight Delays." Transportation Research Part C.

Bazargan, M. (2016). "Airline Operations and Scheduling." Routledge.

Choi, Y., et al. (2020). "Machine Learning Approaches to Flight Delay Prediction." IEEE Transactions on Intelligent Transportation Systems.

Fleurquin, P., et al. (2013). "Systemic Delay Propagation in the US Air Transportation Network." Transportation Research Part C.

Garrow, L. (2012). "Discrete Choice Models of Air Travel Demand." Journal of Air Transport Management.

Hansen, M., & Huang, Y. (2019). "Flight Delay Causality in the US Airline Industry." Transportation Research Part E.

Liu, C., et al. (2021). "Optimization of Airline Scheduling Using Operations Research." Journal of Transportation Science.

Pyrgiotis, N., et al. (2013). "Modelling Delay Propagation in an Airline Network." Transportation Research Part C.

Sun, X., et al. (2022). "Impact of Global Disruptions on Flight Operations." Journal of Aviation Management.

Zhang, H., et al. (2021). "Deep Learning for Time-Series Flight Delay Prediction." IEEE Access.