

Enhancing Flight Operations with AI-Based Delay Forecasting

Mohammed Hamed Alzahrani

Master's Student, College of Informatics, Department of Artificial Intelligence

Midocean University

Supervisor: Dr Khaled Eskaf

1. Project Overview

Flight delays cause operational inefficiencies, financial losses, and passenger dissatisfaction. Traditional forecasting methods fail to account for real-time factors like weather and airport congestion, leading to inaccurate predictions. This project leverages AI and machine learning to analyze large-scale aviation data and accurately predict flight delays by identifying key factors that help airports and airlines optimize scheduling, reduce disruptions, and improve passenger experience, making air travel more efficient and reliable.

Problem: Frequent Flight Delays at Major Airports

Many major airports face frequent flight delays due to unpredictable factors such as weather, air traffic congestion, and operational inefficiencies. Airlines struggle to predict delays accurately, leading to last-minute schedule changes, frustrated passengers, and financial losses. Traditional delay forecasting methods rely on fixed schedules and historical averages, which fail to adapt to changing conditions.

Solution: AI-Based Predictive Analytics for Delays

A leading airline implemented an AI-driven delay prediction system using machine learning models like XGBoost and Prophet. By analyzing past flight data, weather conditions, and airport congestion patterns, the system could predict delays with 80% accuracy. This allowed the airline to:

1. Proactively adjust flight schedules to reduce waiting times.
2. Notify passengers in advance about expected delays.
3. Optimize airport resource allocation for better efficiency.

As a result, the airline reduced flight delays by 20%, improving customer satisfaction and saving operational costs. This real-world case study demonstrates how AI can transform flight scheduling and reduce delays, aligning with the goals of this project.

2. Project Goal

- Develop an AI-based model to predict flight delays using historical aviation data.
- Improve efficiency in flight scheduling by providing accurate delay forecasts.
- Create a data-driven tool to help airlines and airports optimize operations and reduce disruptions.

3. Approach

Data Collection

- **Sources:** Flight delay records and weather data from Kaggle.
- **Stored in Snowflake** for easy access and processing.

Data Processing

- **Remove missing values** and incorrect entries.
- **Standardize formats** (date-time, numerical values).
- **Feature selection:** Identify key factors affecting delays.

Model/Method Selection

- **Prophet (for delay trends):** Predicts delay patterns based on historical data.
- **XGBoost (for delay classification):** Predicts whether a flight will be delayed based on weather, congestion, and scheduling factors.

Testing and Evaluation

- **Mean Absolute Error (MAE) & RMSE:** Measure prediction accuracy for Prophet.
- **F1-score/Accuracy:** Evaluate classification performance for XGBoost.
- **Cross-validation & backtesting.**

4. Project Steps and Timeline

- **Data Collection and Preparation:** 10 weeks
- **Model Building and Training:** Next Term
- **Testing and Evaluation:** Next Term
- **Final Adjustments and Reporting:** Next Term

6. Expected Results

- Improved accuracy in predicting flight delays using AI and machine learning models.
- New insights into key factors causing delays, including weather conditions and airport congestion.
- A working AI-based prediction model that helps airlines and airports optimize flight scheduling.
- Scalable and adaptable forecasting system that can be extended to other regions or integrated with real-time data in the future.

7. Resources Needed

- Datasets from Kaggle and other relevant sources.
- Snowflake for data storage, querying, and management.
- Python Programming as the main language for model development.
- Machine Learning Libraries: TensorFlow, Keras, Scikit-learn, XGBoost, Prophet, and Statsmodels.
- Data Processing Tools: Pandas, NumPy, and Matplotlib for data cleaning and visualization.
- Historical Flight Data: Flight schedules, departure/arrival times, and delay durations.
- Weather Data: Temperature, precipitation, wind speed (if available).
- Public/Open Datasets (if applicable) – FAA or ICAO.

8. References

My personal experience as an Air Traffic Controller. Studies from Since Direct, ICAO (International Civil Aviation Organization) Docs, FAA (Federal Aviation Administration) Docs. Kaggle website for datasets.