



Certificate in Artificial Intelligence and Machine Learning

CAIMLOps

April 2024

A proposal submitted to Refactory Uganda for the Flight Delay Prediction Challenge

Contents

1 Introduction 1

1.1 Problem Description 1

1.2 Main Objective 1

1.2.1 Specific Objective 1

2 Proposed Approach 1

3 Work Schedule 2

4 Expected Outcome 2

5 Conclusion 2

A proposed models II

A.1 LSTM II

A.2 XGBoost II

List of Figures

1 CNN-LSTM Architecture II

2 XGBoost Architecture II

1 Introduction

CAIMLOps, a team of students at refractory in the course of AI/ML aims to predict the estimated duration of flight delays per flight in Tunisia.

1.1 Problem Description

Flight delays have significant repercussions on various stakeholders within the air travel ecosystem [1]. Not only do they inconvenience passengers and disrupt schedules, but they also lead to decreased efficiency, increased capital costs, reallocation of resources, and additional crew expenses. Furthermore, a poor record of flight delays can adversely impact an airline's reputation and passenger demand. To address these challenges, we propose the development of a flight delay predictive model using machine learning techniques. By accurately predicting the duration of flight delays, we aim to empower airlines and other industry players to implement proactive measures to mitigate the impact of delays, minimize losses, and enhance overall operational efficiency [2].

1.2 Main Objective

Develop a Flight Delay Prediction Model: Create a machine learning model capable of accurately predicting the estimated duration of flight delays for individual flights.

1.2.1 Specific Objective

1. Mitigate the Impact of Flight Delays: Empower airlines and industry stakeholders to proactively manage and minimize the adverse effects of flight delays on passengers, operational efficiency, and resource utilization.
2. Enhance Operational Efficiency: Improve the overall efficiency of airline operations by providing timely insights into potential flight delays and enabling proactive decision-making in crew scheduling, aircraft allocation, and route planning.
3. Optimize Resource Allocation: Enable efficient allocation of resources, including flight crews, aircraft, and ground personnel, to mitigate the impact of delays and minimize associated costs.

2 Proposed Approach

1 Data Collection and Preprocessing: Cleanse and preprocess the data to handle missing values, outliers, and inconsistencies. Integrate external datasets such as weather forecasts, airport congestion data, and flight schedules to enrich the feature set and improve model performance.

2 Feature Engineering: Extract relevant features from the dataset, including temporal features (time of day, day of the week), geographical features (airport location, airspace congestion), and categorical features (airline, aircraft type) [3]. Engineer new features that capture the interactions and relationships between different variables to enhance the predictive power of the model.

3 Model Selection and Training: Explore a variety of machine learning algorithms suitable for regression tasks, such as linear regression, decision trees, random forests, gradient boosting, and neural networks. Conduct thorough experimentation to identify the most effective combination of features and algorithms for predicting flight delays accurately. Employ techniques like cross-validation and hyper-parameter tuning to optimize model performance and generalization ability.

4 Evaluation and Validation: Evaluate the trained models using appropriate performance metrics such as mean absolute error, mean squared error, and R-squared.

5 Validate the models on unseen test data to assess their robustness and ability to generalize to new instances. Conduct sensitivity analysis to understand the relative importance of different features and identify potential areas for improvement.

6 Deployment and Integration: Deploy the finalized predictive model into production environments, allowing real-time prediction of flight delays. Integrate the model with existing airline operations systems and decision support tools to facilitate proactive decision-making and resource allocation. Continuously monitor model performance and update it periodically to adapt to evolving patterns and dynamics in air travel.

3 Work Schedule

Action plan	Week 1	Week 2	Week 3	Week 4
Team Registration				
Proposal writing and presentation				
Model training and Testing				
Report writing and presentation				

4 Expected Outcome

This study is intended to correlate all the problem, scope and method for getting most accurate result. Although some features might seem to be the causes of flight delays [4], the study also focuses on digging up other micro causes, this will lead to:

- Reduction in flight delays within the first year of implementation.
- Improved customer satisfaction ratings for airlines and airports.

5 Conclusion

By leveraging data-driven solutions and fostering collaboration across the aviation industry, we can significantly reduce the incidence and impact of flight delays. This proposal outlines a systematic approach to enhance flight delay management, ultimately improving the travel experience for passengers and optimizing operations for airlines and airports.

References

- [1] A. Sternberg, J. Soares, D. Carvalho, E. Ogasawara, A review on flight delay prediction, arXiv preprint arXiv:1703.06118 (2017).
- [2] M. Alfarhood, R. Alotaibi, B. Abdulrahim, A. Einieh, M. Almousa, A. Alkhanifer, et al., Predicting flight delays with machine learning: A case study from saudi arabian airlines, International Journal of Aerospace Engineering 2024 (2024).
- [3] F. Wang, J. Bi, D. Xie, X. Zhao, Flight delay forecasting and analysis of direct and indirect factors, IET Intelligent Transport Systems 16 (7) (2022) 890–907.
- [4] M. F. Yazdi, S. R. Kamel, S. J. M. Chabok, M. Kheirabadi, Flight delay prediction based on deep learning and levenberg-marquart algorithm, Journal of Big Data 7 (1) (2020) 106.

A proposed models

A.1 LSTM

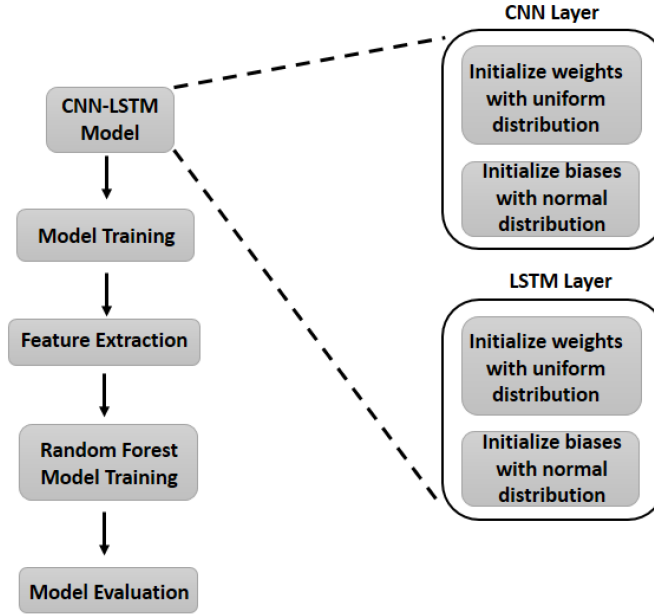


Figure 1: CNN-LSTM Architecture

A.2 XGBoost

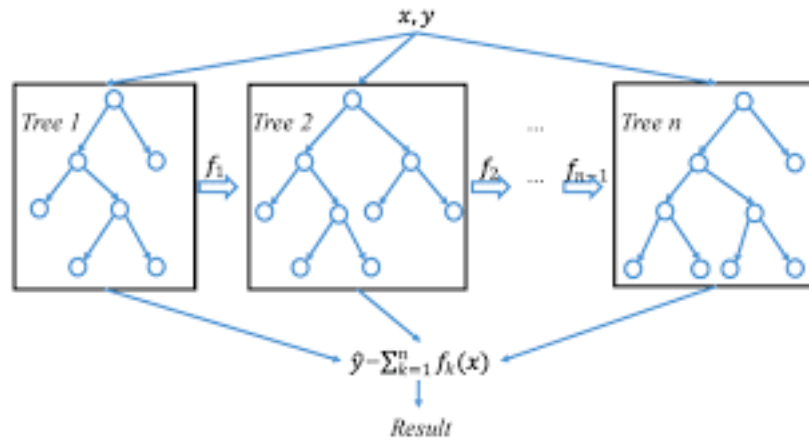


Figure 2: XGBoost Architecture