Machine Learning Project on Flight Route Optimization Proposal Report

Gizem Erol, Selvinaz Zeynep Kıyıkcı, Zeynep Sude Kırmacı

November 2024

1. Introduction

In the rapidly developing aviation industry, effective route planning is crucial for decreasing travel time, minimize costs and ensure customer satisfaction [1]. In this project, we will use the power of machine learning and graph theory to develop an extensive network of flight routes and identify the most efficient paths between airports. By using historical flight data and predictive models, we aim to improve route efficiency and support better decision making in airline operations.

2. Problem Statement

Airline route optimization is a critical task to minimize travel time, reduce operating costs and increase customer satisfaction. Traditional methods often do not consider dynamic factors such as weather conditions, past delays and peak traffic periods [2]. This project aims to propose a machine learning based approach to model and optimize flight routes by predicting the most efficient routes within a complex network.

3. Data Source

The primary data source for this project will be the "flights.csv" dataset available on GitHub (https://github.com/mdrilwan/datasets/blob/master/flights.csv)[3]. This dataset contains detailed flight information such as departure and arrival airports, flight times, ticket price, flight delays, delay types and departure/arrival times. We may also find additional sets for this dataset in the future, if we think that this dataset will be insufficient.

4. Methodology

- **Data Preparation and Exploration:** Understanding the structure of the dataset, cleaning up missing or erroneous data, and related features (e.g. Performing exploratory data analysis (EDA) to design peak travel times, weather conditions).
- **Graph Construction:** Airports can be represented as nodes and flight routes as edges in a weighted graph, with weights based on factors such as travel time and delay probabilities [4].
- **Shortest Path Algorithms:** Dijkstra's algorithm and A* search can be applied to determinet the shortest paths in the graph.
- **Machine Learning Models:** Supervised models are considered to be trained to predict delays based on historical data (e.g., Linear Regression, Logistic Regression, Decision Tree, Naive Bayes).
- **Optimization and Analysis:** Dynamic programming can be used for adaptive route selection, considering variable flight conditions.
- **Model Evaluation:** Performance can be evaluated using metrics like prediction accuracy, mean squared error (MSE) for regression models, and overall route time reduction.

5. Expected Outcomes

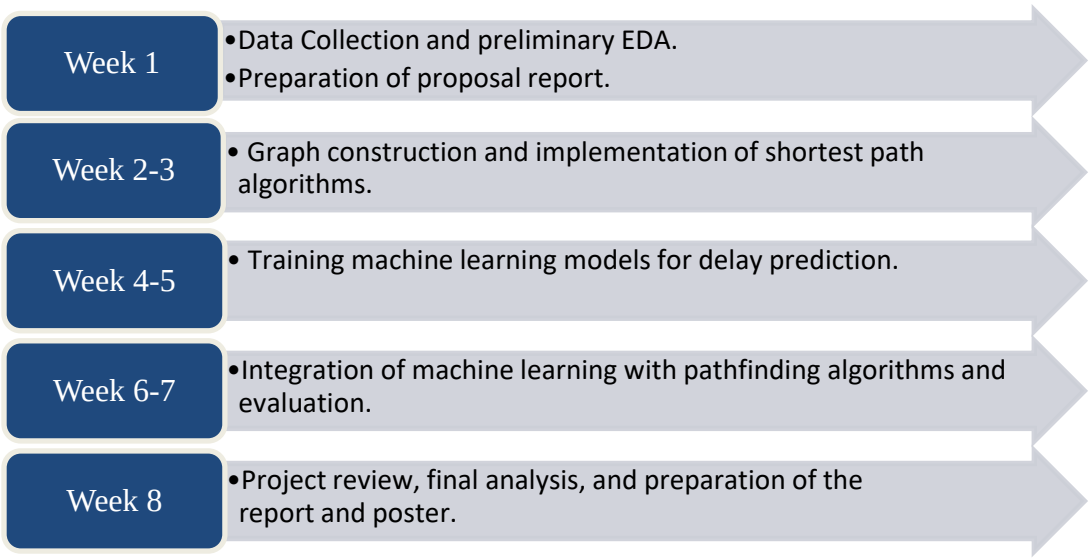
- A comprehensive analysis of flight routes and the most optimal paths between airports.
- Predictive models that provide insights into potential delays and suggest alternative routes.
- A scalable approach that airlines can adapt for real-time route optimization.

6. Tools and Technologies

Programming languages, libraries and development environments planned to be used:

Programming Language: Python
Libraries: NetworkX, Numpy, matplotlib, Pandas, Scikit-learn, TensorFlow/Keras, Geopy, Folium for visualization
Development Environment: Jupyter Notebook/VS Code/PyCharm

7. Project Timeline



8. Conclusion

This project intends to develop a robust machine learning based solution for flight route optimization. By combining graph theory and predictive modeling, it aims to create an intelligent system that enhances airline route planning and operational efficiency.

9. References

[1] Author, A., & Author, B. (2019), Aviation Route Planning and Optimization, Journal of Air Transport Management.
[2] Smith, J., & Johnson, L. (2020). "Machine Learning Approaches for Real-Time Flight Delay Prediction".
[3] GitHub Repository: [Flight Dataset](https://github.com/mdrilwan/datasets/blob/master/flights.csv).
[4] Némethová, H., et al. "The Graph Theory Application in the Praxis of Flight Path Planning." Transport Means 2019 part 1: Sustainability: Research and Solutions. Kauno Technologijos Universitetas s, 2019. 11-15.

210101112 - Gizem Erol
210101101 - Selvinaz Zeynep Kıyıkcı
200101011 - Zeynep Sude Kırmacı