

# Project Report: Data Analytics with Tools

Mika Pärssinen

December 26, 2024

## 1 Abstract

## 2 Introduction and Problem Formulation

Air travel is one of the most popular and efficient form of transportation for long distances. It not only enables people to explore new destinations and experiences but also keep the connection between families, friends, and communities. Millions of passengers rely on airplanes every day to transport them safely and efficiently from one location to another.

However, flight delays are a continuous issue that disrupt this otherwise reliable form of travel. From the traveler's perspective, delays can be inconvenient and frustrating, leading to missed connections, appointments or events. For airlines, delays result in significant economic consequences, including increased operational costs, dissatisfied customers, and damage on their reputation. Airplane companies estimates that these losses counts to billions of dollars annually. Fixing flight delays would not only fix passengers frustration but also save airlines substantial financial resources, contributing to a more efficient and customer friendly air travel experience.

The aim of this project is to analyze historical flight data from American Airlines 2008 dataset, and then develop predictive models to predict flight delays using machine learning techniques. By identifying key factors that contribute to delays, such as departure times, distances, and weather conditions, this research seeks to provide insights for mitigating delays. Ultimately, the goal is to enhance both efficiency for airlines and convenience for passengers.

## 3 Implementation

## 4 Evaluation

## 5 Future work

## 6 Discussion

\* Cleaning

## 7 Conclusion

## 8 References