Enhancing Aviation Safety Through Data Analysis

Fatima Tariq (ft07200) & Muminah Khurram (mk07521) | CS343: Graph Data Science

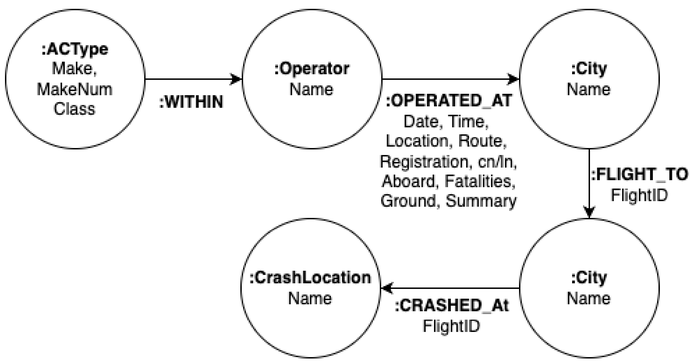
01. Introduction

Aviation safety is a critical concern globally. Understanding the causes and patterns of aviation accidents is crucial for improving safety measures and preventing future incidents. Our work leverages data analysis and machine learning techniques to gain insights into aviation accident data to enhance safety measures and reduce the risk of accidents. With our modeling and machine learning algorithms we provide insights into accident patterns, risk factors, and potential preventive measures for stakeholders in the aviation industry.

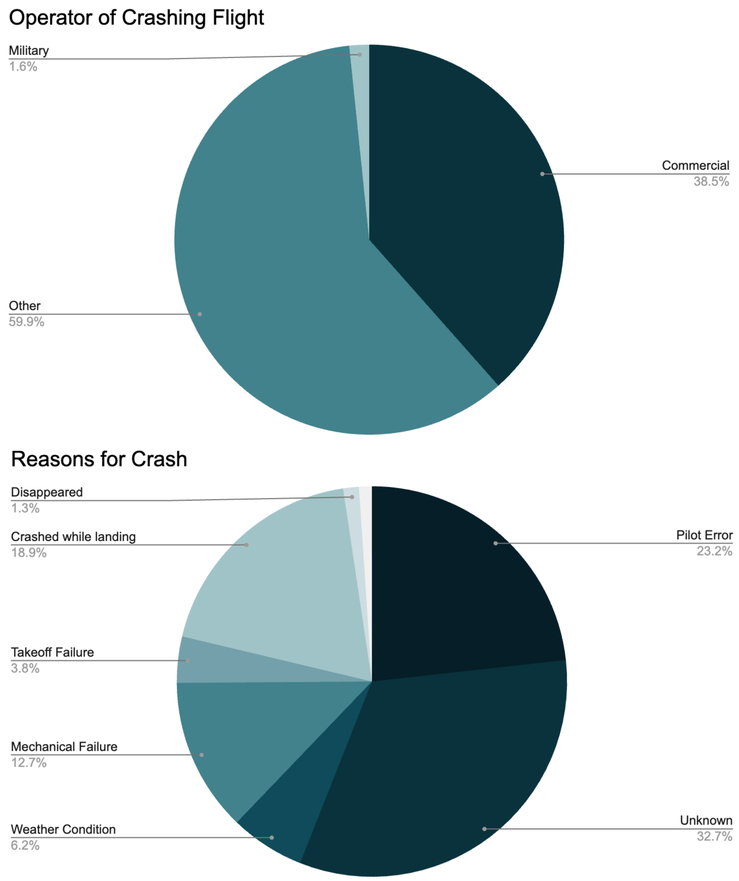
02. Methodology

Data Collection and Preprocessing:
From the two referenced sources using Python to extract and merge data from these sources, we created the dataset for our analysis.

Graph Modeling:



Graph Database and Machine Learning Tasks:
Neo4j was employed for data storage and querying. Its graph structure facilitated efficient retrieval and we used pipelines for supervised machine learning models were trained to predict missing flight routes based on historical data and perform node classification to classify unknown aircraft types using features extracted from the data.



03. Graph Projections and Machine Learning:

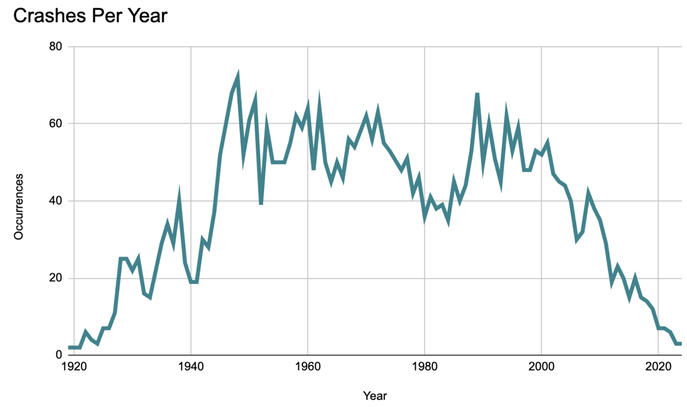
Machine learning techniques were employed for prediction and classification with supervised machine learning models: Logistic Regression, Random Forests (decision tree), Multilayer Perceptron

Link Prediction:

- Winning Model - Link Prediction: Multilayer Perceptron (Highest test score: 0.74)
- Model can predict potential flight routes.

Aircraft Type Classification:

- Winning Model - Link Prediction: Random Forest (Decision trees: 10)
- Model can identify and categorise aircraft types with high accuracy.



04. Results & Analysis

- Centrality analysis reveals crash and route have very few outliers and connectivity is concentrated to what are real world aviation hubs.
- Similarly most frequent crash locations are among the world's largest cities, which also tend to have the highest flight activity.
- Most frequent crashing operators are predominantly American, with an overlap between military and commercial categories (”Other”).
- The timeline chart of accidents has been on a steep decline, indicating a historical trend of decreasing crash occurrences while seeing massive flight frequency increases.
- The most common known crash reason is pilot error and with modern crash rate decrease that we attribute to better safety protocols and auto-pilot programs suggests that regulatory interventions are yielding positive results.

05. References

[1] “Database index,” www.planecrashinfo.com. https://www.planecrashinfo.com/database.htm
[2] “Airplane Crashes Since 1908,” www.kaggle.com. https://www.kaggle.com/datasets/saurograndi/airplane-crashes-since-1908