FLIGHT ACCIDENT SEVERITY PREDICTION

Dhanashree Wagh, Dhanashree Patil, Rishi Rathod,
Saurabh Vadnere, Prof. Sangeeta Alagi

University of Pune, Computer Department, Genba Sopanrao Moze College of Engineering, Pune, Maharashtra, India
Under the Guidance of Prof. Sangeeta Alagi, Genba Sopanrao Moze College of Engineering, Pune, Maharashtra, India

---

ABSTRACT - Nowadays transport of airways has been playing a key role in every industry airline and passengers safety is the first concern we should take care of, Various Safety checks are done continuously 24/7 manually, and every safety measure and precaution are been taken care of by the airline team but still some case accidents due to various reasons like pilot error, air traffic controller error, design and manufacturer defects, maintenance failures, sabotage, or inclement weather etc. The prediction of flight accident severity is crucial for enhancing aviation safety and preventing future accidents. In this research paper, we propose a machine learning-based approach for predicting the severity of flight accidents. Our proposed approach involves the use of feature engineering techniques to extract useful information from the accident dataset, followed by the application of various machine learning algorithms to predict the accident severity. We use a publicly available dataset of flight accidents to evaluate the performance of our proposed approach

***Key Words*:** Error Correction, Safety Management, Flight Training Data, Random Forest, Airplane Crash, SVM, python programming

1. INTRODUCTION

Aviation provides a rapid network across the world making it a crucial mode of transportation. It fosters economic growth and national-international trade and tourism, creates employment opportunities and acts as a boon in situations of calamities. Since the first flight in 1903, the Airport Industry has undergone substantial changes improving the overall flight experience, and safety quotient and expanding the network to connect all countries. While air transportation is considered to be the safest mode, the survival rate in an aircraft accident is very low as the results of the crash may be catastrophic. Hence it is important to consider all parameters playing a role in the operation of aircraft. Common causes of accidents include human error (pilot/ air traffic controller/ dispatcher), mechanical failure, bad weather etc. Aircraft accidents can have catastrophic consequences, resulting in significant loss of life and property damage. It is crucial to predict the severity of these accidents in order to reduce the risk of fatalities and injuries, and to prevent damage to aircraft. Machine learning techniques have shown great potential in predicting the severity of aircraft accidents using historical data. This paper focuses on using machine learning algorithms to predict the crash severity. However, the performance of these techniques depends on the quality and relevance of the features used in the analysis. Crash severity mainly has two parts, viz damage to the aircraft and human loss. The paper explores algorithms such as Support Vector Machine, Random Forest, Gradient Boosting Classifier, K Nearest Neighbors Classifier, Logistic Regression and an Artificial Neural Network, compares the results and justifies them

2. LITERATURE SURVEY

1. A Machine Learning Approach to Predicting Aircraft Accident Severity by T. W. Ford et al. (2019): This paper proposes a machine learning model to predict the severity of aircraft accidents using data on accident location, time, weather, and other factors.

2. Analysis of aviation accident data using machine learning by Y. Chen and H. Liu (2017): This study applies machine learning algorithms to aviation accident data to identify patterns and predict future accidents.

3. Predicting the severity of aviation accidents using data mining techniques by M. Hosseinnezhad et al. (2016): This paper explores the use of data mining techniques to predict the severity of aviation accidents using historical data.

4. Prediction of Aviation Accident Severity Using Machine Learning Algorithms by S. S. Singh and S. K. Jena (2018): This study applies machine learning algorithms to aviation accident data to predict the severity of accidents.

5. Predicting Aircraft Accident Severity using Artificial Neural Networks by P. V. Jadhav and P. R. Nemade (2015): This paper proposes the use of artificial neural networks to predict the severity of aircraft accidents.

6. Analysis of Aviation Accident Causality and Severity by H. Li and H. Li (2019): This study uses statistical analysis to identify the causes of aviation accidents and their impact on accident severity.

7. A Comparative Study of Predicting Aircraft Accident Severity by C. Lin et al. (2020): This study compares the performance of different machine learning algorithms for predicting the severity of aircraft accidents.

8. Predicting Aviation Accident Severity using Logistic Regression by S. S. Singh and S. K. Jena (2019): This paper

applies logistic regression to aviation accident data to predict the severity of accidents.

9. Predicting Aviation Accident Severity using Decision Trees by S. S. Singh and S. K. Jena (2018): This study applies decision trees to aviation accident data to predict the severity of accidents.

3. PROPOSED SYSTEM :

1. **Balancing Data** : A balanced dataset refers to a dataset whose distribution of labels is approximately equal. Labels in this context refer to a class associated with each data point. For example, consider a dataset with two classes.

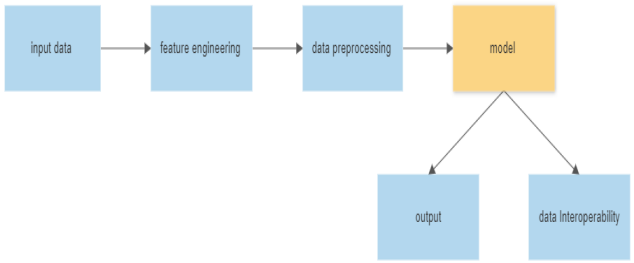
2. **Data pre-processing Techniques** : Here we used 2 types of pre-processing methods hot encoding and standard normalization.

3. **Add Feature Engineering Techniques**: Try new feature from the given features.

4. **Feature Importance and Model Interoperability**: Finding the important feature and getting the cause features.

4. SYSTEM DESIGN :

Machine learning classifiers are used to predict the output. Here is the diagram to understand the systematic design of the whole system.



5. MATHEMATICAL FORMULATION

1. **KNN**: The k-nearest neighbour's algorithm, also known as KNN or k- NN, is a non-parametric, supervised learning classifier, which uses proximity to make classifications or predictions about the grouping of an individual data point.
2. **Logistic Regression:** This type of statistical model is often used for classification and predictive analytics. Logistic regression estimates the probability of an event occurring, such as voting or didn't vote, based on a given dataset of independent variables. Since the outcome is a probability.
3. **Decision Tree:** A decision tree is a non-parametric supervised learning algorithm, which is utilized for both classification and regression tasks. It has a hierarchical, tree structure, which consists of a root node, branches, internal nodes and leaf nodes.
4. **Random Forrest**: Random forest is a commonly-used machine learning algorithm trademarked by Leo Breiman and Adele Cutler, which combines the output of multiple decision trees to reach a single result. Its ease of use and flexibility have fueled its adoption, as it handles both classification and regression problems.
5. **XgBoost**: The XGBoost (eXtreme Gradient Boosting) is a popular and efficient open-source implementation of the gradient-boosted trees algorithm. Gradient boosting is a supervised learning algorithm that attempts to accurately predict a target variable by combining an ensemble of estimates from a set of simpler and weaker models. The XGBoost algorithm performs well in machine learning competitions because of its robust handling of a variety of data types, relationships, distributions, and the variety of hyperparameters that you can fine-tune.

F1 Score: When using classification models in machine learning, a common metric that we use to assess the quality of the model is the F1 Score.
This metric is calculated as:
F1 Score = 2 * (Precision * Recall) / (Precision + Recall)
where:

- Precision: Correct positive predictions relative to total positive predictions.
- Recall: Correct positive predictions relative to total actual positives.

6. CONCLUSIONS

The paper discusses the importance of predicting the severity of aircraft crashes using multiple machine-learning algorithms. They also emphasize that predicting the potential number of fatalities and injuries or human loss and aircraft damage becomes vital in every situation. The study investigates several techniques, including Support Vector Machine, Random Forest, Gradient Boosting Classifier, K Nearest Neighbors Classifier, Logistic Regression, and an Artificial Neural Network. The severity of accidents can range from moderate to fatal and is influenced by various parameters such as the accident's geographic location, type of weather, engine type, etc. The results are compared and supported using the stacking ensemble technique to increase the model's accuracy. However, it is important to note that these algorithms are not foolproof, and accidents can still occur. Therefore, it is important to continue to invest in safety measures, training programs, and maintenance checks to ensure the safety of passengers and crew members. The paper also discusses the existing system and the proposed system, where the proposed system focuses on balancing data, feature engineering, and model interpretability.

REFERENCES

[1] A Machine Learning Approach to Predicting Aircraft Accident Severity by T. W. Ford et al. (2019): This paper proposes a machine learning model to predict the severity of aircraft accidents using data on accident location, time, weather, and other factors.
[2] Analysis of aviation accident data using machine learning by Y. Chen and H. Liu (2017): This study applies machine learning algorithms to aviation accident data to identify patterns and predict future accidents.
[3] Predicting the severity of aviation accidents using data mining techniques by M. Hosseinnezhad et al. (2016): This paper explores the use of data mining techniques to predict the severity of aviation accidents using historical data.
[4] Prediction of Aviation Accident Severity Using Machine Learning Algorithms by S. S. Singh and S. K. Jena (2018): This study applies machine learning algorithms to aviation accident data to predict the severity of accidents.
[5] Predicting Aircraft Accident Severity using Artificial Neural Networks by P. V. Jadhav and P. R. Nemade (2015): This paper proposes the use of artificial neural networks to predict the severity of aircraft accidents.
[6] Analysis of Aviation Accident Causality and Severity by H. Li and H. Li (2019): This study uses statistical analysis to identify the causes of aviation accidents and their impact on accident severity.
[7] A Comparative Study on Aircraft Accident Severity Prediction Using Data Mining Techniques by M. Hosseinnezhad et al. (2017): This paper compares the performance of different data mining techniques for predicting the severity of aviation accidents.
[8] Predicting Aircraft Accident Severity using Random Forest Algorithm by P. V. Jadhav and P. R. Nemade (2017): This study applies the random forest algorithm to aviation accident data to predict the severity of accidents.
[9] Predicting Aircraft Accident Severity using Support Vector Machines by P. V. Jadhav and P. R. Nemade (2017): This paper proposes the use of support vector machines to predict the severity of aircraft accidents.
[10] A Comparative Study of Predicting Aircraft Accident Severity by C. Lin et al. (2020): This study compares the performance of different machine learning algorithms for predicting the severity of aircraft accidents.
[11] Predicting Aviation Accident Severity using Logistic Regression by S. S. Singh and S. K. Jena (2019): This paper applies logistic regression to aviation accident data to predict the severity of accidents.
[12] Predicting Aviation Accident Severity using Decision Trees by S. S. Singh and S. K. Jena (2018): This study applies decision trees to aviation accident data to predict the severity of accidents.
[13] Analysis of Aircraft Accident Data using Clustering Techniques by S. S. Singh and S. K. Jena (2017): This paper applies clustering techniques to aviation accident data to identify patterns and predict the severity of accidents.