A Stacking Model for Aviation High Risk Event Decision Making

Monika¹, Seema Verma², Pardeep Kumar³ and Mansi Kakran⁴

¹ Assistant Professor, Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan
² Professor, Department of Electronics, Banasthali Vidyapith, Jaipur, Rajasthan
³ Associate Professor, Department of Electronics and Communication, SGT University, Gurugram, Haryana
⁴ Research Scholar, Department of Computer Science, Banasthali Vidyapith, Jaipur, Rajasthan

E-mail: mor.monika@gmail.com

Abstract. Aviation accidents are one of the most causes of severe injuries and death worldwide. With the increase in inter-countries movement of the peoples around the world, air transportation is increasing at a much faster rate. As result possibilities of air accidents have also increased which may be due to inter-air collision, technical fault in airplane systems and pilot faults, etc. Nowadays aviation accidents have become one of the major causes of severe injuries and fatalities around the world. This attracts, the research community to look into the factors behind these accidents by applying data analysis techniques based on an advanced machine learning algorithm. This analysis would be beneficial in forecasting the high-risk events based on the available historical accident data. So, in this paper, we classify all aviation abnormal events into three groups with regards to their associated risk: low-risk events, medium-risk events, high-risk events and then deployed ML algorithms like Random Forest, Support Vector Machine (SVM), XGBoost, Logistic Regression, etc., to make forecasts about high-risk events. After analyzing the results, the performance of the hybrid model that is formed by stacking of best-performing algorithms is found to best in quantifying the high-risk events.

Keywords: Air transportation, Air Traffic Congestion, Machine Learning, Aviation Accidents

1. Introduction

Due to increasing globalization, cheap air tickets, and the safest mode of transportation more and more people are opting for air travel. Also, it seems to be double in coming years, which may lead to aviation accident likelihood [1]. The aviation industry is currently struggling to manage the current pressure of traffic. Flight delays, on the way congestion, will worsen owing to enormous growth in the aviation traffic in the bounded airspace [2]. Such situations result in a higher risk of an aviation accident. The increased air traffic congestion puts an immense workload on Air Traffic Control (ATC) operators in keeping the system security level as safe as before [3]. Due to the hazardous outcome of aviation accidents, the safety of aviation transportation has obtained the attention of the research community [4][5]. Various agencies like the NTSB, ASRS, and FAA are maintaining the historical data of aviation accidents. This historical information will be considered as input for the safety decisions making mechanism. Machine learning focuses on algorithms that are capable of learning over the dataset. Such forecasting measures can have a huge role to play in planning and decision making before the flight takes off and can help in reducing air accidents, saving many precious lives [6].

The key motivation of this paper is to analyze the ASRS database from 2010 to 2020 and provide a comprehensive comparison of various machine learning algorithms SVM, XGboost Random Forest,
and Logistic Regression and then forming a hybrid machine learning model by stacking best performing algorithm. For this purpose, the dataset is categorized into three groups: low-risk events, medium-risk events, high-risk events on the basis of the severity of events. The results of the stacking model are good for predictive analysis of high-risk events. This predictive analysis can help in learning from the data pattern and provide help in analyzing the ATC personnel in implementing a proactive safety process.

This manuscript is further organized as follows: in Section 2 presents literature review about the ML techniques applied yet. Section 3 provides detailed insight into the dataset and its categorization used in the prediction model. Experiments and results of the analysis are discussed in Section 4, following concluding remarks and future scope in Section 5.

2. Literature Review
With the advancement in technology, tools, data processing frameworks, etc, it becomes quite easy to implement any novel technique over a dataset. Due to this, the research community develops and uses machine learning tools with special considerations for aviation safety. Apoorv Maheswari et al. [5] carry a comparative overview of supervised machine learning techniques likes Neural Network (NN), regression and classification, etc. to solve the air travel demand problem. Results of SVM and NN provide high accuracy in the cost of a large dataset. Abrar Om Alkhamisi and Rashid Mehmood [7] propose a hybrid model that predicts the likelihood of the risk for aviation safety assurance using deep learning and ML algorithms. Recurrent neural networks were used for exploring interrelationship of data. The performance of the ensemble model is fine tuned with the base classifier diversity score. Di Zhou et al. [8] applied, PSO-SVM and LSTM algorithms for unsafe event identification and their forecasting. Authors had applied this approach to Aircraft Communications and Addressing dataset. They first reduce the dimension of the dataset using RFECV and applies SVM and further optimize the results with the PSO to increase the speed and accuracy of SVM. After analyzing the results, the author concludes that deep learning algorithms prove to be better than SVM.

Anupam Mishra et al. [9] have applied the SVM model for predicting the pilots cognitive condition during the various phases of flight. Ehsan Esmaeilzadeh et al. [10] have used SVM for exploring the non-linear co-relation between flight delay and its outcome. The impact of various parameters like demand capacity, delay, airport ground support operation, flow management, etc., was examined. The correlation of these factors may provide an insight for the reason behind flight delay and delay patterns. Priyam Mathur et al. [11] have proposed a linear forecasting method for those situations that may result in an accident. The logistic regression model is applied over the aviation safety network database.

A. Burnett and D. Si [12] perform predictive analysis of many algorithms based on machine learning approaches like the k-nearest neighbor, Artificial Neural Network (ANN), etc. for analyzing those conditions which are mainly responsible for aviation accidents and incidents. Results proved that a better prediction is possible with the ML algorithm as compared to statistical one.

Xiaomei et al. [13] did a comparative study of the gray neural network, Support Vector Regression (SVR), Deep Belief Network (DBN), and PCA-DBN. After analyzing the results, the authors conclude that PCA-DBN has higher accuracy than other models that prove to be beneficial for the predictive safety of aviation. B. Mathews et al. [14] have applied various data mining algorithms to detect anomalies in the FOQA database. Oza et al. [15] had developed nonnegative matrix factorization and Mariana algorithms to carry out multi-label classification of ASRS dataset. Budalakoti et al. [16] have proposed few algorithms to find out and characterize the anomalies that occur due to values of switch sensors available in the airliner’s cockpits. Lee et al. [17] compares various machine learning algorithms on the American Airlines dataset and analysis shows that with these authors are able to predict the taxi time within 5 minutes after departure. Ukai et al. [18] use NN algorithms for deployment prediction and results prove that techniques provide good accuracy for aviation applications.

The application of ML techniques in the area of aviation safety is high due to insufficient high-quality data. So, there is a big scope to use ML techniques for solving the problems of the aviation community.
3. Dataset

For this study data is collected from ASRS (Aviation Safety Reporting System) database [19]. ASRS operated and controlled by National Aeronautics and Space Administration (NASA) is responsible for collecting, processing and analyzing the data submitted from aviation operations personnel which include aviation incidents, situation reports which contain both insecure occurrences and high risk operations. Every report is thoroughly analyzed by analysts of ASRS and analyst might also contact the reporter to clarify the details provided by the reporter before uploading it to database. For this study, 10 years of data from the ASRS were collected which comprised 47887 records. The dimension of the original data set are shown in Table 1 and the variable types are shown in Table 2. To validate the air traffic data divided into training and testing datasets in the ratio 4:1.

3.1. Features

Table 3 States the feature set that is responsible for Aviation safety prediction from ASRS dataset, which carries information about event characteristics to an event happening location and much more.

3.2. Class Categorization

Since most features are categorical. Which do not indicate the event consequence that might coincide with many of event consequences, i.e. outcome which may belong to different risk levels. Another major problem is the imbalance of class distribution which may lead to maximum algorithms diverted towards the majority class. Also, it is not possible to train a particular model that predicts a specific event result. To resolve these issues, we classify the dataset based on likelihood-based classification where the outcome belongs to one of three available categories: low-risk events, medium-risk events, high-risk events. The high-risk category includes cases where the aircraft has been damaged, there is physical injury, due to problem of engine the flight has shut down in the air, etc. So, there is great loss financially as well as in terms of lives lost in High-risk events. Medium-risk category involves chances of accidents, but can be avoided by aircraft diversion, reorientation, executing go around or missed approach, like canceling, delaying, maintaining the flight etc. These cases might not involve loss in lives, but it will directly affect the reputation of the aviation industry and will have a financial impact. Low-risk event category include minor problem which can be easily solved by taking a few measures which would result in maintaining customer’s trust, no loss of life and would be cost effective. So, based on seriousness each event is related to a particular risk category as shown in Table 5. what if aviation safety is at high risk, but there is no measure taken because our model predicted it as Low Risk. That is a situation we would like to avoid! Therefore, we have prioritized the class. The class distribution with their priority over the dataset is shown in Table 4.

4. Experiments Analysis and Results

In this section, we have implemented and train various machine learning algorithms like SVM, XGBoost, Logistic Regression, and Random Forest on the ASRS dataset. Then we make a hybrid model by stacking the three best-performing algorithms. Here we provide a comparative analysis of these algorithms based
Table 3: Feature Set

| Place               | Locale Reference          | State Reference          | Altitude MSL Single Value |
|---------------------|---------------------------|--------------------------|---------------------------|
| Environment         | Flight Conditions         | Weather Elements/Visibility | Light                    |
| Aircraft-1          | ATC / Advisory            | Make Model Name          | Flight Plan               |
|                     |                           | Mission                  | Flight Phase              |
|                     |                           | Route In Use             |                           |
| Component           | Aircraft Component        | Aircraft Reference       | Problem                   |
| Aircraft-2          |                           | Route In Use             |                           |
| Person-1            | Human Factors             | Communication Breakdown  |                           |
| Person-2            | Location Of Person        | Function                 |                           |
| Events              | Anomaly                   | Detector                 | When Detected             |
| Assessment          | Contributing Factors / Situations | Primary Problem       |

Table 4: Class Distribution

| Class     | No. Of records | Class distribution in % | Class Priority |
|-----------|----------------|-------------------------|----------------|
| High Risk | 16034          | 33                      | High priority  |
| Medium Risk| 21355        | 45                      | Medium priority|
| Low Risk  | 10498          | 22                      | Low priority   |

on their use and ease, which aid in choosing a suitable algorithm for aviation forecasting problems. All of these models are implemented using python. Comparisons of different algorithms are done based on Recall as shown in Figure 1. The recall is the ability of a model to identify all relevant instances. It is calculated as a count of class members classified correctly over the total count of class members. Recall is the ability to specify how much accurately the model is capable to identify the relevant data. We refer to it as Sensitivity or True Positive Rate.

\[
Recall = \frac{TruePositive}{TruePositive + FalsePositive}
\]

On comparing the algorithms based on recall and prioritizing the classes SVM turns out to be the best
algorithm. XGBoost and Random Forest give results close to SVM. So we further combine these three best-performing algorithms using stacking. For improving the results of the model stacking ensemble method is used. In stacking we have implemented SVM and XG Boost as level 1 classifier models and Random forest as a Meta Classifier model. The results of stacking are shown below in Fig1.

Stacking with Level 1 classification models: Logistic SVM, XGBoost and Meta classifier as Random Forest has the highest recall for high risk 0.74. But recall for medium risk is 0.68 which is lower than recall gained from many algorithms as shown in Figure 2. As we have prioritized the classes and highest priority is given to high risk. So this model turns out to be the best model for high-risk events. Generally, there is a not exact way out to algorithm selection, but we provide a comparative analysis for high-risk aviation events forecasting algorithm selection. SVM performs better, but with a large sample size for higher accuracy. XGBoost performance, better due to its high speed and enhanced performance. Random

Table 5: Class categorization on the basis of seriousness of each event

| High Risk                      | Medium Risk                                               | Low Risk                                                   |
|--------------------------------|------------------------------------------------------------|------------------------------------------------------------|
| Flight Crew In-flight Shutdown | Flight Crew Became Reoriented                              | Flight Crew Exited Penetrated Airspace                     |
| Flight Crew Regained Aircraft Control | Flight Crew Executed Go Around Missed Approach             | Flight Crew FLC Overrode Automation                        |
| Flight Crew Landed in Emergency Condition | Flight Crew Overcame Equipment Problem                     | Flight Crew FLC Overrode Automation                        |
| Air Traffic Control Separated Traffic | Flight Crew Rejected Takeoff                              | Flight Crew Requested ATC Assistance Clarification         |
| Aircraft Damaged               | Flight Crew Took Evasive Action                            | Flight Crew Landed As Precaution                          |
| Air Traffic Control Issued Advisory / Alert | General Flight Canceled Delayed                           | Flight Crew Returned To Clearance                         |
| General Declared Emergency     | General Work Refused                                       | Flight Crew Returned To Departure Airport                  |
| General Physical Injury / Incapacitation | General Release Refused Aircraft Not Accepted               | Flight Crew Returned To Gate                              |
| General Evacuated              |                                                             | Flight Crew FLC complied w Automation Advisory             |
|                                |                                                             | Air Traffic Control Provided Assistance                    |
|                                |                                                             | Aircraft Automation Overrode Flight Crew                   |
|                                |                                                             | Aircraft Equipment Problem Dissipated                     |
|                                |                                                             | General Police Security Involved                           |
|                                |                                                             | General None Reported Taken                                |


forest uses average sampling to enhance accuracy and overcome the problem of over-fitting. When all these three are put to stacking then the stacked model performs very well for the high-risk events. This in turn helps in forecasting high-risk-based events during emergency conditions in aviation.

5. Conclusion and Future Scope
In this paper, a diverse set of ML techniques was used to carry out their comparative analysis from the point of view of their consequence to high-risk events in aviation safety problems. We design a risk-based event strategy based on outcome and divide events in three groups: low-risk events, medium-risk events, high-risk events. A hybrid model is developed by stacking best-performing algorithms. The has modeled outcome shows that the stacking model outplay the individual algorithms in terms of recall of the high-risk events. Future work in this direction can be enhanced by including other information like the pilot or operator response in case of abnormal events which further helps the model to learn from the past.

References
[1] Association I A T et al. 2017 International Air Transport Association 24
[2] NextGen 2020 The next generation air transportation system https://www.nasa.gov/sites/default/files/atoms/files/nextgen whitepaper 06 26 07.pdf

[3] Sarter N B and Alexander H M 2000 The international journal of aviation psychology 10 189–206

[4] Hwang I and Seah C E 2008 Proceedings of the IEEE 96 2040–2059

[5] Maheshwari A, DavendraIingam N and DeLaurentis D A 2018 2018 Aviation Technology, Integration, and Operations Conference p 3980

[6] Shyur H J 2008 Computers & Industrial Engineering 54 34–44

[7] Omar Alkhamisi A and Mehmood R 2020 2020 6th Conference on Data Science and Machine Learning Applications (CDMA) pp 54–59

[8] Zhou D, Zhuang X, Zuo H, Wang H and Yan H 2020 IEEE Access 8 103665–103683

[9] Mishra A, Shrivastava K K, B A A and Quadir N A 2019 2019 International Conference on Smart Systems and Inventive Technology (ICSSIT) pp 360–364

[10] Esmaeilzadeh E and Mokhtarimousavi S 2020 Transportation Research Record 2674 145–159

[11] Mathur P, Khatri S K and Sharma M 2017 2017 International Conference on Infocom Technologies and Unmanned Systems (Trends and Future Directions) (ICTUS) pp 725–728

[12] Goodfellow I, Bengio Y, Courville A and Bengio Y 2016 MIT Press, Cambridge, MA

[13] Ni X, Wang H, Che C, Hong J and Sun Z 2019 Safety science 112 90–95

[14] Matthews B, Das S, Bhaduri K, Das K, Martin R and Orza N 2013 Journal of Aerospace Information Systems 10 467–475

[15] Orza N, Castle J P and Stutz J 2009 IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 39 670–680

[16] Budalakoti S, Srivastava A N and Otey M E 2008 IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) 39 101–113

[17] Lee H, Malik W and Jung Y C 2016 16th AIAA Aviation Technology, Integration, and Operations Conference p 3910

[18] Ukai T, Chao H and DeLaurentis D A 2017 17th AIAA Aviation Technology, Integration, and Operations Conference p 3081

[19] 2020 Asrs database online - aviation safety reporting system accessed: 10-12-2020 URL https://asrs.arc.nasa.gov/search/database.html