Journal of Systems Engineering and Electronics
Vol. 35, No. 2, April 2024, pp.396 – 405

# Classification of aviation incident causes using LGBM with improved cross–validation

NI Xiaomei[1,2], WANG Huawei[1,*], CHEN Lingzi[1], and LIN Ruiguan[1]

1. School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China;
2. School of Aeronautical Engineering, Nanjing Vocational University of Industry Technology, Nanjing 210023, China

**Abstract:** Aviation accidents are currently one of the leading causes of significant injuries and deaths worldwide. This entices researchers to investigate aircraft safety using data analysis approaches based on an advanced machine learning algorithm. To assess aviation safety and identify the causes of incidents, a classification model with light gradient boosting machine (LGBM) based on the aviation safety reporting system (ASRS) has been developed. It is improved by *k*-fold cross-validation with hybrid sampling model (HSCV), which may boost classification performance and maintain data balance. The results show that employing the LGBM-HSCV model can significantly improve accuracy while alleviating data imbalance. Vertical comparison with other cross-validation (CV) methods and lateral comparison with different fold times comprise the comparative approach. Aside from the comparison, two further CV approaches based on the improved method in this study are discussed: one with a different sampling and folding order, and the other with more CV. According to the assessment indices with different methods, the LGBM-HSCV model proposed here is effective at detecting incident causes. The improved model for imbalanced data categorization proposed may serve as a point of reference for similar data processing, and the model's accurate identification of civil aviation incident causes can assist to improve civil aviation safety.

**Keywords:** aviation safety, imbalance data, light gradient boosting machine (LGBM), cross-validation (CV).

**DOI:** 10.23919/JSEE.2024.000035

## 1. Introduction

The safety of the air transportation system is becoming increasingly important as the demand for air travel is predicted to skyrocket over the next two decades. Despite the fact that aviation has a high level of safety, accidents and incidents continue to occur [1]. It is a good way to learn from accident and incident reports, which show why accidents happen and why they do not, and can help identify both the dangers and the precautions [2]. A few machine learning studies have been carried out to analyze these reports [3−6].

The Civil Aeronautics Administration, the Traffic Management Bureau, and other agencies have accident reports with extensive civil aviation incident information. Almost every aspect of aircraft operations that could go wrong is covered by the aviation safety reporting system (ASRS) [7]. However, due to the following challenges, developing a model from ASRS data for the prediction of causes linked with accidents is not a simple task:

(i) High-dimensional data. Each incident record consists of more than 50 items ranging from the operational context (weather, visibility, flight phase, and flight conditions) to the characteristics of the anomalous operation (aircraft equipment, malfunction type, and event synopsis). It is likely for the incident to occur at any phase and location due to a variety of factors (i.e., company policies, weather, human factors, etc.), which makes it hard to predict the exact cause of incidents with complex features [8,9].

(ii) Primarily categorical data. Over 99% of the items in the ASRS database are categorical, and only a few attributes (i.e., crew size) are numerical in each record. Since the categorical features are not informative, how to improve the accuracy of the classification of the model trained by the categorical features without compromising the model performance is an issue worthy of investigation.

(iii) Imbalanced class distribution. In the ASRS, the number of records in one class (e.g., the incidents directly caused by aircraft) is significantly larger than that of the others (e.g., the incidents directly caused by company

Manuscript received April 28, 2021.
*Corresponding author.
This work was supported by the National Natural Science Foundation of China Civil Aviation Joint Fund (U1833110), and Research on the Dual Prevention Mechanism and Intelligent Management Technology for Civil Aviation Safety Risks (YK23-03-05).

policies or the environment). Because of the imbalance in the distribution of causes, incidents that were originally caused by airline policies are now being identified as being caused by aircraft failures, and resources in civil aviation safety management are being misdirected to aircraft repair at the expense of actual causes such as management.

Machine learning algorithms that presume a somewhat well-balanced distribution will have a severe problem when the number of data across distinct classes is highly skewed [10,11]. Due to its powerful ability to establish a dense representation of the feature space, which makes it effective in learning high-order features from the raw data [12,13], gradient boosting decision tree might be a good candidate to discover the highly intricate relationship between characteristics and incident causes when dealing with categorical data [14,15]. Since the challenges in ASRS above, an improved cross-validation (CV) method to optimize the classification model with light gradient boosting machine (LGBM) is adopted in this paper. To learn the correlations between incident characteristics and causes, an LGBM [16,17] is first created to analyse categorical data. The imbalance of causes is then addressed using an upgraded CV approach. CV is the most widely used approach for evaluating a model's predicted performance before or after it is generated via a modeling procedure [18–20].

This paper addresses the aforementioned issues and technologies, as well as the vital requirements of aviation safety situational awareness, intelligent management [21], and "the annual implementation plan for strengthening transportation safety production"[22].

## 2. Data analysis

ASRS is a program operated by the National Aeronautics and Space Administration (NASA) with the ultimate goal of increasing aviation system safety by discovering system safety hazards hidden in the multitude of air traffic operations. Over the past few decades, ASRS has become one of the world's largest sources of information on aviation safety and human factors. As one of its primary tasks, ASRS collects, processes, and analyzes voluntarily submitted aviation incident/situation reports from pilots, flight attendants, air traffic controllers, dispatchers, cabin crew, ground workers, maintenance technicians, and others involved in aviation operations. A sample situation record from the ASRS database is shown in Table 1.

**Table 1 An example incident record from ASRS**

| Attribute | Content |
|---|---|
| Time/day | Date: 201801<br>Local time of day: 0601-1200 |
| Place | Local reference.Airport: ZZZ.Airport<br>State reference: US<br>Altitude.Mean sea level (MSL).Single value: 1800 |
| Environment | Flight conditions: visual flight rules<br>Light: daylight |
| Aircraft | ATC/Advisory: TRACON ZZZ<br>Aircraft operator: air carrier<br>Make model name: B737-800<br>Crew size.Number of crew: 2<br>Operating under FAR Part: Part 121<br>Flight plan: IFR<br>Flight phase: approach<br>Airspace: class B ZZZ |
| Component | Aircraft component: landing gear<br>Problem: malfunctioning |
| Event | Were passengers involved in even: N<br>Detector: flight crew<br>When detected: in-flight<br>Result: general maintenance action |
| Assessment | Contributing factors/situations: aircraft<br>Contributing factors/situations: aircraft<br>Primary problem: human factors |
| Synopsis | B737 Captain reported malfunctioning landing gear |

In ASRS, taking the identification of the causes of an incident as an indicator of civil aviation safety, the target has the characteristic of data imbalance. Namely, the occurrence of the incident is usually caused by aircraft or human factors, but factors like company policy and weather cannot be ruled out, so the data structure always presents a characteristic of imbalance. At the same time, the high-dimensional and primarily categorical data mentioned in the background also pose challenges.

Therefore, a composite model is proposed to clean the data while solving the problem of imbalanced classification. The analysis will help the Civil Aviation Administration (CAA) accurately determine the cause of the incident, to take targeted measures to better manage the civil aviation industry. The airlines and the CAA may utilize this to determine the initial cause of the incidents and minimize the number of workers and their degree of knowledge necessary, increasing civil aviation safety management while decreasing resource waste.

### 2.1 Features of the classification model

The indexes include time, place, environment, aircraft, and event assessments, in detail. The time includes date and period, and the place includes airports, air traffic control (ATC) facilities, intersection, and state. The environment includes flight conditions, lighting, and weather. The aircraft includes federal aviation regulations part, flight plan, flight phase, make and mission, where the mission is people. Event assessment includes event type, detector, primary problem, contributing factors, and result. When the aircraft operator is air carrier, the aircraft is always flying under Federal Aviation Regulations (FAR) 121 part [23].

The primary problem of the occurrence is the categorization model's aim. The following are the most prevalent causes of occurrences in the recent decade: aircraft, human factors, weather, company, environment, ambiguity. Furthermore, unusual causes are consistently classified as others.

The distribution of the causes for all the incidents reported from January 2009 to December 2018 is illustrated in Fig. 1.

Therefore, a composite model is proposed to clean the data while solving the problem of imbalanced classification. The analysis will help the CAA accurately determine the cause of the incident, to take targeted measures to better manage the civil aviation industry.

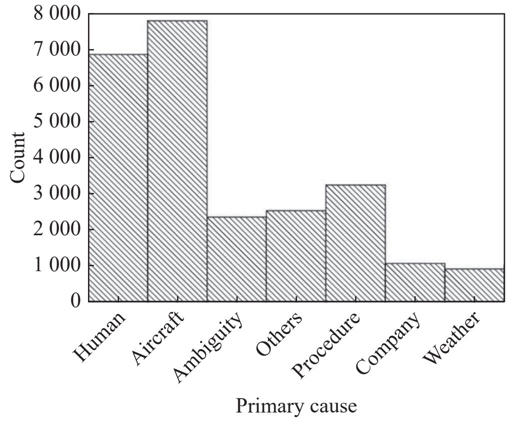



**Fig. 1 Primary causes of incidents in 10 years**

### 2.2 Data pretreatment

#### 2.2.1 Data cleaning

Data cleaning mainly refers to the processing of missing values, unique values, and data filling. For this paper, feature selector [24] with python is used to clean data.

The remaining features after removing the highly missing data include flight phase, detector, above ground level (AGL), crew, local reference, problems, when detected, airspace, flight conditions, ATC, results, anomaly, light, aircraft, contributing causes, flight phase, time, make model name and state.

Then, fill in the original numeric variables according to the average value of each category when filling the missing values and fill the encoded numeric variables according to the mode of the category. After the data is filled, all variables are complete variables, with no missing value.

#### 2.2.2 Feature selection

Feature selection mainly refers to collinear features and feature importance. For each pair of collinear features, the feature that will be removed is the one that comes last in terms of the column ordering in the data frame. No features are high correlated.

Furthermore, feature selection is also selected by feature contribution, and feature importance is selected through eXtreme gradient boosting (XGBoost) and LGBM for the features after data cleaning. The result is shown in Fig. 2. In the two algorithms, crew, discoverer, flight plan, area, and AGL contribute less to the classification model and are removed in the final classification model.

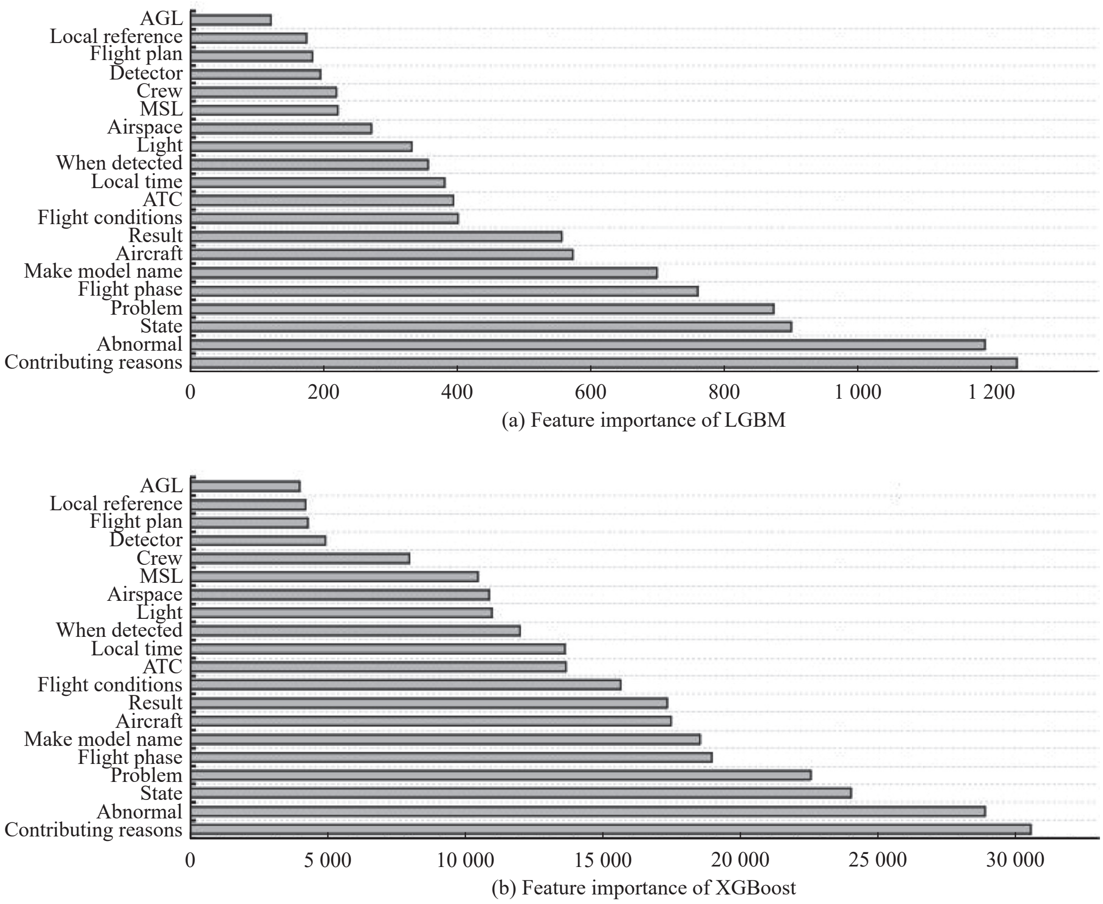



**Fig. 2 Feature contribution of the two methods**

## 3. Methodology

### 3.1 LGBM model

LGBM [25,26] is a new gradient boosting library implemented by Microsoft in April 2017. The LGBM algorithms provide the advantages of being distributed (parallelized), higher efficiency with faster training speeds, lower memory space, better accuracy, and providing scalable solutions.

LGBM aims to make gradient boosting on decision trees faster. This algorithm is used in sorting, classification, regression, and many other machine learning tasks and supports efficient parallel training. The idea is that instead of verifying all of the splits when constructing new leaves, only a subset of them is checked: sort all of the characteristics and bucket the observation into discrete bins. When a leaf in the tree is split, instead of iterating over all of the leaves, all of the buckets are iterated over simply. This implementation is called histogram implementation (as shown in Fig. 3) [26].

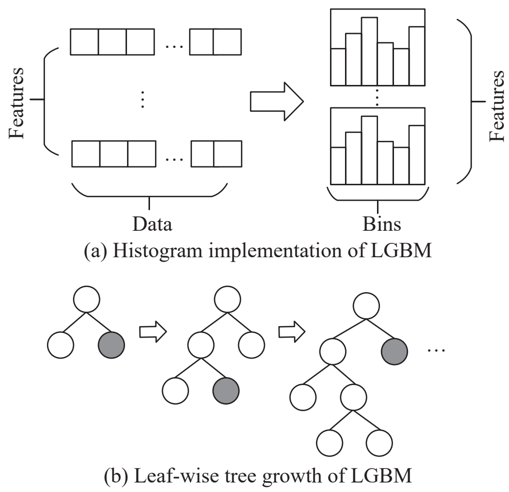



**Fig. 3 Structure of LGBM**

### 3.2 CV

CV [27] is a procedure for estimating the generalization performance. Data is split usually into two parts and based on this splitting, on one part, training is done while the predictive performance is tested on the other part. The training-and-testing scheme works equally well for classification models of machine learning. The schematic diagram of CV is shown in Fig. 4.

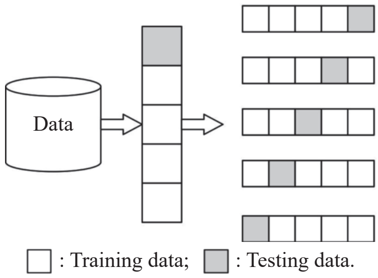



**Fig. 4 Schematic diagram of CV**

We use some of the dataset to train a model and leave some out to test the model once it has been trained. This is the underlying principle in CV. As a result, CV is widely accepted in the data mining and machine learning community and serves as a standard procedure for the sake of model selection or modeling procedure selection.

The method proposed in this paper is to divide the ASRS data according to the proportion of the incident cause, that is, the major causes, aircraft, human factors, and procedure are under-sampling without returning, and the minority causes, weather, company, environment, ambiguity and others are over-sampling with returning. These are the main causes of civil aviation mishaps that have occurred in the recent decade. In essence, it is a sampling method that combines under-sampling and over-sampling. The sample is evenly divided into five subsamples. However, the traditional stratified CV method generally retains the sample an imbalanced situation. The schematic diagram of improved CV is shown in Fig. 5.

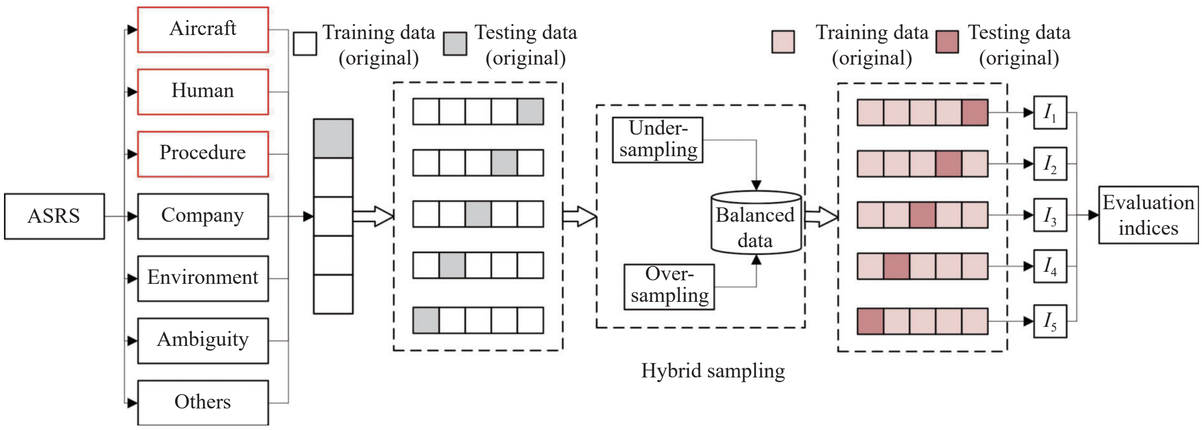



**Fig. 5 Schematic diagram of CV**

It is worth mentioning that the proposed approach is more broad than the usual hybrid sampling method since each subsample is generally balanced and distinct within the sample at this moment.

Appropriate evaluation criteria are the key to evaluate the performance of the classification method. The binary classification system yields the following results: true positive (TP) denotes a right prediction of positive examples as positive, false negative (FN) denotes a prediction of positive examples as incorrect, false positive (FP) denotes a prediction of false examples as positive, and true negative (TN) denotes a prediction of false examples as positive. Referring to the definition of each indicator in the dichotomous classification, a certain category of cause C1 for example, if the confusion matrix can be defined as Table 2.

**Table 2 Confusion matrix of C1**

| Category | C1 | The other cause |
|---|---|---|
| C1 | TP | FP |
| The other cause | FN | TN |

Common evaluation criteria include precision (P), recall (R), accuracy (A), and $F1$ value.

Combined with the actual situation of civil aviation, accuracy is the primary cause of being correctly classified as the proportion of all results. This indicator reflects the discriminative ability of the model from the overall level. The precision refers to the proportion of the correct classification of this category in the model predicted to occur. The recall, which focuses on the model's capacity to identify the primary cause, is the proportion of the main primary that is accurately predicted to all the causes. It represents the model's capacity to distinguish the primary cause. Both the precision and the recall are important reference indicators for judging the civil aviation safety management capability, and $F1$ weights the two to better reflect the model's performance.

The accuracy, precision, recall, and $F1$ of each category are expressed as $A_i$, $P_\mathrm{i}$, $R_i$, and $F1_i$ respectively. The main classification evaluation index in this paper is the weighted average and weight. There are seven categories, aircraft, human, company, weather, procedure, abnormal, and others.

$$A=\sum_{i=1}^{7}\omega_i A_i=\sum_{i=1}^{7}\omega_i\frac{\mathrm{TP}_i+\mathrm{TN}_i}{\mathrm{TP}_i+\mathrm{FP}_i+\mathrm{TN}_i+\mathrm{FN}_i} \tag{1}$$

$$P=\sum_{i=1}^{7}\omega_i P_i=\sum_{i=1}^{7}\omega_i\frac{\mathrm{TP}_i}{\mathrm{TP}_i+\mathrm{FP}_i} \tag{2}$$

$$R=\sum_{i=1}^{7}\omega_i R_i=\sum_{i=1}^{7}\omega_i\frac{\mathrm{TP}_i}{\mathrm{TP}_i+\mathrm{FN}_i} \tag{3}$$

$$F1=\sum_{i=1}^{7}\omega_i F1_i=\sum_{i=1}^{7}\omega_i\frac{2P_iR_i}{P_i+R_i} \tag{4}$$

## 4. Results and discussion

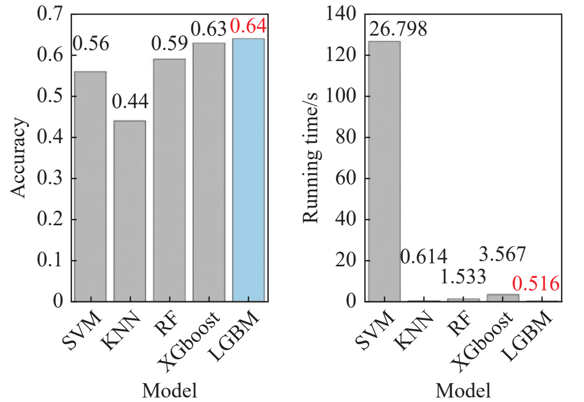
### 4.1 LGBM model

As Fig. 6 illustrates, LGBM has the benefits of high accuracy and rapid speed among the machine learning methods such as support vector machine (SVM), K-nearest neighbor (KNN), and random forest (RF), without being tuned. Therefore, it makes sense to choose LGBM to determine the cause of the incidents.

For complex ASRS data, LGBM can process features quickly and effectively, thus establishing a better incident cause classifier.

### 4.2 Hyperparameter tuning

The hyperparameters tuning involves searching for the best algorithm by adjusting parameters to optimize prediction accuracy [28,29]. The best parameters are identified through trials of several different combinations. The parameters that yield the best performing model are selected for developing a model that is used in the prediction. The current study uses a Bayesian optimization algorithm implemented in the Optuna package to search for the best parameters for LGBM classifiers [30]. The Bayesian optimization algorithm achieves a better performance than random search as it uses past evaluation results to choose the next hyperparameters in the analysis.




(a) Accuracy of models (b) Running time of models

**Fig. 6 Accuracy and runtime of methods**

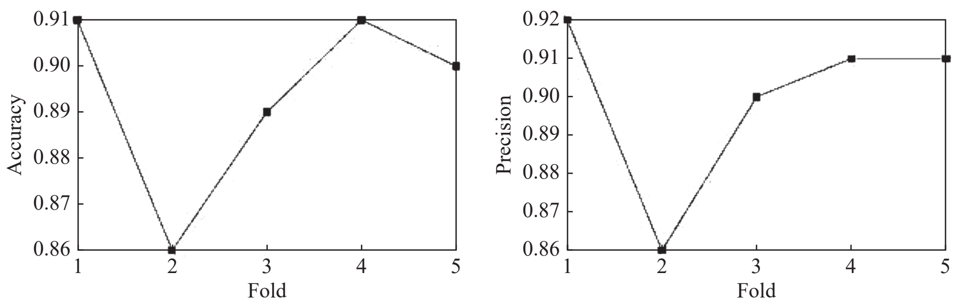
The optimized hyperparameters of the LGBM classifier, range of searches, and final or best parameter values are presented in Table 3.

**Table 3 Set of parameters optimized in LGBM classifier**

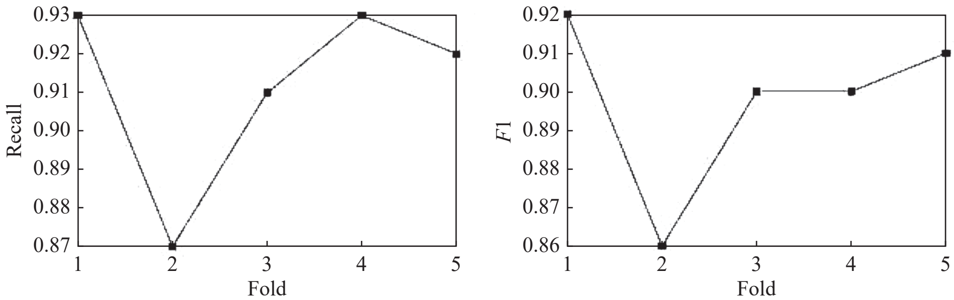
| Parameter | Range of grid | Final parameter | Parameter | Range of grid | Final parameter |
|---|---|---|---|---|---|
| num_leaves | [100, 400] | 143 | feature_fraction | [0.2,1] | 0.63 |
| max_bin | [50, 300] | 147 | bagging_fraction | [0.2,1] | 1 |
| max_depth | [4,20] | 6 | bagging_freq | [1,7] | 7 |
| min_child_weight | [1,6] | 4 | learning_rate | [5, 100] | 68 |
| $\lambda_1$ | [0.001,1] | 0 | learning_rate | [0.01,0.5] | 0.04 |
| $\lambda_2$ | [0.001,1] | 0.75 | n_estimators | (500, 5 000) | 1 045 |

### 4.3 Classification

The improved five-fold LGBM-CV with hybrid sampling (LGBM-HSCV) is employed, whereby five datasets after hybrid sampling for balance are divided and four sets are used for training while the resulting model is validated with the remaining set of the data. The evaluation indices of each fold are shown in Fig. 7.




(a) Accuracy (b) Precision




(c) Recall (d) $F1$

**Fig. 7 Evaluation indeices of LGBM-HSCV**

Then evaluation indices are the average of all folds, that is, the accuracy of the LGBM-HSCV is 0.896, the precision is 0.901, the recall is 0.912 and the $F1$ value is 0.899.

Besides, the confusion matrix is shown in Fig. 8. The classification result of the category is also the average of all folds.

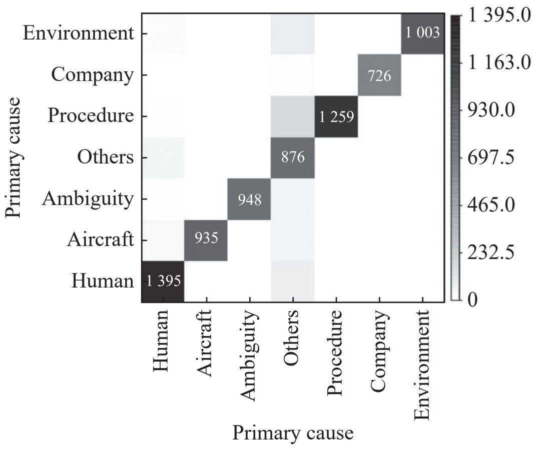


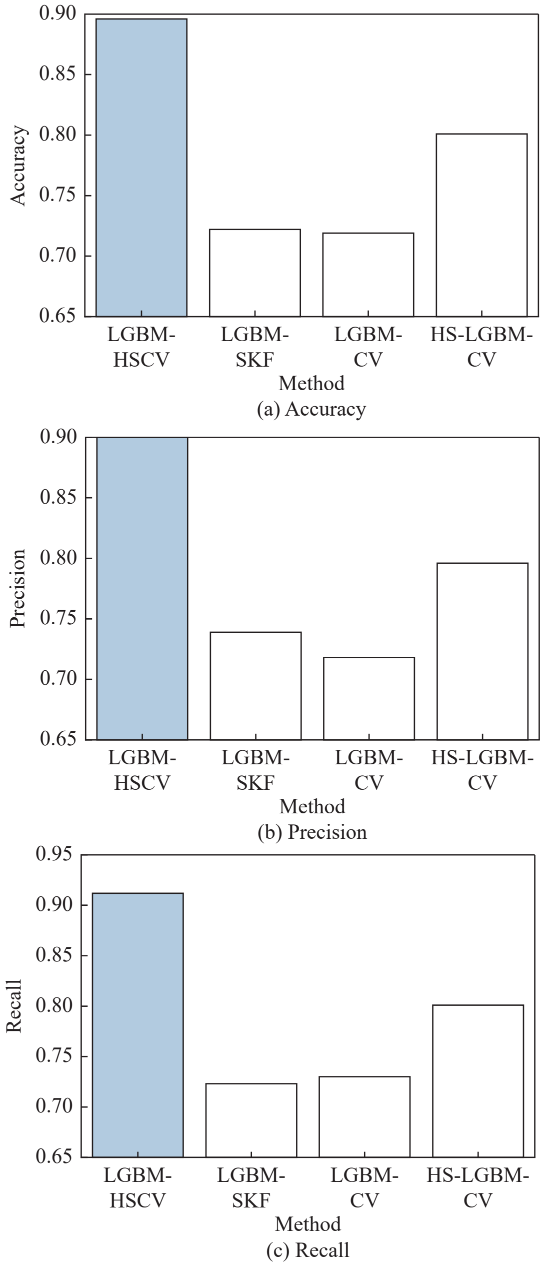

**Fig. 8 Confusion matrix of LGBM-HSCV**

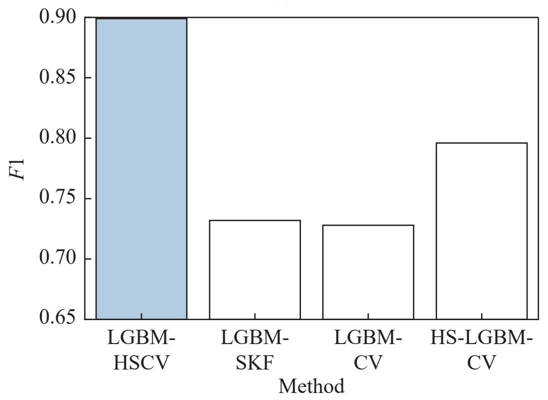
It can be seen that the method proposed in this paper has greatly improved the recognition ability of the minority categories. At the same time, various evaluation indicators also show that it is a good classifier, which can effectively identify the cause of civil aviation incidents.

## 4.4 Comparison

### 4.4.1 Cross validation

This article compares three LGBM models with different CV methods: LGBM-CV, which is the regular $k$-fold CV; LGBM-SKF, which is the general stratified $k$-fold CV; and LGBM-HSCV, which is the hybrid sampling $k$-fold CV. In particular, the hybrid sampling followed by CV approach, HS-LGBM-CV, is also utilized for comparison, and the results prove that the method proposed in this paper is still the best. The comparisons are shown in Fig. 9.





(d) $F1$

**Fig. 9 Evaluation indices of LGBM with different CV models**

The results show that among the three CV methods LGBM-CV, LGBM-SKF, and LGBM-HSCV, LGBM-HSCV is the best method. Therefore, to determine whether it is due to balancing the data before mixing the samples, LGBM-HSCV and HS-LGBM-CV are compared. The results show that although the hybrid sampling to achieve balance improves the accuracy, the performance of LGBM-HSCV put forward in this paper is still better, that is, balancing the sub-sample after $k$-fold is better than balancing the sample and then dividing into $k$ folds.

4.4.2 *K*-fold

The influence with different folds of the LGBM-HSCV is also compared in Fig. 10. The comparison of different folds shows that five-fold is the best.

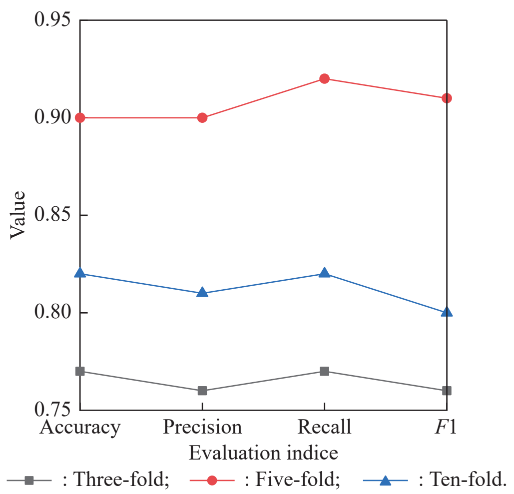



**Fig. 10 Evaluation indeices with different folds**

We also double-check on this basis. Secondary CV is supposed to be performed for each balanced fold, with the fold result being the average of the acquired secondary CV (LGBM-HSCV-CV) results, and then the full five-fold model being averaged. The LGBM-HSCV-CV model's evaluation indices are listed in Table 4. The indices show that without further fold, LGBM-HSCV is appropriate.

**Table 4 Evaluation indices of LGBM-HSCV-CV and LGBM-HSCV**

| Model | Accuracy | Precision | Recall | $F1$ |
|---|---|---|---|---|
| LGBM-HSCV-CV | 0.897 | 0.899 | 0.897 | 0.891 |
| LGBM-HSCV | 0.896 | 0.901 | 0.912 | 0.899 |

## 5. Summary and conclusions

This work develops an improved CV model based on LGBM for analyzing the causes of incidents in the air transportation system. The LGBM model is trained with categorical and numerical data, and the improvement is based on CV hybrid sampling. By using the LGBM-HSCV model, we formulate a classifier model to identify the causes of hazardous events outcomes using 24 753 reports on incidents/accidents reported from January 2009 to December 2018.

Several contributions have been made to this paper. First, the ASRS data are collected and identified in accordance with commercial transportation. Second, following data cleaning and feature selection, a classifier model based on LGBM is created to extract the difficult ASRS data. Third, to improve the LGBM model, Optuna with Bayes is employed for hyperparameter adjustment. Finally, to improve the categorization, an innovative CV algorithm is designed. This approach is compared to several CV methods in particular. It is also compared to the same procedure with different $k$ folds at the same time. The findings show that the developed classifier model outperforms the individual models in terms of accuracy, precision, recall, and $F1$ score when using hybrid sampling $k$-fold CV to balance the ASRS data. This improved classifier will be able to better analyze similar incident reports in the future, resulting in more efficient and accurate incident cause categorization and improved civil aviation safety management.

## References

[1] INSUA D R, ALFARO C, GOMEZ J, et al. Forecasting and assessing consequences of aviation safety occurrences. Safety Science, 2019, 111: 243–252.

[2] PRAMONO A, MIDDLETON J H, CAPONECCHIA C. Civil aviation occurrences in Indonesia. Journal of Advanced Transportation, 2020, 2020(4): 3240764.

[3] TILLER L N, BLISS J P. Categorization of near-collision close calls reported to the aviation safety reporting system. Proceedings of the Human Factors and Ergonomics Society

Annual Meeting, 2017, 61(1): 1866–1870.

[4] ZHANG X G, MAHADEVAN S. Ensemble machine learning models for aviation incident risk prediction. Decision Support Systems, 2019, 116: 48–63.

[5] KUHN K D. Using structural topic modeling to identify latent topics and trends in aviation incident reports. Transportation Research Part C: Emerging Technologies, 2018, 87: 105–122.

[6] REN B, YUE Z F, SI Y, et al. Novel method of aviation safety causality prediction based on forecast. Systems Engineering and Electronics, 2023, 45(3): 762−768. (in Chinese)

[7] Aviation Safety Reporting System. Program briefing. https: asrs.arc.nasa.gov/docs/ASRS_ProgramBriefing.pdf.

[8] CHUNG S H, MA H L, HANSEN M, et al. Data science and analytics in aviation. Transportation Research Part E: Logistics and Transportation Review, 2020, 134: 101837.

[9] HU X, WU J, HE J R. Textual indicator extraction from aviation accident reports. Proc. of the AIAA Aviation 2019 Forum, 2019: 2939.

[10] SCHLOGL M, STUTZ R, LAAHA G, et al. A comparison of statistical learning methods for deriving determining factors of accident occurrence from an imbalanced high resolution dataset. Accident Analysis & Prevention, 2019, 127: 134–149.

[11] JEONG H, JANG Y, BOWMAN P J, et al. Classification of motor vehicle crash injury severity: a hybrid approach for imbalanced data. Accident Analysis & Prevention, 2018, 120: 250–261.

[12] CHRISTENSEN I. The elements of a commercial human spaceflight safety reporting system. Acta Astronautica, 2017, 139: 228–232.

[13] ZHANG C H, KONG L, XU Q, et al. Fault diagnosis of key components in the rotating machinery based on Fourier transform multi-filter decomposition and optimized LGBM. Measurement Science and Technology, 2020, 32(1): 015004.

[14] PARSA A B, MOVAHEDI A, TAGHIPOUR H, et al. Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. Accident Analysis & Prevention, 2020, 136: 105405.

[15] SHI X, WONG Y D, LI M Z F, et al. A feature learning approach based on XGBoost for driving assessment and risk prediction. Accident Analysis & Prevention, 2019, 129: 170–179.

[16] MACHADO M R, KARRY S, DE S I T. LGBM: an effective decision tree gradient boosting method to predict customer loyalty in the finance industry. Proc. of the 14th International Conference on Computer Science & Education, 2019: 1111−1116.

[17] CHEN C, ZHANG Q M, MA Q, et al. LGBM-PPI: predicting protein-protein interactions through LGBM with multi-information fusion. Chemometrics and Intelligent Laboratory Systems, 2019, 191: 54–64.

[18] TAN Y J, JIN B, CUI Q, et al. Generalizing fault detection against domain shifts using stratification-aware cross-validation. https://arxiv.org.abs/2008.08713.

[19] WANG Q, DU J L, LI N. Robust description and recognition of various viewpoint dynamic textures. Journal of Systems Engineering and Electronics, 2016, 27(4): 920–925.

[20] YADAV S, SHUKLA S. Analysis of k-fold cross-validation over hold-out validation on colossal datasets for quality classification. Proc. of the 6th International Conference on Advanced Computing, 2016: 78−83.

[21] XIONG M L, WANG H W, XU Y, et. al. General aviation safety research based on prediction of bird strike symptom. Systems Engineering and Electronics, 2020, 42(9): 2033−2040. (in Chinese)

[22] MOT (Ministry of Transport of the People's Republic of China). Annual implementation plan for strengthening transportation safety production. http://www.gov.cn/zhengce/zhengceku/2022-04/09/content_5684264.htm.

[23] Federal Aviation Administration. Title 14 of the Code of Federal Regulations (14 CFR) part 121 air carriers. Washington D.C.: United States Department of Transport, 2020.

[24] WILL K. Feature selector: simple feature selection in Python. https://github.com/WillKoehrsen/feature-selector.

[25] UKKONEN P, MAKELA A. Evaluation of machine learning classifiers for predicting deep convection. Journal of Advances in Modeling Earth Systems, 2019, 11(6): 1784–1802.

[26] KE G L, MENG Q, FINLEY T, et al. LGBM: a highly efficient gradient boosting decision tree. Advances in Neural Information Processing Systems, 2017: 3146−3154.

[27] BROWNE M W. Cross-validation methods. Journal of Mathematical Psychology, 2000, 44(1): 108–132.

[28] PUTATUNDA S, RAMA K. A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost. Proc. of the International Conference on Signal Processing and Machine Learning, 2018: 6−10.

[29] MASSAOUDI M, REFAAT S S, CHIHI I, et al. A novel stacked generalization ensemble-based hybrid LGBM-XGB-MLP model for short-term load forecasting. Energy, 2020, 214: 118874.

[30] AKIBA T, SANO S, YANASE T, et al. Optuna: a next-generation hyperparameter optimization framework. Proc. of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2019: 2623−2631.

## Biographies

**NI Xiaomei** was born in 1992. She received her Ph.D. degree in aviation engineering from Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2022. She is a lecturer in School of Aeronautical Engineering, Nanjing Vocational University of Industry Technology, Nanjing, China. Her research interests are civil aviation safety engineering and aircraft reliability.
E-mail: 905271004@qq.com

**WANG Huawei** was born in 1974. She received her Ph.D. degree from National University of Defense Technology, Changsha, China in 2003. She is a professor of the School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China. Her research interests include prognostic and health management, civil aviation safety engineering, and aircraft reliability engineering.
E-mail: wang_hw66@163.com

**CHEN Lingzi** was born in 1997. She received her M.S. degree from Nanjing University of Aeronautics and Astronautics in 2022. She is an engineer of the 41st Research Institute of China Electronics Technology Group Corporation. Her research interest is civil aviation safety engineering.
E-mail: 13258035636@163.com

**LIN Ruiguan** was born in 1993. He received his M.S. degree in aviation engineering from Shenyang Aerospace University, Shenyang, China, in 2019. He is currently pursuing his Ph.D. degree at the School of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research interests are civil aviation safety engineering and aircraft reliability.
E-mail: 478636604@qq.com