Journal of Electronic Research and Application, 2026, Volume 10, Issue 2
http://ojs.bbwpublisher.com/index.php/JERA
ISSN Online: 2208-3510
ISSN Print: 2208-3502

IVY-RF-Based Logistics Claims Risk Classification and Prediction

Shufeng Wang*

School of Computer Science and Technology, Taiyuan Normal University, Shanxi 030619, China

**Author to whom correspondence should be addressed.*

Abstract: With the rapid development of e-commerce, the scale and complexity of logistics operations continue to increase, and the claims risks caused by abnormal events such as cargo damage, delays, and loss during transportation are becoming increasingly prominent. The traditional claims processing model, primarily based on manual review, is no longer adequate for the high-frequency, large-scale business demands in terms of processing efficiency, decision consistency, and cost control. There is an urgent need to introduce intelligent methods to achieve accurate identification and hierarchical management of claims risks. Addressing the challenges of diverse feature dimensions, highly imbalanced category distribution, and difficulty in distinguishing different risk types in logistics claims data, this paper proposes a Random Forest Logistics Claims Risk Classification Model (IVY-RF) based on the IVY growth optimization algorithm. This method uses a random forest as the basic classifier, fully leveraging its advantages in nonlinear relationship modeling and feature interaction capture. It also introduces the IVY metaheuristic optimization algorithm to adaptively optimize the model's key hyperparameters globally. Experimental results based on real-world logistics claims datasets demonstrate that the IVY-RF model significantly outperforms comparable models such as IVY-LightGBM and IVY-XGBoost in core evaluation metrics, including macro-average F1 score, weighted precision, and weighted recall, achieving a better performance balance between the majority and minority high-risk categories. The findings indicate that the proposed IVY-RF model exhibits significant advantages in prediction accuracy, stability, and engineering feasibility, providing reliable technical support for logistics companies to conduct intelligent identification and refined management of claims risks.

Keywords: Logistics claims; Anomaly detection; XGBoost; Random forest; Ivy optimization algorithm

Online publication: March 27, 2026

1. Introduction

In recent years, with the rapid development of e-commerce and the continuous growth of online shopping behavior of consumers, the global logistics industry has shown a significant trend of scale and complexity [1]. China's express delivery volume has ranked among the top in the world for many consecutive years. In 2024, the total volume of business exceeded 150 billion pieces. The high-frequency operation of the logistics network has significantly improved the overall efficiency of the industry. However, the continuous expansion of the logistics business scale has also led to the continuous accumulation of potential risks in the transportation process. Abnormal events such as damage, delay and loss of goods occur frequently, and the number of claims disputes caused by this has continued to rise. Logistics claims are not only directly related to the operating costs and profit margins of enterprises, but also have a profound impact on customer satisfaction and brand reputation. They have become a key component of the risk management system of modern logistics enterprises [2].

In the traditional claims processing process, companies usually rely on manual review to judge the rationality of claims. With the rapid increase in the number of claims work orders, the manual review model has gradually exposed problems such as low processing efficiency, strong subjectivity of review standards and high labor costs. On one hand, a large number of routine claims requests occupy the main energy of reviewers, making it difficult to identify abnormal cases with potential high risks in a timely manner. On the other hand, the differences in experience and judgment standards among different reviewers can easily lead to inconsistent handling results for similar claims, affecting the objectivity and repeatability of claims decisions. Therefore, building an efficient, stable and interpretable intelligent claims risk identification model has become an urgent need for the digital transformation and refined management of the logistics industry [3].

Against this backdrop, introducing machine learning technology into the logistics claims risk classification task provides a new technical path for improving the automation level and scientific nature of claims decision-making [4]. However, logistics claims data itself has significant complexity and multi-source heterogeneity. The data usually contains a mixed feature structure with both numerical and categorical variables, such as insured amount, transportation timeliness, commodity category, delivery node and customer historical behavior [5]. Alternatively, there are often complex nonlinear interaction relationships between different features, such as the risk performance differences of specific commodity types under different transportation routes, or the abnormal claims patterns of some outlets under high business load conditions. These characteristics place high demands on the nonlinear modeling ability, feature interaction capture ability and overall robustness of the model.

Compared with deep learning models that are highly sensitive to parameters or have highly complex structures, the random forest algorithm based on ensemble learning shows good stability and generalization ability when processing small-to-medium scale structured data [6]. By constructing multiple decision trees and integrating their prediction results, the random forest effectively reduces the uncertainty of a single model and captures the nonlinear relationship between features without complex feature engineering. In addition, the feature importance evaluation mechanism provided by the random forest also provides a good interpretability basis for the analysis of logistics claims risk factors and business decision support.

However, the classification performance of random forest models largely depends on the hyperparameter configuration [7]. Inappropriate parameter settings may lead to limited model performance or wasted computational resources. To this end, this paper introduces the IVY growth optimization algorithm to adaptively tune the key hyperparameters of the random forest. The IVY algorithm simulates the coordinated growth and phototropic climbing behavior of plants in the environment, and achieves a dynamic balance between global exploration and local optimization during the search process, thereby effectively improving the classification performance and convergence stability of the model. Based on this, an IVY-optimized random forest logistics claims risk classification model is constructed, providing an efficient and feasible solution for risk identification and decision support in complex logistics claims scenarios.

2. Literature review

Risk classification and prediction in logistics claims are receiving increasing attention, and various methods are being explored in different fields. Several studies have highlighted the application of machine learning (ML) technology in improving the accuracy of risk assessment and operational efficiency. For example, Zhang *et al.* used machine learning algorithms to analyze risk factors related to risk assessment in logistics claims and proposed an analysis method based on logistic regression model [8]. They proposed a method for separating urban logistics drones based on trajectory prediction, emphasizing the importance of accurate prediction models for the safety of autonomous vehicles. Their method calculates the necessary separation distance based on trajectory prediction accuracy, demonstrating how predictive modeling can reduce operational risks in logistics scenarios. Recent advances include the application of federated learning frameworks, such as the Fed-SCRP model proposed by Lu *et al.*, which is used for seller claim risk prediction in logistics scenarios [9]. This method addresses data privacy issues while using multi-view learning to improve the prediction accuracy of distributed datasets. Moreover, Liu *et al.* introduced a decoupled graph neural network with pseudo-labels to enhance insurance claim prediction, thereby addressing challenges such as label scarcity and heterogeneity [10]. Their innovative GClaim model demonstrates how advanced neural network architectures can improve claim risk classification under complex data patterns. In addition to models for claims, Wang *et al.* also studied ergonomic risk prediction for manual handling tasks (a common activity in logistics) [11]. They combined video-based posture tracking with machine learning classifiers to assess ergonomic risks, demonstrating the application of multimedia data and machine learning techniques in operational risk prediction.

In summary, these studies demonstrate a wide range of machine learning and statistical methods for classifying and predicting logistics claims and operational risks. However, existing methods are mostly designed for specific application scenarios (such as single logistics claims, drone flight path safety, or ergonomic risks), and their model design is highly dependent on domain assumptions, making it difficult to achieve unified modeling and generalization across different types of logistics claim risks.

3. Materials and methods

3.1. IVY optimization algorithm

Ivy Growth Optimization Algorithm (IVYA) is a metaheuristic optimization method inspired by nature, which is inspired by the adaptive growth behavior of ivy in the natural environment [12]. Instead of relying on a single search mechanism, IVYA simulates multiple stages in the ivy life cycle, including early exploration, growth adaptation, phototactic movement, reproduction and dispersal, and survival-based selection. In this framework, each candidate solution is conceptualized as an ivy branch that dynamically explores the solution space, gradually adjusts its growth direction, and moves to a more favorable region to find the global optimum.

IVYA's underlying design is based on Ivy Growth Optimization's (IVY) remarkable ability to efficiently occupy space and acquire resources in complex and ever-changing environments. IVY can continuously sense external signals such as light, adjust its growth rate, climb surrounding structures, and reproduce in favorable locations [13]. These adaptive strategies enable ivy to balance exploration and utilization throughout its growth process. Accordingly, IVYA organizes its optimization process into five interconnected stages, each of which mathematically simulates a different biological behavior, thus forming a coherent and effective search framework for complex optimization problems.

The IVY algorithm is a novel natural heuristic metaheuristic optimization algorithm, inspired by the growth behavior and survival strategies of ivy plants [14]. This algorithm achieves efficient solutions to complex optimization problems by simulating the adaptive growth mechanism of ivy in the natural environment. The core framework of the IVY algorithm includes the following five stages.

3.1.1. Population initialization

The algorithm first randomly generates an initial population within the search space, with each individual representing a set of candidate solutions. The initialization process adopts a uniform distribution strategy to ensure that the population can widely cover the entire search space and avoid the algorithm getting trapped in local optima due to excessive concentration of initial solutions. This decentralized initial layout lays the foundation for subsequent global exploration and ensures that the algorithm has sufficient search capabilities in the early stages of optimization.

3.1.2. Coordinated growth phase

The coordinated growth phase simulates the biological characteristics of ivy in its natural environment, where it adaptively adjusts its growth rate according to conditions such as light, temperature, and nutrients. The algorithm maintains a dynamic growth vector for each individual, updating this vector by incorporating historical growth trends and random perturbations. This mechanism allows the population to adaptively expand its search range while maintaining search momentum, effectively preserving population diversity and suppressing premature convergence, preventing the algorithm from stagnating near suboptimal solutions in the early stages of optimization.

3.1.3. Phototropic climbing stage

Ivy exhibits significant phototropism in nature, tending to grow towards areas with ample sunlight to obtain more energy for photosynthesis. The IVY algorithm translates this biological characteristic into a neighborhood-guided local search mechanism. Each individual will update its position based on its relative position to its neighbors, moving along a unit direction vector pointing to its neighbors, while simultaneously adding random perturbations based on its historical growth vector. This mechanism cleverly balances local development (learning from high-quality neighbors) and random exploration (normally distributed perturbations), enhancing the algorithm's local optimization capabilities.

3.1.4. The propagation and evolution phase

To simulate the natural tendency of ivy to spread and establish itself in favorable areas, the algorithm introduces a global guidance mechanism, driving each individual to move towards the current global optimum. This phase determines the new position of an individual by combining a random factor and a growth vector, and simultaneously updates the growth vector to support subsequent directional searches. This mechanism ensures that the population as a whole steadily approaches the global optimum while maintaining appropriate diversity during evolution.

3.1.5. Survivor selection phase

Drawing on the biological principles of natural selection, the IVY algorithm employs a fitness-based survival selection mechanism. An individual's survival to the next generation depends on the relative relationship between its fitness value and the current optimal solution; selection pressure is adaptively controlled through a dynamic adjustment factor. This hybrid retention strategy allows elite individuals to coexist with those of moderate fitness, effectively avoiding local optima traps while accelerating convergence, achieving a dynamic balance between exploration and exploitation.

The IVY algorithm achieves a balance between global search capability and local fine-tuning in the hyperparameter space by organically coordinating four mechanisms: growth, light-climbing, propagation evolution, and survival selection. This provides an efficient and robust solution paradigm for complex optimization problems.

3.2. Random forest model

Random forest (RF) is a supervised machine learning algorithm based on the idea of ensemble learning. It was proposed by Breiman. Its core idea is to build multiple independent decision trees and integrate the prediction results of each sub-model to improve the generalization ability and prediction stability of the overall model [15]. Compared with single decision tree models, RF can effectively alleviate the problem of overfitting of decision trees and shows good robustness when dealing with structured data with small and medium scale, high dimension and significant nonlinear features.

In the RF model, each decision tree is constructed by randomly sampling samples from the original training set through bootstrap sampling. The remaining unsampled samples can be used as out-of-bag (OOB) samples for model error estimation. In the process of splitting the nodes of the decision tree, RF does not use all features for optimal splitting, but randomly selects some features from the feature subset to participate in the splitting decision [16]. This dual random mechanism of “sample randomization + feature randomization” effectively reduces the correlation between different decision trees, enabling the ensemble model to significantly reduce variance while maintaining low bias.

In logistics claims risk classification scenarios, data typically exhibits characteristics such as the coexistence of multiple feature types, complex nonlinear interactions between features, and uneven category distribution. RF does not require strict assumptions about data distribution, can directly handle numerical and discrete features, and naturally model high-order nonlinear relationships through tree structures, thus possessing strong adaptability. Furthermore, random forests can output feature importance indices to measure the relative contribution of different input features in the model’s decision-making process, providing interpretability support for logistics claims risk factor analysis and business decision-making.

However, the performance of RF models largely depends on their hyperparameter configuration, such as the number of decision trees, maximum tree depth, minimum number of split samples per node, and feature subset selection strategy. Inappropriate parameter settings can lead to underfitting, overfitting, or wasted computational resources, thus limiting their performance in practical applications. Therefore, how to efficiently and stably optimize the hyperparameters of random forest models is a key issue in further improving their prediction accuracy and generalization ability in logistics claims risk classification tasks.

Based on the above considerations, this paper introduces the IVY growth optimization algorithm to adaptively optimize the key hyperparameters of the random forest model, and constructs the IVY-RF logistics claims risk classification model.

3.3. IVY-RF model

Figure 1 shows the overall technical framework of the IVY-RF proposed in this paper. The framework adopts a top-down pipeline modeling process, which can be divided into four core stages: “data preprocessing-hyperparameter optimization-model training-classification prediction” [17]. Each stage is interconnected and progressively advances to form a complete logistics claim risk classification modeling system.

In the data preprocessing stage, the original waybill and claim data are first cleaned and standardized, mainly including missing value filling, outlier correction and categorical feature encoding conversion, in order to reduce the interference of data noise on model training and improve the consistency of feature expression [18]. Subsequently, a training dataset with relatively balanced categories is constructed to provide a fairer data foundation for subsequent model learning.

Building upon this foundation, the model incorporates the IVY growth optimization algorithm to perform global search and adaptive optimization of the key hyperparameters of the random forest classifier. This process simulates the “coordinated growth” and “light-guided climbing” behavior of ivy in complex environments, achieving a dynamic balance between global exploration and local development in the hyperparameter space, thus effectively avoiding getting trapped in local optima. The optimized hyperparameters mainly include the number of decision trees, maximum tree depth, and minimum number of split samples, which have a significant impact on model performance.

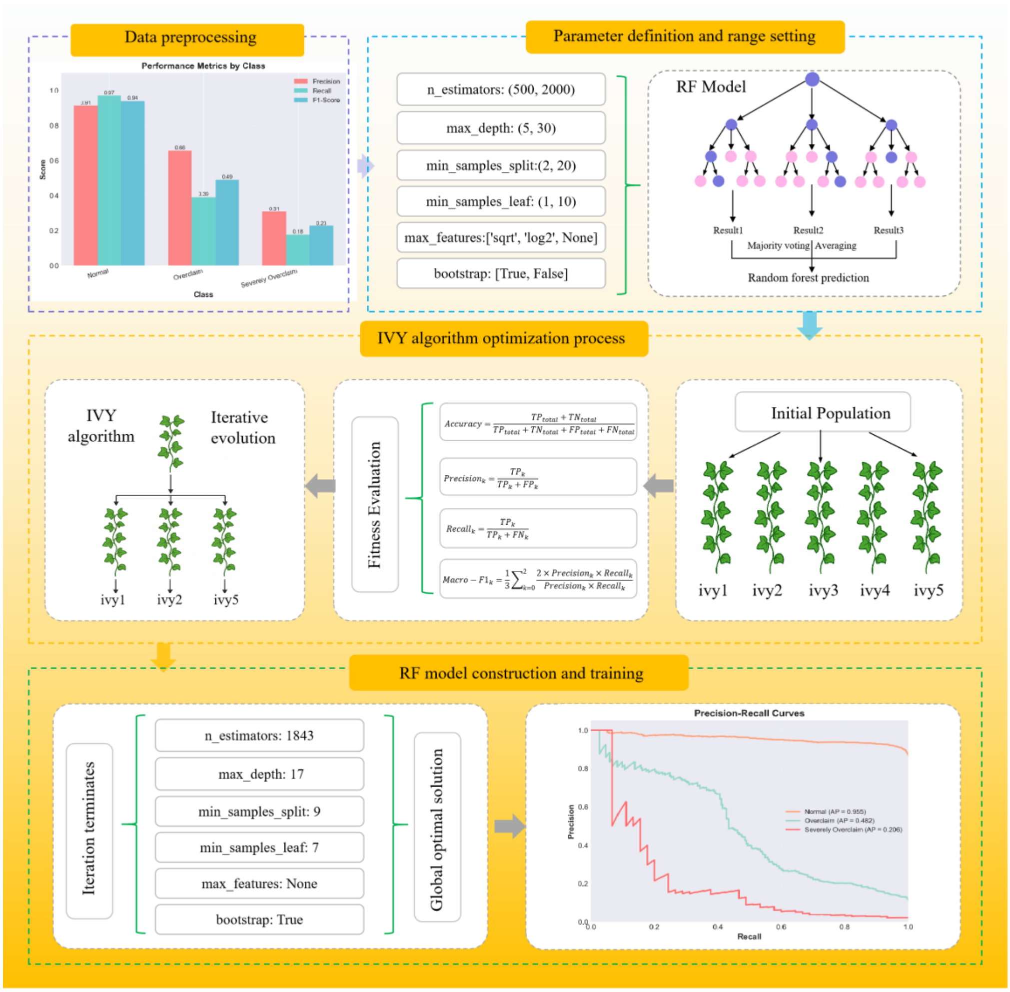


Figure 1. Flowchart of IVY-RF model.

Finally, after obtaining the optimal hyperparameter configuration, the optimized random forest model was trained and applied to the test set samples to achieve accurate classification and prediction of different logistics claim risk levels. Through the above multi-stage collaborative design, the IVY-RF model ensures predictive performance while taking into account stability, robustness, and engineering feasibility, providing an effective technical framework for intelligent assessment of logistics claim risks.

3.4. Evaluation indicators

This article uses the following core evaluation indicator formula:

$$Accuracy = \frac{TP_{total} + TN_{total}}{TP_{total} + TN_{total} + FP_{total} + FN_{total}}$$

$$Precision_k = \frac{TP_k}{TP_k + FP_k}$$

$$Recall_k = \frac{TP_k}{TP_k + FN_k}$$

$$Macro - F1 = \frac{1}{3}\sum_{k=0}^{2} \frac{2 \times Precision_k \times Recall_k}{Precision_k + Recall_k}$$

Where TP_k, TN_k, FP_k and FN_k are the number of true positive, true negative, false positive, and false negative samples in class k, respectively; TP_{total}, TN_{total}, FP_{total} and FN_{total} are the total values for the three classes.

4. Results and discussion

The clean dataset, preprocessed in the second question, is loaded. This dataset has already undergone missing value imputation, outlier correction, feature standardization, and LabelEncoder encoding. The dataset is then divided into a feature matrix X and a target variable y (i.e., risk category labels: "Normal," "Moderate Excess," and "Severe Excess"). Due to the significantly imbalanced distribution of the three labels, to avoid biased model training towards the majority class. The sampling ratio is set to make the number of samples in each class approximately 1:0.2:0.1, thus constructing dataset D_{smote}.

To further improve the generalization performance and classification accuracy of the model, the IVY algorithm is used to adaptively optimize the key hyperparameters of the RF classifier. The IVY algorithm simulates the "growth-diffusion" mechanism of IVY, performing global optimization and local optimization within the hyperparameter space. For the core hyperparameters of the RF, such as n_estimators, max_depth, min_samples_split, min_samples_leaf, max_features, and bootstrap, the optimal hyperparameter configurations obtained after iterative optimization within the specified hyperparameter space using the IVY algorithm are shown in **Table 1**.

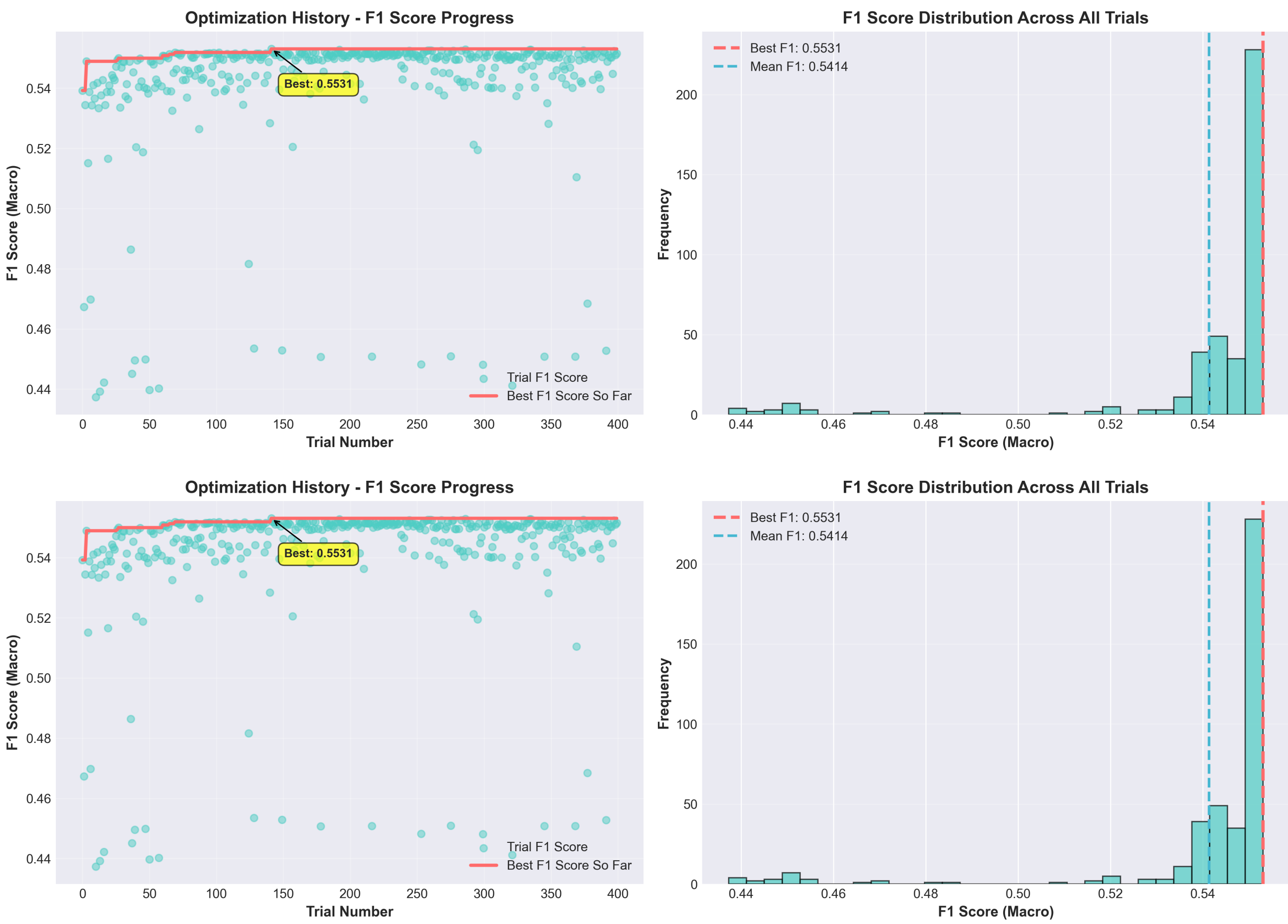
Table 1. Optimal values for hyperparameters

Hyperparameter name	Value
n_estimators	617
max_depth	27
min_samples_split	7
min_samples_leaf	6
max_features	log2
bootstrap	False

This section provides a detailed analysis of the experimental results of the IVY-based optimized RF classifier, including the overall model performance, the performance of each class, the optimal hyperparameter information, and the interpretation of the visualization results, to verify the effectiveness of the proposed method in the three-class classification task of logistics claims risk.

Figure 2 illustrates the hyperparameter optimization process and F1 score distribution of three models: IVY-RF, IVY-XGBoost, and IVY-LightGBM. During hyperparameter optimization (top image), the IVY-RF model rapidly converged and stabilized at a macro-average F1 score of 0.5531, with most experiments showing F1 scores above 0.54, the best F1 score reaching 0.5531, and the mean F1 score at 0.5414. In contrast, IVY-XGBoost (middle image, best F1 score only 0.5348, mean 0.5113) and IVY-LightGBM (bottom image, lower overall score distribution) show that IVY-RF not only significantly outperforms in terms of the best F1 score but also exhibits a more concentrated distribution of F1 scores in the high-value range, fully demonstrating its superior classification performance after hyperparameter optimization.

In terms of convergence speed, IVY-RF stabilized at the optimal F1 score in only about 150 iterations, far faster than other models, indicating that the IVY algorithm is more efficient in optimizing RF hyperparameters. Analysis of score distribution density shows that IVY-RF has a significantly higher proportion of F1 scores above 0.54 compared to the other two models, demonstrating clear advantages in performance stability and reliability. Considering the high accuracy requirements of waybill risk labeling, IVY-RF can more accurately distinguish between the three risk labels, providing stronger technical support for risk management at the business end, further highlighting its applicability and superiority in this task.



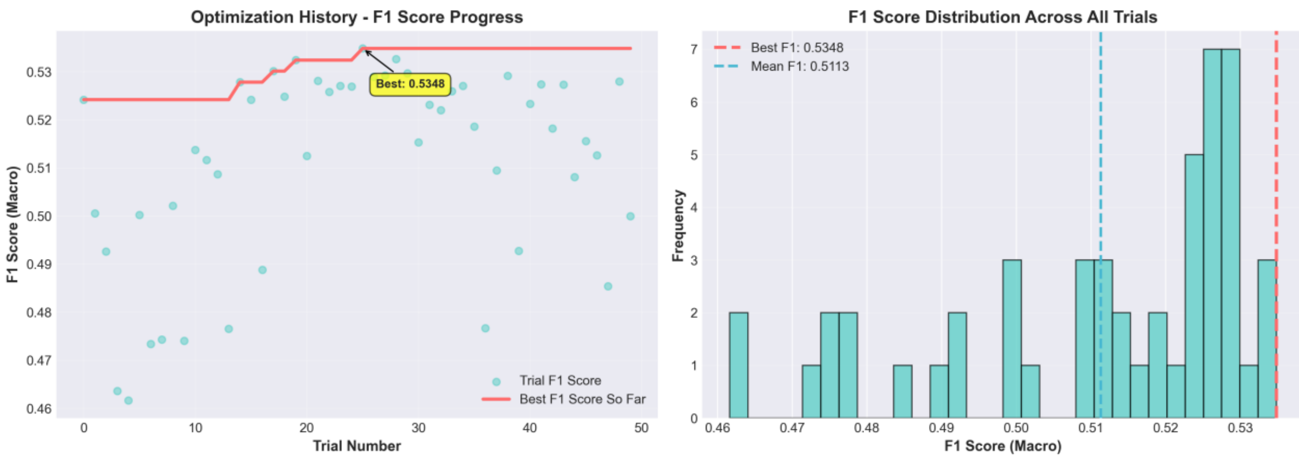
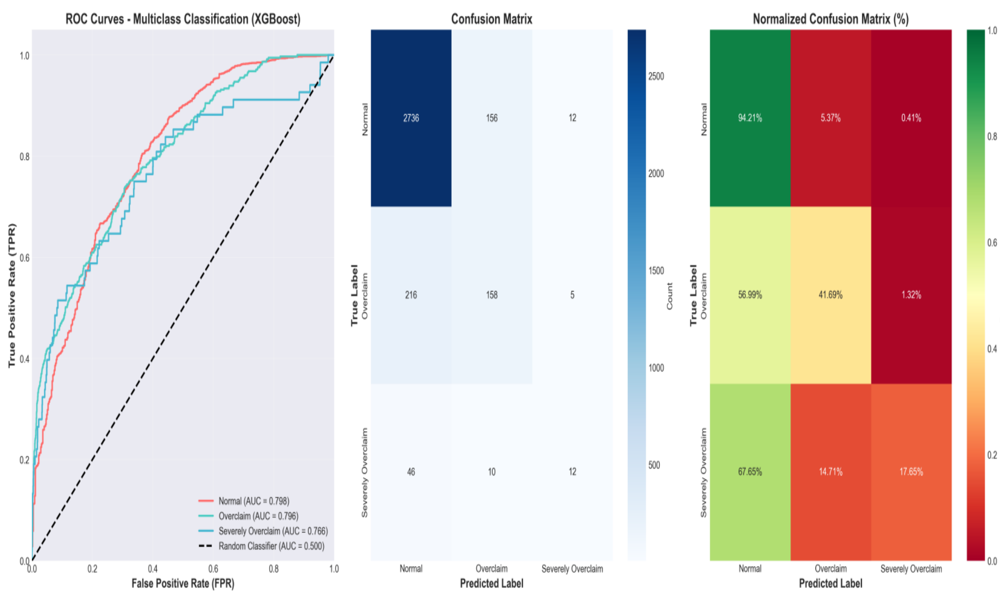


Figure 2. Model hyperparameter optimization process and F1 score distribution (rom top to bottom: IVY-RF, IVY-XGBoost, IVY-LightGBM).

Figure 3 presents the classification performance visualization results of three models: IVY-RF, IVY-LightGBM, and IVY-XGBoost, including ROC curves and confusion matrices (the models are listed from top to bottom). In the IVY-RF ROC curve, the AUC for the Normal class reaches 0.798, the Overclaim class 0.796, and the Severely Overclaim class 0.765, demonstrating outstanding class discrimination ability. Its confusion matrix shows that the Normal class recognition accuracy is as high as 96.95%, and the Overclaim and Severely Overclaim classes also exhibit good classification accuracy. Compared to IVY-LightGBM (whose AUC values across all categories are generally lower than IVY-RF, and whose category recognition rates in the confusion matrix are significantly different) and IVY-XGBoost (whose AUC values in the ROC curve are even lower, and whose classification accuracy in the confusion matrix is insufficient), IVY-RF performs better in both the AUC value of the ROC curve and the category recognition accuracy in the confusion matrix. This fully demonstrates its ability to accurately distinguish between the three types of risk labels on waybills, and it has a significant advantage in the comprehensiveness and reliability of its classification performance.



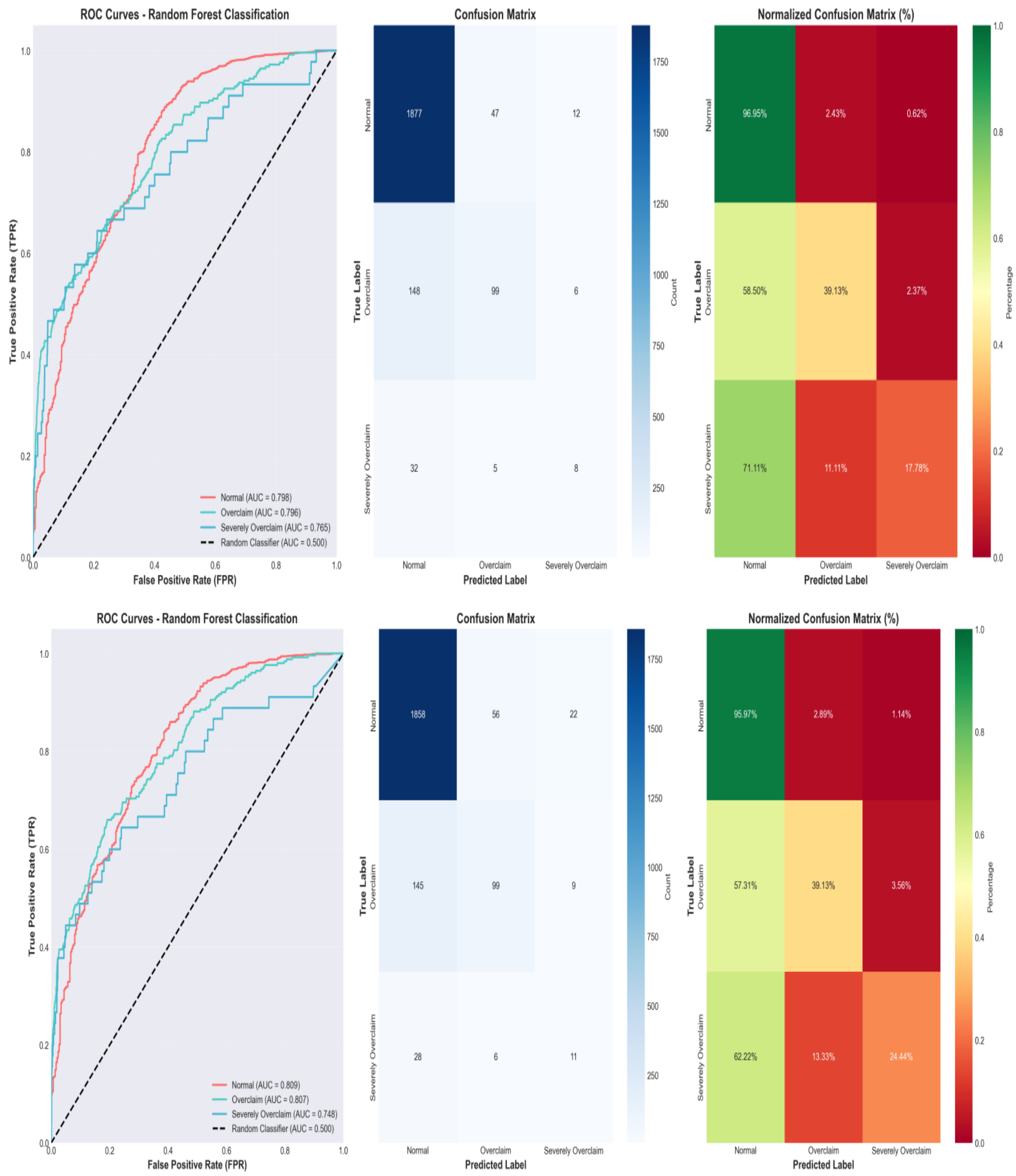


Figure 3. Visualization of model classification performance (ROC curves and confusion matrix. From top to bottom: IVY-RF, IVY-LightGBM, IVY-XGBoost).

Figure 4 contains two subplots: Performance Metrics by Class and Test Set Distribution by Class, visually presenting the class performance and data distribution characteristics of Ivy-RF. In the left subplot, Performance Metrics by Class, the Normal class achieved a precision of 0.91, recall of 0.97, and an F1 score of 0.94, demonstrating extremely strong recognition capabilities for the majority class; the Overclaim class achieved a precision of 0.66, recall of 0.39, and an F1 score of 0.49; and the Severely Overclaim class achieved a precision of 0.31, recall of 0.18, and an F1 score of 0.23. The right subplot, Test Set Distribution by Class, clearly shows the class distribution pattern of the test set, where Normal classes accounted for a high 86.7%, Overclaim classes 11.3%, and Severely Overclaim classes only 2.0%. This extremely imbalanced sample distribution is the key factor leading to the significant performance degradation of Overclaim and Severely Overclaim classes compared to Normal classes, and also reveals the performance challenges of the model in niche risk categories from a data perspective.

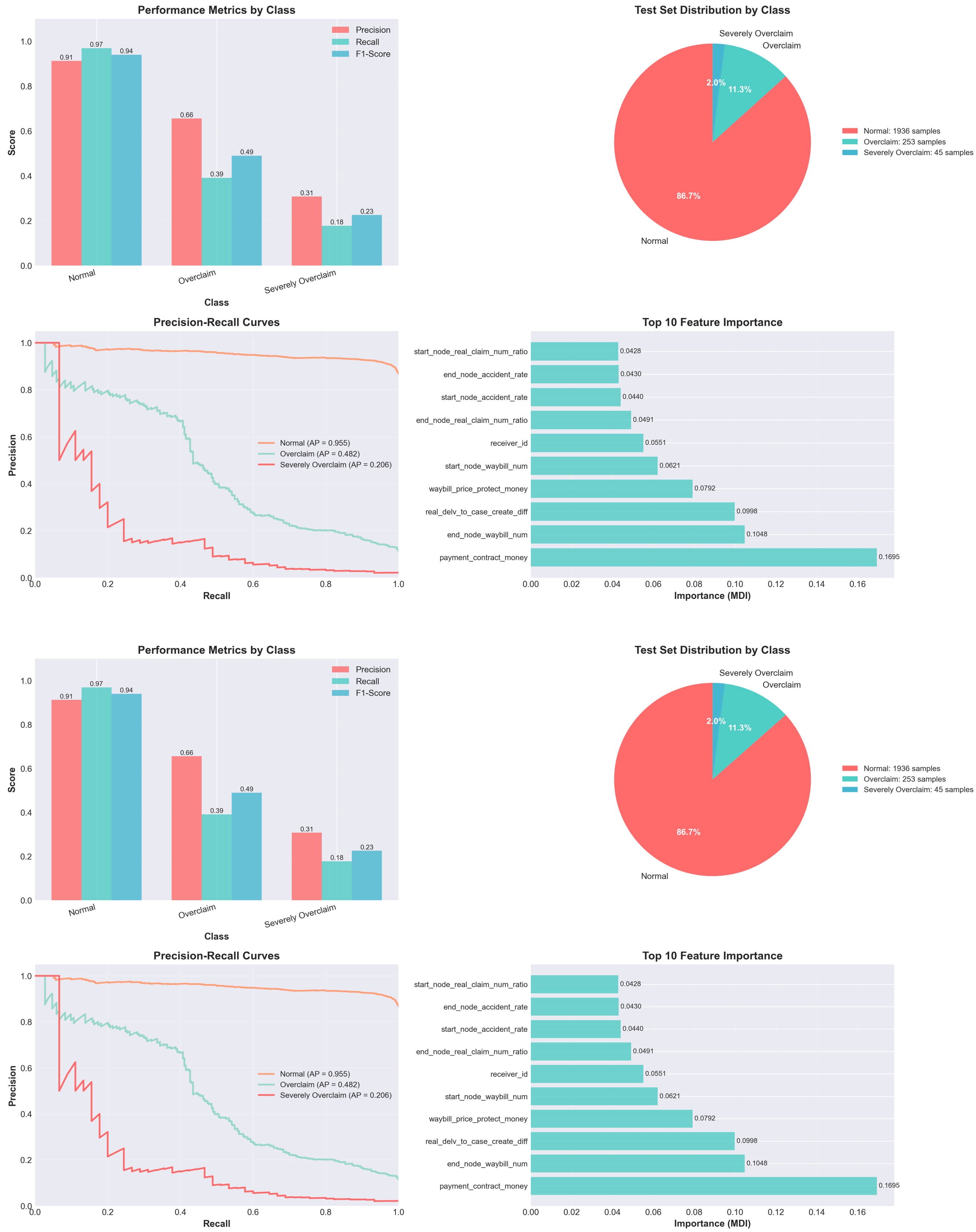


Figure 4. IVY-RF category performance and data distribution.

As can be seen from the comparison of the category metrics and weighted total metrics of IVY-RF, IVY-LightGBM, and IVY-XGBoost in **Table 2**, IVY-RF has a significant advantage in core performance dimensions. In the Normal category (reasonable claims), its F1 score of 0.940 far exceeds IVY-LightGBM's 0.923 and IVY-XGBoost's 0.927, showing better synergy between precision and recall. In the Overclaim category (overly high claims), its F1 score of 0.490 is higher than IVY-LightGBM's 0.463 and IVY-XGBoost's 0.449, demonstrating stronger robustness in identifying this category. Even in the Severely Overclaim category (which accounts for only 2.0%), its F1 score of 0.225 is on par with or even better than the other two models. In terms of weighted overall metrics, IVY-RF leads the pack with a weighted precision of 0.871, a weighted recall of 0.888, and a weighted F1 score of 0.874. This demonstrates that IVY-RF achieves a superior balance between accurate identification of the majority class and effective differentiation of the minority class when addressing the challenge of class imbalance in waybill risk labeling tasks. Its comprehensive and reliable classification performance is significantly better than IVY-LightGBM and IVY-XGBoost.

Table 2. Comparison of performance indicators for waybill risk classification of different IVY optimization models

Model	Various indicators	Precision	Recall	F1-score
IVY-LightGBM	Normal	0.938	0.902	0.923
	Overclaim	0.432	0.504	0.463
	Severely overclaim	0.201	0.243	0.226
	Weighted total	0.857	0.8447	0.850
IVY-XGBoost	Normal	0.912	0.942	0.927
	Overclaim	0.487	0.416	0.449
	Severely overclaim	0.413	0.176	0.247
	Weighted total	0.8544	0.8672	0.8593
IVY-RF	Normal	0.912	0.969	0.940
	Overclaim	0.655	0.391	0.490
	Severely overclaim	0.307	0.177	0.225
	Weighted total	0.871	0.888	0.874

5. Conclusion

To address the challenges of complex risk types, highly imbalanced sample distribution, and insufficient efficiency and consistency of traditional manual review in logistics claims processing, this paper proposes an IVY-RF based on the IVY growth optimization algorithm. This model uses a random forest as the base classifier, fully leveraging its advantages in nonlinear modeling, feature interaction capture, and robustness. The IVY optimization algorithm is introduced to adaptively optimize key hyperparameters globally, effectively improving the model's classification performance and stability. Experimental results on real-world logistics claims datasets demonstrate that the IVY-RF model significantly outperforms comparable models such as IVY-LightGBM and IVY-XGBoost in three-class classification tasks. IVY-RF exhibits superior overall performance in macro-average F1 score, weighted precision and recall, as well as in class discrimination ability reflected by the ROC curve and confusion matrix.

Particularly in the high-proportion normal claims category and the medium-risk excess claims category, the model achieves high recognition accuracy and stability. Even with a very small number of severely excess claims, IVY-RF maintains relatively reasonable classification ability, demonstrating its robustness in extreme class imbalance scenarios. In summary, the proposed IVY-RF model not only demonstrates significant advantages in predictive performance but also balances model stability, interpretability, and engineering feasibility, providing effective technical support for logistics companies to achieve proactive identification and refined management of claims risks. Future research can further expand into areas such as multi-source heterogeneous data fusion, temporal feature modeling, and federated learning to enhance the model's adaptability and generalization capabilities in more complex logistics business scenarios

Disclosure statement

The author declares no conflict of interest.

References

[1] Govindan K, Soleimani H, 2017, A Review of Reverse Logistics and Closed-Loop Supply Chains. Journal of Cleaner Production, 2017(142): 371–384.

[2] Berg C, Deichmann U, Liu Y, et al., 2017, Transport Policies and Development. The Journal of Development Studies, 53(4): 465–480.

[3] Van Hoek R, 2020, Research Opportunities for a More Resilient Post-COVID-19 Supply Chain: Closing the Gap between Research Findings and Industry Practice. International Journal of Operations & Production Management, 40(4): 341–355.

[4] Gao Y, Jiang Y, Peng Y, et al., 2025, Medical Image Segmentation: A Comprehensive Review of Deep Learning-Based Methods. Tomography, 11(5): 52.

[5] Koukopoulos A, Sani G, 2014, DSM-5 Criteria for Depression with Mixed Features: A Farewell to Mixed Depression. Acta Psychiatrica Scandinavica, 129(1): 4–16.

[6] Yuan F, Zuo Z, Jiang Y, et al., 2025, AI-Driven Optimization of Blockchain Scalability, Security, and Privacy Protection. Algorithms, 18(5): 263.

[7] Song Y, Lu Y, 2015, Decision Tree Methods: Applications for Classification and Prediction. Shanghai Archives of Psychiatry, 2015.

[8] Saraswat B, Singhal A, Agarwal S, et al., 2023, Insurance Claim Analysis Using Traditional Machine Learning Algorithms, 2023 International Conference on Disruptive Technologies.

[9] Miao T, Zhang S, Zhang Y, et al., 2024, Combined DeRitis Ratio and Alkaline Phosphatase on The Prediction of Portal Vein Tumor Thrombosis in Patients with Hepatocellular Carcinoma. Med. Public and Global Health.

[10] Zhang J, Li Z, Luo X, et al., 2024, Study of Urban Unmanned Aerial Vehicle Separation in Free Flight Based on Track Prediction. Applied Sciences, 2024.

[11] Lin Y, Xu X, Chen S, 2024, Construction of Nomogram Based on Clinical Factors for The Risk Prediction of Postoperative Complications in Children with Choledochal Cyst. Frontiers in Pediatrics, 2024.

[12] Zhang K, Yuan F, Jiang Y, et al., 2025, A Particle Swarm Optimization-Guided Ivy Algorithm for Global Optimization Problems. Biomimetics, 10(5): 342.

[13] Ghasemi M, Zare M, Trojovský P, et al., 2024, Optimization based on the Smart Behavior of Plants with its

Engineering Applications: Ivy Algorithm. Knowledge-Based Systems, 2024(295): 111850.

[14] Padon O, McMillan K, Panda A, et al., 2016, Ivy: Safety Verification by Interactive Generalization, Proceedings of the 37th ACM SIGPLAN Conference on Programming Language Design and Implementation, 614–630.

[15] Belgiu M, Drăguţ L, 2016, Random Forest in Remote Sensing: A Review of Applications and Future Directions. ISPRS Journal of Photogrammetry and Remote Sensing, 2016(114): 24–31.

[16] Wang K, Truong K, 2025, A Video-Based Assessment Tool Using Machine Learning for Ergonomic Risk Prediction in Manual Lifting Tasks. Ergonomics, 2025.

[17] Ouyang H, Li W, Gao F, et al., 2024, Research on Fault Diagnosis of Ship Diesel Generator System Based on IVY-RF. Energies (19961073), 17(22).

[18] Huang S, Li C, Zhou J, et al., 2025, Use of BOIvy Optimization Algorithm-Based Machine Learning Models in Predicting the Compressive Strength of Bentonite Plastic Concrete. Materials, 2025.