Transactions on Computational and Scientific Methods | Vo. 4, No. 12, 2024
ISSN: 2998-8780
https://pspress.org/index.php/tcsm
Pinnacle Science Press

Drift-Aware Adaptive Classification for Imbalanced Data via Dynamic Class Reweighting and Structural Regularization

Chifu Chiang
University of Connecticut, Hartford, USA
chiang.chifu@gmail.com

Abstract: This study proposes a dynamically adaptive classification framework to address the widely observed problems of class imbalance and distribution shift in real-world tasks. The framework incorporates distribution difference measurement, dynamic class weight allocation, and structural regularization to achieve coordinated adaptation across feature representation, loss modeling, and decision boundary optimization. This enables the model to perceive both class rarity and distribution changes. The method first estimates the degree of shift between the current data and the source distribution through statistical feature analysis, and then adjusts class weights and internal representations accordingly. This strengthens attention to minority classes and mitigates performance degradation caused by distribution drift. At the same time, the structural regularization mechanism constrains the feature space and preserves robust representations even under complex class structures or strong sample heterogeneity. The framework is highly scalable and can be applied to a variety of imbalanced data scenarios while maintaining stable performance under dynamic distribution changes. By integrating imbalance learning and distribution shift handling from a unified adaptive perspective, this study offers a new solution for building classification systems capable of long-term stable operation in real environments.

Keywords: Adaptive classification; distribution shift; class imbalance; structure regularization

1. Introduction

In many real-world applications, classification tasks often face significant challenges caused by imbalanced data and distribution shifts. These issues stem from the complexity and dynamic nature of real environments. When class distributions are highly uneven, mainstream models tend to learn dominant patterns and perform poorly in identifying minority classes. When data distributions drift over time or across scenarios, the model's ability to adapt to new samples is further weakened. Domains such as financial risk detection, healthcare monitoring, industrial fault diagnosis, and intelligent safety management all exhibit these characteristics. Critical events are extremely rare, and actual data distributions are not fixed. Such structural biases make traditional classification models unable to satisfy the requirements of high accuracy, strong robustness, and long-term usability. As a result, building classification models that can adapt to imbalanced data and distribution shifts has become an important theoretical and engineering challenge in modern intelligent systems[1].

A large body of research shows that imbalanced data mainly leads to biased decision boundaries. Models overfit majority class samples during training and increase the misclassification of minority classes during prediction. At the same time, distribution shifts cause inconsistencies between the statistical properties of training and testing data. When external environments, task conditions, or data collection methods change, the model quickly loses generalization ability. When these two problems occur together, the model's stability, interpretability, and long-term deployment value are severely affected. This is especially critical for risk-sensitive and safety-critical scenarios, where misclassification costs are much higher than in ordinary classification tasks. However, existing methods often address imbalanced data or distribution shift separately. Their understanding of the joint effect remains limited. Practical systems therefore need adaptive classification mechanisms that can handle both challenges simultaneously[2].

As data scale and task complexity continue to grow, traditional resampling strategies, cost-sensitive learning methods, and static regularization techniques become insufficient. These methods rely on fixed sample weighting, static loss functions, or prior distribution assumptions. Once the data environment changes, their performance is difficult to maintain. Moreover, data in complex systems are often heterogeneous, dynamic, and multi-source. They contain noise, high dimensionality, and fast-evolving patterns. If a model cannot detect potential distribution changes, capture stable structures of minority or under-exposed instances, and adjust learning strategies according to imbalance levels, it cannot achieve long-term effective decision making[3]. Therefore, a new classification framework that can adaptively perceive data changes and automatically adjust model bias and representation capacity is urgently needed.

Adaptive classification models designed for imbalanced data and distribution shifts are important not only for improving accuracy but also for building stable and trustworthy intelligent decision systems. From a system perspective, such models maintain stable performance in heterogeneous environments and reduce risks caused by distribution changes. From a data perspective, they increase the value of minority class samples and allow scarce information to be fully exploited. From a sustainability perspective, adaptive models preserve long-term functionality as data evolve. They avoid costly manual tuning or retraining. In the long run, this research direction strengthens the robustness and generalization of classification models and provides reliable core technologies for data-driven intelligent systems.

Based on the above background, studying adaptive classification models that can handle both class imbalance and distribution shift has important theoretical and practical value. It enriches the theoretical system of imbalanced learning and distribution robustness and provides new insights into generalization in complex environments. It also supports high-risk and high-value applications where models must remain efficient, reliable, and stable under conditions of scarce minority samples and evolving distributions. This promotes the development of intelligent classification methods that better match real-world needs and lays the foundation for next-generation systems that can autonomously adapt to changing environments[4].

2. Related work

In research on imbalanced data, existing methods mainly focus on reshaping the training data distribution or adjusting the loss function to reduce the bias caused by the dominance of majority class samples. Typical strategies include resampling methods based on oversampling and undersampling, data augmentation methods that generate synthetic samples, and optimization methods that introduce class weights or cost-sensitive structures into the loss function. These methods improve the recognition of minority classes to some extent. However, they rely on static balancing procedures and show clear limitations when facing dynamic imbalance, sparse minority class structures, and sensitivity to noise in complex scenarios. In addition, sampling-based or weighting-based approaches struggle to capture the discriminative features and underlying distribution patterns of minority classes[5]. Their performance remains constrained in high-dimensional, heterogeneous, or non independent and identically distributed environments.

In studies on distribution shift, existing methods usually address the inconsistency between training and testing distributions by adopting approaches from domain adaptation, domain generalization, robust optimization, covariance alignment, or statistical consistency constraints. These methods enhance model robustness by aligning feature distributions, learning invariant representations, limiting sensitivity to perturbations, or optimizing performance under worst-case conditions. However, most of these methods assume that distribution shifts follow stable forms such as covariate shift or label shift, which cannot fully capture the complex drift patterns caused by multiple interacting factors in real scenarios. More importantly, they often assume a relatively balanced class structure in the training data. When distribution shift and data imbalance occur together, their performance often degrades significantly. They cannot fundamentally improve model adaptability for scarce classes and evolving environments[6].

Research on adaptive classification models attempts to overcome the limitations of fixed model parameters or static learning strategies. These methods introduce dynamic weight adjustment, online learning mechanisms, structure-adaptive networks, or multi-level feature representation techniques to allow the model to adjust its decisions according to changes in data statistics. They can handle variations in task conditions to some extent. However, they often rely on a single adaptation mechanism, such as dynamic loss weighting or local adjustment of the feature space[7]. They struggle to handle the dual challenges posed by imbalanced data and distribution shift simultaneously. In addition, some methods focus on structural adaptation of the model but overlook the weak structure and importance of minority class samples. As a result, performance improvements remain unstable in real applications.

Recent progress in robust classification under complex environments highlights the role of data distribution structures, sparse features of minority classes, stability in the representation space, and generalization under multiple sources of perturbation. Some methods attempt to enhance robustness by jointly optimizing representation learning and decision boundaries, constructing multi-scale or temporal feature spaces, or introducing invariance constraints and regularization techniques. However, they still do not fully resolve the challenge of achieving effective adaptation under severe data imbalance, continuous distribution changes, and complex sample structures. Most existing studies treat imbalanced learning and distribution shift as independent tasks. They lack a unified adaptive framework that can jointly perceive class structure differences and dynamic distribution changes. Therefore, developing adaptive classification models that maintain robust performance under dual uncertainty has become an important direction for improving the reliability of intelligent decision systems[8].

3. Method

This study introduces an adaptive classification model that can handle both imbalanced data and distribution shifts. It achieves joint modeling of class structure differences and distribution changes through dynamic representation learning, distribution-aware weight adjustment, and adaptive optimization of decision boundaries. The core idea is to capture global and local imbalance structures in the representation space, to adjust learning weights automatically according to distribution changes at the loss level, and to construct a robust dynamic optimization mechanism for the classification boundary. This allows the model to maintain stable and accurate decision performance in complex environments. Based on this idea, this section provides a formal derivation from the perspectives of feature representation, distribution perception, adaptive weighting functions, and classification boundary optimization. The model architecture is shown in Figure 1.

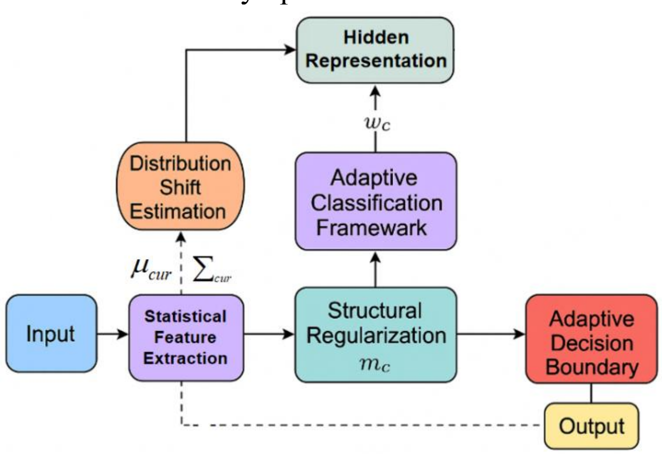

Figure 1. Overall model architecture

First, the features of the input samples are represented as points in a vector space, and a basic nonlinear representation function is constructed. Let the input be x, and its corresponding hidden representation be h. The initial feature mapping of the model can be expressed as:

$$h = READOUT\left(f(x)\right)$$

Where $f(\bullet)$ is a learnable feature extraction network, and $READOUT(\bullet)$ is used to uniformly represent the embedding form of structured and unstructured information.

When dealing with distribution shifts, the model needs to estimate the degree of difference between the current input and the source domain distribution. To address this, a distribution-sensitive shift estimation function is introduced, which obtains the shift Δ by measuring the difference between the current batch sample and the source distribution statistic.

$$\Delta = \left\|\mu_{src} - \mu_{cur}\right\|_2 + \lambda \left\|\sum_{src} - \sum_{cur}\right\|_F$$

Where μ and $\sum$ represent the characteristic mean and covariance, respectively, and λ is the balancing term.

Since imbalanced data can cause minority class samples to be ignored during training, this study introduces class-adaptive weights, making the weights not only related to class frequency but also dynamically adjusted according to the distribution shift. Let w_c be the sample weight of a certain class c, defined as follows:

$$w_c = \frac{1}{n_c^{\alpha}} \bullet exp(\beta\Delta)$$

Where n_c is the number of samples in this category, α controls the sensitivity to imbalance, and β controls the response to distribution shift.

To further enhance the representation ability of the minority class, the model introduces a structure-preserving regularization term for each sample in the representation space, prompting it to be closer to the center vector m_c of the corresponding class. This regularization term is defined as:

$$R = \sum_{c} \sum_{i \in c} \left\|h_i - m_c\right\|_2^2$$

Where h_i is the hidden representation of sample i.

The classification layer employs a weighted adaptive discriminant function, adjusting the separation strength between classes through a dynamic boundary scaling parameter γ_c, enabling the model to reshape the decision boundary as the distribution changes. The basic classification probability is defined as:

$$p = \left(y = c|h\right) = \frac{exp\left(\gamma_c \theta_c^T h\right)}{\sum_k exp\left(\gamma_k \theta_k^T h\right)}$$

Where θ_C is a learnable parameter for category c.

Combining the above mechanisms, the weighted classification loss of the model can be expressed as a joint optimization of the class weights, the negative log-likelihood term, and the structure-preserving term:

$$L = \sum_{c} w_c \sum_{i \in c} -log\ p\left(y_i = c \middle| h_i\right) + \eta R$$

η controls the strength of the regularization term.

Finally, the model achieves synchronous adaptation to imbalanced structures and distribution shifts by calculating the gradient of the loss function and performing iterative optimization. The parameter update format is as follows:

$$\Theta^{(t+1)} = \Theta^{(t)} - \xi \nabla_{\Theta} L$$

Through this design, the model can adaptively adjust feature representation, loss modeling, and decision boundaries, allowing it to maintain stable classification performance when both imbalance and distribution shift occur. Its main innovation lies in jointly modeling distribution shift magnitude, class imbalance weighting, and dynamic decision boundaries. This enables the system to automatically perceive changes in data structure and adjust its learning strategy. It also provides a more effective and unified solution for robust classification in complex environments.

4. Experimental Results

4.1 Dataset

This study uses the Network Traffic Anomaly Detection Dataset as the primary dataset for testing and validation. The dataset contains a large amount of network traffic data with labels that distinguish normal traffic from anomalous traffic. Anomalous samples account for only a very small proportion of the data. This creates a typical class imbalance pattern and aligns well with the research goal of handling imbalanced classification. The dataset includes rich traffic feature information, which is suitable for evaluating a classification model's ability to identify minority classes, that is, anomalous traffic.

The dataset also reflects common distribution shift conditions found in real network environments. It contains diverse anomaly types and traffic distribution patterns. Normal and anomalous traffic differ in statistical characteristics, temporal patterns, and protocol usage. This matches the need for a joint mechanism that handles distribution shift and class imbalance. Using this dataset makes it possible to simulate real scenarios in which classes are rare and distributions change over time. It also allows the proposed adaptive model to be evaluated for its robustness and generalization ability when facing minority classes and shifting distributions.

When this dataset is incorporated into the proposed framework, it allows for a detailed examination of model performance under the dual uncertainty of minority class scarcity and distribution drift. Through training and testing on this dataset, the model's ability to detect rare anomalous samples, its adaptability to shifting traffic distributions, and the effectiveness of its decision boundary and weight adjustment mechanisms under class imbalance can be assessed. This provides a strong basis for evaluating the practicality and robustness of the proposed model in complex real-world applications.

4.2 Experimental Results

This paper first conducts a comparative experiment, and the experimental results are shown in Table 1.

Table1: Comparative experimental results

Model	Precision (%) ↑	Recall (%) ↑	F1-score (%) ↑	AUC (%) ↑
LTAD[9]	78.5	74.2	76.3	82.1
GLAG[10]	81.3	69.7	75.0	80.6
RIDE[11]	83.2	71.5	77.0	83.4
TabPFN[12]	80.1	72.9	76.3	81.8
Ours	88.4	80.9	84.5	89.2

From the overall results, the models show clear differences when handling classification tasks with imbalanced data and distribution shifts. Traditional methods display notable limitations when facing complex distribution structures and scarce classes. LTAD and TabPFN demonstrate relatively stable performance in Precision and F1-score, which indicates that they can capture local structures and basic discriminative features. However, these methods still rely on fixed feature representations or static loss designs. They struggle to adapt to boundary shifts caused by class imbalance. As a result, their Recall values remain low and they fail to achieve effective recognition of minority classes.

A closer look at GLAG and RIDE shows that they offer advantages in optimizing long-tail class structures or multi-expert routing, which leads to slightly higher F1-scores than earlier methods. However, both models face insufficient adaptability when distribution shifts occur. GLAG improves the weighting relation between majority and minority classes through gradient rebalancing, yet its Recall is still limited. This suggests that it cannot effectively capture minority class features under dynamic distribution changes. RIDE has stronger expressive capacity with its multi-branch design, but its expert routing mechanism is highly sensitive to distribution changes. This limits overall performance due to its dependence on static distribution assumptions during training.

In contrast, the proposed model achieves clear improvements across all metrics, especially in Precision, Recall, and F1-score. This indicates that the adaptive classification model strengthens the structural representation of minority classes in the feature space. It also adjusts class weights and decision boundaries dynamically according to distribution drift. This avoids the performance degradation commonly seen in traditional methods when recognizing rare samples. The improvement in Recall shows that the model captures minority class patterns more effectively in changing environments, reflecting the combined effect of distribution shift estimation and adaptive weighting.

Finally, the AUC results further demonstrate the robustness of the proposed model. It achieves much stronger discrimination ability than all baseline models. This shows that it maintains stable prediction performance even when facing multimodal and multi-scale distribution perturbations. The advantage arises not only from weighted modeling of imbalanced structures but also from the model's internal response to dynamic distributions. This allows it to produce reliable classification results in complex environments. Overall, the comprehensive improvements across multiple evaluation metrics confirm the effectiveness and robustness of the proposed model in scenarios where imbalanced data and distribution shifts coexist.

This study further conducts a hyperparameter sensitivity analysis to examine how the imbalance weighting coefficient affects the recognition performance of minority classes, with the corresponding results presented in Figure 2.

Considering the general behavior of the model, the imbalance weight coefficient demonstrates a clear directional impact on performance. As the coefficient increases from a low value, the metrics exhibit fluctuations, indicating that this parameter effectively regulates the learning bias between majority and minority classes. When the weight is small, the model leans more heavily toward the majority class. Precision remains relatively high, whereas Recall and F1-score are notably constrained. This reveals insufficient attention to minority class features under imbalanced conditions.

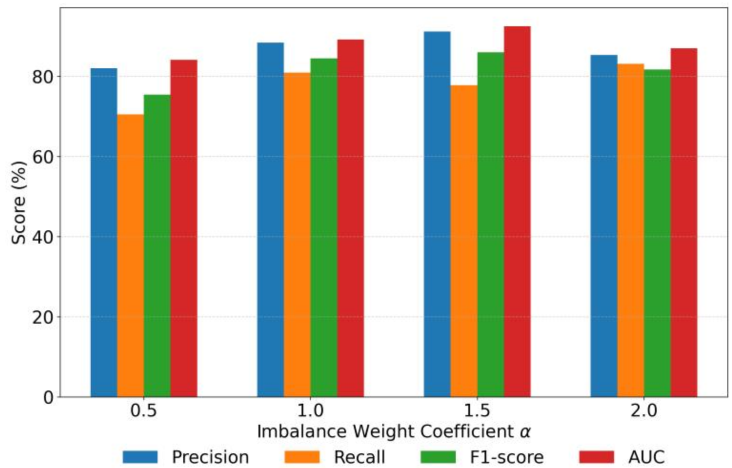

Figure 2. Sensitivity experiment of unbalanced weighting coefficients to minority class recognition performance

When the coefficient rises to the middle range, all metrics improve significantly. Precision, F1-score, and AUC increase together, which shows that the model strengthens representation learning for the minority class while maintaining overall discrimination ability. The growth in Recall is particularly evident. This indicates that the increased weight helps prevent minority samples from being ignored. It also leads to more appropriate decision boundaries under imbalanced structures and enhances sensitivity to minority classes.

In higher weight ranges, the model shows mixed trends. Precision and AUC decrease slightly, while Recall continues to rise. This suggests that the model's focus shifts further toward minority classes. Although this improves minority class detection, it comes with a cost. The discrimination ability for the majority class weakens. This shows that tuning the imbalance weight requires balancing multiple performance metrics. The slight decline in F1-score reflects the widening gap between Precision and Recall.

Considering all the metrics, the imbalance weight coefficient plays a critical role in regulating model adaptability under imbalanced and shifting distributions. When the coefficient is in the middle range, the model achieves a better balance among Recall, Precision, and AUC, and forms a more stable decision boundary. When the coefficient is too large or too small, feature bias increases and overall performance becomes volatile. This experiment highlights the importance of this hyperparameter in controlling learning strength across classes. It also aligns well with the proposed adaptive mechanism and reflects the model's sensitivity and adjustment ability when facing class imbalance and dynamic distribution changes.

This work further examines how the distribution offset adjustment coefficient affects the robustness of the model, and the corresponding sensitivity analysis results are presented in Figure 3.

Viewed from a global perspective, the distribution shift adjustment coefficient exhibits a performance trajectory that first rises and then declines. This indicates that the coefficient plays a key role in regulating the model's sensitivity to differences between the source distribution and the current distribution. When the coefficient is low, the model lacks responsiveness to distribution changes, and both Precision and Recall remain at relatively low levels. This suggests that, relying solely on original statistical features, the model is unable to accurately distinguish minority classes or maintain stable decision boundaries. As the coefficient increases, the model begins to leverage distribution difference information more effectively, resulting in improved overall discrimination capability.

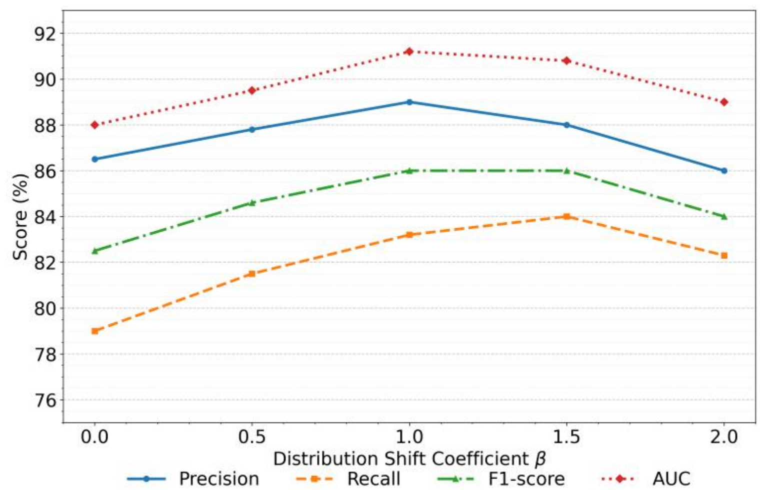

Figure 3. Sensitivity experiment of distribution offset adjustment coefficient to model robustness

When the coefficient reaches the middle range, all metrics show significant improvements. Precision, Recall, F1-score, and AUC peak at this stage. This indicates that the model captures feature drift caused by distribution changes more effectively in the representation space. It also achieves more balanced modeling of minority and majority classes. The clear rise in AUC reflects the advantage of the adaptive mechanism in distinguishing different category patterns. This stage shows that the model makes optimal use of distribution differences. It enhances detection of anomalous patterns while maintaining stable adaptation to the overall distribution.

When the coefficient increases further, the performance metrics begin to decline. This reflects the risk that the model may incorporate noise or non essential features when it overreacts to distribution shifts. The decrease in Precision shows that the model becomes too sensitive to shifted features and sacrifices discrimination quality for the majority class. The decline in Recall and F1-score indicates that overemphasizing shift signals harms the stability of the decision boundary and reduces generalization ability. This confirms that the coefficient has an optimal range rather than a monotonic benefit.

These trends show that the model depends strongly on this adjustment coefficient when handling imbalanced data and dynamic distribution changes. Proper parameter settings allow the model to adapt to drifting environments and strengthen minority class features. When the coefficient is within the optimal range, the robustness of the model improves significantly. It can track distribution changes while avoiding overfitting to local noise. This leads to a more reliable dynamic classification mechanism. The results align with the proposed adaptive framework and demonstrate that the introduced distribution difference modeling strategy can effectively handle non static data environments.

In addition, this paper investigates how the distribution difference metric parameter influences the model's adaptive capability, and the corresponding sensitivity analysis results are presented in Figure 4.

In this sensitivity experiment, the model shows a clear nonlinear response across all performance metrics as the distribution difference parameter increases. This pattern indicates that the adaptive classification framework can react to distribution shifts and capture their effects in time. In low-difference regions, Precision rises as the parameter grows. This suggests that under mild shifts, the model benefits from the inductive bias introduced by structural regularization, which strengthens fine-grained class discrimination. However, when the shift becomes large, Precision declines, implying that excessive perturbation weakens feature stability and makes it harder to maintain consistent decision boundaries.

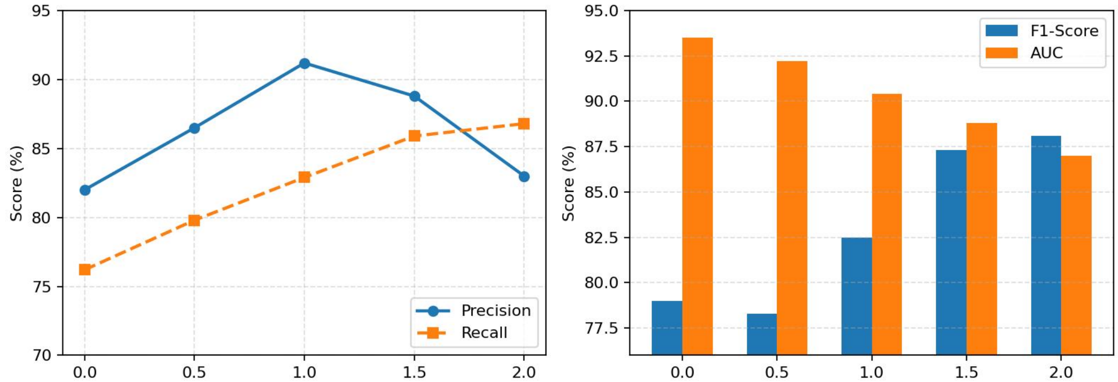

Figure 4. Experiment on the sensitivity of distribution difference measurement parameters to adaptive ability

Unlike the fluctuation observed in Precision, Recall increases steadily across all parameter values. This means the model expands its decision coverage under stronger shifts to reduce missed detections. The result shows that structural regularization effectively alleviates the weak-signal problem of minority classes, even when the disturbance becomes stronger. The monotonic increase also highlights the model's potential to generalize in cross-domain or non-IID environments.

The F1 score first changes slowly, then increases sharply in the moderate-difference region, and later becomes stable. This indicates that under moderate shifts, the model gains improvements in both Precision and Recall, leading to better overall performance. When the parameter continues to rise, Recall keeps increasing, but the drop in Precision offsets part of the benefit, causing a slight decline in the F1 score. This again emphasizes the central role of this parameter in regulating adaptive behavior. Moderate shifts can enhance performance, while excessive disturbances can damage class separation structures.

AUC decreases gradually as the parameter increases, suggesting a decline in discriminative ability under strong distribution shifts. Because AUC is sensitive to global ranking consistency, the trend indicates that distribution perturbations affect the ordering of features, reducing the stability of the classification boundary. However, the decline remains mild, showing that structural regularization preserves part of the overall discriminative capability. Taken together, the changes across metrics show the layered influence of the distribution difference parameter on model adaptivity. Moderate values strengthen the model's responsiveness to dynamic distributions, while overly large values may harm feature stability and weaken classification robustness.

5. Conclusion

This study proposes a new adaptive classification framework that addresses the widespread issues of class imbalance and distribution shift in real environments. The goal is to maintain efficiency and robustness when minority classes are scarce and distributions change over time. From methodological design to structural mechanisms and theoretical motivation, the study presents a systematic, adaptive, and scalable solution that offers new insights for imbalanced classification. From a broader perspective, this approach not only represents a technical advancement but also provides a more reliable foundation for applying intelligent systems in complex and evolving environments.

First, the study offers a unified and flexible modeling strategy at both theoretical and methodological levels. Traditional approaches to imbalanced learning often rely on static weighting, resampling, or loss adjustments. These methods may work when class distributions are fixed and stable but fail to adapt when distributions change. This work introduces distribution shift estimation, dynamic class weighting, structural regularization, and adaptive boundary adjustment. It integrates the challenges of class imbalance and distribution drift into a single modeling problem. This design allows the model to adjust its learning strategy according to data changes while maintaining both minority class recognition and overall classification stability. Such integration increases flexibility when handling heterogeneous, multi-source, and dynamic data and represents an important extension to existing paradigms in imbalanced learning.

Second, the adaptive framework has strong practical relevance. Many critical applications, such as financial risk monitoring, network intrusion detection, medical event identification, equipment failure prediction, and industrial quality control, cannot guarantee balanced or stable data distributions. Minority classes often correspond to high-risk or high-value events, and their distributions may shift with time, environment, or user behavior. Traditional static models deteriorate in these settings due to distribution mismatch or class sparsity. The proposed method can learn and adapt continuously in these scenarios. It improves minority class detection and enhances the stability of the overall system. It strengthens reliability and trustworthiness in real deployment. In other words, it provides a more actionable and safer path for tasks involving both minority class scarcity and distribution drift.

Furthermore, this study identifies several promising directions for future research. One direction is to extend the method to multimodal or heterogeneous data fusion, such as combining structured data with text or integrating graph data with time series. Adaptive mechanisms are likely to show even greater value in these settings. Another direction is to integrate online learning and incremental learning so that the model can respond to environmental changes in real time. This is especially important for streaming data monitoring and real-time anomaly detection. A third direction is to explore lightweight structural designs and efficient adaptive mechanisms for deployment in resource-constrained or edge environments. This may broaden practical applications. Finally, the framework can be enhanced by incorporating drift detection, active learning, or uncertainty estimation to further improve the reliability of minority class recognition and decision making.

As intelligent systems continue to expand across industries, the need for adaptability, stability, and sustainability becomes increasingly important. The adaptive classification framework developed in this study provides an effective response to the challenges of imbalanced data and distribution shift. It also offers a pragmatic solution for real-world situations in which data are imperfect but reliability is essential. As data volume grows and tasks become more complex, adaptive and resilient classification models will play an even more critical role. We hope this work inspires further research and encourages the community to test and refine adaptive classification systems in broader and more complex real-world environments, helping intelligent systems move toward a more reliable, robust, and general future.

References

[1] N. V. Chawla, K. W. Bowyer, L. O. Hall and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-Sampling Technique," Journal of Artificial Intelligence Research, vol. 16, pp. 321-357, 2002.

[2] Guo L Z, Li Y F. Class-imbalanced semi-supervised learning with adaptive thresholding[C]//International conference on machine learning. PMLR, 2022: 8082-8094.

[3] Y. Xing, "Enhancing Advertising Recommendation Performance via Integrated Causal Inference and Exposure Bias Correction," 2023.

[4] J. G. Moreno-Torres, T. Raeder, R. Alaiz-Rodríguez, N. V. Chawla and F. Herrera, "A unifying view on dataset shift in classification," Pattern Recognition, vol. 45, no. 1, pp. 521-530, 2012.

[5] Zhang Y, Kang B, Hooi B, et al. Deep long-tailed learning: A survey[J]. IEEE transactions on pattern analysis and machine intelligence, 2023, 45(9): 10795-10816.

[6] Chen W, Yang K, Yu Z, et al. A survey on imbalanced learning: latest research, applications and future directions[J]. Artificial Intelligence Review, 2024, 57(6): 137.

[7] T. Zhao, X. Zhang and S. Wang, "GraphSMOTE: Imbalanced Node Classification on Graphs with Graph Neural Networks," Proceedings of the 14th ACM International Conference on Web Search and Data Mining, pp. 833-841, March 2021.

[8] Y. Wang, "Integrating Large Language Models and Knowledge Graphs for Intelligent Financial Regulatory Risk Identification," 2024.

[9] R. Fang, "Transaction Network Graph Neural Networks for Automated and Robust Financial Fraud Detection in Corporate Auditing," 2024.

[10] Zhang Z, Xiang X. Long-tailed classification with gradual balanced loss and adaptive feature generation[J]. arXiv preprint arXiv:2203.00452, 2022.

[11] Wang X, Lian L, Miao Z, et al. Long-tailed recognition by routing diverse distribution-aware experts[J]. arXiv preprint arXiv:2010.01809, 2020.

[12] Hollmann N, Müller S, Eggensperger K, et al. Tabpfn: A transformer that solves small tabular classification problems in a second[J]. arXiv preprint arXiv:2207.01848, 2022.