


Graph Sampling-based Cost-Sensitive for Imbalance Classification

Abubakar Isah ^{*2†}, Ibrahim Aliyu [†], Jinsul Kim [†]

Department of Intelligent Electronics and Computer Engineering, Chonnam National University, Gwangju 1[†]

Department of Computer Science, Ahmadu Bello University Zaria 2[†]

Email: abubakarisah@jnu.ac.kr^{*}

Abstract—This paper addresses the challenge of class-imbalanced graph classification, which involves accurately categorizing graphs under skewed class distributions. While graph neural networks (GNNs) have demonstrated strong performance in graph learning tasks, they often suffer from a bias toward majority classes when trained on imbalanced datasets. Existing methods, particularly those applied to failure detection in network systems, frequently overlook the structural dependencies and subgraph semantics that are essential for recognizing minority classes. To mitigate these issues, we propose GASCOS (Graph Sampling-based Cost-Sensitive), a simple yet effective framework that enhances graph-level representation learning in imbalanced settings. GASCOS adaptively samples graphs to form balanced subclasses by mixing minority and majority class instances, ensuring comparability in sample sizes. It further employs the Mixup technique to generate synthetic graphs within each subclass, increasing diversity and preventing representation collapse. Additionally, a cost-sensitive learning strategy is incorporated to guide the model toward learning more discriminative features for minority classes. Experiments conducted on 5G core network datasets demonstrate that GASCOS consistently outperforms state-of-the-art baselines in imbalanced graph classification tasks.

Keywords—Graph neural network, cost-sensitive, boosting, network digital twin, 5G core network

I. INTRODUCTION

Graphs are widely recognized as powerful data structures for modeling complex relationships among entities in various domains, including network digital twins, fraud detection, and recommender systems [1], [2]. With core learning tasks such as node classification [3], graph classification [4], and graph clustering [4], there has been significant interest in exploring the potential of graph-based analysis. Among these, graph classification—which aims to assign class labels to entire graphs—has emerged as a particularly important area, with applications in protein function prediction and molecular property analysis [5].

Class imbalance, especially in multiclass datasets, poses persistent and complex challenges that are yet to be fully addressed. In real-world applications like

human behavior recognition [6], video classification [7], and medical decision-making [8], class imbalance frequently arises. This occurs when some classes are significantly underrepresented (minority classes) while others dominate the dataset (majority classes), complicating model training and reducing classification performance. Without appropriate handling, class imbalance often results in biased models, poor generalization to minority classes, and ultimately, misleading conclusions. These risks are particularly critical in high-stakes domains such as healthcare, where correctly identifying rare conditions is more valuable than maximizing overall accuracy driven by prevalent cases.

To tackle multiclass imbalance, three mainstream strategies are widely adopted: data sampling, cost-sensitive learning, and ensemble methods [9], [10], [11]. Sampling techniques seek to balance class distributions by either oversampling minority classes or undersampling majority ones [12], [13], [14]. While these approaches can improve model performance, they suffer from limitations such as overfitting due to oversampling, loss of informative samples in undersampling [15], and increased computational costs [16]. Cost-sensitive learning methods address imbalance by assigning different weights or misclassification penalties to each class [17]. These are often straightforward to implement and adaptable across learning algorithms [9], though selecting appropriate class weights or cost matrices—especially in multiclass settings—remains non-trivial. Ensemble methods, which combine multiple models to enhance robustness, have also demonstrated success in mitigating class imbalance [18], [19], [20], [21].

To effectively address the challenge of class imbalance in graph classification, we propose GASCOS (Graph Sampling-based Cost-Sensitive), a novel framework that combines adaptive sampling with cost-sensitive learning. The key idea behind GASCOS is to improve graph-level feature representation by integrating principled sampling strategies into a supervised cost-sensitive setting. Specifically, GASCOS adaptively

partitions both minority and majority class graphs into multiple subclasses, ensuring that each subclass contains a balanced number of samples comparable to the minority class. This design prevents the semantic structures of minority classes from being overwhelmed by dominant class patterns. Furthermore, the training process incorporates a cost-sensitive loss function that penalizes misclassification of minority samples more heavily, thereby promoting balanced learning across all categories. Together, these components enable GAS-COS to capture minority substructures more effectively and achieve improved performance in imbalanced graph classification tasks.

II. RELATED WORK

This category includes data-level methods, which aim to mitigate class imbalance by either increasing the number of samples from minority classes or reducing the number of samples from majority classes. Once the dataset is adjusted, learning models can be trained without altering their structure or algorithms [12], [13], [17]. Sampling-based methods typically rely on pairwise distance computations, with Euclidean distance being the most commonly used metric [13], [17], [15], [16]. In more recent studies, alternative distance measures have been explored to capture richer features of the data space, leading to enhanced performance in handling multiclass imbalance. These include entropy-based sampling distances [15], [9], adaptive distance-based oversampling [16], and Hellinger distance-based oversampling [10].

A widely adopted oversampling technique is the Synthetic Minority Over-sampling Technique (SMOTE) [13], which generates synthetic samples for minority classes by interpolating between existing instances. Several variants of SMOTE have been proposed [17], [15], [16], among which ADASYN (Adaptive Synthetic Sampling Approach) is one of the most popular. ADASYN focuses on generating synthetic samples near minority class examples that lie farther from their respective class boundaries, thus enhancing learning in harder regions of the feature space [17]. However, most sampling-based techniques adjust class imbalance by modifying the class ratios in the training data, which typically remain fixed during the iterative learning process.

Sampling techniques are generally straightforward to implement and operate at the data level. By reducing the bias toward majority classes, they can significantly improve the predictive performance of various machine learning models. Nevertheless, they are not without limitations. Undersampling can result in the loss of important information from the majority class, degrad-

ing model performance [12], [15], while oversampling may lead to overfitting due to the replication or interpolation of existing data [12]. Furthermore, sampling methods may encounter challenges in multiclass settings and can be computationally expensive, especially when using complex distance metrics or generating a large number of synthetic samples [15], [9]. In this paper, we propose an advanced strategy for deriving sample weights that leverages both between-class and within-class data distributions to address multiclass imbalance in graph classification. Unlike traditional cost-sensitive methods that rely on fixed or heuristic class weights, our approach dynamically adjusts sample importance based on the statistical separability and internal compactness of class instances.

III. DEFINITION OF THE PROBLEM

Let $\mathcal{G} = \{(G_i, y_i)\}_{i=1}^N$ denote a graph dataset consisting of N graphs, where each graph G_i is paired with its corresponding class label $y_i \in \{1, \dots, K\}$. Denote by N_j the number of graphs in the j -th class, with class indices sorted such that $N_1 \geq N_2 \geq \dots \geq N_K$. The degree of class imbalance in the dataset is characterized by the imbalance ratio (IR), commonly defined as the ratio of the largest class size to the smallest class size. In this work, we consider a class-imbalanced setting where the training set exhibits a skewed class distribution, while the test set remains balanced. More formally, the class proportions in the training data satisfy:

$$\int p(G|y = k_1)dG \geq \int p(G|y = k_2)dG, \quad \forall k_1 \leq k_2, \quad (1)$$

and

$$\lim_{k \rightarrow K} \int p(G|y = k)dG = 0, \quad (2)$$

which reflects the successive reduction in class sizes as the class index increases, with probabilities for higher-index classes tending to zero. The objective is to develop a classifier that can learn effective and unbiased graph representations from the imbalanced training data, and that generalizes well to all classes when evaluated on a balanced test set. Specifically, we aim for the classifier to maintain robust performance across both majority and minority classes, avoiding bias towards the more prevalent classes present during training.

A. GraphSAGE for Imbalance Learning

Given a graph dataset $\mathcal{G} = \{(G_i, y_i)\}_{i=1}^N$ consisting of N graphs, where each G_i is associated with a class label $y_i \in \{1, \dots, K\}$, we seek to learn robust graph-level representations for classification under class imbalance. Let N_j denote the number of graphs in the

j -th class, ordered such that $N_1 \geq N_2 \geq \dots \geq N_K$. The degree of imbalance in the dataset is quantified by the imbalance ratio (IR), typically defined as the ratio between the sizes of the largest and smallest classes. In our context, the training data exhibits such class imbalance, while the test set remains balanced, promoting the need for an unbiased classifier capable of generalization across all classes.

To address scalability and class imbalance, we employ a GraphSAGE-based encoder, which leverages both node features and structural information through a neighborhood sampling mechanism. Specifically, at the l -th layer of the GNN, the embedding for a node v , denoted as \mathbf{h}_v^l , is updated by aggregating the representations of a fixed-size, randomly sampled set of neighbors, $\mathcal{N}_S(v)$, along with the node’s own representation from the previous layer. Formally,

$$\mathbf{h}_v^l = \text{AGGREGATE}^{(l)}(\mathbf{h}_v^{l-1}, \{\mathbf{h}_u^{l-1} \mid u \in \mathcal{N}_S(v)\}), \quad (3)$$

where $\text{AGGREGATE}^{(l)}(\cdot)$ denotes a differentiable aggregation function, and $\mathcal{N}_S(v)$ represents the sampled neighborhood of v . This sampling strategy constrains the computational cost and memory footprint, making the approach suitable for large-scale graphs and stabilizing training under imbalanced conditions.

After L layers of iterative sampling and aggregation, node embeddings $\{\mathbf{h}_v^L\}_{v \in V}$ are pooled via a READOUT function to produce the graph-level representation \mathbf{h}_G :

$$\mathbf{h}_G = \text{READOUT}(\{\mathbf{h}_v^L \mid v \in V\}). \quad (4)$$

The obtained graph representation \mathbf{h}_G is subsequently used for downstream graph classification tasks. The stochasticity introduced by neighborhood sampling in GraphSAGE helps mitigate the risk of majority class dominance during training and encourages the learning of more generalizable features, enhancing the classifier’s performance on minority classes, particularly in the presence of severe imbalance.

IV. SYSTEM MODEL

Our proposed GASCOS (Graph Sampling-based Cost-Sensitive learning), a framework that unifies neighborhood sampling with cost-sensitive learning strategies. The core idea behind GASCOS is to enable effective graph-level representation learning by adaptively mitigating class imbalance during training. In particular, GASCOS begins by partitioning the training data into multiple subclasses, each containing a comparable number of samples, by selectively oversampling minority class instances and undersampling the majority class. This sampling strategy promotes balance at the subclass level, ensuring that structural and semantic patterns of the minority classes are adequately captured.

A. Data sets

We first discuss the 5G core network failure classification dataset used, as indicated in Table I, before providing the data preprocessing techniques. The data consists of 15 network failure case with a normal. These failure have been sample strategically based on their failure class types into 5 failure types. The summary of the datasets description was illustrated in Fig. 1.

Table I: 5G Core Network Dataset Distribution Table

Dataset	5G core network
Failure events	Normal = 2456 Failure = 1186
Features	4,121
Classes	16
Nodes	3,645
Edges	1,328,238
IR	0.4886

B. Data preprocessing

The data preprocessing phase involves several crucial steps to ensure the quality and consistency of the dataset. First, instances containing spaces or null values are removed, and all textual features are converted into their corresponding numeric representations. For classification purposes, a label column is introduced—assigning a value of 0 for normal instances and 1 for attacks in binary classification tasks, while for multi-class scenarios, an `attack_cat` column is added to encode each instance into a categorical label ranging from 0 to 9. To address the varying scales of feature values, the dataset undergoes normalization using the min-max scaling technique, which transforms all feature values into the range $[0, 1]$. This is computed as follows:

$$Z = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

where x is the original feature value, and x_{\min} and x_{\max} represent the minimum and maximum values of that feature, respectively.

C. Evaluation metrics

Performance evaluation in machine learning algorithms is essential for understanding how well a model predicts outcomes relative to the actual targets, particularly in multi-class classification tasks. In this study, we adopt a set of performance metrics suitable for multi-class scenarios. Let $TP_i = p_{ii}$ denote the number of true positives for class i , $FP_i = \sum_{j=1, j \neq i}^r p_{ij}$ the false positives, and $FN_i = \sum_{j=1, j \neq i}^r p_{ji}$ the false negatives, where p_{ij} represents the number of instances of class

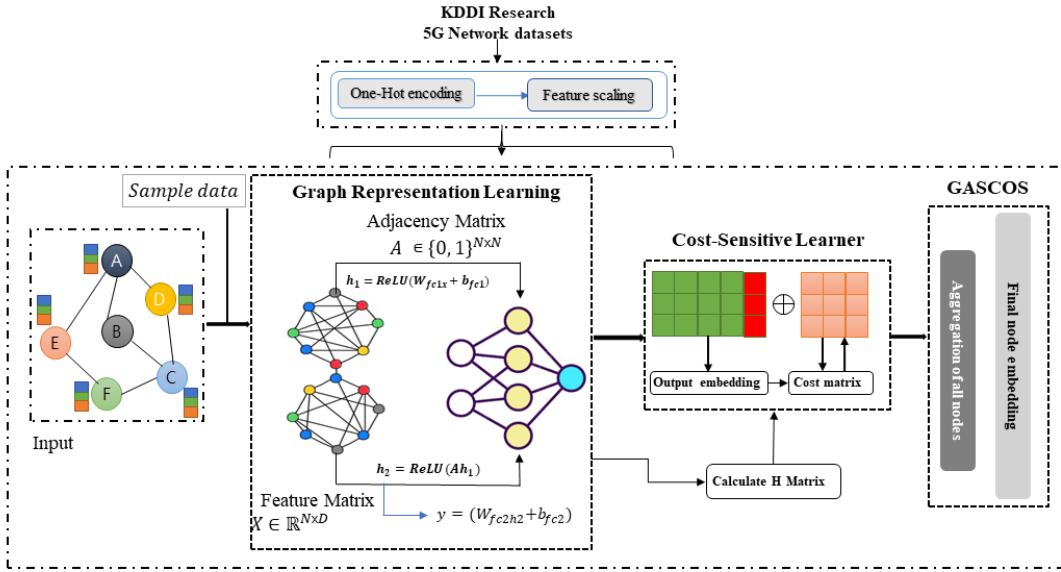


Figure 1: Proposed graph sampling based cost-sensitive for imbalance learning

i predicted as class j , and r is the total number of classes.

To compute precision and recall for each class, we define:

$$\text{Precision}_i = \frac{TP_i}{TP_i + FP_i} = \frac{p_{ii}}{p_{i:}} \quad (6)$$

$$\text{Recall}_i = \frac{TP_i}{TP_i + FN_i} = \frac{p_{ii}}{p_{:i}} \quad (7)$$

The macro-averaged F1-score, which treats all classes equally regardless of their frequency, is given by:

$$\text{Macro-F1} = \frac{2}{r} \sum_{i=1}^r \frac{p_{ii}}{p_{:i} + p_{i:}} \quad (8)$$

Here, $p_{i:}$ denotes the total number of instances predicted as class i , and $p_{:i}$ represents the total number of true instances of class i . These metrics provide a comprehensive view of classification performance across all classes, accounting for both precision and recall.

V. RESULTS AND DISCUSSION

All experiments are implemented using the PyTorch framework and executed in a Python 3.8 environment on a Windows 10 system equipped with a 13th-generation Intel Core i5-13400 processor, 32 GB of RAM, and an NVIDIA GeForce RTX 3060 Ti GPU. For training, validation, and testing, the dataset is split using a 5:2:3 ratio. Model training employs a cross-entropy loss function with class-weighted balancing, optimized using the Adam optimizer with a learning rate of 0.01. Training is conducted for two epochs

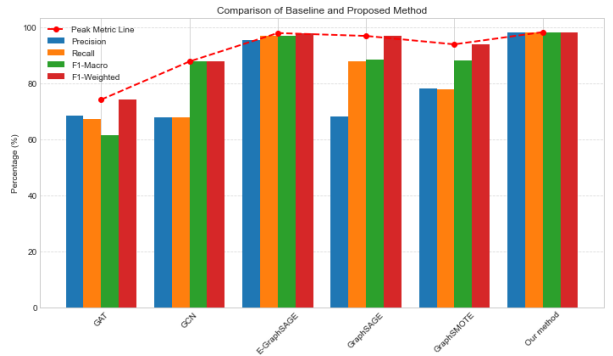


Figure 2: Comparison of the baseline and proposed method

with a batch size of 500. During each training iteration, samples are randomly selected while preserving the original positive-to-negative sample ratio of the complete dataset. The Adam optimizer is used for all model updates, with dataset-specific hyperparameter configurations applied to optimize performance.

Table II: Performance of our proposed based on varying imbalanced ratio sample

Methods	10%	30%	50%	70%
GAT	61.49%	57.32%	68.49%	67.13%
GCN	67.79%	67.75%	67.79%	78.75%
E-GraphSAGE	87.91%	87.88%	88.06%	87.91%
GraphSAGE	88.11%	87.80%	88.11%	87.88%
GraphSMOTE	88.45%	87.35%	88.55%	97.88%
Our method	88.24%	95.23%	98.20%	98.56%

The effectiveness of the proposed GASCOS frame-

Table III: Multiclass classification results for 5G core network

Class types	Precision	Recall	F1-Score	AUC
Normal	96.68%	98.82%	97.75%	95.09%
Bridge-down	62.47%	61.18%	70.16%	75.00%
CPU overload	92.37%	77.31%	84.17%	85.50%
Interface down	62.34%	54.85%	68.11%	70.08%
Memory stress	89.72%	88.57%	85.14%	85.00%
Interface down	65.33%	62.86%	63.59%	64.50%
Macro Average	89.57%	73.35%	75.72%	70.50%
Weighted Average	87.65%	87.67%	88.16%	89.50%

work is evaluated through a series of experiments on a 5G core network digital twin dataset. Table ?? compares GASCOS against several strong GNN baselines including GAT, GCN, GraphSAGE, E-GraphSAGE, and GraphSMOTE. Our method achieves the highest performance across all key metrics: 98.24% Precision, 98.23% Recall, 98.24% F1-Macro, and 98.16% F1-Weighted. These improvements confirm that GASCOS effectively addresses the class imbalance challenge by enhancing representation learning, especially for minority classes. Unlike conventional methods that exhibit bias toward majority classes, GASCOS maintains balance by forming class-comparable subclasses and applying Mixup-based graph augmentation.

To further investigate the robustness of GASCOS, Table II presents model performance across different imbalance ratios (10%, 30%, 50%, 70%). While baseline models show fluctuating or degraded performance as imbalance worsens, GASCOS consistently improves, reaching 98.56% at the 70% imbalance setting. This stability is attributed to the adaptive graph sampling mechanism and the use of subclass-wise synthetic graph generation, which prevents overfitting and representation collapse. In contrast, models like GCN and GAT exhibit reduced performance under high imbalance, highlighting their limitations in handling minority-rich scenarios.

Class-wise performance is detailed in Table III, where GASCOS demonstrates superior detection for minority failure classes such as *CPU Overload* and *Bridge-down*, achieving AUCs of 85.50% and 75.00%, respectively. The Macro F1-score of 75.72% and AUC of 70.50% indicate balanced performance across all classes, including rare events. The cost-sensitive training approach embedded in GASCOS effectively steers learning towards minority classes without sacrificing overall accuracy. These findings collectively demonstrate that GASCOS not only improves classification performance but also provides a principled and interpretable solution for imbalanced graph classification in mission-critical 5G networks.

VI. CONCLUSION

In this paper, we addressed the critical challenge of class imbalance in graph classification tasks, particularly within the context of 5G core network. Existing GNN-based methods often struggle to recognize minority classes due to skewed data distributions, our proposed framework, GASCOS (Graph Sampling-based Cost-Sensitive), offers a principled and effective solution. GASCOS improves graph-level representation learning through a combination of adaptive graph sampling, subclass-based augmentation, and cost-sensitive training. These mechanisms jointly mitigate overfitting to majority classes and promote discriminative learning for minority classes. Extensive experiments on 5G core network datasets demonstrated that GASCOS consistently outperforms state-of-the-art baselines in terms of Precision, Recall, F1-score, and AUC, even under severe imbalance ratios. This work highlights the importance of structural awareness and synthetic data augmentation in tackling imbalanced graph learning. Future research directions include extending the framework to dynamic graphs, incorporating temporal information, and exploring explainability for minority class predictions in critical network scenarios.

REFERENCES

- [1] W. Ju, Z. Fang, Y. Gu, Z. Liu, Q. Long, Z. Qiao, Y. Qin, J. Shen, F. Sun, Z. Xiao *et al.*, "A comprehensive survey on deep graph representation learning," *Neural Networks*, vol. 173, p. 106207, 2024.
- [2] H. Shin, I. Aliyu, A. Isah, and J. Kim, "Improving the real-data driven network evaluation model for digital twin networks," in *2024 IEEE International Conference on Communications Workshops (ICC Workshops)*. IEEE, 2024, pp. 7-12.
- [3] J. Luo, Y. Gu, X. Luo, W. Ju, Z. Xiao, Y. Zhao, J. Yuan, and M. Zhang, "Gala: Graph diffusion-based alignment with jigsaw for source-free domain adaptation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 9038-9051, 2024.
- [4] Z. Mao, W. Ju, Y. Qin, X. Luo, and M. Zhang, "Rahnet: Retrieval augmented hybrid network for long-tailed graph classification," in *Proceedings of the 31st ACM international conference on multimedia*, 2023, pp. 3817-3826.
- [5] W. Ju, Z. Mao, S. Yi, Y. Qin, Y. Gu, Z. Xiao, J. Shen, Z. Qiao, and M. Zhang, "Cluster-guided contrastive class-imbalanced graph classification," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, no. 11, 2025, pp. 11 924-11 932.
- [6] B. A. M'hamed and B. Fergani, "A new multi-class wsvm classification to imbalanced human activity dataset." *J. Comput.*, vol. 9, no. 7, pp. 1560-1565, 2014.
- [7] S. Pouyanfar, S.-C. Chen, and M.-L. Shyu, "Deep spatio-temporal representation learning for multi-class imbalanced data classification," in *2018 IEEE International Conference on Information Reuse and Integration (IRI)*. IEEE, 2018, pp. 386-393.
- [8] L. Lijun, J. Tingting, and H. Meiya, "Feature identification from imbalanced data sets for diagnosis of cardiac arrhythmia," in *2018 11th international symposium on computational intelligence and design (ISCID)*, vol. 2. IEEE, 2018, pp. 52-55.
- [9] G. Haixiang, L. Yijing, J. Shang, G. Mingyun, H. Yuanyue, and G. Bing, "Learning from class-imbalanced data: Review of

- methods and applications," *Expert systems with applications*, vol. 73, pp. 220–239, 2017.
- [10] S. Wang and X. Yao, "Multiclass imbalance problems: Analysis and potential solutions," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 4, pp. 1119–1130, 2012.
- [11] M. Lin, K. Tang, and X. Yao, "Dynamic sampling approach to training neural networks for multiclass imbalance classification," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 24, no. 4, pp. 647–660, 2013.
- [12] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on knowledge and data engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.
- [13] 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.
- [14] H. He, Y. Bai, E. A. Garcia, and S. Li, "Adasyn: Adaptive synthetic sampling approach for imbalanced learning," in *2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence)*. Ieee, 2008, pp. 1322–1328.
- [15] L. Li, H. He, and J. Li, "Entropy-based sampling approaches for multi-class imbalanced problems," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 11, pp. 2159–2170, 2019.
- [16] S. S. Mullick, S. Datta, and S. Das, "Generative adversarial minority oversampling," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1695–1704.
- [17] C. Elkan, "The foundations of cost-sensitive learning," in *International joint conference on artificial intelligence*, vol. 17, no. 1. Lawrence Erlbaum Associates Ltd, 2001, pp. 973–978.
- [18] N. V. Chawla, A. Lazarevic, L. O. Hall, and K. W. Bowyer, "Smoteboost: Improving prediction of the minority class in boosting," in *European conference on principles of data mining and knowledge discovery*. Springer, 2003, pp. 107–119.
- [19] C. Seiffert, T. M. Khoshgoftaar, J. Van Hulse, and A. Napolitano, "Rusboost: A hybrid approach to alleviating class imbalance," *IEEE transactions on systems, man, and cybernetics-part A: systems and humans*, vol. 40, no. 1, pp. 185–197, 2009.
- [20] S. Datta, S. Nag, and S. Das, "Boosting with lexicographic programming: Addressing class imbalance without cost tuning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 5, pp. 883–897, 2019.
- [21] Y. Freund, R. E. Schapire *et al.*, "Experiments with a new boosting algorithm," in *icml*, vol. 96. Bari, Italy, 1996, pp. 148–156.