International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 05 May-2026 | Impact Factor: 3.5

Adaptive Margin Broad Learning System for Imbalanced Data Classification

Chitransh Tomar[1], Bhagat Sing Raghuwanshi[2]

[1] Centre for Artificial Intelligence / Madhav Institute of Technology and Science, Gwalior, M.P., India
[2] Centre for Artificial Intelligence / Madhav Institute of Technology and Science, Gwalior, M.P., India

Corresponding Author Email: tomarchitransh19@gmail.com

How to Cite this Article:

Tomar, C. & Raghuwanshi, B. S. (2026). Adaptive Margin Broad Learning System for Imbalanced Data Classification. International Journal of Creative and Open Research in Engineering and Management, <i>02</i>(05). https://doi.org/10.55041/ijcope.v2i5.031

https://doi.org/10.55041/ijcope.v2i5.031

Abstract—

Generally, class imbalance is a major problem that occurs with various machine learning problems where there are a large number of instances of one class compared to another class. In this case, it has been found that a typical classification model may be biased towards a particular class and may not be able to classify instances of another class properly. This problem is critical while dealing with various real-world problems. In recent times, Broad Learning System (BLS) has gained attention from many researchers due to its fast training speed and efficiency of the flat network. However, it has been found that the conventional BLS model may not be efficient while dealing with imbalanced data sets because it uses a least squares method. Generally, a least squares method may not be efficient while dealing with imbalanced data sets.

This paper proposes an adaptive margin broad learning system (AMBLS) model to improve its efficiency while handling imbalanced data sets in order to address this issue with traditional BLS models. The suggested approach may be more effective when handling imbalanced data sets than traditional BLS models, according to experimental results on a number of benchmark imbalanced data sets.

Keywords- Imbalanced Learning; Broad Learning System; Adaptive Margin; Minority Classification; Machine Learning.

I. INTRODUCTION

The problem of class imbalance is a common problem that occurs in many real-world machine learning problems, where the number of instances in one class is significantly larger than the number of instances in other classes. The problem of class imbalance occurs in many real-world machine learning problems, e.g., fraud detection, disease diagnosis, fault detection, intrusion detection, etc. The majority of classical machine learning algorithms assume that all classes in a given dataset are relatively well-balanced and try to achieve a goal that minimizes the total classification error. Therefore, these algorithms show bias to the majority class, whereas they show poor performance in identifying instances of the minority classes. Since instances belonging to minority classes are often the most important instances, the issue of class imbalance has gained significant research interest [1], [2]. In order to address this problem, several approaches have been suggested in the literature. These approaches have been classified into three categories: data level, algorithm level, and ensemble level. In the first category, the problem is addressed by balancing the dataset with the help of over-sampling and under-sampling. Among these approaches, the Synthetic Minority Oversampling Technique (SMOTE) has been widely used to create synthetic minority class samples [3]. However, oversampling can cause noisy classes and overfitting problems when the number of minority class instances is too small [4]. In algorithm-level methods, the objective of the classifier is modified to tackle imbalance during the training process. For instance, techniques such as cost-sensitive learning and class-weighted learning emphasize more on the minority class during training. This is because these techniques allow the classifier to consider more importance for the minority class during the training process [5]. Recently, more sophisticated techniques such as graph-based learning, ensemble learning, and deep learning have been proposed for improving the performance of classification models in imbalance learning [6]. Despite the good performance of these techniques, it is observed that many of them have high computational complexity. Recently, the Broad Learning System (BLS) has gained much attention in recent years due to its efficient flat network structure and fast training speed. The structure of the Broad Learning System is different from the traditional deep neural network structure (as shown in Fig. 1), which employs multiple hidden layers. In the structure of the Broad Learning System, the network structure is

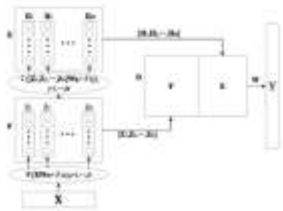

Fig. 1. Structure of the Broad Learning System (BLS). The inputs and outputs are denoted as $X \in \mathbb{R}^{N\times M}$ and $Y \in \mathbb{R}^{N\times C}$, respectively. The component of F consist of feature nodes denoted as Z_i, which are created by a sparse auto-encoder function ϕ, based on lasso regression. E consists of enhancement nodes denoted as H_j, which are created by an activation function ζ. Circles in Z_i and H_j represents row vectors. The components of W_b, β_e can be calculated by ϕ whereas W_h and β_h are random. E and F are integrated into D whereas W is used for output Y.

expanded horizontally through the generation of feature nodes and enhancement nodes via a random mapping method, and the output weights of the network are obtained via a simple least squares solution. However, the conventional BLS method is designed under the assumption of balanced class distributions. In the case of imbalanced datasets, the decision boundary learned from the least squares optimization method will be biased to the majority class. To address this problem of imbalanced learning using the conventional BLS method, several variants of the BLS method have been introduced. For example, incremental weighted ensemble BLS has been developed to enhance classification performance by assigning adaptive weights to training samples and combining multiple BLS models [9]. In addition, cost-sensitive BLS models have been introduced to improve minority class recognition in fault diagnosis applications by incorporating class-dependent cost functions [10]. Moreover, fuzzy-based mixture-of-experts BLS models have been proposed to improve feature representation and classification robustness for imbalanced data [11]. Although these improvements have been made to the existing approaches, the majority of the existing approaches focus on the adjustment of the weights or cost functions while the shape of the decision boundary is almost the same as the original one. Therefore, the risk of misclassifying the minority samples near the decision boundary is inevitable. To address the drawback of the existing methods, the Adaptive Margin Broad Learning System (AMBLS) is put forward to enhance the power of BLS to handle the imbalanced classification problems. The proposed approach is to introduce the class adaptive margin strategy to dynamically expand the decision boundary of the minority classes while the class-dependent weighting strategy is introduced to the output layer of BLS.

II. LITERATURE REVIEW

Learning from imbalanced datasets is a problem that has been extensively explored in the field of machine learning due to its importance in various real-world problems like medical diagnosis, financial fraud detection, and industrial fault detection. In such scenarios, the minority class is comprised of rare instances which hold significant value, whereas the majority class is dominant. However, the traditional classifiers such as decision trees, SVMs, and neural networks generally consider the data to be balanced and thus do not perform well on imbalanced datasets [1], [2]. In order to overcome this challenge, several strategies have been proposed in the literature. The strategies proposed for ad dressing the class imbalance problem can be broadly classified into three categories. They are data level strategies, algorithm level strategies, and ensemble learning strategies. The data level strategies for addressing the class imbalance problem include oversampling and undersampling. In the oversampling method, the minority class is oversampled, and the Synthetic Minority Oversampling Technique (SMOTE) and its variants have been widely used for generating the minority class samples and improving the class distribution [3]. Nevertheless, the oversampling method may lead to the inclusion of noisy data, and the undersampling method may lead to the loss of valuable information from the majority class [4]. In order to address the problems associated with the oversampling and undersampling methods, a number of hybrid strategies have been proposed [2]. Algorithm-level approaches try to modify the learning process of the classifiers to make them more sensitive to the minority classes. One such approach is cost-sensitive learning, which aims to assign higher costs to misclassifying the minority class to recognize the minority samples better [5]. Another approach is the use of margin-based learning algorithms to make the classifiers more sensitive to the minority classes by making the features more separable [6]. Additionally, the use of ensemble learning algorithms such as boosting and bagging is also shown to be effective in imbalanced learning [8]. Recently, the Broad Learning System (BLS) was proposed as a computationally efficient alternative to deep learning models due to their flat network structure [7]. BLS is computationally efficient compared to deep learning models as the deep learning models use multiple hidden layers to learn the features from the data, whereas the BLS uses random mapping to generate a large number of feature nodes as well as enhancement nodes to obtain the output weights through the least squares method [7]. Several research works have proposed extensions of BLS to enhance its performance in imbalanced learning problems. For instance, the Incremental Weighted Ensemble Broad Learning System was proposed as a method of enhancing the classification performance of BLS by combining multiple BLS models with varying sample weights [9]. Additionally, kernel-based and cost-sensitive BLS approaches have also been proposed as tools for solving imbalanced fault diagnosis problems by introducing cost functions in the BLS learning mechanism [10]. Moreover, fuzzy and mixture of experts-based BLS models have also been developed as tools for enhancing the feature representation and classification robustness of BLS in solving imbalanced problems [11]. Although the above approaches enhance the capability of BLS in solving imbalanced problems, most of the existing approaches have focused on weighting, cost adjustment, or ensemble learning methods. Although these approaches increases the representation of minority samples while training BLS, they do not directly affect the margin between classes during the classification mechanism of BLS. As a result, the minority class samples near the classification boundary are still vulnerable to misclassification during the classification mechanism of BLS. In order to address the shortcomings mentioned above, a new Adaptive Margin Broad Learning System (AMBLS) method will be introduced in this paper. In this method, a class adaptive margin method will be adopted to adaptively increase the decision margin of the minority classes, while the weights of the output layer of the BLS model will be adopted to further enhance the separability of the minority class samples without increasing the computational complexity of the BLS model.

III. METHODOLOGY

In this section, the proposed Adaptive Margin Broad Learn ing System (AMBLS) is discussed. The proposed method is expected to enhance the performance of the classifier on unbalanced datasets. The proposed method is an extension of the basic Broad Learning System architecture. The proposed framework is specifically designed in a way that it can enhance the separability of the minority class samples without compromising the computational efficiency of the Broad Learning System. The architecture of the proposed framework is described in Fig.2. The proposed framework has been divided into two steps. In the first step of the framework, the feature representations are generated using random feature nodes and enhancement nodes. The enhancement nodes are used in the framework in order to provide nonlinearity in the feature representations. In the second step of the framework, the adaptive margin is added in the decision boundary between the classes. This can enhance the discrimination between the classes. In the output layer of the framework, the weighted learning scheme has been used. In the weighted learning scheme, a different weight is assigned to the classes.

A. Overview of Broad Learning System-The Broad Learning System (BLS) further extends the network horizontally by creating feature mapping nodes and enhancement nodes. In contrast to deep neural networks, the Broad Learning System can be trained using a closed-form least squares solution.

B. Feature Mapping Layer- The input data is mapped to different groups of feature nodes by using randomly generated transformations.

C. Enhancement Nodes- To improve the nonlinear representation capability, enhancement nodes are generated from the feature mapping layer.

D. Adaptive Margin Mechanism- In the standard BLS framework, all training samples contribute equally to the learning process regardless of their class distribution. This may lead to biased decision boundaries if the classes are imbalanced. To overcome this limitation, we suggest the proposed class-adaptive margin mechanism.

E. Class-Weighted Output Optimization- To further emphasize minority samples during training, a class-weighted learning strategy is proposed.

F. Prediction- Once the output weights are obtained, the prediction for a new sample is computed.

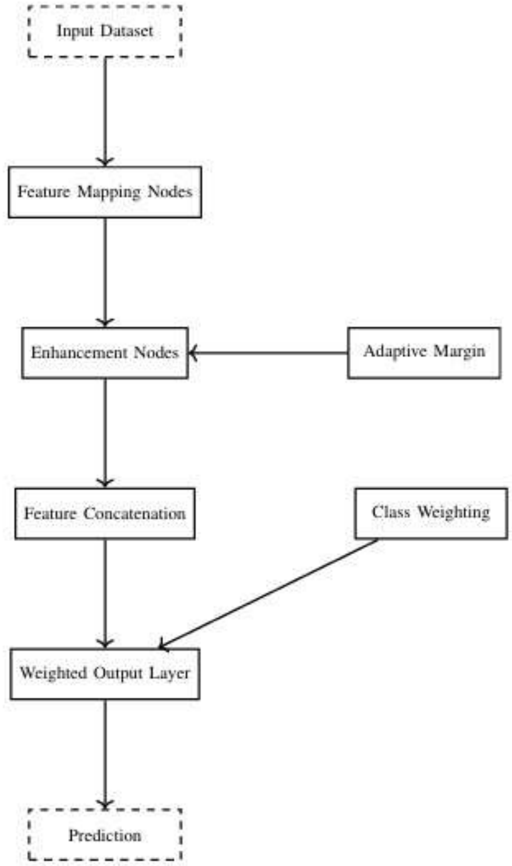

Fig. 2. Architecture of the Adaptive Margin Broad Learning System (AM-BLS). This model is an extension of the traditional Broad Learning System with the incorporation of adaptive margin adjustment and class-weighted learning to effectively recognize minority classes in imbalanced datasets

IV. EXPERIMENTAL SETUP:

In this section, the datasets, the baseline methods, the criteria of the experiments, and the implementation of the experiments for testing the efficacy of the proposed Adaptive Margin Broad Learning System (AMBLS) will be discussed. The experiments will verify the effectiveness of the proposed method in improving the accuracy of the recognition of minority classes without compromising the efficiency of the Broad Learning System.

Datasets: To evaluate the performance of the proposed method on imbalanced learning tasks, experiments were carried out on a number of widely used benchmark datasets. The datasets are different in terms of imbalanced learning ratio and feature dimension. Table I lists the major characteristics of the datasets adopted in the experiments. All datasets were normalized using min-max scaling.

Datasets	Sample	Features	Imb.
Pima	768	8	1.87
Breast Cancer	569	30	1.68
Credit Card	284807	30	492
CWRU Fault	1738	64	10.5
Glass	214	9	3.2
Yeast	1484	8	28.1
Ecoli	336	7	8.6
KDDCup99	494021	41	100+

Table I: Datasets used in the Experiments

A. Data Preprocessing: Prior to the training of the models, a series of preprocessing steps were conducted:

a. Handling of missing values through removal or mean imputation.

b. Normalization of all the features of the datasets using min-max scaling.

c. Random splitting of the datasets into a set of training and testing sets.

d. Stratified sampling was used to avoid any bias in evaluating the performance of the models.

B. Baseline Methods: To demonstrate the effectiveness of the proposed AMBLS model, we compared its performance with several commonly used imbalanced learning methods: Conventional Broad Learning System (BLS), SMOTE + BLS, Cost-Sensitive BLS , Incremental Weighted BLS, MFBLS (Mixture-of-Experts Fuzzy BLS).

C. Evaluation Metrics: As the accuracy of the classifier was not sufficient to evaluate the imbalanced dataset, various metrics were employed for the purpose of evaluating the classifier's performance.

a. Recall

b. F1 Score

c. G-mean.

D. Implementation Details: All experiments were implemented with the Python programming language. The models were executed in the Google Colaboratory environment, which is a cloud-based version of the Jupyter notebook with access to GPU acceleration and high-performance computing. This environment is used for implementing the machine learning models. In the proposed adaptive margin broad learning system (AMBLS), the standard numerical and machine learning libraries available in the Python programming language have been used.

E. Experimental Procedure: The general experimental steps for conducting experiments in the context of the research presented in this study can be outlined as follows:

a. Load the data and then preprocess it.

b. Divide the dataset into training and testing set.

c. Train baseline models on the training data.

d. Train the new model, namely, the AMBLS model.

e. Test all the models on the data and use the defined metrics for evaluating the models.

f. Make a comparison and analyze the performance improvements.

F. Visualization and Analysis: To further evaluate the effectiveness of the proposed method, various visualization results have been produced. These include:

- Performance comparison graphs of Recall, F1-score, and G-mean.
- Comparison of the time taken during the training process using various methods.
- Decision boundary visualization of BLS and AMBLS.

V. RESULTS AND DISCUSSION

In this section, the performance of the proposed model, i.e., the Adaptive Margin Broad Learning System, is evaluated in comparison with other baseline models. Here, experiments have been conducted to evaluate the efficiency of the proposed model in enhancing the recognition of minority classes while keeping the model computationally efficient.

A. Performance Comparison: The performance comparison between the proposed model, namely the Adaptive Margin Broad Learning System, and other baseline methods is presented in Table II. From Table II, it is observed that the AMBLS method has the best performance in terms of all the evaluation metrics. In addition, the improvement in the value of recall and G-mean shows the efficiency of the proposed model in handling the minority class data.

International Journal of Creative and Open Research in Engineering and Management
ISSN: 3108-1754 (Online)
Volume 02 Issue 05 May-2026 | Impact Factor: 3.5

B. Impact of Adaptive Margin: To test the efficiency of the adaptive margin mechanism, experiments with different margin scaling parameters were performed. It is clear from Fig. 3 that, with the increase in adaptive margin, the classification performance is improved. However, over-large margins may cause overfitting.

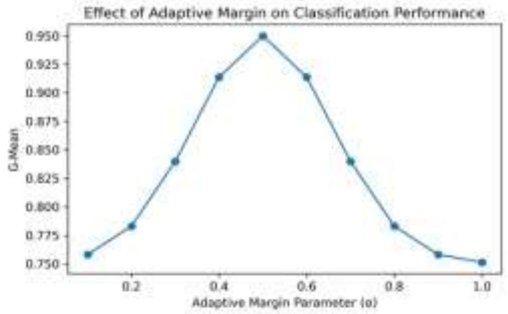

Fig. 3. Effect of adaptive margin parameter on classification performance

C. Training Time Analysis: The BLS has the benefit of efficiency in computation. In order to test the efficiency of the training process for the model, a comparison was made. The results shown in the Table III tell us that the proposed method maintains a relatively low training time while achieving superior classification performance.

Method	Training time(sec)
BLS	1.42
SMOTE+BLS	2.31
CS-BLS	1.87
MFBLS	2.74
Prop AMBLS	1.96

Table III: Training Time Comparison

TABLE II
PERFORMANCE COMPARISON OF DIFFERENT METHODS ACROSS BENCHMARK DATASETS

Dataset	BLS				SMOTE+BLS				CS-BLS				AMBLS			
	Acc	Rec	F1	G	Acc	Rec	F1	G	Acc	Rec	F1	G	Acc	Rec	F1	G
Pima	0.912	0.741	0.768	0.752	0.926	0.804	0.821	0.812	0.931	0.823	0.835	0.827	**0.946**	**0.887**	**0.891**	**0.882**
Breast Cancer	0.938	0.812	0.842	0.835	0.952	0.851	0.872	0.861	0.957	0.864	0.883	0.874	**0.972**	**0.912**	**0.925**	**0.913**
Credit Card	0.971	0.682	0.701	0.689	0.982	0.742	0.762	0.748	0.985	0.763	0.781	0.769	**0.992**	**0.842**	**0.856**	**0.842**
CWRU Fault	0.921	0.768	0.792	0.781	0.938	0.812	0.835	0.822	0.945	0.834	0.853	0.841	**0.962**	**0.889**	**0.901**	**0.889**
Glass	0.812	0.692	0.721	0.705	0.835	0.741	0.763	0.748	0.842	0.763	0.781	0.769	**0.861**	**0.812**	**0.832**	**0.815**
Yeast	0.781	0.641	0.681	0.662	0.812	0.721	0.742	0.728	0.825	0.742	0.761	0.749	**0.861**	**0.801**	**0.829**	**0.812**
Ecoli	0.842	0.712	0.745	0.732	0.861	0.781	0.802	0.791	0.872	0.792	0.818	0.806	**0.894**	**0.842**	**0.871**	**0.858**
KDDCup99	0.951	0.781	0.812	0.798	0.965	0.832	0.854	0.841	0.971	0.854	0.873	0.861	**0.984**	**0.902**	**0.921**	**0.908**

D. Decision Boundary Visualization: To further verify the effectiveness of the proposed method, the decision boundary of the traditional BLS method and the proposed method, namely the AMBLS model, is visualized. As shown in Fig. 4, the decision boundary generated by the proposed AMBLS model better separates minority samples from the majority class.

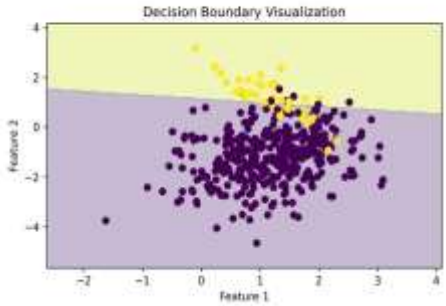

Fig. 4. Decision boundary comparison between conventional BLS and proposed AMBLS

E. Discussion: The experimental evaluation also proves that the AMBLS framework is able to provide an efficient solution to deal with imbalanced classification problems. The adaptive margin can be used to improve the separability of minority classes, and class-weighted optimization can be adopted to guarantee that minority class samples can be better utilized during the training process. In addition, it can also be found that the proposed model also inherits the computational efficiency of the Broad Learning System.

VI. CONCLUSION AND FUTURE WORK

In this paper, a novel Adaptive Margin Broad Learning System, namely AMBLS, was proposed to tackle the challenge of imbalanced learning. The proposed method is an extension of the conventional Broad Learning System. The proposed method introduces a class-adaptive margin strategy and a class-weighted optimization strategy. The proposed method is more efficient in separating the minority class samples by a greater margin and incorporating the minority class samples into the training process. From the experimental results on different benchmark datasets with different domains, it is observed that the proposed AMBLS method performs better than the conventional BLS model and other imbalance learning models. The improvement is more significant in the evaluation metric. The experimental results verify the effectiveness of the proposed method in eliminating the bias of imbalanced learning and the efficiency of the conventional BLS model. The experimental results verify the effectiveness of the proposed method in eliminating the bias of imbalanced learning and the efficiency of the conventional BLS model. Although the proposed model of AMBLS has good results, there are some directions which can be considered to further improve the results. Firstly, the proposed method can be used in big data or real-time data situations where the data distribution is changing over time. Secondly, the use of advanced feature extraction techniques or deep learning representations along with the BLS model can be considered to further improve the results. In addition, the margin or the weights used in the proposed method can be considered to be used dynamically or automatically to further improve the results of the proposed method. Thirdly, it can be applied to real-world problems such as medical diagnosis, fraud detection, or fault detection to further verify the results of the proposed model.

REFERENCES

[1] 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.

[2] A. Fernandez, S. Garcia, F. Herrera, and N. V. Chawla, "Learning from imbalanced data sets," Springer, 2018.

[3] 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.

[4] H. Han, W. Y. Wang, and B. H. Mao, "Borderline-SMOTE: A new over-sampling method in imbalanced data sets learning," International Conference on Intelligent Computing, pp. 878–887, 2005.

[5] Y. Sun, M. S. Kamel, A. K. Wong, and Y. Wang, "Cost-sensitive boosting for classification of imbalanced data," Pattern Recognition, vol. 40, no. 12, pp. 3358–3378, 2007.

[6] S. Wang, Y. Li, and L. Chen, "Intelligent fault diagnosis using machine learning techniques: A review," IEEE Access, vol. 10, pp. 34567–34580, 2022.

[7] C. L. P. Chen and Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," IEEE Transactions on Neural Networks and Learning Systems, vol. 29, no. 1, pp. 10–24, 2018.

[8] Z. H. Zhou, Ensemble Methods: Foundations and Algorithms, Chapman and Hall/CRC, 2012.

[9] X. Zhang, Y. Liu, and H. Li, "Incremental weighted ensemble broad learning system for imbalanced classification," Expert Systems with Applications, vol. 168, 2021.

[10] Y. Liu, J. Wang, and X. Zhang, "Cost-sensitive broad learning system for imbalanced fault diagnosis," Mechanical Systems and Signal Processing, vol. 135, 2020.

[11] L. Wang, H. Zhao, and Y. Chen, "Mixture-of-experts fuzzy broad learning system for imbalanced data classification," Information Sciences, vol. 579, pp. 768–783, 2021.