

A Hybrid Ensemble Model for Corporate Bankruptcy Prediction Based on Feature Engineering Method

Xiaoxia Wu¹, Dongqi Yang^{2, *}, Wenyu Zhang², Shuai Zhang²

¹Department of Financial Accounting, Zhejiang Institute of Economics and Trade, Hangzhou, China

²School of Information, Zhejiang University of Finance and Economics, Hangzhou, China

Email address:

xiyangwxx@163.com (Xiaoxia Wu), scottyong@zufe.edu.cn (Dongqi Yang), wyzhang@e.ntu.edu.sg (Wenyu Zhang),

zhangshuai@zufe.edu.cn (Shuai Zhang)

*Corresponding author

To cite this article:

Xiaoxia Wu, Dongqi Yang, Wenyu Zhang, Shuai Zhang. A Hybrid Ensemble Model for Corporate Bankruptcy Prediction Based on Feature Engineering Method. *International Journal of Information and Communication Sciences*. Vol. 4, No. 3, 2019, pp. 52-58.

doi: 10.11648/j.ijics.20190403.12

Received: August 5, 2019; **Accepted:** September 24, 2019; **Published:** September 27, 2019

Abstract: The bankruptcy of manufacturing corporates is an important factor affecting economic stability. Corporate bankruptcy has become a hot research topic mainly through financial data analysis and prediction. With the development of data science and artificial intelligence, machine learning technology helps researchers improve the accuracy and robustness of classification models. Ensemble learning, with its strong predictive power and robustness, plays an important role in machine learning and binary classification prediction. In this study, we proposed a bankruptcy classification model combining feature engineering method and ensemble learning method, Synthetic Minority Oversampling Technique (SMOTE) imbalanced data learning algorithm is applied to generate balanced dataset, multi-interval discretization filter is applied to enhance the interpretability of the features and ensemble learning method is applied to get an accurate and objective prediction. To demonstrate the validity and performance of the proposed model, we conducted comparative experiments with ten other baseline classifiers, proving that SMOTE imbalanced learning algorithm and feature engineering method with multi-interval discretization was effective. The comparative experiment results show that the ensemble learning method has a good effect on improving the performance of the proposed model. The final results show that the proposed model has achieved better performance and robustness than other baseline classifiers in terms of classification accuracy, F-measure and Area under Curve (AUC).

Keywords: Data Mining, Ensemble Learning, Corporate Bankruptcy, Prediction

1. Introduction

It is of great guiding significance for the national economy to predict the bankruptcy of its corporates, especially the manufacturing corporates. Manufacturing corporates constitute the cornerstone of a country's economic strength and have a significant impact on its overall national strength. Economic decisions about manufacturing corporates also affect countless jobs, suppliers and government taxes. Therefore, it has become a hot topic for researchers to find out the law of bankruptcy of manufacturing corporates and finally predict the plight of corporate.

As is always been, researchers use the most advanced

analytical tools to research the rules of financial statements in order to find out the key to corporate bankruptcy. With the development of data mining and artificial intelligence methods, the data-science approach has entered various research fields, including corporate bankruptcy, and become an effective tool to help decision making.

The main purpose of this research is to explore the effect of different sampling methods on the prediction precision of high imbalance datasets. We purposed an ensemble machine learning model based on Logit Boost [1] algorithm to predict the corporate bankruptcy. We obtained the dataset from the Machine Learning Repository at the University of California Irvine (UCI) <http://archive.ics.uci.edu/ml/index.php>. The dataset are used to validate the model after data balancing and

feature processing steps. The prediction accuracy of the proposed model is compared with Naive Bayes (NB) [2], Logistic Regression (LR) [3], Multilayer Perceptron (MLP) [4], J48 decision tree [5], Random Forests (RF) [6], AdaBoost.M1 [7], Bagging [8], K-Nearest Neighbors (K-NN) [9], Voted Perceptron (VP) [10], proving the superiority of our proposed model.

The remainder of the paper is organized as follows. In Section 2, we review the related work of data mining methods and corporate bankruptcy prediction. In Section 3, an improved ensemble model is proposed for corporate bankruptcy prediction. In Section 4, the experimental results are introduced to prove that the imbalanced data processing algorithm and feature processing steps are effective, and the proposed model performs better than the other baseline models. Sections 5 summarizes and analyzes empirical results and discusses the future work.

2. Related Work

Data mining technology has been widely used in various fields and achieved remarkable performance. In the field of corporate bankruptcy prediction, a large number of corporate financial datasets are collected for analysis and prediction, providing support for corporate operation and decision-making. In 2004, Foster & Stine applied Least-Squares Regression to predict personal bankruptcy [11]. A new predictive approach to corporate failure in a study using AdaBoost learning algorithm to optimize neural network errors and reduce generalized errors by approximately 30% was proposed by Alfaro & Elizondo. (2008) [12]. Nanni & Lumini (2009) [13] conducted a study on the performance of the bankruptcy prediction and credit scoring system based on classifier integration. Li & Miu (2010) proposed a bankruptcy prediction model with dynamic loadings for both the accounting ratio-based z-score and the market-based DD variable [14]. Charitou *et al.* (2011) simplified the Black-Scholes-Merton (BSM) bankruptcy model [15]. Wang *et al.* (2014) proposed a new and improved Boosting, FS-Boosting to predict corporate bankruptcy [16]. Zieba *et al.* (2016) proposed a novel approach for bankruptcy prediction that utilizes Extreme Gradient Boosting for learning an ensemble of decision trees [17]. During the process of prediction, the adverse effect on the performance caused by imbalanced dataset is also one of the important problems. Kim *et al.* (2016) examined the effectiveness of a hybrid method using clustering technique and genetic algorithms based on the artificial neural networks model to balance the proportion between the minority class and majority class [18]. Barboza *et al.* (2017) used an unusually large sample to test the machine learning model, which was used to study corporate bankruptcy, and achieved better prediction accuracy and stability than traditional methods such as Naive Bayes [19]. Jardin (2017) combines method of segmentation with ensemble-based models [20]. Wang *et al.* (2017) proposed a new Kernel Extreme Learning Machine (KELM) parameter tuning strategy [21]. Kliestik *et al.* (2017)

discussed the moral and economic responsibility of corporate bankruptcy and developed a new bankruptcy prediction tool which is superior to traditional models to predict the failure of Slovak corporates [22]. However, few researches combined imbalanced data processing, feature processing and ensemble learning methods together, and most works cannot validate their robustness in multiple aspects.

3. Data Preprocessing and Modeling

In this paper, we proposed an ensemble machine learning model to predict corporate bankruptcy. The dataset about bankruptcy prediction of Polish corporates was collected from Emerging Markets Information Service (EMIS), which is a database containing information on emerging markets around the world. The dataset is available on Machine Learning Repository at the University of California Irvine (UCI) (<http://archive.ics.uci.edu/ml/index.php>) published by Zieba *et al.* (2016) [17]. After imbalanced learning algorithm processing and feature processing of the dataset, we used another ten classifiers to do comparative experiments. Besides, all the experiments are implemented by Weka Data Mining Tool for Java.

3.1. Data Exploration

The dataset contains financial rates from 1st year of the forecasting period and corresponding class label that indicates bankruptcy status after 5 years. The data contains 7027 instances (financial statements), with 65 attributes for each instance. The 65 features and the detailed information is shown in the Table 1. There are 271 bankrupted corporates, 6756 corporates that did not bankrupt in the forecasting period of this dataset. Besides, the imbalance ratio is 24.93.

Table 1. The description of 65 features.

Features	Description
Attr1	Net Profit / Total Assets
Attr2	Total Liabilities / Total Assets
Attr3	Working Capital / Total Assets
Attr4	Current assets / Short-term liabilities
Attr5	[(Cash + Short-term Securities + Receivables-Short-term Liabilities) / (Operating Expenses-Depreciation)] * 365
Attr6	Retained Earnings / Total Assets
Attr7	EBIT / Total Assets
...	...
Attr62	(Short-term Liabilities *365) / Sales
Attr63	Sales / Short-term Liabilities
Attr64	Sales / Fixed Assets
Attr65	Whether or Not the Corporate Ultimately Bankrupt

3.2. Data Preprocessing

In order to evaluate the quality of the process we obtained a dataset of Polish corporates. The selection process involves choosing departments, databases, research phases, the number of firms and the number of financial indicators that will be analyzed. The study samples include bankruptcy and still operating corporates (imbalanced samples).

The quality of data preprocessing is directly related to the

precision of the model. Data preprocessing cannot contaminate data sources. The raw dataset we obtained is a severely imbalanced dataset and contains missing values. So we replaced the missing value with 0.

3.3. Modeling

In this section, we use eleven classification models to predict the bankruptcy status of corporates after 5 years. A flow diagram of our proposed model is provided in the Figure 1.

As shown in the Figure 1, the model is divided into three parts:

Using the Synthetic Minority Oversampling Technique (SMOTE) to deal with the imbalanced dataset in this model.

Processing the high-dimensional dataset by applying the multi-interval discretization filter.

Adopting LogitBoost Classifier for dataset learning and performance evaluation.

The ensemble model is then used to classify and verify on the testing dataset and compare performance with the other ten classifiers.

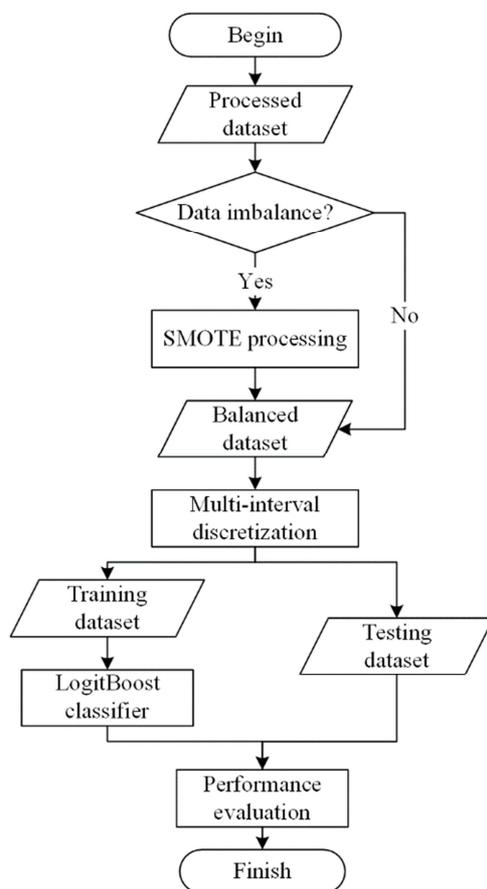


Figure 1. Flow diagram of our proposed model.

3.4. SMOTE Imbalanced Learning Algorithm

In the classification, the imbalance of training data refers to the large number of samples with different features. Actually, imbalanced datasets are common and reasonable. In the case

of corporate bankruptcy, most corporates can survive eventually, and a small part of corporates will fail due to awful management and other reasons.

Chawla et al. (2002) proposed Synthetic Minority Oversampling Technique (SMOTE) as an important technique for dealing with imbalanced datasets [23]. It is an improved scheme based on random oversampling algorithm. The main idea of SMOTE algorithm is to analyze minority samples. New samples are synthesized into the dataset based on the features of minority samples.

3.5. Feature Processing

As the prediction accuracy of the final model will be affected by high-dimensional features, we need to do some processing on the features. The multi-interval discretization as an instance filter was proposed by Fayyad (1993) [24], which divide the range of numerical attributes in the dataset into nominal attributes. The multi-interval discretization builds a better decision tree in the same dataset. Without changing the final result of the algorithm, the heuristic efficiency of the cut point selection can be improved.

A conceptually method is to determine the segmentation point by maximizing the purity of the interval. In practice, however, this method may require manual determination of interval purity and minimum interval size. To solve this problem, some statistical methods are used to separate each attribute value interval. Discretization is applied to attributes which used in classification or association analysis. In general, the effect of discretization depends on the algorithm, as well as other properties used. The transformation of a numerical attribute into a nominal one involves two sub-tasks: determining how many classification values are required, and how to map the numerical attribute values to these nominal values. In the first step, after sorting the numerical attribute values, it divides them into n intervals by specifying n segmentation points. In the second step, all values in an interval are mapped to the same classification value. Therefore, the problem of discretization is to decide how many segmentation points to choose and where to locate the segmentation points.

3.6. Logit Boost

The fundamental idea of the LogitBoost classifier [1] is to integrate multiple simple weak classifiers into a stronger classifier with higher prediction accuracy and better performance. LogitBoost classifier is derived by maximizing logarithmic likelihood function and optimized by Newton iteration.

The performance of many classification algorithms can often be significantly improved by sequentially applying to reweighted versions of the input data, and the weighted majority voting for classifiers sequence. It shows that the phenomenon can be understood as a well-known statistical principle, namely additive modeling and maximum likelihood. For two types of problems, the maximum Bernoulli likelihood can be used as a criterion, and ascension can be regarded as an

approximation of additive modeling on a logical scale. Friedman et al. developed more direct approximations and showed that their results were almost the same [1]. Direct multilevel generalization based on polynomial likelihood proves that, in most cases, the performance is better than other recently proposed multiclass generalization, and in some respects is superior. Its computational speed is faster, making it more suitable for large-scale data mining applications.

4. Experimental Results

In this section, we mainly introduce the experimental setup and comparison of experimental results of our proposed model. The training data and testing data was divided randomly by the ratio of 0.8:0.2. The training data was pre-processed and processed by SMOTE imbalanced learning algorithm. After multi-interval discretization, training data is learned using the LogitBoost classifier and finally validated using testing dataset. After five times experiments, the average performance of our proposed model was obtained. All the experiments were done on a graphics workstation with 3.2GHz Intel Core i5 and 12GB RAM, running Windows 7 professional 64-bit operating system.

4.1. Metrics of Model Performance

The purpose of classification is to construct a classification function or model (that is, classifier) by which data objects are mapped to a given category. The goal of this classifier has only two categories, namely, bankruptcy and non-bankruptcy. For the dichotomy problem, the examples are divided into positive or negative indicators. In the actual situation, there are four cases: 1) true positives (TP): the number of samples that are correctly classified as positive cases; 2) false positives (FP): the number of samples that are wrongly classified as positive cases; 3) false negatives (FN): the number of negative samples that are incorrectly classified; 4) true negatives (TN): the number of negative samples correctly classified.

In this paper, we define bankruptcy as positive and non-bankruptcy as negative. We use the confusion matrix as shown in the Table 2 to represent these four situations.

Receiver Operating Characteristic (ROC) Curve is the abbreviation of receiver operating characteristic curve, also called sensitivity curve. ROC Curve is reflecting the sensitivity and specificity of continuous variable comprehensive index, and is the composition method to reveal the relationship of the sensitivity and specificity.

Table 2. Confusion matrix.

		Prediction	
		Positive (Predict bankruptcy)	Negative (Predict non-bankruptcy)
Real	Positive (Real bankruptcy)	TP	FN
	Negative (Real non-bankruptcy)	FP	TN

We used three metrics in this experiment, they are:

$$Precision = \frac{TP}{FP+TP} \quad (1)$$

$$F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (2)$$

$$FPR = \frac{FP}{FP+TN} \quad (3)$$

$$TPR = \frac{TP}{TP+FN} \quad (4)$$

The Area under Curve (AUC) is a common summary statistic indicator for the goodness of classifier in a binary

classification task. The x-axis of ROC curve represents the false-positive rate (FPR), and the y-axis represents true-positive rate (TPR). Generally, the AUC has a value of between 0.5 and 1.0. The ROC curve has the property that when the distribution of positive and negative samples in the dataset changes, it can remain unchanged. Frequently, the actual dataset is imbalanced, the positive samples are much more than negative samples (or vice versa), and the distribution of positive and negative samples of the test dataset may also change over time. Therefore, AUC value is less sensitive to imbalanced data.

Table 3. Performance of ten baseline classifiers.

Process	Classifier	SMOTE			Non-SMOTE		
		Precision	F-measure	AuROC	Precision	F-measure	AuROC
Discrete	MLP	0.962	0.962	0.945	0.958	0.96	0.842
	AdaBoost. M1	0.957	0.956	0.926	0.938	0.944	0.878
	LR	0.949	0.949	0.913	0.951	0.952	0.844
	VP	0.938	0.932	0.657	0.949	0.946	0.525
	Bagging	0.932	0.921	0.875	0.942	0.944	0.786
	J48	0.924	0.924	0.77	0.95	0.943	0.552
	RF	0.922	0.913	0.888	0.933	0.943	0.781
	K-NN	0.912	0.915	0.702	0.93	0.94	0.615
	NB	0.907	0.801	0.78	0.946	0.824	0.747
	RT	0.899	0.901	0.646	0.933	0.935	0.576

Process	Classifier	SMOTE			Non-SMOTE		
		Precision	F-measure	AuROC	Precision	F-measure	AuROC
Non-Discrete	MLP	0.961	0.961	0.885	0.967	0.968	0.818
	AdaBoost. M1	0.950	0.950	0.930	0.944	0.947	0.866
	LR	0.938	0.931	0.789	0.947	0.950	0.763
	VP	0.869	0.835	0.530	0.929	0.932	0.533
	Bagging	0.937	0.923	0.885	0.954	0.944	0.785
	J48	0.931	0.934	0.766	0.941	0.947	0.672
	RF	0.902	0.903	0.648	0.936	0.943	0.820
	K-NN	0.868	0.890	0.575	0.924	0.940	0.557
	NB	0.868	0.113	0.419	0.910	0.096	0.44
	RT	0.937	0.921	0.904	0.936	0.935	0.613

4.2. Performance of Baseline Classifier

In this section, In order to make a comparison with our proposed model, we presented the performance of the ten baseline classifiers in Table 3. So, the performance of LogitBoost classifier adopted is compared with Naive Bayes (NB) [2], Logistic Regression (LR) [3], Multilayer Perceptron (MLP) [4], J48 decision tree [5], Random Forests (RF) [6], AdaBoost. M1 [7], Bagging [8], K-Nearest Neighbors (K-NN) [9], Voted Perceptron (VP) [10], proving the superiority of our proposed model. We also display the results of dataset that have not been applied by the imbalanced learning algorithm. These results are marked by "Non-SMOTE" label. In addition, the results of discrete feature processing are also shown in the Table 3. These results are marked by "Non-Discrete" label. In order to demonstrate the optimization of the model after imbalanced data learning and discrete feature processing, we present the unprocessed model results as the experimental control group in the Table 3 as well. After five random partitions of the dataset, we obtained the performance results of ten baseline classifiers.

From Table 3, we found that, no matter whether the SMOTE imbalanced learning technique and multi-interval discretization feature processing are applied, MLP and AdaBoost. M1 model perform well.

4.3. The Performance and Analysis of the Proposed Model

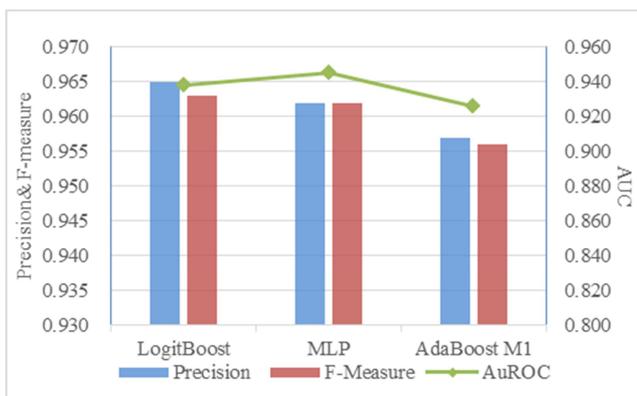
In this section, we tested the performance of our proposed model and make a comparison between our proposed model and the best baseline classifier. Similarly, for the ensemble model we proposed, we randomly split the dataset five times and got the performance results of our proposed model.

In Table 3, we compared the model performance of the raw dataset, the dataset processed by SMOTE imbalanced learning technique, the dataset processed by multi-interval discretization and the dataset processed by both SMOTE imbalanced learning technique and multi-interval discretization.

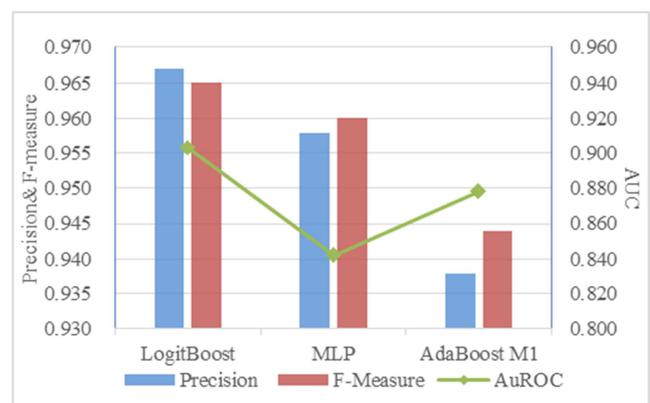
The best result of ten baseline classifiers are recorded in Table 3 and compared with the performance of our proposed model.

Table 4. Performance of our proposed classifier and the best baseline classifiers.

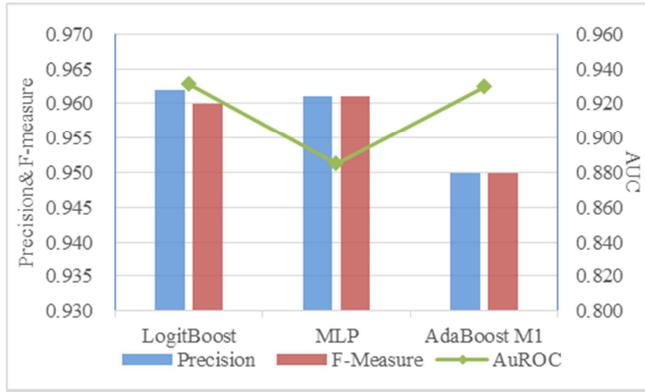
Process	Classifier	SMOTE			Non-SMOTE		
		Precision	F-measure	AuROC	Precision	F-measure	AuROC
Discrete	Best classifier result	0.965	0.963	0.945	0.967	0.965	0.903
		(LogitBoost)	(LogitBoost)	(MLP)	(LogitBoost)	(LogitBoost)	(LogitBoost)
Non-Discrete	Best classifier result	0.962	0.961	0.931	0.967	0.968	0.887
		(LogitBoost)	(MLP)	(LogitBoost)	(LogitBoost / MLP)	(MLP)	(LogitBoost)



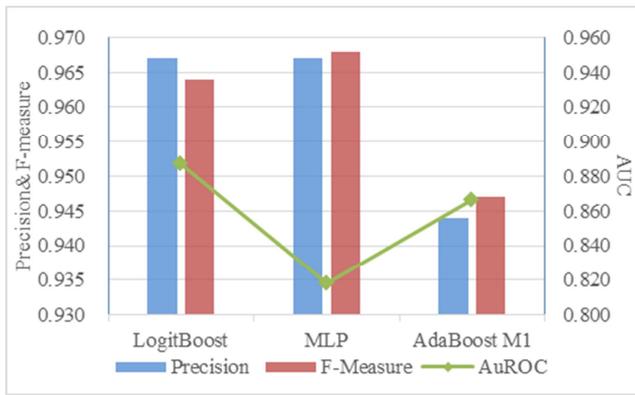
(a). Performance of classifiers after SMOTE algorithm and multi-interval discretization method is applied



(b). Performance of classifiers after multi-interval discretization



(c). Performance of classifiers after SMOTE algorithm is applied



(d). Performance of classifiers on the raw dataset

Figure 2. Performance of classifiers with and without the processing on the raw dataset.

The dataset was processed by the same imbalanced learning technology and discrete feature processing, and the performance results of the dataset were compared with those of the non-discrete processing and imbalanced data learning. The performance and comparison results of our proposed model are shown in Table 4.

We found that the AUC value of the baseline classifier and the model we proposed have been improved after SMOTE treatment. Certainly, the higher the AUC value, the better the classifier is to classify, indicating the higher accuracy of the classifier. Besides, the performance of our proposed model has been improved comprehensively after the raw dataset has been processed by the multi-interval discretization. For the dataset that has been processed by SMOTE imbalanced learning technique, the performance of the model can be improved continuously and comprehensively by using multi-interval discretization method for feature processing. However, for the dataset that has been processed through data discretization only, the SMOTE method can only improve the performance of AUC value, while the precision value and F-measure value are slightly decreased.

Two of the best performing baseline classifiers are selected to compare with our proposed ensemble learning classifier. The results of the comparison are shown in Figure 2.

As we can see, Figure 2 (a) shows the performance of dataset processed by SMOTE imbalanced learning technique

and the multi-interval discretization method processing in the three classifiers. Figure 2 (b) shows the performance of the dataset after multi-interval discretization method processing in the three classifiers. Figure 2 (c) shows the performance of dataset learned through SMOTE imbalanced learning technique in three classifiers, and we find that AUC value has been improved. Figure 2 (d) shows the performance of dataset without any data preprocessing and feature processing in the three classifiers. We find that compared with the dataset without any data processing, the performance of the dataset after multi-interval discretization is improved in three classifiers. Compared with the previous performance, we find that the performance of three classifiers is improved to varying degrees, and the comprehensive performance of our proposed LogitBoost classifier is better than the other two baseline classifiers.

5. Conclusion

Bankruptcy is a difficult problem for the global business community. Manufacturing corporates produce discrete products, which are closely related to production activities and people's life. Nowadays, manufacturing corporates are not only the key link for scientific discovery and technological invention to transform into real scale productivity, but also have close relations to a large number of related industries. How to predict the bankruptcy according to the financial statement data of corporates and avoid the operational risk has become an urgent problem. With the rise and development of data science, it has become a hot topic for researchers to analyze and predict the corporates' bankruptcy by data mining. Machine learning, as a mainstream method of artificial intelligence, is playing an increasingly important role in today's research.

In this paper, we propose a model based on ensemble learning to predict the bankruptcy of corporates. We obtained the dataset about the bankruptcy of Polish manufacturing corporates, and used the ensemble learning model for the prediction after SMOTE imbalanced learning algorithm and multi-interval discretization method processing. In order to verify the validity of LogitBoost classifier, we used ten other models as the baseline classifier for comparative experiments. Compared with the baseline classifiers, the validity of our proposed model in terms of precision, F-measure and AUC value was improved. We further compare the results of the proposed model with the model that have not been processed by SMOTE imbalanced learning algorithm, demonstrating the necessity of the balancing algorithm for serious imbalanced dataset. In addition, we compared the proposed model with the model without multi-interval discretization method process, and the result shows that the performance of the proposed model is improved by the application of multi-interval discretization method.

However, there are some shortcomings in our proposed model. First, the corporate's financial data may contain a large number of missing values and abnormal information, which may affect the performance of classifiers. Due to the high

dimensional characteristics of financial statement data, we should improve the feature processing and selecting method. For example, we could consider using principal component analysis or linear discriminant analysis to first reduce the dimension of features and then the feature selection. In the process of model learning, there may be problems of local optimal solution and over-fitting, which will be found and improved in future work.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 51875503, No. 51475410, No. 51775496) and Zhejiang Natural Science Foundation of China (No. LY17E050010).

References

- [1] Friedman, J., Hastie, T., & Tibshirani, R., "Additive logistic regression: a statistical view of boosting." *The Annals of Statistics*, vol. 28, no. 2, 2000, pp. 400-407.
- [2] John, G. H., & Langley, P., "Estimating continuous distributions in bayesian classifiers." In *Proceedings of the 11th Conference on Uncertainty in Artificial Intelligence*, Montreal, Canada, Aug 18-20, 1995, pp. 338-345.
- [3] Cessie, S. L., & Houwelingen, J. C. V., "Ridge estimators in logistic regression." *Journal of the Royal Statistical Society*, vol. 41, no. 1, 1992, pp. 191-201.
- [4] Back, B., Laitinen, T., & Sere, K., "Neural networks and genetic algorithms for bankruptcy predictions." *Expert Systems with Applications*, vol. 11, no. 4, 1996, pp. 407-413.
- [5] Quinlan, J. R., "C4.5: programs for machine learning." Morgan Kaufmann Publishers Inc., 1993.
- [6] Breiman, L., "Random forests." *Machine Learning*, vol. 45, no. 1, 2001, pp. 5-32.
- [7] Freund, Y., & Schapire, R. E., "Experiments with a new boosting algorithm." In *Proceedings of the 13th International Conference on Machine Learning*, Bari, Italy, July 3-6, 1996, pp. 148-156.
- [8] Breiman, L., "Bagging predictors." *Machine Learning*, vol. 24, no. 2, 1996, pp. 123-140.
- [9] Aha, D. W., Kibler, D., & Albert, M. K., "Instance-based learning algorithms." *Machine Learning*, vol. 6, no. 1, 1991, pp. 37-66.
- [10] Freund, Y., & Schapire, R. E., "Large margin classification using the perceptron algorithm." *Machine Learning*, vol. 37, no. 3, 1999, pp. 277-296.
- [11] Foster, D. P., & Stine, R. A., "Variable selection in data mining: building a predictive model for bankruptcy." *Journal of the American Statistical Association*, vol. 99, no. 466, 2004, pp. 303-313.
- [12] Alfaro, E., & Elizondo, D., "Bankruptcy forecasting: an empirical comparison of adaboost and neural networks." *Decision Support Systems*, vol. 45, no. 1, 2008, pp. 110-122.
- [13] Nanni, L., & Lumini, A., "An experimental comparison of ensemble classifiers for bankruptcy prediction and credit scoring." *Expert Systems with Applications*, vol. 36, no. 2, 2009, pp. 3028-3033.
- [14] Li, M. Y. L., & Miu, P., "A hybrid bankruptcy prediction model with dynamic loadings on accounting-ratio-based and market-based information: a binary quantile regression approach." *Journal of Empirical Finance*, vol. 17, no. 4, 2010, pp. 818-833.
- [15] Charitou, A., Dionysiou, D., Lambertides, N., & Trigeorgis, L., "Alternative bankruptcy prediction models using option-pricing theory." *Journal of Banking & Finance*, vol. 37, no. 4, 2011, pp. 2329-2341.
- [16] Wang, G., Ma, J., & Yang, S., "An improved boosting based on feature selection for corporate bankruptcy prediction." *Expert Systems with Applications*, vol. 41, no. 5, 2014, pp. 2353-2361.
- [17] Zieba, M., Tomczak, S. K., & Tomczak, J. M., "Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction." *Expert Systems with Applications*, vol. 58, no. C, 2016, pp. 93-101.
- [18] Kim, H. J., Jo, N. O., & Shin, K. S., "Optimization of cluster-based evolutionary undersampling for the artificial neural networks in corporate bankruptcy prediction." *Expert Systems with Applications*, vol. 59, 2016, pp. 226-234.
- [19] Barboza, F., Kimura, H., & Altman, E., "Machine learning models and bankruptcy prediction." *Expert Systems with Applications*, vol. 83, 2017, pp. 405-417.
- [20] Jardin, P. D., "Dynamics of firm financial evolution and bankruptcy prediction." *Expert Systems with Applications*, vol. 75, 2017, pp. 25-43.
- [21] Wang, M., Chen, H., Li, H., Cai, Z., Zhao, X., & Tong, C., et al., "Grey wolf optimization evolving kernel extreme learning machine: application to bankruptcy prediction." *Engineering Applications of Artificial Intelligence*, vol. 63, 2017, pp. 54-68.
- [22] Kliestik, T., Misankova, M., Valaskova, K., & Svabova, L., "Bankruptcy prevention: new effort to reflect on legal and social changes." *Science & Engineering Ethics*, vol. 4, 2017, pp. 1-13.
- [23] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P., "Smote: synthetic minority over-sampling technique." *Journal of Artificial Intelligence Research*, vol. 16, no. 1, 2002, pp. 321-357.
- [24] Fayyad, U., "Multi-interval discretization of continuous-valued attributes for classification learning." In *Proceedings of the 13th International Joint Conference on Artificial Intelligence*, Chambery, France, Aug 28-Sept 3, 1993, pp. 1022-1027.