

An Enhanced Bankruptcy Prediction Model Using Fuzzy Clustering Model and Random Forest Algorithm

Sarojini Devi Saladi*, Radhika Yarlagadda

Dept of CSE, GITAM University, Visakhapatnam 530045, India

Corresponding Author Email: ryalavar@gitam.edu



<https://doi.org/10.18280/ria.350109>

ABSTRACT

Received: 25 October 2020

Accepted: 10 January 2021

Keywords:

bankruptcy prediction, random forest classifiers, MapReduce techniques, big data analytics, and fuzzy modeling.

The rapid advancements made in Information Technologies (IT) have evolved in the prediction model for financial data. Research in the prediction of bankruptcy is inclining to owe to the growth of related associations with economic and social phenomena. Financial crises in recent scenarios have influenced the growth of financial institutions. Hence, the need for bankruptcy risk prediction at an earlier stage is of prime importance. Though several prediction algorithms were suggested, the predictive models' accuracy is still a challenging task. In this paper, a bankruptcy prediction model is developed by integrating the Fuzzy clustering model and Multi-objective random forest classifiers. The voluminous number of records of the financial dataset and polish bankruptcy dataset is collected from a public repository. It is pre-processed using a MapReduce technique, one of the Big Data approaches. Benefits given by big data approaches help to achieve better flexibility towards a variable declaration. The collected records are pre-processed and organized under efficient index construction. FCM is employed to cluster the data for analytic purposes. Finally, a multi-objective Random Forest classifier helps to develop a prediction model for bankruptcy. Experimental analysis is carried out with accuracy, precision, sensitivity, and specificity compared with existing, Genetic algorithms. Compared to the existing technique, the proposed technique has obtained 80% accuracy.

1. INTRODUCTION

Bankruptcy is a critical issue in financial concerns. The distress caused by financial failure disturbs the firm's liabilities that are disproportionate to its assets. This loss could degrade the capabilities of stockholders, employees, customers, and investors. Hence, the prediction of bankruptcy plays a significant role in economics and finance research fields. The prediction process is to make decision systems related to business operations in a reliable manner. Most of the business entities wish to have relationships in the long term [1]. Hence, it is valuable to foresee the possibility of the bankruptcy of a business customer (or) partner. To enhance the accuracy of bankruptcy prediction, researchers concentrated on two paths: first, exploring the significant variables and enhancing the methodologies of the bankruptcy prediction and its computation infrastructure.

Decision-makers have suggested several mathematical models, hypothetical models, and soft computing techniques to measure the bankruptcy rate in stipulated periods and depicting the financial ratios. In specific, prediction models such as single and multivariate models are used statistically based on assumptions. However, there is no proper solution to predict bankruptcy due to its characteristics of a data structure [2]. High rates of business failures are observed in the manufacturing and construction field. Continuous improvement in the research for predicting bankruptcy, particularly in constructing the models for predicting bankruptcy, is highly essential. To enhance the prediction, overall reliability, and performance, vast quantities of data are

required [3, 4]. Therefore, big data analytics can be used for this kind of research.

Big data approaches observe a high recognition and attention because of their broad prospects of application and research. It refers to rapid and large-scale data in different types. Sets of data are combined from various sources in a semi-structured, structured, and unstructured data collection, as pointed out by Min and Lee [5]. Apart from these, models developed for predicting bankruptcy are adopted as a baseline model for prediction by the nation's government for determining the construction industry performance, determining complicated areas of the field, and preparing needed policies for tackling these difficulties. Therefore, this research develops a model for predicting bankruptcy based on big data analytics with specific reference to the construction business.

The rest of the paper is organized as follows: Related work is given in Section II; Proposed Methodology is given in Section III; Experimental results and analysis are elaborated in Section IV, and at last, Conclusion of the study is presented in Section V.

2. RELATED WORK

This part describes the prior techniques related to the topic. The effects of bankruptcy were described by several researchers in financial sectors [6]. Most of the studies were carried out by supervised approaches. The data is collected, as per the supervised approach, to design predictive models.

Deep analyses were taken for exploiting the statistical methods such as Partial Least Square Discriminant Analysis (PLS-DA) [7]. Multicollinearity [8] strategies were also utilized to meet the requirements of econometrics with better data availability. The semi-supervised learning algorithms were not analyzed for better prediction systems [9]. It forms multi-view theory [10] for better training algorithms. Generally, it requires two learners, namely, redundant views and multi-views, which expands the training set of other learners. Then, again it is refined by updated L subsets [11]. The Random Subspace Method for CO-training (RASCO) [12] was suggested to enhance random decision-making rules. At last, the estimated relevancy scores of the testing data, data are being classified.

Extreme Learning Machine (ELM) classifiers were suggested to resolve the issues of Artificial Neural Networks (ANNs). It is a kind of neural network using hidden layer feedforward neural networks. It predicts the records for continuous data-flowing environments. Many fields have utilized neural networks to improve the accuracy of the classifier [13]. It has been moved to financial data analysis, corporate lifecycle prediction, and corporate credit ratings [14, 15]. Grey Wolf Optimizer (GWO) [16] has been applied to bankruptcy prediction due to its searching capability. It was extended in training models like q-Gaussian Radial Basis Functional Link using GWO for better meta-heuristic models [17].

The companies that hold high bankruptcy risks were analyzed for the performance estimation of machine learning algorithms such as neural networks, SVM, regression, and Bayesian approaches [18]. By using 10-fold cross-validation, the classification accuracy of the ML algorithms was analyzed. ML algorithms that hold the best accuracy were considered for better prediction models. Some studies checked the performance of ML algorithms with statistical learning models [19]. The results stated that Linear Discriminant Analysis (LDA) performed better than artificial intelligence approaches. Then, a 10-K validation was done to reduce the overfitting issues.

The study found that, contrary to many research findings, there was no significant difference between the two approaches, linear discriminant analysis and regression techniques [20]. The effectiveness of various approaches using active learning models [21] was studied for corporate bankruptcy prediction. Financial data of Greek companies were explored to improve the accuracy of learning and its prediction performance. Likewise, a study done by Situm [22] assessed the firm size and age as predictors variables to find the bankruptcy companies. Finally, the results stated that there is no significant difference in firm size and age to predict bankruptcy.

Table 1. Comparison of existing work

Reference	Technique	Results	Merits/demerits	Scope for further research
[7]	Decision trees	The developed technique performed better than some of the existing methods	Computational complexity	Evaluating the technique on real-world datasets
[12]	C&R Tree model	The model gave an accuracy of 90%	The developed technique is effective in detecting bankruptcy	Investigation of the data to determine the hidden patterns in accountancy, including fraud discovery, risk management, and share basket analysis.
[18]	PSO-SVM	The model gave an accuracy of 94.4%	PSO-SVM achieves better performance compared to SVM, GA-SVM	Implementation is complicated when the number of data sets is increased.
[23]	SPSO-SVM	The model gave an accuracy of 99.2%	In comparison with other techniques, this was better concerning accuracy as well as computation time	Inclusion of the developed SPSO-SVM for the problems of control and filtering.
[24]	Hybrid SVM	The model gave correct classification accuracy of 75%	The computation time of the model is high	Investigation on if the macro-economic data can help improve prediction accuracy in particular segments of the industry.
[26]	PLS-LR	The model gave an accuracy of 94.5%	The unavailability of specific data reduced the performance of regression.	results of our study can be improved by increasing the sample size
[27]	Logit, SVM, and Semi-supervised-SVM	The model gave an accuracy of 78.7%, 89.1%, and 94.1%	S3VM performed better than the other techniques	Introducing more SVM algorithms such as cS3VM and S3VM light to reject inference.
[29]	ANN	The model gave an accuracy of 85.3%	Obtained good performance with better prediction accuracy.	It requires more training data sets.
[30]	FSVM	The model gave an accuracy of 98.28%	For most cases, SVM gives better performance compared to FSVM	Modification in the developed approach to testing it for the dataset of larger size.

Corporate bankruptcy prediction has become a fundamental analysis due to financial debts. Instance case-based learning was suggested for sample estimation using TOPSIS [23], which resolved the disadvantages of k-NN classifiers. Hybrid algorithms and quantization models were constructed, which minimized the costs [24, 25]. The asymmetric costs of two types of errors were combined into the evolutionary classification model used to predict bankruptcy. The genetic algorithm is used to search the solution space for an optimal representation of the LVQ technique.

Since it's time-oriented data, variants of machine learning algorithms have been analyzed for predicting the behaviors of the customers [26]. Social media datasets were discussed to find the behavioral link among the users. Hashtags and sentimental labels were developed to construct recommender systems [27]. Artificial neural networks have a powerful influence over prediction systems. The neural network models have been analyzed for dynamism in bankruptcy that leads to prediction accuracy. Hidden Markov Models (HMM) and Markov chains were combined with ANNs to predict the dynamic behavior of the firms. The transition between different ratings related to rating systems helped to sort out dynamic credit systems. Some models insisted on fine-tuning of component systems. Performing univariate analysis concentrated on leveraging the financial ratios via boosted algorithms. Linear optimization models [28] were derived to classify the companies based on financial assets. Empirical approaches have been enhanced under restrictive models that shall accommodate time-dependent categorical variables. Additionally, statistical techniques such as neural networks, SVM, and genetic algorithms were developed as hybrid techniques to devise the data distribution.

Reviews stated that the application of nearest neighbor algorithms [29] had yielded positive results in bankruptcy prediction. The authors could resolve the data distribution issue and have found non-linear relation sets. This kind of analysis assists in finding likelihood and exploratory variables which is non-linear. Saturation effects between financial ratios and target variables enhanced the prediction accuracy. Multiplicative factors defined in independent variables also have better data prediction systems. Support vector machine (SVM) [30] has also discussed non-linear class boundaries that mapped the lower dimensional into high dimensional feature space. Here, an optimal hyperplane was constructed that defined the decision classes. Since it's data-driven machine learning methods, it helped to resolve various applications. Hybrid intelligent algorithms that developed solutions for ensemble-based prediction models with absolute accuracy. The comparison of research contribution from various researchers is shown in Table 1.

3. PROPOSED METHODOLOGY

In this part, the proposed framework is explained. The study aims to predict bankruptcy using improved random forest classifiers. The proposed phases are explained as follows:

3.1 Data accumulation

Data accumulation is a foremost step that helps to achieve cleaned data. Here, two datasets are collected from a public repository, namely, Machine learning databases and Polish companies' bankruptcy. The Polish companies' bankruptcy

dataset comprises 18 attributes with labels (0,1), whereas machine learning databases consist of 18 attributes with binary labels (0, 1). The list of attributes is shown in Table 2.

Table 2. Presents the details of datasets

S. No	Attribute	S. No	Attribute
1	Company_Profit	10	Share_Holder_Equity_Ratio
2	Company_Asset	11	Gross_Profit
3	Short_Term_Finance	12	Acid_Ratio
4	Current_Ratio	13	Dept_Ratio
5	Short_Term_Year	14	Revenue
6	Retained Earnings Ratio	15	Gross_Profit_Margin
7	Total_Asset	16	Inventory_Turnover_Ratio
8	Share_Holder_Equity	17	Net_Profit_Ratio
9	Asset_Turnover_Ratio	18	Working_Capital

The Description and statistics of the Polish bankruptcy company dataset are shown in Table 3.

Table 3. Details of datasets

Dataset name	Years	No. of attributes	Class label	No. of records
Polish bankruptcy companies	5 years	18	(0,1)	15,470

3.2 Pre-processing

The collected dataset is generally prone to irrelevant data like missing data, repeated data, and empty data. These data are analyzed using the MapReduce technique, which ensures the quality of data. The working of the MapReduce technique is as follows:

Job 1 and job two pre-processing descriptions: Here, there are two tasks, namely, map and reduce. The map task in job one works based on key fields while the other works on acceptance rules. In job 2, the reducer task includes non-key fields, which are processed in two steps: comparing with involved fields and comparing with non-involved fields.

Indexing: It works in a single key field which lookup for duplicates by comparing with previous fields. Likewise, missing data and empty data are replaced and registered with Map of Job 1. In some cases, if a record is being matched with several key fields, then set it as the primary key. This process is followed for all records.

Indexing by more than one key field: Map task of job one computed as N pairs (key, value) for each record, where N is the number of selected vital fields. In case of similarity between two records, it is labeled as duplicate data. The combiner function in the map will not permit redundant data to job 2. In some cases, a combiner in reduces function lowers the impacts of defining key fields to similar data.

Gathering matching process: It just gathers the records by comparing its field to previous records.

3.3 Fuzzy clustering Model (FCM)

Here, pre-processed data is clustered into its relevant groups using the FCM technique. The membership is defined for all clusters. Selecting a relevant cluster for data is a difficult task. Data is clustered using fuzzifier constants which are defined

by minimum change and maximum row process. When any changes prevail in objective functions, then the entire algorithm is terminated. Data is being checked based on objective functions whether it falls under defined membership (or) not. It consists of four clusters. Namely, the company didn't pay, the company can't pay, the company paid (without profit), and the company paid (with profit). Each record belongs to one or more clusters. The steps to be followed in the fuzzy clustering process are:

- (1) Each record is partitioned into n elements, such as $X = \{X_1, X_2 \dots X_n\}$ with C clusters.
- (2) For a given set of records, the fuzzy modeling returns clusters $C = \{C_1 \dots C_c\}$ and its partition matrix.
- (3) The weight matrix is estimated as $W = W_{ij}$, $j \in [0, 1]$ where W_{ij} , j tells the degree of record to its clusters.
- (4) Then, the objective function is minimized as follows:

$$\arg \min_c \sum_{i=1}^n \sum_{j=1}^c w_{ij}^m \| \mathbf{x}_i - \mathbf{c}_j \|^2 \quad (1)$$

where,

$$w_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\| \mathbf{x}_i - \mathbf{c}_j \|^2}{\| \mathbf{x}_i - \mathbf{c}_k \|^2} \right)^{\frac{2}{m-1}}} \quad (2)$$

The cluster centroid analysis helps grouping similar data points nearer to centroid value with weights under the degree of membership. Increased weight value defines that data point closer to the clusters and vice versa. Control parameters and termination criteria determine the efficiency of the cluster centroids. This algorithm ensures the flexibility of the data points that belong to more than one cluster. At last, the records are clustered according to their labels.

Prediction process: Here, a novel multi-objective random forest classifier is employed to predict bankruptcy and non-bankruptcy firms. The aggregation of decision trees is known as Random forest. At first, the clustered records are classified into trained and tested records. Based on given binary values (0,1), then the relevant classes are predicted. Before moving onto the multi-objective, let us first define the single objective for supervised estimation problems.

Preparing the random trees in a single objective: Let us assume a set of N records, which is denoted as a d-dimensional vector, and c is the label of the single-objective random trees. It is given as:

$$S_0 = \{(\mathbf{x}_i, y_i) \mid i = 1, \dots, N\} \quad (3)$$

where, x and y are the label characteristics of the single objectives. The random tree is formed by the splitting criterion, which compares and the split into two dimensions, as

$$h(\mathbf{x}; \theta_1, \theta_2) = \begin{cases} 1 & \text{if } x_{\theta_1} > x_{\theta_2} \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

The random tree is constructed by splitting the possible sets based on mutual information. The left and right child of a tree are initialized by splitting functions

$$\begin{aligned} S_i^L(\theta) &= \{ \mathbf{x} \in S_i \mid h(\mathbf{x}; \theta_1, \theta_2) = 1 \} \\ S_i^R(\theta) &= \{ \mathbf{x} \in S_i \mid h(\mathbf{x}; \theta_1, \theta_2) = 0 \} \end{aligned} \quad (5)$$

The above Eq. (5) assures the quality of the data in the sets S_i . It is measured by the information gain, which is given in Eq. (6).

$$I_i(\theta) = H(S_i) - \sum_{j \in \{L,R\}} \frac{|S_i^j(\theta)|}{|S_i(\theta)|} H(S_i^j(\theta)) \quad (6)$$

where, characteristic entropy in set S is given as H(S)

By estimating the information gain for all possible sets S_i , the split function generates the trees for each class. In the continuous sampling of the Eqns. (5) & (7), the random trees are constructed. In each internal tree, the nodes are arranged by maximum information gain. For depth of a tree, $T = d(d-1)$, then the overall split of the tree is given in Eq. (7)

$$\theta_i^* = \arg \max_t I_i(\theta_t) \quad (7)$$

Formation of a forest from the random trees: Based on the splitting criterion, the nodes under random tree T is collectively represented as,

$$\mathcal{T} = \{ \theta_i \}_{i=1}^{i=|\mathcal{T}|} \quad (8)$$

To avoid the diversity and the fitting issues over the tree, it is grouped like a forest. A forest of trees is denoted as $F = (T_1, T_f)$. The regularization of each tree under a forest is independently trained.

Measuring the test data under single-objective RF: Let X be the unlabelled test data. A node N to be placed in a tree T takes hierarchical tests under the arrival of leaf nodes. The leaf node denotes it as $l(\mathbf{x}; T)$. Finally, the unlabelled y of X with the training elements S_i is represented as

$$\text{label}(\mathbf{x}; \mathcal{F}) = \arg \max_y \frac{1}{|\mathcal{F}|} \sum_{T \in \mathcal{F}} p(y \mid L(\mathbf{x}; T)) \quad (9)$$

where, $L(\mathbf{x}; T)$ = Possible set of training samples under leaf L in tree T. Formation of multi-objective estimation with random forests: The training sample x_i of the sets S_i under single-objective $C > 1$, will be estimated as,

$$S_0 = \left\{ \left(\mathbf{x}_i, \{y_i^j\}_{j=1}^c \right) \mid i = 1 \dots N \right\} \quad (10)$$

The classification problem becomes simplified when the trees under a single-objective are regularized. The measure of Information gain on sets S_i distinguishes the multi-objective from single-objective. It is also to be assured that the information gain in one set S_i is not related to information gain in another set S_i . Thus, information gain for the multi-objective random forest is given as, Ilw

$$I_{lw}(S_i, \theta) = \max_c \frac{H_c(S_i)}{H_c(S_0)} I^c(S_i, \theta) \quad (11)$$

The above Eq. (11) represents a normalized measure of weighted information gain. Here, the characteristic c of each objective will be analyzed between the local entropy S_i and the entropy of a root, $c(S_0)$. The split function θ will frequently update the local information gain during training samples. The characteristic c is also intelligently selected by each node of each tree. Finally, the frequent updation of Ilw will help to

classify the training samples. Figure 1 presents the workflow of the proposed study.

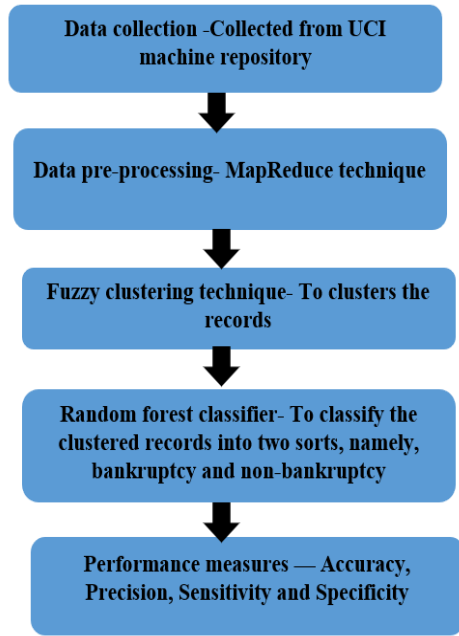


Figure 1. Process flow of the proposed framework

4. EXPERIMENTAL ANALYSIS

Here, the experimental settings of the proposed framework are explained. The dataset is collected from a public repository, namely, polish bankruptcy companies that compose five years of records with 18 attributes and implemented in the Weka tool. The proposed MORF classifier with and without clustering is compared with the different existing models like Naive Bayes, RF without clustering, and RF with clustering.

The followings are the performance indexes used for analytic purposes. It is explored in confusion matrix such as,

- True Positive (TP) = If the observed cases and predicted cases are matched correctly, it is known as True Positive.
- True Negative (TN) = If both the observed and predicted case is negative, it is known as True Negative.
- False Positive (FP) = If the observed case is negative and the predicted case is positive, it is known as False Positive (FP).
- False Negative (FN) = If both the observed and predicted case is negative, it is known as False Negative (FN).

The performance metrics studied are explained as follows:

Accuracy: It is a testing model that differentiates the bankruptcy and non-bankruptcy companies correctly. It estimates the proportion of true positive and true negative in all evaluated cases. It is given as follows:

- Accuracy= (TP+TN) / (TP + TN + FP + FN)

Sensitivity: It is the ability to determine the bankruptcy companies correctly. The proportion of true positives determines the sensitivity of the systems.

- Sensitivity: (TP)/ (TP+FN)

Specificity: It is the ability to determine the non-bankruptcy companies correctly. The proportion of true negatives determines the specificity of the systems. It is given as follows:

- Specificity= (TN) / (TN+FP)

Precision: It is the ability to determine the true positive rate, and it's given as:

- Precision= (TP) / (FP+TP)

Below Table 4 presents the comparative analysis between existing and proposed systems. It is inferred that proposed Multi-objective Random Forest classifiers work better than existing models. The performance results obtained are shown in Figure 2.

Table 4. Comparison values between existing and proposed systems

Datasets	Model	ACCURACY (%)	SENSITIVITY (%)	PRECISION (%)	SPECIFICITY (%)	Remarks
Financial datasets from Data.gov.in (250 records, 6 attributes & 2 class labels)	Naïve Bayes Model	78.66	53.72	46.25	44.8	Unable to analyze the vast datasets
Polish bankruptcy dataset (mixed Data of all countries), From UCI Machine Learning Repository (15458 records, 18 attributes & 2 class labels)	RF without Clustering	75.47	52	48.204	47.4	Needs clustering the data to improve the accuracy
Polish bankruptcy dataset (mixed Data of all countries) From UCI Machine Learning Repository (15458 records, 18 attributes & 2 class labels)	RF with Clustering	76.28	54.23	46.369	42.5	Needs to enhance the random forest algorithm to get high accuracy
Polish bankruptcy dataset (mixed Data of all countries) From UCI Machine Learning Repository (15458 records, 18 attributes & 2 class labels)	MORF without clustering	78.25	55.68	48.36	43.2	Clustering needed to get high accuracy
Polish bankruptcy dataset (mixed Data of all countries) From UCI Machine Learning Repository (15458 records, 18 attributes & 2 class labels)	MORF with clustering	80.821	56.234	49.314	44.3	Got high accuracy than compare to other models

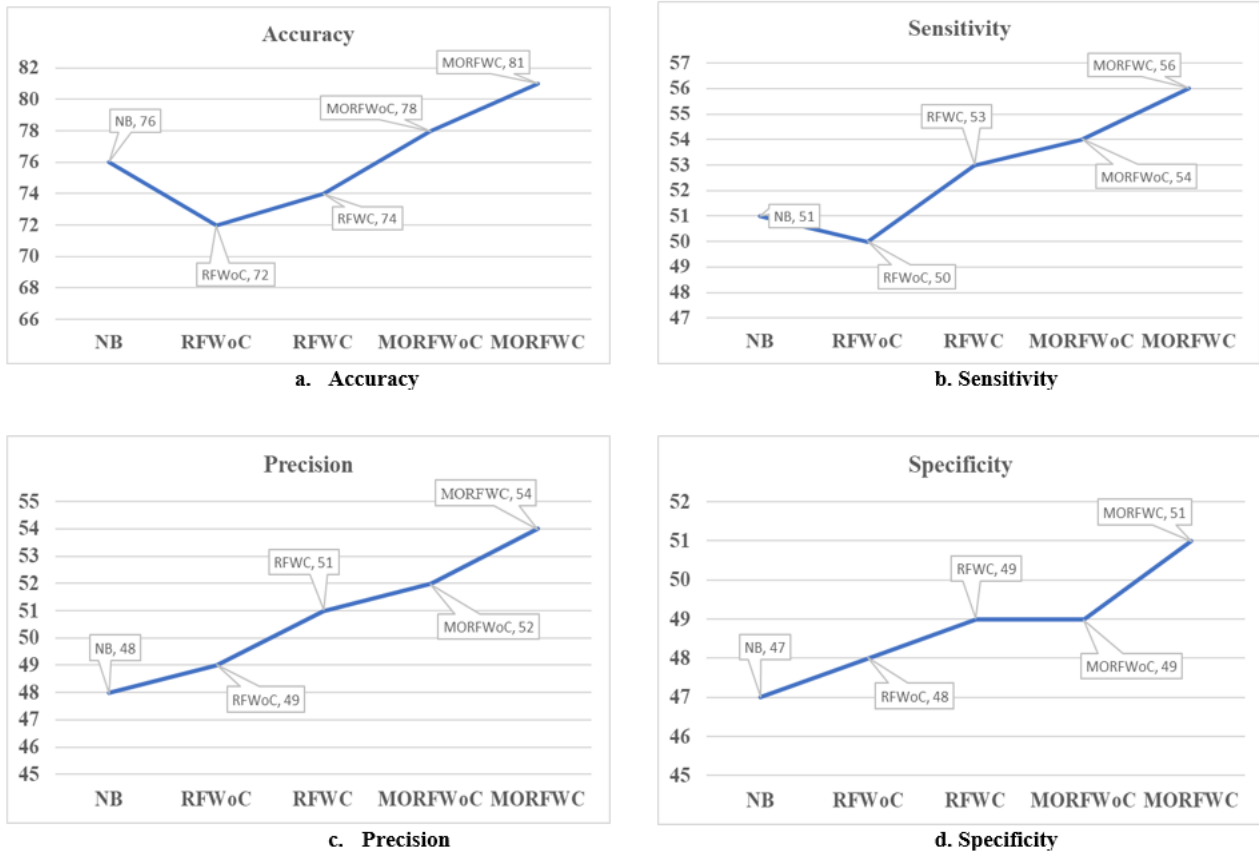


Figure 2. Comparison analysis

5. CONCLUSIONS

This paper is an attempt at predicting bankruptcy specific to the construction business. Financial datasets and polish bankruptcy datasets were collected from public machine repositories. Here, the MapReduce technique is a pre-processing process that eliminates the missing, empty, and redundant data. This helps to speed up the prediction process. Then, fuzzy clustering is employed to categorize the processed data into four sorts; namely, the company didn't pay, the company can't pay, the company paid (without profit), and the company paid (with profit). This clustering process assists in fixing the class labels, as 'bankruptcy' and 'non-bankruptcy'. Finally, a Multi-objective Random Forest (MORF) classifier is adopted to predict the class labels by constructing efficient random trees with single-objectives. This predicted system is further compared with different models in which the proposed classifier has given 80% accuracy and proves its effectiveness in real-time possibilities.

REFERENCES

- [1] Zoričák, M., Gnip, P., Drotár, P., Gazda, V. (2020). Bankruptcy prediction for small-and medium-sized companies using severely imbalanced datasets. *Economic Modelling*, 84: 165-176. <https://doi.org/10.1016/j.econmod.2019.04.003>
- [2] Drotar, P., Gnip, P., Zoričák, M., Gazda, V. (2019). Dataset of financial ratios of Slovak companies. *Mendeley Data*, V2. <https://doi.org/10.17632/j89csb932y.2>
- [3] Labrinidis, A., Jagadish, H.V. (2012). Challenges and opportunities with big data. *Proceedings of the VLDB Endowment*, 5(12): 2032-2033. <https://doi.org/10.14778/2367502.2367572>
- [4] Tseng, F.M., Hu, Y.C. (2010). Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks. *Expert Systems with Applications*, 37(3): 1846-1853. <https://doi.org/10.1016/j.eswa.2009.07.081>
- [5] Min, J.H., Lee, Y.C. (2005). Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters. *Expert Systems with Applications*, 28(4): 603-614. <https://doi.org/10.1016/j.eswa.2004.12.008>
- [6] García, V., Marqués, A.I., Sánchez, J.S. (2015). An insight into the experimental design for credit risk and corporate bankruptcy prediction systems. *Journal of Intelligent Information Systems*, 44(1): 159-189. <https://doi.org/10.1007/s10844-014-0333-4>
- [7] Tian, S., Yu, Y., Zhou, M. (2015). Data sample selection issues for bankruptcy prediction. *Risk, Hazards & Crisis in Public Policy*, 6(1): 91-116. <https://doi.org/10.1002/rhc3.12071>
- [8] Serrano-Cinca, C., Gutiérrez-Nieto, B. (2013). Partial least square discriminant analysis for bankruptcy prediction. *Decision Support Systems*, 54(3): 1245-1255. <https://doi.org/10.1016/j.dss.2012.11.015>
- [9] Kennedy, K., Mac Namee, B., Delany, S.J. (2013). Using semi-supervised classifiers for credit scoring. *Journal of the Operational Research Society*, 64(4): 513-529. <https://doi.org/10.1057/jors.2011.30>
- [10] Xu, C., Tao, D., Xu, C. (2013). A survey on multi-view learning. *arXiv preprint arXiv:1304.5634*.

- [11] Tsai, C.F., Hsu, Y.F., Yen, D.C. (2014). A comparative study of classifier ensembles for bankruptcy prediction. *Applied Soft Computing*, 24: 977-984. <https://doi.org/10.1016/j.asoc.2014.08.047>
- [12] Kasper, K., Gentry, J., Long, L., Gentleman, R., Falcon, S., Hahne, F., Sarkar, D. (2016). Rgraphviz: Provides plotting capabilities for R graph objects. R package version 2.15.
- [13] Cao, J., Zhang, K., Luo, M., Yin, C., Lai, X. (2016). Extreme learning machine and adaptive sparse representation for image classification. *Neural Networks*, 81: 91-102. <https://doi.org/10.1016/j.neunet.2016.06.001>
- [14] Yu, Q., Miche, Y., Séverin, E., Lendasse, A. (2014). Bankruptcy prediction using extreme learning machine and financial expertise. *Neurocomputing*, 128: 296-302. <https://doi.org/10.1016/j.neucom.2013.01.063>
- [15] Zhong, H., Miao, C., Shen, Z., Feng, Y. (2014). Comparing the learning effectiveness of BP, ELM, I-ELM, and SVM for corporate credit ratings. *Neurocomputing*, 128: 285-295. <https://doi.org/10.1016/j.neucom.2013.02.054>
- [16] Mirjalili, S., Mirjalili, S.M., Hatamlou, A. (2016). Multi-verse optimizer: A nature-inspired algorithm for global optimization. *Neural Computing and Applications*, 27(2): 495-513. <https://doi.org/10.1007/s00521-015-1870-7>
- [17] Muangkote, N., Sunat, K., Chiewchanwattana, S. (2014). An improved grey wolf optimizer for training q-Gaussian Radial Basis Functional-link nets. In 2014 International Computer Science and Engineering Conference (ICSEC), pp. 209-214. <https://doi.org/10.1109/ICSEC.2014.6978196>
- [18] Nagaraj, K., Sridhar, A. (2015). A predictive system for detection of bankruptcy using machine learning techniques. *International Journal of Data Mining & Knowledge Management Process (IJDKP)*, 5(1): 29-40. <https://doi.org/10.5121/ijdkp.2015.5103>
- [19] Wagemans, F.A.J., van Koppen, C.S.A., Mol, A.P.J. (2013). The effectiveness of socially responsible investment: A review. *Journal of Integrative Environmental Sciences*, 10(3-4): 235-252. <https://doi.org/10.1080/1943815X.2013.844169>
- [20] Barboza, F., Kimura, H., Altman, E. (2017). Machine learning models and bankruptcy prediction. *Expert Systems with Applications*, 83: 405-417. <https://doi.org/10.1016/j.eswa.2017.04.006>
- [21] Kostopoulos, G., Karlos, S., Kotsiantis, S., Tampakas, V. (2017). Evaluating active learning methods for bankruptcy prediction. In *International Conference on Brain Function Assessment in Learning*, 10512: 57-66. https://doi.org/10.1007/978-3-319-67615-9_5
- [22] Situm, M. (2014). The relevance of employee-related ratios for early detection of corporate crises. *Economic and Business Review for Central and South-Eastern Europe*, 16(3): 279-314.
- [23] Kelleher, J.D., Mac Namee, B., D'Arcy, A. (2015). *Fundamentals of machine learning for predictive data analytics: Algorithms, Worked Examples, and Case Studies*.
- [24] Liu, Y., Wang, Y., Zhang, J. (2012). New machine learning algorithm: Random forest. In *International Conference on Information Computing and Applications*, 7473. 246-252. https://doi.org/10.1007/978-3-642-34062-8_32
- [25] Zięba, M., Tomczak, S.K., Tomczak, J.M. (2016). Ensemble boosted trees with synthetic features generation in application to bankruptcy prediction. *Expert Systems with Applications*, 58: 93-101. <https://doi.org/10.1016/j.eswa.2016.04.001>
- [26] Elliott, R.J., Siu, T.K., Fung, E.S. (2014). A Double HMM approach to Altman Z-scores and credit ratings. *Expert Systems with Applications*, 41(4): 1553-1560. <https://doi.org/10.1016/j.eswa.2013.08.052>
- [27] Petropoulos, A., Chatzis, S.P., Xanthopoulos, S. (2016). A novel corporate credit rating system based on Student's t hidden Markov models. *Expert Systems with Applications*, 53: 87-105. <https://doi.org/10.1016/j.eswa.2016.01.015>
- [28] Sousa, M.R., Gama, J., Brandão, E. (2016). A new dynamic modeling framework for credit risk assessment. *Expert Systems with Applications*, 45: 341-351. <https://doi.org/10.1016/j.eswa.2015.09.055>
- [29] Zhou, L. (2013). Performance of corporate bankruptcy prediction models on imbalanced dataset: The effect of sampling methods. *Knowledge-Based Systems*, 41: 16-25. <https://doi.org/10.1016/j.knosys.2012.12.007>
- [30] Wang, L., Gopal, R., Shankar, R., Pancras, J. (2015). On the brink: Predicting business failure with mobile location-based checkins. *Decision Support Systems*, 76, 3-13. <https://doi.org/10.1016/j.dss.2015.04.010>