ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2023.0441
SPECIAL ISSUE ON INNOVATIONS AND TECHNOLOGY IN FINTECH (ITFT-2023)

ENSEMBLE MODEL - BASED BANKRUPTCY PREDICTION

J. Arumugam[1], S. Raja Sekar[2] and V. Prasanna Venkatesan[3]
[1]*Department of Computer Science, Pondicherry University, India*
[2,3]*Department of Banking Technology, Pondicherry University, India*

Abstract

Bankruptcy prediction is a crucial task in the determination of an organization's economic condition, that is, whether it can meet its financial obligations or not. It is extensively researched because it includes a crucial impact on staff, customers, management, stockholders, bank disposition assessments, and profitableness. In recent years, Artificial Intelligence and Machine Learning techniques have been widely studied for bankruptcy prediction and Decision-making problems. When it comes to Machine Learning, Artificial Neural Networks perform really well and are extensively used for bankruptcy prediction since they have proven to be a good predictor in financial applications. various machine learning models are integrated into one called the ensemble technique. It lessens the bias and variance of the ml model. This improves prediction power. The proposed model operated on quantitative and qualitative datasets. This ensemble model finds key ratios and factors of Bankruptcy prediction. LR, decision tree, and Naive Bayes models were compared with the proposed model's results. Model performance was evaluated on the validation set. Accuracy was taken as a metric for the model's performance evaluation purpose. Logistic Regression has given 100% accuracy on the Qualitative Bankruptcy Data Set dataset, resulting in the Ensemble model also performing well.

Keywords:

Machine Learning, Ensemble Model, Bankruptcy Prediction, Qualitative Bankruptcy Data, Ensemble Blending

1. INTRODUCTION

For a very long time, auditors, financial institutions, bankruptcy specialists, and scholars have all been very concerned with estimating the probabilities of economic collapse. Bankers and other creditors might be affected because of financial failure. They could not return back their investment sufficiently [1]. Mostly, failure means the circumstance that causes to organization's bankruptcy following a payment default. Watson et al. depicted three types of risks for failures. The first one is the national economy-related risk, industry-oriented risks, and unique risks that are to the business itself [2]. Furthermore, the nature of failures may differ. Laitinen et al. [3] have noted that the failure of firms may fall immediately, gradually degrade, or perform unacceptably in the long run. Debtors enter into bankruptcy system shelter when business failures suddenly happened. Even through the reorganization, they can endure financial failure [4]-[6].

To manage the risk in corporate, many things are taken into consideration. one thing that poses a threat to risk management, is corporate failure prediction. Banks and other financial institutions have taken this as a serious issue. They may come with an effective alert algorithm to predict bankruptcy. In the modern era, vast amounts of present economic data about firms are gathered from the sources of Big data emerging, Information Technology, and Social media. Decision-makers do not give proper direction to attain goals by using enormous amounts of information. Pre-processing is needed to find crucial information among enormous amounts of information. thus, a prediction model in an effective way needs to be constructed without affecting desired quality output. The feature selection step is employed in Machine Learning (ML) techniques to attain crucial information. It is one of the pre-processing algorithms in general [7].

eXtreme Gradient Boosting (XGBoost) adopted feature selection, which helps to predict bankruptcy prediction. In this paper, we proposed AS-XGBoost which contains XGBoost with Attribute Selection. It is a distributed, scalable, gradient-boosting decision tree technique. An efficient way of predicting bankruptcy is through the use of this technique. The contribution of this paper is devised in two ways. The first one is an XGBoost tree incorporating important features that enhance accuracy. It is a suitable machine learning model to determine financial distress. Another contribution is to compare the proposed AS-XGBoost with established machine learning models like SVM, NB, Decision Tree, and Logistic Regression(LR) to identify ML that is sensitive to feature selection from comparisons made. This gives guidelines to banks and financial institutions to identify suitable models for bankruptcy.

In this paper, Section 2 deals with the literature review. Section 3 explains the methodology utilized. Part 4 gives the results of the proposed method along with its analysis. Section 5 concludes.

2. LITERATURE REVIEW

A crucial area of finance is the identification of bankruptcy. The possibility of a company becoming bankrupt is, in fact, a concern for numerous players for obvious factors, including executives, shareholders, or financiers. As a result, numerous research on the subject of bankruptcy prediction have been conducted. Beaver (1966) provided a univariate analysis in the late 1960s, giving financial ratios their first statistical explanation for their capacity to account for default [8]. In Altman's research, he employed multiple discriminant analysis (MDA) techniques to calculate the probability of bankruptcy for a sample of businesses. Due to its widespread use and popularity, Altman's Z-Score model is regularly used by auditors, accountants, courts, banks, and other creditors. The MDA technique assumes that there is a Gaussian distribution for the variables which was afterwards embraced by many other researchers [9]-12]. The idea that parameters have multiple normal distributions is then challenged in support of the hypothesis that the factors that explain something have distinct distributions. The probit, as well as logit models, were subsequently often applied to the prediction of bankruptcy. The second phase of the narrative started in the 1990s with the development of AI techniques, particularly those in the machine learning field like neural networks or genetic algorithms [13-15]. They achieved impressive, predicted outcomes without any statistical constraints. Indeed, using data from North American firm's data from 1985 to 2013, Barboza et al. assessed five machine learning models and contrasted their ability to forecast bankruptcy with more established statistical methods (discriminant analysis and logistic regression) [16]. From the literature survey, [17] reached this conclusion.

The use of financial ratios can improve bankruptcy prediction accuracy. Using principal component analysis (PCA) or least absolute shrinkage and selection operator (LASSO), one can identify the most relevant predicted features based on a subset of explanatory variables [18]. In order to form a list of 50 ratios, detailed information was taken from the Balance Sheet [19]. Du Jardin et al. employ feature selection [20]. To predict bankruptcy, Glen et al. used quarterly data rather than annual data [21]. Data about the bankruptcy process may include other attributes besides financial information, such as relational and textual data, market evaluations, and corporate governance information. In their study, Retznakova et al. examined average ratios for a number of years prior to bankruptcy. In addition, Pump et al. found that Bankruptcy model outcomes have an impact on economic limit periods as well. A bankruptcy prediction model is commonly used in small businesses and listed companies.

Table.1. Existing methods in the prediction of Bankruptcy

Authors	Algorithms	Characteristics	Limitations
[21]	XGBoost, SVM, and DNN	Three financial ratios are taken to classify bankruptcy from non-bankruptcy firms	No Hybrid techniques were used to categorize bankruptcy firms
[22]	Mathematical model	The model to assess financial risk	It estimates risks and then forms clusters and does not classify them.
[23]	SVM and KNN	Noisy training samples were removed with the help of SVM and KNN	Compared model with only traditional SVM
[24]	Magnetic Optimization Algorithm (MOA) and Particle Swarm Optimization (PSO)	Bio-inspired algorithms and ANN applied to bankruptcy data. It reduces training time	Standard Data Repository was not taken for the experiment
[25]	multi-criteria decision-making (MCDM) based approach	The standard ML classifier's performance was evaluated on an imbalanced dataset	Four evaluation criteria were not considered imbalanced data ratios
[26]	Hybrid switching particle swarm optimization (SPSO) and support vector machine (SVM)	Finding optimal parameter values of the Radial basis function of SVM	Comparison made against hybrid SVM and GA only
[27]	A hybrid ANN based on variables selection techniques	Multivariate discriminant analysis (MDA), Logistic Regression, and Decision Tree (DT) combined with ANN are used to distinguish bankruptcy or not.	The Moroccan firm's financial data statement was utilized. It may be an imbalanced classifier.

In order to improve bankruptcy prediction accuracy, a robust machine learning method is needed that can generalize well on financial data. As a result, many models rely on categorization methods like the Support Vector Machine (SVM). However, in financial situations, especially in bankruptcy prediction [23], it is difficult to make conclusions and choose effective solutions based on inadequate, imprecise, and noisy data. These authors offer a method for filtering out noisy training data by combining a Support Vector Machine with a K-nearest neighbor (KNN-SVM). The experimental findings demonstrate that, when applied to engineering tasks, the proposed method greatly improves generalization performance and classification accuracy by 12% over the standard SVM classifier. Potentially beneficial in computerized system applications, a composite classifier based on these variables may improve outcomes in company bankruptcy prediction.

Bankruptcy prediction is a paramount thing for financiers, investors, and also organizations. For an efficient prediction model to be constructed, Machine learning and other factors are utilized. The trained dataset containing financial ratios as features are acquired from the financial statements of various companies. Using Genetic Algorithms that determine the most weightage financial ratios helps in bankruptcy prediction. The input as financial ratios have been given to the random forest model, implemented in R. This predicts accurate results on various test cases [22].

Predicting whether or not a corporation will declare bankruptcy is a crucial step in establishing the viability of a business. For this reason, Artificial Neural Networks (ANNs) and other machine learning approaches have become increasingly popular in recent years.

For the purpose of predicting financial risk, many classifiers have been proposed. This research developed a multi-criteria decision-making (MCDM) based technique for rating insolvency prediction models, which considers numerous performance measures concurrently [24]. Seven financial unbalanced binary data sets were obtained from the UCI Machine Learning repository and used in an experiment aiming to test the suggested method. In this study, we apply four common classifiers (LR, SVM, MLP, and C4.5) in conjunction with three sets of unbalanced techniques: cost-sensitive learning, resampling (RUS and SMOTE), and hybrid methods.

Strength training is a crucial step in the process of learning a network. Strength training with ANN is more effective. Many recent works have used metaheuristic algorithms including Evolutionary Algorithms (EA) and Swarm Intelligence (SI) techniques to enhance ANN's weight training in order to better forecast insolvency [25].

This research improves upon two existing metaheuristics algorithms - the Magnetic Optimisation Algorithm (MOA) and the Particle Swarm Optimisation (PSO) - by proposing a hybrid of the two. It has been shown that hybrid algorithms can solve optimization issues more quickly and accurately. An improved prediction speed of up to 99.7 percent is demonstrated by the suggested hybrid MOA-PSO method. The next step is to test the method with more up-to-date data sets that are just as reliable. For bankruptcy forecasting, other MOA variants such the Functional Sized Population MOA (FSMOA) should be explored.

In this study, we forecast insolvency using a PSO and SVM hybrid algorithm [26]. At first, they analyzed data from the UCI Machine Learning Repository's sample bankruptcies. Then, a switching PSO technique is used to optimize the SVM's parameters. The included model has been effectively used to provide bankruptcy forecasts.

The noisy based tolerant method has not been updated to include more current datasets of equivalent quality. Instead of analyzing each and every issue, focus on the ones that matter most when making a bankruptcy determination. When compared to another hybrid algorithm, the accuracy is drastically lower. It was determined that the KNN-SVM classifier, and not the hybrid methods, would benefit most from the use of the five financial ratios chosen. The average accuracy of other algorithms is just 92.5%, which is much lower. Furthermore, a hybrid method consisting of neural networks and two optimization techniques obtained 99.728% accuracy. However, this investigation relied on historical data with only a few observations rather than more current, reliable datasets.

This paper's primary contribution is Bank collapses may be predicted with the highest accuracy possible. To lay forth the factors of becoming bankrupt. Finding the Critical Ratios for bankruptcy

3. METHODOLOGY

Machine learning techniques are often utilized for bankruptcy forecasting. Support Vector Machines, Artificial Neural Networks, Gaussian Process, Classification and Regression Trees, Logistic Regression, Decision Tree, Random Forest, Linear Discriminant Analysis, and Ensemble Learning Techniques are some of the most used methods. Additionally, several recent research agrees on the merits of combining mechanisms from various search strategies. In both operations research and AI, the development of hybrid methods is a current trend.

In the realm of managing financial risks, bankruptcy is the single most important procedure. Predicting whether or not a company will go bankrupt is an important step in figuring out the health of an organization's finances. Because it has such far-reaching consequences for a bank's personnel, customers, management, shareholders, asset-sale valuations, and bottom line, it has been the subject of many studies. The findings have implications for the lending decisions made by financial organizations and their bottom lines. A bank's ability to foresee an organization's potential is critical for preventing loan defaults. This means that banks urgently need factors related to bankruptcies and more accurate data from the present to anticipate bankruptcies.

The Ensemble models are meta-algorithms that integrate many machine learning approaches into a single predictive model to either reduce variance (bagging), bias (boosting), or enhance predictions (stacking). Using a preset qualitative bankruptcy data set and a quantitative bankruptcy data set in an ensemble model with several models, the proposed model will give important causes and key ratios relating to insolvency. After that, we'll examine the data with a number of popular categorization models including Naive Bayes, Support Vector Machines, and Logistic Regression. Finally, we use multiple measures (accuracy, precision, recall, etc.) to evaluate the models' performance on the validation datasets and rank them appropriately. As a result, banks will have a higher level of awareness and access to highly connected aspects.

3.1 DATA MODELING

With the goal of developing a predictive model that can accurately forecast the bankruptcy state of a given (unseen) bank, we will examine the various categorization models that we have examined for training on both datasets independently in this section. The following 5 models were taken into consideration: Logistic Regression, Support Vector Machine, Gaussian Naïve Bayes, Decision Tree and Random Forests.

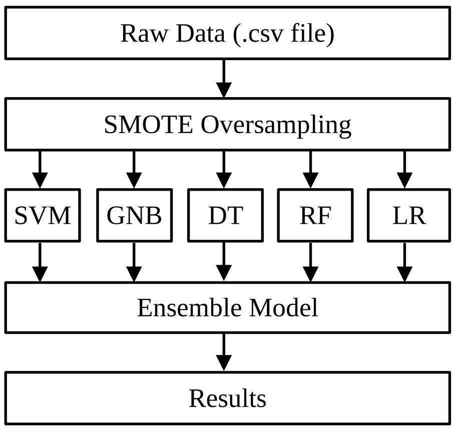

Fig.1. Pipeline for data modeling

3.1.1 Logistic Regression Classifier:

The purpose of logistic regression in statistics is to make predictions about discrete categories. The outcome or metric of interest is a binary variable. In a binary classification system, there are just two options. It's a kind of linear regression when the dependent variable is a discrete set of categories. The dependent variable is a log of the odds. The logit function is used in Logistic Regression to make predictions about the likelihood of a binary event occurring.

$$p = 1/(1+e^{-(\beta 0+\beta 1\ X1+\beta 2X2\ldots\ldots\beta nXn)}) \tag{1}$$

3.1.2 Support Vector Machine:

Although Support Vector Machines are more commonly used for classification issues, they may also be used for regression. It works well with both continuous and categorical data. To differentiate between groups, support vector machines (SVMs) create a hyperplane in a dimensional space. The ideal hyperplane for minimizing the error is generated via SVM in an iterative fashion. Finding the maximum marginal hyperplane (MMH) that most effectively separates the dataset into classes is important to SVM.

3.1.3 GNB Classifier:

A supervised learning method, the Naive Bayes classifier takes the 'naive' assumption of independence between each pair of characteristics and applies Bayes' theorem to the data. For a categorical outcome y and a set of features $x_1...x_n$,

$$p(y|x_1,\ldots x_n)= (P(y)P(x_1,\ldots,x_n)|y))/(P(x_1,\ldots,x_n)) \tag{2}$$

Based on the simplistic notion of autonomy,

$$P(x_i|y,x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_n) = P(x_i|y) \tag{3}$$

for all i, this relationship is simplified to:

$$P(y|x_1,\ldots,x_n) = \frac{P(y)\prod_{i=1}^{n} P(x_i|y)}{P(x_i,\ldots,x_n)} \tag{4}$$

The following categorization rule may be applied since P is independent of the input:

$$P(y|x_1,\ldots,x_n) \propto P(y)\prod_{i=1}^{n} P(x_i|y) \tag{5}$$

$$\hat{y} \propto \arg\max P(y)\prod_{i=1}^{n} P(x_i|y) \tag{6}$$

The Gaussian Naive Bayes classification algorithm is implemented in Gaussian Naive Bayes. The feature probabilities are assumed to follow a Gaussian distribution.

$$P(x_i|y) = \frac{1}{\sqrt{2\pi\sigma^2 y}}\exp\left(-\frac{(x_i-\mu_y)^2}{2\sigma^2 y}\right) \tag{7}$$

Maximum likelihood is used to estimate the values of the parameters σ_y and μ_y .

3.1.4 Decision Tree Classifier:

A decision tree is a type of flowchart in which each node represents an individual feature or characteristic, each branch an individual decision rule, and the individual leaf nodes the final results. The 'root' node is the node at the very top of a decision tree. It figures out how to divide things apart according to their attribute values. Recursive partitioning is a method through which the tree is divided in a self-referential fashion. This decision-making 'flowchart' will serve you well. A flowchart-like visual representation that can be used to represent complex ideas in the same way a human brain can.

In Random Forest, it works in four steps: Gather shuffled representations of data. Generate a forecast from each sample by building a decision tree. Put each expected outcome to a vote. Choose the most popular forecast as the final forecast.

3.1.5 Attribute Selection – XGBoost (Ensemble Model):

By combining many models into a single, more accurate prediction, ensemble learning aims to boost predictive model performance. The term 'Ensemble Learning' refers to a technique whereby numerous machine learning models (such as classifiers) are systematically built to address a specific issue.

Diverse Ensemble Learning methods exist, distinguished mostly by the models they employ (homogeneous vs. heterogeneous models), the methods they use to sample data (replacement vs. non-replacement, k-fold, etc.), and the decision function they employ (voting, average, meta-model, etc.). Therefore, there are several ways to categorize Ensemble Learning methods: Stacking, Blending, Voting and Blending.

The concept of blending may be traced back to the more general method of stacking. The sole distinction is that in Blending, the meta-model's training data is not generated using the k-fold cross-validation method. In order to blend, a 'one-holdout set' is used.

Predictions made using a fraction (validation) of the whole training set are 'stacked' to create the meta-model's training set. The meta-model test data is also formed by making predictions based on the test data.

The Fig.2 depicts a Blending architecture, which consists of a final classifier and three basic models (weak learners). The meta-model (yellow boxes) is formed by using predictions (blue boxes) from the training data. The predictions made from the green boxes are utilized to create the purple boxes of meta-model test data.

4. RESULTS OF ANALYSIS

4.1 DATA

In this study, we examined two datasets: the qualitative one to identify causal variables in bank failure and the quantitative one to identify correlations between those causal factors and ratios derived from bank financial statements.

4.2 QUALITATIVE BANKRUPTCY DATA

The Qualitative bankruptcy data from the UCI Machine Learning Repository has been considered for the bankruptcy prediction problem. This repository is a large collection of freely available datasets which can be used in different domains such as Machine Learning and Data Science community.

Data and questionnaires were used to separate quantitative and qualitative aspects of the analysis of the failed banks. The bankruptcy prediction data set is ideal for our study since it contains several advantageous econometric indicators as characteristics (features).

4.3 QUANTITATIVE BANKRUPTCY DATA

This data set was generated with the help of Altman Bankruptcy Model and Ratios. The bankruptcy equation of Altman bankruptcy model is given below,

$$X = 0.012\ A1 + 0.014\ A2 + 0.033\ A3 + 0.006\ A4 + 0.999\ A5$$

where A1 is ratio of working capital to total assets, A2 is the ratio of retained earnings to total assets, A3 is called as earnings before interest divided by total assets, A4 is the ratio of equity market value to total liabilities, A5 is division of sales and total assets, and X is Altman Bankrupt value.

4.4 IMPLEMENTATION

Python v3.6 is used as a working programming environment in this work. We used an Intel Core i3 Core processor with 4 GB Memory (RAM) and 1 TB of storage (disk space) to run our experiments. Our code workflow exactly mimics the data modeling pipeline shown in Fig.6. We used the libraries listed in Table.3 to run our experiments and achieve our results. Libraries mentioned in Table.3 are imported. Dataset in the form of raw data (.csv files) as pandas data frames is loaded. Features are numeric in this dataset and labels for each class are binary. These data types need to be converted into appropriate ones for data frames to store and process data efficiently. So, the numeric data type was changed to float and binary data type was changed to integer type.

Now we start the data analysis. We apply SMOTE oversampling on quantitative dataframes to get fresh dataframes of oversampled dataframes and store them in a dictionary.

Table.2. Comparative analysis with existing works

Authors	G-mean	F-measure	AUC	Accuracy
Song et al. [25]	0.9193	0.9189	0.9619	-
Ansari et al. [26]	-	-	-	99.728
Y. Lu et al. [27]	-	-	-	99.2063
Fatima et al. [28]				99.2063
Proposed Method	-	-	-	100

Instantiation of the 5 classifier models (GNB, LR, SVM DT, RF, XGB, BB) is done and stored in a dictionary. We iterate the best classifier in ensemble models - validation using Metrics.

4.5 MODEL PERFORMANCES

Accuracy and ROC of seven models for Qualitative datasets was given below: Fig.2 and Fig.5. In Qualitative data, every model depicts 100% accuracy. It is pure data for Bankruptcy prediction.

4.6 FEATURES

From Qualitative data set, we can infer from tree chart and correlation matrix that Financial Flexibility, Credibility, Competitiveness has high correlation. These factors have a high impact on bankruptcy prediction. And these factors are relatable to Altman ratios. i.e., Management Risk is a core feature of Working capital / Total assets (X1), Financial Flexibility is a core feature of Retained earnings / Total assets (X2) and Earnings before interest and taxes/ Total assets (X3), Operating Risk is a core feature of Market value of equity / Book value of total liabilities(X4) and Sales / Total assets(X5). From the relation, we can infer that X2, X3 has high impact on bankruptcy prediction. The decision tree structure is formed to make classification shown in Fig.3. The dataset correlation is shown in Fig.4. Accuracy and ROC of seven models for Quantitative datasets is given in Fig.5.

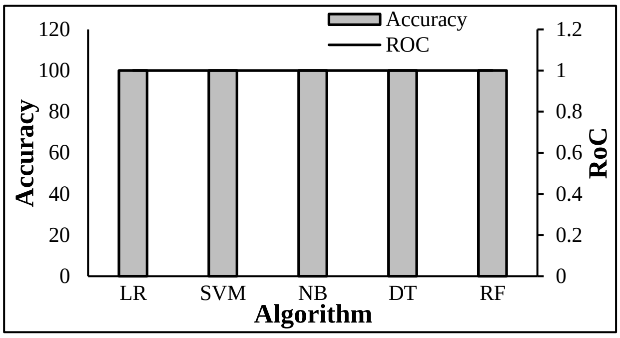

Fig.2. Qualitative Accuracy Chart

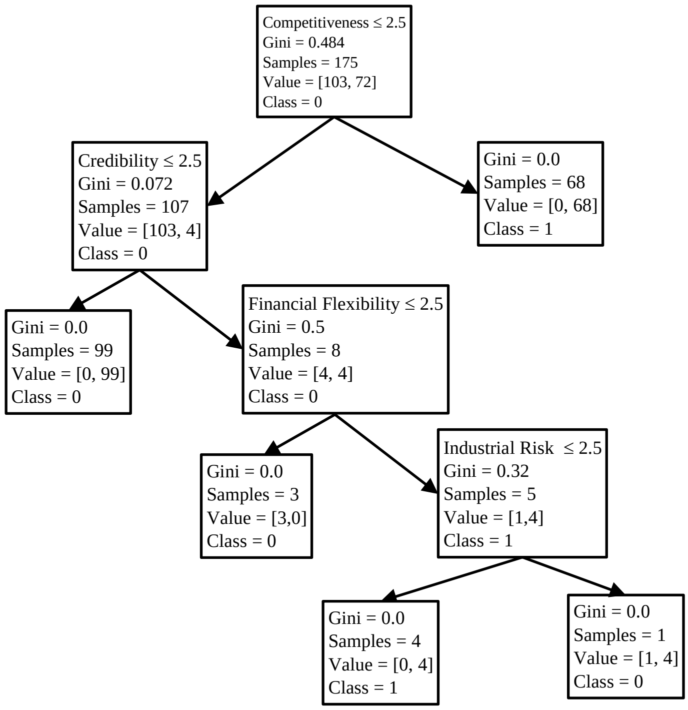

Fig.3. Tree Structure

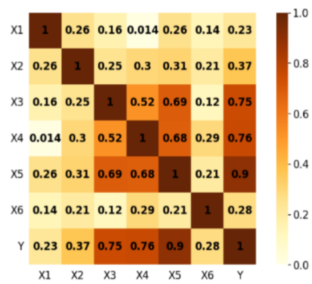

Fig.4. Qualitative Features Correlation Score

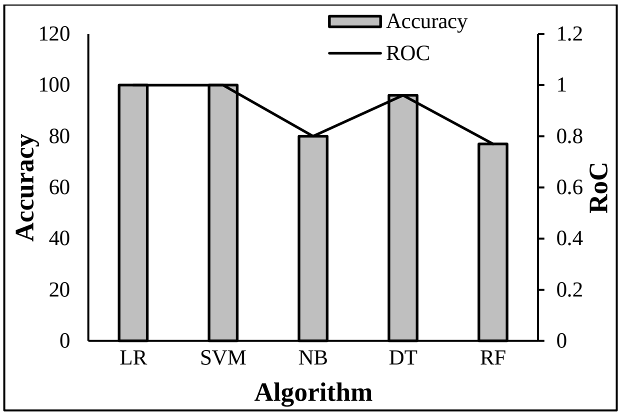

Fig.5. Quantitative Accuracy Chart

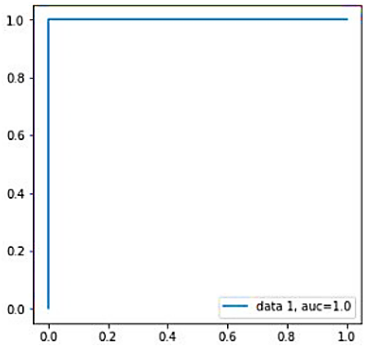

Fig.6. Blending Result of AUC curve

From the results, we can infer that Logistic Regression, Support Vector Machine, Random Forest are the best classifiers. We can infer there is drastic change in accuracy of every model due to involvement of low correlation. This model achieved accuracy of 100 due to combination of qualitative and quantify nature of dataset. For the nature of the structured dataset and ensemble technique, classification of bankruptcy from non-bankruptcy was performed more accurately shown in Fig.6.

5. CONCLUSION

Classification models, including the Gaussian Naive Bayes, Support Vector Classifier, Logistic Regression, Decision Tree, Random Forest, Extreme Gradient Boosting, and Balanced Bagging classifiers, were used to inform the formulation of the ensemble model. Using the Synthetic Minority Oversampling method, we ensured that the training sets had an equitable distribution of class labels. Predicting insolvency using indicators other than financial numbers present in firms' balance sheets requires extensive research and validation. We've done a thorough job of documenting our findings and offering our best recommendation for a bankruptcy prediction model The Deep-learning model may be used for large unstructured datasets that will be feature work.

ACKNOWLEDGMENT

The work was supported by the University Grants Commission (UGC). Ministry of Education, India. No.3674 (NET-DEC2018).

REFERENCES

[1] Karen Hopper Wruck, "Financial Distress, Reorganization, and Organizational Efficiency", *Journal of Financial Economics*, Vol. 27, No. 2, pp. 419-444, 1990.

[2] Jim Everett and John Watson, "Small Business Failure and External Risk Factors", *Small Business Economics*, Vol. 11, No. 4, pp. 371-390, 1998.

[3] Erkki K. Laitinen, Oliver Lukason and Arto Suvas, "Are Firm Failure Processes Different? Evidence from Seven Countries", *Investment Management and Financial Innovations*, Vol. 11, pp. 1-13, 2014.

[4] Regis Blazy and Nicolae Stef, "Bankruptcy Procedures in the Post-Transition Economies", *European Journal of Law and Economics*, Vol. 50, No. 7, pp. 7-64, 2020.

[5] Nicolae Stef, "Bankruptcy and the Difficulty of Firing", *International Review of Law and Economics*, Vol. 54, pp. 85-94, 2018.

[6] N. Stef, "Correction to: Institutions and Corporate Financial Distress in Central and Eastern Europe", *European Journal of Law and Economics*, Vol. 53, pp. 145-146, 2022.

[7] D. Liang and H.T. Wu, "The Effect of Feature Selection on Financial Distress Prediction", *Knowledge-Based* Systems, Vol. 73, No. 1, pp. 298297-298301, 2015.

[8] W.H. Beaver, "Financial Ratios as Predictors of Failure", *Journal of Accounting Research*, Vol. 4, pp. 71-111, 1966.

[9] Edward I. Altman, "Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", *The Journal of Finance*, Vol. 23, No. 4, pp. 589-609, 1968.

[10] Edward B. Deakin, "A Discriminant Analysis of Predictors of Business Failure", *Journal of Accounting Research*, Vol. 10, No. 1, pp. 167-179, 1972.

[11] R.O. Edmister, "An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction", *The Journal of Financial and Quantitative Analysis*, Vol. 7, No. 2, pp. 1477-1493, 1972.

[12] Edward I. Altman, Robert G. Haldeman and P. Narayanan, "ZETATM Analysis a New Model to Identify Bankruptcy Risk of Corporations", *Journal of Banking and Finance*, Vol. 1, No. 1, pp. 29-54, 1977.

[13] James A. Ohlson, "Financial Ratios and the Probabilistic Prediction of Bankruptcy", *Journal of Accounting Research*, Vol. 18, No. 1, pp. 109-131, 1980.

[14] M.E. Zmijewski, "Methodological Issues Related to the Estimation of Financial Distress Prediction Models", *Journal of Accounting Research*, Vol. 22, pp. 59-82, 1984.

[15] Clive Lennox, "Identifying Failing Companies: A Re-Evaluation of the Logit, Probit and Mda Approaches", Available at https://ssrn.com/abstract=67888, Accessed at 1998.

[16] Flavio Barboza and Edward Altman, "Machine Learning Models and Bankruptcy Prediction", *Expert Systems with Applications*, Vol. 83, pp. 1-15, 2017.

[17] M. Aziz and H. Dar, "Predicting Corporate Bankruptcy: Where We Stand?", *Corporate Governance*, Vol. 6, pp. 18-33, 2006.

[18] Jianqing Fan and Runze Li, "Variable Selection via Nonconcave Penalized Likelihood and its Oracle Properties", *Journal of the American Statistical Association*, Vol. 96, pp. 456-467, 2001.

[19] Philippe Du Jardin, "Bankruptcy Prediction using Terminal Failure Processes", *European Journal of Operational Research*, Vol. 242, pp. 286-303, 2015.

[20] M. Papik and L. Papíkova, "Impacts of Crisis on SME Bankruptcy Prediction Models' Performance", *Expert Systems with Applications*, Vol. 214, pp. 119072-119083, 2023.

[21] S. Shetty and M. Musa, "Bankruptcy Prediction using Machine Learning Techniques", *Journal of Risk and Financial Management*, Vol. 15, No. 1, pp. 35-48, 2022.

[22] R. Joshi, R. Ramesh and S. Tahsildar, "A Bankruptcy Prediction Model using Random Forest", *Proceedings of International Conference on Intelligent Computing and Control Systems*, pp. 1-6, 2018.

[23] Z. Gao, M. Cui and L.M. Po, "Enterprise Bankruptcy Prediction using Noisy-Tolerant Support Vector Machine", *Proceedings of International Seminar on Future Information Technology and Management Engineering*, pp. 153-156, 2008.

[24] A. Ansari, I.S. Ahmad, A.A. Bakar and M.R. Yaakub, "A Hybrid Metaheuristic Method in Training Artificial Neural Network for Bankruptcy Prediction", *IEEE Access*, Vol. 8, pp. 176640-176650, 2020.

[25] Y. Song and Y. Peng, "A MCDM-Based Evaluation Approach for Imbalanced Classification Methods in Financial Risk Prediction", *IEEE Access*, Vol. 7, pp. 84897-84906, 2019.

[26] Y. Lu, J. Zhu, N. Zhang and Q. Shao, "A Hybrid Switching PSO Algorithm and Support Vector Machines for Bankruptcy Prediction", *Proceedings of International Conference on Mechatronics and Control*, pp. 1329-1333, 2014.

[27] Fatima Zahra Azayite, Said Achchab, Fatima Zahra Azayite and Said Achchab, "A Hybrid Neural Network Model based on Improved PSO and SA for Bankruptcy Prediction", *International Journal of Computer Science*, Vol 16, No. 2, pp. 1-6, 2019.