2020 IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS) | 978-1-6654-0478-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/iSES50453.2020.00017

An Ensemble Classifier Model to Predict Credit Scoring – Comparative Analysis

Safiya Parvin A[1], Saleena B.[2]
[1,2] School of Computer Science and Engineering
VIT, Chennai Campus

Abstract

Credit scoring is a way of analyzing statistical data used in financial organizations and banks to acquire a person's creditworthiness. The bestowers generally manipulate it to decide to widen or retract credit. The score plays a significant role in determining the creditworthiness of a person and if he/she can be sanctioned a loan or not. Machine learning techniques help us to predict the credit score more accurately using classification algorithms. Few base and ensemble classification algorithms were used in this research to perform a comparative analysis. The ensemble method incorporates several base classification algorithms like Decision trees, Logistic Regression, Nearest neighbor, Support Vector Machine, etc. to achieve better results. The objective of this paper is to predict the credit score based on different classifier models and evaluate the performance of each model based on the metrics. A comparative analysis is done to identify the best classifier to predict the credit score. The evaluation metrics used for evaluating the model are Recall, Precision, F-measure, and Accuracy. Error measures like MAE and RMSE of the model were also used to evaluate the model. This helps us to improve the decision in identifying the more accurate classifier model. The dataset used for this analysis is the Australian credit dataset from the UCI Machine learning repository. Experimental results prove that the Random Forest and Extratree classifier model produces better accuracy in ensemble classifiers and the SVM model furnishes better accuracy in the base classifier.

Keywords:

Credit Score, Ensemble classifier, Decision trees, Logistic Regression, Nearest neighbor, Support Vector Machine, Bagging, Boosting, Voting, Random Forest, Extra Trees Classifier, Ada Boost, Bagged Decision Tree.

1. Introduction

A credit score is a number based on investigating a person's credit files, to represent the trustworthiness of an individual. The score computes the creditworthiness of borrowers based on their existing information. It is often used by banking, mobile manufacturers, insurance companies, landlords, government departments, and financial institutions such as online lenders. The statistical efforts behind credit scoring and credit benchmark stated by Hand et al [8].

The way of credit score creation and finding the key factors in the establishment of a scoring template by various statistical techniques and performance evaluation criteria was discussed by Abdou et al [1]. The techniques for scoring can be enlarged to add distinct areas and can aid decision makers, in banking, to predict the behaviour of clients.

Machine learning algorithms are mainly used for predicting the behavior of existing real world data. The framework along with credit-related data can be used to forecast the future. Initially, the credit applications are analyzed with base (single) and ensemble classifier techniques which were described in [2] (Adnan Dželihodžić , 2016). Ensemble classifier model enhances the prediction and it can be created based on nature of any application. Ensemble strategy converts an imbalanced dataset into balanced dataset discussed in Xu et al[21]. The main focus of this research is to perform a comparative analysis to predict the best classifier model. Each model is evaluated based on the metrics called precision, recall, F-measure (F1-Score), and Accuracy.

Section 2 discusses the related work carried out for predicting the credit score using different classifier models (base and ensemble). Section 3 discusses the methodology used for performing the comparative analysis. Section 4 summarizes the different algorithms used for comparative analysis and tabulates the score of the confusion matrix. Section 5 deals with the experimental results and discussion and analyses an efficient algorithm to predict the credit score based on the evaluation metrics and error measures.

2. Related Work

This section discusses the different research contributions for predicting the credit score and other similar applications. This related work was divided into two segments based on the research done using base classification techniques and ensemble classification techniques.

978-1-6654-0478-5/20/$31.00 ©2020 IEEE
DOI 10.1109/iSES50453.2020.00017

He et al [9] used logistic regression with a stacking ensemble method to construct a complex framework for credit score based data. Another ensemble model called Gradient Descent based Logistic regression for binary classification was developed by Cheon J H et al [5]. Support Vector Machines proposed by Cortes and Vapnik[6].This model is used to find a hyperplane to separate two different types of training samples with a small error rate. The SVM model was applied by Li et al [13] to credit prediction for evaluating consumer loans. The neural network algorithm was used by Jensen et al [10] to predict the loan approval process. The classification results achieved resemble the same as traditional credit scoring results. Cheng-An Li [4] used KNN to provide a novel algorithm for predicting credit score inquiry. Siami et al [16] tested the locally linear model tree algorithm to compute the Decision tree and achieved better results. Okesola et al [15] used the Naïve Bayesian approach to predict loan requests as good or bad.

Ala'raj et al [3] followed a consensus approach and ranking to distinguish classifiers. The proposed combination method refines prediction performance against all base classifiers. The ensemble approach gave better results than the base classifiers by applying several classifiers and seven real-world datasets on retail credit scoring in Lessmann S et al[12]. The random and bagging framework evaluated the highest accuracies in sentiment mining suggested by Whitehead M et al[20]. To test the performance of credit score prediction, multi classifiers stacking framework was worked out by Guo et al [7]. A stack-based approach by differentiating the performance of traditional time series framework and deep learning methods in malaria case predictions recommended in Wang M et al [19]. A hybrid neural network works well for financial distress prediction than ensemble by Tsai et al [18]. The complex fusion schemes using heterogeneous ensemble strategies analyzed by Tabik et al[17]. Soft computing methods applied were applied by Lahsasna et al[11] to calculate the credit score model. To predict the loan default in credit based organization, decision rules with heuristic methods used by Odeh et al[14].

It was evident from the literature review that few algorithms like Random forest ,Voting ensemble, Bagged Decision Tree, and Extra tree classifier framework were not considered in calculating the credit score. The proposed research has considered these algorithms together with few existing algorithms to perform a comparative analysis to identify a more accurate algorithm in predicting the credit score.

3. The methodology used for the Comparative Analysis

This section discusses the methodology as shown in Figure 3.1. The dataset used for the comparative analysis was the Australian credit dataset from the UCI machine repository with 14 features and 690 instances. The methodology can be described in a 4-step process.

1. Load the respective dataset and preprocess the dataset.
2. Apply various classification algorithms to the dataset.
3. Evaluate the derived models with performance metrics and error measures.
4. Find out the best model with high accuracy and low error rate in base and ensemble algorithms.

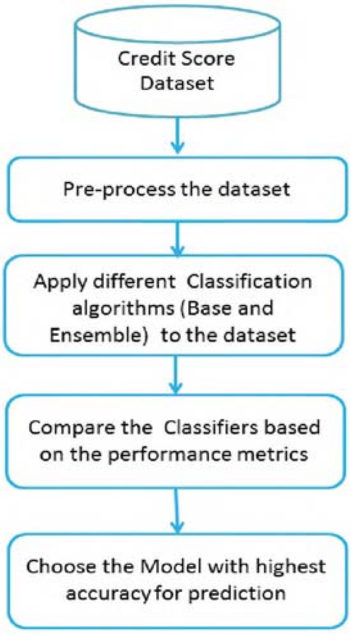

Fig 3.1: Methodology for comparative Analysis

The dataset has some missing values and filled by arithmetical mean for continuous variables and mode for categorical variables. The dataset has a combination of categorical, continuous variables and is converted to numeric attributes. Since the dataset is a multivariate dataset standard transformations are done for each feature. Each feature is subtracted by its mean and divided by the standard deviation to avoid numerical stabilities. In this way, each feature is standardized into a unit scale. Principal Component Analysis (PCA) dimensionality reduction is used to convert high dimension to low dimension. The dataset with 14-dimensional space converted into 2-dimensional space. The explained variance ratio attained for the first principal component as 66.06%

and the second component as 33.93%. The two components contain 99.99% of the information. Then this data is fed to the base and ensemble classifiers simultaneously.

4. Summary of Algorithms analyzed to predict Credit Scoring

This section summarizes the different classification algorithms used to perform the comparative analysis.

Base Classifiers

Logistic Regression (LR),Decision Tree (DT),Support Vector Machine (SVM), K-Nearest Neighbour Algorithm (k-NN),Multilayer Perceptron (MLP) and Naïve Bayes classifier (NB) are used.

Ensemble Classifiers

Random Forest (RF),Bagged Decision Tree, ExtraTreesClassifier(ETC), Adaboosting, Gradient Boosting (GB) and Voting Ensemble (LR+DT+SVM) are used.

5. Results and discussion

The proposed system was tested on the Australian dataset from the UCI machine repository with 690 instances. The original dataset is partitioned into training and test sets in the ratio of 80:20 respectively. The performance and error measures used for evaluating the classifier model are discussed below and their values obtained from the model are shown in Table 1.

Table 1: Performance Metrics and Error Measure Comparison of classifiers

Classifier	Performance Metrics			Error Measures		
	Accuracy	Precision	Recall	F1-Score	MAE	RMSE
Logistic Regression	0.8333	0.85	0.78	0.82	0.1667	0.4082
Decision Tree	0.8261	0.83	0.80	0.81	0.1594	0.3993
SVM	**0.8768**	0.89	0.85	0.87	**0.1232**	**0.351**
KNN	0.7174	0.75	0.60	0.67	0.2826	0.5316
MLP	0.7246	0.75	0.62	0.68	0.2754	0.5247
Naïve Bayes	0.8043	0.95	0.62	0.75	0.1957	0.4423
Random Forest	**0.8841**	**0.87**	**0.80**	**0.84**	**0.116**	**0.3405**
Bagged Decision Tree	0.8638	0.86	0.88	0.87	**0.1232**	**0.351**
Extra Trees Classifier	**0.8841**	0.89	0.86	0.88	0.1304	0.3612
Ada Boosting	0.8551	0.86	0.83	0.84	0.1449	0.3807
Gradient Boosting	0.8696	0.87	0.85	0.86	0.1304	0.3612
Voting Ensemble (Logistic+Decision tree+ SVM)	0.7754	0.87	0.62	0.72	0.2246	0.474

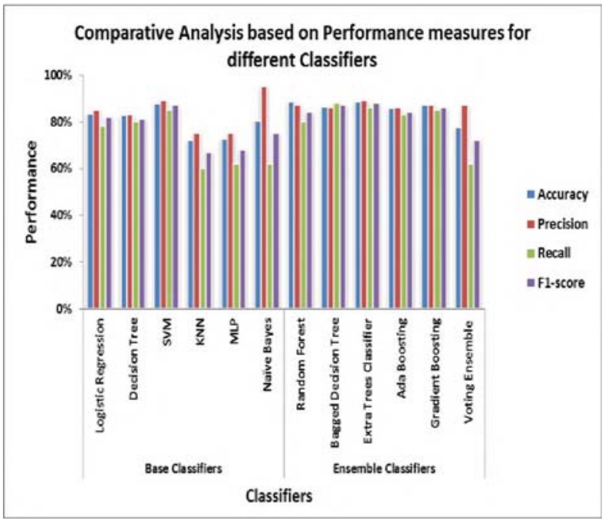

Figure 5.1: Performance Metric Comparison of All Classifiers

Since the dataset has balanced data, the main focus is on the Accuracy measure. The highest accuracy is achieved by Random forest, Extratree classifier followed by SVM. Figure 5.1 depicts the comparison of these classifiers. This enhances the data to achieve better results. Figure 5.2 denotes that the error rate is minimal in Random forest Model and Bagged tree classifier.

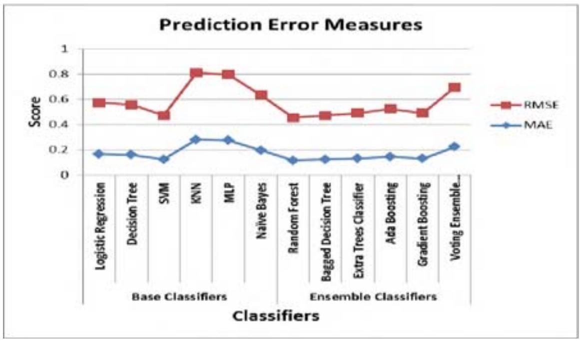

Fig 5.2: Comparison of Error Measure of Classifiers

6. Conclusion

To predict the trustworthiness of a person, the statistical analysis of the data is required. Twelve different machine learning models are employed to perform this task. In this evaluation, performance metrics and error measures are derived for both types of classifiers. SVM one of the base classifier achieves high accuracy than other base classifiers. Even though SVM a good framework for linearly separable data, it faces several risks in choosing soft margin or kernel function with nonlinear data. Random forest, Extra Tree classifier, and Bagged Decision tree are the ensemble classifiers that achieve high accuracy and less error rate than other classifiers. This happens by choosing arbitrary data and construction of trees and combining the outcome. The future directions of this work will be to estimate the computational time taken for each classifier model and to adapt optimization techniques to better improvement on the resultant model.

References

[1] Abdou, HAH and Pointon, J. (2011). Credit scoring, statistical techniques and evaluation criteria: A review of the literature, Intelligent Systems in Accounting, Finance & Management. 18 (2-3), pp. 59-88.

[2] Adnan Dželihodžić , Dženana Đonko ,(2016), "Comparison of Ensemble Classification Techniques and Single Classifiers Performance for Customer Credit Assessment" Modeling of Artificial Intelligence, Vol.(11), Is. 3 pp.140-150.

[3] Ala'raj M and Abbod M.F, Jul. 2016,"Classifiers consensus system approach for credit scoring" ,*Knowl.-Based Syst.*, vol. 104, pp. 89-105.

[4] Cheng-An Li (2013), "Credit Scoring Analysis Using B-Cell Algorithm and K-Nearest Neighbor Classifiers" ,Part II, LNCS 7929, pp. 191–199, Springer-Verlag Berlin Heidelberg.

[5] Cheon J H, Kim D, Kim Y and Song Y, (2018),"Ensemble Method for Privacy-Preserving Logistic Regression Based on Homomorphic Encryption", IEEE Access, VOLUME 6,pp. 46938-46948.

[6]Cortes C and Vapnik V, (Sep. 1995),"Support-vector networks" , *Mach. Learn.*,vol. 20, no. 3, pp. 273_297.

[7] Guo S,He H, Huang X, (2019) "A Multi-Stage Self-Adaptive Classifier Ensemble Model With Application in Credit Scoring" , IEEE Access,Volume 7,pp.78549-78559.

[8] Hand D. J and Henley W. E., (Sep. 1997),"Statistical classification methods in consumer credit scoring: A review," *J. Roy. Stat. Soc., A*, vol. 160, pp. 523-541.

[9] He H, Zhang W., and Zhang S., (May 2018),"A novel ensemble method for credit scoring: Adaption of different imbalance ratios", *Expert Syst. Appl.*, vol. 98,pp. 105-117.

[10] Jensen L., (1992), "Using Neural Networks for Credit Scoring," *Computer Journal of Managerial Finance*, vol. 18, no. 15, pp. 26-29.

[11] Lahsasna A, Ainon R N, Wah T Y,(April 2010), "Credit Scoring Models Using Soft Computing Methods: A Survey", The International Arab Journal of Information Technology.

[12] Lessmann S, Baesens B, Thomas L. C., and Seow H.-V,(Nov. 2015), "Benchmarking state-of-the-art classi_cation algorithms for credit scoring: An update of research," *Eur. J. Oper. Res.*, vol. 247, no. 1, pp. 124-136.

[13] Li T., Shiue W., and Huang H.,(2006), "The Evaluation of Consumer Loans Using Support Vector Machines" ,*Computer Journal of Expert Systems with Applications*, vol. 30, no. 4, pp. 772-782.

[14] "Odeh O, Koduru P, Das S, Allen M., Stephen M., (2011), "A multi-objective approach for the prediction of loan defaults", ACM Press,New York, NY, USA.

[15] Okesola, Okokpujie, Kennedy O. and Adewale,,Samuel N. John(2017)," An improved Bank Credit Scoring Model A Naïve Bayesian Approach",International Conference on Computational Science and Computational Intelligence,IEEE,pp.228-233.

[16] Siami M, Gholamian M.R, and Basiri J,(Oct. 2014.),``An application of locally linear model tree algorithm with a combination of feature selection in credit scoring," *Int. J. Syst. Sci.*, vol. 45, pp. 2213-2222.

[17] Tabik S, Alvear-Sandoval R.F., Ruiz M.M., Sancho-Gómez J.L., Figueiras-Vidal A.R., Herrera F. , (January 2020),"A Tutorial On Ensembles And Deep Learning Fusion With Mnist As Guiding Thread: A Complex Heterogeneous Fusion Scheme Reaching 10 Digits error",arxiv.org.

[18] Tsai C.-F. and Hung C., 2014), Modeling credit scoring using neural network ensembles," *Kybernetes*, vol. 43, no. 7, pp. 1114-1123.

[19] Wang M , Wang H, Wang J, Liu H, Lu R, Duan T, Gong X, Feng S, Liu YS, Zhuang Cui, Changping Li, Jun Ma, (2019)," A novel model for malaria prediction based on ensemble algorithms",PLOS One.

[20] Whitehead M,Yaeger L,(2010),"Sentiment Mining Using Ensemble Classification Models", Innovations and Advances in Computer Sciences and Engineering, Springer Science+Business Media B.V, pp.509-519.

[21] Dayu Xu , Xuyao Zhang , Junguo Hu,1and Jiahao Chen,(2020),"A Novel Ensemble Credit Scoring Model Based on Extreme Learning Machine and Generalized Fuzzy Soft Sets", Hindawi ,Mathematical Problems in Engineering.