International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07, Special Issue | June 2020 www.irjet.net p-ISSN: 2395-0072
International Conference on Recent Trends in Science & Technology-2020 (ICRTST - 2020)
Organised by: ATME College of Engineering, Mysuru, INDIA

An Experimental Study on Comparison of Credit Scoring Models

Anil Kumar C J[1], B.K. Raghavendra[2]

[1]Research Scholar, Department of Computer Science and Engineering, BGS Institute of Technology, B G Nagara, Mandya, Karnataka, India.
[2]Professor & Head, Department of Computer Science and Engineering, BGS Institute of Technology, B G Nagara, Mandya, Karnataka, India.

---***---

***Abstract* -** Credit Scoring is widely used in banking and risk assessment systems are more important for decision making. Creditworthiness of the credit applicants are assessed using Credit scoring technique. The main purpose of credit scoring is to categorize applicants into two groups, i.e., good or bad customer based on their payment actions. Credit Scoring is a process to measure risk. Credit scoring model building is concerned with the developed of experimental models to aid financial decision making for banking industries. Applicant's statistical and historical data or artificial intelligence (AI) techniques are used for evaluating the risk involved in connection with an applicant. We, in this paper implemented four credit scoring models: Artificial neural networks (ANN), Random Forest (RF), support vector machines (SVM) and Logistic Regression (LR). From the results it is evident that RF gives better result for both credit datasets when compared with LR, RF and ANN. Based on this experiment, Random Forest is considered as the best classification model.

Keywords: SVM, Logistic regression, ANN, Random Forest, Credit Scoring.

I. INTRODUCTION

Credit Scoring is universally used and important for providing decisions and profitability in banking industry. For deciding on whether to or not to grant credit to new applicants Credit Scoring models are used. To support credit approval decisions various credit scoring models are studied.

Credit scoring models are used in credit industry for credit approval decision. For the credit scoring problem, numerous models based on credit scoring has been proposed and developed. Credit scoring models are classified into two classes: Traditional mining models and Modern mining models. Traditional mining models use conventional statistical methods where as modern mining models uses artificial intelligence methods. Random forest, Support vector machines, artificial neural networks, K- nearest neighbor, Logistic regression and genetic programming are among the popular models. Logistic regression is widely used and is appropriate for variety of distribution functions including credit scoring problem.

Four models using Random Forest, Logistic Regression (LR) , Artificial Neural Networks (ANN), and Support Vector Machines (SVM) are built for the study of classification. The aim of building a simplified form of the problem is to determine whether this credit scoring model provides a workable model. If this simple model gives a promising result, then it will be worth exploring these techniques on multiclass problems.

II. RELATED WORK

Xingzhi Zhang, Yan Yang, Zhurong Zhou [4] used an optimized random forest algorithm based on feature selection and grid search to build a novel credit scoring model. In order to get the higher prediction accuracy, irrelevant and duplicate features are reduced. To select the optimal feature, the information entropy is used as heuristic in NCSM.

Seyedeh Maryam Anaei, Mohsen Moradi [5] suggest that, if two steps of feature selection and regulation of SVM parameters are performed simultaneously, the performance will be improved greatly compared with cases where the two steps are performed separately. They also suggest that the accuracy of model in terms data classification is greater than other models and it reduces classification errors. Genetic algorithm has a better performance with respect to precision and data distribution.

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07, Special Issue | June 2020 www.irjet.net p-ISSN: 2395-0072
International Conference on Recent Trends in Science & Technology-2020 (ICRTST - 2020)

Organised by: ATME College of Engineering, Mysuru, INDIA

Ali Ghasemy Armaki, Feiz Fallah, Mahmoud Alborzi, Amir Mohammadzadeh [6] suggests a hybrid meta-learner model that is a combination of traditional and ensemble methods. A meta-learner model comprises of 3 base classifiers followed by meta-learner followed by clustering method.

Sofia Ayouche, Rajae Aboulaich, Rachid Ellaia [7], proposed ANN approach for credit scoring. Partnership contracts of credit applicants are assessed using this model. Sixteen predictor variables are used in multilayer perceptron neural network for credit scoring. Compared to conventional models, MP neural network gives better results. Conventional techniques like discriminant analysis and linear regression are used for comparison. Compared to other techniques MP model gives highest CCR in both training and testing subset.

Sang Ha Van, Nam Nguyen Ha, Hien Ngugen Thi Bao [8] discussed about credit scoring model built using feature selection approach. The authors have constructed the model using GBM, Filter and Wrapper approaches. This model is used to evaluate the applicant's credit score. Optimal subsets of relevant features are selected using backward sequential scheme. This subset is evaluated by GBM classifier.

From the literature survey, it is observed that the competition between financial organizations has arrived at a conflicting stage and hence credit scoring is very important. Most of the companies are looking for better strategies. In this regard, different credit scoring techniques are used in the credit assessment process.

III. CREDIT SCORING MODELS

3.1 Neural Network

A neural network (NN) is an information processing mathematical model. ANN consists of one input layer, one output layer and one or more hidden layers. Every layer has set of nodes.

In back propagation a set of training samples are processed iteratively for learning. The predicted class of each sample is compared with actual class label. The weights are modified for each training sample where mean squared error between the actual class label and network's prediction is minimized. Different activation functions like RBF, logistic and sigmoid are applied in hidden layer. Mean squared error (MSE) equation is as given below.

$$\mathbf{MSE} = \frac{1}{2}\sum_{i=1}^{n}(\tilde{y}_i - y_i)^2 \quad (1)$$

Where, $\tilde{y}_i$ is a predicted class and n is the total number of pattern pairs.

Every input pattern passes the network with a given initial weights and threshold, and produces output pattern. MSE is used to find the error between the predicted class and target pattern. Then based on the error rate, weights and thresholds are adjusted. For every pattern pair which is assigned for training network, the process of adjusting weights and threshold is repeated till the train error drops to the range of acceptance level.

3.2 Logistic Regression

Logistic regression is a special form of the linear regression. Relationship between one or more independent variables and dependent variable is measured using logistic regression. It is used in the fields of social sciences, medical and machine learning. The model consists of n predictors and one dichotomous output (response) variable. The response variable has only two possible outcomes: 1 and 0.

The equation is as shown below,

$$\log\frac{p}{(1-p)} = \boldsymbol{\beta_1 * x_1 + \beta_2 * x_2 + \beta_3 * x_3 + \ldots + \beta_k * x_k} \quad (2)$$

Where, **p** represents specific customers probability.

3.3 Support vector machines

Support vector machine(SVM) is a classification technique and was proposed by Vapnik. SVM for two-class classification is briefly described. For a training set of labeled pair(x_i, y_i) instances, y_i belongs to {+1,-1} and x_i belongs to R^n. For two class classification, separating two classifications can be done through finding a hyper plane using the following equation:

$$wx + b = 0 \quad (3)$$

An optimal separating hyper plane with maximum margin can be obtained using dual lagrangian function.

$$\mathbf{L_D(\alpha)} = sgn\left(\sum_{i=1}^{n} y_i \alpha_i^* \langle x_i, x \rangle + b^*\right) \quad (4)$$

While applied to classification, a small set of lagrange multipliers α_i have values greater than zero. These set of values are close to the optimal hyperplane. These are called support vectors.

We can extend the above concept to the non-separable case, by including slack variables as shown in the following equation.

$$y_i [(w \cdot x) + b] - 1 + \xi_i \geq 0, i = 1, 2, \ldots, m \quad (5)$$

Where, ξ_i is a positive slack variable. This can be used to find the less number of training errors.

3.4 Random Forest

Random Forest is an ensemble classification method. Each classifier in the ensemble is built using decision tree classifier. It is a collection of classifiers which forms a forest. Individual decision trees are constructed by using attributes randomly selected at each node. Each individual tree votes during classification. Output is based on votes, where the most voted class is considered.

Bagging or Bootstrap aggregation techniques are applied for random forest during training. Consider a training set $\mathbf{x_i} \in \mathbf{X}$ with class output $\mathbf{yi} \in \mathbf{Y}$. Bagging fits decision trees to the samples of the training dataset based on repeated selection of a random sample. To classify a new instance, multiple trees are created and based on attributes; each tree assigns a particular class. Based on voting, the algorithm chooses the classification, that is, to which class the new instance belongs to.

IV. MODEL EVALUATION

We have carried out experiments on four different classifiers and the results are shown. For the experiment purpose two credit scoring datasets Australian and German datasets are used to compare the performance of classifiers implemented. The credit datasets used are freely available in UCI, a repository for machine learning. The classifiers considered are Support Vector Machine, Logistic Regression, Random Forest and Artificial Neural Network. All four classifiers used are implemented using Python.

In order to evaluate the predictive power of the developed models, prediction accuracy rate and f-measure are used. These evaluation methods are calculated using equation (6) and (7) respectively.

$$\textbf{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (6)$$

F-Measure = $\frac{2TP}{2TP+FP+FN}$ (7)

where,

TP: Number of True positives

FP: Number of False positives

TN: Number of True Negatives

FN: Number of False Negatives

V. EXPERIMENTAL RESULTS

A survey on credit scoring using various machine learning algorithms is done. As an experiment, 4 different models are implemented using Python. Two publicly available credit datasets – Australian and German data are used. Classification accuracy and f-measure are used to evaluate the performance of the classification models LR, ANN, RF and SVM. Table 1 shows the specification for the datasets. The results of the evaluation are given in Table 2 and Table 3 respectively. From the results it is evident that RF gives better result for both credit datasets when compared with LR, RF and ANN. Based on this experiment, Random Forest is considered as the best classification model. Figure 1 and 2 gives classification accuracy and f-measure details after evaluation process.

Table 1. Specification for the Dataset

Sl. No	Dataset	No of instances	Total no. of attributes	No of classes
1	Australian	690	15	2
2	German	1000	21	2

Table 2. Classification accuracy and F-Measure for Australian credit dataset

	Accuracy %	F-Measure %
LR	87.39	85.25
ANN	68.50	66.65
SVM	85.79	83.46
RF	95.94	93

Table 3. Classification accuracy and F-Measure for German credit dataset

	Accuracy	F-Measure
LR	78	84.24

ANN	79.1	76.19
SVM	82.2	81.07
RF	87.1	85

Figure 1: Classification Accuracy and f-measure for German Data Set

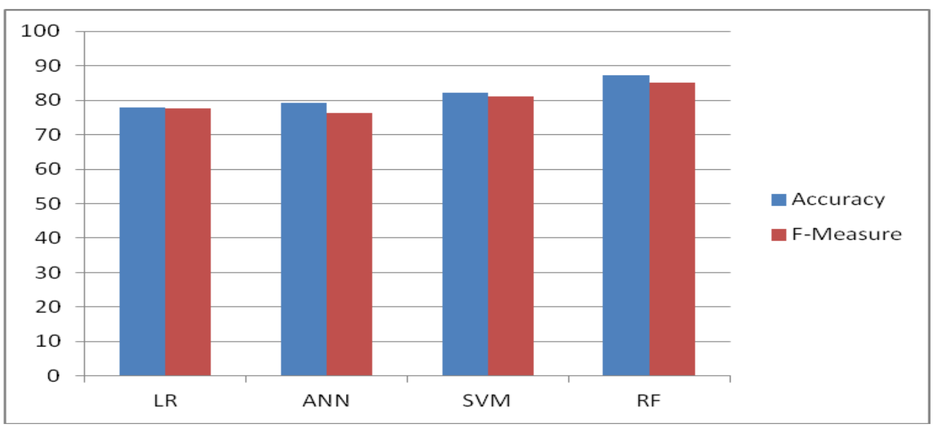

Figure 2: Classification Accuracy and f-measure for Australian Data Set

VI. CONCLUSION

Credit applications are processed efficiently through automatic credit approval systems. More and more financial institutes are seeking better strategies for decision making in financial domain. Logistic regression is most widely utilized technique in the credit scoring domain. LR gives good results when the data is linear in nature and at the same time for both linear and non-linear data AI techniques promises to produce good results. In this research work an attempt has been made to evaluate LR, ANN, RF and SVM for the credit scoring problem. From the experimental result it is evident that RF performs well and has higher accuracy than other three approaches in the credit risk analysis.

References

[1] Thomas, L. C. A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. International Journal of Forecasting, 2000, 16 (2), 149–172.

[2] Baesens, B., Van Gestel, T., Viaene, S., Stepanova, M., Suykens, J., & Vanthienen, J. Benchmarking state-of-the-art classification algorithms for credit scoring. Journal of the Operational Research Society, 2003, 54(6), 627–635.

[3] Henley, W. E. Statistical aspects of credit scoring. Dissertation,The Open University, Milton Keynes, UK.1995.

[4] Xingzhi Zhang, Yan Yang, Zhurong Zhou, "A Novel Credit Scoring Model based on Optimized Random Forest", IEEE transactions, 2018.

[5] Yajiao Tang, 2 Junkai Ji, Shangce Gao, Hongwei Dai, Yang Yu, Yuki Todo, "A Pruning Neural Network Model in Credit Classification Analysis", Computational Intelligence and Neuroscience, 2018.

[6] Ali Ghasemy Armaki, Mir Feiz Fallah, Mahmoud Alborzi, Amir Mohammadzadeh, Engineering, "A Hybrid Meta-Learner Technique for Credit Scoring of Banks' Customers", Engineering, Technology & Applied Science Research, Vol. 7, No. 5 , 2073-2082, 2017.

[7] Sofia Ayouche, Rajae Aboulaich, Rachid Ellaia, "Partnership Credit Scoring Classification Problem: A Neural Network Approach", International Journal of Applied Engineering Research, 2017.

[8] Sang Ha Van, Nam Nguyen Ha, Hien Ngugen Thi Bao, "A Hybrid Feature Selection Method for Credit Scoring" , EAI Endorsed transactions on context-aware systems and applications, 2017.

[9] Mohammad Khandabaei, Mohmood Albori " The use of genetic algorithm, clustering and feature selection techniques in construction of decision tree model s for credit scoring", International journal of managing informational technology, Vol. 5, No. 4, 2013.

[10] Defu Zhang, Hongyi Huang, Qingshan Chen, Yi Jiang, "A Comparison Study of Credit scoring Models" , Third International Conference on Natural Computation (ICNC 2007) 0-7695-2875-9/07.

[11] Weimin Chen, Guocheng Xiang, Youjin Liu, Kexi Wang, "Credit risk Evaluation by hybrid data mining technique", Systems Engineering Procedia 3, pp. 194 – 200, 2012.

[12] Xiao-Lin Li, Yu Zhong " An overview of personal credit scoring: Techniques and Future work", Journal of Intellenge Science, Vol. 2,pp. 181-189, 2012.

[13] Apapan Pumsirirat, Liu Yan, " Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 1, 2018.

[14] Bing Zhu, Wenchuan Yang, Huaxuan Wang, Yuan Yuan, "A Hybrid Deep Learning Model for Consumer Credit Scoring", IEEE transactions, 2018.

[15] Yufei Xia, Chuanzhe Liu, Bowen Da, Fangming Xie, "A novel heterogeneous ensemble credit scoring model based on bstacking approach", Expert systems with applications, pp.182-190,2018.

[16] M. Z. Abedin, C. Guotai, F. Moula, A. S. M. S. Azad, and M. S. U. Khan, "Topological applications of multilayer perceptrons and support vector machines in financial decision support systems" Int. J. Financ. Econ., vol. 24, no. 1, pp. 474–507, 2019.

[17] S. Guo, H. He, and X. Huang, "A Multi-Stage Self-Adaptive Classifier Ensemble Model With Application in Credit Scoring" IEEE Access, 2019.

[18] Y. Z. Xu, J. L. Zhang, Y. Hua, and L. Y. Wang, "Dynamic credit risk evaluation method for e-commerce sellers based on a hybrid artificial intelligence model" Sustain., vol. 11, no. 19, 2019.