IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-7, July 2018 E-ISSN: 2347-2693

Modeling Customer's Credit Worthiness using Enhanced Ensemble Model

M. Singh [1*], G. K. Dixit [2]

[1,2]Computer Science & Engineering, Oriental Institute of Science & Technology, RGPV, Bhopal, India

**Corresponding Author: mukeshtit007@gmail.com, Tel.: +91-8818965776*

Available online at: www.ijcseonline.org

Accepted: 24/Jul/2018, Published: 31/Jul/2018

Abstract— The financial organizations such as banks assess credit worthiness of borrowers before providing new loans. Certain assessments are done on the basis of 'probability of default' for the potential borrower which is based on credit scoring of customer. Banks are having customer's portfolios that are likely to go through current crisis without much difficulty, but at the same time it is difficult to determine the risks arising due to various factors. Most of the features in customer databases have little predictive effect on the credit worthiness of the customer. So, determining features that affects the credit worthiness is another important task before knowledge discovery. The predictive models, based on Ensemble Methods (EM): a ML paradigm under Data Mining (DM), have the capability to determine the relevant features and customer's credit worthiness in efficient manner. In this paper, a good number of prediction models proposed in the literature are surveyed. To deal with major challenges like data incompleteness, noise and its vastness while building predictive models, this work proposes a predictive framework (classifier) for detection of Credit Worthiness. It employs a Supervised Attribute (Feature) selection method to select the worth of subset, filters and an Enhanced Ensemble Method based on Boosting for precise convergence. The proposed Enhanced Ensemble based classifier is implemented in Weka 3.8.2. Later in this work, to know the performance of proposed model, evaluation and comparison against several model is presented. The proposed framework selects minimum number of attributes and is good in terms of False Positive, Recall, True Positive and Prediction Rate than original base classifiers. Also, the model is good for applying on large datasets due the benefit of ensembles.

Keywords— Artificial Intelligence, Boosting, Classification, Credit Risk, Credit Worthiness, Ensemble Methods, Machine Learning, Predictive Models, Weka

I. INTRODUCTION

One of the popular definitions of Credit risk tells that a borrower may not settle up a loan and the provider may lose the interest of the loan and sometimes even principal amount of it. The situation occurs when borrowers' use their future savings to pay their current debts; and it's always a risk associated with the borrowers. The credit worthiness of customer may turn down over time due to various factors like increased competition, rising inflation, volatility in asset value and weaker exchange rates. So, financial Organizations like banks are facing challenges while estimating probability of default and sanctioning the loans to current customers. Banks assess credit worthiness of borrowers before providing new loans. Certain assessments are done on the basis of 'probability of default' for the potential borrowers. The assessments are usually based on subjective analysis of credit experts of financial institutes.

The banks saves large number of features/attributes of the customers and are normally taken into consideration, while advancing loans, but most of these features have little predictive effect on the credit worthiness or otherwise of the customer. Another issues like data incompleteness, vastness and presence of noise in data put hurdles in the path of analysis and prediction. That is why even the robust and automated system of predicting credit worthiness of customers can't provide accurate decision on calculating credit risk or determining worthiness.

Artificial Intelligence (AI) is an emerging field for building intelligent models to "learn" from data and be able to do prediction for various analytical purposes. Machine Learning algorithms which are heart of AI have the capability to deal such problems in efficient way. Most of the organizations are relying on ML models to analyze the current and potential risks associated with the domain. Modeling the Credit worthiness is an active research field and facilitates in dealing Credit risks on loans.

Ensemble learning is a ML paradigm where several learners of same or different types are trained to resolve the same problem [1]. Then results from individual learners are pooled to get better accuracy. Learners that are composed in an ensemble are usually called base learners [2]. The area of research in Ensemble Methods (EMs) is one of the recent areas of research in Data Mining and ML. DM techniques are actually based on mathematics and ML algorithms [3]. Many recent researches indicate that EMs lead to a major improvement in performance of basic classifiers.

So, in this work a survey on good number of prediction models proposed in the literature is presented in Section II. To deal with major challenges like data incompleteness, noise and its vastness while building predictive models, this work proposes a predictive framework (classifier) for detection of Credit Worthiness. The proposed model is presented in Section III. The proposed Enhanced Ensemble based classifier is implemented in Weka 3.8.2. . Simulation setup, dataset description is discussed in Section IV. The proposed framework is efficient in terms of False Positive, Recall, True Positive and Prediction Rate. The performance evaluation and comparison against several base classifiers is presented in Section V. Also, the model is good for applying on large datasets due the benefit of ensembles. The paper is concluded with remarks and future directions in Section VI.

II. LITERATURE REVIEW

Over the last decade, sophisticated systems has been developed by number of the world's major banks in view to model the credit risk that arises from their business lines [4]. Most methods for determining credit worthiness require historical data to build and validate the models. Existing applications of single AI technique can be further improved by ensemble methods [5]. The usual Machine Learning techniques attempt to build a model from the training data on one proposition, while Ensemble Methods tries to build a set of rules or propositions to use [6]. A number of prediction models are proposed in the literature that uses EMs are surveyed and depicted in Table 1.

Table 1. Survey based on Ensemble Approaches

Authors	Work / Methods undertaken	Achievements	Dataset used	Tool used
[7]	Applied 15 different ML methods. Proposed a model based on Linear Regression.	Accuracy between 76 – 80% in all except in 2 methods. Features selected = 5	Taiwan Bank Credit Card dataset [8]	Scikit - Learn [8] and MATLAB
[2]	Performance comparison of ensemble of classifiers for credit scoring and bankruptcy prediction is done.	Random Subspace (RS) ensemble method, with Neural Net classifier, performed better than other EMs.	Australian credit, German credit and Japanese credit dataset.	Different suitable toolkits
[9]	03 EMs namely Bagging, AdaBoost and Random Forest combined with three ML algorithm, to asses credit risk. Feature selection is applied to derive the important attributes.	assessment on performance of the ensemble classifiers	German credit dataset [10]	Weka [11, 12]
[6]	3 popular EMs, i.e., Boosting, Bagging, and Stacking, based on 4 basic learners: Logistic Regression, Decision Tree (DT), ANN and SVM, are compared.	Performance evaluation shown that ensemble methods improve base learners.	German credit and Australian credit dataset [10]	Weka toolkit [11, 12]
[13]	Proposed a Model based on Bagging EM with REP Tree, for Credit risk prediction.	Accuracy of more than 81% with minimum number of features selected, i.e. 3.	Taiwan Bank Credit Card dataset [8]	Weka toolkit [11, 12]

[14]	Experimented with batch & incremental classifiers. Pre-processing and feature extraction using ID3 are applied to reduce the dataset and to get highest gain ratio.	Author's concluded that sample and partition size of training and testing has an effect on accuracy of prediction.	Malaysian Bank Dataset.	Weka [11, 12]

III. Proposed Work

A. Proposed Framework
The proposed framework for detection of Credit Worthiness is depicted in Figure 1. The proposed framework having 5 phases:

B. Model Phases
In phase I, Pre-Processing of data using suitable filters and methods is done.

In phase II, we applied 'CfsSubsetEval' method to evaluate worth of a subset of features by taking contribution of the prediction ability, into account. The method reduces model building time by selecting only features that are correlated with final class. The algorithm/method selected for this work deduces the minimum features i.e. 3 only. K-fold cross validation Test Split method is applied for good convergence of model. Here K=10 is taken.

In phase III and IV, Model building process using training data and performance evaluation is carried out respectively. The best model for classifying the data is chosen in this phase by iteratively applying various models (phase III) and evaluating the model (phase IV) with Test data on the basis of performance metrics. The phase I and II are also reapplied to get better prediction accuracy. The process of applying algorithm and evaluation is applied as many times the new classifier is tested.

Finally, Boosting EM with Decision Stump as base classifier is chosen and applied. Decision Stump is briefly discussed here. The name of Boosting method is real 'AdaBoost'. A decision tree with one split point is called a decision stump, this is because there is little tree to speak of.

In last phase (V), the model built is used as Prediction model for prediction purpose, over new Data. The predictive model is deployed and used as a tool on new unclassified data. The model is tested with new data to get responses. The predicted responses are finally evaluated and compared with other models. The overall correct response out of total test instances provided gives the accuracy of model. The evaluation is done on basis of few more metrics to check the models features or attributes.

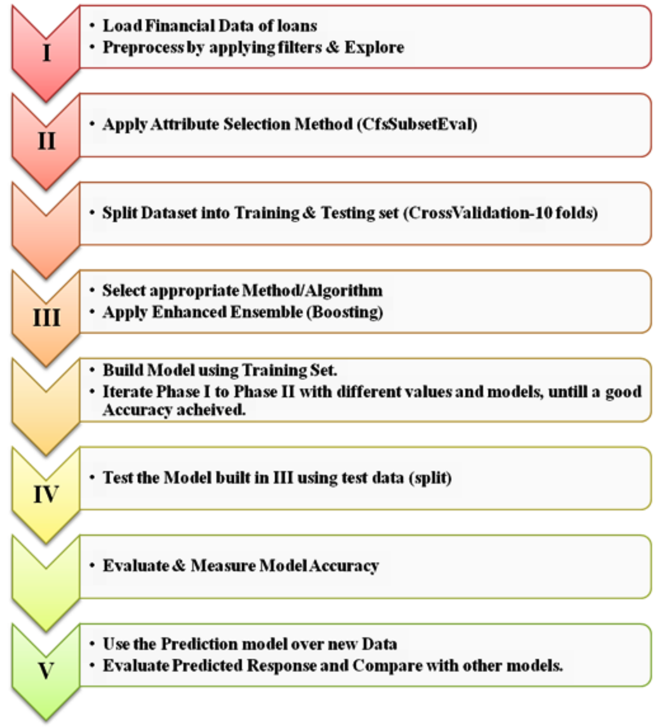

Figure 1. Proposed Predictive Framework

IV. Experimental Setup and Methodology

A. Tool Used
Weka 3.8.2 tool [11, 12], is used for simulating the proposed work, It is a famous DM tool that includes tools for all Knowledge Discovery steps. It also allows development and implementation of new ML schemes.

B. About Credit Dataset
For this work the dataset of credit card clients, in Taiwan from April 2005 to September 2005, is used. The dataset and description is available at UCI – ML repository [15]. The number of reliable (1 = yes) and number of defaulters (0 = no) out of total instances in dataset are shown in Table 2.

Table 2. Details of Dataset

Dataset	No. of Features	Total Instances	No. of Instances (Yes)	No. of Instances (No)
Credit Card	23	30000	23364 (77.88%)	6636 (22.12%)

C. Simulation
An experiment configuration scenario is shown in Figure 2. To Classify, choose Classify tab and click choose button to select AdaBoost ensemble in Meta folder. Within Boosting select the Decision Stump as base classifier.

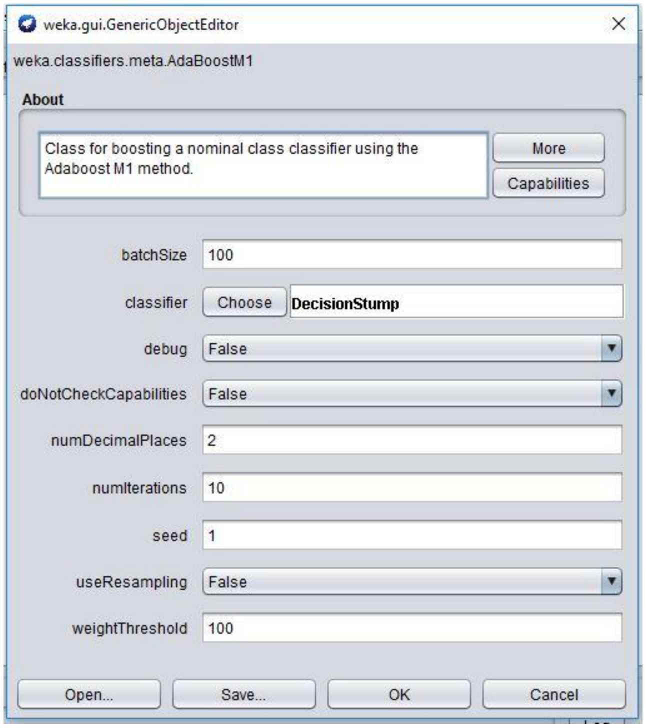

Figure 2. Configuring Classifier for Experiment

V. SIMULATION RESULTS AND ANALYSIS

The simulation is run several times with different configurations i.e. for base classifiers, and then with proposed work. To evaluate the performance using confusion matrix, the classifier is evaluated on the basis of correct predictions and several other parameters. The results are shown in Table 3. Some prominent results are depicted graphically in Figure 3 through Figure 6. The results of following algorithms / models are compared with that of proposed model.

1. Naïve Bayes
2. Random Forest
3. Logistic Regression
4. Neural Network
5. Decision Stump

Table 2. Comparison of Results

Algorithm	Accuracy (%)	True Positive	False Positive	Recall	Prediction Rate
Naive Bayes	81.95	22423	4474	0.960	0.912
Random Forest	81.93	22425	4480	0.960	0.912
Logistic Regression	81.96	22421	4469	0.960	0.912
Neural Network	81.93	22428	4485	0.960	0.912
Decision Stump	80.6	22400	4900	0.961	0.926
Proposed Work	82	22541	4792	0.965	0.924

A. *Accuracy*

The accuracy of model is measured generally on basis of correctly classified instances. Figure 3 shows the comparison of accuracy of popular models and proposed model.

$$\text{Accuracy} = \frac{TP + TN}{No.of\ Instances} * 100$$

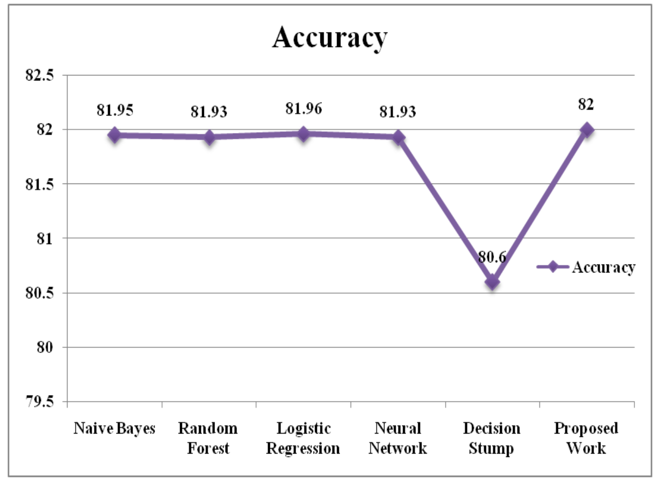

Figure 3. Comparison of Accuracy

B. *True Positive*

It represents number of correctly identified instances from among the total number of correct instances. Figure 4 shows the comparison of True Positives of 5 popular model and proposed models. The TP of proposed work is remarkably high.

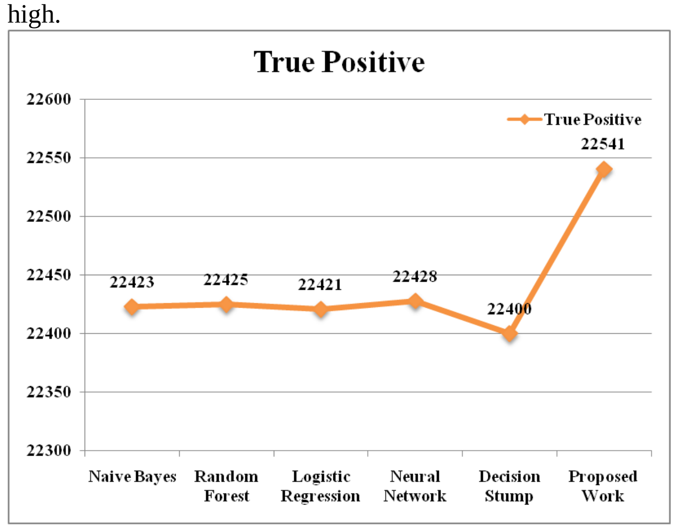

Figure 4. Comparison of True Positive

C. *Recall*

It is defined as number of positive cases that are correctly identified. Some paper refers it Sensitivity and it should be higher. Figure 5 shows the comparison of Recall of 5 popular models and proposed work.

$$\text{Recall} = \frac{TP}{TP + FN}$$

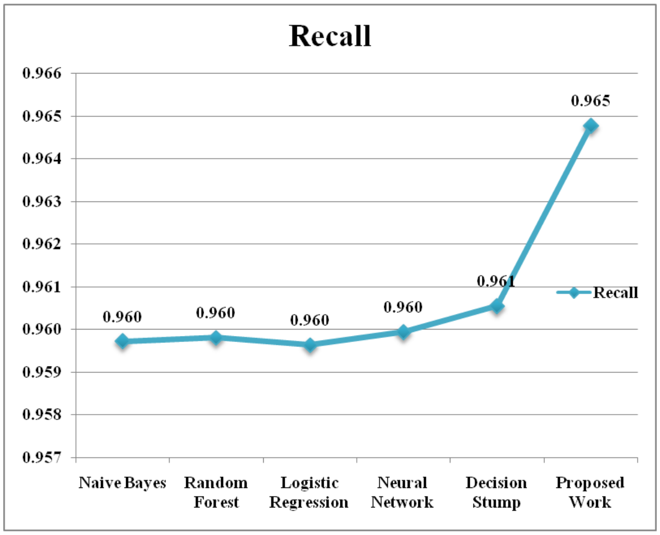

Figure 5. Comparison of Recall

D. False Positives

The FP means those negative instances that are predicted or classified as positive. Many researchers are constantly working to reduce the number significantly. Figure 6 shows the comparison of False Positive (FP) of 5 popular models and proposed work. In this work a significant reduction in FP is seen.

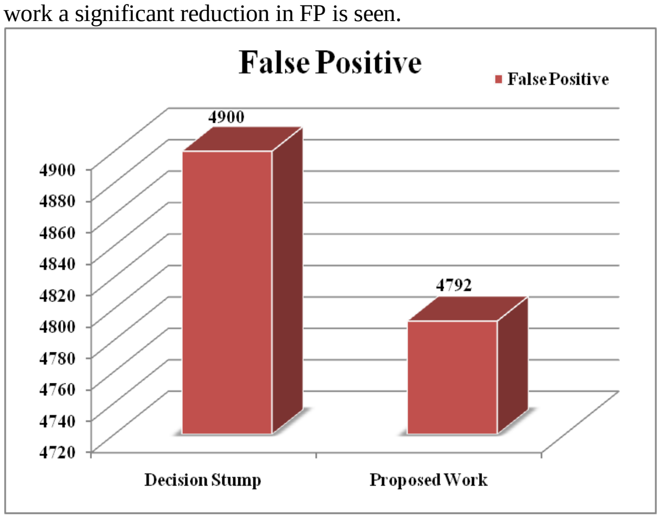

Figure 6. Comparison of False Positive

VI. CONCLUSION AND FUTURE WORK

This work explores many ML techniques and Ensembles to build predictive model using real credit card dataset. The proposed model predicts customer's credit worthiness using enhanced ensemble method. In this work, the proposed model for predicting credit worthiness is compared with 5 base classifiers on the basis of parameters of prime importance in ML. The proposed framework employs suitable filters and methods for feature selection as number of features selected generally affects training and model building time. The number of features selected in the proposed technique is 3, which is the lowest. Furthermore, the model has better generalization ability in terms of True-Positive and False-Positives. The other parameters like recall are better than other base classifiers.

The credit worthiness identification and analysis based on Ensemble Learning paradigm of ML has outstanding advantages and are suitable for studying and applying with credit datasets. The inferences drawn from this work can certainly help the researchers working in the field of financial domain. The work leaves lot of scope for researchers as they can experiment on various feature selection methods, ML models and EMs for getting better results. To build up a ML system for automation of identifying financial and credit risks over cloud can be seen as future work. The development of cloud based predictive model to determine credit worthiness of customers will be future work.

REFERENCES

[1] Polikar, R. (2006). Ensemble based systems in decision making. IEEE Circuits and Systems Magazine, 6(3), 21–45.
[2] Loris Nanni , Alessandra Lumini, "An experimental comparison of ensemble of classifiers for bankruptcy prediction and credit scoring", Elsevier, Expert Systems with Applications 36 (2009) 3028–3033.
[3] Stuti Nathaniel, Anand Motwani, Arpit Saxena, "Cloud based Predictive Model for Detection of 'Chronic Kidney Disease' Risk", International Journal of Computer Sciences and Engineering, Vol.6, Issue.4, pp.185-188, 2018.
[4] Bhekisipho Twala, "Multiple classifier application to credit risk assessment", Expert Systems with Applications 37 (2010) 3326–3336.
[5] Gang Wang, Jinxing Hao, Jian Mab, Hongbing Jiang, "A comparative assessment of ensemble learning for credit scoring", Expert Systems with Applications 38 (2011) 223–230.
[6] Zhou, Z. H. (2009). Ensemble. In L. Liu & T. Özsu (Eds.), Encyclopedia of database systems. Berlin: Springer.
[7] Regina Esi Turkson, Edward Yeallakuor Baagyere, Gideon Evans Wenya, "A Machine Learning Approach for Predicting Bank Credit Worthiness", ISBN: 978-1-4673-9187-0, IEEE 2016
[8] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, vol. 12, pp. 2825–2830, 2011.
[9] C.R.Durga devi, Dr.R.Manicka chezian, "A Relative Evaluation of the Performance of Ensemble Learning in Credit Scoring", IEEE International Conference on Advances in Computer Applications (ICACA), 978-1-5090-3770-4/16, 2016
[10] Zhang, "Z., Research of credit risk of commercial bank's personal loan based on CHAID decision tree", Artificial Intelligence, Management Science and Electronic Commerce (AIMSEC), 2011.
[11] http://www.cs.waikato.ac.nz/~ml/weka/index.html
[12] Weka, University of Waikato, Hamilton, New Zealand.
[13] Anand Motwani, Goldi Bajaj, Sushila Mohane, "Predictive Modelling for Credit Risk Detection using Ensemble Method", International Journal of Computer Sciences and Engineering, Vol.6, Issue.6, pp.863-867, 2018.
[14] Ling Kock Sheng, and Teh Ying Wah, "A comparative study of data mining techniques in predicting consumers' credit card risk in banks", African Journal of Business Management Vol. 5 (20), pp. 8307-8312, 16 September, 2011
[15] http://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients

