IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Implementation And Analysis Of Loan Credibility Prediction System Using Data Mining Techniques

Simrandeep Kaur , M.Tech CSE
St. Soldier Institute of Engg. & Technology, Near NIT, Jalandhar
Gurpreet Singh
Professor, Department of CSE
St. Soldier Institute of Engg. & Technology, Near NIT, Jalandhar
Neha
Assistant Professor, Department of Computer Sc. & Engg.
St. Soldier Institute of Engg. & Technology, Near NIT, Jalandhar

Abstract: As we know that now-a-days there is a rapid growth in banking sector, resulting lots of people are applying for bank loans. Finding out the applicant to whom the loan will be approved is a difficult process. Data mining techniques are becoming very popular nowadays because of the wide availability of huge quantity of data and the need for transforming such data into knowledge. Techniques of data mining are implemented in various domains such as retail industry, telecommunication industry, biological data analysis, etc. In this research, we proposed a model which predicts loan approval/rejection of an applicant using data mining techniques. This can be done by training the model with the data of the previous records of the people applied for loan.

Keyword: KDD, DM, AdvAdaboost

1. Introduction:

Data mining algorithms are used to study the loan-approved data and exact patterns, which would help in predicting the reasonable defaulters, thereby helping the banks for making better choices in the future. Loan Prediction is extremely useful for employee of banks and for the applicant also. The main aim of this model is to provide a speedy, immediate and simple approach to pick the deserving applicants. The model proposed in

has been built using data from banks to predict the status of loans. This model uses three classification algorithms namely DecisionStump, AdvAdaboost and UserClassifier. The model was implemented using Weka. In a decision tree model was utilized as a classifier and for feature selection genetic algorithm is utilized. The model was tried utilizing Weka.

2. Proposed Methodology (AdvAdaBoost)

AdvAdaBoost is short for Adaptive Boosting and is a very popular boosting technique that combines multiple “weak classifiers” into a single “strong classifier”.

1. Initialise the dataset and assign equal weight to each of the data point.
2. Provide this as input to the model and identify the wrongly classified data points.
3. Increase the weight of the wrongly classified data points.
4. if (got required results)

   Goto step 5

   else

   Goto step 2
5. End

3. Implementation and Result

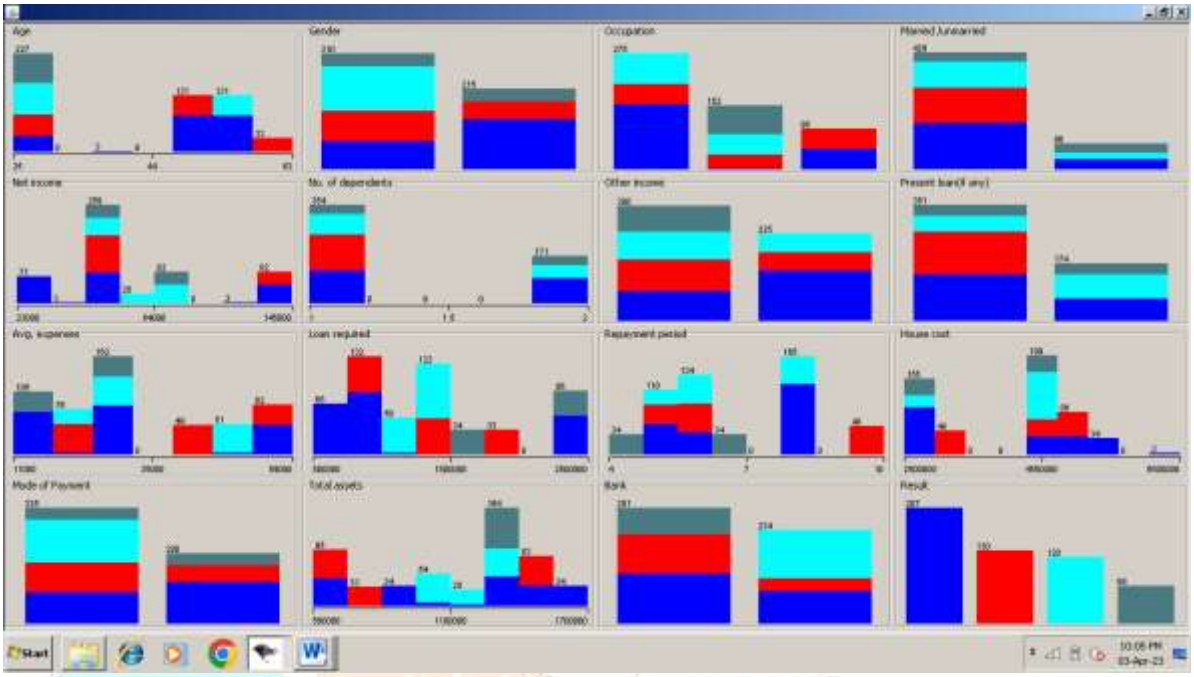

Fig. 1.1 Presents graphical representation of all the attributes

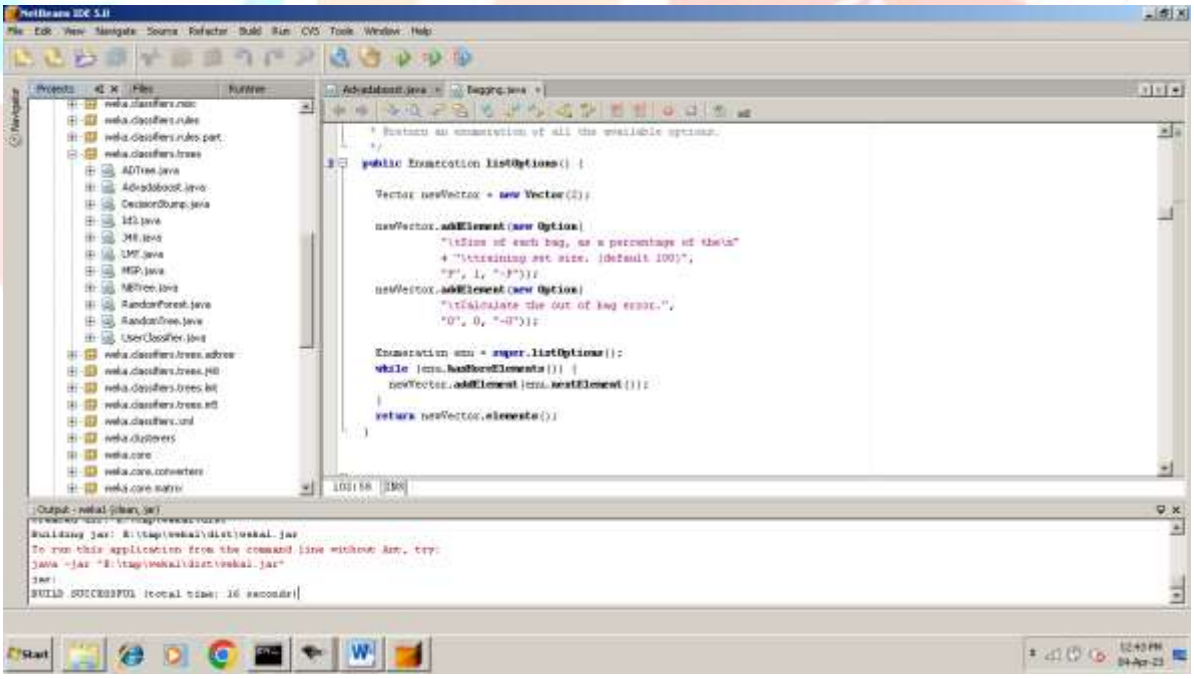

Fig. 1.2 Shows the codding of proposed Algorithm in Net Beans

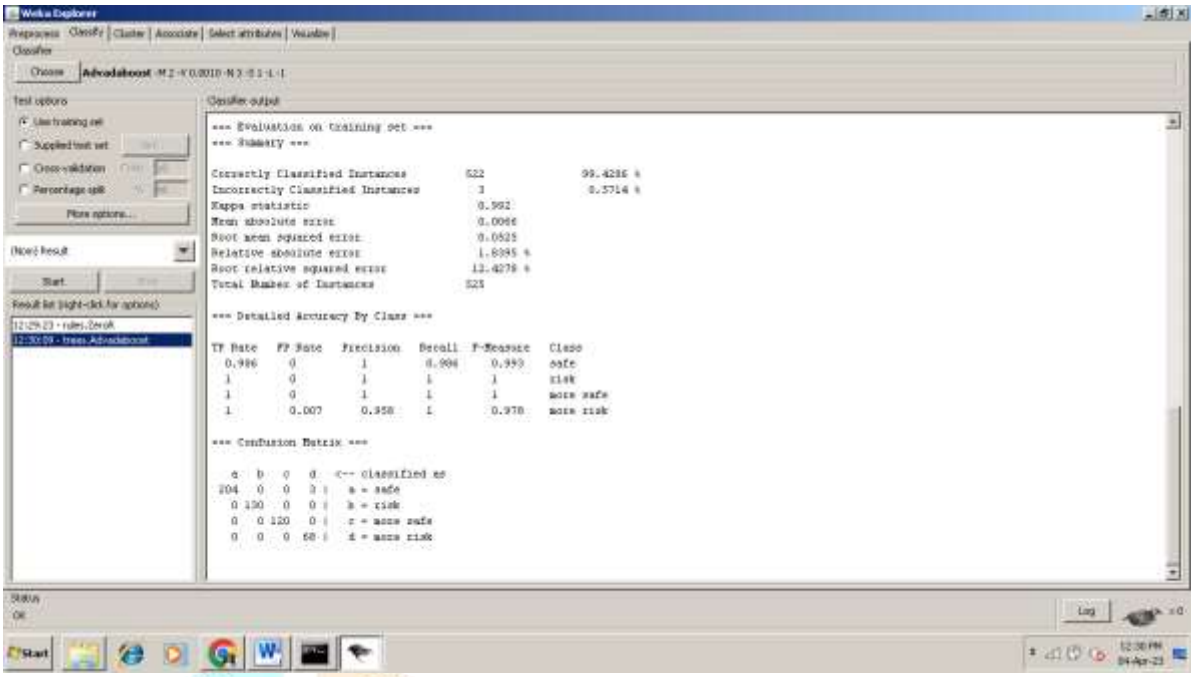

Fig. 1.3: Model Evaluation on training data using Advadaboost algorithm

Table 1.1: Comparison of Accuracy(%) of classifiers DecisionStump, AdvAdaboost and UserClassifier

Algorithms	DecisionStump	AdvAdaBoost	UserClassifier
Correctly Classified Instance	52.76	99.42	39.42
Incorrect Classified Instance	47.23	0.57	60.57
Error Rate	88.91	12.42	99.99
Total Time Taken to build the model (Sec)	0.05	0.05	20.07

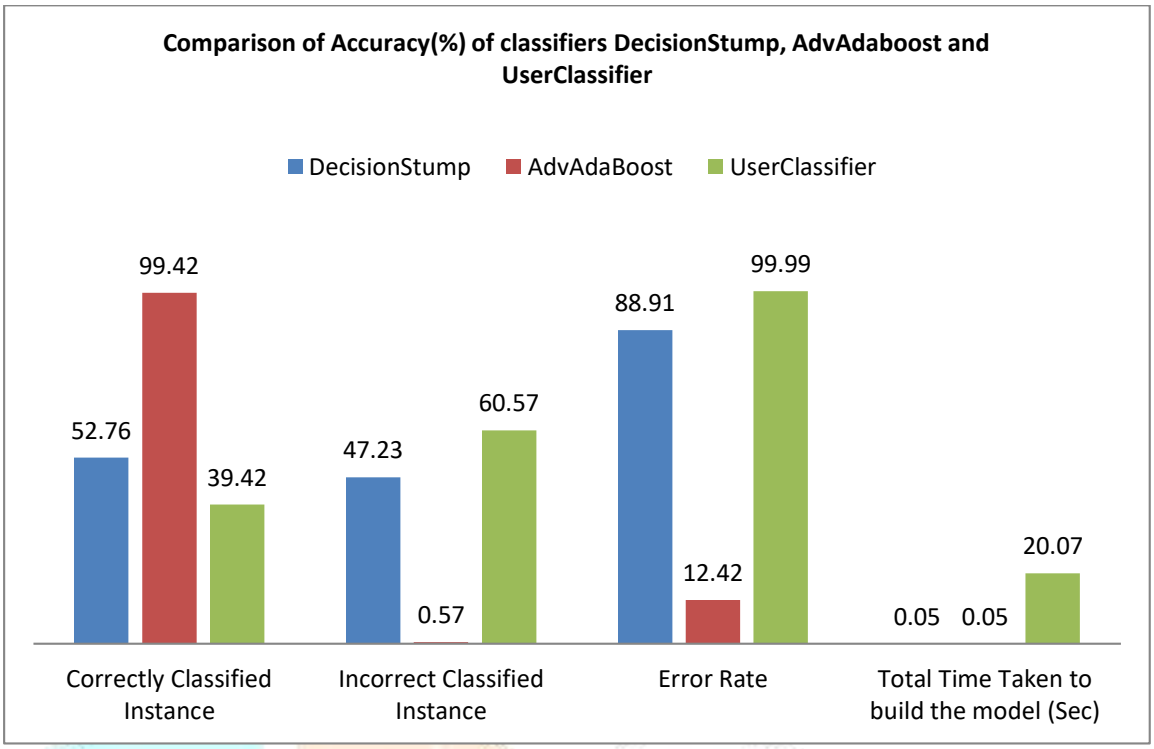

Fig. 1.4: Comparison of Accuracy(%) of classifiers DecisionStump, , Proposed Algorithm (AdvAdaboost) and UserClassifier

4. Conclusion

From the above experiments performed on credit risk dataset using different classification algorithms using different testing methods it was observed that there is proposed technique or algorithm which gives the best results in terms of more classification accuracy and less model development time for all situations.

In this research work we have developed a model using Advadaboost of data mining from heterogeneous classifiers which improves the accuracy of classifiers

References

[1] Lin Zhu, Dafeng Qiu, Daji Ergu, Cai Ying, Kuiyi Liu, " A study on predicting loan default based on the random forest algorithm", 7th International Conference on Information Technology and Quantitative Management (ITQM 2019) , Elsevier 2019.

[2] Xingyun Li,Daj Ergu,Di Zhang,Dafeng Qiu,Ying Caib,Bo Ma, "Prediction of loan default based on multi-model fusion", The 8th International Conference on Information Technology and Quantitative Management (ITQM 2020 & 2021, Elsevier 2019.

[3] Bhoomi Patel, Harshal Patil, Jovita Hembram, Shree Jaswal, "Loan Default Forecasting using Data Mining", 2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020.

[4] Sefik Ilkin Serengil, Salih Imece, Ugur Gurkan Tosun, Ege Berk Buyukbas, and Bilge Koroglu, "Comparative Study of Machine Learning Approaches for Non Performing Loan Prediction with Explainability", International Journal of Machine Learning and Computing, Vol. 12, No. 5, September 2022.

[5] Md. Rafiqul Islam and Md. Ahsan Habib, "A Data Mining Approach To Predict Prospective Business Sectors For Lending in Retail Banking Using Decision Tree", International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.5, No.2, March 2015.

[6] Aboobyda Jafar Hamid and Tarig Mohammed Ahmed, "Developing Prediction Model Of Loan Risk in Banks Using Data Mining" Machine Learning and Applications: An International Journal (MLAIJ) Vol.3, No.1, March 2016.

[7] Fabio Martinelli, Francesco Mercaldo, Domenico Raucci and Antonella Santone, "Bank Credit Risk Management based on Data Mining Techniques", 6th International Conference on Information Systems Security and Privacy (ICISSP 2020).

[8] Varun S Kumar, Abhishek Rokade, Srinath MS, "Bank Loan Approval Prediction Using Data Mining Technique", International Research Journal of Modernization in Engineering Technology and Science Volume:02/Issue:05/May-2020.

[9] Soni P M, Varghese Paul, "Algorithm For the Loan Credibility Prediction System", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019.