



ISSN: 2321-2152

IJMECE

*International Journal of modern
electronics and communication engineering*

E-Mail

editor.ijmece@gmail.com

editor@ijmece.com

www.ijmece.com

Forecasting A Modernized Loan Approval System Using Machine Learning Techniques

J Madhavi ¹

Asst. Professor

Department of CSE (DS)

TKR College of Engineering & Technology

jmadhavi@tkrcet.com**E. Sai Sriraj ²**

B. Tech (Scholar)

Department of CSE(DS)

TKR College of Engineering & Technology

21k91a6734@tkrcet.com**G Himavanth³**

B. Tech (Scholar)

Department of CSE (DS)

TKR College of Engineering & Technology

21k91a6743@tkrcet.com**J Aravind ⁴**

B. Tech (Scholar)

Department of CSE(DS)

TKR College of Engineering & Technology

21k91a6751@tkrcet.com**B Naveen ⁵**

B. Tech (Scholar)

Department of CSE(DS)

TKR College of Engineering & Technology

21k91a6711@tkrcet.com

ABSTRACT

Technology has boosted the existence of humankind the quality of life they live. Every day we are planning to create something new and different. We have a solution for every other problem we have machines to support our lives and make us somewhat complete in the banking sector candidate gets proofs backup before approval of the loan amount. The application approved or not approved depends upon the historical data of the candidate by the system. Every day lots of people applying for the loan in the banking sector but Bank would have limited funds. In this case, the right prediction would be very beneficial using some classes-function algorithm. An example the logistic regression, random forest classifier, support vector machine classifier, etc. A Bank's profit and loss depend on the amount of the loans that is whether the Client or customer is paying back the loan. Recovery of loans is the most important for the banking sector. The improvement process plays an important role in the banking sector. The historical data of candidates was used to build a machine learning model using different classification algorithms. The main objective of this paper is to predict whether a new applicant granted the loan or not using machine learning models trained on the historical data set.

Keywords :- Machine learning, Data, Loan, Training, Testing, Prediction

I. INTRODUCTION

Prediction of modernized loan approval system based on machine learning approach is a loan approval system from where we can know whether the loan will pass or not. In this system, we take some data from the user like his monthly income, marriage status, loan amount, loan duration, etc. Then the bank will decide according to its parameters whether the client will get the loan or not. So there is a classification system, in this system, a training set is employed to make the model and the classifier may classify the data items into their appropriate class. A test dataset is created that trains the data and gives the appropriate result that, is the client potential and can repay the loan.

Prediction of a modernized loan approval system is incredibly helpful for banks and also the clients. This system checks the candidate on his priority basis. Customer can submit his application directly to the bank so the bank will do the whole process, no third party or stockholder will interfere in it. And finally, the bank will decide that the candidate is deserving or not on its priority basis. The only object of this research paper is that the deserving candidate gets straight forward and quick results.

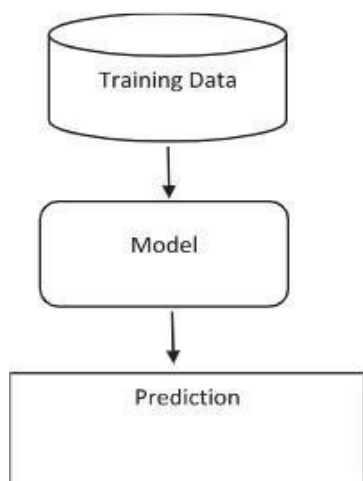


Fig. 1. Basic Machine Learning Model

II. MACHINE LEARNING ALGORITHMS

In this research paper we are using three Machine Learning algorithms which are used to find out the correct prediction of Data set.

- (a) **XG Boost** – XG Boost is a Decision tree based opensource software library. It implements machine learning algorithms that uses a gradient boosting framework. It works on Linux, Windows, and macOS.

- (b) **Random Forest**– Random forests is a classification algorithm which builds big number of Decision tree, whose prediction is more accurate than any of individual decision tree.
- (c) **Decision Tree** – A Decision tree split the dataset in to smaller parts. And then predict the every chances.

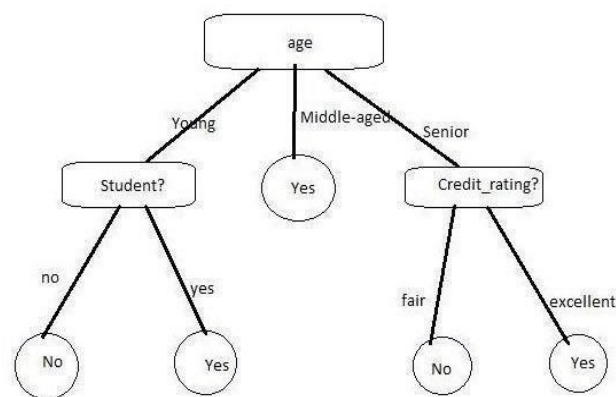


Fig. 2. Decision Tree

III. PROBLEM FORMULATION

There is a major problem that many people not able to back the loans to banks. And banks are going in losses. Banks received many applications for loan approval day by day and not everyone gets approved. Most of the banks have their own credit score and risk assessment techniques so as to check that the loan is approved or not. Why this loan problem arises this question will get resolved in just a few minutes. The main reason to get a loan is to fulfill the needs of something. For a businessman he/she wants to increase the business or if that company is at loss to get over from that he/she needs a loan. In middle-class people wants to fulfill their needs so they want a loan. So, the main thing of this to fulfill the needs of someone or for something. Again the question arises that what are the problems that are forming in providing the loans. The answer to this question that not everybody can loan because if he/she is not able to return then who is providing the loan he/she or the company or the bank that is providing the loan will get in the loss. So, first who is providing the loan they have to verify or set some criteria that who is taking the loan is able to return or not. Like in banks like we have a credit card facility but not everybody gets a credit card. For that, a credit score is there to check whether eligible or not. For credit score one should have a good credit score then he/she be able to get a loan. Some criteria like a source of income should be

there for getting a credit card. Banks provide loans on behalf of one who is taking the loan he/she should provide some documents and verify. Like some company not able to provide the loans then banks get in loss and they called it NBFC's. During this project data processing algorithms are going to study loan-approved data might help in predicting the like defaulters thereby helping the banks for creating better decisions within the future.

IV. REQUIRED TOOLS

- MS Office
- Jupyter notebook
- Python3
- Data set
- Numpy
- Pandas
- XG Boost
- Machine learning algorithms
- Matplotlib

i

V. LOAN PREDICTION DATA ANALYSIS

The question arises that on what basis we analyze that we should provide the loan or not. We have two target variables on that basis we provide the loan to our customer. We have to check all the formalities like income proof, address proof, id proof, etc. Then we provide the loan that the customer is eligible to return or not. In the middle class, there is a major need for loans as parents need for their child's education, for business also.

In some cases, people suddenly undergo financial crises, while some try and scam money from banks. So, the reason we have to check all the things because banks are not undergoing an NPA loan. Better the customer, chances of loan to be back are high. Background verification should be high so that we can expect a return of the loan at the perfect time. So, we analysis on several bases and these are called our target variable

	Description	Type
Dependents	Number of dependents	Integer
Education	Graduate/ Under Graduate	String
Self_Employed	Self Employed	(Y/N) Character
Applicant_Income	Applicant income	Integer
Co_Applicant_Income	Coapplicant income	Integer
Loan_Amount	Loan amount in thousands	Integer
Loan_Amount_Term	Term of loan in months	Integer

Credit_History	credit history guidelines	Integer
Property_Area	Urban/ Semi Urban/ Rural	String
Loan_Status	Loan Approved(Y/N)	Character

VI. LOAN PREDICTION METHODOLOGY

This proposed model will characterize the behavior of customers on the Basis of their record. These records is taken from the customers, and create a data set. With the help of These data sets and training machine learning model, we predict that the customer's loan will passed or not. This Machine algorithms predict the possibility of a customer would be able to repay the loan or not

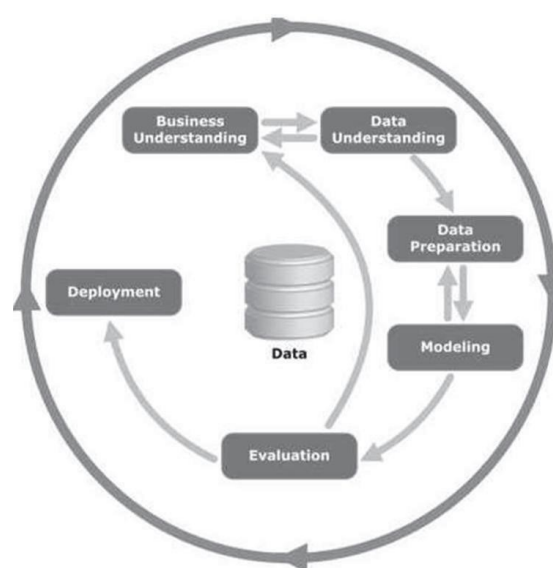


Fig. 3. Process diagram

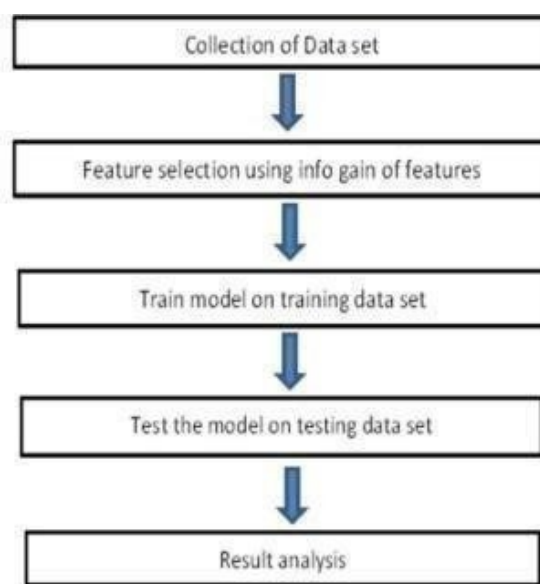


Fig. 4. Loan Prediction Methodology

This Research paper helps the banks to minimize the possible losses and can increase the volume of credits.

VII. ARCHITECTURE DIAGRAM FOR PROPOSED METHOD

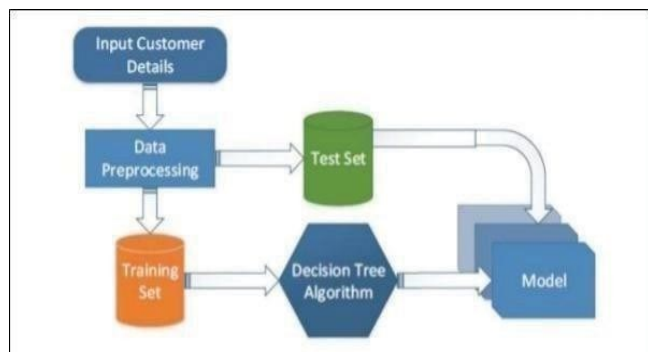


Fig. 5. Architecture Diagram

VIII. CONCLUSION

According to this research paper prediction accuracy is sweet for both datasets. In some situations like client going through some disaster so here the algorithm cannot predict the appropriate result. This conditions by setting the algorithms and just by evaluating the details, we get to know eligibility criteria that client is eligible or not.

This system may be built which is able to takevarious inputs from the users like salary, address, loan amount, loan duration, etc and provide a prediction of whether their application will be approved by the bank or not. I would like to thank Galgotias University for proposing the Capstone Project in our curriculum so that we can learn new concepts easily by doing hands-on encouraging us to learn more by doing practical.

REFERENCE

[1] E. T. Lau, L. Sun, and Q. Yang, "Modelling, prediction and classification of student academic performance using artificial neural networks," *Social Netw. Appl. Sci.*, vol. 1, no. 9, pp. 1–9, Sep. 2019.

[2] W. Wang, Y. Zhao, Y. J. Wu, and M. Goh, "Factors of dropout from MOOCs: A bibliometric review," *Library Hi Tech*, vol. 41, no. 2, pp. 432–453, Jun. 2023.

[3] X. Li and S. He, "Research and analysis of student portrait based on campus big data," in *Proc. IEEE 6th Int. Conf. Big Data Analytics (ICBDA)*, Mar. 2021, pp. 23–27.

research paper can find out the client is potential and repay the loan and the accuracy is good. loan duration, loan amount, age, income are the most important factors for finding out there (whether the client would have been). 'zip code' and 'credit history' are the foremost important factors for predicting the category of the loan Applicant.

Submission message	Score	Code File	Solution File	Final Solution
XGBoost	0.7777777777777778	---	Download	<input type="radio"/>
Random Forest	0.7638888888888889	---	Download	<input type="radio"/>
Decision Tree	0.6458333333333333	---	Download	<input type="radio"/>
Third submission	0.7777777777777778	---	Download	<input type="radio"/>
Second submission	0.7777777777777778	---	Download	<input type="radio"/>
First submission	0.7847222222222222	---	Download	<input type="radio"/>

[4] F. Yang and F. W. B. Li, "Study on student performance estimation, student progress analysis, and student potential prediction based on data mining," *Comput. Educ.*, vol. 123, pp. 97–108, Aug. 2018.

[5] Z. Mingyu, W. Sutong, W. Yanzhang, and W. Dujuan, "An interpretable prediction method for university student academic crisis warning," *Complex Intell. Syst.*, vol. 8, no. 1, pp. 323–336, Feb. 2022.

[6] L. E. Adamova and O. O. Varlamov, "Logic artificial intelligence application for the students individual trajectories introduction," in *Proc. ITM Web Conf.*, vol. 35, 2020, p. 2001.

[7] A. Muhammad, Q. Zhou, G. Beydoun, D. Xu, and J. Shen, "Learning path adaptation in online learning systems," in *Proc. IEEE 20th Int. Conf. Comput. Supported Cooperat. Work Design (CSCWD)*, May 2016, pp. 421–426.

[8] H. Lim, S. Kim, K.-M. Chung, K. Lee, T. Kim, and J. Heo, "Is college students' trajectory associated with academic performance?" *Comput. Educ.*, vol. 178, Mar. 2022, Art. no. 104397.

[9] L. Meng, W. Zhang, Y. Chu, and M. Zhang, "LD-LP generation of personalized learning path based on learning diagnosis," *IEEE Trans. Learn. Technol.*, vol. 14, no. 1, pp. 122–128, Feb. 2021.

- [10] A. A. Mubarak, H. Cao, and W. Zhang, "Prediction of students' early dropout based on their interaction logs in online learning environment," *Interact. Learn. Environ.*, vol. 30, no. 8, pp. 1414–1433, Jul. 2022.
- [11] N. Tomasevic, N. Gvozdenovic, and S. Vranes, "An overview and comparison of supervised data mining techniques for student exam performance prediction," *Comput. Educ.*, vol. 143, Jan. 2020, Art. no. 103676.
- [12] B. K. Bhardwaj and S. Pal, "Data mining: A prediction for performance improvement using classification," 2012, arXiv:1201.3418.
- [13] Y. Zhao, Q. Xu, M. Chen, and G. M. Weiss, "Predicting student performance in a master of data science program using admissions data," in *Proc. Int. Educ. Data Mining Soc.*, Jul. 2020, pp. 1–9.
- [14] G. Su-Hui, B. Cheng-Jie, and W. Quan, "Hadoop-based college student behavior warning decision system," in *Proc. IEEE 3rd Int. Conf. Big Data Anal. (ICBDA)*, Mar. 2018, pp. 217–221.
- [15] A. Akram, C. Fu, Y. Li, M. Y. Javed, R. Lin, Y. Jiang, and Y. Tang, "Predicting students' academic procrastination in blended learning course using homework submission data," *IEEE Access*, vol. 7, pp. 102487–102498, 2019.
- [16] N. Hidayat, R. Wardoyo, A. Sn, and H. Dwi, "Enhanced performance of the automatic learning style detection model using a combination of modified K-means algorithm and naive Bayesian," *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 3, pp. 638–648, 2020.
- [17] N. T. Nghe, P. Janecek, and P. Haddawy, "A comparative analysis of techniques for predicting academic performance," in *Proc. 37th Annu. Frontiers Educ. Conf. Global Eng., Knowl. Without Borders, Opportunities Without Passports*, Oct. 2007, p. 7.
- [18] D. Kabakchieva, "Predicting student performance by using data mining methods for classification," *Cybern. Inf. Technol.*, vol. 13, no. 1, pp. 61–72, Mar. 2013.
- [19] J. L. Rastrollo-Guerrero, J. A. Gómez-Pulido, and A. Durán-Domínguez, "Analyzing and predicting students' performance by means of machine learning: A review," *Appl. Sci.*, vol. 10, no. 3, p. 1042, Feb. 2020.
- [20] R. Asif, A. Merceron, S. A. Ali, and N. G. Haider, "Analyzing undergraduate students' performance using educational data mining," *Comput. Educ.*, vol. 113, pp. 177–194, Oct. 2017.
- [21] M. Yağcı, "Educational data mining: Prediction of students' academic performance using machine learning algorithms," *Smart Learn. Environ.*, vol. 9, no. 1, p. 11, 2022.
- [22] J. Niyogisubizo, L. Liao, E. Nziyumva, E. Murwanashyaka, and P. C. Nshimyumukiza, "Predicting student's dropout in university classes using two-layer ensemble machine learning approach: A novel stacked generalization," *Comput. Educ., Artif. Intell.*, vol. 3, Jan. 2022, Art. no. 100066.
- [23] L. Huang, C.-D. Wang, H.-Y. Chao, J.-H. Lai, and P. S. Yu, "A score prediction approach for optional course recommendation via cross-userdomain collaborative filtering," *IEEE Access*, vol. 7, pp. 19550–19563, 2019.
- [24] A. Alshanqiti and A. Namoun, "Predicting student performance and its influential factors using hybrid regression and multi-label classification," *IEEE Access*, vol. 8, pp. 203827–203844, 2020.
- [25] M. Akour, H. A. Sghaier, and O. Al Qasem, "The effectiveness of using deep learning algorithms in predicting students achievements," *Indonesian J. Electr. Eng. Comput. Sci.*, vol. 19, no. 1, p. 388, Jul. 2020.