

Loan Eligibility Prediction Using Machine Learning

Gorantla Lavanya¹, Bobbala Naga Sunitha², Konkala Sai Kalpana³, Ravinutala V P SaiViswanadh Sarma⁴, B. Sravani⁵, and Nedunchezian⁶

^{1,2,3,4,5,6} Department of Computer Science and Engineering, PACE Institute of Technology & Sciences, Vallur, Ongole, Andhra Pradesh, India

Correspondence should be addressed to Gorantla Lavanya; 18kq1a0523@pace.ac.in

Copyright © 2022 Made Gorantla Lavanya et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT- Banks and other financial institutions compete for customers by providing a wide range of services and products. Most banks, however, make the vast majority of their money from their credit portfolio. Loans accepted by borrowers might lead to interest charges. The loan portfolio, and customers' repayment habits in particular, can have a substantial impact on a bank's bottom line. The financial institution's Non-Performing Assets can be reduced if it can accurately predict which borrowers are likely to default on their loans. Therefore, there is substantial scholarly value in exploring the prediction of loan endorsement. In order to make accurate predictions, it is crucial to use Machine Learning methods. Based on a person's past loan qualification history, this research uses a machine learning methodology to predict the person's likelihood of consistently making loan repayments. The primary aim of this research is to foretell how likely it is that a given individual will be granted a loan.

KEYWORD- Loan, Machine Learning, Prediction, Testing, Training

I. INTRODUCTION

One of the most important factors which affects our country's budget and the financial state is the credit system governed by the banks. The process of bank credit risk assessment is recognizing at banks across the globe. "As we know credit risk assessment is very crucial, there are variety of techniques are used for risk level calculation. Small loan is an important aspect of our everyday life: it allows ambitious entrepreneurs to get started on ideas that could be full-fledged into business; it allows curious students to afford higher education that is otherwise unattainable without a stable income; more importantly, it allows ordinary people who have no friends or relatives for support to procurement short-term financial assistance and get back on their feet to fight for the American Dream. Nevertheless, with loan it comes with the opportunity of default as well. Default is a financial term describing the failure of meeting the legal responsibility of a loan - paying back the principal and interest. It's a common problem in the financial trades and one of the major risks of offering loans. Of course, default does not happen the mainstream of the time and the

lending banks usually able to make up the loss from a avoidance loan from other fully paid loans and their accompanied interests. Furthermore, banks issuing loans with higher interest rate to individuals with high probability of default - the financial organizations are trading off an increased chance of default with an increased profit from the high interest. All things considered, default is a fact of life and most financial institutions have a entrenched practice to minimize its impact and absorbing the loss. But what about a condition where instead of a single bank is issuing the loan, the loan is contained of funds from several investors? Lending Club is one of the many peer-to-peer lending company that gives rise to this abnormal situation.

In plain words, peer-to-peer lending company acts as a broker between pledges and investors. The company creates a platform where borrowers can create small unsafe personal loans, and investors can seek out these loans and decide which loans to invest from. Borrowers obtain the loan they want, financiers get to profit from the loan interest, and the company gets a cut from both parties (origination fee from pledges and service fee from investors). This also means that when a loan goes default, it's no longer a single bank that is absorbing the loss - single or multiple individual investors will be riveting it instead. The overall profit might be positive fall the loans were originated from a single end era so their fully paid loans could cover the loss. An obvious resolution to this problem is to predict whether a particular loan will go default based on initial information provided by the pledgers and their credit report.

This paper will explore the process and result on formulating and machine learning model that could calculate a loan default but more importantly, the model will focus on minimizing the overall loss in investment of bad loans in order to reduce the burden passed onto individual investors. As aside note, the paper will also explore privacy-preserving mechanism on sensitive information provided from the derives credit report. The end goal is to evaluate a simplified version of RAPPOR (Randomized Aggregately Privacy Preserving Ordinal Response) and decide whether data that have been hashed by this algorithm could still be use to predict loan default as stated before.

II. PROBLEM STATEMENT

Account firms and banks need to automatism the credit qualification action (continuously) essentially dependent on data given by customers when turning out an online structure. Sex, Marital Status, Education, Number of Wards, Salary, Loan Amount, Credit History, and different subtleties are integrated. To digitize this interaction, they made an issue to group the client sections that can apply for a credit sum, approving them to focus on these clients explicitly. They have presented a fractional informational collection for this situation.

“Approval of Loan is a very common real-life problem that every company faces in their lending processes. If the loan approval process is automated, it can save a lot of man hours and improve the speed of service to the clients. “However, the rewards can only be realized if the bank has a sturdy model in place to perfectly forecast which client's loans it should accept and which it should reject, in order to reduce potential risk”.

Variable	Description	Type
<u>Loan_ID</u>	<u>UniqueLoanID</u> (Primary Key)	Integer
Gender	Female or Male	Character
Married	Yes or No	Character
Dependents	Number of dependencies	Integer
Education	Graduate /Un-graduate	String
Self Employed	YesorNo	Character
Applicant Income	Applicant's Income	Integer
<u>CoApplicant Income</u>	Co-applicant's Income	Integer
<u>Loan_Amount</u>	Loan Amount	Integer
<u>Loan_Amount Term</u>	<u>Termofloan(In months)</u>	Integer
<u>Credit_History</u>	<u>Applicant'sCredit history</u>	Integer
Property_Area	Urban/Semi Urban/Rural	String
<u>Loan_Status</u>	<u>YesorNo</u>	String

A. Dataset

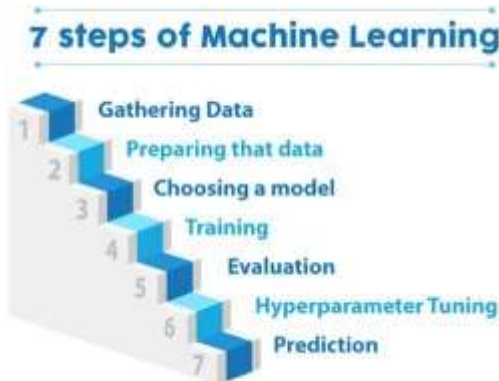


Figure 1: Methodology for Typical Machine Learning approach

III. MACHINE LEARNING ALGORITHMS

A. Algorithms used for Prediction

a) Logistic Regression

It is a classification set of rules used to assign observation to a discrete set of guidelines. Logistic regression is also a predictive analysis, like other regression analyses systems. Logistic regression is basically used for define the relationship between dependent binary variable and token or other independent variable. Now a day's logistic regression is used in many investigation areas like medical science, machine learning and social science. It also used by many e-commerce presentations to predict them in d set of customer to buy the product.

B. Random Forest

Random Forest is a robust system learning algorithm that is used For a ramification of accountabilities along with classification and regression. Random forests method overcome the over fitting issue of decision trees throughout training. It is an ensemble method made up of a large number of small decision trees called estimators where each tree products the prediction. There and random forest model combines the predictions of the estimators to harvest a more accurate prediction.

C. Correlation between Parameters

Heat map is a data visualization technique that shows the greatness of a phenomenon scalar in two sizes as shown in figure2. Color intensity shows the relationship between each other. The color variation may be by hue or intensity, giving an obvious visual to the reader about how the occurrence is clustered or varies over space. From this heat map, it is negative relation in Loan amount term with presentation come attribute.

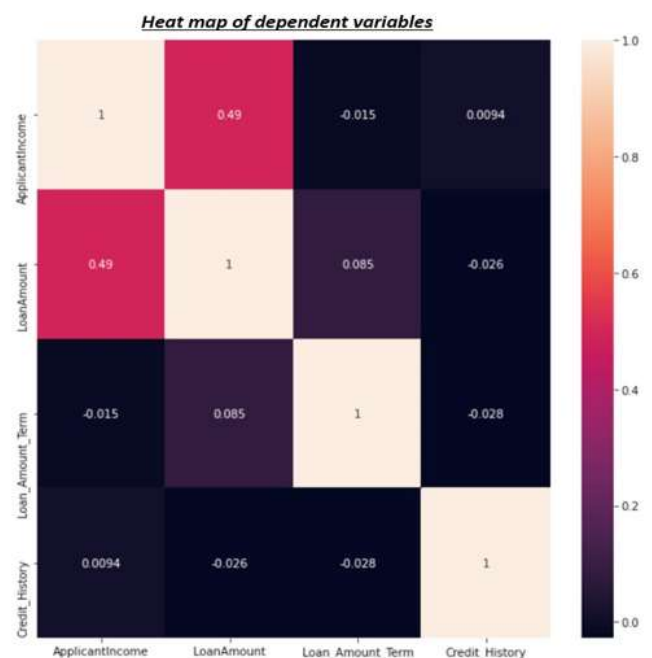


Figure 2: Heat Map

IV. EXPERIMENT AND RESULT ANALYSIS

In this section, we use a machine learning algorithm on a loan prediction dataset and organize the result using HTML,CSS, Django at the limited server. Figure 3 shows the loan prediction system of the aspirants based on the value enter by the bank employee.

Figure 3: Final Layout

The first attribute is to select the gender of the aspirant either male or female. The second is the marital status the independent attributes mean that the aspirant is dependent financially on someone or not. Other attributes are the education of the customer, employment status, aspirant income, loan amount term, credit history, residential area, etc. of the applicant. Finally, it shows the position of the loan i.e. is safe or risky as shown in fig 4.

Figure 4: Final Result

V. CONCLUSION

In this paper, machine learning was used to predict loan recognition. The prediction method begins with data pre-processing, filling the missing values, tentative data analysis. This model is used for the banking system or anyone who needs to apply for a loan. It will be very supportive in bank management. From the analysis of the data, it is very clear that it moderates all the frauds done at the time of loan approval. Time is also very exquisite for everyone through this not only the bank but also the waiting time of the applicant will also reduce.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

- [1] Ashwini S.Kadam, Shraddha RNikam ,AnkitaA .Aher, GayatriV.Shelke, AmarS.Chand gude.
- [2] "Prediction for Loan Approval using Machine Learning Algorithm", Apr2021 International Research Journal of Engineering and Technology(IRJET).
- [3] Mohammad Ahmad Sheikh, Amit Kumar Goel, Tapas Kumar. "An Approach for Prediction of Loan Approval using Machine Learning Algorithm", 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020
- [4] X.FrencisJensy, V.P.Sumathi,Janani Shiva Shri,"An exploratory Data Analysis for Loan Predictionbasedonnatureofclients",InternationalJournal of Recent Technology and Engineering(IJRTE),Volume-7Issue-4S,November2018
- [5] J. Tejaswini1, T. MohanaKavya, R. Devi NagaRamya, P. Sai Triveni VenkataRao Maddumala. "Accurate Loan Approval Prediction Based On Machine Learning Approach" Vol 11, www.jespublication.com, page 523 Issue4, April/ 2020
- [6] "Prediction for Loan Approval using Machine Learning Algorithm" International Research Journal of Engineering and Technology (IRJET)e-ISSN: 2395-0056 Volume: 08 Issue: 04 | Apr2021 www.irjet.netp-ISSN: 2395-0072
- [7] Vaidya,"Predictive and probabilistic approach using logistic regression: Application to prediction of loan approval," 2017 8th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Delhi,2017,pp.1-6.doi:10.1109/ICCCNT.2017.8203946
- [8] M. Bayraktar, M. S. Aktaş, O. Kalıpsız, O. SusuzandS.Bayracı,"Credit risk analysis with classification Restricted Boltzmann Machine," 2018 26th Signal Processing and Communications Applications Conference (SIU), Izmir, 2018, pp. 1-4.doi: 10.1109/SIU.2018.8404397
- [9] Y. Shi and P. Song, "Improvement Research ontheProjectLoanEvaluationofCommercialBankBasedontheRiskAnalysis,"2017 10th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, 2017,pp.3-6.doi:10.1109/ISCID.2017.60
- [10] V. C. T. Chan et al., "Designing a Credit Approval System Using Web Services, BPEL, and AJAX,"2009 IEEE International Conference on e-Business Engineering, Macau,2009,pp.287-294.doi:10.1109/ICEBE.2009.46