Smart Lender-Applicant Credibility Prediction for Loan Approval

Team Leader :

Mr. Ragul V

Team Members :

Mr. Aswath P

Mr. Murali Krishnan N

Mr. Nivas Saravana N

Guide :

Dr. Ramya J

1. INTRODUCTION :

The primary source of revenue for the banking industry and source of financial risk for banks is a loan. The interest collected on loans disbursed directly accounts for significant amounts of a bank's assets. The ability of the borrower to repay the loan within the allotted period is one of the major hazards associated with the lending of loans. It's known as "credit risk." A candidate's credit score served as the basis for determining whether or not to approve a loan on them. Thus, the purpose of this study is to describe the use of various machine learning approaches that effectively identify who to lend money to and assist banks in identifying loan defaulters for significantly lower credit risk.

The fact that our banking system offers a wide range of goods, a bank's credit line is its primary source of income. As a result, they are able to profit from interest on the loans they credit. Lenders always seek to lower their credit risk since commercial loans have historically represented a significant portion of the banking sector. The function that banks play in the

modern market economy is substantial. Loans, or whether clients repay or don't return them, significantly impact a bank's profitability. Before granting loans to borrowers, banks must determine if they are good (non-defaulters) or bad (defaulters).

The creditworthiness of the borrowers is one of the most critical issues in commercial loan financing.

The probability that borrowers may default on their loan commitments is referred to as credit risk. For any bank or institution, determining whether a borrower will be good or bad is a very difficult process. The banking system employs a manual procedure to determine whether or not a borrower has defaulted. The manual method will undoubtedly be more precise and efficient, but it will not be able to handle a high volume of loan applications at once. When a situation like this arises, it will take a very long time to make decisions and a lot of labour will be needed.

2. OBJECTIVE :

Loans account for a large portion of bank profits. For financial companies, the loan approval process is crucial. Because loan defaults are occurring more frequently and it is becoming more challenging for banking authorities to properly access loan requests and address the dangers of people defaulting on loans, it is very difficult to forecast if clients will be able to pay back the loan. Numerous scholars have been focusing on loan approval system prediction in recent years. For vast amounts of data, the machine learning technique is highly helpful in predicting outcomes. Four algorithms, including Random Forest, Decision Tree, Naive Bayes, and Logistic Regression, are employed in this study to forecast whether or not clients would be approved for loans.The same information will be analysed for all four methods, and the most accurate algorithm will be chosen to deploy the model. From this point forward, we create a machine learning-based bank loan prediction system that chooses the qualified applicants for loan approval on its own.

3. LITERATURE SURVEY:

M. Cary Collins ,Sivasree M S & Rekha Sunny T :

DESCRIPTION: There is a lot of room for improvement in information quality and data problem prevention in bank data management for loan approval processes, particularly with relation to fair lending and fair pricing policies. Prior to that, they quickly went over the standard data collection procedures used by many financial institutions for loan approval and loan pricing. A subset of these data methods must comply with federal requirements. They demonstrated some of the basic critical processes currently required for increasing information quality for all stakeholders involved while discussing the data collecting and analysis for fair lending.

Data mining techniques are growing in popularity as a result of the abundance of data and the demand for turning that data into knowledge. Numerous industries, including the retail sector, the study of biological data, the telecommunications sector, intrusion detection, and other scientific applications, use data mining techniques. In order to compete effectively in the market, data mining techniques are frequently applied in the banking sector. They established a prediction model for bankers in this research that will aid them in determining which consumers who have asked for loans are credible. The Decision Tree Algorithm is being used to forecast the credibility-relevant qualities.In this paper, a prototype of the model is provided that can be applied by companies to decide whether to approve or deny loan requests from clients.

Kumar Arun, Garg Ishan, Kaur Sanmeet:

DESCRIPTION: Due to the expansion of the banking industry, a large number of people have applied for bank loans. However, because banks have limited resources, they can only lend to a certain number of borrowers. In order to save a lot of bank resources and work, they sought to lessen this risk in this paper by choosing the safe person. The process

involved mining the prior loan recipients' records, and the machine was trained using the machine learning model that produced the best accurate result based on this information.This essay's primary objective is to determine whether or not assigning a loan to a particular person will be secure. The four components of this work are as follows: I data collection (ii) comparison of machine learning models using the data gathered (iii) system training using the most promising model (iv) system testing.

Advantages:

1.The algorithms' efficiency and accuracy can be calculated and compared.

2.Machine learning techniques can be used to address the class imbalance.

Disadvantages:

1.They did not employ machine learning methods; instead, they offered a mathematical model.

2.The issue of class imbalance was not addressed, nor were the necessary steps made.

5. CONCLUSION:

We have selected the machine learning technique to analyse the bank dataset in order to forecast the loan approval status of the applied client. We tested a number of machine learning algorithms to see which one would work the best on the dataset to get the most accurate results. Using this strategy, we discovered that, aside from logistic regression, the other algorithms delivered accurate results in a suitable manner.

The remaining algorithms' accuracy ranged from 75% to 85%. The logistic regression, however, provided us with the highest level of accuracy (88.70%) after comparison of all the methods.

We also identified the key elements that have the greatest impact on the loan approval status. The performance accuracy of these most crucial characteristics applied to a few chosen algorithms is then compared to the case where all features were employed. The banks may use this model to determine what elements are crucial for the loan approval process. Based on their accuracy, the comparison research clearly identifies which algorithm will be the best and disregards the others.

6. REFERENCES:

[1]   Arun Kumar, Ishan Garg, and Sanmeer Kaur, \"Loan Approval Prediction Using Machine Learning Approach,\" 2018.

[2]   K. Hanumantha Rao, G. Srinivas, A. Damodhar, and M. Vikas Krishna at International Journal of Computer Science and Telecommunications published an article titled \"Implementation of Anomaly Detection Technique Using Machine Learning Algorithms\" (Volume2, Issue3, June 2011).

[3]   G. Arutjothi and C. Senthamarai, \"Prediction of loan status in commercial banks using machine learning classifier,\" International Conference on Intelligent Sustainable Systems (ICISS), 2017.

[4]   \ ``AzureML based analysis and prediction of loan applicants creditworthy,\" by Alshouiliy K, Alghamdi A, and Agrawal D P I n 2020, Third International conference on information and computer technologies.

[5]   \"Developing prediction model of loan risk in banks using data mining Machine Learning and Applications,\" Hamid A J and Ahmed T M, 2016.

[6]   M. Li, A. Mickel, and S. Taylor \"Should this loan be approved or denied?\" published a paper in the Journal of Statistics Education in 2018.

[7]   A. Vinayagamoorthy, M. Somasundaram, and C. Sankar, \"Impact of Personal Loans Offered by Banks and Non-Banking Financial Companies in Coimbatore City,\" 2012.

[8]   M. Cary Collins, Ph.D., and Frank M. Guess, Ph.D., MIT\'s Information Quality Conference, 2000, \"Improving information quality in loan approval processes for fair lending and fair pricing.\"

[9]   Arun Kumar, Ishan Garg, and Sanmeet Kaur, \"Loan approval prediction based on machine learning approach,\" National Conference on Recent Trends in Computer Science and Information Technology, 2016.

[10]   Sivasree M S and Rekha Sunny T, \"Loan Credibility Prediction System Using Decision Tree Algorithm,\" International Journal of Engineering Research & Technology (IJERT), Vol. 4 Issue 09, September-2015.

[11]   Ji?í Doležal, Ji?í Šnajdr, Jaroslav Belás, Zuzana Vincúrová, "Model of the loan process in the context of unrealized income and loss prevention", Journal of International Studies, Vol. 8, No 1, 2015, pp. 91-106. DOI: 10.14254/2071-8330.2015.

[12]   M. A. Sheikh, A. K. Goel, and T. Kumar, "An Approach for Prediction of Loan Approval using Machine Learning Algorithm," 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC), 2020, pp. 490-494, DOI: 10.1109/ICESC48915.2020.9155614

[13]   K. Alshouiliy, A. AlGhamdi, and D. P. Agrawal, "AzureML Based Analysis and Prediction Loan Borrowers Creditworthy," 2020 3rd International Conference on Information and Computer Technologies (ICICT), 2020, pp. 302-306, DOI: 10.1109/ICICT50521.2020.00053.