



Loan Eligibility Prediction

Rushikesh S. Dwarke, Kunal K. Mane, Saurabh S. Maharanwar, Nadeem B. Mulani, Prof. Shiv Shankar

Department of Computer Engineering, Savitribai Phule Pune University

ABSTRACT

The standard of living for humans has improved thanks to technology. We aim to produce something fresh and novel every day. There is an answer for every other issue. In the financial sector, an applicant receives proofs or backup before the loan amount is approved, providing us with tools that sustain our existence and help us feel somewhat complete. The system's past data on the applicant determines whether or not the application is accepted. Numerous individuals ask for loans every day in the banking industry, but the bank's resources are constrained. A classification algorithm would be very helpful in this situation if the correct forecast could be made. Examples include support vector machine, logistic regression, and random forest classifiers. The quantity of loans a bank makes or loses depends on whether the client or customer pays back the debt. The most crucial task for the banking industry is debt recovery. 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 is granted the loan or not using machine learning models trained on the historical data set.

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

Introduction

A loan clearance method using machine learning where we can determine whether or not the credit will be approved. We gather details about the customer for this method, such as his monthly salary, marital situation, loan size, length of loan, etc. The bank will then determine if the customer will receive the credit based on its criteria. Thus, a classification system exists. In this system, a training set is used to create the model, and the classifier is then able to categorize the data objects into the correct classes. In order to train the data and produce the correct outcome, which is the client's potential and ability to repay the debt, a test dataset is produced. For banks and customers alike, predicting the outcome of a modernized credit approval procedure is very useful. This method evaluates each candidate based on their significance.

The customer can send his application immediately to the bank, in which case the bank will handle the entire process without interference from a third party or stockholder. Finally, based on its priority system, the bank will determine whether the applicant is worthy or not. This research paper's sole goal is to ensure that the worthy applicant receives straightforward answers right away.

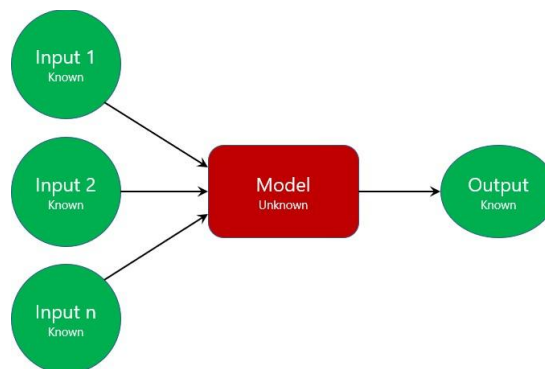


Figure 1: Machine Learning Model

Problem Statement

There is a significant issue that prevents many individuals from supporting financial debts. Additionally, institutions are losing money. For the loan clearance day, banks get a lot of applications, by day, and not everyone is granted permission. The majority of banks use their own risk evaluation methods and credit scoring systems to determine whether or not a loan is authorized. In a matter of minutes, the issue of why there is a debt problem will be answered. The primary purpose of obtaining a credit is to meet a need. A credit is necessary for a merchant if they want to grow their company or if they need to recover from a loss.

Related Work

Loan price prediction is a common problem in the field of finance and economics, and as such, there is a significant body of related work in this area. Here are a few examples:

"Predicting Loan Default Probabilities with Machine Learning" by Edward Robinson and Mary Clare Peate. In this paper, the authors use machine learning algorithms to predict the likelihood of loan defaults. They compare several algorithms and find that random forest and gradient boosting are the most effective.

"Predicting Personal Loan Default Probability Using Machine Learning Techniques" by Rajesh Kumar and Vipin Pal. This paper examines the effectiveness of several machine learning techniques in predicting the likelihood of personal loan defaults. They find that support vector machines (SVMs) and artificial neural networks (ANNs) are the most effective.

"Loan Default Prediction Using Machine Learning Techniques: A Review" by Muhammad Khurram Khan and Muhammad Bilal Amin. This paper provides a comprehensive review of various machine learning techniques used for loan default prediction. The authors discuss the strengths and weaknesses of each technique and provide insights into how they can be applied to this problem.

"A Comparative Study of Machine Learning Techniques for Loan Default Prediction" by Rakesh

Kumar, Rajesh Kumar, and Ruchika Malhotra. In this paper, the authors compare the effectiveness of various machine learning techniques for predicting loan defaults. They find that SVMs and decision trees are the most effective, and that feature selection techniques can significantly improve the accuracy of predictions.

"Loan Performance Prediction Using Deep Learning" by Guozhu Liu, Liang Zhang, and David G. Luenberger. In this paper, the authors use deep learning techniques to predict loan performance. They propose a deep neural network architecture that takes into account both borrower and loan features, and find that their model outperforms traditional machine learning algorithms.

These are just a few examples of the many papers that have been published on loan price prediction. By studying the related work, you can gain a better understanding of the techniques and algorithms that have been successful in this area and how they can be applied to your specific problem.

Problem Formulation

A significant issue is that many people are unable to repay bank debts. Additionally, banks are losing money. Every day, banks receive a large number of loan applications, but not all of them are granted. The majority of banks use their own risk assessment methods and credit scoring systems to determine whether or not a loan is authorized. In a matter of minutes, the issue of why there is a loan difficulty will be answered. The primary purpose of obtaining a loan is to meet a need. A loan is necessary for a businessman if they want to grow their company or if they need to recover from a loss.

Once more, the issue of what issues are developing in credit provision is raised. The answer to this question is that not everyone qualifies for loans because if the borrower is unable to repay the loan, either they themselves or the business or bank that provided it would suffer a loss. Therefore, the person granting the loan must first confirm or establish some criteria

as to whether the person receiving it is able to repay it or not. Like with banks, we have a credit card option, but not everyone qualifies for one. For that, a credit score is available to determine eligibility. To be eligible for a loan, a person must first have a strong credit score. A source of income and other requirements should be present in order to obtain a credit card. Banks offer loans, but the person applying for the loan must submit supporting documentation and undergo verification. For example, when a business is unable to deliver loans, banks suffer losses and refer to them as NBFCs.

In this project, data processing algorithms will analyse loan-approved data in order to identify similar defaulters, which will aid banks in making better judgments in the future.

Proposed Work

System Architecture

In this, we are going to discuss the advantage of loan prediction. With the use of this technique, we can determine whether or not a loan applicant will be able to repay the loan. We forecast that the client will be qualified for a loan if they can repay the amount. Additionally, we forecast that the customer is ineligible if the candidate fails. The benefit of this approach is that we can determine if a client meets the eligibility requirements simply by assessing the facts once we've set some conditions through the use of algorithms. This system may be created to forecast if

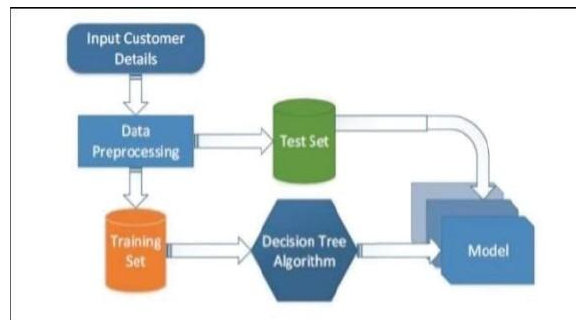


Figure 2: Architecture Diagram

a user's loan application will be granted by the bank based on inputs from the user such as salary, address, loan amount, and loan term.

Loan Prediction Methodology

This suggested model will evaluate consumer behaviour based on their past behaviour. These client records are collected to form a data set. We forecast whether the customer's loan will be approved or not using these data sets and a training machine learning model. This Machine algorithms predict the possibility of a customer would be able to repay the loan or not

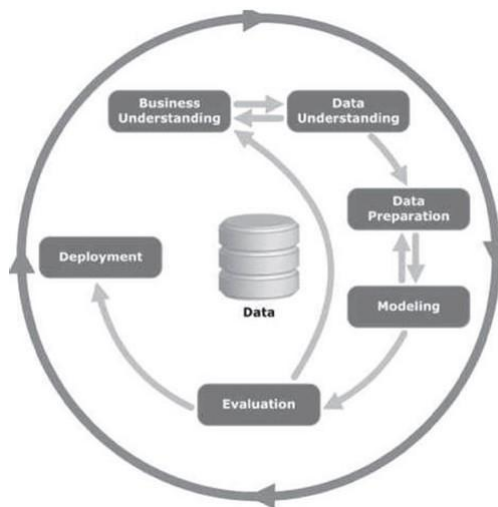


Figure 3: process diagram

Loan Prediction Data Analysis

The question of how we evaluate whether or not to provide the loan emerges. We supply the loan to our consumer based on two goal criteria. We must verify all the requirements, including evidence of income, residence, and identification. The consumer is then given the loan, whether they are qualified to return it or not. The middle class has a significant demand for loans since parents need them for their children's education as well as for their businesses.

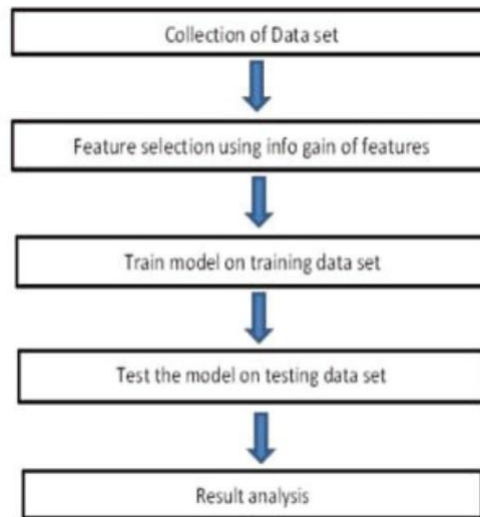


Figure 4: Loan Prediction Methodology

Some people have abrupt financial crises, while others attempt to defraud banks of their money. As a result, we must double-check everything because banks are not experiencing NPA loans.

Higher possibilities of loan repayment are associated with better customers. Background checks should be thorough so that we can anticipate receiving the loan back at the ideal moment. As a result, we analyse on many bases, and they are referred to as our target variables.

Advantages

Increased Efficiency: With loan eligibility prediction models in place, loan officers can quickly assess whether a potential borrower meets the eligibility criteria for a loan. This can save time and resources that would otherwise be spent manually reviewing applications.

Improved Accuracy: Machine learning algorithms can be trained on historical loan data to identify patterns and predict the likelihood of a borrower defaulting on a loan. This can lead to more accurate assessments of loan eligibility and reduce the risk of bad loans.

Better Customer Experience: By automating the loan eligibility process, lenders can provide customers with faster loan approvals and reduce the likelihood of human error in the decision-making process. This can result in a better customer experience and increased customer satisfaction.

Reduced Costs: The use of loan eligibility prediction models can help lenders identify high-risk loans and avoid lending to borrowers who are unlikely to repay the loan. This can reduce the number of bad loans and associated costs, such as collection costs and legal fees.

Required Tools

MS Office

Microsoft Office is a suite of productivity software developed and published by Microsoft Corporation. It includes various applications that are commonly used in business, education, and personal settings. Some of the key applications included in Microsoft Office are:

Microsoft Word: A word processing application that is used to create and edit text-based documents such as letters, reports, and resumes.

Microsoft Excel: A spreadsheet application that is used to organize and analyze data, create charts, and perform complex calculations.

Microsoft Outlook: An email client and personal information manager that is used to manage email, calendars, contacts, and tasks.

Jupyter notebook

Jupyter Notebook is an open-source web application that allows users to create and share interactive code, visualizations, and text-based documents. Originally developed for the Python programming language, Jupyter Notebook now supports over 40 programming languages.

Jupyter Notebook is organized into cells, which can contain code, text, or visualizations. Code cells allow users to write and execute code directly in the notebook. Text cells can be used to document code, write notes, or provide context for the code. Visualization cells allow users to create and display charts, graphs, and other visualizations.

Jupyter Notebook has several key features that make it a popular tool for data analysis, scientific computing, and education.

Python3

Python 3 is the latest version of the Python programming language. It was released in 2008 and is the successor to Python 2. Python 3 introduced several key changes and improvements over Python 2, including: Print function: In Python 3, the print statement was replaced with a print function, which makes it easier to format and print output.

Unicode support: Python 3 supports Unicode natively, which makes it easier to work with international characters and languages.

Numpy

NumPy (Numerical Python) is a Python library for numerical computing. It provides support for large, multi-dimensional arrays and matrices, along with a wide range of mathematical functions for working with these arrays.

NumPy is a core library for scientific computing in Python, and is widely used for tasks such as data analysis, machine learning, and scientific simulations.

Pandas

Pandas is a Python library for data manipulation and analysis. It provides support for handling data in a variety of formats, including CSV, Excel, SQL databases, and more. Pandas is built on top of NumPy, and provides additional functionality for working with labeled data, such as data frames and series.

Conclusion

This study claims that both datasets' prediction accuracy is excellent. When a customer is experiencing a disaster, for example, the algorithm may not be able to forecast the right outcome. This study can determine if a potential consumer would return a loan, and its accuracy is good. The most crucial variables for determining (if the customer would have been) are loan length, loan amount, age, and income. Zip code and credit history are the two most crucial variables in determining the loan applicant's category.

References

- Amruta S. Aphale and R. Prof. Dr. Sandeep. R Shinde, "Predict Loan Approval in Banking System Machine Learning Approach for Co-operative Banks Loan Approval", International Journal of Engineering Trends and Applications (IJETA), vol. 9, issue 8, 2020).
- Loan Prediction Using Ensemble Technique, International Journal of Advanced Research in Computer and Communication Engineering Vol. 5, Issue 3, March 2016