The Web Based Loan Prediction App Using Machine Learning

Jaganathan K T[1], Janarthanan S[2], Karthik R[3], Sanath Pandian R[4], Mr.Nagarasan[5]

[1, 2, 3,4]Dept of CSE

[5]Assistant professor, Dept of Mechanical Engineering

[1, 2, 3, 4, 5]Info Institute of Engineering

Abstract- *Here in this work efforts were made to predict whether the applicant is eligible or not to get the loan in the bank by using Machine learning Algorithms.Over the last decade, digital credit has been the fastest growing financial innovation in India. This has largely been attributed to by technological innovations and mobile phone penetration enabling expanded access to financial services to individuals who were previously unbanked. The main Abstract of this project is to predict loan default by applying machine learning algorithms. The proposed methodology involves data collection, data pre-processing, data analysis, model selection and performance evaluation. This project takes data of previous customers to whom on a set of parameters loan were approved. The machine learning model is then trained on that record to get accurate results. The main machine learning algorithms applied are logistic regressions, naïve bayes and decision trees. The performance of the machine learning models are then compared using performance metrics and the best machine learning algorithm is selected to predict the loan default.*

I. INTRODUCTION

This system is designed and implemented to predict and provide the eligibility of the applicant, whether he/she gets a Loan in Bank or not. Loan Prediction is very helpful for employee of banks as well as for the applicant also. The aim of our project is to provide quick, immediate and easy way to choose the deserving applicants. Dream housing Finance Company deals in all loans. They have presence across all urban, semi urban and rural areas. Customer first apply for loan after that company or bank validates the customer eligibility for loan. Company or bank wants to automate the loan eligibility process (real time) based on customer details provided while filling application form.These details are Gender, Marital Status, Education, Number of Dependents, Income, Loan Amount, Credit History and other. This project has taken the data of previous customers of various banks to whom on a set of parameters loan were approved.

1.1 Objective

Here the objective of the work was to predict the accuracy between seven algorithms to find the best model with higher accuracy and use that algorithm to predict the loan status of the applicant. The Main objective of the problem is to pick out which customer will be able to pay the debt and which customer is likely will not be able to pay the debts. Clearly we have to create a classification model here. We have to use algorithms like logistic regression, decision tree, random forest,etc. We need to create a model that is accurate and the error percentage should be less. The main objective of this project is to predict whether assigning the loan to particular person will be safe or not. ... In this project we predict the loan data by using some machine learning algorithms they are classification, logic regression, Decision Tree, etc. A classification model is run on data attempting to classify whether the person or client is eligible for get loan from any bank with good accuracy of statement. Our objectives included some points about this Loan Status Prediction.

1.2 Problem definition

Automate the loan eligibility process based on customer details provided while filling online application form Based on the details provided by customers, we have to create a model that can decide where or not their loan should be approved. This completes the problem definition part of the first stage of the machine learning lifecycle. The next step is to generate hypotheses and point out the factors that will help us to predict whether the loan for a customer should be approved or not.

II. PROPOSED SYSTEM

Loan Approval System is software used for approval of loan in banking sector. In this proposed system we have used machine learning algorithm. Machine Learning is process in which a symmetric model is build from the existing dataset; this model is applied for the testing of the new dataset. The system consists of trained dataset and test dataset. The trained dataset is used for construction of model. This model is

applied on testing dataset for the required result. We have used Ensemble approach for building of the model.
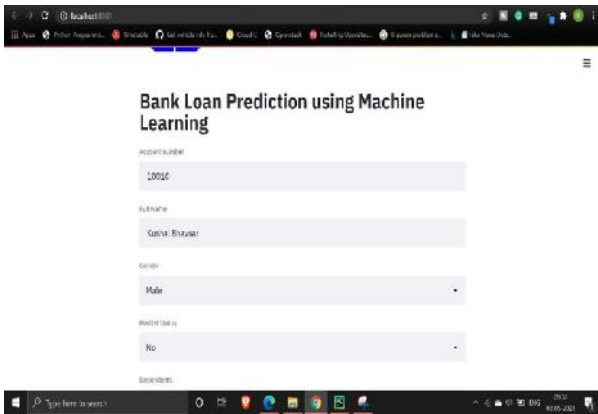
To deal with the problem, we developed automatic loan prediction using machine learning techniques. We will train the machine with previous dataset, so machine can analyse and understand the process. Then machine will check for eligible applicant and give us result. Eligible applicant will be sanctioned loan without any delay. Whole process will be automated , so human error will be avoided.

In this project we try to develop user interface flexibly graphics concepts in mind, associated through a browser interface. Our goal is to implement machine learning model so as to classify, to the best potential degree of accuracy, master card fraud from a dataset gathered from Kaggle. once initial knowledge exploration, we have a tendency to knew we might implement different types of machine learning algorithm for best accuracy.

We used UI based **streamlit** web application for predicting the loan eligibility which will be very useful for the user experiencence and they can handle it easily.

The different steps that are involved in creating our machine learning Loan prediction project are, Broadly, the entire machine learning lifecycle can be described as a combination of 6 stages.

Stage 1: Problem Definition

The first and most important part of any project is to define the problem statement. Here, we want to describe the aim or the goal of our project and what we want to achieve at the end.

Stage 2: Hypothesis Generation

Once the problem statement is finalized, we move on to the hypothesis generation part. Here, we try to point out the factors/features that can help us to solve the problem at hand.

Stage 3: Data Collection

After generating hypotheses, we get the list of features that are useful for a problem. Next, we collect the dataaccordingly. This data can be collected from different sources.

Stage 4: Data Exploration and Pre-processing

After collecting the data, we move on to explore and pre-process it. These steps help us to generate meaningful insights from the data. We also clean the dataset in this step, before building the model

Stage 5: Model Building

Once we have explored and pre-processed the dataset, the next step is to build the model. Here, we create predictive models in order to build a solution for the project.

Stage 6: Model Deployment

Once you have the solution, you want to showcase it and make it accessible for others. And hence, the final stage of the machine learning lifecycle is to deploy that model.
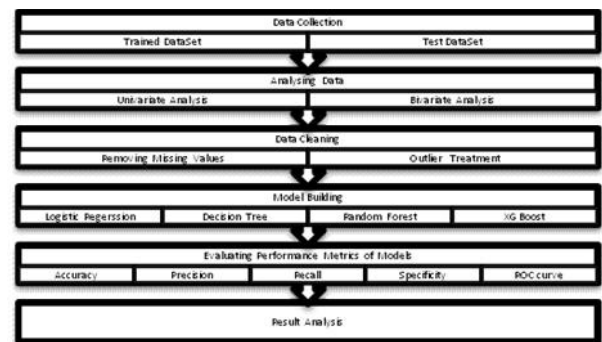
Chart: System architecture

III. CONCLUSIONS

This research explores using machine learning algorithms to improve the accuracy of predicting loan default. This model will be instrumental to mobile lending institutions in evaluating their customer credit risk. The best performing model in the research which is LOGISTIC REGRESSION achieves an accuracy of about 80%. This is a fair performance and can further be improved through different methods of

parameter tuning and feature selection which may possibly yield improvements in the model performance. It may also be beneficial to do a cross validation with other sources of open dataset as they become more accessible to compare the performance of the model. Since the research is also limited to the probability of default in a default state , further exploration may be made in determining the expected return of the loan based on borrower's characteristics , loan characteristics the recollection of loans process.

REFERENCES

[1] Amira Kamil Ibrahim Hassan, Ajith Abraham, "Modeling Consumer Loan Default Prediction Using Ensemble Neural Networks", Aug. 2013.

[2] D. Fantazzini and S. Figini, "Random Survival Forests Models For SME Credit Risk Measurement," Methodology and Computing in Applied Probability, vol.11,no.1,2009.

[3] Gwer, F., Odero, J., & Totolo, E. (2020). Digital credit audit report: Evaluatingthe conduct and practice of digital lending in Kenya – Financial Sector Deepening Kenya. Retrieved 18 October 2020, from https://fsdkenya.org/publication/digital-credit-audit-reportevaluating-the-conduct-and-practice-of-digital-lending-in-kenya/

[4] Credit scoring approaches guidelines - World Bank. (n.d.). Retrieved October 19,2020, from http://pubdocs.worldbank.org/en/935891585869698451/C REDIT-SCORING APPROACHES - GUIDELINES-FINAL-WEB.pdf

[5] Madaan, M., Kumar, A., Keshri, C., Jain, R., & Nagrath, P. (2021). Loan default prediction using decision trees and random forest : A comparative study. IOP Conference Series: Materials Science And Engineering, 1022, 012042. HTTPS://10.1088/1757-899x/1022/1/012042