PREDICTING PERSONAL LOAN APPROVAL USING MECHINE LEARNING

Submitted in partial fulfillment of requirement for the award of the Degree

Bachelor of Computer Science

In the faculty of Computer Science of Bharathiar University, Coimbatore

Submitted by

TEAM ID: NM2023TMID17454

S.DIVIYA

S.K.DHANASHREE

K.VASTHAMANI

S.SWETHA

Under the guidance of

Dr.V.PUNITHAVATHI MCA,M.Phil,DOA.,Ph.D.

Lecturer, Department of Computer Science

DEPARTMENT OF COMPUTER SCIENCE

L.R.G GOVERNMENT ARTS COLLEGE FOR WOMEN

(Affiliated ToBharathiar University)

TIRUPUR-4

APRIL-2023

LRG GOVERNMENT ARTS COLLEGE FOR WOMEN

NAAN MUDHALVAN PROJECT WORK

(AFFILIATED TO BHARATHIAR UNIVERSITY)

TIRUPUR-641602

TITLE: PREDICTING PERSONAL LOAN APPROVAL USING MACHINE LEARNING

This is to certify that this is a bonafide record of work done by the above students of III B.Sc (CS) Degree **NAAN MUDHALVAN PROJECT** during the year 2022-2023

Submitted for the Naan Mudhalvan project work held on April 2023

CLASS MENTOR **HEAD OF DEPARTMENT**

[Dr.V.Punithavathi] **[Dr. R. Parimala]**

TABLE OF THE CONTENT

1.INTRODUCTION

LOANS are the major requirement of the modern world. By this only, Banks get a major part of the total profit. It is beneficial for students to manage their education and living expenses, and for people to buy any kind of luxury like houses, cars, etc.

But when it comes to deciding whether the applicant's profile is relevant to be granted with loan or not. Banks have to look after many aspects.

So, here we will be using Machine Learning with Python to ease their work and predict whether the candidate's profile is relevant or not using key features like Marital Status, Education, Applicant Income, Credit History, etc.

Loan prediction involves the lender looking at various background information about the applicant and deciding whether the bank should grant the loan. Parameters like credit score, loan amount, lifestyle, career, and assets are the deciding factors in getting the loan approved.

It is done by predicting if the loan can be given to that person on the basis of various parameters like credit score, income, age, marital status, gender, etc. The prediction model not only helps the applicant but also helps the bank by minimizing the risk and reducing the number of defaulters.

2. PROBLEM DEFINITION

1. Credit Score: A borrower's credit score is one of the most critical factors in determining loan approval. It is a numerical representation of a borrower's creditworthiness, based on their credit history and financial behavior.

2.Income: Lenders consider a borrower's income to assess their ability to repay the loan. Higher income means that the borrower is more likely to repay the loan on time.

3.Debt-to-Income Ratio: This is the ratio of a borrower's total debt to their income. It helps lenders determine whether the borrower can afford to repay the loan.

4.Employment Status: Lenders prefer borrowers who have a stable job and a consistent income stream. Therefore, a borrower's employment status is an important factor in loan approval.

5.Loan Amount: The amount of the loan requested is an essential factor in loan

approval. Lenders are more likely to approve smaller loan amounts than larger ones.

6.Loan Purpose: The purpose of the loan can also affect the likelihood of loan approval. For example, a borrower who is taking a loan to start a business may be considered riskier than one who is taking a loan to pay off credit card debt.

7.Collateral: Collateral is an asset that a borrower pledges as security for the loan. Lenders may require collateral to reduce the risk of default. The type and value of the collateral can affect loan approval.

3.IDEATION

1. **Logistic Regression Model:** One of the most commonly used models for predicting loan approval is logistic regression. It uses various factors such as credit score, income, employment status, and loan amount to predict whether a borrower will be approved for a loan or not.

2. **Decision Tree Model:** Another model that could be used for predicting loan approval is a decision tree. Decision trees are tree-like structures that use a series of if-then statements to predict loan approval. For example, if the borrower's credit score is above a certain threshold, they may be more likely to be approved for a loan.

3. **Random Forest Model**: Random forest is an ensemble model that uses multiple decision trees to make predictions. It combines the predictions of each tree to generate a final prediction. Random forest models can be useful for predicting loan approval because they can handle multiple input variables and non-linear relationships.

4. **Gradient Boosting Model:** Gradient boosting is another ensemble model that combines multiple weak models to create a stronger one. It is particularly useful for predicting loan approval because it can handle missing data and outliers. Gradient boosting models can be trained using various features such as credit score, income, employment status, loan amount, and loan purpose.

5. **Neural Network Model:** Neural networks are a type of deep learning model that can be used for predicting loan approval. They are particularly useful for handling complex data with non-linear relationships. Neural networks can be trained using various features such as credit score, income, employment status, loan amount, and loan purpose.

4.REQUIREMENT ANANLYSIS

Understand the problem: It is important to understand the business problem that needs to be solved. In this case, the problem is to predict whether a personal loan application will be approved or not based on various factors.

Collect data: Collect relevant data from various sources. This data can include information about the applicant such as their age, income, employment status, credit score, and other financial information. Additionally, information about the loan itself such as the amount requested, the loan purpose, and the loan term may also be relevant.

Clean and preprocess data: The collected data may contain errors or inconsistencies that need to be cleaned before analysis. Additionally, data may need to be preprocessed to prepare it for modeling. This can include tasks such as removing missing values, scaling features, and encoding categorical variables.

Identify key features: Identify the most important features that are likely to affect the loan approval decision. This can be done through exploratory data analysis or using domain knowledge. Examples of key features might include credit score, income level, and employment status.

Build a predictive model: Select an appropriate machine learning algorithm to build a predictive model. This can be done using techniques such as logistic regression, decision trees, or random forests. The model should be trained on a portion of the available data and evaluated on a separate validation set to ensure accuracy.

Test and refine the model: Test the model on new data and refine it as necessary. This can involve tweaking hyperparameters, selecting different features, or using a different algorithm altogether.

Deploy the model: Once the model has been refined and tested, it can be deployed into production for use in predicting loan approvals.

5.INPUT DESIGN

The Input design is the process of entering data to the system. The input design goals is to enter to the computer as accurate as possible .In this proposed work input are designed effectively so that errors made by operations are minimized.

The input to the system have been designed in such a way that manual forms and the input are coordinated where the data elements are common to the sources document and to the input.

Input design is the process of converting user organized input to a computer-based format input data are collected and organized into group of similar data once identified appropriate input media are selected for processing.

Input design means the physical and performance requirements of a device that are used as a basis for device design. Input is the raw data that is processed to produced output

5.1 INPUT DESIGN DESCRIPTION

HOME

In the home page is simple concept of prediction personal loan approval is displayed an it is show a predict option to get the predict page.

PREDICT

In this page it asks for input certain data that is necessary to predict to the prediction personal loan approval type is also shows a submit button which redirects it to the submit page.

6.OUTPUT DESIGN

A design output is a drawing or specification or manufacturing instruction.

Design outputs and describe all the components,parts, and pieces that go into your device. design outputs describe all assembles and subassemblies of product.

Output is the process of converting data into hardcopy that is understood by all.

The varies outputs have been in such a way that they represent the same format that the office and management used to.

6.1OUTPUT DESIGN DESCRIPTION

SUBMIT

In this page after enter necessary data in predict page then click the submit button.

It displays the result of the data entered like (personal loan, business loan, educational, house)

7.DESCRIPTION OF MODULES

Modules are unit of code written in accesses in basic language.

- HOME
- PREDICT
- SUBMIT

HOME

In the home page is simple concept of prediction personal loan approval is displayed an it is show a predict option to get the predict page.

PREDICT

In this page its asks for input certain data that is necessary to predict to the prediction personal loan approval type its also show a submit button which redirects it to the submit page.

SUBMIT

In this page after enter necessary data in predict page then click the submit button.

It displays the result of the data entered like (personal loan, business loan, educational)

8.PROJECT PLANNING PHASE

1.Define project scope: Clearly define the scope of the project, including the problem statement, data sources, and expected outcomes.

2.Identify project resources: Identify the resources required to complete the project, including personnel, hardware, and software.

3.Define project timeline: Create a timeline that outlines the major milestones and deadlines for the project, including data collection, model development, testing, and deployment.

4.Define project deliverables: Define the deliverables of the project, including the final model, documentation, and any other necessary artifacts.

5.Identify potential risks: Identify potential risks that may impact the project, including technical issues, data quality issues, and personnel issues.

6.Define project budget: Define the budget required to complete the project, including any expenses related to data collection, software, and hardware.

7.Define project management approach: Define the project management approach that will be used to manage the project, including communication channels, project meetings, and project documentation.

8.Identify project stakeholders: Identify the stakeholders involved in the project, including project sponsors, project team members, and end-users of the final product.

9.Define project success criteria: Define the criteria that will be used to determine the success of the project, including model accuracy, usability, and business value.

9. PROJECT DESIGNING

Define the problem statement: The problem statement should define the objective of the project, such as "To predict whether a personal loan application will be approved or not based on various factors".

Gather and preprocess the data: Collect the relevant data from various sources and preprocess it to ensure it is clean and ready for modeling. This can include tasks such as removing missing values, scaling features, and encoding categorical variables.

Identify key features: Identify the most important features that are likely to affect the loan approval decision. This can be done through exploratory data analysis or using domain knowledge. Examples of key features might include credit score, income level, and employment status.

S**plit the data**: Split the data into training, validation, and test sets. The training set will be used to train the model, the validation set will be used to tune hyperparameters and evaluate model performance, and the test set will be used to evaluate the final model.

Choose a modeling approach: Choose an appropriate machine learning algorithm to build a predictive model. This can include techniques such as logistic regression, decision trees, or random forests.

Train the model: Train the model on the training set and evaluate its performance on the validation set. Tweak hyperparameters as necessary to improve performance.

Test and refine the model: Test the model on the test set and refine it as necessary. This can involve tweaking hyperparameters, selecting different features, or using a different algorithm altogether.

Deploy the model: Once the model has been refined and tested, it can be deployed into production for use in predicting loan approvals.

10.PROJECT DEVELOPMENT PHASE

1.Data Collection and Preprocessing: Collect the relevant data from various sources and preprocess it to ensure it is clean and ready for modeling. This can include tasks such as removing missing values, scaling features, and encoding categorical variables.

2.Feature Selection: Identify the most important features that are likely to affect the loan approval decision. This can be done through exploratory data analysis or using domain knowledge. Examples of key features might include credit score, income level, and employment status.

3.Model Selection and Training: Choose an appropriate machine learning algorithm to build a predictive model. This can include techniques such as logistic regression, decision trees, or random forests. Train the model on the training set and evaluate its performance on the validation set. Tweak hyperparameters as necessary to improve performance.

4.Model Evaluation and Improvement: Test the model on the test set and refine it as necessary. This can involve tweaking hyperparameters, selecting different features, or using a different algorithm altogether.

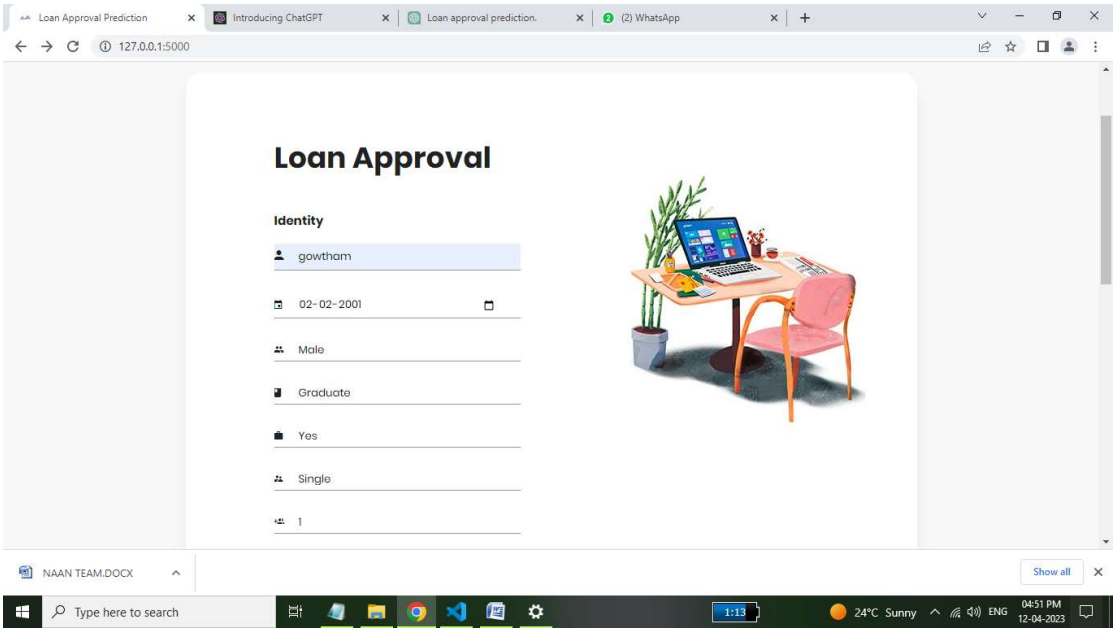
5.Model Deployment: Once the model has been refined and tested, it can be deployed into production for use in predicting loan approvals. This may involve integrating the model with other systems or applications.

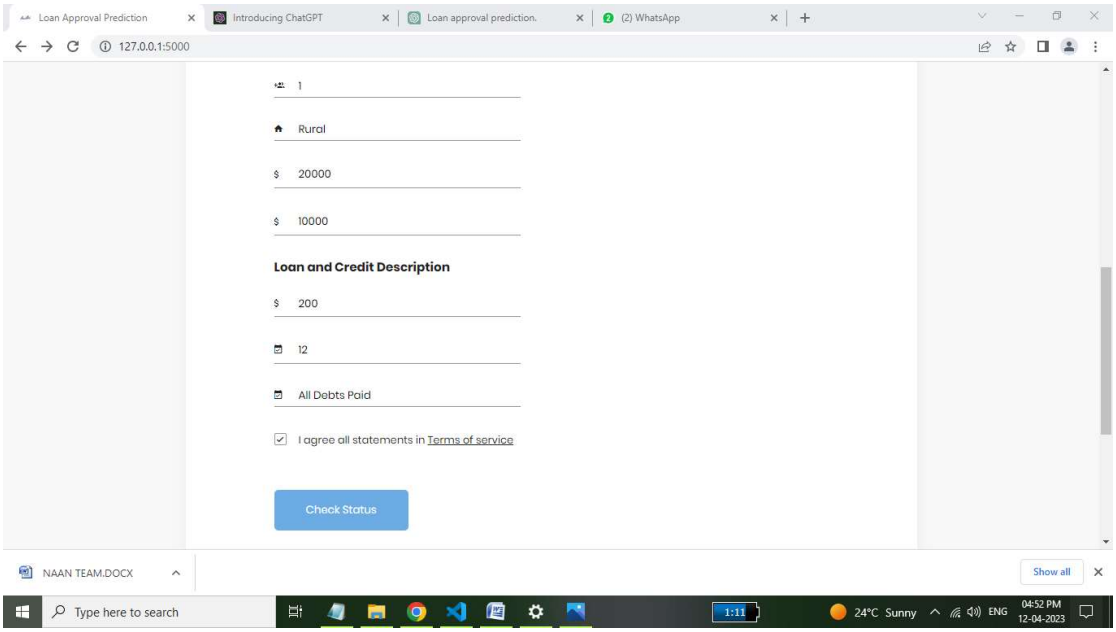
6.Monitoring and Updating: Monitor the performance of the model in production and update it as necessary to ensure continued accuracy and reliability.

11. CONCLUSION

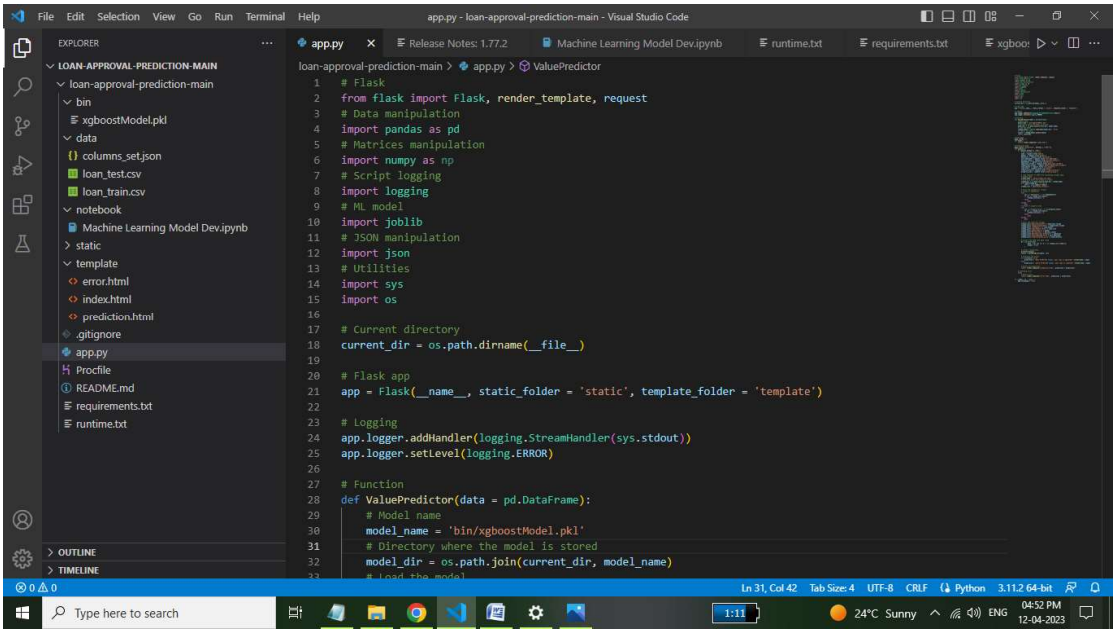
Machine Learning plays a very deciding role in loan prediction. Machine learning algorithm such as SVM, Random Forest Classifier, K-NN, logestic regression and DT Classifier are used to predict the risk of not getting appropriate information about getting loan. The web page is created to get data from users to predict the type of personal loan approval.

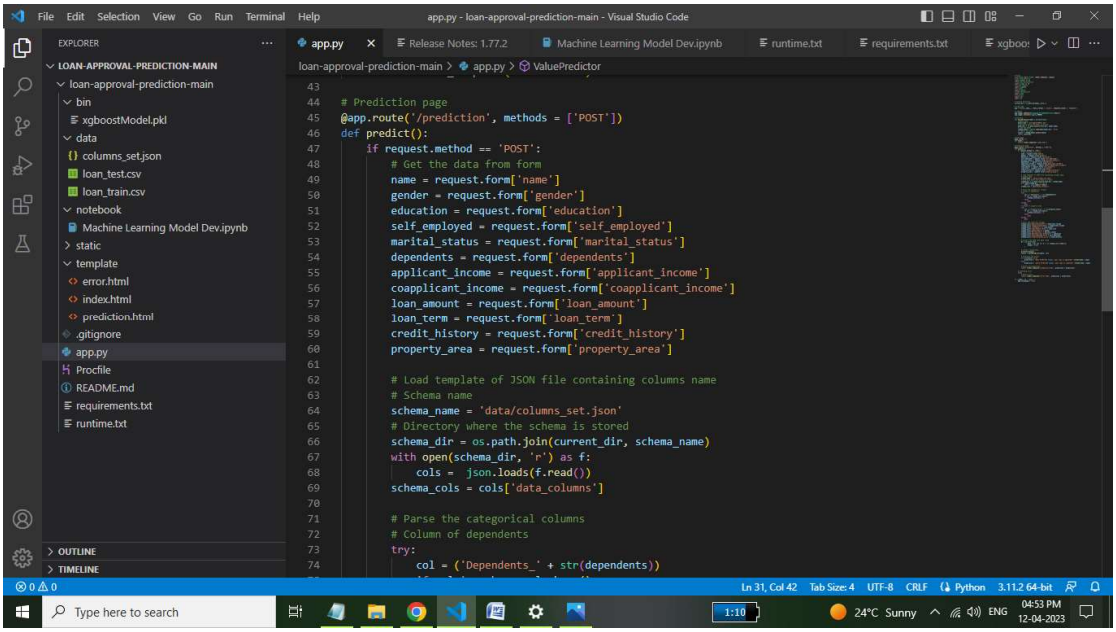
11.APPENDICES





SAMPLE CODING





Project Link:

Video link:

https://youtu.be/NztkmHcSHT0

Project link:

https://github.com/Dhanushyalrg/PREDICTING-PERSONAL-LOAN-APPROVAL-USING-MACHINE-LEARNING