1.INTRODUCTION

1.1 PROJECT OVERVIEW

A precise credit risk assessment system is always vital to any financial institution for impeccable and gainful functioning. In such an ever-changing economy as the rate of loan defaults are gradually increasing, authorities of financial institutions are finding it more and more difficult to correctly assess loan requests and tackle the risks of loan defaulters. In light of these events this paper proposes a machine learning model which can precisely assess credit risk and predict possible loan defaulters for credit lending institutions. A comparative analysis has been made using tuned supervised learning algorithms such as Support Vector Machine, Random Forest, Extreme Gradient Boosting and Logistic Regression for identifying defaulters. Recursive Feature Elimination with Cross-Validation and Principal Component Analysis have been used for dimensionality reduction. Metrics such as F1 score, AUC score, prediction accuracy, precision and recall have been used to evaluate each model. Among all the models, the combination of a tuned Support Vector Machine and Recursive Feature Elimination with Cross-Validation have shown great promise in identifying loan defaulters. The proposed model, therefore, can assist financial institutions in accurately identifying loan defaulters and prevent them from incurring further loss.

1.2 PURPOSE

Loans are the core business of banks. The main profit comes directly from the loan's interest. The loan companies grant a loan after an intensive process of verification and validation. However, they still don't have assurance if the applicant is able to repay the loan with no difficulties.

The Company wants to automate the loan eligibility process (real time) based on customer detail provided while filling online application form. These details are Gender, Marital Status, Education, Number of Dependents, Income, Loan Amount, Credit History and others. To automate this process, they have given a problem to identify the customers segments, those are eligible for loan amount so that they can specifically target these customers.

2. LITERATURE SURVEY

2.1 Existing Problem

Up until now many notable findings have been made which can make credit risk analysis much more efficient and precise but it is to be mentioned that in such ever-changing times in modern science there is always room for improvement. Now, the paper exhibits some brilliant work done in the early years of credit risk assessment using machine learning. The paper showcases a comparison of different models for credit risk evaluation where a multiagent based system created using an adaptive linear neural network achieved the best results with an accuracy of 71.19 percent. The paper proposes a multistage reliability-based neural network ensemble learning model for credit risk assessment. When using a credit card approval dataset of 653 instances the highest accuracy reached by the model is 88.08 percent.

2.2 References

[1] D. Faggella, “What is Machine Learning? | Emerj - Artificial Intelligence Research and Insight,” Emerj. [Online]. Available: https://emerj.com/ai-glossary-terms/what-is-machine-learning/. [Accessed: 15-Jan-2019].

[2] Press Trust of India, “9 million loan defaulters blacklisted in China; $27 billion frozen,” https://www.hindustantimes.com/, 13-Jan-2018. [Online]. Available: https://www.hindustantimes.com/world-news/9- million-loan-defaulters-blacklisted-in-china-27-billion-frozen/storyZ9KBvxffasUwsD8LRLjpPJ.html. [Accessed: 15-Jan-2019].

[3] G. Mowla, “Default loans plague banking sector,” Dhaka Tribune, 22-Mar-2018. [Online]. Available: https://www.dhakatribune.com/business/banks/2018/03/22/defaultloan-plagues-banking-sector. [Accessed: 15-Jan-2019].

[4] S. Islam, “Bad loans cripple the banking sector,” Dhaka Tribune, 08- Oct-2017. [Online]. Available:

https://www.dhakatribune.com/business/banks/2017/10/06/defaultedloans-amount-12-gdp. [Accessed: 15-Jan-2019].

[5] A. Nova, “More than 1 million people default on their student loans each year,” CNBC, 13-Aug-2018. [Online]. Available: https://www.cnbc.com/2018/08/13/twenty-two-percent-of-studentloan-borrowers-fall-into-default.html. [Accessed: 15-Jan-2019].

[6] A. Thakur, “India's wilful defaulters owe more than Rs 1 lakh crore to banks - Times of India,” The Times of India, 23-Feb-2018. [Online]. Available: https://timesofindia.indiatimes.com/business/indiabusiness/indias-wilful-defaulters-owe-more-than-rs-1-lakh-crore-tobanks/articleshow/63035851.cms. [Accessed: 15-Jan-2019].

2.3 Problem Statement Definition

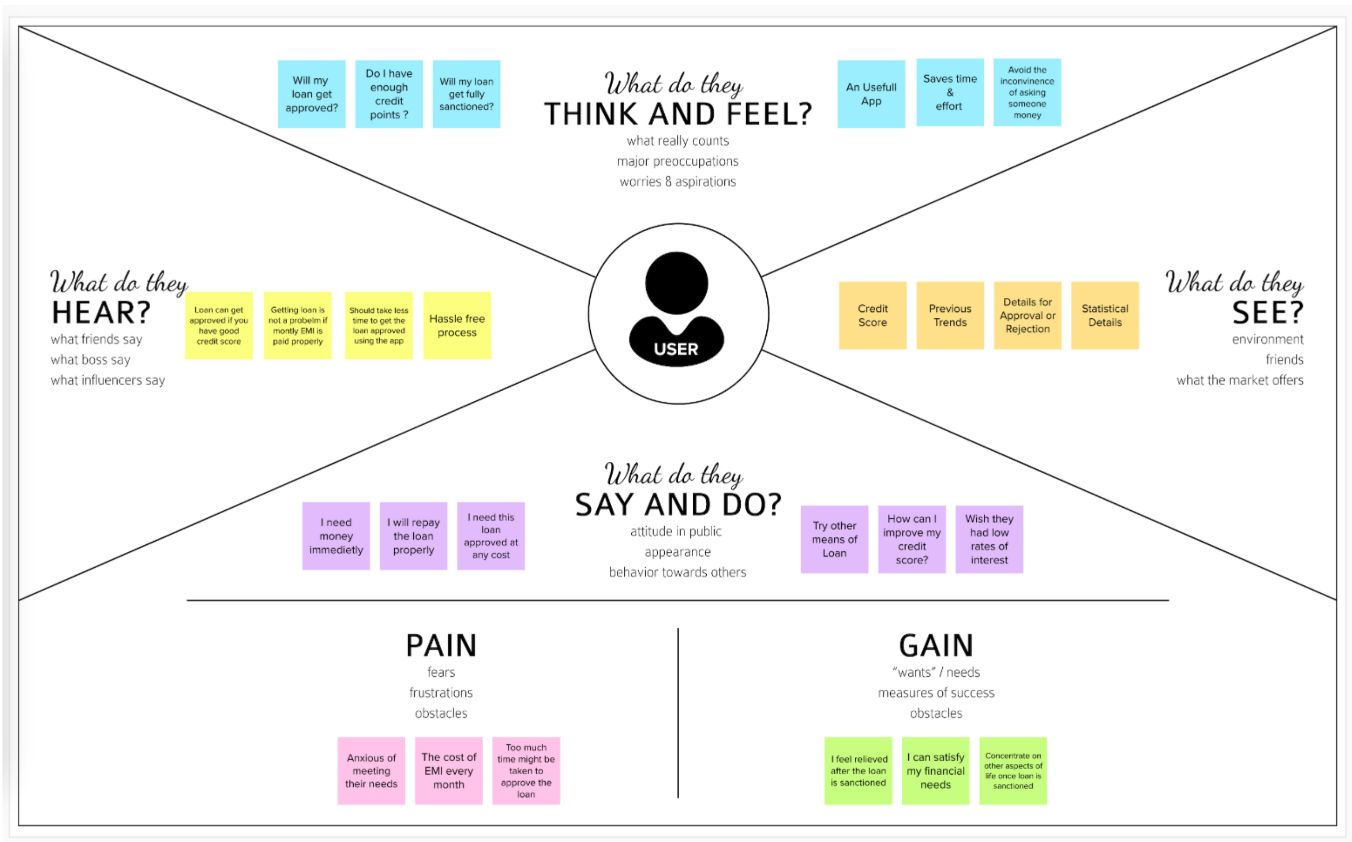
One of the most important factors which affect our country’s economy and financial condition is the credit system governed by the banks. The process of bank credit risk evaluation is recognized at banks across the globe. “As we know credit risk evaluation is very crucial, there is a variety of techniques are used for risk level calculation. In addition, credit risk is one of the main functions of the banking community.

The prediction of credit defaulters is one of the difficult tasks for any bank. But by forecasting the loan defaulters, the banks definitely may reduce their loss by reducing their non-profit assets, so that recovery of approved loans can take place without any loss and it can play as the contributing parameter of the bank statement. This makes the study of this loan approval prediction important. Machine Learning techniques are very crucial and useful in the prediction of these types of data.

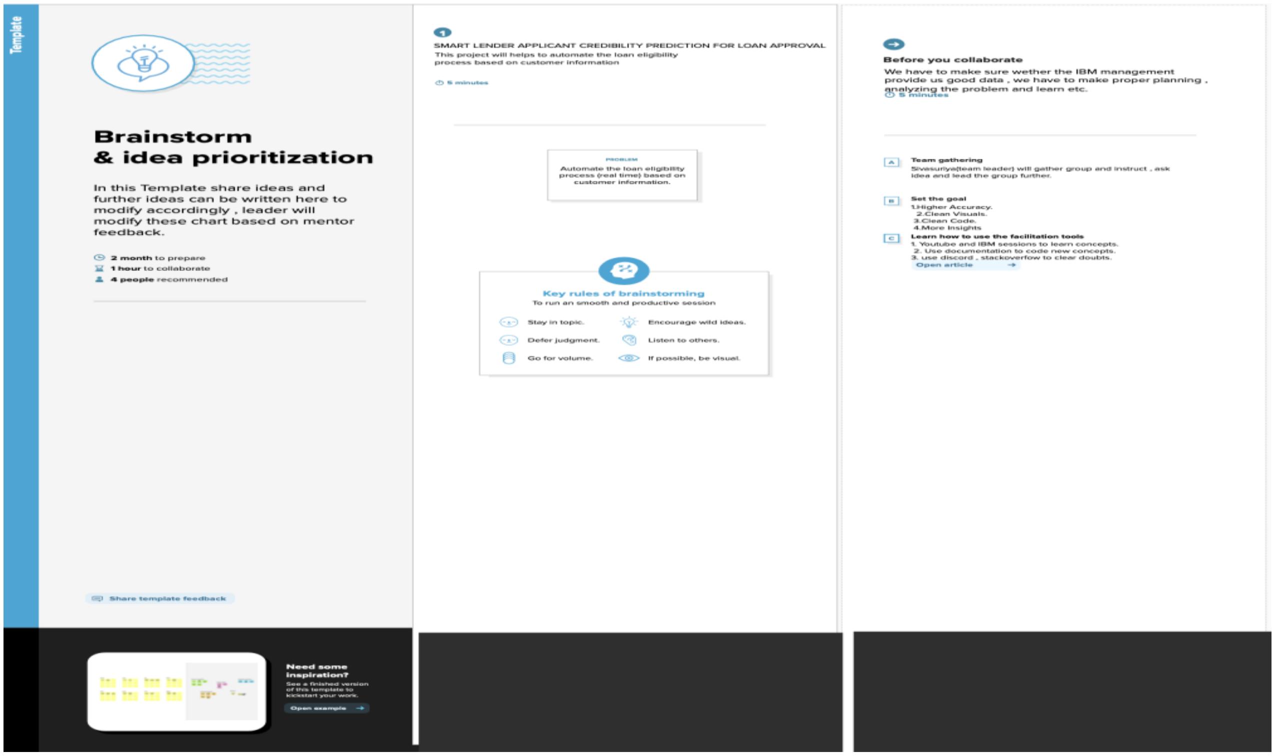
We will be using classification algorithms such as Decision tree, Random forest, KNN, and xgboost. We will train and test the data with these algorithms. From this best model is selected and saved in pkl format. We will be doing flask integration and IBM deployment.

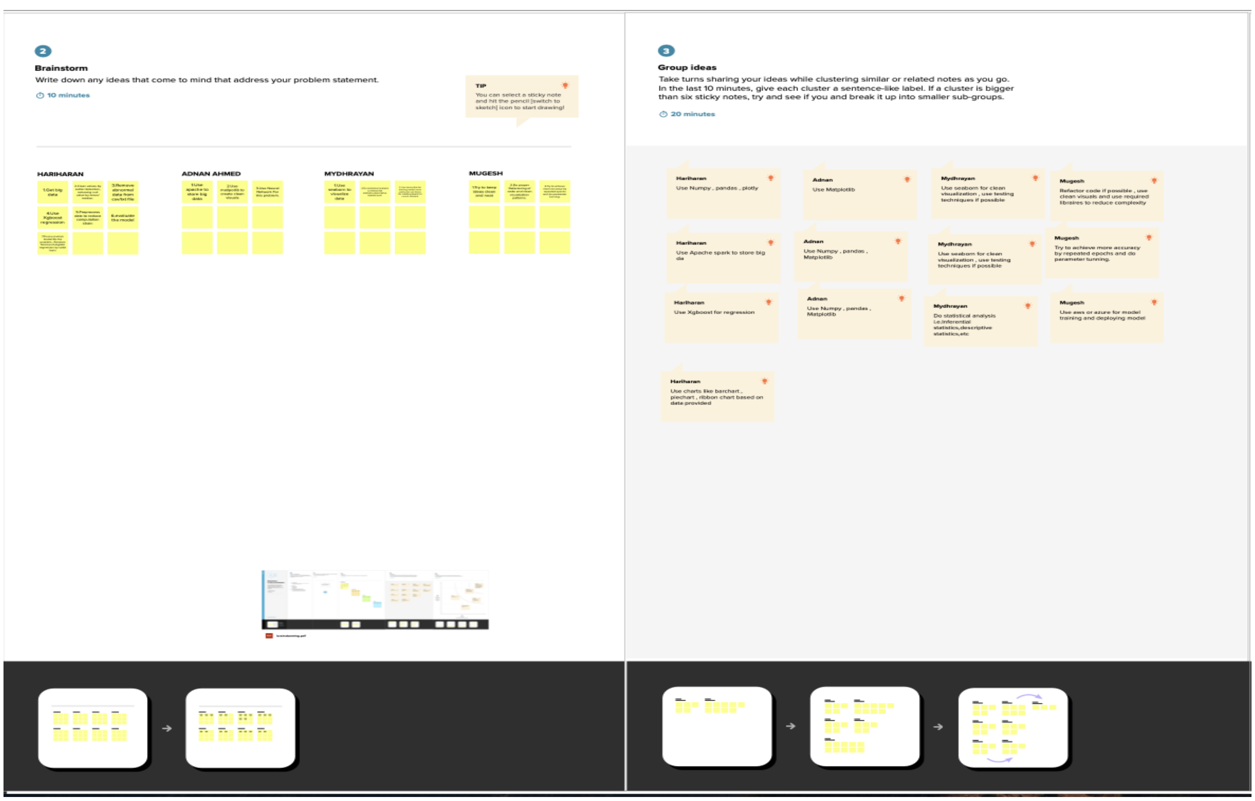
3.IDEATION AND PROPOSED SOLUTION

3.1 EMPATHY MAP CANVAS



3.2 IDEATION AND BRAINSTORMING



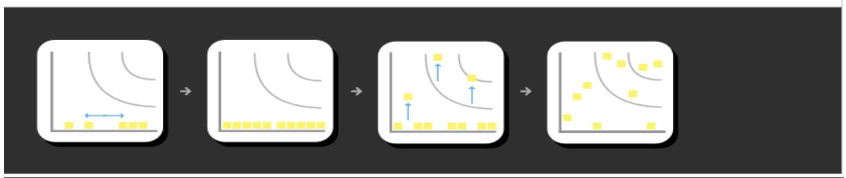


4

Prioritize

Your team should all be on the same page about what's important moving forward. Place your ideas on this grid to determine which ideas are important and which are feasible.

20 minutes



3.3 PROPOSED SOLUTION

Proposed Solution Template:

S.No.	**Parameter**	**Description**
1.	Problem Statement (Problem to be solved)	Banking is a crucial sector, it deals with financial transactions which can be availed by everyone, but banks are not able to resolve the queries of customers at all times related to the products or services insatisfactory way which in turn hinders the customer satisfaction. Dream Housing Finance company deals in all kinds of home loans
2.	Idea / Solution description	We are building a machine learning model that uses several data points such as credit history, duration, credit mix to predict whether a customer is eligible for loan
3.	Novelty / Uniqueness	The machine learning model uses several data points to make an accurate prediction of the credit eligibility of the person.
4.	Social Impact / Customer Satisfaction	Using credit score as a basis to judge a individuals loan taking capacity makes our country a credit based society and such a society has spending power
5.	Business Model (Revenue Model)	The AI based prediction model can be monetized by a subscription model that charges banks and other financial institutions a fee
6.	Scalability of the Solution	AI models can be easily scaled and software as a service(Saas). This software can be banking app

3.4 PROBLEM SOLUTION FIT

Problem-Solution Fit Canvas:

Project Title: SMART LENDER - APPLICANT CREDIBILITY PREDICTION **Project Design Phase-I - Solution Fit Template** **Team ID:** PNT2022TMID27812

Define CS, fit into CC

1. CUSTOMER SEGMENT(S) CS

Who is your customer?
i.e. working parents of 0-5 y.o. kids

Our customers are people who require a loan. Our application helps provide credit to customers in a safe way.

6. CUSTOMER CONSTRAINTS CC

What constraints prevent your customers from taking action or limit their choices of solutions? i.e. spending power, budget, no cash, network connection, available devices.

Availing credit from a financial institution is a difficult process requiring several documents. A smart prediction model makes it easy for them

5. AVAILABLE SOLUTIONS AS

Which solutions are available to the customers when they face the problem
or need to get the job done? What have they tried in the past? What pros & cons do these solutions have? i.e. pen and paper is an alternative to digital notetaking

Excessive documentation and work is required to provide loans. Also loans are given at the discretion of the branch manager thus a lucrative software helps ease this.

Explore AS, differentiate

Focus on J&P, tap into BE, understand RC

2. JOBS-TO-BE-DONE / PROBLEMS J&P

Which jobs-to-be-done (or problems) do you address for your customers? There could be more than one; explore different sides.

1. Easy loan provision using a smart software
2. Safety for the bank or financial institution to prevent NPA.

9. PROBLEM ROOT CAUSE RC

What is the real reason that this problem exists?
What is the back story behind the need to do this job?
i.e. customers have to do it because of the change in regulations.

Assessing and providing credit to customers is a difficult process involving several evaluations

7. BEHAVIOUR BE

What does your customer do to address the problem and get the job done?
i.e. directly related: find the right solar panel installer, calculate usage and benefits; indirectly associated: customers spend free time on volunteering work (i.e. Greenpeace)

The customer visits the bank and undergoes several processes to avail a loan. They use other contacts to see if they can get credit.

Focus on J&P, tap into BE, understand RC

Identify strong TR & EM

3. TRIGGERS TR

What triggers customers to act? i.e. seeing their neighbour installingsolar panels, reading about a more efficient solution in the news.

Customers act by visiting their bank or consulting anyone working ina financial institution

4. EMOTIONS: BEFORE / AFTER EM

How do customers feel when they face a problem or a job and afterwards?
i.e. lost, insecure > confident, in control - use it in your communication strategy & design.

Customers in need of debt feel very sad not getting credit and thus visit smaller financialinstitutions or pledge gold.

10. YOUR SOLUTION SL

If you are working on an existing business, write down your current solution first,fill in the canvas, and check how much it fits reality.
If you are working on a new business proposition, then keep it blank until you fill inthe canvas and come up with a solution that fits within customer limitations, solves a problem and matches customer behaviour.

Our solution uses a smart machine learning model which gives both financial institutions and customers safety and instant service. This model uses several parameters like credit score, previous banking history, on time payments etc tohelp the bank assess the customer.

The customer on the other hand gets instant approval based on their capabilities.

8.CHANNELS of BEHAVIOUR CH

1. ONLINE
What kind of actions do customers take online? Extract online channels from #7

2. OFFLINE
What kind of actions do customers take offline? Extract offline channels from #7and use them for customer development.

ONLINE
Customers enter their details to get their creditworthiness which tells them theireligibility to get a loan.

OFFLINE

No offline work is requires.

Identify strong TR & EM

4.REQUIREMENT ANALYSIS

4.1 FUNCTIONAL REQUIREMENT

Functional Requirements:

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	User Registration	Registration through Form Registration through Gmail Registration through LinkedIN
FR-2	User Confirmation	Confirmation via Email Confirmation via OTP
FR-3	User credit score check	Confirmation via our proprietary software which evaluates.
FR-4	User enters loan details	Validated by bank or financial institution.
FR-5	Fund transfer By the bank to customer	Payment sent through systems such as NEFT or IMPS.

4.2 NON-FUNCTIONAL REQUIREMENTS

Non-functional Requirements:

Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	**Usability**	The application will be easy to use with a neat and clean interface.
NFR-2	**Security**	Safe encryption of data is done to ensure customer data safety.
NFR-3	**Reliability**	The machine learning model provides an accurate credit check.
NFR-4	**Performance**	Sleek and higher order functions ensure fast running and also low time complexity.
NFR-5	**Availability**	All banks, financial institutions and customers will be able to use the application.
NFR-6	**Scalability**	The application is very scalable and runs across operating systems.

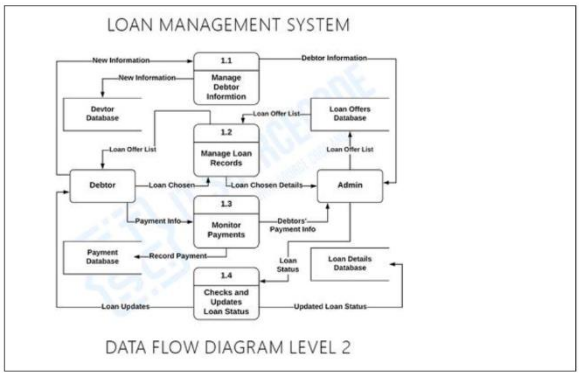
5. PROJECT DESIGN

5.1 DATA FLOW DIAGRAM

Data Flow Diagrams:

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It shows how data enters and leaves the system, what changes the information, and where data is stored.

Smart Lender - Applicant Credibility Prediction for Loan Approval



5.2 SOLUTION AND TECHNICAL ARCHITECTURE

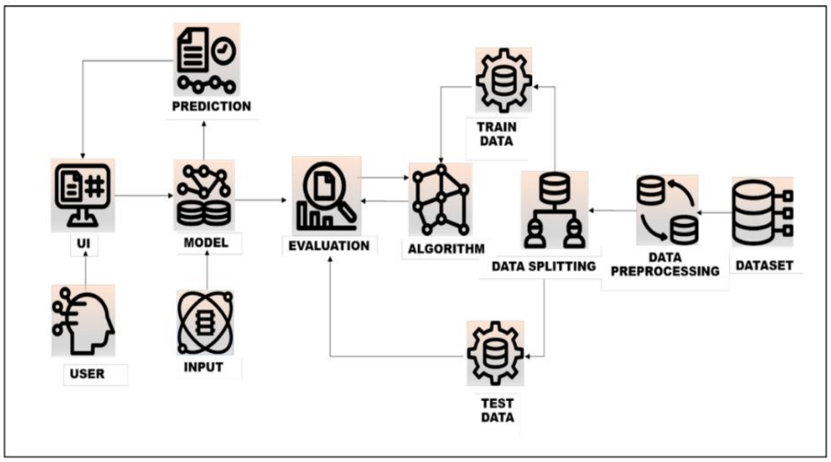
SOLUTION ARCHITECTURE

One of the most important factors which affect our country's economy and financial condition is the credit system governed by the banks. The process of bank credit risk evaluation is recognized at banks across the globe. "As we know credit risk evaluation is very crucial, there is a variety of techniques are used for risk level calculation. In addition, credit risk is one of the main functions of the banking community. The prediction of credit defaulters is one of the difficult tasks for any bank. But by forecasting the loan defaulters, the banks definitely may reduce their loss by reducing their non-profit assets, so that recovery of approved loans can take place without any loss and it can

play as the contributing parameter of the bank statement. This makes the study of this loan approval prediction important. Machine Learning techniques are very crucial and useful in the prediction of these types of data.

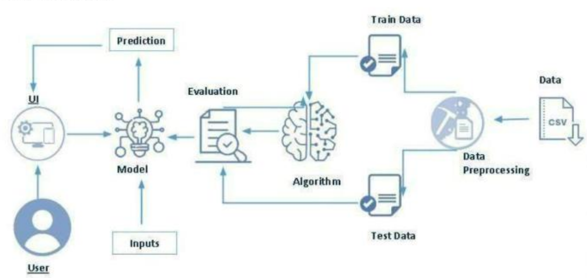
We will be using classification algorithms such as Decision tree, Random forest, KNN, and XG-boost. We will train and test the data with these algorithms.

SOLUTION ARCHITECTURE



TECHNICAL ARCHITECTURE

Technical Architecture :



5.3 USER STORIES

User Stories

Use the below template to list all the user stories for the product.

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	I can access my account / dashboard	High
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can receive confirmation email & click confirm	High
		USN-3	As a user, I can register for the application through Facebook	I can register & access the dashboard with Facebook Login	Low
		USN-4	As a user, I can register for the application through Gmail		Medium
	Login	USN-5	As a user, I can log into the application by entering email & password		High
	Dashboard	USN-6	As a user, credit score can be checked	I can check my score by providing PAN	High
Customer (Web user)		USN-7	As a user, I can interact with financial institutions		High
Customer Care Executive		USN-8	As a user, I can solve my queries by interacting with helpline		Medium
Administrator		USN-9	As a user, I can avail loans and credit from financial institutions	Credit score	High

6. PROJECT PLANNING AND SCHEDULING

6.1 SPRINT PLANNING AND ESTIMATION

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-1	Dataset	USN-4	Downloading the dataset	1	High	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-1		USN-5	Visualizing the dataset	2	Low	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-1		USN-6	Pre-process the dataset	3	Medium	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-1	Machine Learning Model	USN-7	KNN model building	5	High	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-2		USN-8	Decision Tree model building	5	High	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M

Sprint	Functional Requirement (Epic)	User Story Number	User Story / Task	Story Points	Priority	Team Members
Sprint-2		USN-9	Naive Bayes model building	5	High	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-2		USN-10	Fine Tuning the model	3	Low	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-2		USN-11	Evaluation and saving of the models	5	High	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-3	Customer User Interface	USN-12	Model Integration with flask	5	High	Hariharan A G Mydhrayan S G Mugeshraja M Adnan Ahmed S
Sprint-3		USN-1	As a user, I should be able to access the dashboard.	3	Medium	Hariharan A G Mydhrayan S G Mugeshraja M Adnan Ahmed S
Sprint-3		USN-2	Select the type of loan	3	Low	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-3		USN-3	Fill the application and check the eligibility of loan approval	5	High	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-4	Deployed the website	USN-13	Register on IBM Cloud	3	Low	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M
Sprint-4		USN-14	Train the ML model on IBM Cloud	5	Medium	Hariharan A G Adnan Ahmed S Mydhrayan S G Mugeshraja M

6.2 SPRINT DELIVERY SCHEDULE

Sprint	Total Story Points	Duration	Sprint Start Date	Sprint End Date (Planned)	Story Points Completed (as on Planned End Date)	Sprint Release Date (Actual)
Sprint-1	11	6 Days	24 Oct 2022	29 Oct 2022	11	29 Oct 2022
Sprint-2	18	6 Days	31 Oct 2022	05 Nov 2022	18	05 Nov 2022
Sprint-3	16	6 Days	07 Nov 2022	12 Nov 2022	16	12 Nov 2022
Sprint-4	16	6 Days	14 Nov 2022	19 Nov 2022	16	19 Nov 2022

6.3 REPORTS FROM JIRA

Velocity:
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

$$AV = \frac{sprint\ duration}{velocity} = \frac{20}{10} = 2$$

Our Project velocity

Sprint-1 = 11/6 = 1.833

Sprint-2 = 18/6 = 3

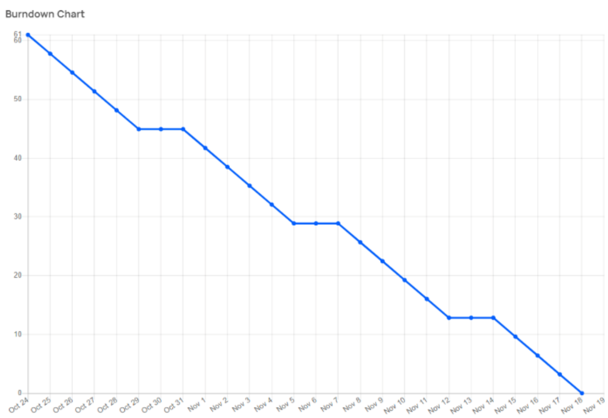
Sprint-3 = 16/6 = 2.67

Sprint-4 = 16/6 = 2.67

Total Velocity = 61/24 = 2.54

Burndown Chart:

A burn down chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.

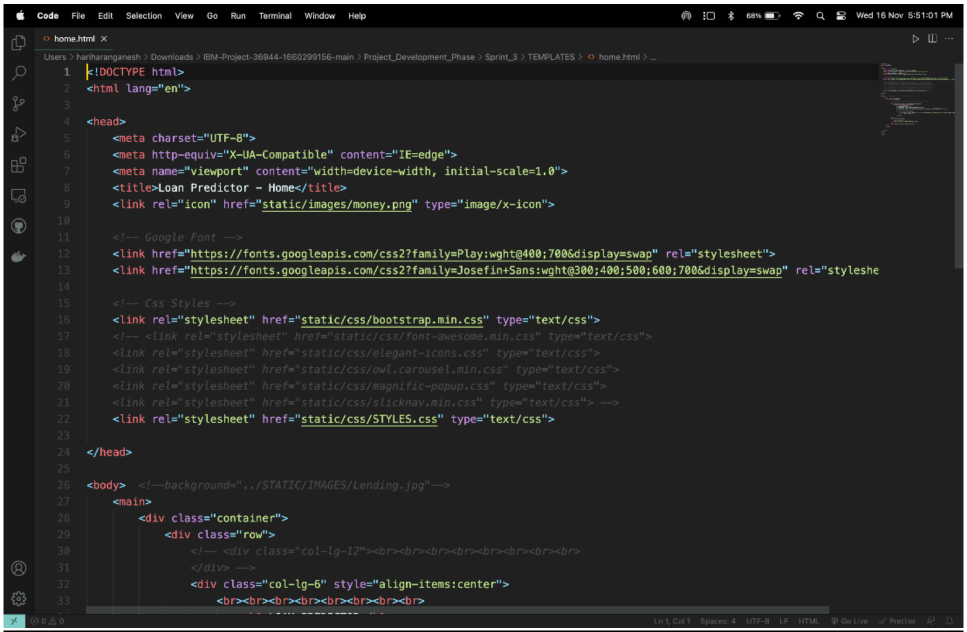


7. CODING AND SOLUTIONING

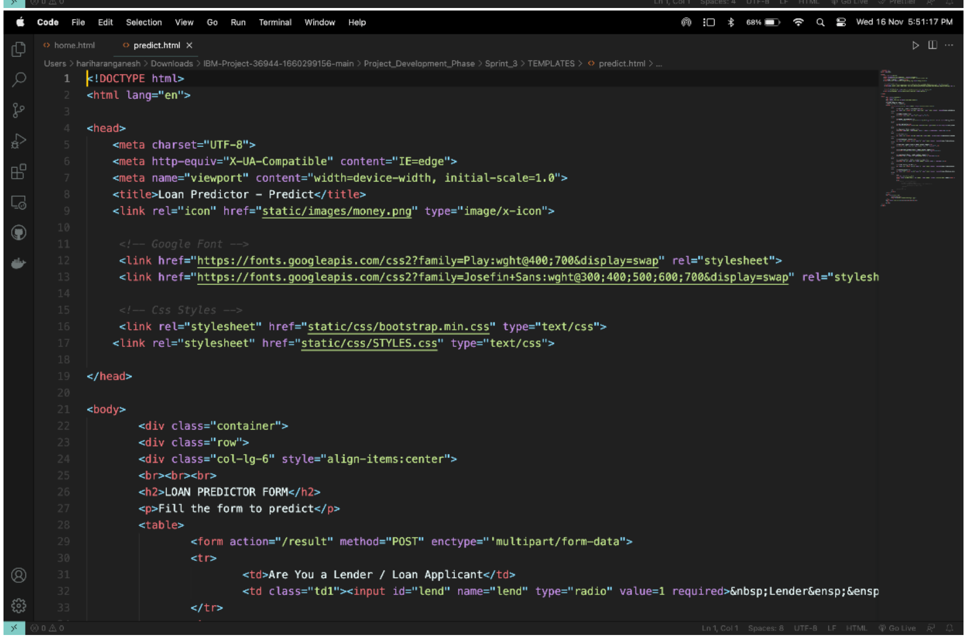
7.1 FORM FILLING BT APPLICANT TO CHECK HIS CREDIT ELIGIBILITY

```python
import numpy as np
import os
from PIL import Image
from flask import Flask, request, render_template, url_for
from werkzeug.utils import secure_filename, redirect
from gevent.pywsgi import WSGIServer
from flask import send_from_directory
from joblib import Parallel,delayed
import joblib
import pandas as pd
from scipy.sparse import issparse


app = Flask(__name__)


@app.route('/')
def index():
    return render_template('home.html')

@app.route('/predict')
def predict():
    return render_template('predict.html')

@app.route('/result', methods=['GET', 'POST'])
def upload():
    if request.method == "POST":
        lend_data = request.form.get('lend')
        data = [[request.form.get('gender'),request.form.get('married'),request.form.get('dep'),request.form.get('edu'),request.form.get('se'),request.form.get('ai')
                ,request.form.get('cai'),request.form.get('la'),request.form.get('lat'),request.form.get('ch'),request.form.get('pa')]]

        df = pd.DataFrame(data, columns=['Gender','Married','Dependents','Education',
                                         'Self_Employed','ApplicantIncome','CoapplicantIncome',
                                         'LoanAmount','Loan_Amount_Term','Credit_History','Property_Area'])

        gh=joblib.load('/Users/hariharanganesh/Downloads/IBM/PROJECT/rdf.pkl')
        num=gh.predict(df)
        a=''
        lend_data=int(lend_data)
        if(num==0):
            if(lend_data==1):
                a='It is not advisable to provide loan for this applicant.'
            else:
                a='Your Loan application will be Rejected.'
        else:
            if(lend_data==1):
                a='This applicant can be provided with the loan amount requested.'
            else:
                a='Your Loan application will be succesfull.'
        return render_template('submit.html', num=a)


if __name__ == '__main__':
    app.run(debug=True, threaded=False)
```

Ln: 1 Col: 0




```html
<!DOCTYPE html>
<html lang="en">

<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Loan Predictor - Home</title>
    <link rel="icon" href="static/images/money.png" type="image/x-icon">

    <!-- Google Font -->
    <link href="https://fonts.googleapis.com/css2?family=Play:wght@400;700&display=swap" rel="stylesheet">
    <link href="https://fonts.googleapis.com/css2?family=Josefin+Sans:wght@300;400;500;600;700&display=swap" rel="styleshe

    <!-- Css Styles -->
    <link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">
    <!-- <link rel="stylesheet" href="static/css/font-awesome.min.css" type="text/css">
    <link rel="stylesheet" href="static/css/elegant-icons.css" type="text/css">
    <link rel="stylesheet" href="static/css/owl.carousel.min.css" type="text/css">
    <link rel="stylesheet" href="static/css/magnific-popup.css" type="text/css">
    <link rel="stylesheet" href="static/css/slicknav.min.css" type="text/css"> -->
    <link rel="stylesheet" href="static/css/STYLES.css" type="text/css">

</head>

<body> <!--background="../STATIC/IMAGES/Lending.jpg"-->
    <main>
        <div class="container">
            <div class="row">
                <!-- <div class="col-lg-12"><br><br><br><br><br><br><br><br>
                </div> -->
                <div class="col-lg-6" style="align-items:center">
                    <br><br><br><br><br><br><br><br>
```

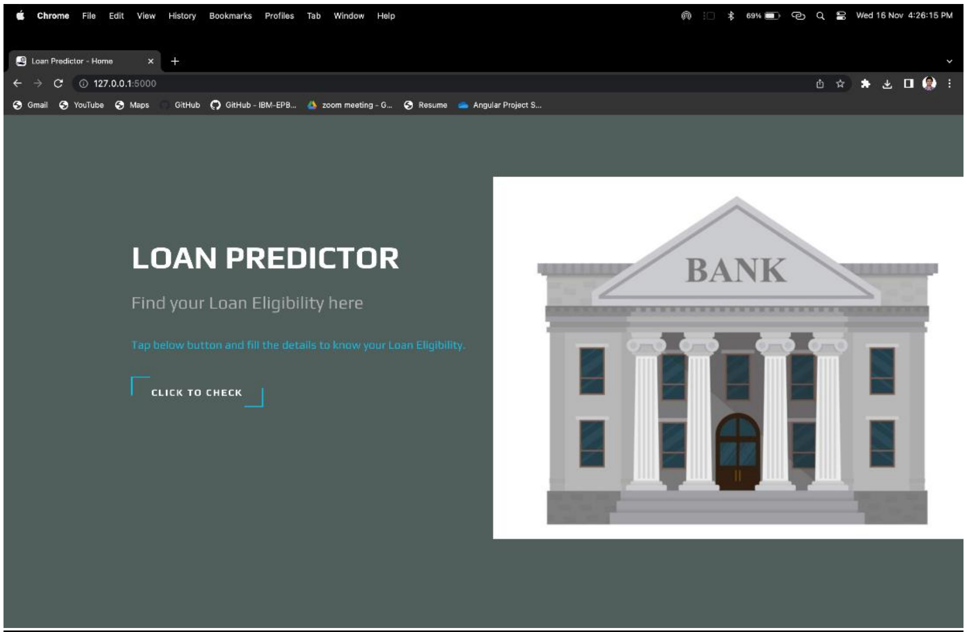


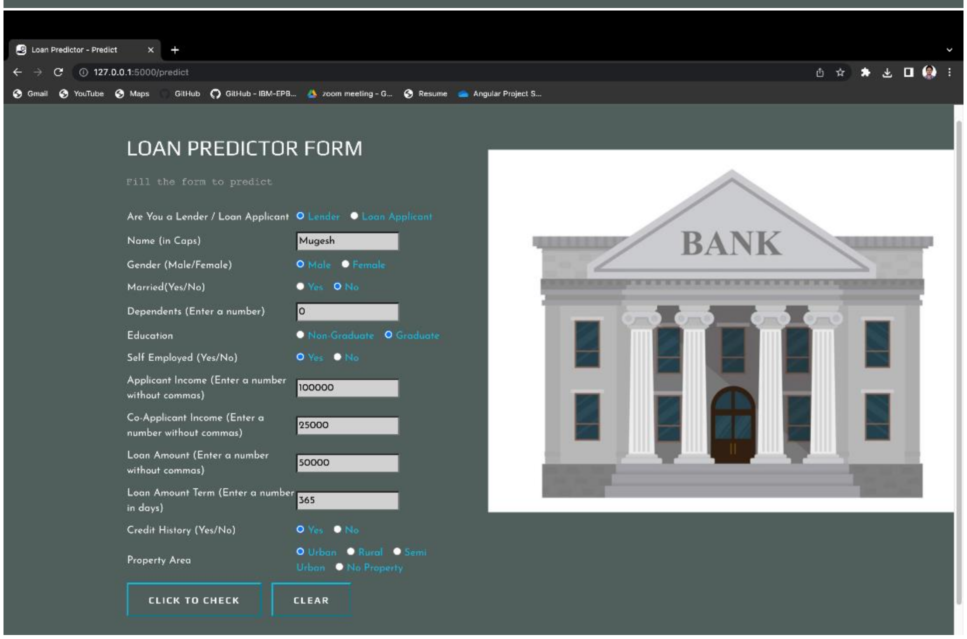

```html
<!DOCTYPE html>
<html lang="en">

<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Loan Predictor - Predict</title>
    <link rel="icon" href="static/images/money.png" type="image/x-icon">

     <!-- Google Font -->
     <link href="https://fonts.googleapis.com/css2?family=Play:wght@400;700&display=swap" rel="stylesheet">
     <link href="https://fonts.googleapis.com/css2?family=Josefin+Sans:wght@300;400;500;600;700&display=swap" rel="stylesh

     <!-- Css Styles -->
     <link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">
    <link rel="stylesheet" href="static/css/STYLES.css" type="text/css">

</head>

<body>
        <div class="container">
        <div class="row">
        <div class="col-lg-6" style="align-items:center">
        <br><br><br>
        <h2>LOAN PREDICTOR FORM</h2>
        <p>Fill the form to predict</p>
        <table>
                <form action="/result" method="POST" enctype="'multipart/form-data">
                <tr>
                        <td>Are You a Lender / Loan Applicant</td>
                        <td class="td1"><input id="lend" name="lend" type="radio" value=1 required> Lender  
                </tr>
```

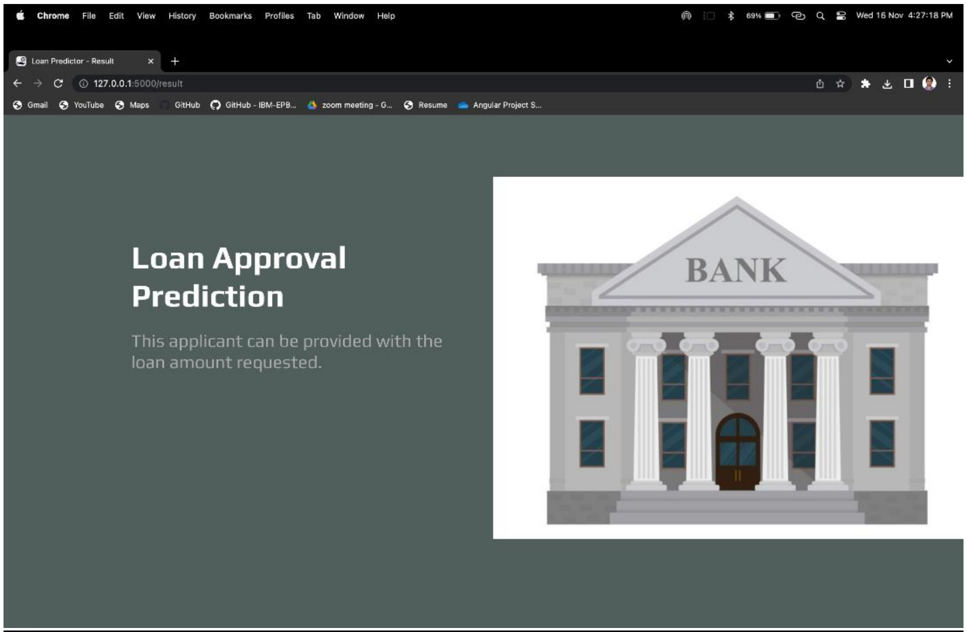
```html
<!DOCTYPE html>
<html lang="en">

<head>
    <meta charset="UTF-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1.0">
    <title>Loan Predictor - Result</title>
    <link rel="icon" href="static/images/money.png" type="image/x-icon" style="border-radius: 5px;">

     <!-- Google Font -->
     <link href="https://fonts.googleapis.com/css2?family=Play:wght@400;700&display=swap" rel="stylesheet">
     <link href="https://fonts.googleapis.com/css2?family=Josefin+Sans:wght@300;400;500;600;700&display=swap" rel="stylesh

     <!-- Css Styles -->
     <link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">
     <link rel="stylesheet" href="static/css/STYLES.css" type="text/css">

</head>

<body>
    <main>
        <div class="container">
            <div class="row">
                <!-- <div class="col-lg-12"><br><br><br><br><br><br><br><br>
                </div> -->
                <div class="col-lg-6" style="align-items:center">
                    <br><br><br><br><br><br><br><br>
            <h1>Loan Approval Prediction</h1>
            <h3>{{num}}</h3>
            </div>
            <div class="col-lg-6">
                <br><br><br><br>
```

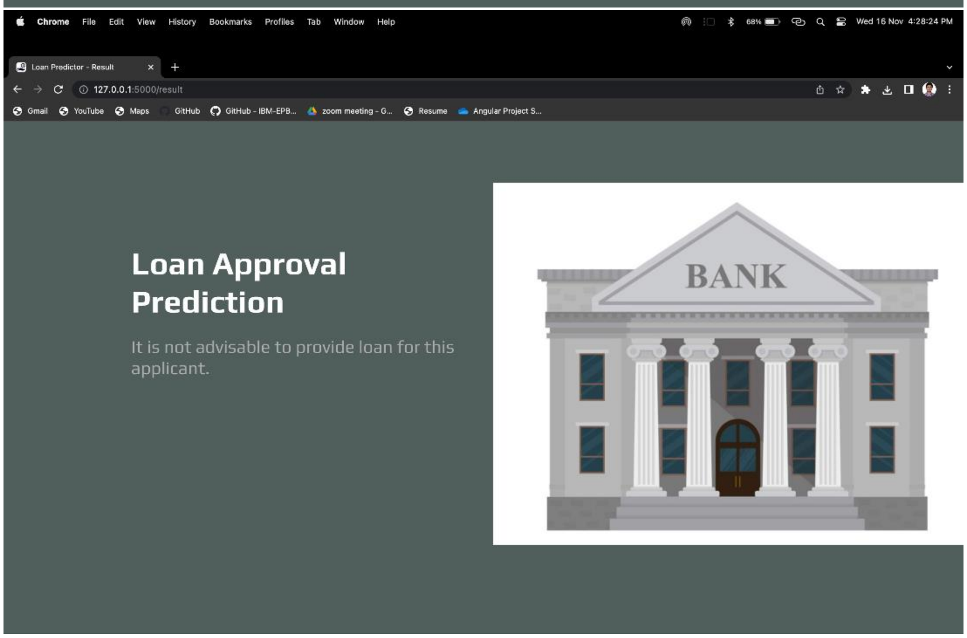
7.2 THE MACHINE LEARNING MODEL PREDICTS THE APPLICANT’S ELIGIBILITY

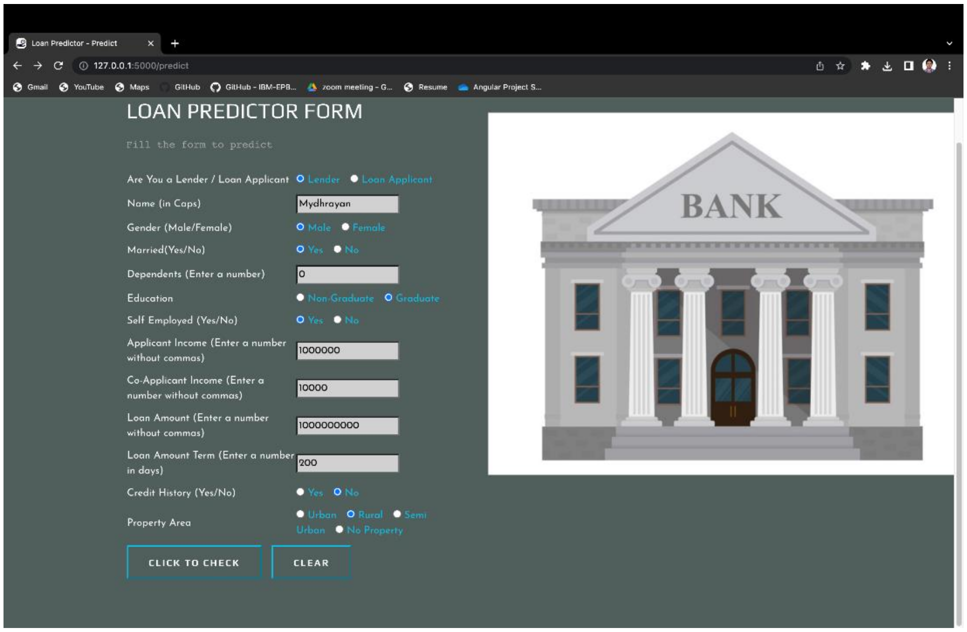
```
Project_Development_Phase — python • python APP.py — 111×31
IBM-Project-25357-1659960296-main
IBM-Project-36944-1660299156-main
IITM
Kaar Tech
MINI PROJECT
PLACEMENT
Screenshot 2022-10-30 at 10.48.08 PM.png
Sprint 3
bank-getty (1).jpg
bank-getty.jpg
seeya
shoumo2019.pdf
(base) hariharanganesh@Hariharans-MacBook-Pro Downloads % cd IBM-Project-36944-1660299156-main
(base) hariharanganesh@Hariharans-MacBook-Pro IBM-Project-36944-1660299156-main % ls
Assignments                       Project Design & Planning
Final Deliverables                Project_Development_Phase
Other_Files                       README.md
(base) hariharanganesh@Hariharans-MacBook-Pro IBM-Project-36944-1660299156-main % cd Project_Development_Phase
(base) hariharanganesh@Hariharans-MacBook-Pro Project_Development_Phase % ls
APP.py          Sprint_1        Sprint_2        Sprint_3        Sprint_4
(base) hariharanganesh@Hariharans-MacBook-Pro Project_Development_Phase % python APP.py
 * Serving Flask app 'APP' (lazy loading)
 * Environment: production
   WARNING: This is a development server. Do not use it in a production deployment.
   Use a production WSGI server instead.
 * Debug mode: on
 * Running on http://127.0.0.1:5000 (Press CTRL+C to quit)
 * Restarting with stat
 * Debugger is active!
 * Debugger PIN: 127-308-674
```


Chrome File Edit View History Bookmarks Profiles Tab Window Help
Wed 16 Nov 4:26:15 PM
Loan Predictor - Home
127.0.0.1:5000
Gmail YouTube Maps GitHub GitHub - IBM-EPB... zoom meeting - G... Resume Angular Project S...
LOAN PREDICTOR
Find your Loan Eligibility here
Tap below button and fill the details to know your Loan Eligibility.
CLICK TO CHECK
BANK


Loan Predictor - Predict
127.0.0.1:5000/predict
Gmail YouTube Maps GitHub GitHub - IBM-EPB... zoom meeting - G... Resume Angular Project S...
LOAN PREDICTOR FORM
Fill the form to predict
Are You a Lender / Loan Applicant Lender Loan Applicant
Name (in Caps) Mugesh
Gender (Male/Female) Male Female
Married(Yes/No) Yes No
Dependents (Enter a number) 0
Education Non-Graduate Graduate
Self Employed (Yes/No) Yes No
Applicant Income (Enter a number without commas) 100000
Co-Applicant Income (Enter a number without commas) 25000
Loan Amount (Enter a number without commas) 50000
Loan Amount Term (Enter a number in days) 365
Credit History (Yes/No) Yes No
Property Area Urban Rural Semi Urban No Property
CLICK TO CHECK
CLEAR
BANK


Chrome File Edit View History Bookmarks Profiles Tab Window Help
Wed 16 Nov 4:27:18 PM
Loan Predictor - Result
127.0.0.1:5000/result
Gmail YouTube Maps GitHub GitHub - IBM-EPB... zoom meeting - G... Resume Angular Project S...
Loan Approval Prediction
This applicant can be provided with the loan amount requested.
BANK


Chrome File Edit View History Bookmarks Profiles Tab Window Help
Wed 16 Nov 4:28:24 PM
Loan Predictor - Result
127.0.0.1:5000/result
Gmail YouTube Maps GitHub GitHub - IBM-EPB... zoom meeting - G... Resume Angular Project S...
Loan Approval Prediction
It is not advisable to provide loan for this applicant.
BANK


Loan Predictor - Predict
127.0.0.1:5000/predict
Gmail
YouTube
Maps
GitHub
GitHub - IBM-EPB...
zoom meeting - G...
Resume
Angular Project S...
LOAN PREDICTOR FORM
Fill the form to predict
Are You a Lender / Loan Applicant
Lender
Loan Applicant
Name (in Caps)
Mydhrayan
Gender (Male/Female)
Male
Female
Married(Yes/No)
Yes
No
Dependents (Enter a number)
0
Education
Non-Graduate
Graduate
Self Employed (Yes/No)
Yes
No
Applicant Income (Enter a number without commas)
1000000
Co-Applicant Income (Enter a number without commas)
10000
Loan Amount (Enter a number without commas)
1000000000
Loan Amount Term (Enter a number in days)
200
Credit History (Yes/No)
Yes
No
Property Area
Urban
Rural
Semi Urban
No Property
CLICK TO CHECK
CLEAR
BANK

8. TESTING

8.1 TEST CASES

Loan _ID	Gender	Married	Dependents	Education	Self-employed	Applicant Income	Co-applicant Income	Loan Amount	Loan Amount Term	Credit History	Property Area	Loan Status
LP001002	Male	No	0	Graduate	No	5849	0		360	1	Urban	Y
LP001003	Male	Yes	1	Graduate	No	4583	1508	128	360	1	Rural	N
LP001005	Male	Yes	0	Graduate	Yes	3000	0	66	360	1	Urban	Y
LP001006	Male	Yes	0	Not Graduate	No	2583	2358	120	360	1	Urban	Y
LP001008	Male	No	0	Graduate	No	6000	0	141	360	1	Urban	Y
LP001011	Male	Yes	2	Graduate	Yes	5417	4196	267	360	1	Urban	Y
LP001013	Male	Yes	0	Not Graduate	No	2333	1516	95	360	1	Urban	Y
LP001014	Male	Yes	3	Graduate	No	3036	2504	158	360	0	Semiurban	N
LP001018	Male	Yes	2	Graduate	No	4006	1526	168	360	1	Urban	Y
LP001020	Male	Yes	1	Graduate	No	12841	10968	349	360	1	Semiurban	N

8.2 USER ACCEPTANCE TESTING

1. Purpose of Document

The purpose of this document is to briefly explain the test coverage and open issues of the Smart Lender - Applicant Credibility Prediction for Loan Approval project at the time of the release to User Acceptance Testing (UAT).

2. Defect Analysis

This report shows the number of resolved or closed bugs at each severity level, and how they were resolved

Resolution	Severity 1	Severity 2	Severity 3	Severity 4	Subtotal
By Design	10	4	2	3	20
Duplicate	1	0	3	0	4
External	2	3	0	1	6
Fixed	11	2	4	20	37
Not Reproduced	0	0	1	0	1
Skipped	0	0	1	1	2
Won't Fix	0	5	2	1	8
Totals	24	14	13	26	77

3. Test Case Analysis

This report shows the number of test cases that have passed, failed, and untested

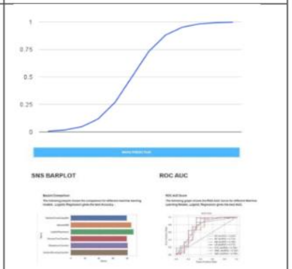
Section	Total Cases	Not Tested	Fail	Pass
Print Engine	7	0	0	7
Client Application	51	0	0	51
Security	2	0	0	2
Outsource Shipping	3	0	0	3
Exception Reporting	9	0	0	9
Final Report Output	4	0	0	4
Version Control	2	0	0	2

9. RESULTS

9.1 PERFORMANCE METRICS

Model Performance Testing:

Project team shall fill the following information in model performance testing template.

S.No	Parameter	Values	Screenshot
1.	Metrics	**Regression Model:** MAE - , MSE - , RMSE - , R2 score -**Classification Model:** Confusion Matrix - , Accuray Score- & Classification Report -	
2.	Tune the Model	Hyperparameter Tuning - Validation Method -	

For the purpose of predicting the loan approval status of the applied customer, we have chosen the machine learning approach to study the bank dataset. We have applied various machine learning algorithms to decide which one will be the best for applying on the dataset to get the result with the highest accuracy. Following this approach, we found that apart from the logistic regression, the rest of the algorithms performed satisfactory in terms of giving out the accuracy. The accuracy range of the rest of the algorithms were from 75% to 85%. Whereas the logistic regression gave us the best possible accuracy (88.70%) after the comparative study of all the algorithms.

We also determined the most important features that influence the loan approval status. These most important features are then used on some selected algorithms and their performance accuracy is compared with the instance of using all the features. This model can help the banks in figuring out which factors are important for the loan approval procedure. The comparative study makes us clear about which algorithm will be the best and ignores the rest, based on their accuracy.

10. ADVANTAGES AND DISADVANTAGES

ADVANTAGES

The benefits of loan tech servicing software for lenders include:

- Eliminating human error

It's no secret, that calculations are something that algorithms handle better than we, humans. In a lending system, there are just too many variables, which is why it is error-prone. The best loan servicing software, however, is created to completely rule out any errors, which is, undoubtedly, beneficial from every standpoint.

- Preventing delays in payment

Not being able to collect a debt is something that most lenders are especially wary of. However, if they leverage a traditional loan management approach, they may not see it coming. Loan servicing systems, on the other hand, integrate analytic modules capable of detecting even the most subtle fluctuations in clients' credibility and preventing payment delays in a timely manner.

- Saving time

Loan management requires a great level of meticulousness and attention to detail. As a rule, a full-fledged team is required to deal with every aspect of a loan process. Needless to say, loan management carried out manually and based on paperwork takes up a lot of time. A digital lending system, on the other hand, automates the routines and enables your team to dedicate time to other important tasks.

Read also: [Use Trunk Based Development for Product Agility]

Learn how to deliver a product in real-time easy and fast

- Automated reporting

Automated report generation is another invaluable feature offered by a digital loan servicing platform. Accounting, tax reports, and invoices are often requested by regulatory bodies, borrowers and investors. These high urgency reports should be provided on demand, and contain information, which is 100% accurate. Loan tracking software enables lenders to quickly generate reports of different types and submit them urgently, in the required formats.

- Increased revenue

This stems from all of the above: an automated loan processing system enables lenders to process more applications, assign and manage more loans, and see them all the way through closing all while detecting scams and preventing delays. The staff is free to oversee the process and focus on client relationships and look for new business opportunities. This enables financial companies to gain a distinct competitive edge and increase revenue.

DISADVANTAGES

1. Accessibility

An organization looking to build loan software may not have enough on-premise infrastructure capacities to ensure its non-disruptive operation, updates, and support. Scaling during peak workloads and handling an increase in the number of users and subscriptions may also be quite challenging. Using cloud infrastructure is best to ensure optimal scalability and availability.

2. Servicing different loan types

The more types of loans your money lending software is capable of servicing, the better. Lending applications that have a wide range of use cases, will surely attract more users than apps targeting only one specific loan type. A loan Tech software to create loan app estimation, for example, may have a broad range of applications from student loan tech calculations to estimating business loans and mortgages.

3. Centralized data storage

Every stage of the lending process involves working with customer data. The best loan servicing software stores this data in centralized storage accessible during every loan processing stage. A legacy loan management system, on the other hand, uses a siloed approach to data storage, which makes loan processing more laborious and lengthier.

4. Integrated credit assessment capabilities

Modern loan servicing software for private lenders should be able to instantly connect with credit bureaus and any other bodies responsible for credibility assessment. Such platforms should receive regular credit data updates and leverage big data analytics to assess the trustworthiness of applicants. The client's social media activity, for example, can be a valid source of alternative assessment of credibility.

5. Automation of routine processes

Using robotic process automation to streamline simple rule-based processes is another must-have feature of a loan management platform. Automation accelerates loan origination and processing and accounts for increasing client satisfaction. On top of that, it helps to avoid human error.

6. In-built analytic modules

Leveraging artificial intelligence (AI) and big data is another hallmark of excellent loan servicing software for lenders. Not only does it help to generate reports but also enables companies to evaluate market trends, detect patterns in customer behavior and come up with new products and offerings.

7. Third-party integration

Another feature that most organizations find especially attractive in a loan processing system, is its capability to integrate with other enterprise software. ERP and CRM solutions are capable of enriching the lending system with data and insights. Systems integrating lending modules with software for remote sales personnel are also enjoying high popularity among lenders.

8. Security

Finance company software works with classified and highly sensitive data, and for both lenders and customers, security is a matter of paramount importance. An excellent lending system should possess advanced security capabilities, and ensure the highest level of customer, data, and network protection.

11. CONCLUSION

In the debate over which supervised learning model to use for credit risk assessment, we have come to the conclusion that support vector machines can outperform other tree-based models or regression models if the setup of the experiment is similar to that of ours. Furthermore, in the debate over which dimensionality reduction technique to use, our model has shown us that recursive feature elimination with cross-validation can outperform models based on principal component analysis. For future improvements we would like to use more current data and from different sources for illustrating a better understanding of the trends present in this field. Datasets similar to the above-mentioned experiments from previous works will be used to test this model for better comparison. It has been mentioned that in order to reduce computational cost and complexity we have omitted the idea of using neural networks. But as we are looking forward to work with even larger amount of data, we would like to make a comparative analysis using neural networks as well. It is a known fact that neural networks tend to perform better with large datasets and we would like to test this hypothesis in our future works. Again, as we are also discussing the contributions of feature selection/extraction techniques, we would like to implement other dimensionality reduction techniques such as genetic algorithm, univariate feature selection methods, tree-based feature selections etc. to gauge their performances and further improve the efficiency of the credit lending sector. Therefore, this paper can be concluded with the statement that this model illustrates an interesting approach in identifying loan defaulters in this ever-changing economy. Using the dataset from Lending Club our model has brought about remarkable results which in turn can play a major role in assessing the credit risk of borrowers, assist credit lending institutions and enable financial institutions to keep operating in a transparent and profitable way.

12. FUTURE SCOPE

In this section, based on various performance metrics, a comparative analysis will be made of all the generated models. A precise classifier is the backbone of any machine learning model. Four supervised algorithms: Support vector machine (SVM), Logistic Regression (LR), Extreme Gradient Boosting (XGB) and Random Forest (RF) have been selected for the analysis. The hyperparameters of these algorithms will be tuned using GridSearchCV to select the best set of values for each model. The results will be discussed in two categories and will be illustrated in both a graphical and tabular manner. Firstly the models will be evaluated on a holdout test set using a train test split. Then another comparative analysis will be made of the same models but using a 5 fold cross-validation and GridSearchCV.

Z-score has been chosen over normalization (min-max scaling) for scaling the features. Classifiers such as support vector machine, logistic regression or neural network prefer standardization over normalization. Additionally, this paper proposes to use such feature extraction methodologies where maximizing the variance is highly preferred. This can be achieved using standardization. Furthermore GridSearchCV has been used to optimize the hyperparameters of each classifier. Studies done in perfectly show the effectiveness of GridSearchCV in maximizing the performance of classifiers.

13. APPENDIX

SOURCE CODE

APP.PY (PYTHON FILE)

```
import numpy as np
import os
from PIL import Image
from flask import Flask, request, render_template, url_for
from werkzeug.utils import secure_filename, redirect
from gevent.pywsgi import WSGIServer
from flask import send_from_directory
from joblib import Parallel,delayed
import joblib
import pandas as pd
from scipy.sparse import issparse



app = Flask(__name__)




@app.route('/')
def index():
    return render_template('home.html')

@app.route('/predict')
def predict():
```

```
    return render_template('predict.html')


@app.route('/result', methods=['GET', 'POST'])

def upload():

    if request.method == "POST":

        lend_data = request.form.get('lend')

        data =
[[request.form.get('gender'),request.form.get('married'),request.form.get('dep'),request.form.get('edu'),request.form.g
et('se'),request.form.get('ai')

,request.form.get('cai'),request.form.get('la'),request.form.get('lat'),request.form.get('ch'),request.form.get('pa')]]


        df = pd.DataFrame(data, columns=['Gender','Married','Dependents','Education',

                                'Self_Employed','ApplicantIncome','CoapplicantIncome',

                                'LoanAmount','Loan_Amount_Term','Credit_History','Property_Area'])


        gh=joblib.load('/Users/hariharanganesh/Downloads/IBM/PROJECT/rdf.pkl')

        num=gh.predict(df)

        a=''

        lend_data=int(lend_data)

        if(num==0):

            if(lend_data==1):

                a='It is not advisable to provide loan for this applicant.'

            else:

                a='Your Loan application will be Rejected.'

        else:

            if(lend_data==1):

                a='This applicant can be provided with the loan amount requested.'

            else:

                a='Your Loan application will be succesfull.'

        return render_template('submit.html', num=a)
```

```python
if __name__ == '__main__':
    app.run(debug=True, threaded=False)
```

HOME.HTML

```html
<!DOCTYPE html>

<html lang="en">



<head>

<meta charset="UTF-8">

<meta http-equiv="X-UA-Compatible" content="IE=edge">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<title>Loan Predictor - Home</title>

<link rel="icon" href="static/images/money.png" type="image/x-icon">



<!-- Google Font -->

<link href="https://fonts.googleapis.com/css2?family=Play:wght@400;700&display=swap" rel="stylesheet">

<link href="https://fonts.googleapis.com/css2?family=Josefin+Sans:wght@300;400;500;600;700&display=swap"
rel="stylesheet">



<!-- Css Styles -->

<link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">

<!-- <link rel="stylesheet" href="static/css/font-awesome.min.css" type="text/css">
```

```
<link rel="stylesheet" href="static/css/elegant-icons.css" type="text/css">

<link rel="stylesheet" href="static/css/owl.carousel.min.css" type="text/css">

<link rel="stylesheet" href="static/css/magnific-popup.css" type="text/css">

<link rel="stylesheet" href="static/css/slicknav.min.css" type="text/css"> -->

<link rel="stylesheet" href="static/css/STYLES.css" type="text/css">



</head>



<body> <!--background="../STATIC/IMAGES/Lending.jpg"-->

<main>

<div class="container">

<div class="row">

<!-- <div class="col-lg-12"><br><br><br><br><br><br><br><br>

</div> -->

<div class="col-lg-6" style="align-items:center">

<br><br><br><br><br><br><br><br>

<h1>LOAN PREDICTOR</h1>

<h3>Find your Loan Eligibility here</h3><br>

<h5>Tap below button and fill the details to know your Loan Eligibility.</h5><br>



<div class="portfolio__btn">
```

```
<a href="{{ url_for('predict') }}" class="tm-nav-link primary-btn" data-hover="Loan Predictor">

Click to Check

</a>

</div>

</div>

<div class="col-lg-6">

<br><br><br><br>

<img src="static/images/Lending_1.jpg">

</div>

<div class="col-lg-12"><br><br></div>

</div>

</div>




</main>

</body>



</html>
```

PREDICT.HTML

```
<!DOCTYPE html>
```

```html
<html lang="en">

<head>
<meta charset="UTF-8">
<meta http-equiv="X-UA-Compatible" content="IE=edge">
<meta name="viewport" content="width=device-width, initial-scale=1.0">
<title>Loan Predictor - Predict</title>
<link rel="icon" href="static/images/money.png" type="image/x-icon">

<!-- Google Font -->
<link href="https://fonts.googleapis.com/css2?family=Play:wght@400;700&display=swap" rel="stylesheet">
<link href="https://fonts.googleapis.com/css2?family=Josefin+Sans:wght@300;400;500;600;700&display=swap"
rel="stylesheet">
<!-- Css Styles -->
<link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">
<link rel="stylesheet" href="static/css/STYLES.css" type="text/css">

</head>

<body>
<div class="container">
```

```
<div class="row">

<div class="col-lg-6" style="align-items:center">

<br><br><br>

<h2>LOAN PREDICTOR FORM</h2>

<p>Fill the form to predict</p>

<table>

<form action="/result" method="POST" enctype="'multipart/form-data">

<tr>

<td>Are You a Lender / Loan Applicant</td>

<td class="td1"><input id="lend" name="lend" type="radio" value=1 required> Lender  <input
id="lend" name="lend" type="radio" value=0> Loan Applicant</td>

</tr>

<tr>

<td>Name (in Caps)</td>

<td><input name='name' type="text" required=""></td>

</tr>

<tr>

<td>Gender (Male/Female)</td>

<td class="td1"><input id="gender" name="gender" type="radio" value=1
required> Male  <input id="gender" name="gender" type="radio" value=0> Female</td>

</tr>

<tr>

<td>Married(Yes/No)</td>
```

```
<td class="td1"><input id="married" name="married" type="radio" value=1 required> Yes  <input id="married" name="married" type="radio" value=0> No</td>
</tr>
<tr>
<td>Dependents (Enter a number)</td>
<td><input name='dep' type="number" min="0" step='1' placeholder="" required=""></td>
</tr>
<tr>
<td>Education</td>
<td class="td1"><input id='edu' name='edu' type="radio" value=1 required> Non-Graduate  <input id="edu" name="edu" type="radio" value=0> Graduate</td>
</tr>
<tr>
<td>Self Employed (Yes/No)</td>
<td class="td1"><input id="se" name="se" type="radio" value=1 required> Yes  <input id="se" name="se" type="radio" value=0> No</td>
</tr>
<tr>
<td>Applicant Income (Enter a number without commas)</td>
<td><input id='AI' name='ai' type="number" min='0' required=""></td>
</tr>
<tr>
<td>Co-Applicant Income (Enter a number without commas)</td>
<td><input id='CAI' name='cai' type="number" min='0' required=""></td>
```

```
</tr>
<tr>
<td>Loan Amount (Enter a number without commas)</td>
<td><input id='la' name='la' type="number" min='0' required=""></td>
</tr>
<tr>
<td>Loan Amount Term (Enter a number in days)</td>
<td><input id='lat' name='lat' type="number" min='0' step="1" required=""></td>
</tr>
<tr>
<td>Credit History (Yes/No)</td>
<td class="td1"><input id="ch" name="ch" type="radio" value=1 required> Yes  <input id="ch" name="ch" type="radio" value=0> No</td>
</tr>
<tr>
<td>Property Area</td>
<td class="td1"><input id="pa" name="pa" type="radio" value=2 required> Urban  <input id="pa" name="pa" type="radio" value=0> Rural  <input id="pa" name="pa" type="radio" value=1> Semi Urban  <input id="pa" name="pa" type="radio" value=0> No Property</td>
</tr>
<tr>
<td>
<div class="portfolio__btn">
```

```
<input class="primary-btn-1" id="submit" type="submit" value="Click to Check"> <input class="primary-btn-1" id="reset" type="reset" value="Clear">

</div>

</td>

<!-- <td><div class="container">

<a href="./submit.html">

<button class="btn" data-hover="Loan Predictor">

<div>Click to Check</div>

</button>

</a>

</div></td> -->

</tr>

</form>

</table>

</div>

<div class="col-lg-6">

<br><br><br><br>

<img src="static/images/Lending_1.jpg">

</div>

<div class="col-lg-12"><br><br><br><br><br></div>

</div>

</div>

</body>

</html>
```

SUBMIT.HTML

```html
<!DOCTYPE html>

<html lang="en">



<head>

<meta charset="UTF-8">

<meta http-equiv="X-UA-Compatible" content="IE=edge">

<meta name="viewport" content="width=device-width, initial-scale=1.0">

<title>Loan Predictor - Result</title>

<link rel="icon" href="static/images/money.png" type="image/x-icon" style="border-radius: 5px;">



<!-- Google Font -->

<link href="https://fonts.googleapis.com/css2?family=Play:wght@400;700&display=swap" rel="stylesheet">

<link href="https://fonts.googleapis.com/css2?family=Josefin+Sans:wght@300;400;500;600;700&display=swap"
rel="stylesheet">

<!-- Css Styles -->

<link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">

<link rel="stylesheet" href="static/css/STYLES.css" type="text/css">



</head>
```

```
<body>

<main>

<div class="container">

<div class="row">

<!-- <div class="col-lg-12"><br><br><br><br><br><br><br><br>

</div> -->

<div class="col-lg-6" style="align-items:center">

<br><br><br><br><br><br><br><br>

<h1>Loan Approval Prediction</h1>

<h3>{{num}}</h3>

</div>

<div class="col-lg-6">

<br><br><br><br>

<img src="static/images/Lending_1.jpg">

</div>

<div class="col-lg-12"><br><br></div>

</div>

</div>

</main>

</body>

</html>
```

GITHUB LINK: https://github.com/IBM-EPBL/IBM-Project-19649-1659703271

PROJECT DEMO LINK: Project Demo Video