## **SMART LENDER -**

## APPLICANT CREDIBILITY PREDICTION FOR LOAN

## 1.INTRODUCTION

## 1.1 PROJECT OVERVIEW

Accurate credit risk rating systems have always been essential for the proper and profitable functioning of financial institutions. In an everchanging economy where loan default rates are slowly rising, financial institutions are finding it increasingly difficult to adequately evaluate loan applications and manage the risk of loan default. I'm starting to notice. In light of these events, this white paper proposes a machine learning model that can accurately assess credit risk and predict potential loan defaults for lending institutions. We performed comparative analysis and identified defaults using corresponding supervised learning algorithms such as support vector machines, random forests, extreme gradient boosting, and logistic regression. Recursive feature elimination with cross-validation and principal component analysis was used for dimensionality reduction. We evaluated each model using metrics such as F1 score, AUC score, prediction accuracy, precision, and recall. Among all models, a combination of recursive feature removal using tuned support vector machines and cross-validation has shown promise in identifying loan delinquents. Therefore, the proposed model will help financial institutions to accurately identify loan defaulters and prevent them from suffering further losses.

## 1.2 PURPOSE

Credit is a bank's core business. The main winnings come directly from his interest on the loan. The lending company issues loans after intensive screening and his verification process. However, we are still unsure as to whether the applicant will be able to repay her loan without issue.

One of the most important factors affecting our country's economic and financial situation is the credit system administered by banks. Bank credit risk assessment processes are recognized by banks around the world. "We know that credit risk assessment is very important, so different techniques are used to calculate the risk level. In addition, credit risk is one of the main functions of the banking industry.

Predicting loan defaults is one of the most difficult tasks for any bank. However, by anticipating loan defaults, banks can ensure that losses are reduced by reducing charitable assets, so recovery of approved loans can be done loss, and bank statement contributions can be reduced. Can act as a parameter. This makes the study of this forecast important for credit approval. Machine learning techniques are very important and useful in predicting these types of data.

The company would like to automate (real time) the credit eligibility process based on customer data provided when completing an online application form. These details include gender, marital status, education, number of dependents, income, loan amount, and credit history. To automate this process, they combined the problem of identifying her customer segments eligible for loan amounts so they could target those customers.

## 2. LITERATURE SURVEY

## 2.1 Existing Problem

So far there have been many notable insights that can make credit risk analysis much more efficient and accurate, but in the ever-changing era of modern science there is always room for improvement. One thing should be mentioned. Well, paper shows some great early work done in credit risk assessment using machine learning. This paper presents a comparison of different credit risk assessment

models and a multi -agent based system built using adaptive linear neural network achieved the best results with an accuracy of 71.19%. This paper proposes a multi-level reliability-based neural network ensemble learning model for credit risk assessment. Using 653 instances of credit card authorization records,

the best accuracy achieved by the model is 88.08%.

## 2.2 References

- [1] D. Faggella, "What is Machine Learning? | Emerj Artificial Intelligence Research and Insight," Emerj. [Online]. Available: <a href="https://emerj.com/ai-glossary-terms/what-is-machine-learning/">https://emerj.com/ai-glossary-terms/what-is-machine-learning/</a>. [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/defaultlon-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

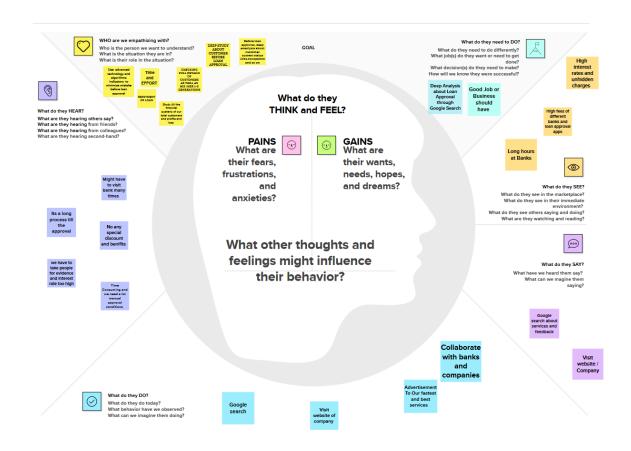
One of the most important factors affecting our country's economic and financial situation is the credit system administered by banks. Bank credit risk assessment processes are recognized by banks around the world. "We know that the assessment of credit risk is very important, so different techniques are used to calculate the risk level. Moreover, credit risk is one of the key functions of the banking industry.

Predicting loan defaults is one of the most difficult tasks for any bank. However, by anticipating loan defaults, banks can ensure that losses are reduced by reducing charitable assets, so that recovery of approved loans can be made without loss, and bank statement parameters can function as This makes the study of this forecast important for credit approval. Machine learning techniques are very important and useful in predicting these types of data.

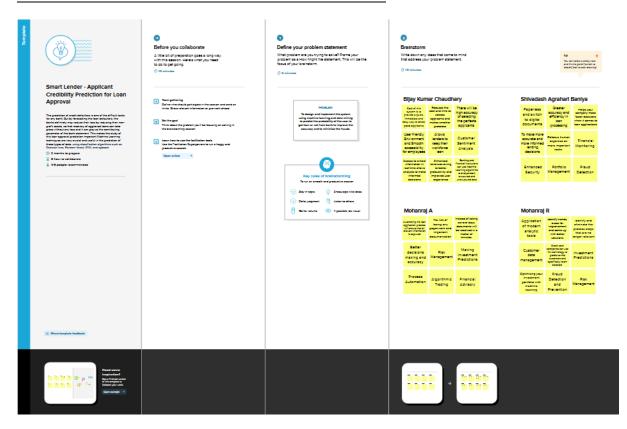
Use classification algorithms such as decision trees, random forests, KNN, and xgboost. Use these algorithms to train and test data. The best model is selected from this and saved in pkl format. Flask integration and IBM deployment.

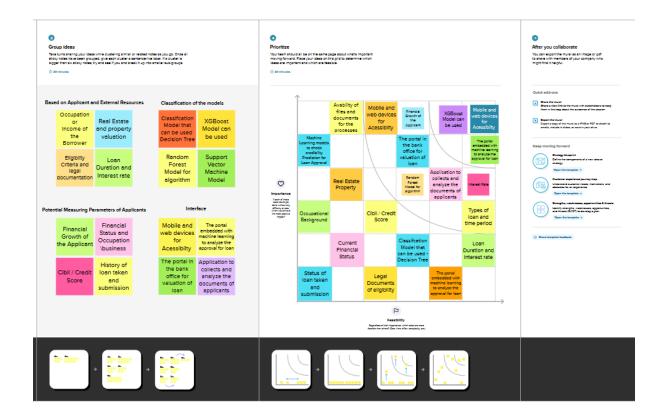
## 3.IDEATION AND PROPOSED SOLUTION

## 3.1 EMPATHY MAP CANVAS



## 3.2 IDEATION AND BRAINSTORMING





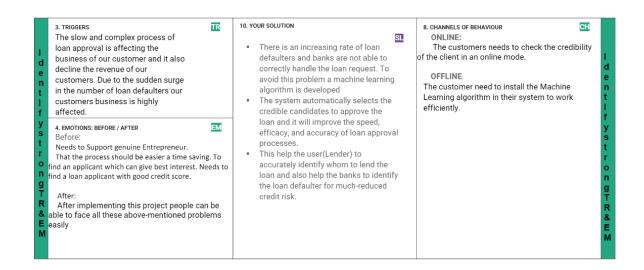
## 3.3 PROPOSED SOLUTION

#### **Proposed Solution**

S.No.	Parameter	Description
1.	Problem Statement (Problem to be solved)	Loan Prediction is very helpful for employee of banks as well as for the applicant also. The aim of this Search is to provide quick, immediate and easy way to choose the deserving applicants. It can provide special advantages to the bank. The Loan Prediction System can automatically calculate the weight of each features taking part in loan processing and on new best data same features are processed with respect to their associated weight.
2.	Idea / Solution description	A loan is a form of debt incurred by an individual or other entity. The lender— usually a corporation, financial institution, or government—advances a sum of money to the borrower. In return, the borrower agrees to a certain set of terms including any finance charges, interest, repayment date, and other conditions.
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

## 3.4 PROBLEM SOLUTION FIT

#### 6. CUSTOMER CONSTRAINTS 5. AVAILABLE SOLUTIONS 1. CUSTOMER SEGMENT(S) Our target customers are mostly Banks are not to correctly handle the loan Random forest, Logistic request. People within a protected class banking firm, small financial firms regression, Decision tree and being clearly treated differently than those of that lends out loan and credit Naive bayes algorithm are non-protected classes for loan. There is an card companies because of the increasing rate of loan defaults. Banks increasing rate of loan defaulter identify the loan defaulters for much-Using data pre-processing and also to increase the slow reduced credit risk as large portions of a data mining and data filtering process of the loan approval. bank's assets directly come from the interest Algorithms such as naïve earned on loans given. bayes, k-nearest neighbors are used. 2. JOBS-TO-BE-DONE / PROBLEMS 9. PROBLEM ROOT CAUSE RC 7. BEHAVIOUR BE Needs to Support genuine Entrepreneur. The root cause of this problem is the Directly related: That the process should be easier a time The customers who lends the loan and banks identify the loan defaulters for saving. To find an applicant which can the banks that checks the credibility much-reduced credit risk as large give best interest. Needs to find a loan seek to do the process faster. portions of a bank's assets directly applicant with good credit score come from the interest earned on Indirectly associated: loans given. . People within a protected The small finance sector that deals with class being clearly treated differently middle class and poor class people seek than those of non-protected classes for to find the credibility.



## **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 Requirement	To check the loan eligibility using the credit score, prediction for loan approval.
FR-2	User Confirmation	Through one time verification and using captcha etc
FR-3	Profile Updation	The user can update their profile when their is need to add any add-ons to it.
FR-4	User Registration	The user gets login or signup using Gmail account or by using mobile number.
FR-5	User Authentication	By OTP or verification code the user gets authenticated and OTP is used for mobile number registration.
FR-6	Feedback Evaluation	The user provided feedbacks are used for evaluation of app performance and updation is made over that.

## **4.2 NON-FUNCTIONAL REQUIREMENTS**

#### Non-functional Requirements:

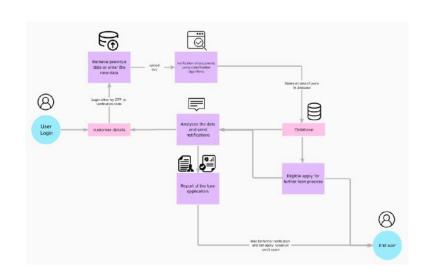
Following are the non-functional requirements of the proposed solution.

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	This application is mainly used to analyse cibil score and predict the eligibility for users to avail for loan approval by following community guidelines.
NFR-2	Security	It uses OTP and verify code verification for each user and uses hybrid security features over internet to safely maintain the updated documents of user .
NFR-3	Reliability	Maintaining the app up to date for reliant features ,durability and efficiency of the mobile app by releasing patch fix and software updates.
NFR-4	Performance	It has a user friendly interface and can check multiple persons cibil score parallels irrespective of server traffic .It stores the data collected over in a efficient database
NFR-5	Availability	It is platform independent and it is available where the users are able to wish it want to be. Depending upon the user requirements all services get offered.

## 5. PROJECT DESIGN

#### 5.1 DATA FLOW DIAGRAM

Data Flow Diagrams:

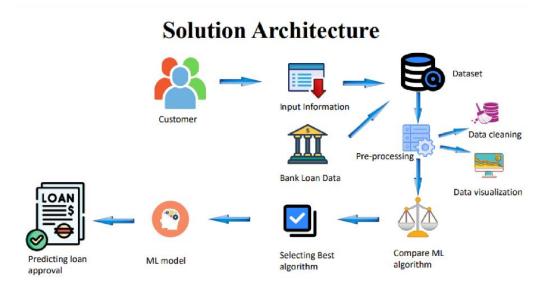


# 5.2 SOLUTION AND TECHNICAL ARCHITECTURE SOLUTION ARCHITECTURE

One of the most important factors affecting our country's economic and financial situation is the credit system administered by banks. Bank credit risk assessment processes are recognized by banks around the world. "We know that assessing credit risk is very important, so different techniques are used to calculate risk levels. Moreover, credit risk is one of the key functions of the banking industry.

Predicting loan defaults is one of the most difficult tasks for any bank. However, by anticipating loan defaults, banks can ensure that losses are reduced by reducing charitable assets, so that recovery of approved loans can be made without loss, and bank statement parameters can function as This makes the study of this forecast important for credit approval. Machine learning techniques are very important and useful in predicting these types of data.

Use classification algorithms such as decision trees, random forests, KNN, and xgboost. Use these algorithms to train and test data. The best model is selected from this and saved in pkl format. Flask integration and IBM deployment.



## **5.3 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	Release
Customer (Mobile user)	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	As a user I can enter Gmail and set a password	High	Sprint-1
		USN-2	As a user, I will receive confirmation email once I have registered for the application	I can get a code for confirmation	High	Sprint-1
		USN-3	Registration as a user can be confirmed using OTP or verification code.	As a user can get OTP or verification code	Low	Sprint-1
	Login	USN-4	Users can log into the web/mobile interface by storing or using the registered login credentials.	Able to login	Medium	Sprint-1
		USN-5	As a user, I can log into the application by entering email & password	Can be able to login using Gmail	Medium	Sprint-1
	Dashboard	USN-6	As a user,I should be able to login the profile or status dashboard	Able to access dashboard account	Medium	Sprint-2
Customer care executive		USN-7	Checks the user feedbacks and provide essential technical support	Access the account/ able to access the dashboard		Sprint-2
Loan approval Executive	Automated analysis of cibil-score	USN-8	As a loan approval officer I can make decisions by checking and monitoring all the feeded applications and getting to a prediction.	Get a decision for loan prediction based on the details provided in the loan application	High	Sprint-3
		USN-9	As a admin cibil score which represents credit history plays major role in analysis	Cibil score /credit history plays major role	High	Sprint-3
Admin	Login/Register	USN-10	As an admin I should be able to login with a unique email and password.	Able to get logged in	High	Sprint-4
	Dashboard	USN-11	As an admin I need the access of full authority towards the dashboard.	Access the dashboard	Medium	Sprint-4

## 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	Registration	USN-1	As a user, I can register for the application by entering my email, password, and confirming my password.	2	High	BIJAY KUMAR CHUADHARY
Sprint-1		USN-2	As a user, I will receive confirmation email once I have registered for the application	1	High	MOHANRAJ A MOHANRAJ R
Sprint-1	Login	USN-3	As a user, I can login the application through email	2	Low	SHIVDASH AGRAHARI BANIYA
Sprint-2	Documents Upload	USN-4	As a user,I can upload all my details and documents	3	High	BIJAY KUMAR CHAUDHARY
Sprint-2	View Details	USN-5	As a user ,I can view all the details and procedure	2	Low	MOHANRAJ A MOHANRAJ R
Sprint-3	View user details	USN-6	As a Loan Approval administrator,I can view the user details	3	Medium	BIJAY KUMAR CHAUDHARY MOHANRAJ A
Sprint-3	Credit score verification	USN-7	As a Loan Approval administrator ,I can verify the cibil score details		High	SHIVDASH AGRAHARI BANIYA BIJAY KUMAR CHAUDHARY

Sprint-4	Verify Documents	USN-8	As a Bank Administrator ,I can verify all the documents and proofs uploaded by the user.	3	High	BIJAY KUMAR CHAUDHARY
Sprint-4	Loan Approval Details	USN-9	As a Bank Administrator ,I can Approve or Reject the loan for the customer based on the uploaded details	3	High	BIJAY KUMAR CHAUDHARY SHIVDASH AGRAHARI BANIYA
Sprint-4		USN-10	As a user,I can get the confirmation mail through email or messages.	through email or 2 Medium SHIVDAS AGRAHA MOHANR		SHIVDASH AGRAHARI BANIYA MOHANRAJ R MOHANRAJ A

## **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	5	6 Days	24 Oct 2022	29 Oct 2022	5	29 Oct 2022
Sprint-2	5	6 Days	31 Oct 2022	05 Nov 2022	5	05 Nov 2022
Sprint-3	6	6 Days	07 Nov 2022	12 Nov 2022	6	12 Nov 2022
Sprint-4	8	6 Days	14 Nov 2022	19 Nov 2022	8	19 Nov 2022

## 6.3 REPORTS FROM JIRA

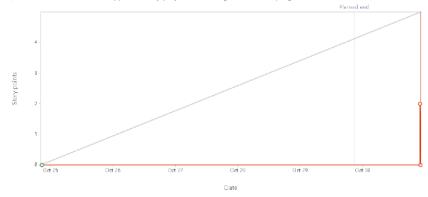
#### Velocity:

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)

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

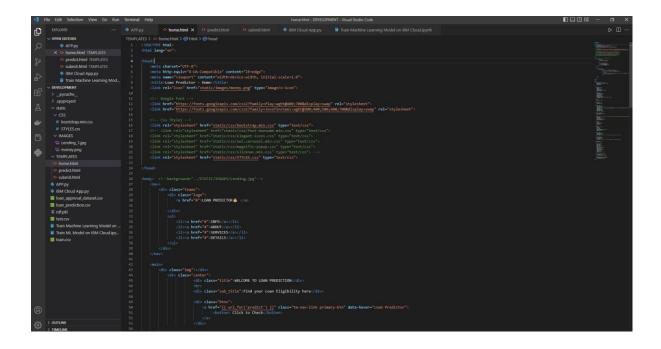
#### **Burndown Chart:**

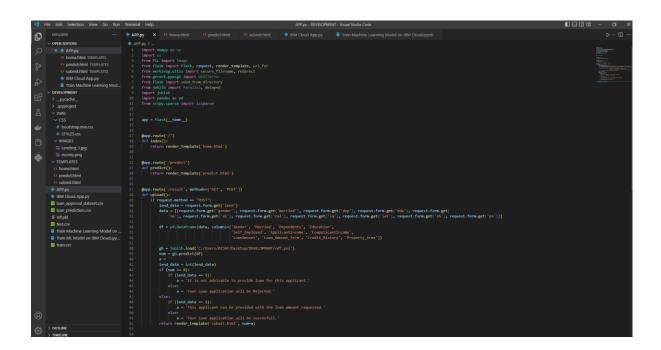
A burndown chart is a graphical representation of work left to do versus time. It is often used in agile software development methodologies such as Scrum. However, burn down charts can be applied to any project containing measurable progress over time.

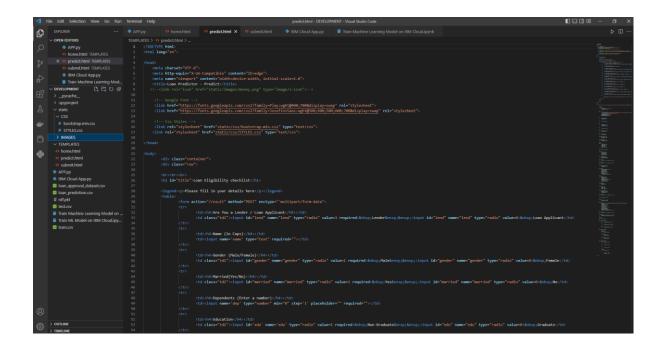


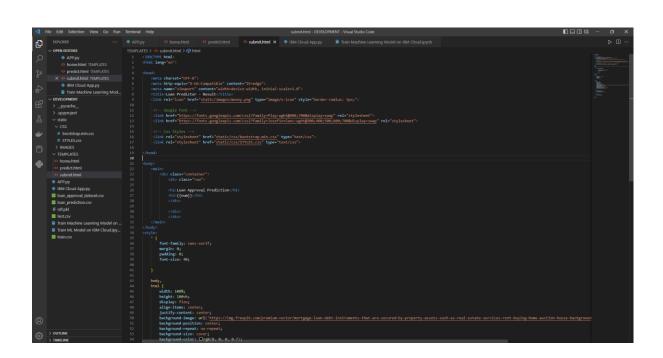
## 7. CODING AND SOLUTIONING

## 7.1 FORM FILLING BT APPLICANT TO CHECK HIS **CREDIT ELIGIBILITY**





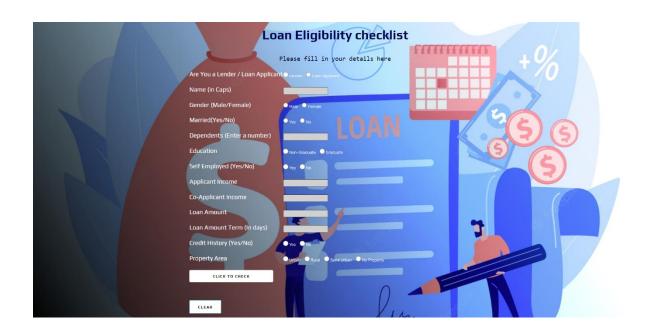


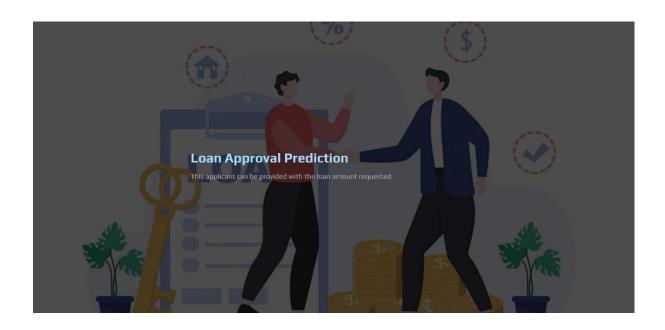


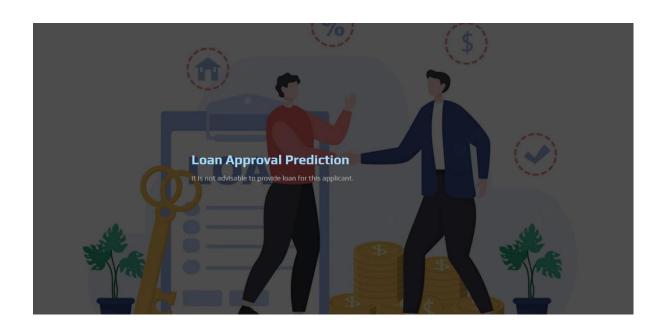
```
| No. Col. | No. Col. | No. Col. | No. Section | No. Col. | No. Co
```

## 7.2 THE MACHINE LEARNING MODEL PREDICTS THE APPLICANT'S ELIGIBILITY









## 8. TESTING

## 8.1 TEST CASES

Loan _ID	Gen der	Mar ried	Depend ents	Educati on	Self- employed	Applicant Income	Co-applicant Income	Loan Amount	Loan Amount Term	Credit History	Property Area	Loan Status
LP00 1002	Mal e	No	0	Graduate	No	5849	0		360	1	Urban	Y
LP00 1003	Mal e	Yes	1	Graduate	No	4583	1508	128	360	1	Rural	N
LP00 1005	Mal e	Yes	0	Graduate	Yes	3000	0	66	360	1	Urban	Y
LP00 1006	Mal e	Yes	0	Not Graduate	No	2583	2358	120	360	1	Urban	Y
LP00 1008	Mal e	No	0	Graduate	No	6000	0	141	360	1	Urban	Y
LP00 1011	Mal e	Yes	2	Graduate	Yes	5417	4196	267	360	1	Urban	Y
LP00 1013	Mal e	Yes	0	Not Graduate	No	2333	1516	95	360	1	Urban	Y
LP00 1014	Mal e	Yes	3	Graduate	No	3036	2504	158	360	0	Semiurba n	N
LP00 1018	Mal e	Yes	2	Graduate	No	4006	1526	168	360	1	Urban	Y
LP00 1020	Mal e	Yes	1	Graduate	No	12841	10968	349	360	1	Semiurba n	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	Pas s
Print Engine	7	0	0	7

#### Page 2 of 2

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 -	10 10 10 10 10 10 10 10 10 10 10 10 10 1
2.	Tune the Model	Hyperparameter Tuning - Validation Method -	Day 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

To predict the credit approval status of requested customers, we used a machine learning approach to look at bank records. We applied various machine learning algorithms to determine the best algorithm to apply to the dataset to obtain results with the highest accuracy. Following this approach, apart from logistic regression, we found the rest of the algorithms to be satisfactory in terms of accuracy reporting. The remaining algorithms had an accuracy range of 75% to 85%. On the other hand, logistic regression gave the best accuracy of (88.70%) among all algorithms compared.

We also identified the main characteristics that influence the approval status of a loan. 4468 We then apply these most important features to a few selected algorithms and compare the accuracy of their performance with all features. This model helps the bank identify the factors that are important to his loan approval process. Comparative studies reveal which algorithm is the best and the rest are ignored based on accuracy.

## 10. ADVANTAGES AND DISADVANTAGES

## **ADVANTAGES**

Advantages of Loan Tech Servicing Software to lenders include:

The rental system has too many variables and is error prone. However, the best his loan management software is designed to eliminate errors completely and this is definitely an asset of in every respect.

## 1. Preventing Late Payments

Inability to collect debt is particularly skeptical of most lenders. But if they are using a traditional credit management approach, they may not see it coming. On the other hand, the Loan Service System integrates analysis modules that can detect the most subtle fluctuations in customer confidence and prevent payment delays in a timely manner.

## 2. Time Savings

Managing loans requires a high degree of meticulousness and attention to detail. A full team is typically required to handle all aspects of the loan process. Needless to say, 's loan management is manual and paper-based, which is labor intensive. Meanwhile, the digital circulation system automates routines, allowing the team to devote his hours to other important tasks. READ ALSO: Use Trunk-Based Development for Product Agility Learn How to Deliver Your Products Easily and Fast in Real Time

## 3. Automated Reports One valuable feature.

Accounting, tax reports and invoices are frequently requested by regulators, borrowers and investors. These high priority reports must be provided upon request and contain 100% accurate information. Credit tracking software allows the lender to quickly generate various types of his reports and submit them urgently in the required format.

#### 4. Increased Revenue

This is the result of all of the above: Automated loan processing systems enable lenders to process more applications, allocate and manage more loans, detect fraud and defer They can be closed while 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

Organizations wishing to create rental software may not have sufficient infrastructure capacity onsite to ensure uninterrupted operations, updates, and support. Scaling peak load hours and keeping up with increasing numbers of users and subscriptions can also be very challenging. For optimal scalability and availability, it's best to use a cloud infrastructure.

## 2. Offers a Variety of Loan Types The more loan types

Money Lending Software can handle, the better. Lending apps with different use cases will definitely attract more users than apps that are only for specific types of loans. For example, the Loan Tech software that makes the loan app quotes has a wide range of uses, from technical calculators for student loans to his quotes for business loans and mortgages.

## 3. Centralized Data Storage

Customer data is used at every stage of the credit process. The best loan processing software stores this data in a central repository and can be accessed at any stage of the loan processing process. On the other hand, traditional loan management systems use a siled approach to data storage, making loan processing more tedious and time consuming.

#### 4. Built-in Credit

Check Functionality Modern credit management software for individual lenders must be able to instantly connect to credit bureaus and other entities responsible for credit checks. Such platforms should receive regularly updated her credit data and use big data analytics to assess the credibility of her applicants. For example, a customer's social media activity can be a valid source of another credibility score.

#### 5. Routine Process Automation

The use of robotic process automation to streamline simple rule-based processes is another key feature of credit management platforms. Automation speeds up loan origination and processing and helps improve customer satisfaction. Additionally, it helps avoid human error.

## 6. Built-in analytics module

The use of artificial intelligence (AI) and big data is another feature of the excellent loan management software for lenders. It not only helps in the preparation of reports, but also enables the companies to assess market trends, identify patterns of customer behavior and develop new products and offerings.

## 7. Third Party Integration

Another feature that most organizations find particularly attractive in a loan processing system is its ability to integrate with other business software. ERP and CRM solutions can enrich your credit system with data and insights. Systems that integrate credit modules with distance selling software are also popular with lenders.

## 8. Security

Financial Company's software deals with sensitive and sensitive data, and security is of utmost importance to both his lenders and customers. A good credit system should have advanced security features to ensure the highest level of customer, data and network protection.

## 11. CONCLUSION

During the debate about which supervised learning model to use for credit risk assessment, support vector machines outperform other treebased or regression models when the experimental design is similar to ours. I concluded that In addition, our model presented us in a discussion about which dimensionality reduction technique to use. For future improvements, more recent data and data from a variety of sources. We will use a data set similar to the previous experiment above to test this model and improve comparisons. It was mentioned that they abandoned the idea of using neural networks to reduce computational cost and complexity. However, we are looking forward to working with even larger amounts of data for the, so we also want to do a comparative analysis with the neural network. It is a well-known fact that neural networks tend to perform better on large datasets, and we would like to test this hypothesis in future work. As we also discuss the contribution of feature selection/extraction techniques, we have implemented other dimensionality reduction techniques, such as genetic algorithms, univariate feature selection methods, and tree-based feature selection, to measure their performance, and also to the lending sector. improve the efficiency of the paper can therefore conclude by stating that the model represents an interesting approach to identifying credit defaults in this ever-changing economy. Using the Lending Club dataset, our model yielded impressive results. It can play a key role in assessing the credit risk of borrowers, assisting lending institutions in making loans, and enabling financial institutions to operate transparently 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 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 proposes use such feature to 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 effectiveness perfectly show the 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.get(
       'se'), request.form.get('ai'), request.form.get('cai'), 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('C:/Users/BIJAY/Desktop/DEVELOPMENT/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.'
         a = 'Your Loan application will be Rejected.'
       if (lend_data == 1):
         a = 'This applicant can be provided with the loan amount requested.'
         a = 'Your Loan application will be succesfull.'
    return render_template('submit.html', num=a)
if __name__ == '__main__':
  app.run(debug=True, threaded=False)
```

#### HOME.HTML

```
k rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">
<!-- < link rel="stylesheet" href="static/css/font-awesome.min.css" type="text/css">
k rel="stylesheet" href="static/css/magnific-popup.css" type="text/css">
link rel="stylesheet" href="static/css/STYLES.css" type="text/css">
  <div class="teams">
    <div class="logo">
      <a href="#">LOAN PREDICTOR ($) </a>
      <a href="#">INFO</a>
      <a href="#">ABOUT</a>
      <li><a href="#">SERVICES</a></li>
      <a href="#">DETAILS</a>
  <div class="img"></div>
    <div class="center">
           <div class="title">WELCOME TO LOAN PREDICTION</div>
           <div class="sub_title">Find your Loan Eligibility here</div>
           <div class="btns">
             <a href="{{ url_for('predict') }}" class="tm-nav-link primary-btn" data-hover="Loan Predictor">
                <button> Click to Check</button>
nav {
  background-position: fixed;
  background-color: rgb(107, 144, 192);
  width: 100%;
```

```
padding: 12px 0;
 nav .teams {
    max-width: 1250px;
    margin: auto;
    display: flex;
    align-items: center;
    justify-content: space-between;
    padding: 0 20px;
  .teams .logo a {
    text-decoration: none;
    color: rgb(10, 5, 5);
    font-size: 30px;
    background: white;
    padding: 6px 12px;
    border-radius: 30px;
    font-family: Verdana, Geneva, Tahoma, sans-serif;
    font-weight: 530;
 .teams ul {
    display: inline-flex;
 .teams ul li {
    list-style: none;
    margin-left: 8px;
 .teams ul li a {
    text-decoration: none;
    color: rgb(0, 0, 0);
    font-size: 15px;
    font-family: system-ui, -apple-system, BlinkMacSystemFont, 'Segoe UI', Roboto, Oxygen, Ubuntu, Cantarell,
'Open Sans', 'Helvetica Neue', sans-serif;
    font-weight: 500;
    background: white;
    padding: 8px 15px;
    border-radius: 30px;
    transition: all 0.3s ease;
 .teams ul li a:hover {
    background: rgb(0, 16, 49);
    color: white;
  .img {
```

```
background: url('https://static.vecteezy.com/system/resources/previews/005/085/280/original/a-new-startup-has-
successfully-applied-for-a-loan-from-a-bank-entrepreneur-successful-business-negotiations-and-approved-loan-free-
vector.jpg') no-repeat;
    height: 95vh;
    width: 100%;
    background-size: cover;
    background-position: center;
    position: relative;
  .img::before {
    content: ";
    position: absolute;
    height: 100%;
    width: 100%;
  .center {
    position: absolute;
    top: 58%;
    left: 50%;
    width: 100%;
    transform: translate(-50%, -50%);
    padding: 0 20px;
    text-align: center;
  .center .title {
    color: rgb(0, 20, 31);
    font-size: 55px;
    font-weight: 900;
    font-family: 'Gill Sans', 'Gill Sans MT', Calibri, 'Trebuchet MS', sans-serif;
  .center .sub_title {
    color: rgb(5, 21, 65);
    font-size: 18px;
    font-weight: 900;
     font-family: 'Gill Sans', 'Gill Sans MT', Calibri, 'Trebuchet MS', sans-serif;
  .center .btns {
    margin-top: 20px;
  .center .btns button {
    height: 45px;
    width: 150px;
    border-radius: 25px;
    border: none;
    margin: 0 10px;
    border: 1px solid white;
```

font-size: 22px; font-weight: 500;

```
font-family: fantasy;
padding: 0 10px;
cursor: pointer;
outline: none;
transition: all 0.3s ease;
}

.center .btns button:first-child {
   color: rgb(8, 8, 8);
   background: rgb(249, 250, 249);
}

.center .btns button:first-child:hover {
   background: rgb(37, 71, 221);
   color: rgb(236, 247, 247);
}

</style>
```

#### PREDIT..HTML

```
!DOCTYPE html>
 <html lang="en">
  <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 href="https://fonts.googleapis.com/css2?family=Play:wght@400;700&display=swap" rel="stylesheet">
   kref="https://fonts.googleapis.com/css2?family=Josefin+Sans:wght@300;400;500;600;700&display=swap"
rel="stylesheet">
  link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">
  k rel="stylesheet" href="static/css/STYLES.css" type="text/css">
    <div class="container">
    <div class="row">
    <h1 id="title">Loan Eligibility checklist</h1>
    <le>egend>Please fill in your details here</legend>
         <form action="/result" method="POST" enctype="'multipart/form-data">
```

```
Are You a Lender / Loan Applicant</hd>
          <input id="lend" name="lend" type="radio" value=1
required> Lender  <input id="lend" name="lend" type="radio" value=0>&nbsp;Loan
Applicant
          Name (in Caps)</hd>
          <input name='name' type="text" required="">
           <h4>Gender (Male/Female)</h4>
          <input id="gender" name="gender" type="radio" value=1
required> Male  <input id="gender" name="gender" type="radio" value=0>&nbsp;Female
          <h4>Married(Yes/No)</h4>
          <input id="married" name="married" type="radio" value=1
required> Yes  <input id="married" name="married" type="radio" value=0>&nbsp;No
           <h4>Dependents (Enter a number)</h4>
           <input name='dep' type="number" min="0" step='1' placeholder="" required="">
          <h4>Education</h4>
           <input id='edu' name='edu' type="radio" value=1 required>&nbsp;Non-
Graduate  <input id="edu" name="edu" type="radio" value=0>&nbsp;Graduate
          Self Employed (Yes/No)</hd>
          <input id="se" name="se" type="radio" value=1
required> Yes  <input id="se" name="se" type="radio" value=0>&nbsp;No
           <h4>Applicant Income </h4>
           <input id='AI' name='ai' type="number" min='0' required="">
          <h4>Co-Applicant Income </h4>
          <input id='CAI' name='cai' type="number" min='0' required="">
          <h4>Loan Amount </h4>
          <input id='la' name='la' type="number" min='0' required="">
          Loan Amount Term (In days)</hd>
           <input id='lat' name='lat' type="number" min='0' step="1" required="">
           <h4>Credit History (Yes/No)</h4>
          <input id="ch" name="ch" type="radio" value=1
required> Yes  <input id="ch" name="ch" type="radio" value=0>&nbsp;No
```

```
Property Area</h4>
             <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>&nbsp;Semi
Urban  <input id="pa" name="pa" type="radio" value=0>&nbsp;No Property
             <input class="primary-btn-1" id="submit" type="submit" value="Click to Check">&emsp;<input</pre>
class="primary-btn-1" id="reset" type="reset" value="Clear">
    @import
url(https://fonts.googleapis.com/css2?family=Exo:ital,wght@1,300&family=Rubik:wght@300&display=swap');
    *::before,
    *::after {
        font-family: 'Rubik', sans-serif;
        color: white;
        box-sizing: border-box;
    body {
        background: url("https://img.freepik.com/free-vector/bank-credit-finance-management-loan-agreement-
signing-mortgage-money-credit_335657-
3136.jpg?w=1380&t=st=1668591228~exp=1668591828~hmac=c22cafbc069cb7d25be2c9be6636a081239e7a9ca2fb6
6e0e6b7cd1109882eca"),
             linear-gradient(60deg, #000000, #abd4eee7);
        background-size: cover;
        background-repeat: no-repeat;
        background-blend-mode: overlay;
        background-position: center;
    h1 {
        font-weight: 600;
        line-height: 2.6;
        font-size: 1.65rem;
```

```
p {
    text-align: center;
    margin-top: 0;
    margin-bottom: 0.5rem;
    text-decoration: none;
label {
    line-height: 2.5rem;
    font-size: 25px;
input {
    color: black;
    font-weight: 600;
input[type="checkbox"] {
    min-height: 1.1rem;
    min-width: 1.1rem;
input[type="radio"] {
    min-height: 1.3rem;
    min-width: 1.3rem;
option {
    color: black;
    font-weight: 600;
    text-align: start;
    width: 100%;
    height: 35px;
    border-radius: 5px;
option {
    font-size: 1.013rem;
textarea {
    min-width: 100%;
    min-height: 150px;
    color: black;
    font-weight: 600;
    border-radius: 5px;
```

```
textarea::placeholder {
    padding-left: 10px;
    font-size: 1.031rem;
textarea:focus {
    outline: 3px solid rgb(0, 15, 255, 0.6);
select:focus {
    outline: 3px solid rgb(0, 15, 255, 0.6);
select {
    font-size: 1.013rem;
    color: grey;
    padding-left: 13px;
#survey-form {
    background: rgb(0, 0, 0, 0.7);
    min-height: 180vh;
    margin: 10vh auto;
    max-width: 750px;
    border-radius: 10px;
    display: flex;
    flex-direction: column;
    align-items: center;
    justify-content: space-around;
#content {
    width: 90%;
    min-height: 140vh;
    margin: 1.13rem auto;
    display: grid;
    grid-template-columns: 100%;
    grid-gap: 1.5rem;
    justify-items: center;
    border-radius: 10px;
.form-element {
    width: 90%;
    display: block;
.form-con {
    width: 100%;
    height: 35px;
    border: none;
    border-radius: 5px;
```

```
.form-con::placeholder {
     font-size: 1.013rem;
     padding-left: 15px;
     font-weight: 600;
.form-con:focus {
     outline: 3px solid rgb(0, 15, 255, 0.6);
.long {
     line-height: 1.6rem;
.check {
     min-height: 1.013rem;
     min-width: 1.013rem;
#submit {
     background: linear-gradient(-60deg, #ff0000, #da2727ba, #2658ccba, #150a8ae7);
     width: 90%;
     border-radius: 5px;
     min-height: 2.031rem;
     border: none;
     font-weight: 600;
     letter-spacing: 0.1rem;
     cursor: pointer;
     margin-bottom: 6vh;
     font-size: 1.031rem;
     background-size: 250% 250%;
     animation: hover-effect 5s ease infinite;
#submit:hover {
     outline: 3px solid rgb(0, 15, 255, 0.6);
#submit:focus {
     animation: hover-effect 5s ease infinite;
     outline: 3px solid rgb(0, 15, 255, 0.6);
textarea,
select {
     margin-top: 1vh;
@keyframes hover-effect {
     0% {
          background-position: 0% 50%;
```

```
50% {
         background-position: 100% 50%;
    100% {
         background-position: 0% 50%;
@media(max-width:900px) {
    #survey-form {
         max-width: 650px;
    #content {
        margin: auto;
        max-width: 500px;
    #submit {
         max-width: 500px;
         margin: 7vh 0;
@media(max-width:500px) {
    .long {
         line-height: 2.0rem;
    h1 {
         font-size: 1.7rem;
    p {
         font-size: 1.3rem;
    textarea,
    select {
         margin-top: 8px;
         font-size: 1.031rem;
         margin: auto 0;
```

#### SUBMIT.HTML

```
!DOCTYPE html>
 <html lang="en">
  <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>
  k rel="icon" href="static/images/money.png" type="image/x-icon" style="border-radius: 5px;">
  k 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"</pre>
rel="stylesheet">
  link rel="stylesheet" href="static/css/bootstrap.min.css" type="text/css">
  link rel="stylesheet" href="static/css/STYLES.css" type="text/css">
    <div class="container">
      <div class="row">
       <h1>Loan Approval Prediction</h1>
       <h3>{{num}}}</h3>
    font-family: sans-serif;
    margin: 0;
    padding: 0;
    font-size: 40;
  body,
  html {
    width: 100%;
    height: 100vh;
    display: flex;
     align-items: center;
    justify-content: center;
```

```
background-image: url('https://img.freepik.com/premium-vector/mortgage-loan-debt-instruments-that-are-secured-by-property-assets-such-as-real-estate-services-rent-buying-home-auction-house-background-vector-illustration_2175-1342.jpg?w=1060');

background-position: center;
background-repeat: no-repeat;
background-size: cover;
background-color: rgb(0, 0, 0, 0.7);
background-blend-mode: overlay;
}

h1 {
    text-align: center;
    color: rgb(166, 219, 255);
    font-size: 3.5rem;
}

h2 {
    color: white;
    font-size: 2.5rem;
}
```

### **GITHUB LINK:**

https://github.com/IBM-EPBL/IBM-Project-26307-1660024383

## Project Demo Link:

https://www.youtube.com/watch?v=2oZFQvaPUeU