

SMART LENDER- APPLICANT CREDIBILITY PREDICTION FOR LOAN APPROVAL

Team ID: PNT2022TMID34269

Team Members: Pon Malar J, Poornima S, Prathiba P, Chonali Nachiyar M

1.INTRODUCTION

1.1. PROJECT OVERVIEW

Loan Prediction is very helpful for employee of banks as well as for the applicant also. The aim of this Paper is to provide quick, immediate and easy way to choose the deserving applicants. Dream housing Finance Company deals in all loans. They have presence across all urban, semi urban and rural areas. Customer first apply for loan after that company or bank validates the customer eligibility for loan. Company or bank wants to automate the loan eligibility process (real time) based on customer details provided while filling application form. These details are Gender, Marital Status, Education, Number of Dependents, Income, Loan Amount, Credit History and other. This project has taken the data of previous customers of various banks to whom on a set of parameters loan were approved. So the machine learning model is trained on that record to get accurate results. Our main objective of this project is to predict the safety of loan. To predict loan safety, the SVM and Naïve bayes algorithm are used. First the data is cleaned so as to avoid the missing values in the data set.

1.2. PURPOSE

The project report is the basic document for analyzing the future performance of a component and is a very important document for forming an opinion about the future of a company or business component.

1. To obtain capital loans, term loans and other loans from a bank or financial institution.
2. To make a presentation to get investors' equity participation.
3. To structure / restructure the firm's bank loan / financial and business policies.
4. To buy, own or start a new business.

5. for proper disposal of existing business.

6. To evaluate the value of a project or business

2.LITERATURE SURVEY

2.1. EXISTING PROBLEM

Bank data management on loan approval processes has great room for improvements of information quality and data problems prevention especially with regards to fair lending and fair pricing practices. They first reviewed briefly typical data collection protocols deployed at many financial institutions for loan approval and loan pricing. Federal regulations mandate portions of these data protocols. While discussing the data capture and analysis for fair lending, they illustrated some initial key steps currently needed for improving information quality to all parties involved.

With the enhancement in the banking sector, lots of people apply for bank loans but the bank has its limited assets which it grants to only limited people , so finding out to whom the loan can be granted is a typical process for the banks. So, in this paper , they tried to reduce this risk by selecting the safe person so as to save lots of bank efforts and assets.

2.2. REFERENCES

| Author | Title | Source | Findings |
|----------------------------------|--|---|---|
| M.Cary Collins (2013) | Improving Information Quality in loan Approval processes for Fair Lending & Fair Pricing | Ijrasnet Journal for Research in Applied Data Science &Engineering Technology | They illustrated some initial key steps currently needed for improving information quality to all parties involved. |
| Sivasree M S, Rekha Sunny (2015) | Loan credibility Prediction System Based on Decision Tree Algorithm. | Asian Journal of Applied Science and Technology (AJAST) | In this model ,they introduced a prediction model for bankers that will help them predict the credible customers who have applied for a loan. |

| | | | |
|--|---|---|---|
| Kumar Arun,Garg Ishan ,Kaur Sanmeet (2016) | Loan Approval Prediction based on Machine Learning Approach | IEEE | It was done by mining the previous records of the people to whom the loan was granted before on the basics of these records the machine was trained using this model. |
| Goyal and Kaur (2018) | Loan Prediction Procedures | International Journal of Current Engineering and Scientific Research (IJCESR) | The sub process in this method includes data collection ,filtering the data,feature extraction,applying the model and finally analysis the result. |

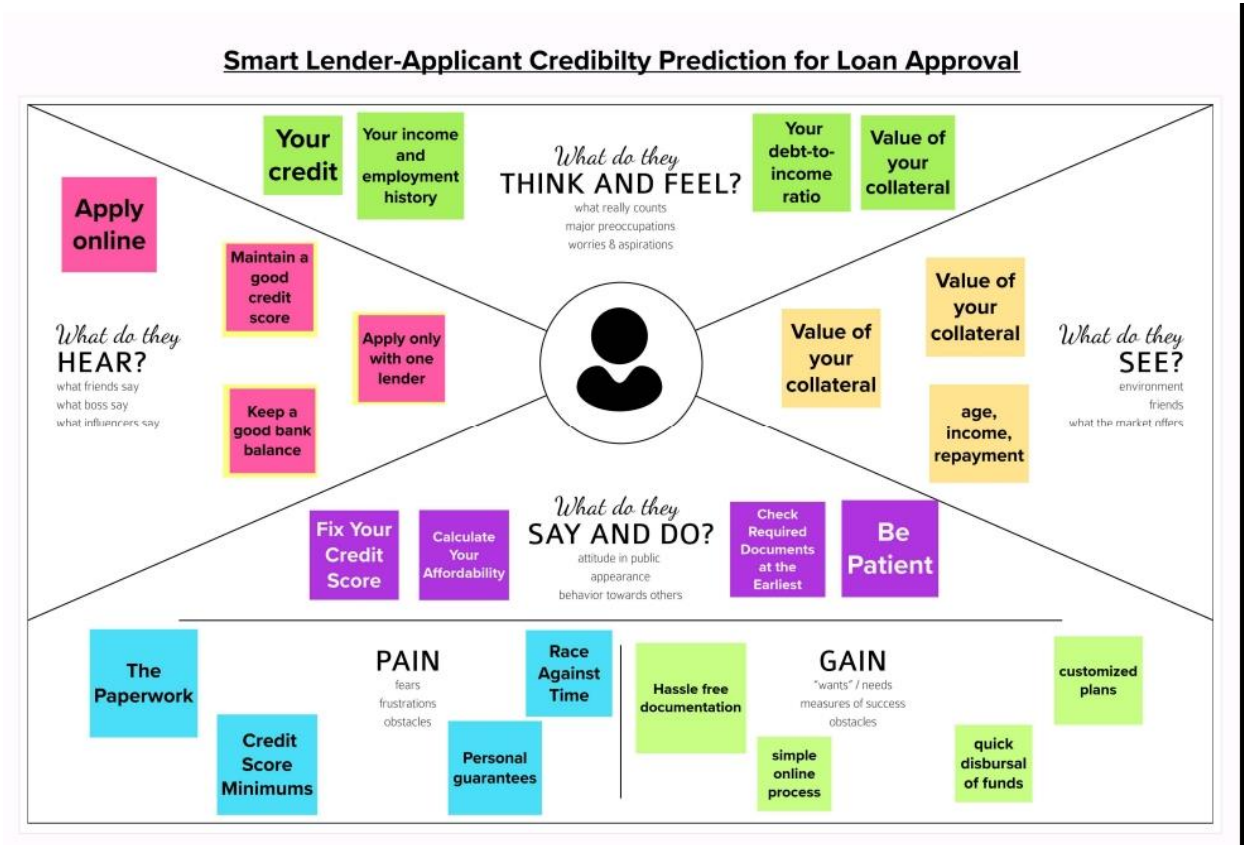
2.3.PROBLEM STATEMENT DEFINITION

- Loan prediction is a very common real life problem that each bank faces at least once in its lifetime.
- It can save a lot of man hours at the end of a retail bank.
- It is specifically built to give you walkthrough of the loan prediction problem.

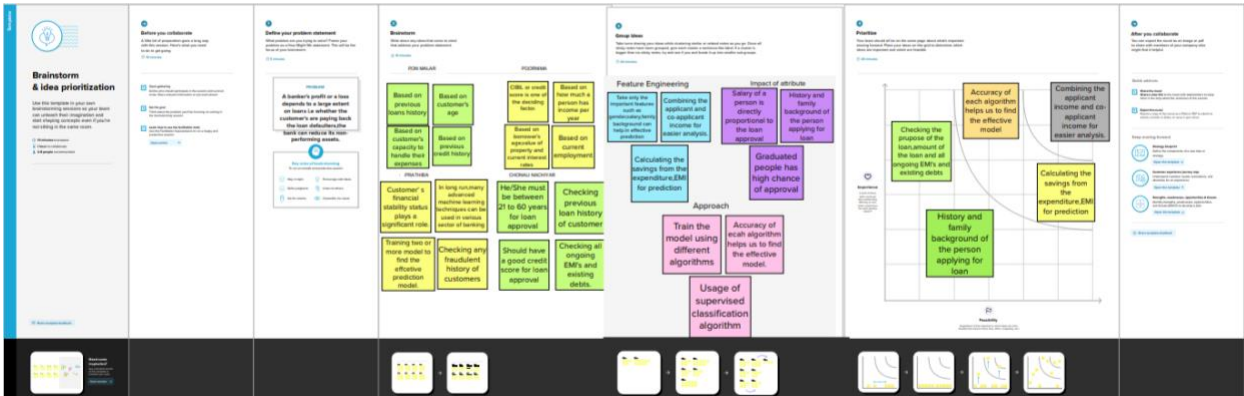
3.IDEATION& PROPOSED SOLUTION

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.

3.1.EMPATHY MAP CANVAS



3.2.IDEATION& BRAINSTORMING



3.3.PROPOSED SOLUTION

| S.NO | PARAMETER | DESCRIPTION |
|------|-----------|-------------|
| . | | |

| | | |
|-----------|---|--|
| 1. | Problem Statement (Problem to be solved) | <p>The credit system governed by the banks is one of the most important factors which affect our country's economy and financial condition. Also, 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. This problem occurs when the banks need to provide loans to the customers who are in need of the money. But by forecasting the loan defaulters, the banks definitely may reduce their loss by reducing their non-profit assets.</p> |
| 2. | Idea /Solution description | <p>This solution uses Machine Learning techniques which can be used to perform such classifications of the credit defaulters as they are very crucial and useful in the prediction of these types of data. Classification algorithms such as Decision tree, Random forest, KNN, Xgboost and SVM will be used. The data is trained and tested with these algorithms and finally, the best model is selected and saved in pkl format. Then, flask integration and IBM deployment will be done.</p> |
| 3. | Novelty /Uniqueness | <p>The solution tries to use the best model from the mentioned five models and classify the applicants with least error.</p> |
| 4. | Social Impact /Customer Satisfaction | <p>This application will help the bank employees to classify the credit defaulters accurately with minimum error. So, the non - profit losses of the banks have been reduced. Thus, they may recover the approved loans with minimum losses.</p> |

| | | |
|-----------|---------------------------------------|--|
| 5. | Business Model (Revenue Model) | The model can be implemented as a pay per month use model. The bank employees can pay the monthly or yearly subscription. Another option is to sell the model to the bank that pays the amount which is most profitable to developers. |
| 6 | Scalability of the Solution | The front end of the application is modular. Python Web Framework is used to do so. The bank end uses the flask integration. Therefore different features can be implemented and new pages can be added easily. |

3.4.PROBLEM SOLUTION FIT

Project Title: Smart lender –applicant credibility for loan prediction

Project Design Phase-I - Solution Fit Template

Team ID: PNT2022TMD34269

| | | | | |
|------------------------|---|---|---|---------------------------|
| Define CS, fit into CC | 1. CUSTOMER SEGMENT(S) CS Our target customers are mostly banking firm, small financial firms that lends out loan and credit card companies because of the increasing rate of loan defaulter and also to increase the slow process of the loan approval. | 6. CUSTOMER CONSTRAINTS CC Banks are not to correctly handle the loan request. People within a protected class being clearly treated differently than those of non-protected classes for loan. There is an increasing rate of loan defaults. Banks identify the loan defaulters for much-reduced credit risk as large portions of a bank's assets directly come from the interest earned on loans given. | 5. AVAILABLE SOLUTIONS AS <ul style="list-style-type: none"> Random forest, Logistic regression, Decision tree and Naive bayes algorithm are used Using data pre-processing data mining and data filtering Algorithms such as naïve bayes, k-nearest neighbors are used. | Explore AS, differentiate |
| | 2. JOBS-TO-BE-DONE / PROBLEMS J&P Needs to Support genuine Entrepreneur. That the process should be easier a time saving. To find an applicant which can give best interest. Needs to find a loan applicant with good credit score | 9. PROBLEM ROOT CAUSE RC The root cause of this problem is the banks identify the loan defaulters for much-reduced credit risk as large portions of a bank's assets directly come from the interest earned on loans given. . People within a protected class being clearly treated differently than those of non-protected classes for loan. | 7. BEHAVIOUR BE Directly related: The customers who lends the loan and the banks that checks the credibility seek to do the process faster. Indirectly associated: The small finance sector that deals with middle class and poor class people seek to find the credibility. | |

| | | | | |
|-------------------------|--|--|---|-------------------------|
| Identify strong TR & EM | 3. TRIGGERS TR The slow and complex process of loan approval is affecting the business of our customer and it also decline the revenue of our customers. Due to the sudden surge in the number of loan defaulters our customers business is highly affected. | 10. YOUR SOLUTION SL <ul style="list-style-type: none"> There is an increasing rate of loan defaulters and banks are not able to correctly handle the loan request. To avoid this problem a machine learning algorithm is developed The system automatically selects the credible candidates to approve the loan and it will improve the speed, efficacy, and accuracy of loan approval processes. This help the user(Lender) to accurately identify whom to lend the loan and also help the banks to identify the loan defaulter for much-reduced credit risk. | 8. CHANNELS OF BEHAVIOUR CH ONLINE: The customers needs to check the credibility of the client in an online mode. OFFLINE: The customer need to install the Machine Learning algorithm in their system to work efficiently. | Identify strong TR & EM |
| | 4. EMOTIONS: BEFORE / AFTER EM Before: Needs to Support genuine Entrepreneur. That the process should be easier a time saving. To find an applicant which can give best interest. Needs to find a loan applicant with good credit score. After: After implementing this project people can be able to face all these above-mentioned problems easily | | | |

4.REQUIREMENT ANALYSIS

4.1.FUNCTIONAL REQUIREMENTS

| FR No. | Functional Requirement (Epic) | Sub Requirement (Story / Sub-Task) |
|--------|-------------------------------|---|
| FR-1 | Home Page | <ul style="list-style-type: none">• Smart Lender Applicant Credibility description• Information about Credibility details required for loan approval• if new user , REGISTER• if already exists, SIGN IN |
| FR-2 | User Registration | Enter Mail Id and other personal details required for Registering |
| FR-3 | User login | User Mail Id and Password for Login |
| FR-4 | Loan Approval form | Credibility details should be entered for prediction |
| FR-5 | Result | if Approved - It display the information about what is done to be next. if Not Approved - It display the information about what rejection criteria you are not eligible for the loan. |

4.2.NONFUNCTIONAL REQUIREMENTS

| FR No. | Non-Functional Requirement | Description |
|--------|----------------------------|---|
| NFR-1 | Usability | If the customer is eligible,he/she should be able to receive the acknowledgement receipt for loan application within 7 days from the bank.If not,then he/she should be intimated with the reason for rejection. |
| NFR-2 | Security | Check if the customer has any fraudulent history. |
| NFR-3 | Reliability | Customer's financial status plays a crucial role.So the customer should have a good credit score. |
| NFR-4 | Performance | By training the model using different ML algorithms ,the performance of the system can be increased. |
| NFR-5 | Availability | The loan will be available easily to those persons who have high income and to those who assure to repay the high sum within short period of time. |
| NFR-6 | Scalability | The customer should be between 21 to 60 age.And based on customer's capacity to handle their expenses. |

5.PROJECT DESIGN

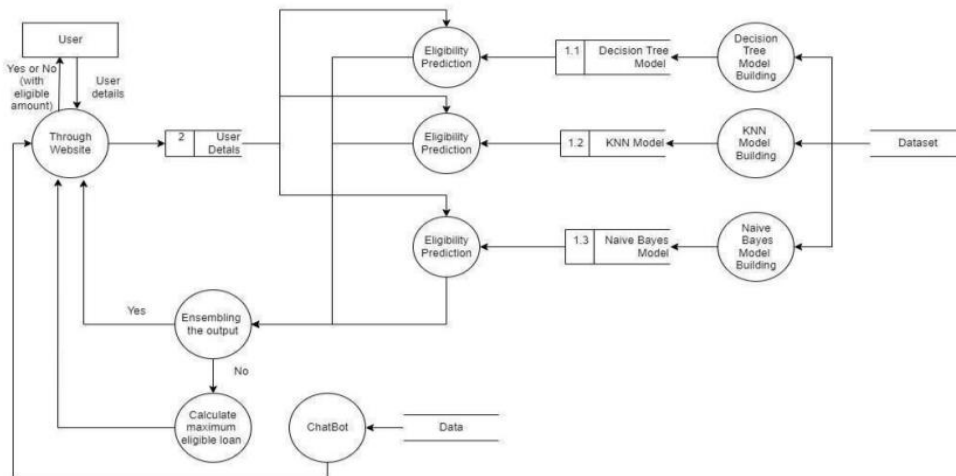
5.1.DATAFLOW DIAGRAM

Borrowing Process

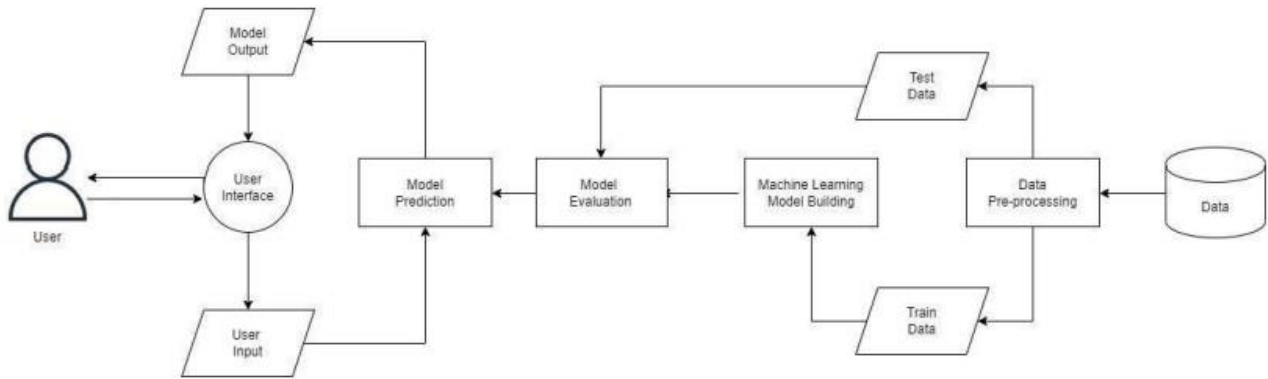


DFD Level 2 (Industry Standard)

DFD - LEVEL 2








5.2 SOLUTION & TECHNICAL ARCHITECTURE:



5.3.USER STORIES

| 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. | I can access my account /dashboard | High | Sprint-1 |
| | | 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 | Sprint-1 |
| | | USN-3 | Automatically, the information are stored in the Customer database | | Low | Sprint-3 |
| | | USN-4 | As a user, I can register for the application through Gmail | | Medium | Sprint-2 |
| | Login | USN-5 | As a user, I can log into the application by entering email & password | I can login to the application using login credentials | High | Sprint-1 |
| Customer (Webuser) | | USN-6 | After logging in, I have to select loan chosen, then the loan offer list is displayed from the Loan Offers Database. | I have select the types of loan | High | Sprint-1 |
| | | USN-7 | The payment information has to be given then the monitor payments stores the information. | | Medium | Sprint-2 |
| | | USN-8 | The record payments are stored in the payment database | | Low | Sprint-3 |
| Administrator | | USN-9 | The web application gives the loan status and the updates the loan to the customer and the updated loan is also stored in Loan Details in Database | Then I have shown with the loan status | High | Sprint-1 |

Creating a user journey is a quick way to help you and your team gain a deeper understanding of who you're designing for, aka the stakeholder in your project. The information you add here should be representative of the observations and research you've done about your users. 

| Phases | Awareness | Consideration | Service | Loyalty |
|--|---|--|--|---|
| High-level steps your user needs to accomplish from start to finish. | | | | |
| Steps | View online ad, see social media campaign, hear about from friends about benefits | Conduct feedback session for customers, compare features and benefits of loan | Loan eligibility check and comparing with other documentation | Make additional benefits and approval of higher amount |
| Detailed actions your user has to perform. | | | | |
| Feelings |  |  |  |  |
| What your user might be thinking and feeling at the moment. | In case the user does not use the social media it must be hard. | If the income less the expected value then loan prediction will detect only for low amount | You are offline the application does not show the any information. | Sometimes to difficult the predict the amount based on income. |
| | | | | |
| Pain points | - Is not aware of all loan benefits - Doesn't know what to choose | - Doesn't know where to start - Doesn't want to spend a lot of time on research | - Hard to get response - Buffering issues | - No discount - Not enough other incentives |
| Problems your user runs into. | | | | |
| Opportunities | Increase awareness interest, marketing on loan ,communications- awarness | Customer weights multiple offerings that could solve the problem | Respond to customer inquiries and concerns in a timely manner to improve experiences | Reward long-term loyalty to keep your customer wanting more |
| Potential improvements or enhancements to the experience. | | | | |

Share your feedback

6.PROJECT PLANNING &SCHEDULING:

6.1 Sprint Planning& Estimation:

Activity List:

In Project Management Planning is an important task to scheduling the phrase of the project to the Team Member.

In this Activity can shows the various activity are allocated and Done by the Team Members! In Project we can Split into the Four Step of Phrases are

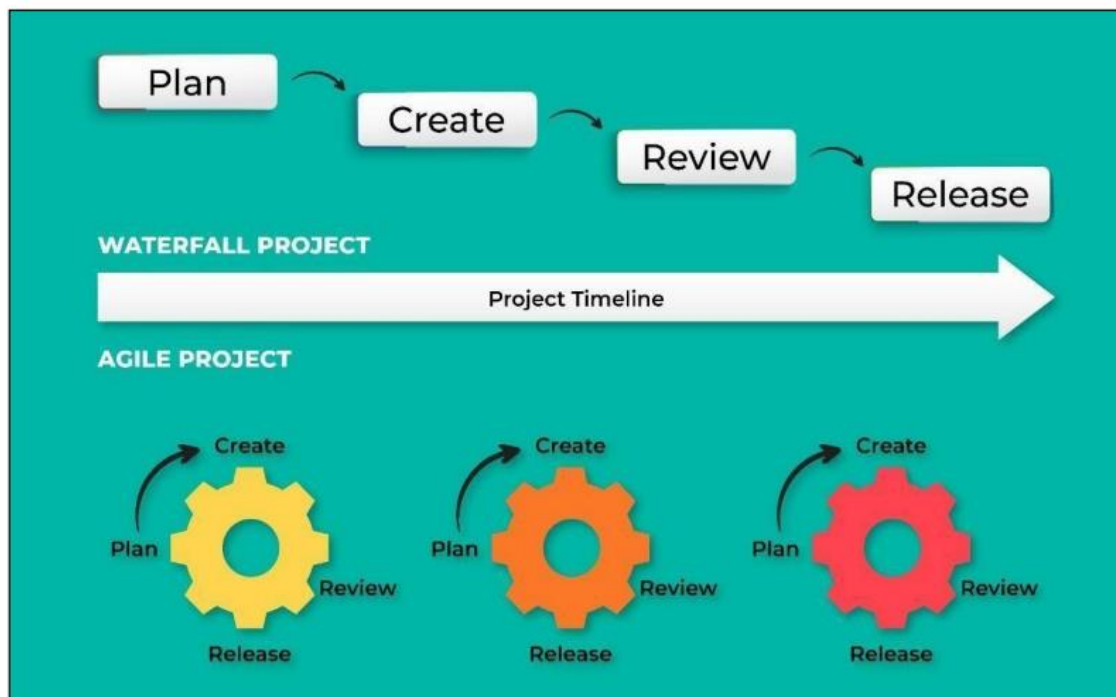
Phrase 1: Information Collection and Requirement Analysis

Phrase 2: Project Planning and Developing Modules

Phrase 3: Implementing the High Accuracy Deep Learning Algorithm to Perform

Phrase 4: Deploying the Model on Cloud and Testing the Model and UI Performance

Agile Methodology For Activity Planning



6.2 Sprint Delivery Schedule:

| Release | Functional Requirement (Epic) | User Story Number | User Story / Task | Story Points | Priority | Team Members |
|------------|------------------------------------|-------------------|---|--------------|-----------|---|
| Sprint - 1 | Forms | USN - 1 | As a user, I can enter the data which I have and also the data which the website asks to me | 6 | Very High | Pon Malar. J, Poornima. S Prathiba. P, Chonali Nachiyar. M |
| Sprint - 3 | Prediction | USN - 2 | As I have given the data into the webpage, now the data can be predicted for the loan avail | 4 | Medium | Pon Malar. J, Poornima. S Prathiba. P, Chonali Nachiyar. M |
| Sprint - 4 | Deployment of the Webpage in Cloud | USN - 3 | As a user, I require global access to the web page as a user | 3 | Low | Pon Malar. J, Poornima. S Prathiba. P, Chonali Nachiyar. M |

| | | | | | | |
|------------|-------------------------------------|---------|---|---|--------|---|
| Sprint - 4 | Deployment of AI model in the cloud | USN - 4 | Model would be running on the Cloud | 3 | Low | Pon Malar. J, Poornima. S Prathiba. P, Chonali Nachiyar. M |
| Sprint - 2 | Model building | USN - 5 | I require an ML model that can categorise Credit defaulters | 5 | High | Pon Malar. J, Poornima. S Prathiba. P, Chonali Nachiyar. M |
| Sprint - 3 | User Interface building | USN - 6 | As a User, I need a medium to enter my data | 4 | Medium | Pon Malar. J, Poornima. S Prathiba. P, Chonali Nachiyar. M |

Velocity:

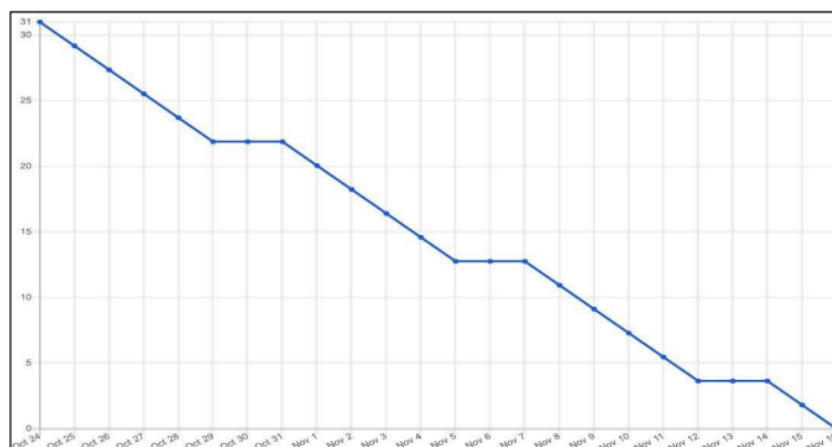
Imagine we have a 10-day sprint duration, and the velocity of the team is 20 (points per sprint). Let's calculate the team's average velocity (AV) per iteration unit (story points per day)

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

6.3 Reports from JIRA:

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.



7.1.FEATURE 1

home.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" />
    <link rel="stylesheet" href="style.css" type="text/css" />
    <title>Loan Predictor</title>
  </head>
  <body>
    <main>
      <div class="mail">
        <center>
          <h1>Loan Prediction</h1>
          <h3>Find your Loan Eligibility</h3>
          <h5>
            Click below button and fill the details to know your Loan
            Eligibility.
          </h5>
          <div class="container">
            <a href="index.html">
              <button class="btn" data-hover="Loan Predictor">
                <div>Click to Check</div>
              </button>
            </a>
          </div>
        </center>
      </div>
    </main>
```



```
</body>
```

```
</html>
```

```
index.html
```

```
<!DOCTYPE html>
```

```
<html>
```

```
<head>
```

```
<style type="text/css">
```

```
.header{
```

```
background: linear-gradient(pink);
```

```
background:-webkit-linear-gradient(pink);
```

```
background:-moz-linear-gradient(pink);
```

```
background:-o-linear-gradient(pink);
```

```
position: absolute;
```

```
left: 0px;
```

```
top: 0px;
```

```
width: 100%;
```

```
height: 100px;
```

```
text-align: center;
```

```
text-transform: capitalize;
```

```
}
```

```
.reg_img
```

```
{
```

```
height: 1000px;
```

```
margin-top: 0px;
```

```
background-image: url("../static/33.jpeg");
```

```
}
```

```
.box2
```

```
{
```

```
height: 960px;
```

```
width: 600px;
```

```
background-color: black;
```

```
margin: 70px auto;
```

```
opacity: .7;
color: white;
padding: 20px;
}

.
.reg{
margin-top: 130px;
width: 50%;
text-align: center;
text-decoration-style: smooth;
font-family: verdana;
color: white;
text-transform: capitalize;
background: rgb(10, 34, 56);
border-radius: 10px;
overflow: scroll;
}

.sco table{
text-overflow: scroll;
}
```

</style>

</head>

<head>

<title>Loan Registration using Machine Learning </title>

</head>

<body>

<header>

<div class="header", id="ls">

<h1>Smart Lender - Applicant Credibility Prediction for Loan Approval 💰</h1>

</div>

</header>

```
<div class="reg_img">
```

```
<div class="box2">
```

```
<h1 style="text-align: center; font-size: 25px;">APPLICANT DETAILS  
FORM</h1>
```

```
<form action="{{url_for('prediction')}}" method="POST">
```

```
<br><br>
```

```
<label>Select Gender</label>
```

```
<br><br>
```

```
<input type="radio" value="Male" name="gender">Male
```

```
<input type="radio" value="Female" name="gender">Female
```

```
<br><br>
```

```
<label>Status</label>
```

```
<br><br>
```

```
<input type="radio" value="Married" name="status">Married
```

```
<input type="radio" value="Single" name="status">single
```

```
<br><br>
```

```
<label>Enter number of dependants</label>
```

```
<br><br>
```

```
<input type="text" name="dependants", placeholder="3">
```

```
<br><br>
```

```
<label>Education Level</label>
```

```
<br><br>
```

```
<select name = "education">
```

```
<option value="Graduate" value="Graduate">Graduate</option>
```

```
<option value="Not Graduate" value="Not Graduate"> Not  
Graduate</option>
```

```
</select>
```

```
<br><br>
```

```
<label>Employment status</label>
```

```
<br><br>
```

```
<select name="employ">
```

```
<option value="No">Employed</option>
<option value="Yes"> Self Employed</option>
<option value="Yes">I do both</option>
<option value="No">None of above</option>
</select>
<br><br>
<label>Enter your annual income </label>
<br><br>
<input type="text" name="aincome", placeholder="5849">
<br><br>
<label>Enter your Coincome </label>
<br><br>
<input type="text" name="coincome", placeholder="0">
<br><br>
<label>Loan amount </label>
<br><br>
<input type="text" name="Lamount", placeholder="128">
<br><br>
<label>Loan amount Term </label>
<br><br>
<input type="text" name="Lamount_term", placeholder="360">
<br><br>
<label>Enter your credit history </label>
<br><br>
<input type="text" name="credit", placeholder="1">
<br><br>
<label>select your property area </label>
<br><br>
<select name = "property_area">
<option value="urban">urban</option>
<option value="Semiurban">Semi urban</option>
<option value="Rural">Rural</option>
```

```
</select>
<br><br>
<input type="submit" name="sumbit" value="Register">
</form>
</div>
```

```
</div>
<footer>
<div>
</div>
</footer>
</body>
</html>
```

output.html

```
<!DOCTYPE html>
<html>
<head>
<style type="text/css">
.header{
background: linear-gradient( white);
background:-webkit-linear-gradient( white);
background:-moz-linear-gradient( white);
background:-o-linear-gradient( white);
position: absolute;
left: 0px;
top: 0px;
width: 100%;
height: 100px;
text-align: center;
text-transform: capitalize;
}
.log_img
```

```
{
height: 650px;
margin-top: 0px;
background-image: url("../static/22.jpeg");
}

.app {
height: 500px;
width: 600px;
background-color: rgba(0, 0, 0, 0.945);
margin: 100px auto ;
opacity: .7;
color: white;
padding: 20px;
}

</style>

<title> Classification results</title>

</head>

<body>

<div>

<header class="header">

<h1>The Application result </h1>

</header>

</div>

<div class="log_img">

<div class="app">

<section>

<table align="center">

<th>

<h1> Your information has been Submitted Please wait </h1>

<table align="center">

<th>
```

```

{% if output== 1 %}

<br><br>

<h1 style="text-align: center; font-size: 35px;font-family: Lucida
Console;">Congratulations 👍👍👍 You are eligible for the Loan 💰💰💰</h1>



<br><br>

{% elif output==0 %}

<h1 style="text-align: center; font-size: 35px;font-family: Lucida Console;">We are sad
😞😞😞 to inform that you are not eligible for the Loan</h1>



{% endif %}

</th>

</table>

</th>

</table>

</section>

</div>

<div>

<footer>

</footer>

</div>

</div>

</body>

</html>

```

7.2.FEATURE 2

app.py

```

import flask

import joblib

import numpy as np

from flask import render_template, request

```

```

from flask_cors import CORS
app = flask.Flask(__name__)
CORS(app)
@app.route('/', methods=['GET'])
def sendHomePage():
    return render_template('index.html')
@app.route('/output', methods = ['POST'])
def prediction():
    if request.method == 'POST':
        gender = request.form['gender']
        married = request.form['status']
        dependat =request.form['dependants']
        education = request.form['education']
        employ = request.form['employ']
        annual_income = request.form['aincome']
        co_income = request.form['coincome']
        Loan_amount = request.form['Lamount']
        Loan_amount_term = request.form['Lamount_term']
        credit = request.form['credit']
        proper = request.form['property_area']
        gender = gender.lower()
        married= married.lower()
        education = education.lower()
        employ = employ.lower()
        proper = proper.lower()
        if(employ=='yes'):
            employ = 1
        else:
            employ = 0
        if(gender=='male'):
            gender = 1
        else:

```



```

gender = 0
if (married=='married'):
    married=1
else:
    married=0
if (proper=='rural'):
    proper=0
elif (proper=='semiurban'):
    proper=1
else:
    proper=2
if (education=='graduate'):
    education=0
else:
    education=1

dependat = int(dependat)
annual_income = int(annual_income)
co_income = int(co_income)
Loan_amount = int(Loan_amount)
Loan_amount_term = int(Loan_amount_term)
credit = int(credit)
x =np.array([[gender, married,
dependat,education,employ,annual_income,co_income,Loan_amount,Loan_amount_term
,credit
,proper]])
model = joblib.load('Forest.pkl')
ans = int(model.predict(x)[0])
if (ans==1):
    print("Congratulations your eligible for this Loan")
else:
    print("We are sad to inform that your request has not been accepted")

```

```
return render_template('output.html', output=ans)

if __name__ == '__main__':
    app.debug = True

    app.run()
```

8.TESTING

8.1.TEST CASES

For checking the loan application, We have two testcase

- Eligible
- Not Eligible

This is based on the training and testing the model we used in our application. This eligibility can be checked by using the details entered by the users. This includes the details like

- Gender
- Status
- Dependants
- Education
- Employ
- Income
- Co-income(additional income)
- Loan amount
- Loan amount term(in days)
- Credit history
- Property area(type of location)

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 project - Smart Lender - Applicant Credibility Prediction for Loan Approval 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 resolve

| Resolution | Severity 1 (High) | Severity 2 (Moderate) | Severity 3 (Low) | Subtotal |
|-------------------|------------------------------|----------------------------------|-----------------------------|-----------------|
| By design | 1 | 3 | 2 | 6 |
| Duplicate | 1 | 0 | 3 | 4 |
| External | 2 | 3 | 0 | 5 |
| Fixed | 4 | 6 | 4 | 14 |
| Not Reproduced | 0 | 0 | 1 | 1 |
| Totals | 8 | 12 | 10 | 30 |

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 | 6 | 0 | 0 | 6 |
| Client Application | 16 | 0 | 0 | 16 |
| Security | 2 | 0 | 0 | 2 |
| Exception Reporting | 3 | 0 | 0 | 3 |
| Final Report Output | 4 | 0 | 0 | 4 |
| Version Control | 1 | 0 | 0 | 1 |

9.RESULTS

9.1.PERFORMANCE METRICES

In our project we used XG-Boost model for prediction.

| S.No | Parameters | Values | Screenshot |
|------|----------------|---|------------|
| 1 | Metrics | Classification Model: Confusion Matrix,Accuracy Score-&Classification Report | Fig 1 |
| 2 | Tune the Model | Hyperparameter Tuning Validation Method | Fig 2 |

| | | | | | |
|--|-----------|--------|----------|---------|--|
| In [52]: xgboost(x_train, x_test, y_train, y_test) | | | | | |
| ****Gradient BoostingClassifier**** | | | | | |
| Confusion matrix | | | | | |
| [[74 29] | | | | | |
| [12 108]] | | | | | |
| Classification report | | | | | |
| | precision | recall | f1-score | support | |
| 0 | 0.86 | 0.72 | 0.78 | 103 | |
| 1 | 0.79 | 0.90 | 0.84 | 120 | |
| accuracy | | | 0.82 | 223 | |
| macro avg | 0.82 | 0.81 | 0.81 | 223 | |
| weighted avg | 0.82 | 0.82 | 0.81 | 223 | |
| Testing accuracy: 0.8161434977578476 | | | | | |
| Training accuracy: 0.9466666666666667 | | | | | |
| From the four model Xgboost is performing well. Xgboost is giving the accuracy of 94% with training data , 81% accuracy for the testing data.so we considering xgboost and deploying this model. | | | | | |

Fig 1

Evaluating Performance Of The Model

```
In [53]: from sklearn.model_selection import cross_val_score

In [54]: # Xgboost Model is selected
         xg = GradientBoostingClassifier()

In [55]: xg.fit(x_train,y_train)

Out[55]: ▾ GradientBoostingClassifier
         GradientBoostingClassifier()

In [56]: yPred = xg.predict(x_test)

In [57]: f1_score(yPred,y_test, average='weighted')

Out[57]: 0.8183313193520658

In [58]: cv = cross_val_score(xg,x,y,cv=5)

In [59]: np.mean(cv)

Out[59]: 0.7230974276955885
```

Fig 2

10.ADVANTAGES & DISADVANTAGES

• Advantages:

- Fast and highly accurate result
- Easy handling of the problem
- Less risk and more convenient to use
- Reliability is pretty high
- Better choice for responsive result
- Better user interface

• Disadvantages:

- Machine Learning model in general is little complex
- Prediction sometime not reliable because the model is build on the old data.
- The prediction result is more depend on the model.

11.CONCLUSION

We 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. The 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. We have developed a model which can easily predict that the person

will repay its loan or not. we can see our model has reduced the efforts of bankers. Machine learning has helped a lot in developing this model which gives precise results.

12.FUTURE SCOPE

Future enhancement of this research work on training bots to predict the loan eligibility areas by using machine learning techniques. Since, machine learning is similar to data mining advanced concept of machine learning can be used for better prediction. The data privacy, reliability, accuracy can be improved for enhanced prediction From the encouraging results, we believe that crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis.

Visual and intuitive criminal and intelligence investigation techniques can be developed for loan credibility pattern. As we have applied machine learning technique of data mining for loan prediction we can also perform other techniques of data mining such as classification. Also we can perform analysis on various dataset such as enterprise survey dataset, poverty dataset, aid effectiveness dataset, etc.

13.APPENDIX:

GitHub Link:<https://github.com/IBM-EPBL/IBM-Project-50687-1660921442>

Project Demo Link:<https://youtu.be/weTjx2JcsGQ>