

Project Report

title - Smart Lender - Applicant Credibility Prediction for Loan Approval

1. INTRODUCTION

a. Project Overview

To determine the success rate of lone approval.

b. Purpose

To find the exact ability of a person for loan eligible.

2. LITERATURE SURVEY

1. Existing problem

Anomaly detection relies on individuals' behaviour profiling and works by detecting any deviation from the norm. When it is used for online banking fraud detection, it suffers from three disadvantages. First, for an individual, the historical behaviour data are often too limited for profiling his/her behaviour pattern. Second, because of the heterogeneous nature of transaction data, there is no uniform treatment to various attribute values, which will become a potential barrier for development of the model and for further usage. Third, the transaction data are highly skewed, and it becomes a challenge for utilizing the label information effectively. Anomaly detection often suffers from poor generalization ability and a very high false alarm rate. We argue that individuals' limited historical data for behaviour profiling and fraud data's highly skewed nature could account for this defect. Since it is straightforward to use information from other similar individuals, similarity measurement itself becomes a great challenge due to heterogeneous nature of attribute values.

2. References

[1] Arun Kumar, Ishan Garg, and Sanmeer Kaur, "Loan Approval Prediction Using Machine Learning Approach," 2018. [2] K. Hanumantha Rao, G. Srinivas, A. Damodhar, and M. Vikas Krishna at International Journal of Computer Science and Telecommunications published an article titled "Implementation of Anomaly Detection Technique Using Machine Learning Algorithms" (Volume2, Issue3, June 2011).

3. Problem Statement Definition

They have a presence across all urban, semi-urban and rural areas.

The customer first applies for a home loan and after that, the company validates the customer eligibility for the loan. The company wants to automate the loan eligibility process (real-time) based on customer detail provided while filling out online application forms. These details are Gender, Marital Status, Education, and number of Dependents, Income, Loan Amount, Credit History, and others. To automate this process, they have provided a dataset to identify the customer segments that are eligible for loan amounts so that they can specifically target these customers. We need to predict our Target label which is “Loan Status”. Loan status can have two values: **Yes** or **NO**.

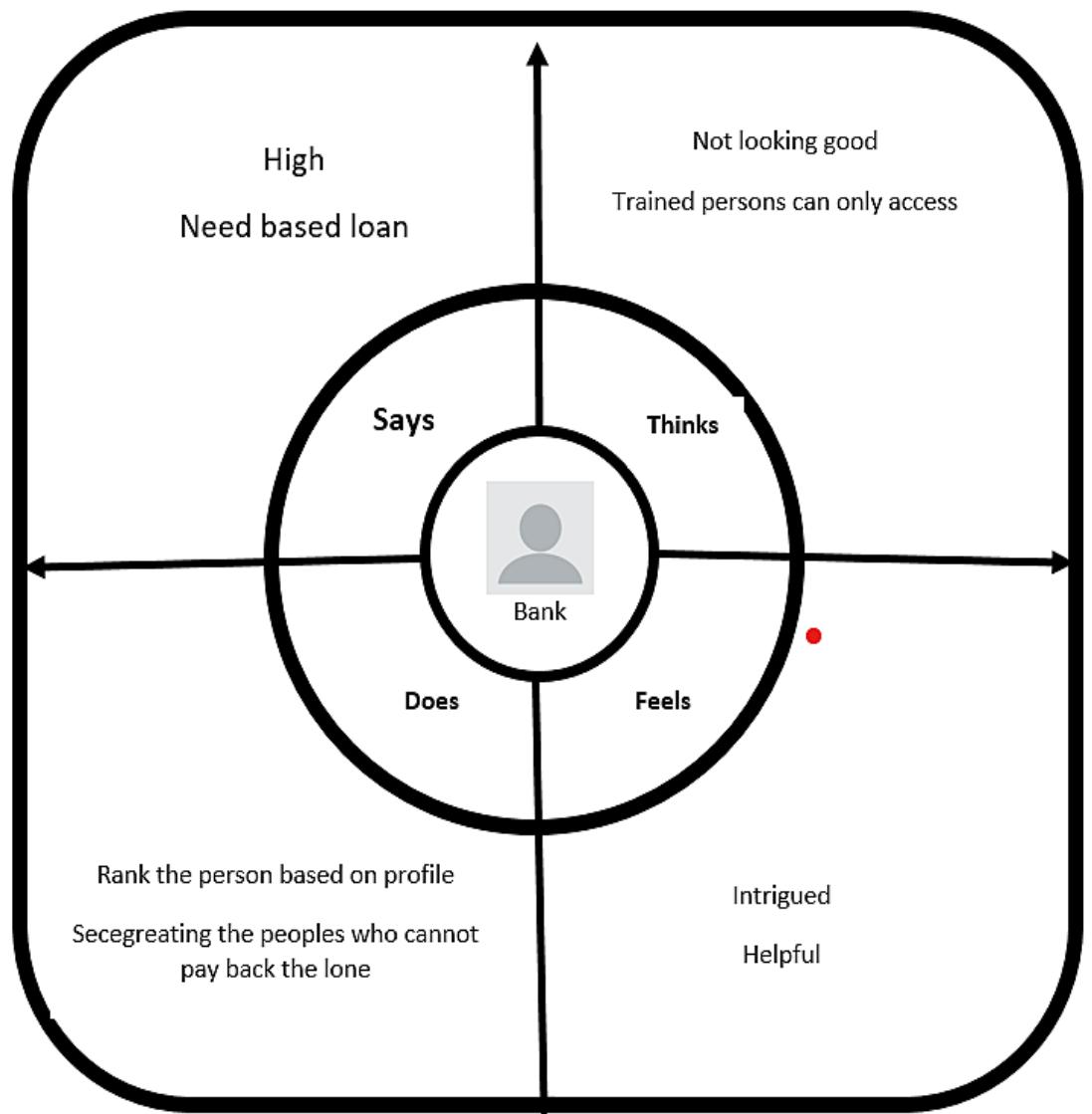
Yes: if the loan is approved

NO: if the loan is not approved

3. IDEATION & PROPOSED SOLUTION

1. Empathy Map Canvas

Empathy map is a collaborative tool teams can use to gain a deeper insight into their customers. Much like a user persona, an empathy map can represent a group of users, such as a customer segment. The empathy map was originally created by Dave Gray and has gained much popularity within the agile community.



2. Ideation & Brainstorming

Step-1: Team Gathering, Collaboration and Select the Problem Statement

Template



Smart Lender - Applicant Credibility Prediction for Loan Approval

Use this template in your own brainstorming sessions so your team can unleash their imagination and start shaping concepts even if you're not sitting in the same room.

⌚ 10 minutes to prepare
⌚ 1 hour to collaborate
👤 2-8 people recommended

[Share template feedback](#)



Need some inspiration?
See a finished version of this template to kickstart your work.
[Open example →](#)

Before you collaborate

A little bit of preparation goes a long way with this session. Here's what you need to do to get going.

⌚ 10 minutes

A Team gathering
Define who should participate in the session and send an invite. Share relevant information or pre-work ahead.

B Set the goal
Think about the problem you'll be focusing on solving in the brainstorming session.

C Learn how to use the facilitation tools
Use the Facilitation Superpowers to run a happy and productive session.

[Open article →](#)

Define your problem statement

What problem are you trying to solve? Frame your problem as a How Might We statement. This will be the focus of your brainstorm.

⌚ 5 minutes

They have a presence across all urban, semi-urban and rural areas. The customer first applies for a home loan and after that, the company validates the customer eligibility for the loan. The company works on the loan application process (real-time) based on customer detail provided while filling out online application forms. These details are Gender, Marital Status, Education, and number of Dependents, Income, Loan Amount, Credit History, and others.

Key rules of brainstorming
To run a smooth and productive session

- Stay in topic.
- Encourage wild ideas.
- Defer Judgment.
- Listen to others.
- Go for volume.
- If possible, be visual.

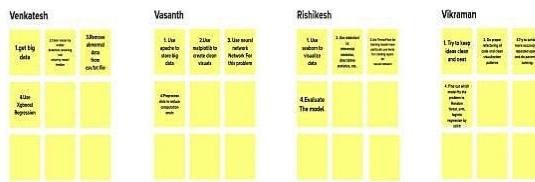
Step-2:Brainstorm, Idea Listing and Grouping

2

Brainstorm

Write down any ideas that come to mind that address your problem statement.

⌚ 10 minutes



3

Group ideas

Take turns sharing your ideas while clustering similar or related notes as you go. Once all sticky notes have been grouped, give each cluster a sentence-like label. If a cluster is bigger than six sticky notes, try and see if you can break it up into smaller sub-groups.

⌚ 20 minutes

1.Venkatesh
Use Numpy ,
pandas , plotly.

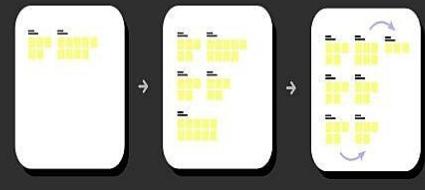
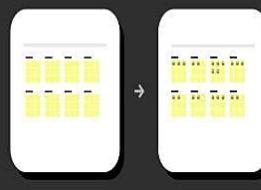
2.Vasanth
use matplotlib

3.Rishikesh
Use seaborn for
clean
visualization , use
testing
techniques if
possible.

Use Apache
spark to
store big data

Use Numpy ,
pandas ,
Matplotlib

4.Vikraman
Refactor code if
possible ,
use clean visuals
and use
required libraires
to
reduce complexity



Step-3:Idea Prioritization

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



5

After you collaborate

You can export the mural as an image or pdf to share with members of your company who might find it helpful.

Quick add-ons

- [A] **Share the mural**
Share a view link to the mural with stakeholders to keep them in the loop about the outcomes of the session.
- [B] **Export the mural**
Export a copy of the mural as a PNG or PDF to attach to emails, include in slides, or save in your drive.

Keep moving forward

- Strategy blueprint**
Define the components of a new idea or strategy.
[Open the template →](#)
- Customer experience journey map**
Understand customer needs, motivations, and obstacles for an experience.
[Open the template →](#)
- Strengths, weaknesses, opportunities & threats**
Identify strengths, weaknesses, opportunities, and threats (SWOT) to develop a plan.
[Open the template →](#)

[Share template feedback](#)



3. Proposed Solution

S.NO	PARAMETER	DESCRIPTION
1.	Problem statement	In our banking system, banks have many products to sell but main source of income of any banks is on its credit line. So, they can earn from interest of those loans which they credit. A bank's profit or a loss depends to a large extent on loans i.e., whether the customers are paying back the loan or defaulting. By predicting the loan defaulters, the bank can reduce its Non-Performing Assets. Banks typically process an applicant's loan after screening and verifying the applicants' eligibility, which is a difficult and time consuming process.

2.	Solution Description	<p>we developed automatic loan prediction using machine learning techniques.</p> <p>We will train the machine with previous dataset. so machine can analyse and understand the process.</p> <p>Then machine will check for eligible applicant and give us result.</p>
----	-----------------------------	---

		<p style="text-align: center;"><u>ADVANTAGES OF THE PROPOSED SYSTEM</u></p> <ol style="list-style-type: none"> 1. Feed-forward back propagation neural network is used to forecast the credit defaults The method in which two or more classifiers are combined together to produce Prediction results 2. Ensemble model for the better prediction 3. Compromised with noise and outlier data of classification. 4. Removal of such above data was done.
3.	Uniqueness	<ol style="list-style-type: none"> 1. Time period for loan sanctioning will be reduced. 2. Whole process will be automated, so human error will be avoided. 3. Eligible applicant will be sanctioned loan without any delay.

4.	Social impact	<ol style="list-style-type: none"> 1. Easy and fastloan approval process for thecustomer. 2. Approves Loan to a trustable person. 3. Bank can find a genuine customer for loanand hence assure that the mentioned loan amount will be repaid.
5.	Scalability	<ol style="list-style-type: none"> 1. It can be provided as software application. 2. Both banking sectorand Lender can use thissoftware application. 3. Customer can use this software anytime andanywhere. 4. This systemis easily scalability and efficient. 5. Easy and user friendly software to all.

4. Problem Solution fit

<p>1. CUSTOMER SEGMENT(S)</p> <p>Public is our customer And their Age Should Between minimum of 21 Years and maximum of 60</p> <p>Focus on CS, tap into CC.</p>	<p>6. CUSTOMER CONSTRAINTS</p> <p>Our solution is completely online so people Who has no Knowledge about online should depend up On someone to help Them to upload the required Bank details.</p> <p>Focus on CC, tap into AS.</p>	<p>5. AVAILABLE SOLUTIONS</p> <p>There is no need to collect details like PAN and Aadhar card</p> <p>Explore AS, differentiate.</p>
<p>2. JOBS TO BE DONE / PROBLEMS</p> <p>Check the eligibility for the loan approval Predicating process can be time consuming and tedious</p> <p>Focus on J&P, tap into BE, understand RC.</p>	<p>9. PROBLEM ROOT CAUSE</p> <p>Predication delay due to lack of Customer details</p> <p>Focus on RC, tap into BE, understand RC.</p>	<p>7. BEHAVIOUR</p> <p>Directed related: find the appropriate Bank manager and ask their queries And perform the process Indirect associated: customer can Discuss with their friends and family</p>
<p>3. TRIGGERS</p> <p>Collect the information from the nearby bank. Gaining more details about loan approval from the neighbour, family and friends. Searching in the website may give detail report about the system</p> <p>Identify strong TR & EM</p>	<p>10. YOUR SOLUTION</p> <p>It is quite efficient and reliable model which will not only reduce the time of the bank but also reduce the waiting time of the customers</p> <p>Identify strong TR & EM</p>	<p>8. CHANNELS OF BEHAVIOUR</p> <p>8.1 ONLINE Search about local approval in website. Search for the popular bank.</p> <p>8.1 OFFLINE Visit the bank manager. Fill the application form with appropriate details.</p> <p>Identify strong TR & EM</p>

1. REQUIREMENT ANALYSIS

a. Functional requirement

FR No.	Non-Functional Requirement	Description
NFR-1	Usability	It describes the context Who,What,When,Where and Why. The specific activities the requirements describe should reflect the both range of goal that the system must support and business goals for creating new system.
NFR-2	Security	Security functionality that ensures one of many different security properties of software is being satisfied. Security requirements are derived from industry standards, applicable laws, and a history of past vulnerabilities.

Following are the functional requirements of the proposed solution.

FR No.	Functional Requirement (Epic)	Sub Requirement (Story / Sub-Task)
FR-1	Home Page	<ol style="list-style-type: none"> 1. Smart Lender Applicant Credibility description 2. Information about Credibility details required for loan approval 3. if new user , REGISTER 4. if already exist , 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.

b. Non-Functional requirements

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

NFR-3	Reliability	It is the measure of the stability or consistency of the test score
-------	--------------------	---

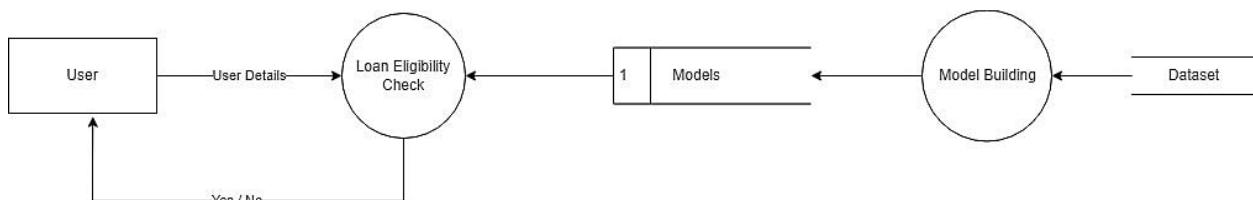
NFR-4	Performance	It defines how well the software system accomplishes certain functions under specific condition.
NFR-5	Availability	It defines how long the IT system can be unavailable without impacting operations.
NFR-6	Scalability	It is the measure of a system ability to increase or decrease in performance and cost in response

PROJECT DESIGN

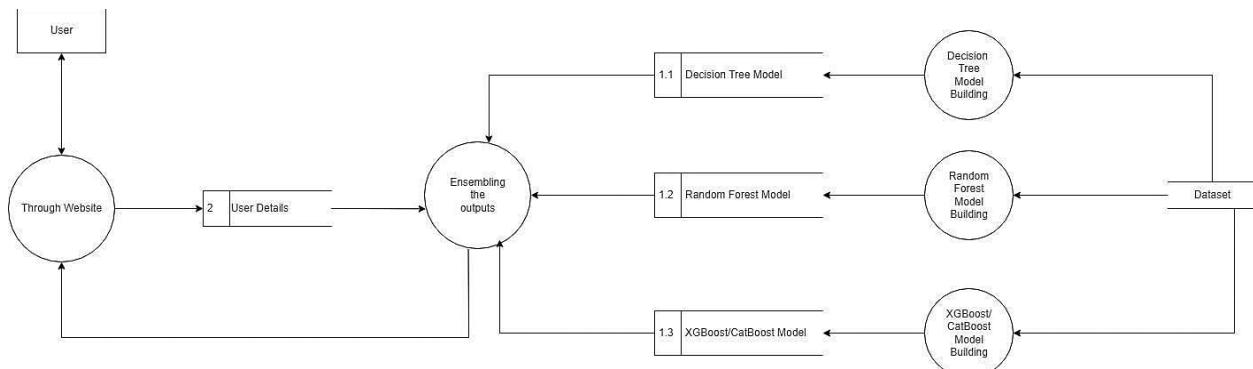
a. Data Flow Diagrams

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.

DFD LEVEL 0



DFD LEVEL 1



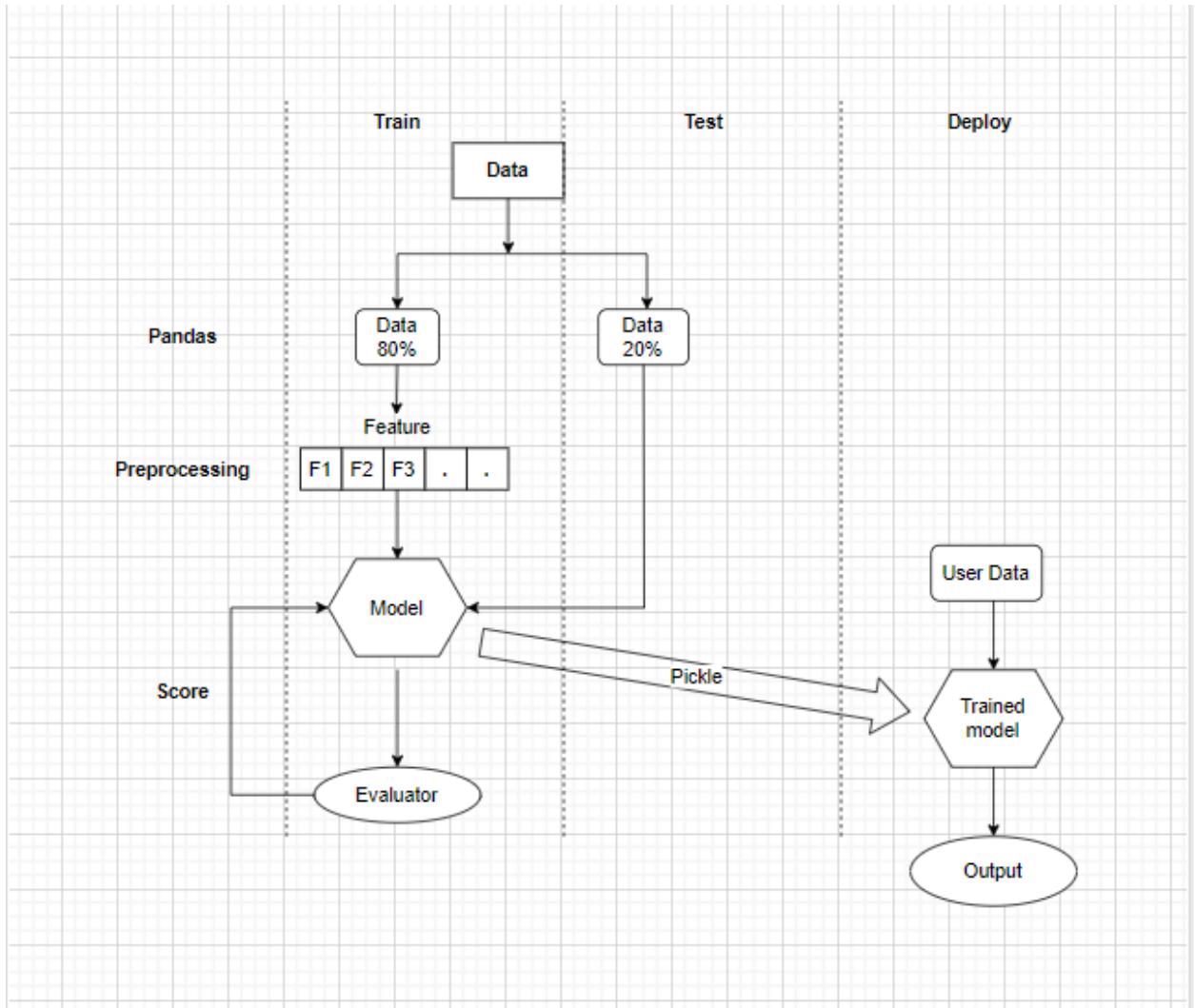
User Stories

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
Money Lender (Web user)	Dashboard	USN-1	As a user, I should be able to access the dashboard.	Access the dashboard	Low	Sprint-3
		USN-2	Select the type of loan	Select the type of loan	Medium	Sprint 3

User Type	Functional Requirement (Epic)	User Story Number	User Story / Task	Acceptance criteria	Priority	Release
	Check for loan eligibility	USN-3	Fill the application with the details of the borrower.	Check the eligibility of the borrower.	High	Sprint-4
Borrower	Dashboard	USN-4	Should be able to access the dashboard.	Access the dashboard.	Low	Sprint-3
		USN-5	Choose the type of loan	Choose the type of loan	Medium	Sprint-3

	Prediction of loan approval	USN-6	Fill the application to check if the loan will be approved or not.	Check for the chances of loan approval. If the loan will not be approved for the current amount, find the amount that will be approved by the lender	High	Sprint-4
	Chatbot	USN-7	Ask doubts to the chatbot	Getting clarified on the doubts	Medium	Sprint-2

b. Solution & Technical Architecture



c. User Stories

A **user story** is an informal, general explanation of a software feature written from the perspective of the end user. Its purpose is to articulate

2. PROJECT PLANNING & SCHEDULING

a. Sprint

Planning & Estimation

discuss with the team and plan how to do the work and discuss about the issues in the project and try to solve the problem.

b. Sprint

Delivery Schedule

The delivery schedule is **an agreement between buyer and seller as to when and how often merchandise will be delivered**. It is a plan that outlines the specifics of future delivery periods. This can be a mutually determined schedule or one that is dictated by the buyer.

- C. Reports
from JIRA

1. CODING & SOLUTIONING (Explain the features added in the project along with code)

- 1. Feature 1
Having the high percentage to predict whether the loan will approve or not
- 2. Feature 2
Has higher accuracy.
- 3. Database Schema (if Applicable)

1. RESULTS

1. Performance Metrics

Performing high accuracy on loan prediction

1. CONCLUSION

The analysis starts from data cleaning and processing missing value, exploratory analysis and finally model building and evaluation of the model. The best accuracy on public test set is when we get higher accuracy score and other performance metrics which will be found out. This paper can help to predict the approval of bank loan or not for a candidate.

1. FUTURE SCOPE

We can make the Bank Loan Approval prediction to connect with Cloud for

future use to optimize the work to implement in Artificial Intelligence environment.

1. APPENDIX

Source Code

<https://github.com/IBM-EPBL/IBM-Project-smart-lender/tree/main/Pre%20development/Project%20Development%20Phase>

GitHub & Project Demo Link

https://drive.google.com/file/d/1lMo2kKp6Ne_yH2xpIUV26MLTeeyg3-i6/view?usp=share_link

Works Done By :

Venkatesh Prasad.S - Ideation phase, Project Design phase-1, Project Design Phase-2, Train the IBM on model, milestone Activity list, Sprint-1.

Vasantha kumar.R - Model bulding, Application building, Data prepossing, projecSprint-2, milestone Activity list,Sprint-3,Sprint-4.

Rishikesh.B - litrature survay, prior knowledge, project objective.

vikraman.GD -Pre-Requisites, projectflow, project stucture.

