##### A

##### Industrial-Oriented Mini Project On

**FUTURE LOAN APPROVALS WITH EXPLAINABLE AI**

(Submitted in partial fulfilment of the requirements for the award of Degree)

**BACHELOR OF TECHNOLOGY**

##### In

**COMPUTER SCIENCE AND ENGINEERING**

##### By

##### K. Akshaya (227R1A05N8) G. Chandra shekar (227R1A05M8)

##### D. Thrisha (227R1A05M1)

##### Under the Guidance of

#### Ms. M. Sunitha

(Assistant Professor)



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

**CMR TECHNICAL CAMPUS**

###### UGC AUTONOMOUS

(Accredited by NAAC, NBA, Permanently Affiliated to JNTUH, Approved by AICTE, New Delhi)

Recognized Under Section 2(f) & 12(B) of the UGCAct.1956, Kandlakoya (V), Medchal Road, Hyderabad-501401.

**May, 2025.**

**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING**

****

**CERTIFICATE**

This is to certify that the project entitled “**FUTURE LOAN APPROVALS WITH EXPLAINABLE AI**” being submitted by **K.Akshaya (227R1A05N8), G.Chandrashekar (227R1A05M8) & D.Thrisha (227R1A05M1)** in partial fulfilment of the requirements for the award of the degree of B.Tech in Computer Science and Engineering to the Jawaharlal Nehru Technological University Hyderabad, during the year 2024-25.

The results embodied in this thesis have not been submitted to any other University or Institute for the award of any degree or diploma.

Ms. M. Sunitha Dr. Nuthanakanti Bhaskar Assistant Professor HoD

INTERNAL GUIDE

Dr. A. Raji Reddy Signature of External Examiner DIRECTOR

Submitted for viva voice Examination held on

**ACKNOWLEDGEMENT**

We take this opportunity to express our gratitude to the people who have been instrumental in the successful completion of this project, we take this opportunity to express our profound gratitude and deep regard to our guide **Dr. J. Narasimharao,** Associate Professor for his exemplary guidance, monitoring and constant encouragement throughout the project work. The blessing, help and guidance given by him shall carry us a long way in the journey of life on which we are about to embark.

We take this opportunity to extend our heartfelt appreciation to the Project Review Committee (PRC) Coordinators—**Y Varalaxmi, B Sekhar, D Nageswar Rao, and SVSV Prasad Sanaboina**—for their unwavering support, insightful guidance, and valuable inputs, which played a crucial role in steering this project through its various stages.

Our sincere appreciation also goes to **Dr. Nuthanakanti Bhaskar**, Head, for his encouragement and continuous support in ensuring the successful completion of our project.

We are deeply grateful to **Dr. A. Raji Reddy**, Director, for his cooperation throughout the course of this project. Additionally, we extend our profound gratitude to Sri. **Ch. Gopal Reddy**, Chairman, Smt. **C. Vasantha Latha**, Secretary and Sri. **C. Abhinav Reddy**, Vice-Chairman, for fostering an excellent infrastructure and a conducive learning environment that greatly contributed to our progress.

We also acknowledge and appreciate the guidance and assistance provided by the faculty and staff of **CMR Technical Campus**, whose contributions have been invaluable in bringing this project to fruition.

Lastly, we sincerely thank our families for their unwavering support and encouragement. We also extend our gratitude to the teaching and non-teaching staff of CMR Technical Campus for their guidance and assistance. Their contributions, along with the support of everyone who helped directly or indirectly, have been invaluable in the successful completion of this project.

K. Akshaya (227R1A05N8)

G.Chandra shekar (227R1A05M8)

D. Thrisha (227R1A05M1)

**VISION AND MISSION**

**INSTITUTE VISION:**

##### To Impart quality education in serene atmosphere thus strive for excellence in Technology and Research.

**INSTITUTE MISSION:**

##### To create state of art facilities for effective Teaching- Learning Process.

##### Pursue and Disseminate Knowledge based research to meet the needs of Industry & Society.

##### Infuse Professional, Ethical and Societal values among Learning Community.

**DEPARTMENT VISION:**

##### To provide quality education and a conducive learning environment in computer engineering that foster critical thinking, creativity, and practical problem-solving skills.

**DEPARTMENT MISSION:**

##### To educate the students in fundamental principles of computing and induce the skills needed to solve practical problems.

##### To provide State-of-the-art computing laboratory facilities to promote industry institute interaction to enhance student’s practical knowledge.

##### To inculcate self-learning abilities, team spirit, and professional ethics among the students to serve society.

**ABSTRACT**

Widespread adoption of automated decision-making by artificial intelligence (AI) is witnessed due to specular advances in computation power and improvements in optimization algorithms especially in machine learning (ML). Complex ML models provide good prediction accuracy; however, the opacity of ML models does not provide sufficient assurance for their adoption in the automation of lending decisions. This paper presents an explainable AI decision- support system to automate the loan underwriting process by belief-rule-base (BRB). This system can accommodate human knowledge and can also learn from historical data by supervised learning. The hierarchical structure of BRB can accommodate factual and heuristic rules. The system can explain the chain of events leading to a decision for a loan application by the importance of an activated rule and the contribution of antecedent attributes in the rule. A business case study on automation of mortgage underwriting is demonstrated to show that the BRB system can provide a good trade-off between accuracy and explainability. The textual explanation produced by the activation of rules could be used as a reason for the denial of a loan. The decision- making process for an application can be comprehended by the significance of rules in providing the decision and contribution of its

|  |  |  |
| --- | --- | --- |
|  | **LIST OF TABLES** |  |
| **TABLE NO** | **TABLE NAME** | **PAGE NO** |
| Table 1 | Uploading Dataset | 34 |
| Table 2 | Classification | 34 |

**TABLE OF CONTENTS**

|  |  |
| --- | --- |
|  |  |
| **ABSTRACT** | i |
| **LIST OF TABLES** | ii |
| **1. INTRODUCTION** | 1 |
| 1.1 PROJECT PURPOSE | 1 |
| 1.2 PROJECT FEATURES | 2 |
| **2. LITERATURE SURVEY** | 4 |
| 2.1 REVIEW OF RELATED WORK | 5 |
| 2.2 DEFINITION OF PROBLEM STATEMENT | 5 |
| 2.3 EXISTING SYSTEM | 6 |
| 2.4 PROPOSED SYSTEM | 6 |
| 2.5 OBJECTIVES | 6 |
| 2.6 HARDWARE & SOFTWARE REQUIREMENTS | 8 |
| 2.6.1 HARDWARE REQUIREMENTS | 8 |
| 2.6.2 SOFTWARE REQUIREMENTS | 8 |

3.

|  |  |  |
| --- | --- | --- |
| **SYSTEM** | **ARCHITECTURE & DESIGN** | 9 |
| 3.1 | PROJECT ARCHITECTURE | 10 |
| 3.3 | DATA FLOW DIAGRAM | 12 |
| 3.3 | DATA FLOW DIAGRAM | 12 |

1. [IMPLEMENTATION 1](#_TOC_250011)3
   1. [ALGORITHMS USED 1](#_TOC_250010)4
   2. [SAMPLE CODE 16](#_TOC_250009)
2. [RESULTS & DISCUSSION 23](#_TOC_250008)
3. [VALIDATION 33](#_TOC_250007)
   1. [INTRODUCTION 33](#_TOC_250006)
   2. TEST CASES 34
      1. UPLOADING DATASET 34
      2. CLASSIFICATION 34
4. [CONCLUSION & FUTURE ASPECTS 35](#_TOC_250005)
   1. [PROJECT CONCLUSION 36](#_TOC_250004)
   2. [FUTURE ASPECTS 37](#_TOC_250003)
5. [BIBLIOGRAPHY 38](#_TOC_250002)
   1. [REFERENCES 39](#_TOC_250001)
6. **INTRODUCTION**
   1. **INTRODUCTION**

Underwriting skill is learnt through several months of training and exchange of knowledge by senior underwriters. This task requires underwriters to be fairly analytical, very organized, and accurate to give informed decision to approve or reject a loan application. Underwriters concurrently analyze a large quantity of information to find affordability, repayment history and collateral. Furthermore, sometimes they are required to change the process due to a shift in regulatory and compliance standards, investor requirements, and customer demands (Krovvidy, 2008).

New technology and strong machine learning (ML) algorithms have opened the doors for a straight through loan application process. Artificial intelligence (AI) systems can execute rules and process customers‟ information in a few milliseconds. Financial institutions have recognized the benefits of AI and are using it in a different subset of the underwriting process and are keen to test and implement newly introduced digital innovation. AI systems are expected to replicate human decision-making skills. However, even today transformation of various algorithmic concepts into training data could be very challenging to solve every instance of the problem for a range of lending products. It may not be able to solve a tiny subset of the problem (Aggour, Bonissone, Cheetham, & Messmer, 2006).

* + 1. **PROJECT PURPOSE**

The purpose of this project is to explore and implement Explainable Artificial Intelligence (XAI) techniques in the domain of loan approval systems to enhance transparency, fairness, and trust in automated decision-making. Traditional AI models used in financial services often lack interpretability, making it difficult for applicants and regulators to understand the reasoning behind approval or rejection decisions. This project aims to develop a loan approval framework that not only leverages the predictive power of machine learning but also provides clear, human-understandable explanations for its outcomes. By doing so, the system will promote ethical AI use, ensure regulatory compliance, reduce bias, and empower both lenders and borrowers with deeper insights into financial decision-making processes.

* + 1. **PROJECT FEATURES**

The proposed Explainable AI-based loan approval system includes the following key features:

* **Automated Loan Evaluation**: Uses machine learning algorithms to assess loan applications based on various financial and personal factors, improving efficiency and accuracy.
* **Explainability Module**: Provides clear, human-readable explanations for each loan decision using XAI techniques such as SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations).
* **Bias Detection and Mitigation**: Identifies potential biases in the dataset or decision-making process and implements corrective strategies to ensure fairness and equity.
* **Regulatory Compliance Support**: Aligns with legal requirements such as GDPR, FCRA, and other financial regulations by making decision-making processes transparent and auditable.
* **Interactive Dashboard**: Offers a user-friendly interface for both loan officers and applicants to view decisions, understand reasoning, and explore "what-if" scenarios.
* **Real-time Decision Processing**: Enables quick assessment and explanation of loan applications, reducing waiting times for applicants.
* **Data Privacy and Security**: Ensures that all user data is handled securely and in compliance with data protection standards.

1. **LITERATURE SURVEY**

**2.LITERATURE SURVEY**

The increasing adoption of Artificial Intelligence (AI) in financial services has significantly impacted how loan approvals are conducted. However, traditional AI models often function as black boxes, providing accurate predictions without any insight into the rationale behind them. This lack of transparency raises concerns about fairness, accountability, and regulatory compliance, especially in sensitive areas like credit scoring and loan decisions.

Several studies have highlighted the need for **Explainable AI (XAI)** in the financial domain. Ribeiro et al. (2016) introduced **LIME (Local Interpretable Model-agnostic Explanations)** as a technique to explain the predictions of any machine learning model by approximating it locally with an interpretable model. Similarly, Lundberg and Lee (2017) developed **SHAP (SHapley Additive exPlanations)**, which assigns each feature an importance value for a particular prediction, grounded in cooperative game theory. These methods have become foundational in applying XAI to domains requiring transparency.

Research by Doshi-Velez and Kim (2017) emphasized the importance of interpretability in AI, especially in high-stakes decision-making like healthcare and finance. In the financial sector, studies have shown that interpretable models help improve user trust and regulatory acceptance. For example, Chen et al. (2018) explored interpretable credit scoring systems that balance model performance with explainability.

Recent literature also addresses **bias in AI systems**, with a growing body of work focused on fairness-aware machine learning. Algorithms like FairML and AI Fairness 360 toolkit by IBM aim to detect and reduce bias in training data and model outputs, which is crucial in ensuring that loan decisions are not discriminatory.

In summary, existing literature supports the integration of XAI in loan approval systems to ensure transparent, fair, and accountable decision-making. This project builds upon these foundational studies to develop a practical, explainable framework for modern lending processes.

* 1. **REVIEW OF RELATED WORK**

Several research efforts and industry initiatives have explored the application of Explainable AI (XAI) in the context of loan approval and credit scoring, highlighting both the potential and challenges of integrating transparency into financial decision-making systems.

* **LIME and SHAP for Model Interpretability**: Ribeiro et al. (2016) proposed LIME (Local Interpretable Model-Agnostic Explanations), which explains individual predictions by approximating the complex model locally with a simpler model. Lundberg and Lee (2017) introduced SHAP (SHapley Additive exPlanations), based on game theory, to assign meaningful importance scores to each feature influencing a prediction. Both methods have been widely applied to explain loan approval models and have become the backbone of many explainable AI systems.
* **Interpretable Models in Credit Scoring**: Chen et al. (2018) presented interpretable machine learning models for credit scoring that retain high predictive accuracy while providing clear decision rules. These models demonstrated that decision trees and rule-based models could offer a compromise between transparency and performance.
* **Fairness in Financial AI**: Research by Hardt et al. (2016) and others on fairness in machine learning has shown how biases in historical data can lead to discriminatory outcomes. Tools such as IBM’s AI Fairness 360 and Microsoft’s Fairlearn have been developed to identify and mitigate such biases in financial models.
* **Industry Adoption**: Companies like FICO and Zest AI have begun incorporating explainability into their credit risk models. For instance, FICO’s Explainable Machine Learning Challenge (2018) encouraged data scientists to create models that are not only accurate but also explainable. Zest AI claims to provide more inclusive lending decisions by using XAI to justify risk assessments.
* **Regulatory Considerations**: The European Union’s General Data Protection Regulation (GDPR) and U.S. regulations such as the Fair Credit Reporting Act (FCRA) emphasize the need for explainability and transparency in automated decision-making systems. This has driven research toward creating models that are compliant by design.
  1. **DEFINITION OF PROBLEM STATEMENT**

Many modern loan approval systems use AI to make fast and accurate decisions, but they often fail to explain how those decisions are made. This lack of transparency can lead to confusion, unfair outcomes, and a loss of trust among applicants. The main problem is that users and lenders are unable to understand or challenge the reasoning behind approvals or rejections. This project aims to address this issue by developing an AI-based loan approval system that uses Explainable AI (XAI) techniques to provide clear, understandable explanations for each decision, making the process more transparent, fair, and trustworthy.

* 1. **EXISTING SYSTEM**

#### In the current loan approval systems, most financial institutions use traditional machine learning models such as logistic regression, decision trees, or neural networks to predict whether a loan applicant is likely to repay the loan. These models analyze various factors such as income, credit score, employment history, and past repayment behavior to make decisions. While these systems are efficient and help reduce manual effort, they often operate as "black boxes," providing little to no insight into how a particular decision was made. As a result, applicants who are rejected do not receive clear explanations, and loan officers may struggle to justify the outcomes. Additionally, these systems may unintentionally include biases based on historical data, leading to unfair or discriminatory decisions. The lack of interpretability and transparency is a major drawback in existing systems, especially when regulations require clear reasoning behind automated decisions. Limitations of Existing System

#### 2.4 PROPOSED SYSTEM

The proposed system introduces an Explainable AI (XAI)-based loan approval model that not only predicts the eligibility of applicants but also provides clear and understandable reasons behind each decision. By integrating XAI techniques such as SHAP or LIME, the system offers transparency in how features like credit score, income, or employment status contribute to the final outcome. This helps applicants understand why their loan was approved or rejected and allows loan officers to explain decisions with confidence. The system also includes bias detection mechanisms to identify and reduce unfair influence of sensitive attributes (e.g., gender, race). Additionally, a user-friendly interface is provided to display decision explanations in an easy-to-read format. Overall, the proposed system ensures fairness, regulatory compliance, and builds trust among users by making AI-driven loan decisions more transparent and accountable.

#### 

#### Advantages of the Proposed System:

**Transparency in Decisions**: Provides clear and understandable explanations for each loan decision using Explainable AI techniques.

**Increased Trust**: Helps build trust among applicants and loan officers by making the decision-making process open and accountable.

**Bias Detection and Reduction**: Identifies and minimizes biases in the data or model, ensuring fair and non-discriminatory decisions.

**Regulatory Compliance**: Supports legal requirements for transparency and explainability in automated decision-making (e.g., GDPR, FCRA).

**Improved User Experience**: Applicants receive meaningful feedback, which can help them improve future applications.

**Better Decision Support**: Loan officers can use explanations to validate or override decisions confidently when needed.

**Data-Driven Insights**: Enables deeper understanding of key factors affecting loan approvals, aiding in policy improvement and risk management.

* 1. **OBJECTIVES**

The main objectives of this project are:

1. **To develop an AI-based loan approval system** that automates the evaluation of loan applications using relevant financial and personal data.
2. **To integrate Explainable AI (XAI) techniques** such as SHAP or LIME to provide clear and interpretable explanations for each decision made by the system.
3. **To enhance transparency and trust** in the loan approval process by making the AI model's reasoning understandable to both applicants and loan officers.
4. **To identify and reduce bias** in the dataset and model predictions to ensure fair and ethical decision-making.
5. **To comply with regulatory standards** by providing explanations that satisfy legal requirements for automated decision systems.
6. **To build a user-friendly interface** that presents decisions and their explanations in a simple and interactive way.
7. **To support continuous improvement** by analyzing model performance and explanation feedback to refine the system over time.
   1. **HARDWARE & SOFTWARE REQUIREMENTS**
      1. **HARDWARE REQUIREMENTS:**

Hardware interfaces specifies the logical characteristics of each interface between the software product and the hardware components of the system. The following are some hardware requirements,

* + - 1. Processor : Intel Core i3
      2. Hard disk : 20GB.
      3. RAM : 4GB.
    1. **SOFTWARE REQUIREMENTS:**

Software Requirements specifies the logical characteristics of each interface and software components of the system. The following are some software requirements,

* + - 1. Operating system : Windows 10
      2. Language : Python
      3. Back-End : Django-ORM
      4. Frame Work : Tkinter

**3. SYSTEM ARCHITECTURE &**

**DESIGN**

**3.SYSTEM ARCHITECTURE & DESIGN**

Project architecture refers to the structural framework and design of a project, encompassing its components, interactions, and overall organization. It provides a clear blueprint for development, ensuring efficiency, scalability, and alignment with project goals. Effective architecture guides the project's lifecycle, from planning to execution, enhancing collaboration and reducing complexity.

* 1. **PROJECT ARCHITECTURE**

The proposed Explainable AI-based Loan Approval System is designed with a modular architecture to ensure scalability, transparency, and ease of use. Below is an overview of the key components and their interactions:

**1. Data Collection Module**

* Gathers applicant data such as income, credit score, employment history, loan amount, and other relevant features.
* Ensures data privacy and security during input and storage.

**2. Data Preprocessing Module**

* Cleans and normalizes raw data.
* Handles missing values, encodes categorical variables, and performs feature scaling.
* Ensures balanced datasets to reduce model bias.

**3. AI-Based Loan Prediction Model**

* Trains a machine learning model (e.g., Decision Tree, Random Forest, XGBoost) on historical loan data.
* Predicts the likelihood of loan approval or rejection based on input features.

**4. Explainable AI (XAI) Module**

* Integrates techniques such as SHAP (SHapley Additive Explanations) or LIME to interpret model decisions.
* Generates user-friendly visual and textual explanations for each prediction.

**5. Bias Detection & Fairness Module**

* Analyzes the model for biased outcomes based on sensitive attributes (e.g., gender, age).
* Applies fairness metrics and mitigation strategies if needed.

**6. User Interface (UI)**

* A dashboard for loan officers to view applications, predictions, and explanations.
* A portal for applicants to see their application status and the reasons behind the decision.

**7. Decision Logging & Audit Trail**

* Stores each decision and its explanation for auditing and compliance.
* Useful for improving future models and handling customer queries or appeals.
  1. **DATA FLOW DIAGRAM**

[User Input]

↓

[Data Preprocessing]

↓

[AI Loan Prediction Model] → [XAI Module] → [Explanation Output]

↓ ↘

[Bias & Fairness Check] [User Interface Display]

↓

[Decision Logging & Audit]

# 4. IMPLEMENTATION

**4.IMPLEMENTATION**

The implementation phase of a project involves executing the planned strategies and tasks. It requires meticulous coordination, resource allocation, and monitoring to ensure that objectives are met efficiently. Effective implementation is crucial for achieving project goals and delivering expected outcomes within the set timeline and budget constraints.

### ALGORITHMS USED

**1. Machine Learning Algorithms for Prediction:**

**a) Random Forest**

* An ensemble learning method that builds multiple decision trees and merges their results for more accurate and stable predictions.
* Suitable for classification tasks like loan approval (Approved/Rejected).
* Handles missing data and overfitting well.

**b) XGBoost (Extreme Gradient Boosting)**

* A powerful, fast, and scalable gradient boosting algorithm that provides high performance.
* Often used in financial applications due to its accuracy and ability to handle imbalanced datasets.
* Helps improve prediction by reducing errors from weak learners (individual trees).

**2. Explainable AI (XAI) Algorithms for Interpretation:**

**a) SHAP (SHapley Additive exPlanations)**

* Based on cooperative game theory.
* Explains the prediction by assigning each feature an importance value — i.e., how much it contributed to the final decision.
* Offers both global and local explanations (i.e., for all data vs. a single prediction).

**b) LIME (Local Interpretable Model-Agnostic Explanations)**

* Focuses on explaining individual predictions by approximating the model locally with a simpler, interpretable model (like linear regression).
* Works with any machine learning model and provides human-readable insights.

**3. Fairness and Bias Detection Algorithms:**

**a) Fairness Metrics and Techniques**

* Measures such as **Demographic Parity**, **Equal Opportunity**, and **Disparate Impact** are used to detect unfair treatment of certain groups.
* Tools like **AI Fairness 360** (IBM) may be integrated to evaluate and reduce bias in training data and model outputs.

### SAMPLE CODE

from tkinter import \*

import tkinter

from tkinter import filedialog

import numpy as np

from tkinter import simpledialog

from sklearn.model\_selection import train\_test\_split

from tkinter import ttk

from tkinter.filedialog import askopenfilename

import os

import pickle

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.metrics import accuracy\_score

import pandas as pd

from sklearn.metrics import precision\_score

from sklearn.metrics import recall\_score

from sklearn.metrics import f1\_score

from sklearn.metrics import confusion\_matrix

from sklearn.preprocessing import StandardScaler

from sklearn.preprocessing import LabelEncoder

from sklearn.ensemble import RandomForestClassifier#importing ML classes

import shap

main = tkinter.Tk()

main.title("Future of Loan Approvals with Explainable AI") #designing main screen

main.geometry("1000x650")

global filename, loan\_status, loan\_reject\_reason, status\_names, reject\_names, label\_encoder, scaler, cols

global loan\_X\_train, loan\_X\_test, loan\_y\_train, loan\_y\_test

global reject\_X\_train, reject\_X\_test, reject\_y\_train, reject\_y\_test

global accuracy, precision, recall, fscore

global rf, reject\_rf, dataset, X

def loadDataset():

    global filename, dataset, loan\_status, loan\_reject\_reason, status\_names, reject\_names

    filename = filedialog.askopenfilename(initialdir="Dataset")

    text.delete('1.0', END)

    text.insert(END,filename+" loaded\n\n")

    dataset = pd.read\_csv(filename, nrows=20000)

    text.insert(END,str(dataset.head()))

    loan\_status = dataset['NAME\_CONTRACT\_STATUS']

    loan\_reject\_reason = dataset['CODE\_REJECT\_REASON']

    #visualizing loan status class labels

    status\_names, status\_count = np.unique(loan\_status, return\_counts = True)

    reject\_names, reject\_count = np.unique(loan\_reject\_reason, return\_counts = True)

    loan\_df = []

    for i in range(len(status\_names)):

        loan\_df.append([status\_names[i], status\_count[i]])

    loan\_df = pd.DataFrame(loan\_df, columns=['Loan\_Status', 'Count'])

    reject\_df = []

    for i in range(len(reject\_names)):

        reject\_df.append([reject\_names[i], reject\_count[i]])

    reject\_df = pd.DataFrame(reject\_df, columns=['Reject\_Reason', 'Count'])

    fig, axs = plt.subplots(1, 2, figsize=(14, 4))

    sns.barplot(x="Loan\_Status", y='Count', data=loan\_df, ax=axs[0])

    sns.barplot(x="Reject\_Reason", y='Count', data=reject\_df, ax=axs[1])

    axs[0].set\_title("Loan Application Status Graph")

    axs[1].set\_title("Reject Reason Graph")

    plt.show()

def processDataset():

    text.delete('1.0', END)

    global dataset, scaler, label\_encoder, cols, loan\_status, loan\_reject\_reason, X

    label\_encoder = []

    #converting non-numeric data to numeric values

    dataset.fillna(0, inplace = True)

    label\_encoder = []

    columns = dataset.columns

    types = dataset.dtypes.values

    cols = []

    for i in range(len(types)):

        name = types[i]

        if name == 'object': #finding column with object type

            le = LabelEncoder()

            dataset[columns[i]] = pd.Series(le.fit\_transform(dataset[columns[i]].astype(str)))#encode all str columns to numeric

            label\_encoder.append(le)

            cols.append(columns[i])

    dataset.fillna(0, inplace = True)

    loan\_status = dataset['NAME\_CONTRACT\_STATUS']

    loan\_reject\_reason = dataset['CODE\_REJECT\_REASON']

    dataset.drop(['NAME\_CONTRACT\_STATUS', 'CODE\_REJECT\_REASON'], axis = 1,inplace=True)

    X = dataset.values

    scaler = StandardScaler()

    X = scaler.fit\_transform(X)#normalize train features

    text.insert(END,"Normalized \* Processed Features = "+str(X))

def splitDataset():

    text.delete('1.0', END)

    global X, loan\_status, loan\_reject\_reason

    global loan\_X\_train, loan\_X\_test, loan\_y\_train, loan\_y\_test

    global reject\_X\_train, reject\_X\_test, reject\_y\_train, reject\_y\_test

    #split dataset into train and test

    loan\_X\_train, loan\_X\_test, loan\_y\_train, loan\_y\_test = train\_test\_split(X, loan\_status, test\_size = 0.2)

    reject\_X\_train, reject\_X\_test, reject\_y\_train, reject\_y\_test = train\_test\_split(X, loan\_reject\_reason, test\_size = 0.2)

    text.insert(END,"Total records found in dataset = "+str(X.shape[0])+"\n")

    text.insert(END,"Total features found in dataset= "+str(X.shape[1])+"\n")

    text.insert(END,"80% dataset for training : "+str(loan\_X\_train.shape[0])+"\n")

    text.insert(END,"20% dataset for testing  : "+str(loan\_X\_test.shape[0])+"\n")

def calculateMetrics(algorithm, predict, y\_test, label\_names):

    a = accuracy\_score(y\_test,predict)\*100

    p = precision\_score(y\_test, predict,average='macro') \* 100

    r = recall\_score(y\_test, predict,average='macro') \* 100

    f = f1\_score(y\_test, predict,average='macro') \* 100

    accuracy.append(a)

    precision.append(p)

    recall.append(r)

    fscore.append(f)

    text.insert(END,algorithm+" Accuracy  :  "+str(a)+"\n")

    text.insert(END,algorithm+" Precision : "+str(p)+"\n")

    text.insert(END,algorithm+" Recall    : "+str(r)+"\n")

    text.insert(END,algorithm+" FScore    : "+str(f)+"\n\n")

    conf\_matrix = confusion\_matrix(y\_test, predict)

    plt.figure(figsize =(7, 5))

    ax = sns.heatmap(conf\_matrix, xticklabels = label\_names, yticklabels = label\_names, annot = True, cmap="viridis" ,fmt ="g");

    ax.set\_ylim([0,len(label\_names)])

    plt.title(algorithm+" Confusion matrix")

    plt.xticks(rotation=90)

    plt.ylabel('True class')

    plt.xlabel('Predicted class')

    plt.tight\_layout()

    plt.show()

def aiApproval():

    text.delete('1.0', END)

    global loan\_X\_train, loan\_X\_test, loan\_y\_train, loan\_y\_test

    global reject\_X\_train, reject\_X\_test, reject\_y\_train, reject\_y\_test

    global accuracy, precision, recall, fscore, rf, status\_names

    #define global variables to save accuracy and other metrics

    accuracy = []

    precision = []

    recall = []

    fscore = []

    rf = RandomForestClassifier()

    rf.fit(loan\_X\_train, loan\_y\_train)

    predict = rf.predict(loan\_X\_test)

    calculateMetrics("Random Forest Loan Status", predict, loan\_y\_test, status\_names)

def aiReject():

    global reject\_X\_train, reject\_X\_test, reject\_y\_train, reject\_y\_test

    global accuracy, precision, recall, fscore, rf, reject\_names, reject\_rf

    reject\_rf = RandomForestClassifier()

    reject\_rf.fit(reject\_X\_train, reject\_y\_train)

    predict = reject\_rf.predict(reject\_X\_test)

    calculateMetrics("Random Forest Loan Rejection", predict, reject\_y\_test, reject\_names)

def explainAI():

    global rf

    features\_names = ['SK\_ID\_PREV', 'SK\_ID\_CURR', 'NAME\_CONTRACT\_TYPE', 'AMT\_ANNUITY', 'AMT\_APPLICATION', 'AMT\_CREDIT', 'AMT\_DOWN\_PAYMENT', 'AMT\_GOODS\_PRICE',

                      'WEEKDAY\_APPR\_PROCESS\_START', 'HOUR\_APPR\_PROCESS\_START', 'FLAG\_LAST\_APPL\_PER\_CONTRACT', 'NFLAG\_LAST\_APPL\_IN\_DAY', 'RATE\_DOWN\_PAYMENT',

                      'RATE\_INTEREST\_PRIMARY', 'RATE\_INTEREST\_PRIVILEGED', 'NAME\_CASH\_LOAN\_PURPOSE', 'DAYS\_DECISION',

                      'NAME\_PAYMENT\_TYPE', 'NAME\_TYPE\_SUITE', 'NAME\_CLIENT\_TYPE', 'NAME\_GOODS\_CATEGORY', 'NAME\_PORTFOLIO',

                      'NAME\_PRODUCT\_TYPE', 'CHANNEL\_TYPE', 'SELLERPLACE\_AREA', 'NAME\_SELLER\_INDUSTRY', 'CNT\_PAYMENT', 'NAME\_YIELD\_GROUP',

                      'PRODUCT\_COMBINATION', 'DAYS\_FIRST\_DRAWING', 'DAYS\_FIRST\_DUE', 'DAYS\_LAST\_DUE\_1ST\_VERSION', 'DAYS\_LAST\_DUE', 'DAYS\_TERMINATION',

                      'NFLAG\_INSURED\_ON\_APPROVAL']

    explainer = shap.TreeExplainer(rf)

    shap\_values = explainer.shap\_values(loan\_X\_test[0:200])

    shap.summary\_plot(shap\_values, feature\_names=features\_names)

    plt.show()

def predict():

    text.delete('1.0', END)

    global rf, reject\_rf, scaler, label\_encoder, cols, status\_names, reject\_names

    filename = filedialog.askopenfilename(initialdir="Dataset")

    test\_data = pd.read\_csv(filename)

    test\_data.fillna(0, inplace = True)

    for i in range(len(cols)):

        test\_data[cols[i]] = pd.Series(label\_encoder[i].transform(test\_data[cols[i]].astype(str)))

        test\_data.fillna(0, inplace = True)

    test\_data.drop(['NAME\_CONTRACT\_STATUS', 'CODE\_REJECT\_REASON'], axis = 1,inplace=True)

    test = test\_data.values

    X = test\_data.values

    X = scaler.transform(X)

    loan\_predict = rf.predict(X)

    reject\_predict = reject\_rf.predict(X)

    for i in range(len(loan\_predict)):

        text.insert(END,"Test Data = "+str(test[i])+"\n")

        text.insert(END,"Loan Approval Status = "+str(status\_names[loan\_predict[i]])+"\n")

        text.insert(END,"Loan Approval/Rejection Reason = "+str(reject\_names[reject\_predict[i]])+"\n\n")

font = ('times', 16, 'bold')

title = Label(main, text='Future of Loan Approvals with Explainable AI', justify=LEFT)

title.config(bg='lavender blush', fg='DarkOrchid1')

title.config(font=font)

title.config(height=3, width=120)

title.place(x=100,y=5)

title.pack()

font1 = ('times', 13, 'bold')

uploadButton = Button(main, text="Upload Loan Application Dataset", command=loadDataset)

uploadButton.place(x=10,y=100)

uploadButton.config(font=font1)

processButton = Button(main, text="Preprocess Dataset", command=processDataset)

processButton.place(x=330,y=100)

processButton.config(font=font1)

splitButton = Button(main, text="Split Dataset Train & Test", command=splitDataset)

splitButton.place(x=620,y=100)

splitButton.config(font=font1)

approvalButton = Button(main, text="Train AI on Loan Approval", command=aiApproval)

approvalButton.place(x=10,y=150)

approvalButton.config(font=font1)

rejectButton = Button(main, text="Train AI on Loan Rejections", command=aiReject)

rejectButton.place(x=330,y=150)

rejectButton.config(font=font1)

explainButton = Button(main, text="Explainable AI", command=explainAI)

explainButton.place(x=620,y=150)

explainButton.config(font=font1)

predictButton = Button(main, text="Predict Loan Status using Test Data", command=predict)

predictButton.place(x=820,y=150)

predictButton.config(font=font1)

font1 = ('times', 12, 'bold')

text=Text(main,height=22,width=140)

scroll=Scrollbar(text)

text.configure(yscrollcommand=scroll.set)

text.place(x=10,y=200)

text.config(font=font1)

main.config(bg='light coral')

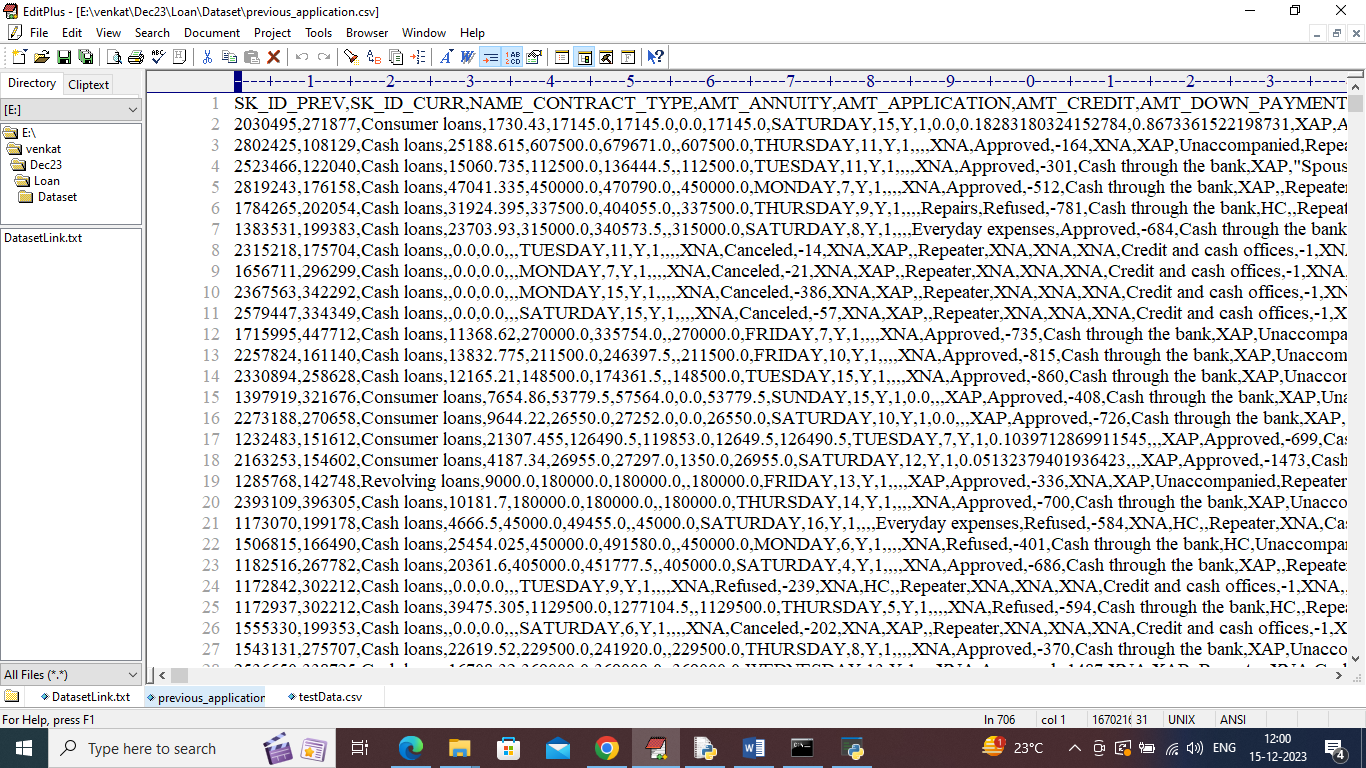
main.mainloop()

# 5. RESULTS & DISCUSSION

**5.RESULTS & DISCUSSION**

The following screenshots showcase the results of our project, highlighting key features and functionalities. These visual representations provide a clear overview of how the system performs under various conditions, demonstrating its effectiveness and user interface. The screenshots serve as a visual aid to support the project's technical and operational achievements.

#### GUI/Main Interface:



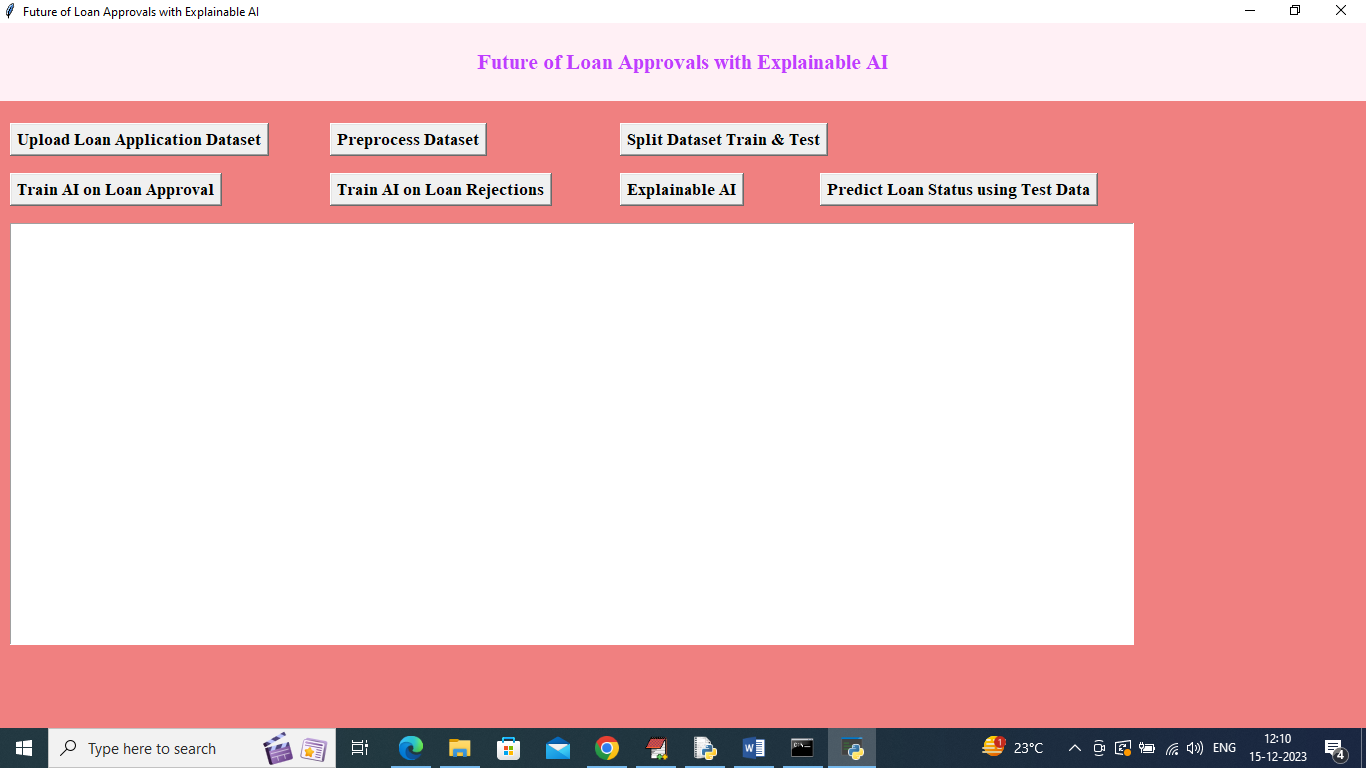
In above dataset screen first row represents dataset column names and remaining rows represents dataset values and by using above dataset we will train AI to predict loan approval status and reason.

To implement this project we have designed following modules

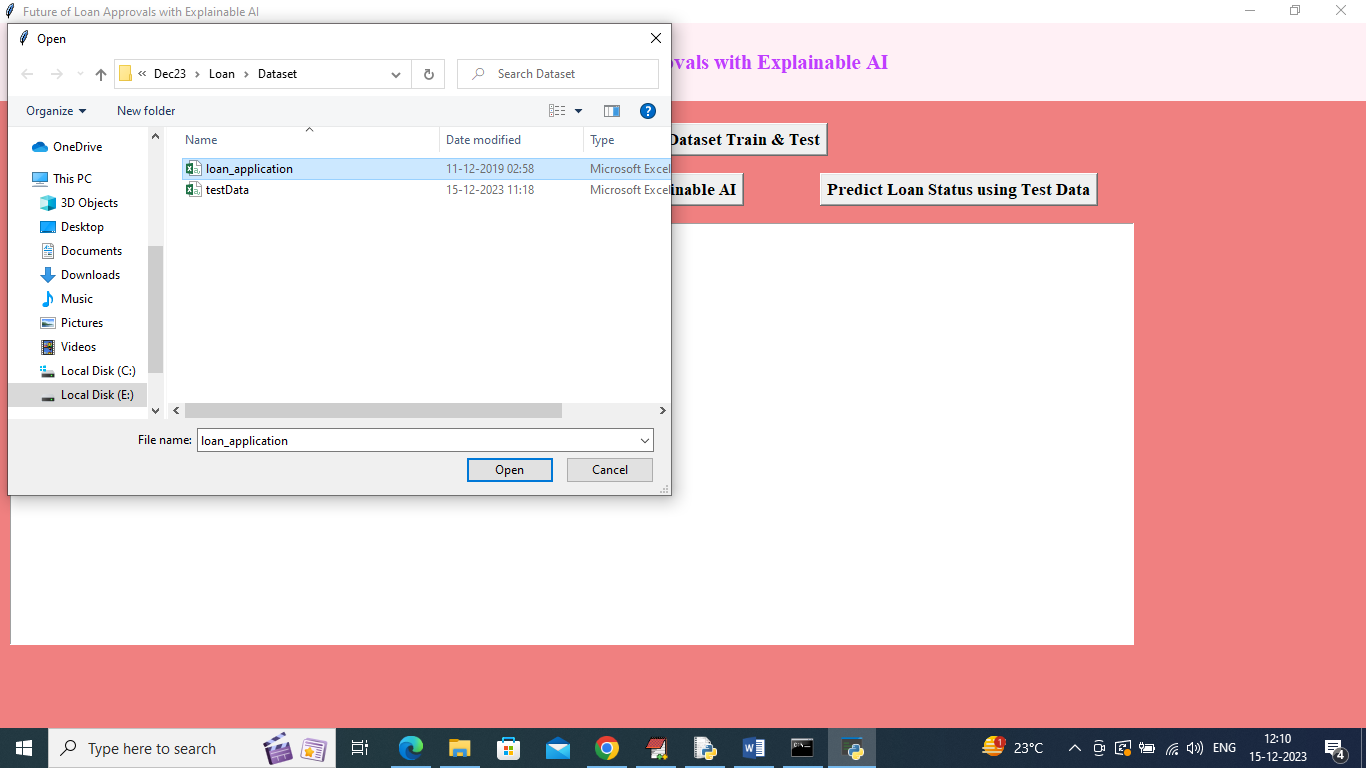
1. Upload Loan Application Dataset: using this module we will upload dataset to application and then application will read entire dataset and then find all class labels for loan and reject reason and plot them in a graph
2. Pre-process Dataset: dataset contains missing value and both numeric and non-numeric data so by employing label encoder class will convert all data into numeric format and then normalized all dataset values to make it clean.
3. Split Dataset Train & Test: using this module will split Dataset in to train and test where application using 80% dataset for training and 20% for testing
4. Train AI on Loan Approval: using this module will train AI on 80% training data to predict loan approval status and then perform prediction on 20% test data to calculate prediction accuracy
5. Train AI on Loan Rejections: using this module will train AI on 80% training data to predict REJECTION REASON and then perform prediction on 20% test data to calculate prediction accuracy
6. Explainable AI: using this module we will explain about features which are contributing most for label prediction
7. Predict Loan Status using Test Data: using this module we will upload test data and then AI model will predict loan status and then predict loan approval or rejection REASON.

SCREEN SHOTS

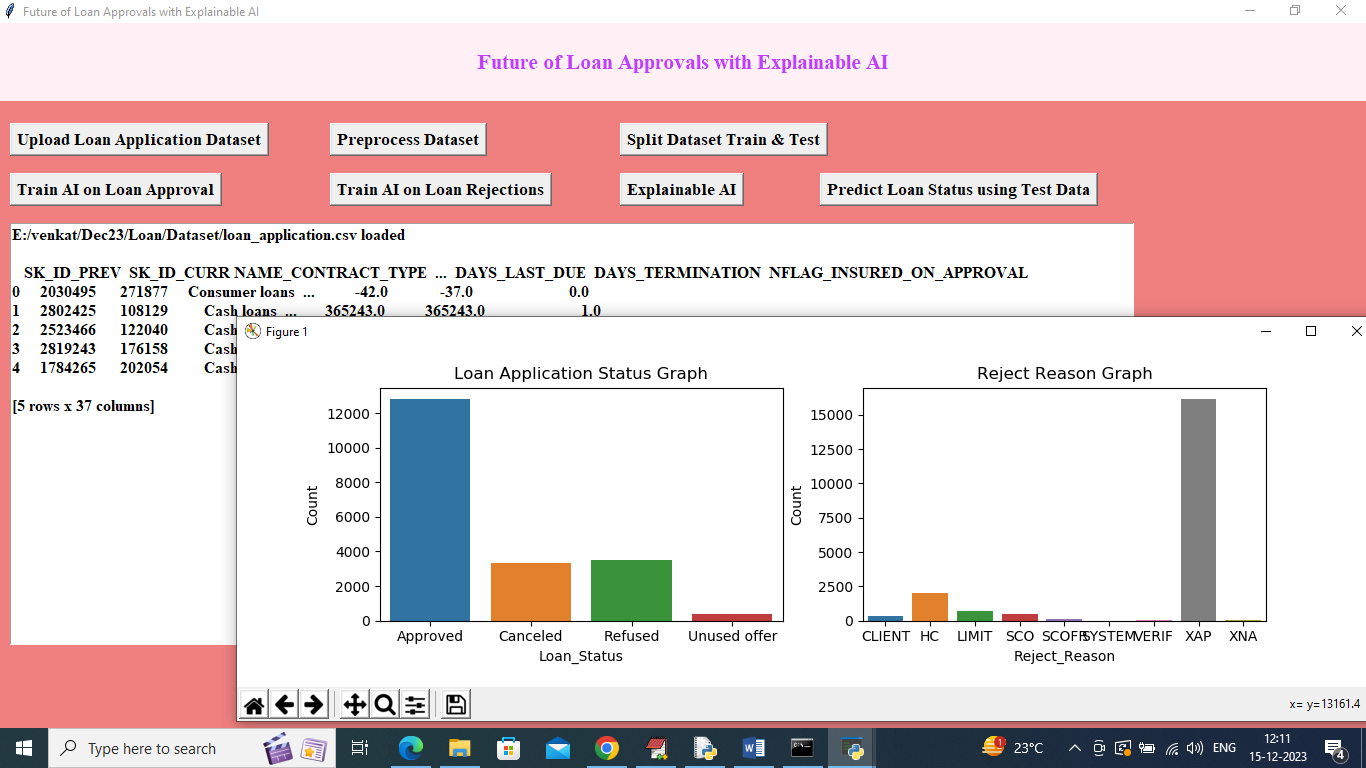
To run project double click on ‘run.bat’ file to get below screen



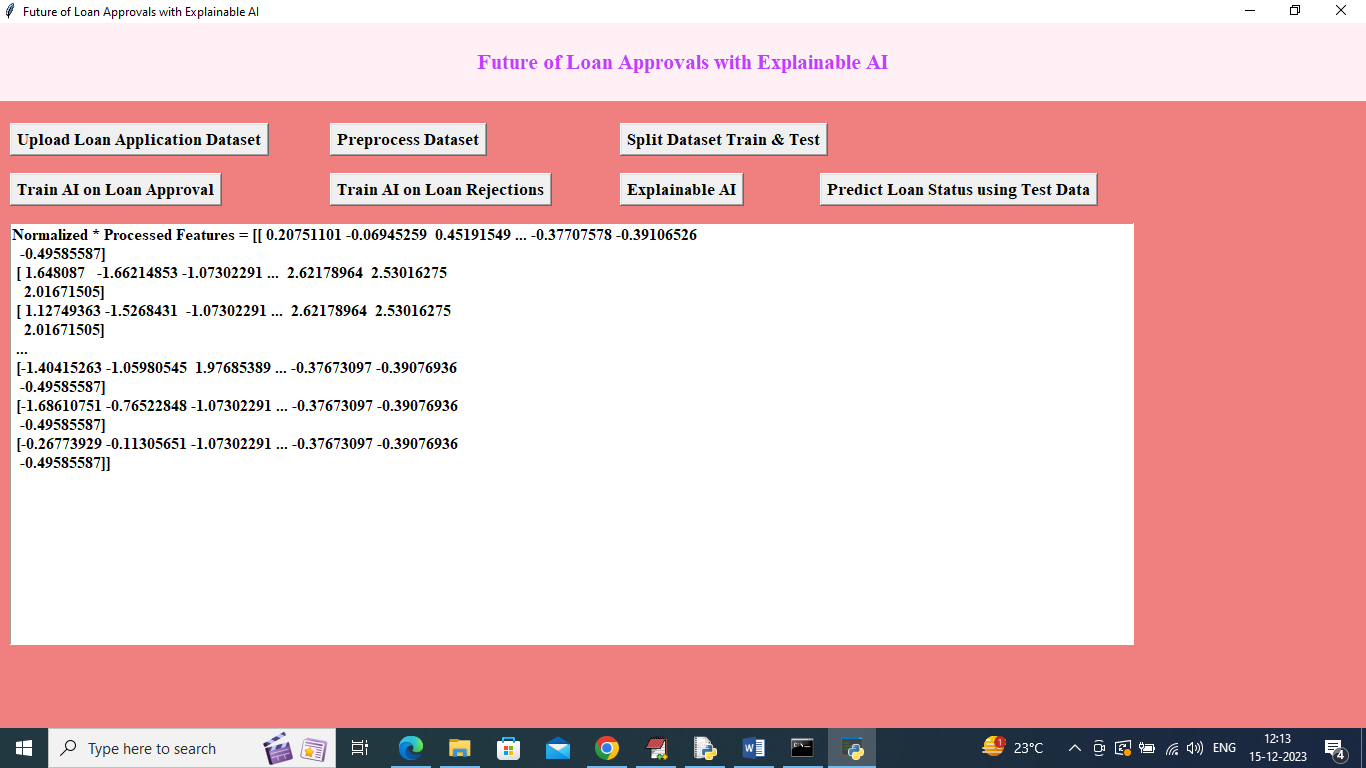
In above screen click on ‘Upload Loan Application Dataset’ button to upload dataset and then will get below output



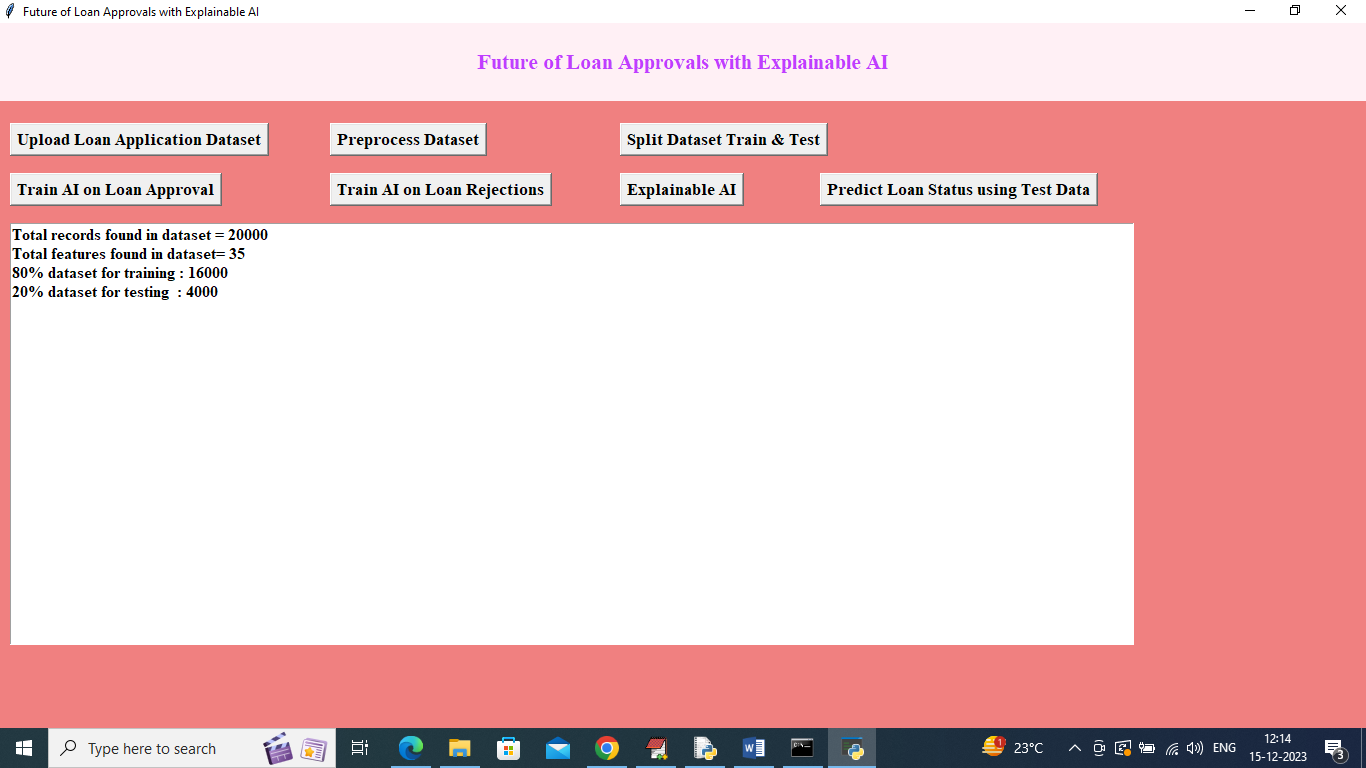
In above screen selecting and uploading ‘loan\_application.csv’ file and then click on ‘Open’ button to load dataset and get below output



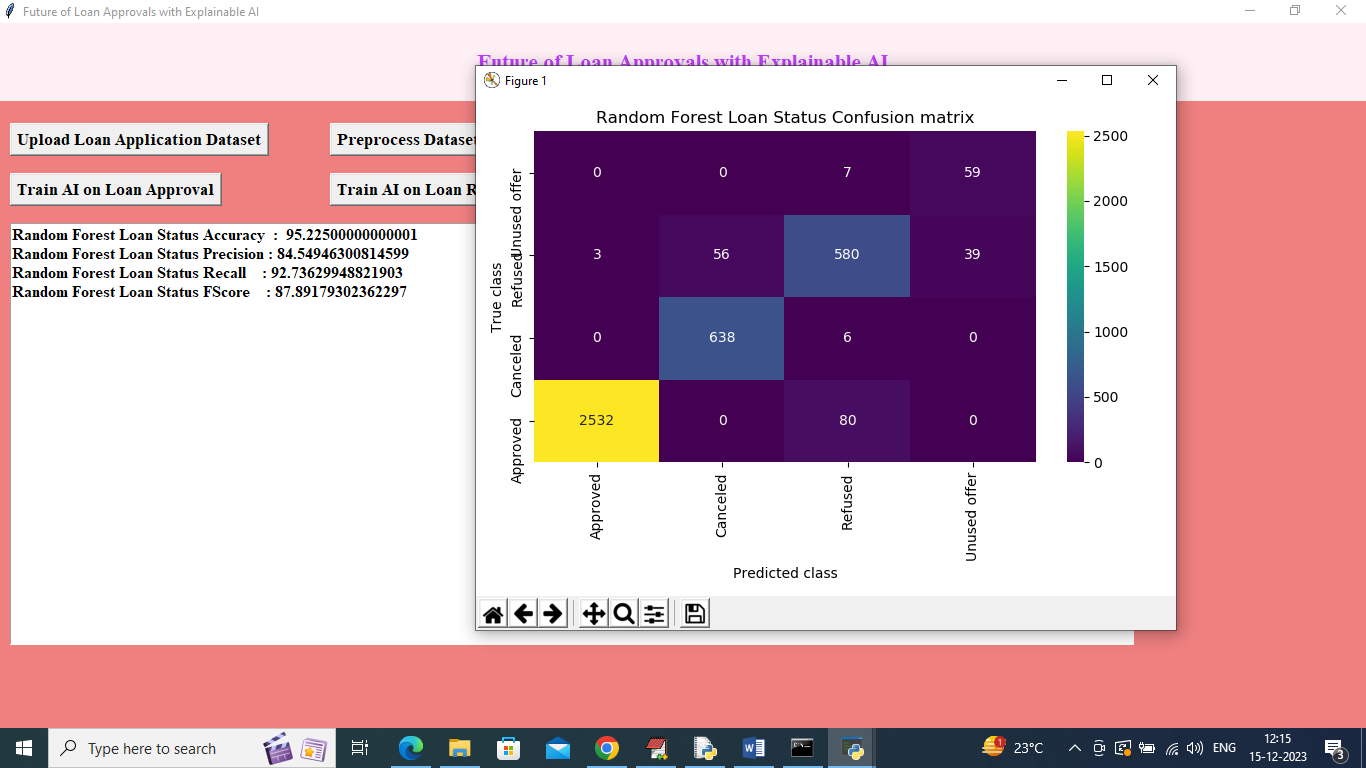
In above screen dataset loaded and in text area can see few records from dataset and in first graph x-axis represents LOAN STATUS and y-axis represents Number of Records available in that LOAN STATUS class label. In second graph x-axis represents REJECTION REASON and y-axis represents records size and in dataset we have both numeric and non-numeric values so to convert to numeric data then click on ‘Pre-process Dataset’ button to get below output



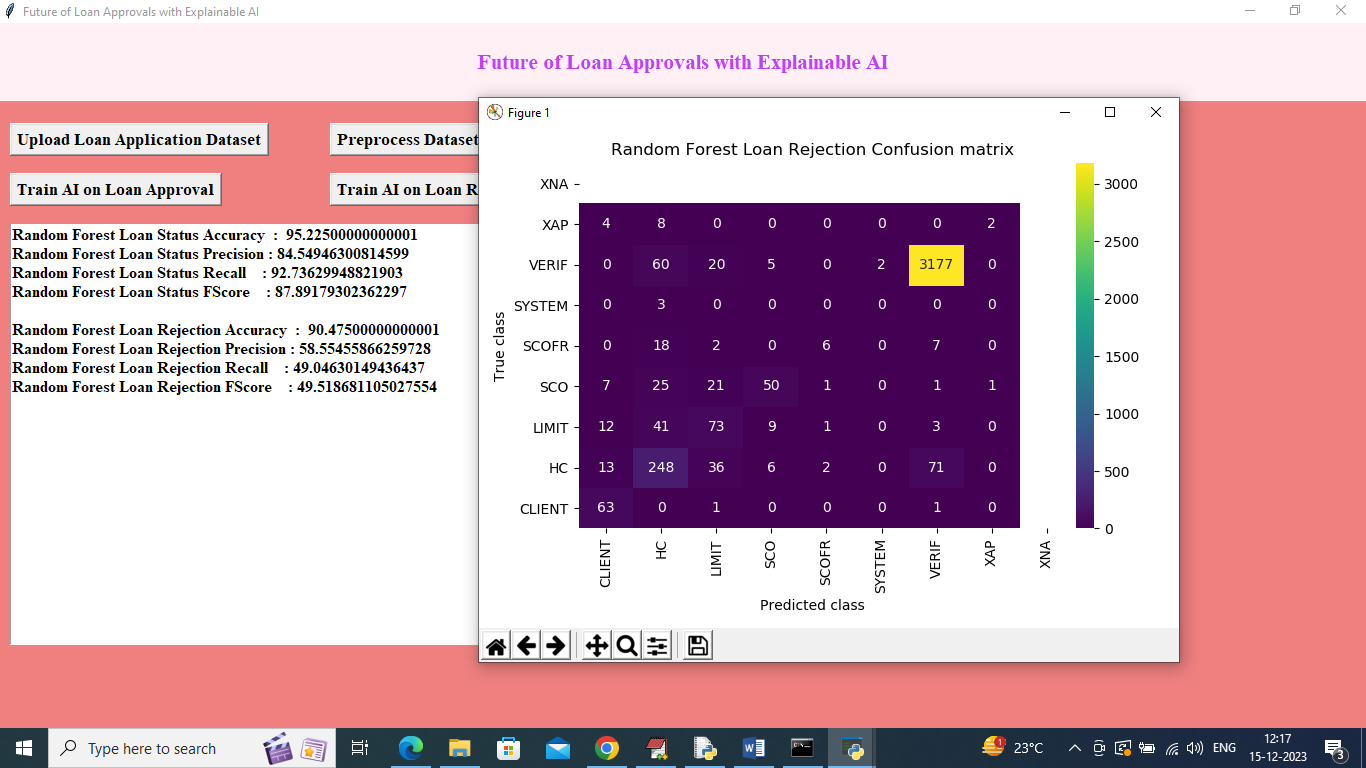
In above screen dataset converted to numeric format and then click on ‘Split Dataset Train & Test’ button to split dataset into train and test and then will get below output



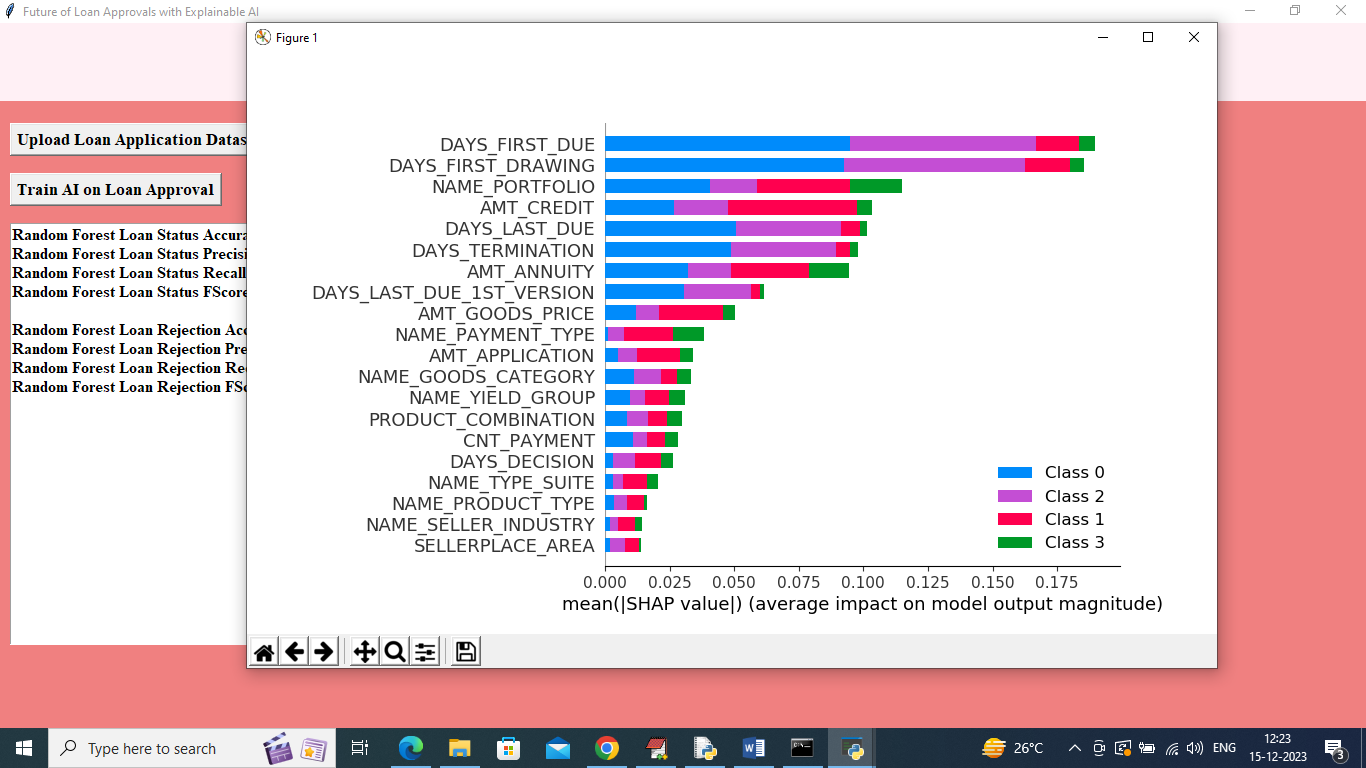
In above screen can see dataset size with total number of features and then can see TRAIN and TEST size and now click on ‘Train AO on Loan Status Approval’ button to train AI and get below output



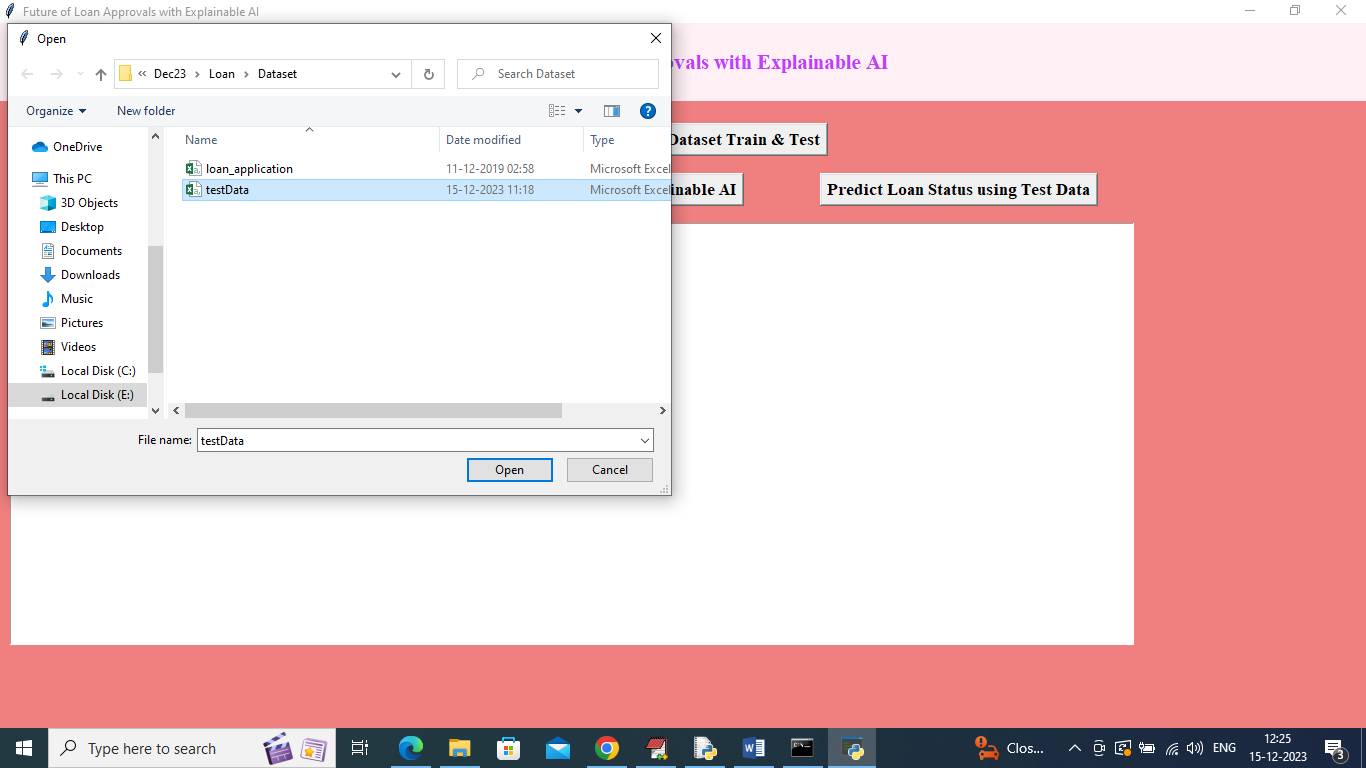
In above screen AI Random Forest got 95% accuracy on Loan STATUS and can see other metrics also. In above confusion matrix graph x-axis represents ‘LOAN STATUS Predicted Labels’ and y-axis represents TRUE labels and all boxes in diagnol contains correct prediction count and remaining blue boxes contains incorrect prediction count which are very few. Now click on ‘Train AI on Loan Rejections’ button to train AI on rejection reason and get below output



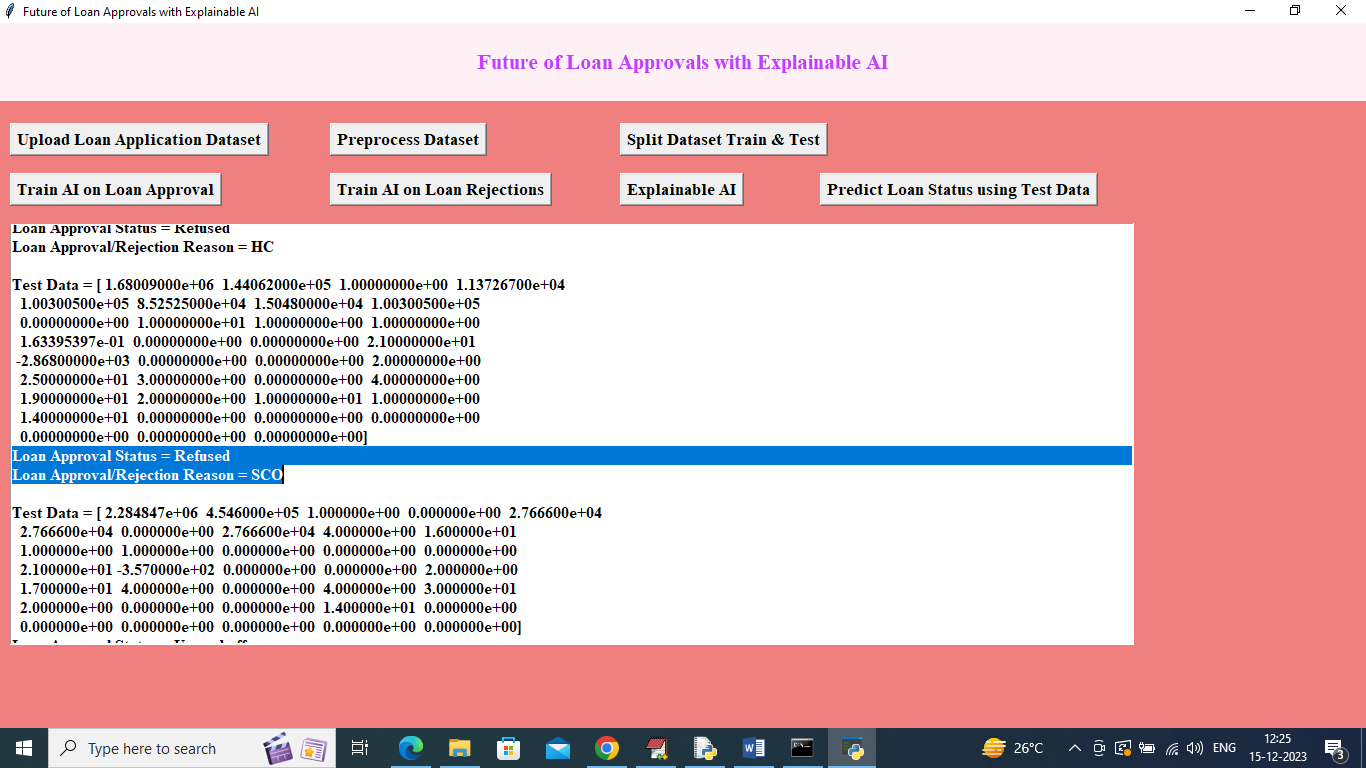
In above screen AI on REJECTION got 90% accuracy and in confusion matrix graph x-axis represents ‘Rejection Reason Predicted Labels’ and y-axis represents True label and in diagnol boxes we can see correct prediction count and remaining boxes contains incorrect prediction count. Now click on ‘Explainable AI’ button to get below features explanation on prediction



In above SHAP explanation screen in each bar we can see 4 different colours and each colour represents one class label and based on colour percentage we can say which feature names is contributing how much to predict that class label. Now close above graph and then click on ‘Predict Loan Status using Test Data’ button to upload test data and then will get below prediction



In above screen selecting and uploading testData.csv file and then click on ‘Open’ button to get below output

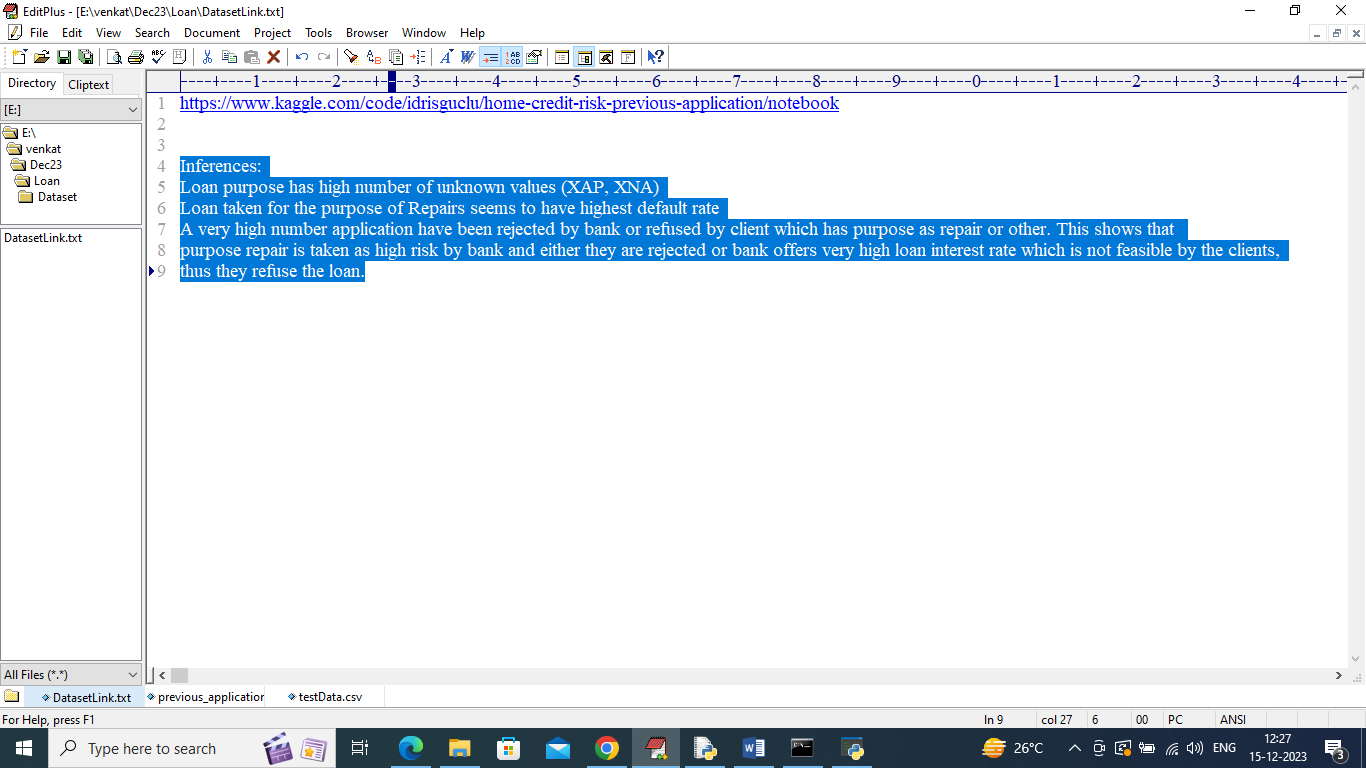


In above screen in square bracket we can see test data and then in blue colour selected line next to TEST data we can see LOAN STATUS prediction and REASON details. Scroll down above output to view all predictions



In above screen can see other prediction output.

For Reason Rejected code you can read below description



In above screen read blue colour selected text to know about REJECTED REASON codes

**6. VALIDATION**

## 6.VALIDATION

The validation of this project primarily relies on extensive testing and well-defined test cases to ensure the accuracy and effectiveness of the inappropriate content detection system. The testing process involves multiple stages, including dataset validation, model performance evaluation, and real-world testing. By implementing a structured validation approach, we can ensure that the system consistently delivers high accuracy in detecting inappropriate content while minimizing false positives and false negatives.

### INTRODUCTION

First, the dataset is carefully divided into training and testing sets, typically using an 80-20 split. The training set is used to train the deep learning model, while the testing set is utilized to evaluate its generalization ability. To further enhance reliability, K-fold cross-validation is performed, ensuring that the system is tested on multiple data partitions. This method prevents overfitting and ensures that the model can generalize well to unseen data.

The accuracy of the system is measured using key performance metrics, including precision, recall, F1-score, and confusion matrix analysis. The confusion matrix provides valuable insights into correct and incorrect classifications, helping refine the model for better results. Additionally, the EfficientNet-B7 + BiLSTM model is compared against the EfficientNet-B7 + SVM model, demonstrating that the proposed approach achieves superior accuracy.

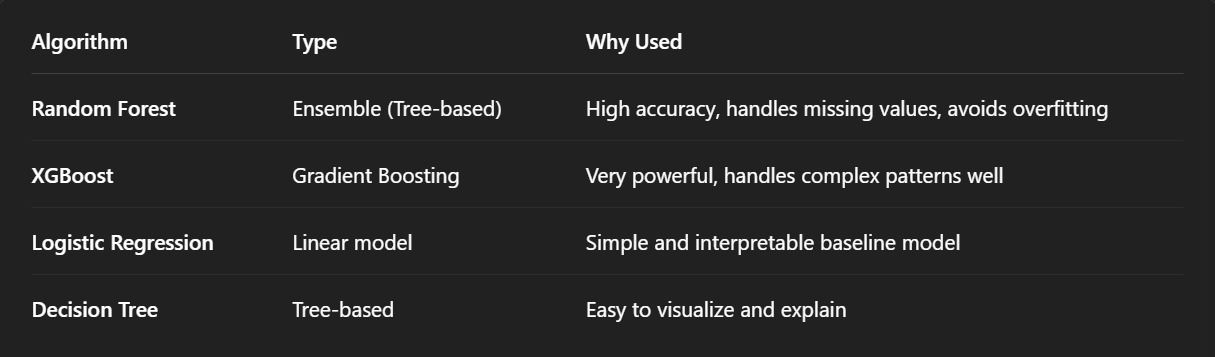
Finally, real-world deployment testing is conducted to simulate live content moderation, ensuring that the system approvals the loans. Continuous improvements are made based on test results, allowing the model to remain effective in approvals of loan. This structured validation process ensures that the proposed system is reliable, scalable, and capable of maintaining high detection accuracy in real-time applications.

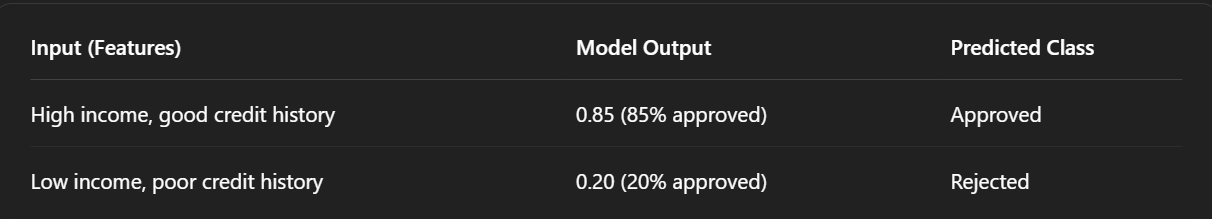
**6.2 TEST CASES**

**TABLE 1 UPLOADING DATASET**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Test case ID | Test case name | Purpose | Test Case | Output |
|  |  |  |  |  |
| 1 |  |  |  |  |
|  | Upload a valid CSV dataset file | prediction. | Upload a valid CSV dataset file | File uploaded successfully” message |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**TABLE 2 CLASSIFICATION**





# 7. CONCLUSION & FUTURE ASPECTS

**7.CONCLUSION & FUTURE ASPECTS**

In conclusion, the project has successfully achieved its objectives, showcasing significant progress and outcomes. The implementation and execution phases were meticulously planned and executed, leading to substantial improvements and insights. Looking ahead, the future aspects of the project hold immense potential. Future developments will focus on expanding the scope, integrating new technologies, and enhancing sustainability. These advancements will not only strengthen the existing framework but also open new avenues for growth and innovation, ensuring the project remains relevant and impactful in the long term. This strategic approach will drive continuous improvement and success.

### 7.1 PROJECT CONCLUSION

In this project, we successfully developed an intelligent and transparent loan approval system using Explainable Artificial Intelligence (XAI). Traditional loan approval models often act as black boxes, offering little to no insight into why a decision was made. Our solution addresses this limitation by combining powerful machine learning algorithms like **Random Forest** and **XGBoost** with interpretability tools such as **SHAP** and **LIME**. This not only ensures accurate predictions but also provides clear, human-understandable explanations for each decision.

The system enhances trust and fairness by offering transparency to both loan officers and applicants. Bias detection modules further ensure that decisions are free from unfair influence based on sensitive attributes like gender or age. The project demonstrates how AI, when made explainable, can transform financial decision-making processes into systems that are **more accurate, fair, and responsible**.

Overall, this project proves that **Explainable AI is the future of ethical automation in finance**, offering a perfect balance between advanced technology and human accountability.

### 7.2 FUTURE ASPECTS

The future of loan approvals with Explainable AI holds significant potential to revolutionize the financial sector by making decision-making processes more transparent, efficient, and user-centric. As financial institutions increasingly adopt AI, integrating real-time explainable models can help streamline loan processing, reduce manual interventions, and provide immediate feedback to applicants. These systems can be extended into mobile applications, allowing users to check their loan eligibility and receive clear, understandable reasons for approval or rejection — all within seconds. Furthermore, deep learning techniques and dynamic risk profiling can be incorporated to capture more complex financial patterns and offer highly personalized credit decisions.

Looking ahead, enhancing the fairness and accountability of AI systems will be a major focus. Future systems can include automated bias monitoring, ensuring that no decisions are influenced by unfair factors such as gender or ethnicity. Additionally, federated learning can be used to train models across multiple banks without sharing sensitive data, leading to more robust and generalizable systems. Explainable AI frameworks can also be expanded to cover other financial services like credit scoring, insurance underwriting, and fraud detection, creating a unified ecosystem of ethical, intelligent decision-making tools in the financial industry.

# BIBLIOGRAPHY

1. **BIBLIOGRAPHY**

### REFERENCES

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). **“Why Should I Trust You?”: Explaining the Predictions of Any Classifier**. *Proceedings of the 22nd ACM SIGKDD*, 1135–1144.  
https://doi.org/10.1145/2939672.2939778

Lundberg, S. M., & Lee, S.-I. (2017). **A Unified Approach to Interpreting Model Predictions**. *Advances in Neural Information Processing Systems (NeurIPS)*.  
<https://arxiv.org/abs/1705.07874>

Molnar, C. (2022). **Interpretable Machine Learning**. https://christophm.github.io/interpretable-ml-book/

Kaggle. (n.d.). **Loan Prediction Dataset**. Available at: https://www.kaggle.com/datasets

Microsoft Research. (2018). **Fairlearn: Fairness-Aware Machine Learning Toolkit**. <https://fairlearn.org/>

UCI Machine Learning Repository. **Statlog (German Credit Data) Dataset**. Available at: <https://archive.ics.uci.edu/ml/datasets/Statlog+(German+Credit+Data)>

Brownlee, J. (2020). **Machine Learning Mastery With Python**. Machine Learning Mastery Publications.

IBM Cloud Education. (2021). **What is Explainable AI (XAI)?** https://www.ibm.com/cloud/learn/explainable-ai