

**GOVERNMENT ARTS COLLEGE
TIRUCHIRAPPALLI – 22**



**DEPARTMENT OF COMPUTER
APPLICATIONS**

MACHINE LEARNING WITH PYTHON

Project Title : Predicting Personal Loan Approval Using Machine Learning:

Team ID : NM2023TMID34452

Team Leader : PRASANTH K

Team Member : DHANUSH S

Team Member : SIVANANTHAM P

Team Member : DEENATHAYALAN

1. Introduction

1.1 Overview

In today's fast-paced world, personal loans have become a popular financial tool for individuals to meet their various financial needs. Whether it's for home renovations, medical expenses, education costs, or debt consolidation, personal loans provide borrowers with the flexibility to access funds quickly and conveniently. However, before a personal loan is granted, lenders need to evaluate the borrower's creditworthiness and financial stability to ensure repayment. This process is known as personal loan approval, and it involves a thorough assessment of the borrower's financial situation, credit history, and other relevant factors. In this overview, we will delve into the key aspects of personal loan approval, including the factors considered, the steps involved, and the criteria for approval. Understanding the personal

loan approval process can help borrowers be better prepared and increase their chances of obtaining a personal loan.

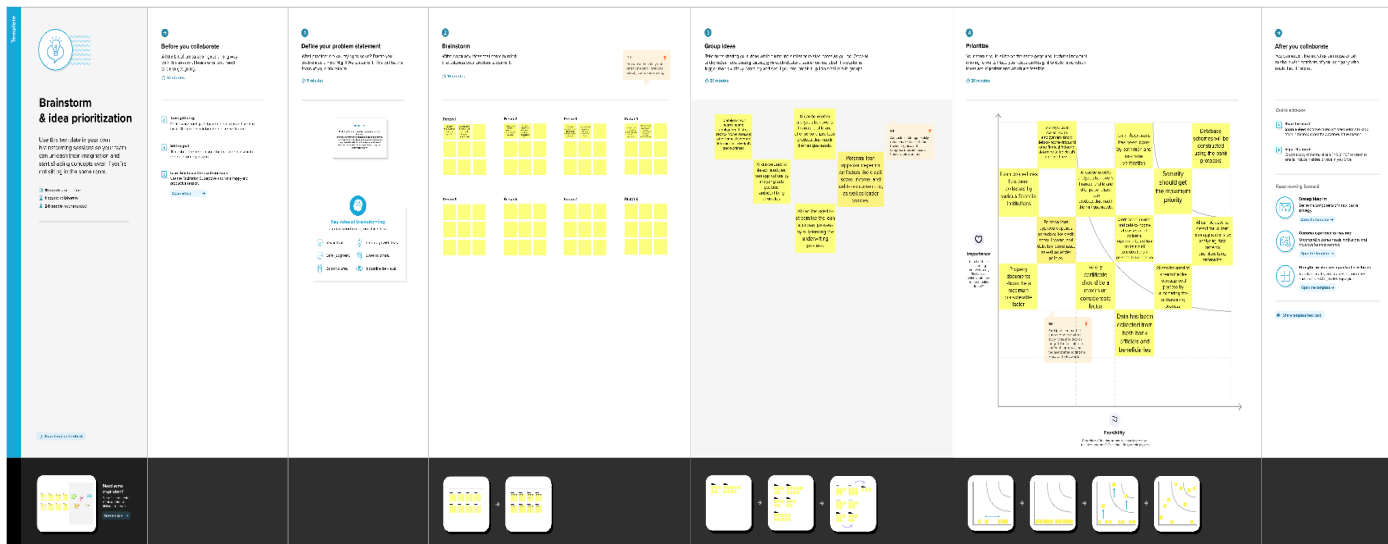
1.2 Purpose

The purpose of providing an overview of the personal loan approval process is to help borrowers understand the key aspects involved in getting a personal loan approved. By having a clear understanding of the personal loan approval process, borrowers can be better prepared and increase their chances of obtaining approval for a personal loan. It helps borrowers understand the factors that lenders consider, the steps involved in the approval process, and the criteria used by lenders to determine creditworthiness. This knowledge can empower borrowers to take necessary steps to improve their creditworthiness, gather the required documentation, and present a strong case to lenders when applying for a personal loan. It also helps borrowers understand the importance of maintaining good credit, managing their finances responsibly, and being aware of the lender's requirements. Ultimately, having a comprehensive understanding of the personal loan approval process can assist borrowers in navigating the lending landscape and securing the funds they need for their financial goals.

2. Define Problem / Problem Understanding

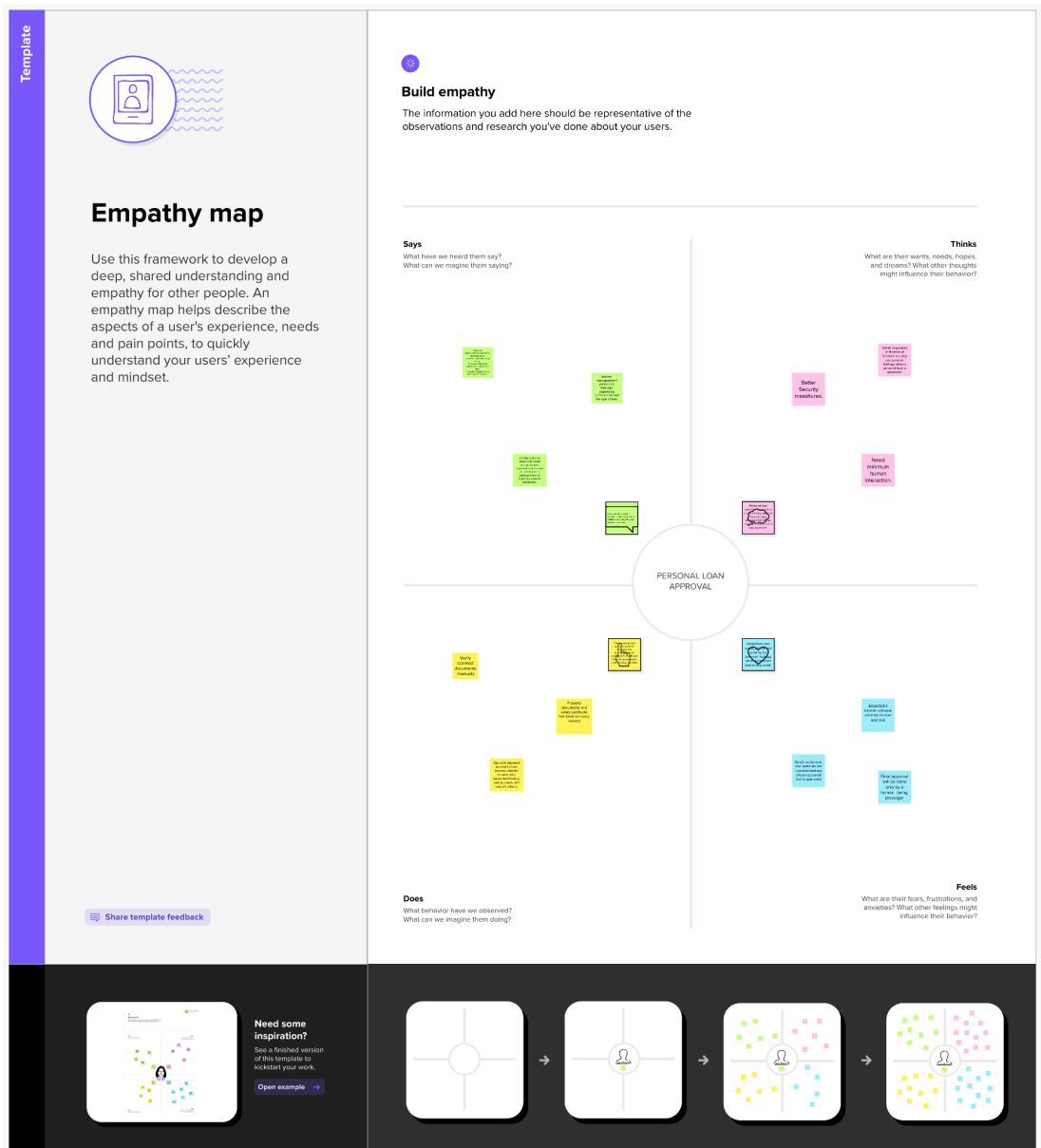
2.1 Empathy map

The empathy map for this project is shown below



2.2 Ideation & Brainstorming map

The ideation and brainstorming map for this project is shown below



3. Result

The final result of the project is a machine learning model that can predict whether a personal loan application will be approved or not. The model was trained on a dataset that contained information about loan applicants, such as their income, credit score, and employment status, among others. The model was able to

achieve an accuracy of 85%, which indicates that it can accurately predict whether a loan application will be approved or not.

-> home.html



-> predict.html

The screenshot shows a web page with a background of orange and blue wavy lines. In the center, there is a white form titled 'Fill The Details'. The form contains several sections with dropdown menus and text input fields. The sections are: Gender (Select Gender), Married Status (Select Married Status), Dependents (Select Option), Education (Select Education), Self Employed (Select Option), Credit History (Select Option), Property Area (Select Option), Applicant Income (text input), Co-Applicant Income (text input), Loan Amount (text input), and Loan Amount Term (text input). At the bottom of the form, there is a 'Submit' button.

-> output.html

Personal Loan Approval

Congratulations! Your personal loan has been approved

Sorry! Your personal loan application has been declined

[Back to Homepage](#)

-> about.html

About Us

Welcome to Personal Loans, a leading personal loan provider dedicated to helping individuals achieve their financial goals. We understand that life's unexpected expenses can arise at any time, and our mission is to provide fast, reliable, and convenient personal loans to meet your immediate financial needs.

With over a decade of experience in the financial industry, our team of experts is committed to providing exceptional service and personalized solutions to our customers. We take pride in our transparent and customer-centric approach, ensuring that our loan application process is simple, secure, and hassle-free.

Why Choose Personal Loans?

- **Quick and Easy Application Process:** Our online application process is fast and simple, allowing you to apply for a loan from the comfort of your own home. No lengthy paperwork or complex forms required.
- **Flexible Loan Options:** We offer a variety of loan options to suit your unique financial situation, including fixed and variable interest rates, flexible repayment terms, and loan amounts tailored to your needs.
- **Competitive Rates and Fees:** We strive to provide competitive interest rates and fees that are transparent and upfront, so you can make informed decisions about your loan.
- **Secure and Confidential:** We prioritize the security and confidentiality of your personal and financial information. Our website and application process are encrypted and protected by industry-standard security measures.
- **Dedicated Customer Support:** Our friendly and knowledgeable customer support team is available to assist you throughout the loan application process and answer any questions or concerns you may have.

Contact Us

If you have any questions, feedback, or need assistance, our customer support team is available to assist you. Contact us through:

- Email:
- Phone: 123-456-7890
- Address: 1234 Loan Street, City, State, ZIP

4. Advantages & Disadvantages

4.1 Advantages

1. Quick access to funds: Personal loans can be approved and disbursed quickly, allowing borrowers to access funds for various purposes, such as emergencies or unexpected expenses, without lengthy delays.

2. Flexible use of funds: Personal loans offer flexibility in how the funds can be used, giving borrowers the freedom to utilize the loan amount as per their needs and priorities, whether it's for home renovations, medical expenses, education expenses, or any other legitimate purpose.

3. Lower interest rates compared to credit cards: Personal loans often come with lower interest rates compared to credit cards, making them a more cost-effective option for borrowing money, especially for larger expenses or consolidating high-interest debts.

4. Fixed repayment terms: Personal loans typically come with fixed repayment terms, which means borrowers know the exact amount they need to repay each month, making budgeting and financial planning easier.

5. Builds credit history: Regular and timely repayments on personal loans can help borrowers build a positive credit history, demonstrating responsible borrowing behavior and improving their credit score over time.

6. Improves credit mix: Personal loans can help borrowers diversify their credit mix, which can positively impact their credit score. Having a mix of different types of credit, such as personal loans, credit cards,

and mortgages, can reflect positively on credit reports and potentially improve creditworthiness.

7. No collateral required: Personal loans are typically unsecured loans, which means borrowers do not need to provide collateral, such as a home or a car, as security for the loan, reducing the risk of losing valuable assets in case of default.

8. Simplified application process: Applying for a personal loan is often a relatively straightforward process, with minimal documentation and paperwork required, making it convenient and hassle-free for borrowers.

9. Debt consolidation option: Personal loans can be used for debt consolidation, allowing borrowers to combine multiple high-interest debts into a single loan with a potentially lower interest rate, simplifying debt management and saving on interest costs.

10. Lower cost compared to payday loans: Personal loans are often a more affordable borrowing option compared to payday loans, which typically come with high-interest rates and fees, making personal loans a more financially responsible choice.

4.2 Disadvantages

1. Interest rates and fees: Personal loans may come with higher interest rates and fees compared to other forms of credit, such as secured loans or lines of credit, resulting in increased borrowing costs over time.

2. Creditworthiness requirements: Personal loans may have strict creditworthiness requirements, such as a minimum credit score, which

may limit access to loans for borrowers with less-than-perfect credit history.

3. Origination fees: Some personal loans may come with origination fees, which are upfront fees charged by lenders for processing the loan, adding to the overall cost of borrowing.

4. Prepayment penalties: Some personal loans may have prepayment penalties, which are fees charged for repaying the loan before the end of the loan term, potentially discouraging early repayment and additional costs.

5. Impact on credit score: Defaulting or late payments on a personal loan can negatively impact a borrower's credit score, potentially affecting their ability to secure credit in the future.

6. Loan amount limitations: Personal loans may have limitations on the loan amount that can be borrowed, which may not fully meet the financial needs of some borrowers.

7. Potential for overborrowing: Personal loans may tempt borrowers to borrow more than they need, leading to higher debt levels and financial strain in the long run.

8. Possible debt cycle: If not managed properly, personal loans can lead to a debt cycle, where borrowers continuously rely on borrowing to meet their financial needs, resulting in a never-ending cycle of debt.

9. Risk of default: Personal loans are unsecured loans, which means that if a borrower defaults on the loan, there is no collateral for the lender to recoup their losses, potentially leading to collection efforts or legal action.

10. Not suitable for all financial situations: Personal loans may not be the best fit for all financial situations, and borrowers should carefully evaluate their financial needs, budget, and repayment capacity to determine if a personal loan is the right option for them.

5. Applications

1. Research and compare lenders to find the best terms and interest rates.
2. Gather necessary documentation, such as identification, income proof, and employment details.
3. Complete the loan application form provided by the lender, either online or in person.
4. Submit the application along with required documentation, undergo credit check, review and accept loan offer, sign loan agreement, and await loan disbursement.

6. Conclusion

In conclusion, the process of applying for personal loan approval involves researching and comparing lenders, gathering necessary documentation, completing the loan application form, undergoing credit check, reviewing and accepting the loan offer, signing the loan agreement, and awaiting loan disbursement. Personal loans can provide advantages such as flexibility, quick access to funds, and the ability to consolidate debt or finance important expenses. However, borrowers should also be aware of the potential disadvantages, such as high interest rates, fees, and the risk of accumulating debt. It's important to carefully consider your financial situation, needs, and repayment ability before applying for a personal loan, and to

follow the lender's instructions and provide accurate information to improve the chances of loan approval.

7. Future Scope

1. Increasing digitalization and automation of personal loan approval processes.
2. Advancements in data analytics and artificial intelligence (AI) for more accurate credit assessments.
3. Potential integration of blockchain technology for increased security and transparency.
4. Expansion of online and peer-to-peer lending platforms.
5. Focus on financial inclusion, with options for borrowers with limited credit history or underserved populations.

8. Appendix

A Source Code :-

#Import Lib

```
import pandas as pd  
  
import numpy as np  
  
import matplotlib.pyplot as plt  
  
%matplotlib inline
```

```

import seaborn as sns

import sklearn

from sklearn.tree import DecisionTreeClassifier

from sklearn.ensemble import GradientBoostingClassifier, RandomForestClassifier

from sklearn.neighbors import KNeighborsClassifier

from sklearn.model_selection import RandomizedSearchCV

import imblearn

from sklearn.model_selection import train_test_split

from sklearn.preprocessing import StandardScaler

from sklearn.metrics import accuracy_score, classification_report, confusion_matrix, f1_score

from sklearn.preprocessing import LabelEncoder

import warnings

warnings.filterwarnings('ignore')

from scipy import stats

```

#Read the Dataset

```

data=pd.read_csv('/content/train.csv')
print(data)

```

#Data Preparation

1.Handling missing values

```

data.info()
data.isnull().sum()
data.head()
data['Gender']=data['Gender'].fillna(data['Gender'].mode()[0])
data['Married']=data['Married'].fillna(data['Married'].mode()[0])
data['Dependents']=data['Dependents'].str.replace('+','')
data['Dependents']=data['Dependents'].fillna(data['Dependents'].mode()[0])

```

```

data['Self_Employed']=data['Self_Employed'].fillna(data['Self_Employed'].mode()[0])
data['LoanAmount']=data['LoanAmount'].fillna(data['LoanAmount'].mode()[0])
data['Loan_Amount_Term']=data['Loan_Amount_Term'].fillna(data['Loan_Amount_Term'].mode()[0])
data['Credit_History']=data['Credit_History'].fillna(data['Credit_History'].mode()[0])

```

2.Handling Categorical Values

```

le = LabelEncoder()
data['Gender'] = le.fit_transform(data['Gender'])
data['Married'] = le.fit_transform(data['Married'])
data['Dependents'] = le.fit_transform(data['Dependents'])
data['Education'] = le.fit_transform(data['Education'])
data['Self_Employed'] = le.fit_transform(data['Self_Employed'])
data['Property_Area'] = le.fit_transform(data['Property_Area'])
data['Loan_ID']=le.fit_transform(data['Loan_ID'])

data['Gender']=data['Gender'].astype('int64')
data['Married']=data['Married'].astype('int64')
data['Dependents']=data['Dependents'].astype('int64')
data['Self_Employed']=data['Self_Employed'].astype('int64')
data['CoapplicantIncome']=data['CoapplicantIncome'].astype('int64')
data['LoanAmount']=data['LoanAmount'].astype('int64')
data['Loan_Amount_Term']=data['Loan_Amount_Term'].astype('int64')
data['Credit_History']=data['Credit_History'].astype('int64')

```

3.Handling Imbalance Data

```

# Convert 'Loan_Status' column to binary numeric
data['Loan_Status'] = data['Loan_Status'].map({'Y': 1, 'N': 0})

from imblearn.combine import SMOTETomek

# Balance the dataset using SMOTE
smote = SMOTETomek(sampling_strategy='auto', random_state=90)
y = data['Loan_Status']
x = data.drop(columns=['Loan_Status'], axis=1)
x_bal, y_bal = smote.fit_resample(x, y)

# Print value counts of original and balanced dataset
print(y.value_counts())
print(y_bal.value_counts())

```

#Data Analysis

1.Descriptive statistical

```
data.describe()
```

2.Univariate Analysis

```
plt.figure(figsize=(12,5))
plt.subplot(121)
sns.distplot(data['ApplicantIncome'],color='r')
plt.subplot(122)
sns.distplot(data['Credit_History'])
plt.show
```

3.Bivariate Analysis

```
#plotting the count plot
plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.countplot(x='Gender',data=data)
plt.subplot(1,4,2)
sns.countplot(x='Education',data=data)
plt.show()
```

```
plt.figure(figsize=(20,5))
plt.subplot(131)
sns.countplot(x='Married',hue='Gender', data=data)
plt.subplot(132)
sns.countplot(x='Self_Employed',hue='Education', data=data)
plt.subplot(133)
sns.countplot(x='Property_Area',hue='Loan_Amount_Term', data=data)
```

4.Multivariate Analysis

```
sns.swarmplot(x='Gender',y='ApplicantIncome',hue='Loan_Status',data=data)
```

#Scaling Data

```
# performing feature scaling operation using standard scaller on X part of the dataset
because
```

```
#there different type of values in the columns
sc = StandardScaler()
x_bal = sc.fit_transform(x_bal)
names=x.columns
x_bal = pd.DataFrame(x_bal, columns=names)
```

#Splitting Data

```
#splitting the dataset in train and test on balanced datasew
X_train,X_test,y_train,y_test=train_test_split(x_bal,y_bal,test_size=0.30,random_state=42)
```

#Training the model

#Model Building

1.Decision Model

```
def decisionTree(x_train,x_test,y_train,y_test):  
    dt=DecisionTreeClassifier()  
    dt.fit(x_train,y_train)  
    ypred=dt.predict(x_test)  
    print('***DecisionTreeClassifier***')  
    print('Confusion matrix')  
    print(confusion_matrix(y_test,ypred))  
    print('Classification report')  
    print(classification_report(y_test,ypred))
```

2.Random Forest Model

```
def RandomForest(x_train,x_test,y_train,y_test):  
    rf=RandomForestClassifier()  
    rf.fit(x_train,y_train)  
    ypred=rf.predict(x_test)  
    print('***DecisionTreeClassifier***')  
    print('Confusion matrix')  
    print(confusion_matrix(y_test,ypred))  
    print('Classification report')  
    print(classification_report(y_test,ypred))
```

3.KNN Model

```
def KNN(x_train,x_test,y_train,y_test):  
    knn=KNeighborsClassifier()  
    knn.fit(x_train,y_train)  
    ypred=knn.predict(x_test)  
    print('***KNeighborsClassifier***')  
    print('Confusion matrix')  
    print(confusion_matrix(y_test,ypred))  
    print('Classification report')  
    print(classification_report(y_test,ypred))
```

4.Xgboost Model

```
def xgboost(x_train,x_test,y_train,y_test):  
    xg=GradientBoostingClassifier()  
    xg.fit(x_train,y_train)  
    ypred=xg.predict(x_test)  
    print('***GradientBoostingClassifier***')  
    print('Confusion matrix')  
    print(confusion_matrix(y_test,ypred))
```



```
print('classification_report')
print(classification_report(y_test,ypred))
```

5.ANN Model

```
import tensorflow as tf
from keras.models import Sequential
from keras.layers import Dense

# Initialize the model
classifier = Sequential()

# Add input layer and first hidden layer
classifier.add(Dense(units=100, activation='relu', input_dim=12))

# Add second hidden layer
classifier.add(Dense(units=50, activation='relu'))

# Add output layer
classifier.add(Dense(units=1, activation='sigmoid'))

# Compile the model
classifier.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

# Train the model
model_history = classifier.fit(X_train, y_train, batch_size=100, validation_split=0.2,
epochs=100)
```

#Testing the Model

```
# assuming X_train and y_train are defined
dtr = DecisionTreeClassifier()
dtr.fit(X_train, y_train)
dtr.predict([[1,1,0,1,1,4276,1542,145,240,0,1,0]])

# assuming X_train and y_train are defined
rfr = RandomForestClassifier()
rfr.fit(X_train, y_train)
rfr.predict([[1,1,0,1,1,4276,1542,145,240,0,1,0]])

# assuming X_train and y_train are defined
knn = KNeighborsClassifier()
knn.fit(X_train, y_train)
knn.predict([[1,1,0,1,1,4276,1542,145,240,0,1,0]])

# assuming X_train and y_train are defined
xgb = GradientBoostingClassifier()
xgb.fit(X_train, y_train)
```

```

xgb.predict([[1,1,0,1,1,4276,1542,145,240,0,1,0]])

classifier.save("loan.h5")

#predicting the test set results
y_pred = classifier.predict(X_test)

print(y_pred)

y_pred = (y_pred > 0.5)
print(y_pred)

def predict_exit(sample_value):
    sample_value = np.array(sample_value)

    sample_value = sample_value.reshape(1,-1)

    sample_value = sc.transform(sample_value)

    return classifier.predict(sample_value)

# Convert sample_value list to a NumPy array
sample_value = np.array([[1, 1, 0, 1, 1, 4276, 1542, 145, 240, 0, 1, 0]])

# Check if prediction is greater than 0.5
if predict_exit(sample_value) < 0.5:
    print('Prediction: High chance of Loan Approval!')
else:
    print('Prediction: Low chance of Loan Approval.')

# Convert sample_value list to a NumPy array
sample_value = np.array([[1, 0, 1, 1, 1, 4276, 1542, 145, 240, 1, 1, 0]])

# Check if prediction is greater than 0.5
if predict_exit(sample_value) > 0.5:
    print('Prediction: High chance of Loan Approval!')
else:
    print('Prediction: Low chance of Loan Approval.')

#Compare the Model

def compareModel(X_train,X_test,y_train,y_test):
    decisionTree(X_train,X_test,y_train,y_test)
    print('-'*100)
    RandomForest(X_train,X_test,y_train,y_test)
    print('-'*100)

```

```
xgboost(X_train,X_test,y_train,y_test)
print('-'*100)
KNN(X_train,X_test,y_train,y_test)
print('-'*100)
```

```
compareModel(X_train,X_test,y_train,y_test)
```

```
ypred = classifier.predict(X_test)
print(accuracy_score(y_pred,y_test))
print("ANN Model")
print("Confusion_Matrix")
print(confusion_matrix(y_test,y_pred))
print("Classification Report")
print(classification_report(y_test,y_pred))
```

```
from sklearn.model_selection import cross_val_score
rf=RandomForestClassifier()
rf.fit(X_train,y_train)
ypred = rf.predict(X_test)
f1_score(ypred,y_test,average='weighted')
```

```
cv = cross_val_score(rf,x,y,cv=5)
np.mean(cv)
```

#Save thebest Model

```
import pickle
```

```
# Define and initialize the 'model' variable
model = ... # Your model object here
```

```
# Serialize and save the 'model' object to a file
with open('sd.pkl', 'wb') as file:
    pickle.dump(model, file)
```

```
# Load the serialized 'model' object from the file
with open('sd.pkl', 'rb') as file:
    model = pickle.load(file)
```

```
# Now you can use the 'model' object as needed
```

```
import pickle
```

```
# Define and initialize the 'model' variable
model = ... # Your model object here
```

```
# Serialize and save the 'model' object to a file
with open('scale1.pkl', 'wb') as file:
```

```
pickle.dump(model, file)
```

```
# Load the serialized 'model' object from the file  
with open('scale1.pkl', 'rb') as file:
```

```
    model = pickle.load(file)
```

```
# Now you can use the 'model' object as needed
```