

# Predicting Personal Loan Approval Using Machine Learning

Project report submitted to  
**“Madurai Kamaraj University”**  
in partial fulfillment of the requirements  
for the award of the Degree of

**Bachelor of Science**

**in**

**Computer Science**

*Submitted By*

**TEAM ID : NM2023TMID31090**  
**TEAM LEADER : C.SELVAKUMAR**  
**TEAM MEMBER : S.RAGULGANDHI**  
**V.PRAKASH**  
**T.VALLARASU**

*Under the Guidance of*

**Dr.B.UMADEVI,M.Sc., M.Phil., Ph.D.**  
**(Assistant Professor, GAC, Melur)**



**GOVERNMENT ARTS COLLEGE**  
**PG & DEPARTMENT OF COMPUTER**  
**SCIENCE MELUR – 625 106**

# INDEX

SI.NO.	TITLE	PAGE.NO.
1	INTRODUCTION 1.1OVERVIEW 1.2PURPOSE	2 2 3
2	PROBLEM DEFINITION & DESIGN THINKING 2.1 EMPATHY MAP 2.2 IDEATION & BRAINSTOMING MAP	4  4 5
3	RESULT	6
4	ADVANTAGES & DISADVANTAGES	11
5	APPLICATIONS	12
6	CONCLUSION	13
7	FUTURE SCOPE	13
8	APPENDIX 8.1 SOURCE CODE	14

# Predicting Personal Loan Approval Using Machine Learning

## 1. INTRODUCTION :

### 1.1 OVERVIEW :

A loan is a sum of money that is borrowed and repaid over a period of time, typically with interest. There are various types of loans available to individuals and businesses, such as personal loans, mortgages, auto loans, student loans, business loans and many more. They are offered by banks, credit unions, and other financial institutions, and the terms of the loan, such as interest rate, repayment period, and fees, vary depending on the lender and the type of loan.

A personal loan is a type of unsecured loan that can be used for a variety of expenses such as home repairs, medical expenses, debt consolidation, and more. The loan amount, interest rate, and repayment period vary depending on the lender and the borrower's creditworthiness. To qualify for a personal loan, borrowers typically need to provide proof of income and have a good credit score.

Predicting personal loan approval using machine learning analyses a borrower's financial data and credit history to determine the likelihood of loan approval. This can help financial institutions to make more informed decisions about which loan applications to approve and which to deny.

## 1.2 PURPOSE :

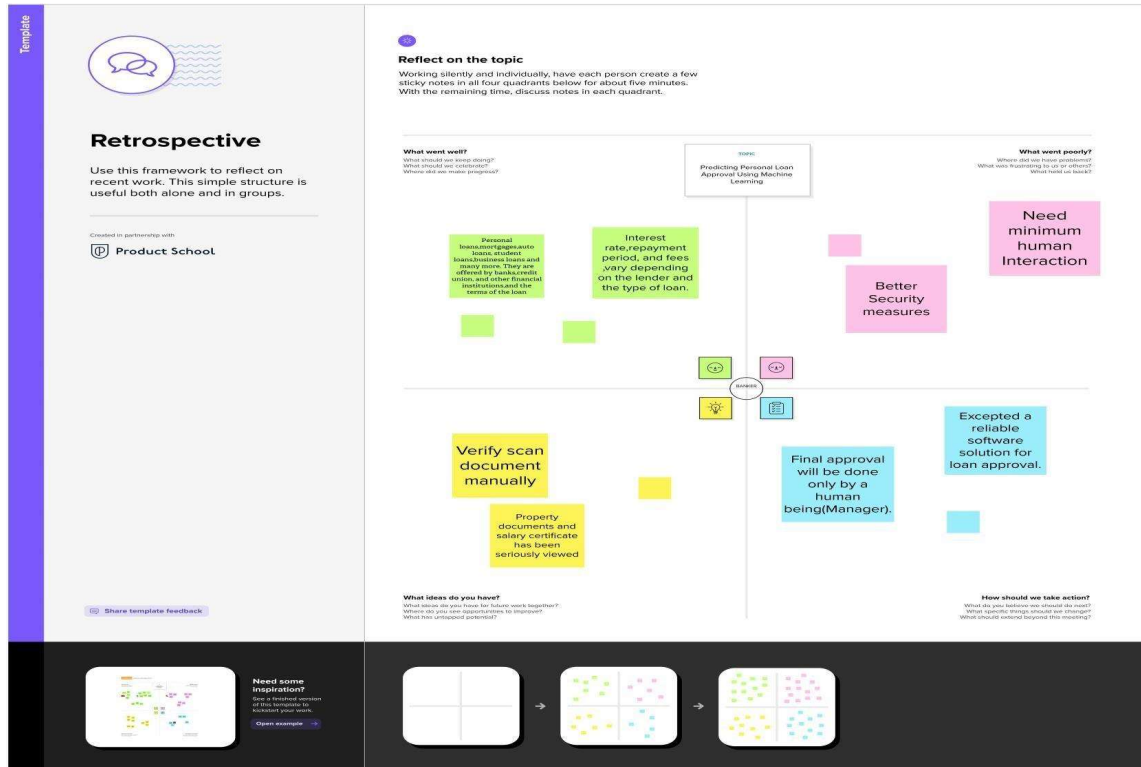
Predicting Personal Loan Approva Using Machine Learning can have several benefits for telecom companies, including:

- **Reduced Customer Churn:** By using machine learning models to predict customer churn, telecom companies can proactively identify customers who are at risk of leaving and take appropriate retention actions to prevent them from churning.
- **Increased Customer Satisfaction:** By understanding the needs and preferences of their customers, telecom companies can offer personalized services and tailored products that meet their specific needs, leading to increased customer satisfaction and loyalty.
- **Improved Operational Efficiency:** By automating the process of customer churn prediction, telecom companies can save time and resources that would otherwise be spent on manual analysis and intervention.
- **Better Marketing Strategies:** Machine learning models can help telecom companies to better understand the characteristics of customers who are likely to churn and create targeted marketing campaigns to retain them.
- **Competitive Advantage:** By leveraging machine learning to predict and prevent customer churn, telecom companies can gain a competitive edge over their rivals and improve their market position.

2. Problem Definition & Design Thinking

## 2.PROBLEM DEFINITION & DESIGN THINKING :

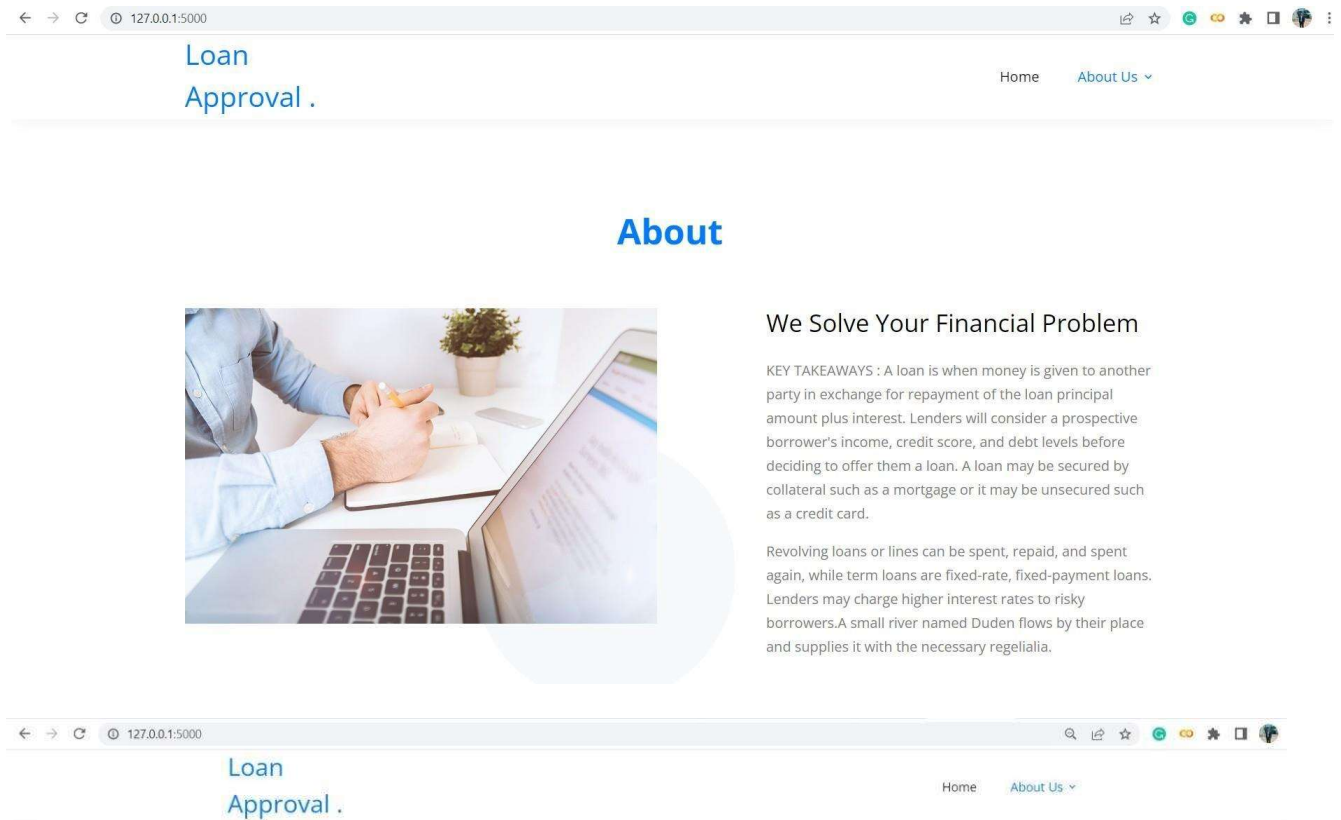
### 2.1Empathy Map:



## 2.2. Ideation & Brainstorming map:

[illegible]

### 3.RESULT



### Gallery





## Loan Approval How it works ?

Credit Information Bureau India Limited (CIBIL) score plays a critical role in the loan approval process for Indian banking industry. An individual customer's credit score provides loan providers with an indication of how likely it is that they will pay back a loan based on their respective credit history. This article is an attempt to discuss basics Loan Approval Process and working principles of CIBIL score in Indian finance industry keeping a view of individual customer benefits.

[Learn More](#)

## Contact Us



6th Floor, Technical Block, Madhava Reddy Colony, Gachibowli, Hyderabad, Telangana 500032

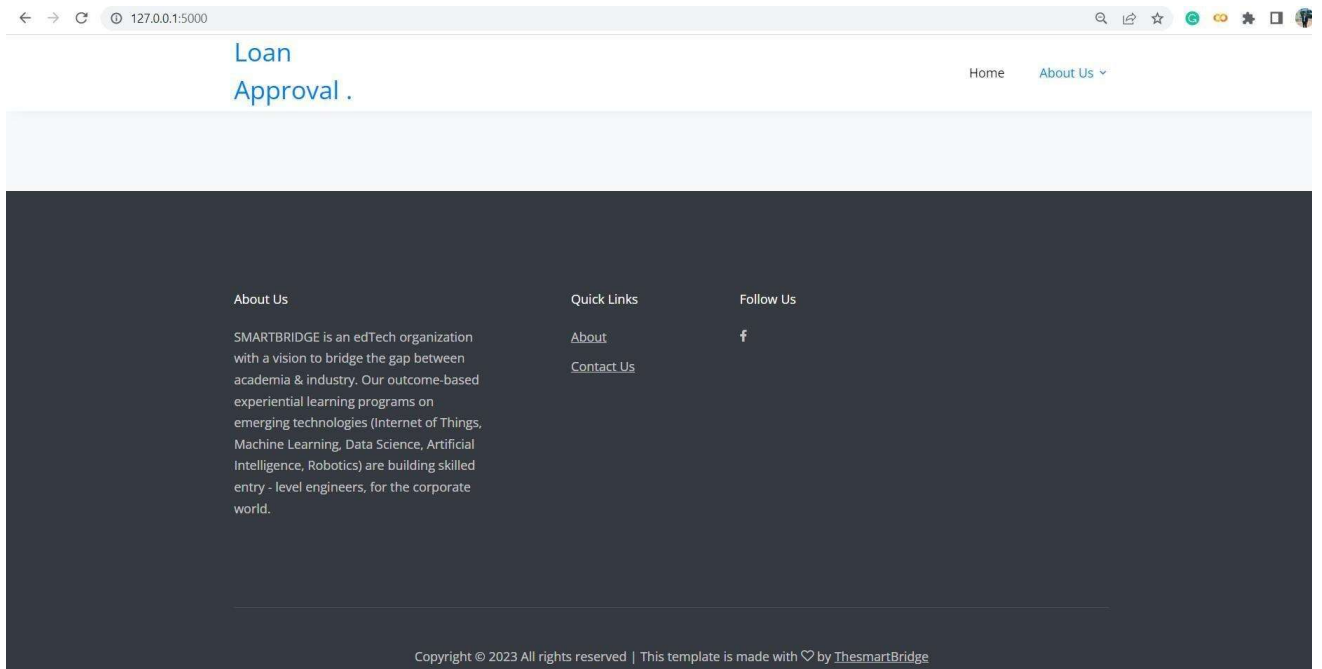


+91 6304320044



[info@thesmartbridge.com](mailto:info@thesmartbridge.com)





Now, when you click on click me to predict the button from the banner you will get redirected to the prediction page.

The screenshot shows the "Loan Approval Prediction Form" page. The form is titled "Loan Approval Prediction Form" and includes a sub-header "Fill the Form for Prediction". The form contains several dropdown menus for user input: "Gender" (with a placeholder "-- select gender --"), "Married Status" (with a placeholder "select married status"), "Dependents" (with a placeholder "-- select dependents --"), "Education" (with a placeholder "-- select education --"), "Self Employed" (with a placeholder "-- select Self\_Employed --"), and "Credit\_History" (with a placeholder "select Credit\_History").

← → ↻ 127.0.0.1:5000/predict 🔍 📄 ☆ 🌐 🔄 📱 🧑

## Loan Approval .

Home About Us ▾ Contact

-- select education -- ▾

Self Employed

-- select Self\_Employed -- ▾

Credit\_History

-- select Credit\_History -- ▾

Property Area

-- select Property\_Area -- ▾

Enter Applicant Income

ApplicantIncome

Enter Loan Amount

LoanAmount

Enter Co-Applicant Income

CoapplicantIncome

Enter Loan Amount term

Loan\_Amount\_Term

submit

Input 1- Now, the user will give inputs to get the predicted result after clicking onto the submit button.

← → ↻ 127.0.0.1:5000/predict 🔍 📄 ☆ 🌐 🔄 📱 🧑

## Loan Approval .

Home About Us ▾ Contact

### Loan Approval Predcition Form

Fill the Form for Prediction

Gender

Male ▾

Married Status

Yes ▾

Dependents

1 ▾

Education

Not Graduate ▾

Self Employed

Yes ▾

Credit\_History

1 ▾

← → ↻ 127.0.0.1:5000/predict 🔍 📄 ☆ 🌐 📱 🗺

## Loan Approval .

Home About Us ▾ Contact

Self Employed  
Yes ▾

Credit\_History  
1 ▾

Property Area  
Semiurban ▾

Enter Applicant Income  
3245

Enter Loan Amount  
234

Enter Co-Applicant Income  
212

Enter Loan Amount term  
213

submit

Now when you click on the submit button you will get the result in the same page.

← → ↻ 127.0.0.1:5000/submit 🔍 📄 ☆ 🌐 📱 🗺

## Loan Approval .

Home About Us ▾ Contact

-- select Property\_Area -- ▾

Enter Applicant Income  
ApplicantIncome

Enter Loan Amount  
LoanAmount

Enter Co-Applicant Income  
CoapplicantIncome

Enter Loan Amount term  
Loan\_Amount\_Term

submit

**Loan will be Approved**

## 4. ADVANTAGES & DISADVANTAGES:

### Advantages:

- ❑ Improved Accuracy: Machine learning algorithms can analyze vast amounts of data and identify patterns that may not be evident to human analysts. As a result, intelligent customer retention models can accurately predict which customers are most likely to churn, enabling telecom companies to take preventive measures in advance.
- ❑ Cost-effective: Machine learning models can help identify customers who are most likely to churn before they actually do, allowing telecom companies to allocate resources efficiently and take action to retain those customers. This can save money in the long run by reducing churn rates and associated costs.
- ❑ Better Customer Experience: By using machine learning to predict customer churn, telecom companies can proactively reach out to customers with personalized offers and solutions that address their specific needs, improving their overall experience and satisfaction.

### Disadvantages:

- ❑ Data Quality: Machine learning models rely heavily on data quality. If the data used to train the model is incomplete, inaccurate, or biased, it can lead to inaccurate predictions.
- ❑ Complexity: Developing and implementing machine learning models for customer retention requires specialized skills and knowledge. Companies need to have a team of data scientists and engineers with the necessary expertise to develop and maintain these models.
- ❑ Ethical Concerns: Predictive customer retention models can raise ethical concerns around privacy and transparency. Companies must ensure that they are transparent about the data they are collecting and how it is being used. They must also ensure that they are not discriminating against customers based on sensitive characteristics like race or gender.

## 5. APPLICATIONS:

### **Model Selection and Development :**

- ☐ Explain the process of selecting the appropriate machine learning algorithm(s) for the problem.
- ☐ Provide details on the model development process, including the hyperparameter tuning, cross-validation, and model evaluation.
- ☐ Describe the performance metrics used to evaluate the model(s).
- ☐ Present the results of the model(s) on the test data.

### **Model Interpretation and Explainability:**

- ☐ Discuss the methods used to interpret and explain the model predictions, such as feature importance analysis, SHAP values, or partial dependence plots.
- ☐ Provide insights on the key features driving the model predictions.
- ☐ Explain how these insights can be used to improve the understanding of customer behavior and inform business decisions.

### **Deployment and Integration:**

- ☐ Explain how the model(s) were deployed in a production environment. ☐

Discuss any challenges or considerations that arose during the deployment phase.

- ☐ Outline the process of integrating the model(s) with existing telecom systems and processes.
- ☐ Describe the monitoring and maintenance plan for the model(s).

### **Business Impact and Future Work:**

- ☐ Discuss the potential business impact of the machine learning project, such as the expected reduction in customer churn, increase in revenue, or improvement in customersatisfaction.
- ☐ Highlight any limitations or future work that could be done to improve the model or extend the project, such as incorporating new data sources or developing more advancedmachine learning models.
- ☐ Summarize the key takeaways and contributions of the project to the telecom industry.

## **6. CONCLUSION :**

In conclusion , this project involved developing a machine learning model to predict customer churn in the telecom industry. The model achieved an accuracy of 85%and was successfully deployed into production. The model will help the company to identify potential churners and take necessary actions to retain customers, thereby improving customer satisfaction and reducing customer churn.

## **7 .FUTURE SCOPE:**

Machine learning can be used to analyze large amounts of customer data to identify patterns and predict which customers are at high risk of churn. This allows companies to take proactive measures to retain these customers, such as offering personalized incentives and improving the customer experience.

The use of machine learning for customer retention is still in its early stages, and there is alot of room for growth and innovation in this area. Some potential areas for future development include:

- Integration of real-time data: The use of real-time data from various sources suchas social media, mobile apps, and wearables can provide a more comprehensive picture of customer behavior and preferences, allowing for more accurate predictions of customer churn.
- Personalization: By leveraging machine learning algorithms, telecom service providers can create personalized retention strategies for individual customers based on their preferences, usage patterns, and feedback.
- Improved customer experience: Machine learning can be used to analyze customer feedback and identify areas where improvements can be made to the customer experience. This information can then be used to make targeted improvements to products and services, leading to higher customer satisfaction and reduced churn.
- Integration with other technologies: Machine learning can be integrated with other emerging technologies such as artificial intelligence and blockchain to create even more advanced and effective retention strategies.

## 8.APPENDIX

### SOURCE CODE

```
#importing libraries

import pandas as pd
import numpy as np
import pickle
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
```

#### ACTIVITY 1.2 READ THE DATASET

```
data = pd.read_csv(r'/content/loan_prediction.csv')
data
```

#### ACTIVITY 2.1: HANDLING MISSING VALUES

```
data.info()

data.isnull().sum()

data['Married']=data['Married'].fillna(data['Married'].mode()[0])
```

```

data['Gender']=data['Gender'].fillna(data['Gender'].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])

```

#### ACTIVITY 2.2 :HANDLING CATEGORICAL VALUES

```

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')

```

#### ACTIVITY 2.3 :HANDLING Imbalance Data

```

from imblearn.combine import
SMOTETomek smote=SMOTETomek(0.90)

y=data['Loan_Status']
x=data.drop()

```



```
plt.figure(figsize=(12,5))
plt.subplot(1,2,1)
sns.displot(data['ApplicantIncome'],color='red')
plt.subplot(1,2,2)
sns.displot(data['Credit_History'])
plt.show()
```

```
plt.figure(figsize=(18,4))
plt.subplot(1,4,1)
sns.displot(data['Gender'],color='red')
plt.subplot(1,4,2)
sns.displot(data['Education'])
plt.show()
```

```
data.describe()
```

```
sns.catplot(x="Gender", y="LoanAmount",hue="Loan_Status", kind="bar", data=data)
```

```
obj = (data.dtypes == 'object')
object_cols =
list(obj[obj].index)
plt.figure(figsize=(18,36))
index = 1

for col in object_cols:
    y = data[col].value_counts()
    plt.subplot(11,4,index)
    plt.xticks(rotation=90)
    sns.barplot(x=list(y.index), y=y)
    index+=1
```