Project Report

Predicting Credit Card Approval Using Machine Learning

1. INTRODUCTION

1.1 Overview

Nowadays, banks receive a lot of applications for issuance of credit cards. Many of them are rejected for many reasons, like high-loan balances, low-income levels, or too many inquiries on an individual’s credit report. Manually analyzing these applications is error-prone and a time-consuming process. Luckily, this task can be automated with the power of machine learning and pretty much every bank does so nowadays. In this project, we will build an automatic credit card approval predictor using machine learning techniques, just like the real banks do.

In this project, we will be using regression algorithms such as Decision tree, Random forest, KNN, and xgboost. We will train and test the data with these algorithms. From this the best model is selected and saved in pkl format. We will be doing flask integration and IBM deployment.

Banks receive a lot of credit card applications. Many of the applications do not get approved for a variety of reasons, like increased loan balances or poor-income levels. Manually analyzing these applications can be very time-consuming and full of human errors. Thankfully, we can automate this task with the help of machine learning. Below are the concepts and theories that helped understand the project solution and are an integral part of this process. A thorough understanding of them facilitated the development process.

There are several manual steps involving when granting a credit card to a customer. Assessing applicant’s creditworthiness and checking the eligibility are the key factors and decisions the bank would take about a credit worthiness will not always be accurate. Application of machine learning techniques can eliminate manual paperwork, time-consuming processes and most importantly data driven decision making before granting a credit card to a customer. In this research, different supervised machine learning algorithms were used to develop models.

1.2 Purpose

Banks receive a lot of credit card applications. Many of the applications do not get approved for a variety of reasons, like increased loan balances or poor-income levels. Manually analyzing these applications can be very time-consuming and full of human errors.The main purpose of this classification model is to decide whether credit card should be approved, i.e to classify the application into two categories (approved and not approved)

2. LITERATURE SURVEY

2.1 Existing problem

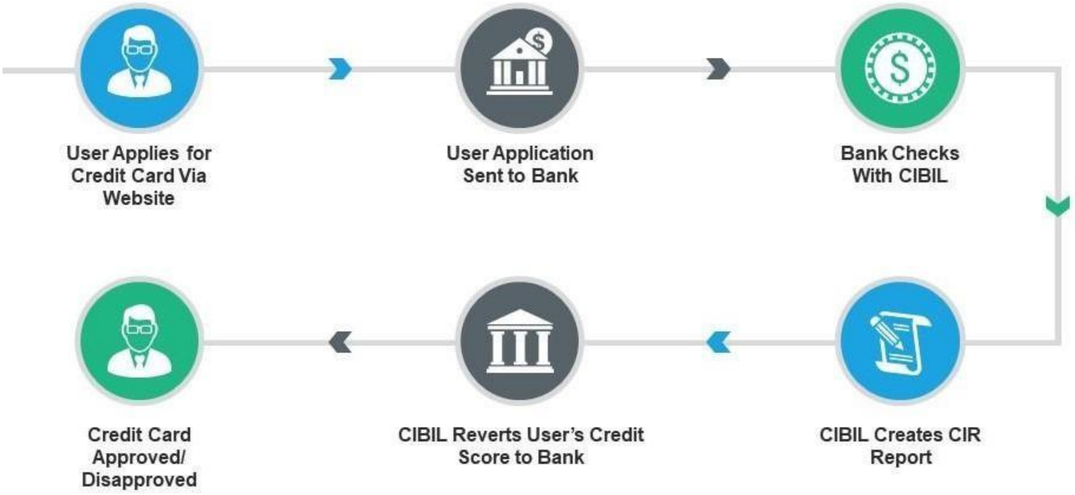
Credit risk as the boards in banks basically revolves around determining the probability of default or the creditworthiness of a customer, collapse, and the cost, assuming it happens. It is important to consider key factors and anticipate the likelihood of consumer default, given the circumstances. This is where machine learning models come into play. This allows banks and large financial institutions to predict whether their customers will default on their loans. This project uses Python to create machine-learning models with the highest possible accuracy. First, we load the dataset and take a glimpse. The data set is a combination of mathematical and non-mathematical elements, with various ranges of values and some missing points. We pre-process the dataset so that the selected ML model meets high expectations. Once the information looks good, an exploratory information check is performed to glean instincts. Finally, we created a machine learning model that can predict whether an individual's credit card application will be approved. This project uses the Jupyter Python programming notebook to create a machine-learning model. This project used data analytics and machine learning to determine the most important parameters for credit card approval. The machine learning model we built is based on the idea that a credit card will get approved or not, considering various factors listed in the credit cardholder's application.

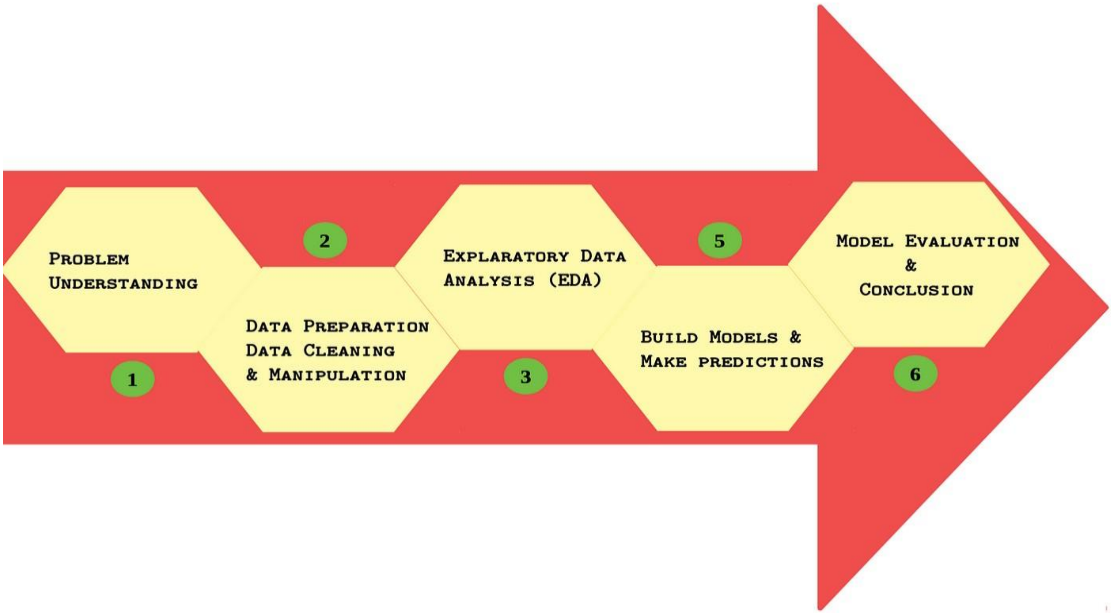
2.2 Proposed solution

we will be building a web application that is integrated to the model we built. A User Interface is provided for the uses where he/she has to enter the values for predictions. The entered values are given to the saved model and prediction is showcased on the User Interface.

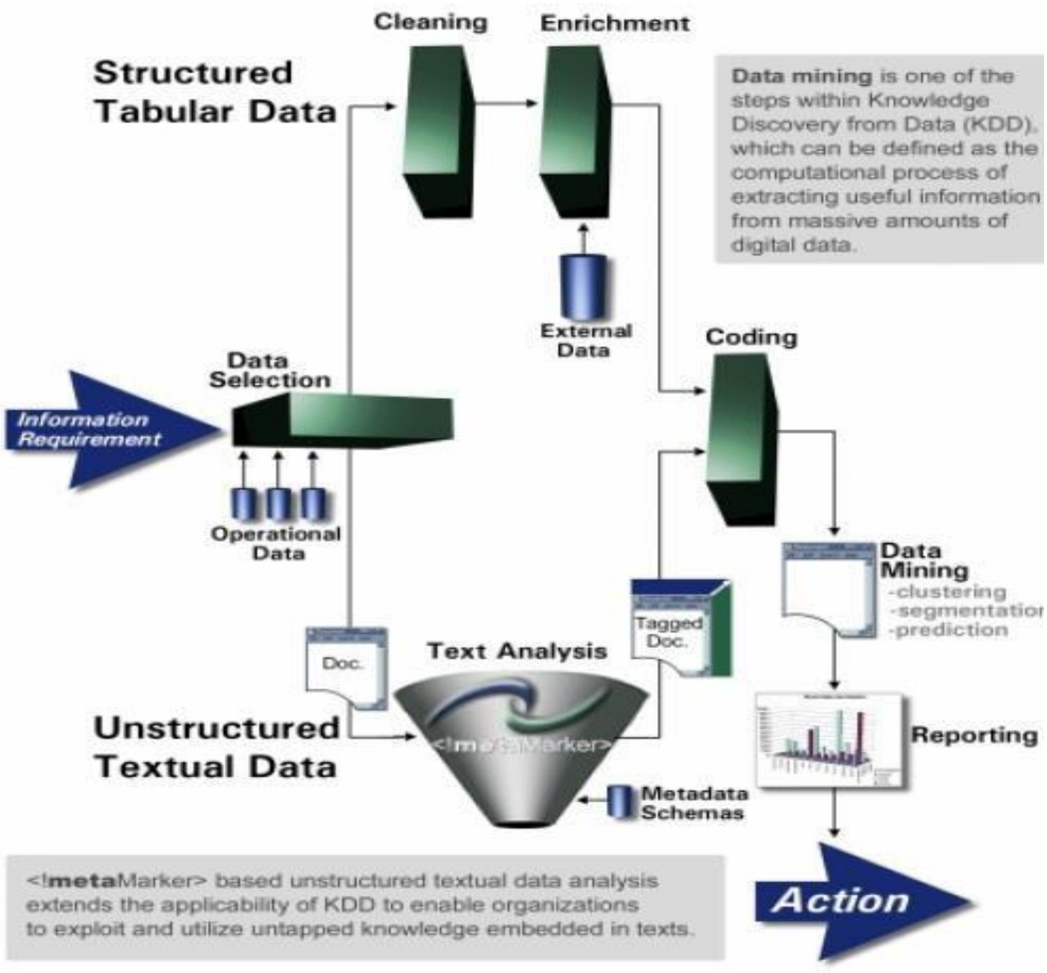
3. THEORITICAL ANALYSIS

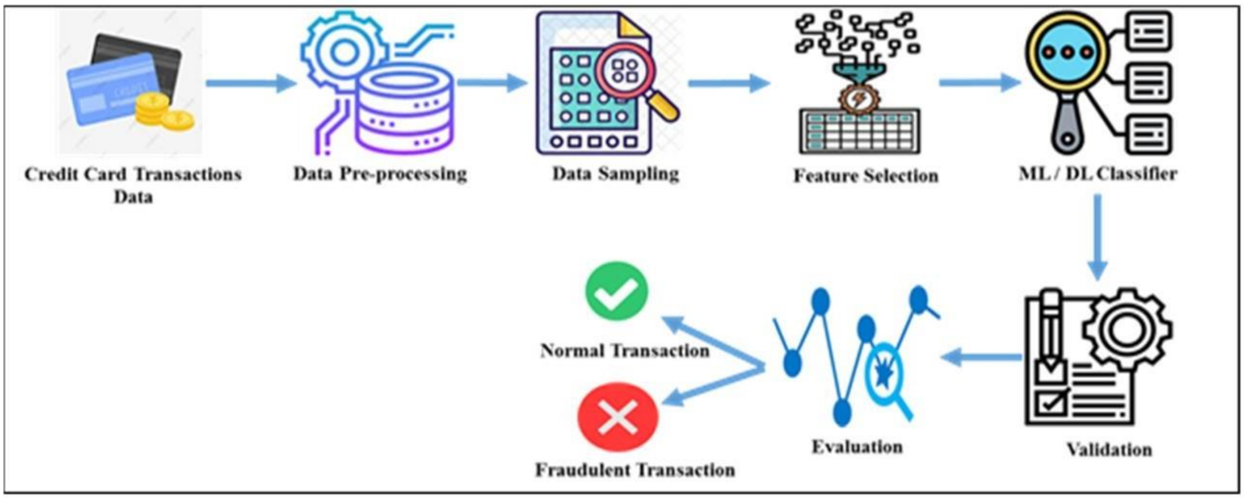
3.1 Block diagram


User Applies for Credit Card Via Website
User Application Sent to Bank
Bank Checks With CIBIL
CIBIL Creates CIR Report
CIBIL Reverts User's Credit Score to Bank
Credit Card Approved/ Disapproved


Problem Understanding
1
2
Data Preparation Data Cleaning & Manipulation
Explaratory Data Analysis (EDA)
3
5
Build Models & Make predictions
Model Evaluation & Conclusion
6

3.2 Hardware / Software designing


Cleaning
Enrichment
Structured Tabular Data
Data mining is one of the steps within Knowledge Discovery from Data (KDD), which can be defined as the computational process of extracting useful information from massive amounts of digital data.
External Data
Coding
Data Selection
Information Requirement
Operational Data
Data Mining
-clustering
-segmentation
-prediction
Tagged Doc.
Doc.
Text Analysis
Unstructured Textual Data
<!metaMarker>
Reporting
Metadata Schemas
<!metaMarker> based unstructured textual data analysis extends the applicability of KDD to enable organizations to exploit and utilize untapped knowledge embedded in texts.
Action



4. EXPERIMENTAL INVESTIGATIONS

- User interacts with the UI to enter the input.
- Entered input is analyzed by the model which is integrated.
- Once model analyses the input the prediction is showcased on the UI

To accomplish this, we have to complete all the activities listed below,

- Data Collection.
    - Collect the dataset or Create the dataset
- Data Visualization
    - Univariate analysis
    - Multivariate analysis
    - Descriptive analysis
- Data Pre-processing
    - Checking for null values
    - Drop unwanted features
    - Data Cleaning and merging
    - Handling categorical data
    - Splitting Data into Train and Test.
- Model Building
    - Import the model building Libraries
    - Initializing the model
    - Training and testing the model
    - Evaluation of Model
    - Save the Model
- Application Building
    - Create an HTML file
    - Build a Python Code

The task of predicting whether a credit card application will be approved or rejected based on values of feature variables is a supervised machine learning classification task. We need to separate the dataset into features and target variables. Following the popular convention, we call the dataframe with feature variables as X and the one with target variable as y. To implement machine learning algorithms we use the popular python library scikit-learn.

Data Pre-Processing

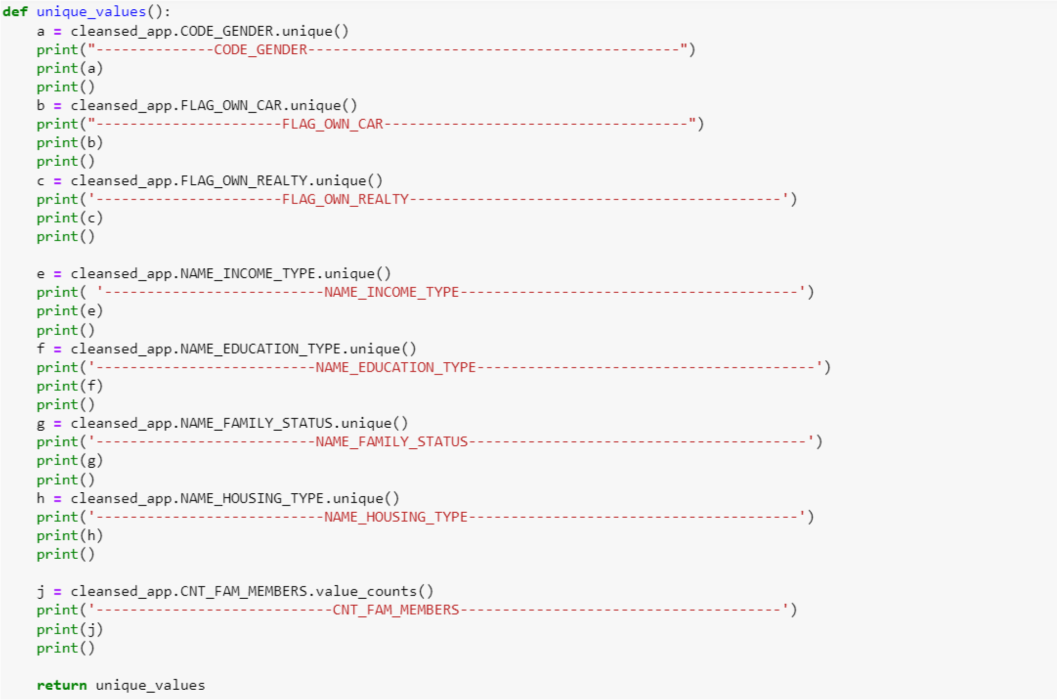
As we have understood how the data is. Let's pre-process the collected data.

The download data set is not suitable for training the machine learning model as it might have so much randomness so we need to clean the dataset properly in order to fetch good results. This activity includes the following steps.

- Handling missing values
- Handling categorical data
- Handling outliers
- Scaling Techniques
- Splitting dataset into training and test set.
- To find the data type of columns info() function is used. It gives small information about the features.

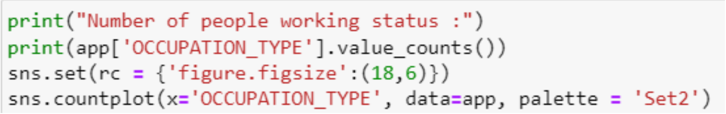
Code

- Finding unique values



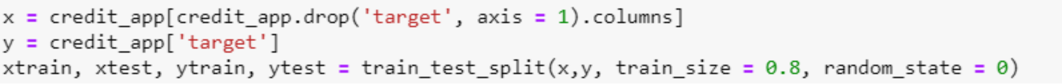
```
def unique_values():
    a = cleansed_app.CODE_GENDER.unique()
    print("--------------CODE_GENDER-------------------------------------------")
    print(a)
    print()
    b = cleansed_app.FLAG_OWN_CAR.unique()
    print("----------------------FLAG_OWN_CAR-----------------------------------")
    print(b)
    print()
    c = cleansed_app.FLAG_OWN_REALTY.unique()
    print('----------------------FLAG_OWN_REALTY-------------------------------------------')
    print(c)
    print()

    e = cleansed_app.NAME_INCOME_TYPE.unique()
    print( '-------------------------NAME_INCOME_TYPE----------------------------------------')
    print(e)
    print()
    f = cleansed_app.NAME_EDUCATION_TYPE.unique()
    print('--------------------------NAME_EDUCATION_TYPE---------------------------------------')
    print(f)
    print()
    g = cleansed_app.NAME_FAMILY_STATUS.unique()
    print('--------------------------NAME_FAMILY_STATUS---------------------------------------')
    print(g)
    print()
    h = cleansed_app.NAME_HOUSING_TYPE.unique()
    print('---------------------------NAME_HOUSING_TYPE---------------------------------------')
    print(h)
    print()

    j = cleansed_app.CNT_FAM_MEMBERS.value_counts()
    print('----------------------------CNT_FAM_MEMBERS--------------------------------------')
    print(j)
    print()

    return unique_values
```

- Data Visualization



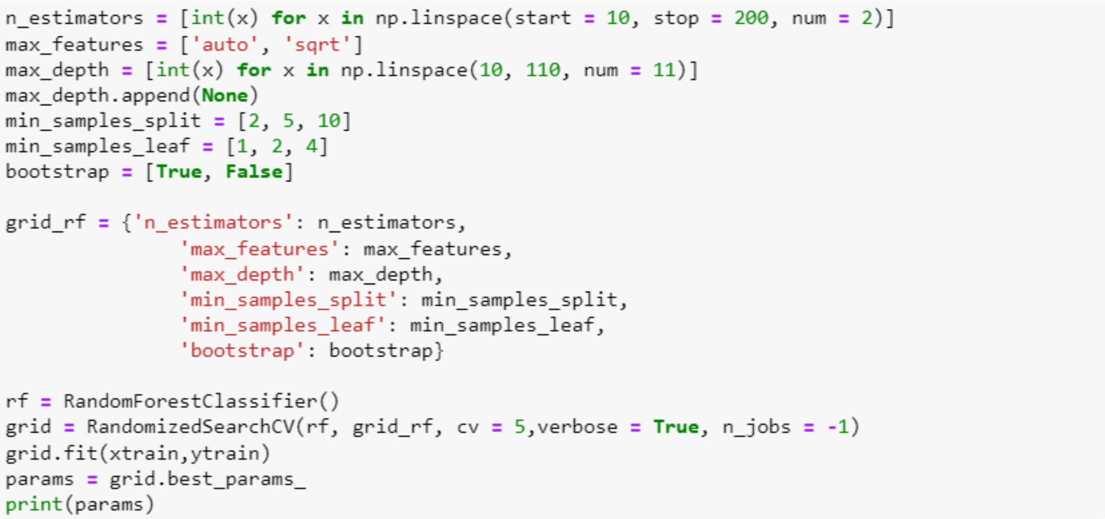
```
print("Number of people working status :")
print(app['OCCUPATION_TYPE'].value_counts())
sns.set(rc = {'figure.figsize':(18,6)})
sns.countplot(x='OCCUPATION_TYPE', data=app, palette = 'Set2')
```

- Splitting Test and Train Data

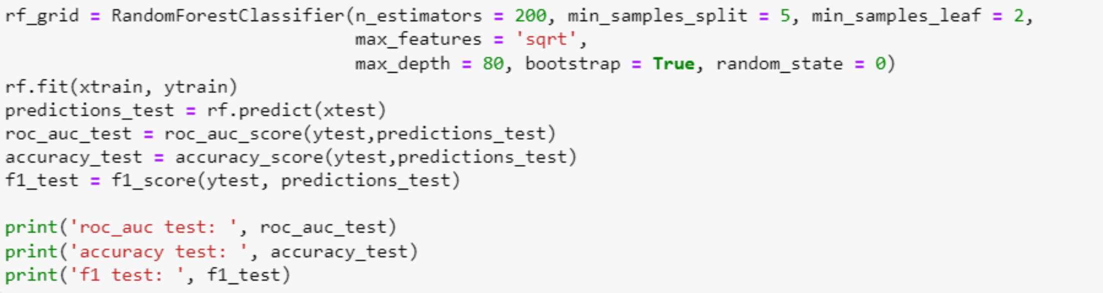


```
x = credit_app[credit_app.drop('target', axis = 1).columns]
y = credit_app['target']
xtrain, xtest, ytrain, ytest = train_test_split(x,y, train_size = 0.8, random_state = 0)
```

- Model Building

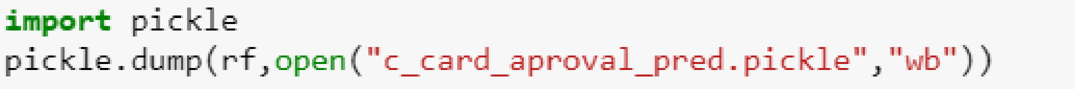


```
n_estimators = [int(x) for x in np.linspace(start = 10, stop = 200, num = 2)]
max_features = ['auto', 'sqrt']
max_depth = [int(x) for x in np.linspace(10, 110, num = 11)]
max_depth.append(None)
min_samples_split = [2, 5, 10]
min_samples_leaf = [1, 2, 4]
bootstrap = [True, False]

grid_rf = {'n_estimators': n_estimators,
           'max_features': max_features,
           'max_depth': max_depth,
           'min_samples_split': min_samples_split,
           'min_samples_leaf': min_samples_leaf,
           'bootstrap': bootstrap}

rf = RandomForestClassifier()
grid = RandomizedSearchCV(rf, grid_rf, cv = 5,verbose = True, n_jobs = -1)
grid.fit(xtrain,ytrain)
params = grid.best_params_
print(params)
```



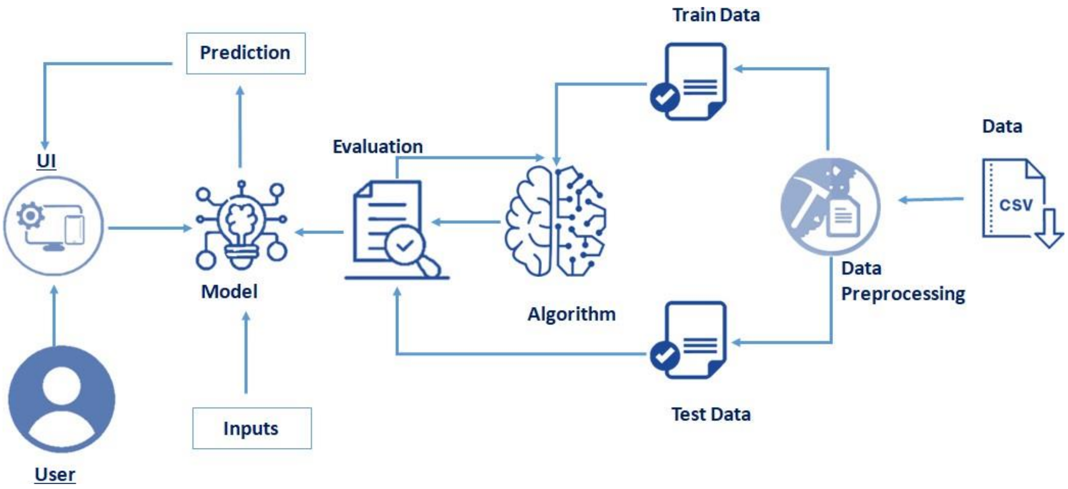
```
rf_grid = RandomForestClassifier(n_estimators = 200, min_samples_split = 5, min_samples_leaf = 2,
                                 max_features = 'sqrt',
                                 max_depth = 80, bootstrap = True, random_state = 0)
rf.fit(xtrain, ytrain)
predictions_test = rf.predict(xtest)
roc_auc_test = roc_auc_score(ytest,predictions_test)
accuracy_test = accuracy_score(ytest,predictions_test)
f1_test = f1_score(ytest, predictions_test)

print('roc_auc test: ', roc_auc_test)
print('accuracy test: ', accuracy_test)
print('f1 test: ', f1_test)
```

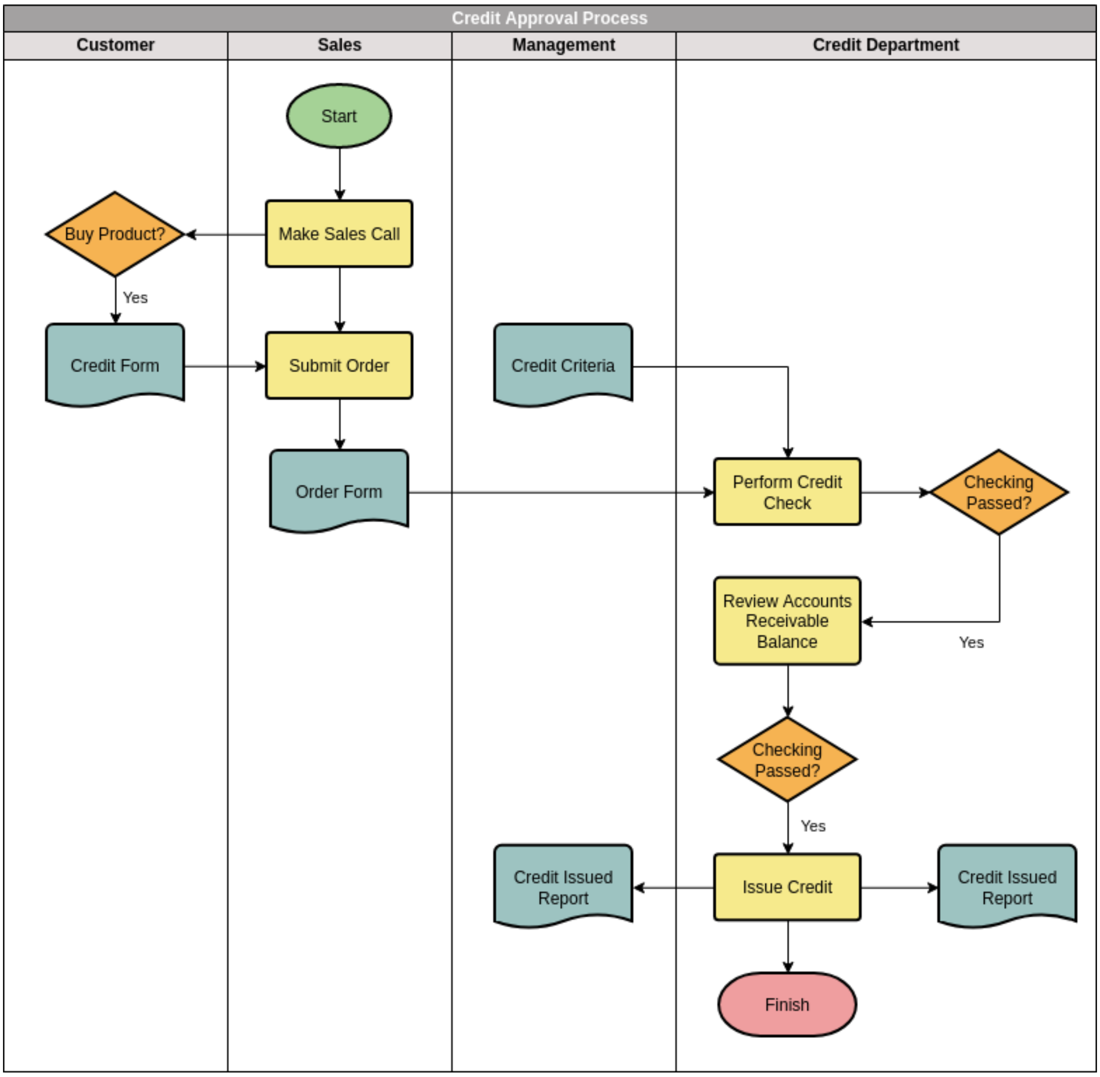
- Save the model



```
import pickle
pickle.dump(rf,open("c_card_aproval_pred.pickle","wb"))
```

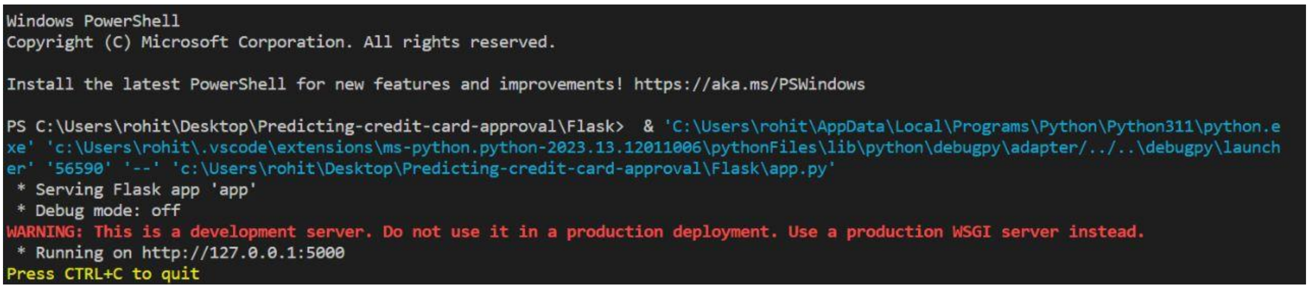
5. FLOWCHART


Prediction
Train Data
UI
Evaluation
Data
CSV
Model
Algorithm
Data Preprocessing
Test Data
Inputs
User


Credit Approval Process
Customer
Sales
Management
Credit Department
Start
Buy Product?
Make Sales Call
Yes
Credit Form
Submit Order
Credit Criteria
Order Form
Perform Credit Check
Checking Passed?
Review Accounts Receivable Balance
Yes
Checking Passed?
Yes
Credit Issued Report
Issue Credit
Credit Issued Report
Finish

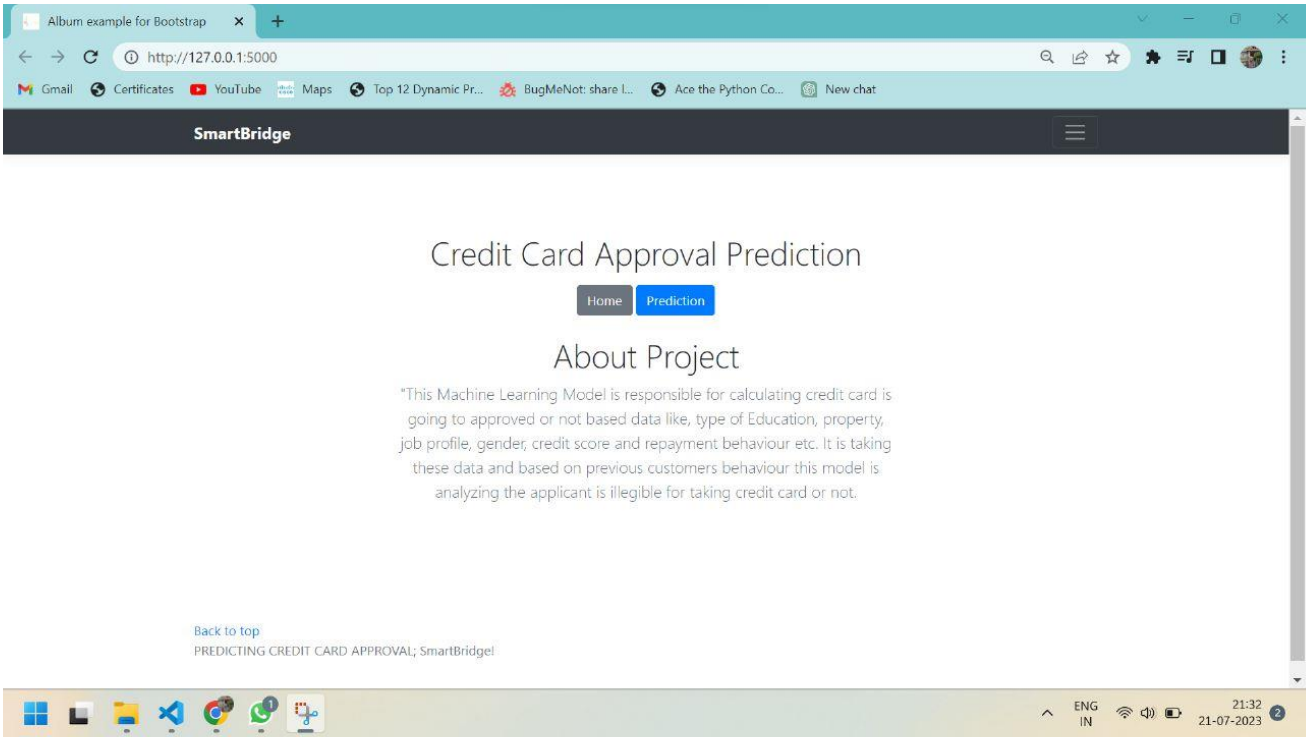
6 RESULT

- Open anaconda prompt from the start menu
- Navigate to the folder where your python script is.
- Now type “python app.py” command
- Navigate to the localhost where you can view your web page.
- Click on the predict button from the top right corner, enter the inputs, click on the submit button, and see the result/prediction on the web.



```
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Install the latest PowerShell for new features and improvements! https://aka.ms/PSWindows

PS C:\Users\rohit\Desktop\Predicting-credit-card-approval\Flask>  & 'C:\Users\rohit\AppData\Local\Programs\Python\Python311\python.e
xe' 'c:\Users\rohit\.vscode\extensions\ms-python.python-2023.13.12011006\pythonFiles\lib\python\debugpy\adapter/../..\debugpy\launch
er' '56590' '--' 'c:\Users\rohit\Desktop\Predicting-credit-card-approval\Flask\app.py'
 * Serving Flask app 'app'
 * Debug mode: off
WARNING: This is a development server. Do not use it in a production deployment. Use a production WSGI server instead.
 * Running on http://127.0.0.1:5000
Press CTRL+C to quit
```

Now paste the URL on the browser, you will redirect to index.html page. Let’s look our index page



Credit Card Approval Prediction

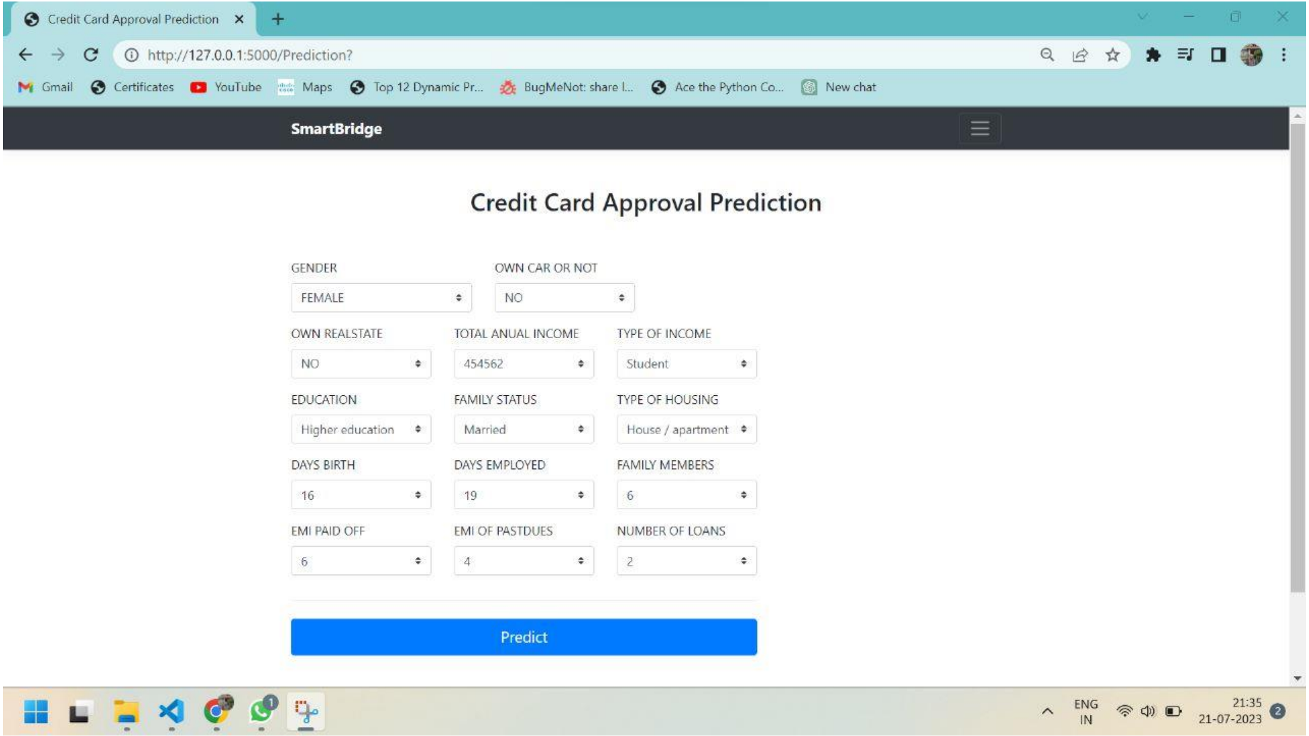
About Project

"This Machine Learning Model is responsible for calculating credit card is going to approved or not based data like, type of Education, property, job profile, gender, credit score and repayment behaviour etc. It is taking these data and based on previous customers behaviour this model is analyzing the applicant is illegible for taking credit card or not.

Back to top

PREDICTING CREDIT CARD APPROVAL; SmartBridge!

Click on predict button to give your inputs. It routes to index1.html page.

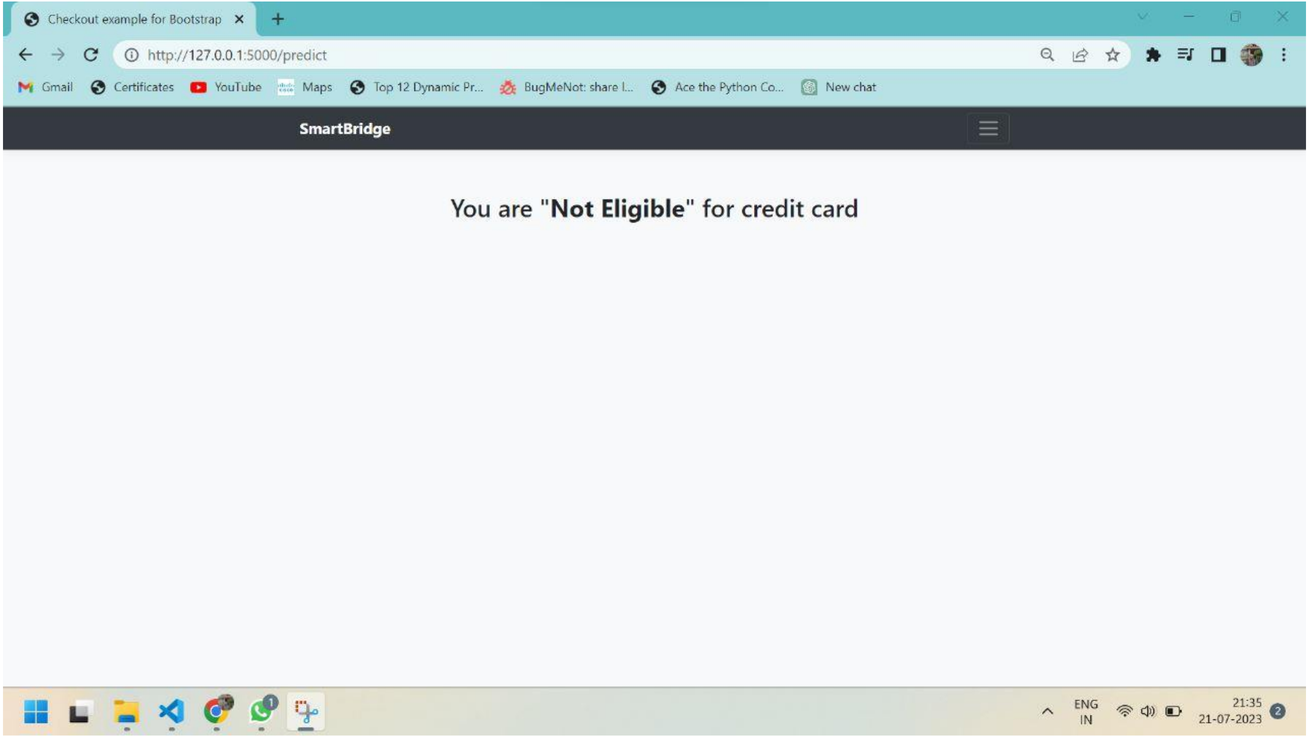


Credit Card Approval Prediction

GENDER: FEMALE	OWN CAR OR NOT: NO	
OWN REALSTATE: NO	TOTAL ANUAL INCOME: 454562	TYPE OF INCOME: Student
EDUCATION: Higher education	FAMILY STATUS: Married	TYPE OF HOUSING: House / apartment
DAYS BIRTH: 16	DAYS EMPLOYED: 19	FAMILY MEMBERS: 6
EMI PAID OFF: 6	EMI OF PASTDUES: 4	NUMBER OF LOANS: 2

To predict your credit card eligibility, click on predict button. The output will be displayed on result.html page.

Output:



You are "**Not Eligible**" for credit card

7 ADVANTAGES & DISADVANTAGES

- The correct assessment for credit card approval is very important for banks and organisations who lend a credit card to the people. The recent years have seen a huge growth in credit cards and loans. The exact judgement of person to be approved for credit cards allows the organisations to minimize losses and the same time make suitable credit arrangements as per requirement. Due to the huge growth in the number of applicants, there is a need for a more sophisticated method to automate the process and speed it up.

- Credit card approval can be beneficial for organisations that lend credit cards, and due to increase in a huge number of the applicant, there is need to automate the task and classify the applicants into if they are eligible for a credit card or not. This helps to avoid organisation losses by avoiding potential defaulters. Here we are not just looking into bank balance but into there personal attributes like gender, married, age, Occupation etc. We account for these personal attributes to evaluate if the given applicant is a good customer. This can also help cut down the weeks-long process into a few days. This gives benefit by cutting down costs on credit analysis and faster credit decisions.

8. APPLICATIONS

◆ **Buy on credit:**

What makes a credit card attractive is the credit limit allowed to the cardholder. You can buy anything within that limit and pay later. Your monthly budget will not affect, even if you buy items of high value on credit. One among the most important benefits of **credit card** is you can convert the total amount of your purchases into low-cost EMIs to enable you to repay it easily over a period of time. This has helped revolutionize the shopping experience.

◆ **Most accepted method of payment:**

You can travel anywhere, without carrying much money if you have this card. Being the most accepted method of payment, you can use a credit card to pay anything.

- **Interest-free cash withdrawals:**

There are a few credit cards that allow you to withdraw money up to a certain limit in case of emergency, with no interest charged up to 45 to 50 days. You can make use of it in times of financial emergency.

- **Unlimited reward points:**

These cards come with reward points when you use them. For instance, IDFC FIRST Bank credit cards offer unlimited and never-expiring reward points, which are easily redeemable.

- **5. Insurance coverage:**

You get personal accident coverage, as well as comprehensive travel insurance coverage and this is one of the significant benefits of credit cards, which make them attractive.

- **Make travel easy:**

The uses of credit cards in travel make them important. When it comes to IDFC FIRST Bank credit cards, they give you a unique experience through complimentary lounge access at the airports and railway stations in India and priority check-in. Other than these, you can also enjoy discounts on food in more than 280 restaurants.

- **Discounts and cashbacks:**

The advantages of credit cards extend to discounts on some of your favourite entertainment and dining outlets, travel and shopping apps, etc. You can also enjoy fuel surcharge waivers at petrol pumps across the country.

- **Improve your credit score:**

The benefits of credit cards do not limit to shopping on credit; instead, it helps improve your credit score. If you know how to use a credit card and how to make use of the credit period, and repay the amount used on time, you can boost your CIBIL score. This will help you obtain loans, without any difficulty in future.

- **Offers safety:**

You don’t have to carry much money if you have a credit card.

- **Keep track of your expenses:**

The statements you get every month from net banking helps you check your expenses and plan the repayment without any delay.

9. CONCLUSION

we have mentioned various machine learning methods to predict whether a credit card will be approved for an individual or not. Several parameters were taken into consideration as these parameters make the model more effective and help institutions make better decisions to avoid fraud and losses. We applied a lot of data pre-processing techniques as good amount of data pre-processing contributes effectively to developing better performance of traditional machine learning models. During Exploratory Data Analysis, we plotted a lot of graphs and charts to study the dataset deeply so that we can get a better understanding of the dataset. This was done so that we can decide which models to apply which can perform well on this dataset and can correctly predict whether to approve a credit card or not. This prediction system can be helpful to various banks as it makes their task easier and increases efficiency as compared to the manual system which is currently used by many banks and this system is cost effective.

10. FUTURE SCOPE

To further improve our system, we can use deep learning models as it can increase our accuracy. Neural networks can be used as it can discover hidden patterns and correlations in raw data, cluster and classify it, and continuously learn and improve over time. In the future, this credit card approval system will be able to be optimized and implemented in an artificial intelligence environment. By displaying the prediction result on a web or desktop application, the system can also be automated. Thus, this work has a good future scope and can be enhanced by adding other various feature for better predictions.

11. BIBILOGRAPHY

https://www.bajajfinserv.in/insights/6-reasons-why-your-credit-card-application-can-be-rejected

https://www.ijraset.com/research-paper/credit-card-approval-prediction-using-classification-algorithms

APPENDIX A.

https://drive.google.com/drive/folders/196huzcfmNyHc7q4GESjQVrpHnx2GvINI?usp=drive_link