MICRO CREDIT LOAN

Submitted by:

NEHA GUPTA

**ACKNOWLEDGMENT**

This includes mentioning of all the references, research papers, data sources, professionals and other resources that helped you and guided you in completion of the project.

**INTRODUCTION**

* Business Problem Framing

This can help organization to identify the defaulter and can help organization

* Conceptual Background of the Domain Problem

Each and every field was important except few which are removed.

* Review of Literature

Unnecessary columns have been removed to clear the data. EDA has applied to clear the data. Skewness has been removed.

* Motivation for the Problem Undertaken

This will help organization to minimize the risk of losses.

**Analytical Problem Framing**

* Mathematical/ Analytical Modeling of the Problem

Decision tree regressor has applied with greateR

accuracy.

* Data Sources and their formats

Column names - Index(['Unnamed: 0', 'label', 'msisdn', 'aon', 'daily\_decr30', 'daily\_decr90', 'rental30', 'rental90', 'last\_rech\_date\_ma', 'last\_rech\_date\_da', 'last\_rech\_amt\_ma', 'cnt\_ma\_rech30', 'fr\_ma\_rech30', 'sumamnt\_ma\_rech30', 'medianamnt\_ma\_rech30', 'medianmarechprebal30', 'cnt\_ma\_rech90', 'fr\_ma\_rech90', 'sumamnt\_ma\_rech90', 'medianamnt\_ma\_rech90', 'medianmarechprebal90', 'cnt\_da\_rech30', 'fr\_da\_rech30', 'cnt\_da\_rech90', 'fr\_da\_rech90', 'cnt\_loans30', 'amnt\_loans30', 'maxamnt\_loans30', 'medianamnt\_loans30', 'cnt\_loans90', 'amnt\_loans90', 'maxamnt\_loans90', 'medianamnt\_loans90', 'payback30', 'payback90', 'pcircle', 'pdate'], dtype='object')

Removed “pcircle” and “pdate” whereas applied label encoder to tranform the pdate column in to int data type

* Data Preprocessing Done

EDA and Applied skewness methods

* Hardware and Software Requirements and Tools Use

Python as a software

Libraries:

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn.linear\_model import LogisticRegression

from sklearn.metrics import accuracy\_score

from sklearn.metrics import confusion\_matrix,classification\_report

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

**Model/s Development and Evaluation**

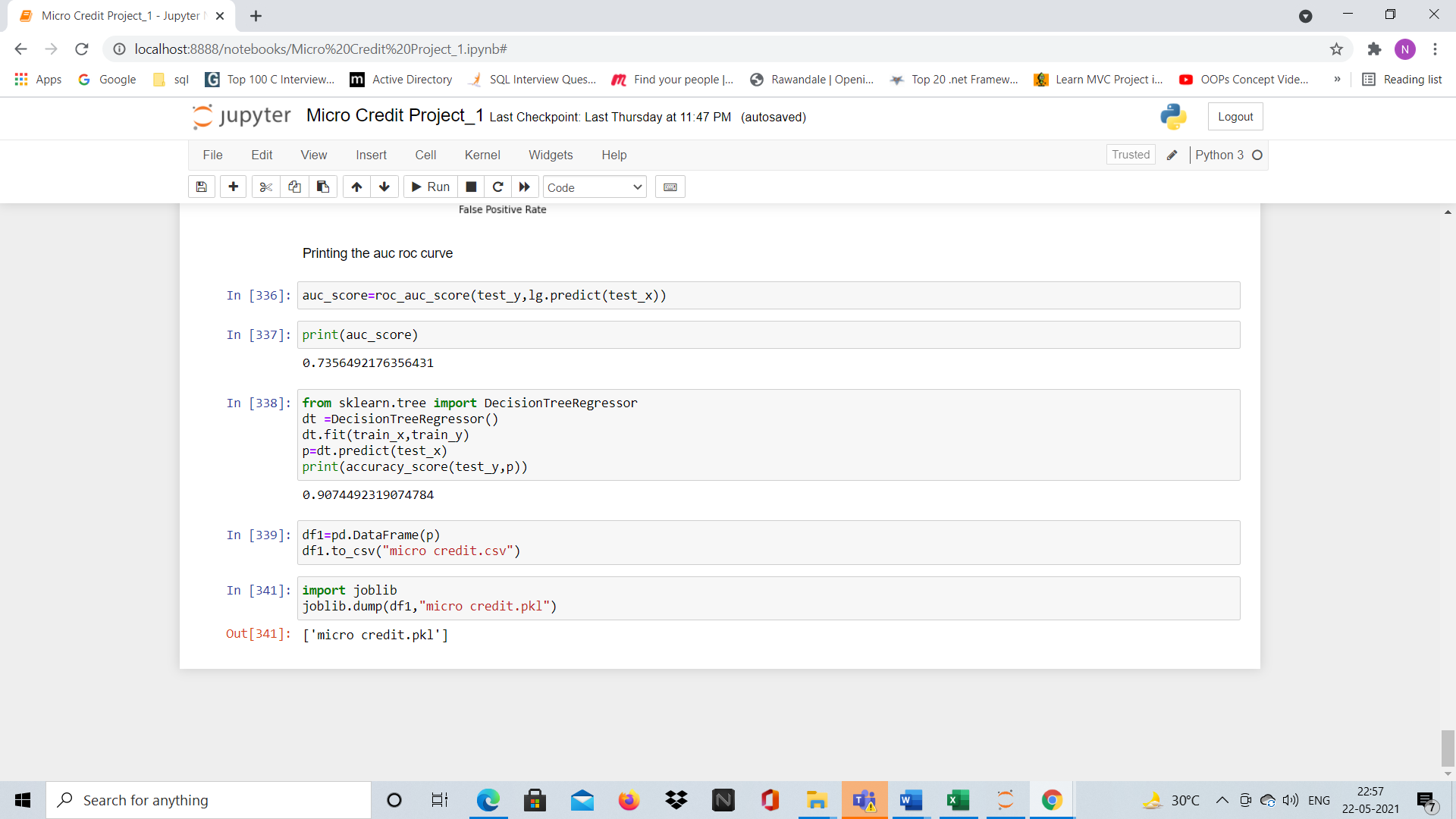
* Identification of possible problem-solving approaches (methods)

All the types of regression methods

* Testing of Identified Approaches (Algorithms)

Decision tree regressor

* Run and Evaluate selected models



* Key Metrics for success in solving problem under consideration

Removed the skewness.

**CONCLUSION**

* Key Findings and Conclusions of the Study

Unnecessary columns to be removed.

* Learning Outcomes of the Study in respect of Data Science

Decision Tree Regressor