Credit Card Approval Prediction

Batch - 8
Guide: Dr.J.Hyma

Rakesh Naidu	121810305023
Mani Chandra	121810305032
Shashidhar	121810305043
Vishnu	121810305008

Abstract

Credit Card Approval Prediction is a Machine Learning Solution to Automate the process of Credit Card approval to an account. The model is trained with multiple attributes and the accuracy is measured with both train and test data.

Introduction

The accurate assessment of consumer credit risk is of uttermost importance for lending organizations. Credit scoring is a widely used technique that helps financial institutions evaluates the likelihood for a credit applicant to default on the financial obligation and decide whether to grant credit or not.

Introduction

The goal here is to build an end to end automated Machine Learning solution where a user will be able to predict whether a bank customer should be approved for attaining the credit card or not. The user is only need to give the value of feature variables and the model will able to predict the binary outcome (Approve/ Not Approve).

Introduction

The model will be able take care of all intermediate functionalities like cross validation, hyper parameter tuning, algorithm selection etc.

This project shall be delivered in two phases:

Phase 1: All the functionalities with PyPi packages.
Phase 2: Integration of UI to all the functionalities.

Note: All the code will be written in python version 3.6

About Dataset

This file concerns credit card applications. All attribute names and values have been changed to meaningless symbols to protect confidentiality of the data.

This dataset is interesting because there is a good mix of attributes -- continuous, nominal with small numbers of values, and nominal with larger numbers of values. There are also a few missing values.

Data Source: UCI, Kaggle

Attribute Information

A1	b, a	Gender
A2	continuous	Age
A3	continuous	Debt
A4	u, y, l, t	Marital status
A5	g, p, gg	Bank
A6	c, d, cc, i, j, k, m, r, q, w, x, e, aa, ff	Education Level
A7	v, h, bb, j, n, z, dd, ff, o	Ethnicity
A8	continuous	Years Employed

Attribute Information

A9	t, f	Prior default
A10	t, f	Employed
A11	continuous	Credit score
A12	t, f	Drivers license
A13	g, p, s	Citizen
A14	continuous	Zip Code
A15	continuous	Income
A16	+,- (class attribute)	Approved

Existing Project

The Existing project is done with some data cleaning and then training and prediction which led to lower accuracy. Machine learning models used in the existing project are:

- Logistic Regression
- Decision Tree
- Random Forest

Proposed Project

The proposed project is built end to end. Starting from Data Preprocessing to Deployment. This project includes the features like:

- Statistical analysis
- Hyper parameter tuning
- Best algorithm selection
- Deployment in Heroku using flask.

Limitations of the existing project.

Some of the areas where the existing project is lacking are:

- High accuracy
- Data analysis
- Training with more models
- Different model selection criteria
- Deployment

Objectives of proposed project

The main objectives of the proposed project are to:

- Increase the accuracy
- Do Exploratory data analysis
- Test the model with different algorithms
- Try different model selection criteria
- Do Hyperparameter tuning
- Deploy the project for easy use

System requirements

For Model Training	For Model Testing
• 8 GB RAM	• 4 GB RAM
• 2 GB of Hard Disk Space	• 2 GB of Hard Disk Space
• Intel Core i5 Processor	• Intel Core i5 Processor

Note: These are just Recommended.

Algorithms to be used in the project

Different Machine Learning algorithms used in the project are:

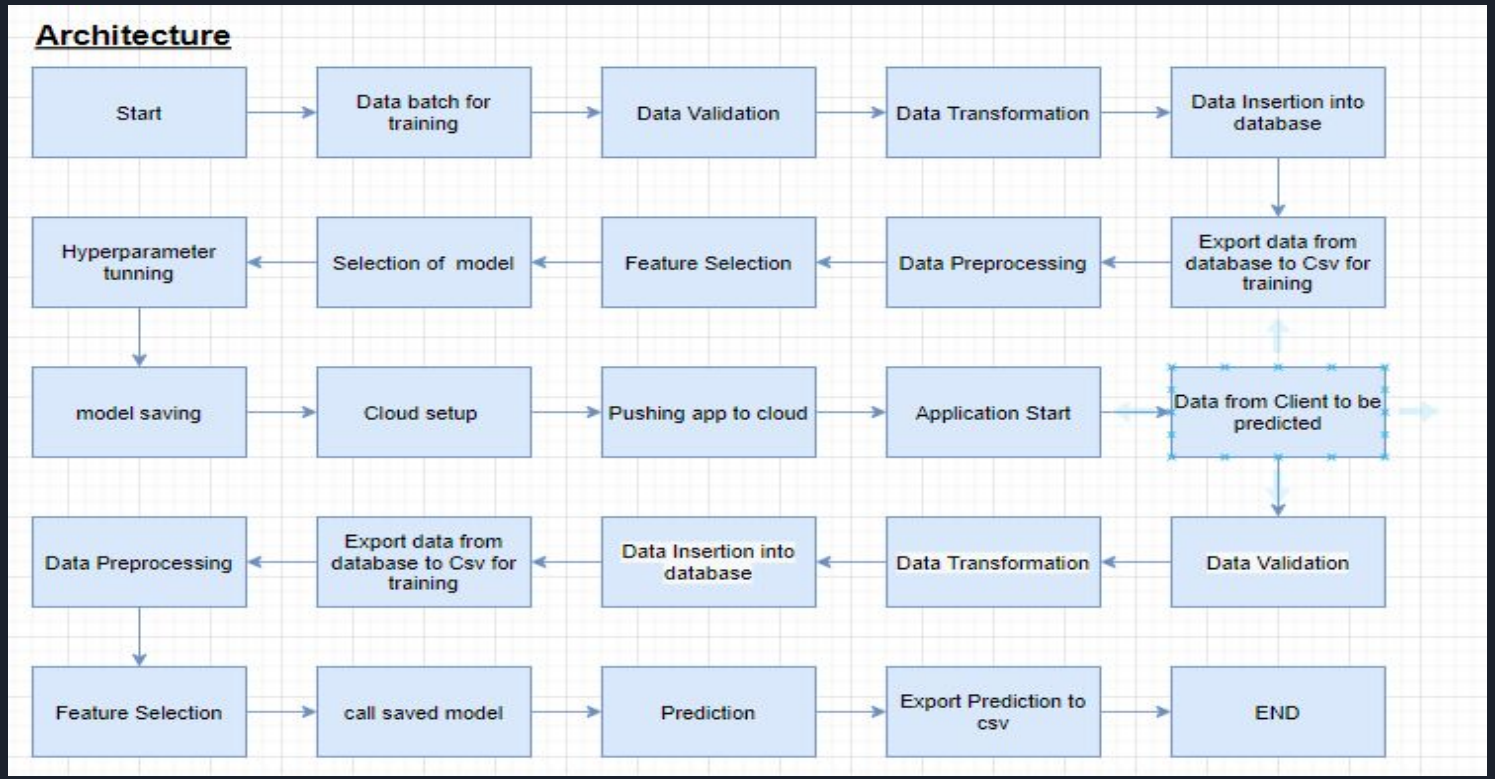
- Decision Tree
- Random Forest
- XG Boost
- KNN Classifier
- Naïve Baye’s

Technologies to be used in the project

Different Technologies used in the project are:

- numpy, pandas
- matplotlib, seaborn
- scipy, scikit learn
- xgboost
- html, css
- flask, gunicorn

Flow Chart of the proposed project



References

1. https://www.ijrar.org/papers/IJRAR190B030.pdf
2. https://www.ijeat.org/wp-content/uploads/papers/v9i4/D7293049420.pdf
3. https://www.researchgate.net/publication/321002603_Credit_Approval_Analysis_using_R
4. https://rstudio-pubs-static.s3.amazonaws.com/73039_9946de135c0a49daa7a0a9eda4a67a72.html

Thank You