Final Mini Project Demonstration

Project Title : Delinquency flagging and pd of loan
Project Guide : Prof.V R Badri Prasad
Project Team : Mahin Mohan PES1UG19CS253
Raghav V Pandit PES1UG19CS364




Project Abstract and Scope

Introduction

In the modern world, we notice a lot of credit card fraudulent activities all around the world and banks spend a tremendous amount of time and money in identifying such crimes .On the other hand banks also need to deal with providing loans by checking the client's/customer's credibility which is based on multiple criteria, the most important one amongst them is the probability of default . Our project makes the job of banks easy by providing both these services as one package/unit . Our project aims to flag the fraudulent transactions and simultaneously provide the probability of default of a particular customer.
The Scope of our project is to provide the functionalities such as identification of fraudulent transactions and computing the probability default of a customer as one combined unit.

Set the context…

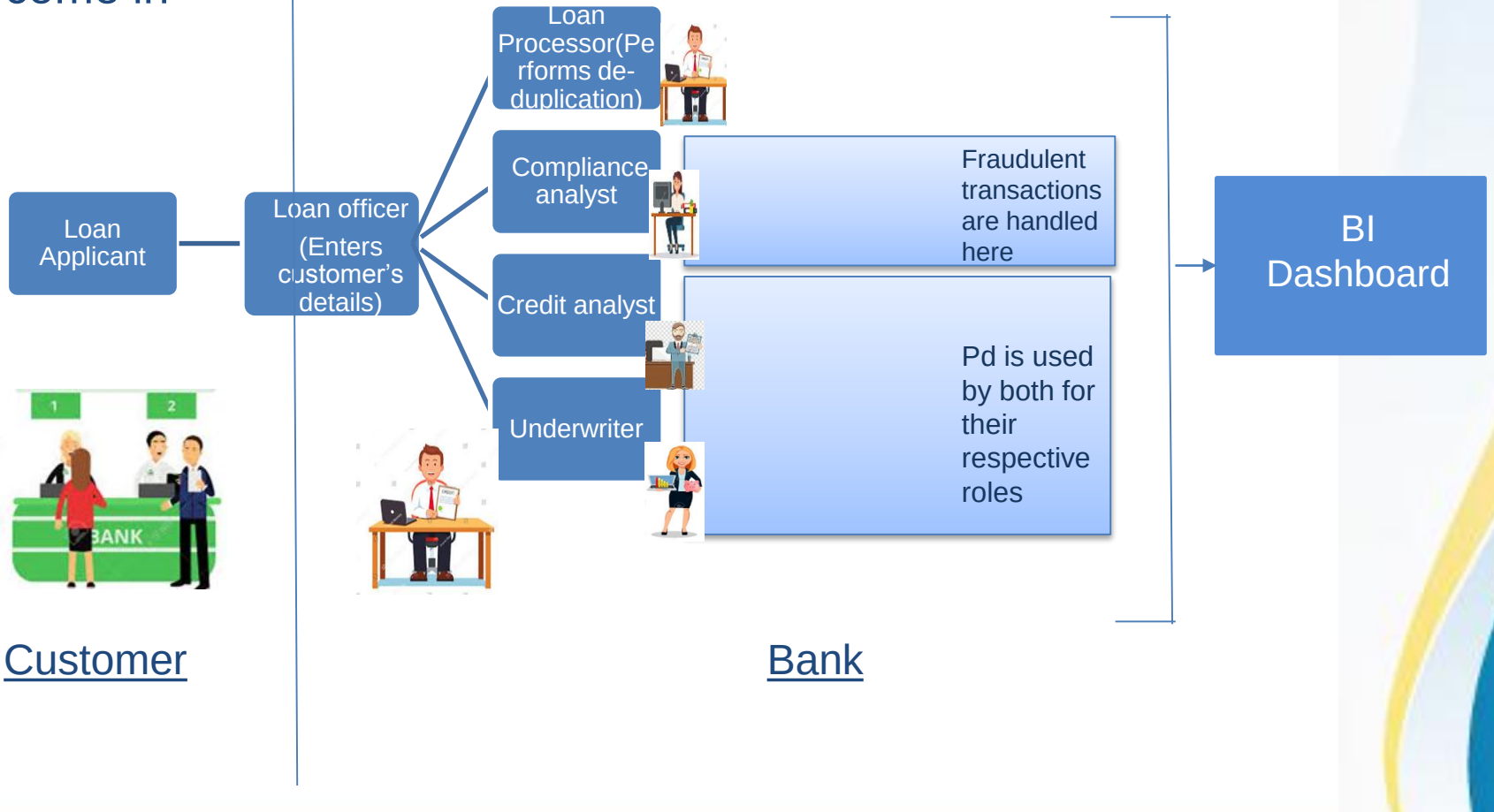
Our project aims to develop a machine learning model that is used to detect credit card fraudulent activity and side-by-side based on the linear regression of the probability of default it provides information regarding the client's/customer's credibility and whether the person will default or not

Literature Survey

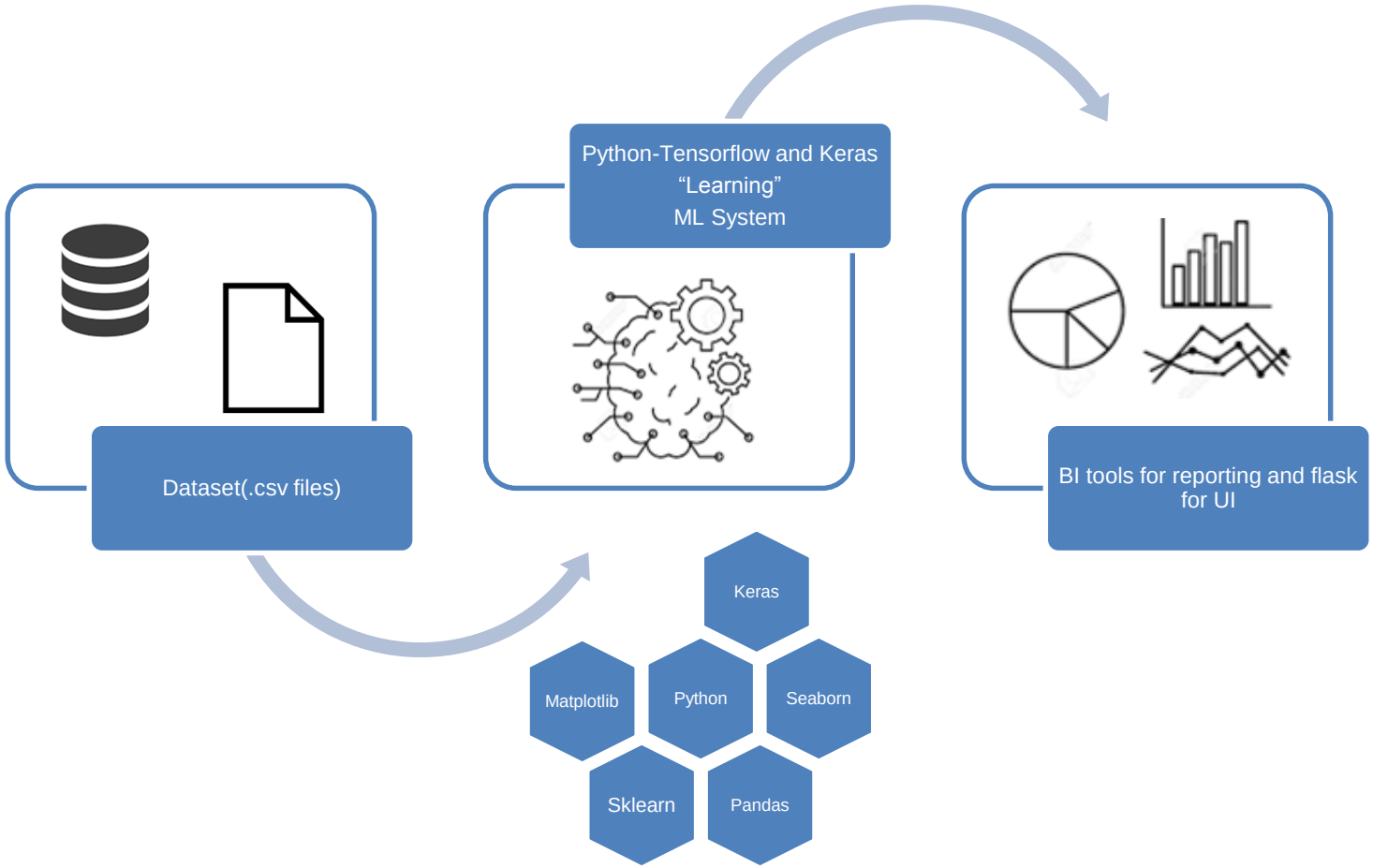
Paper Title	Inference
Cyber Risk for the Financial Sector: A Framework for Quantitative Assessment	Domain knowledge regarding the security risks in the financial sector
Predicting the Probability of Loan-Default: An Application of Binary Logistic Regression	The importance of PD was shown and different possible methods of implementation
Credit Card Fraud Detection using Deep Learning	Method of deep learning was ventured in this paper and performance comparisons of different models to detect fraud was seen

Design Description/UI Design

Bird's eye view of the roles involved in the process and where we come in



Technologies Used



Implementation Details

Code explanation.

Logistic Regression Model for Probability Default

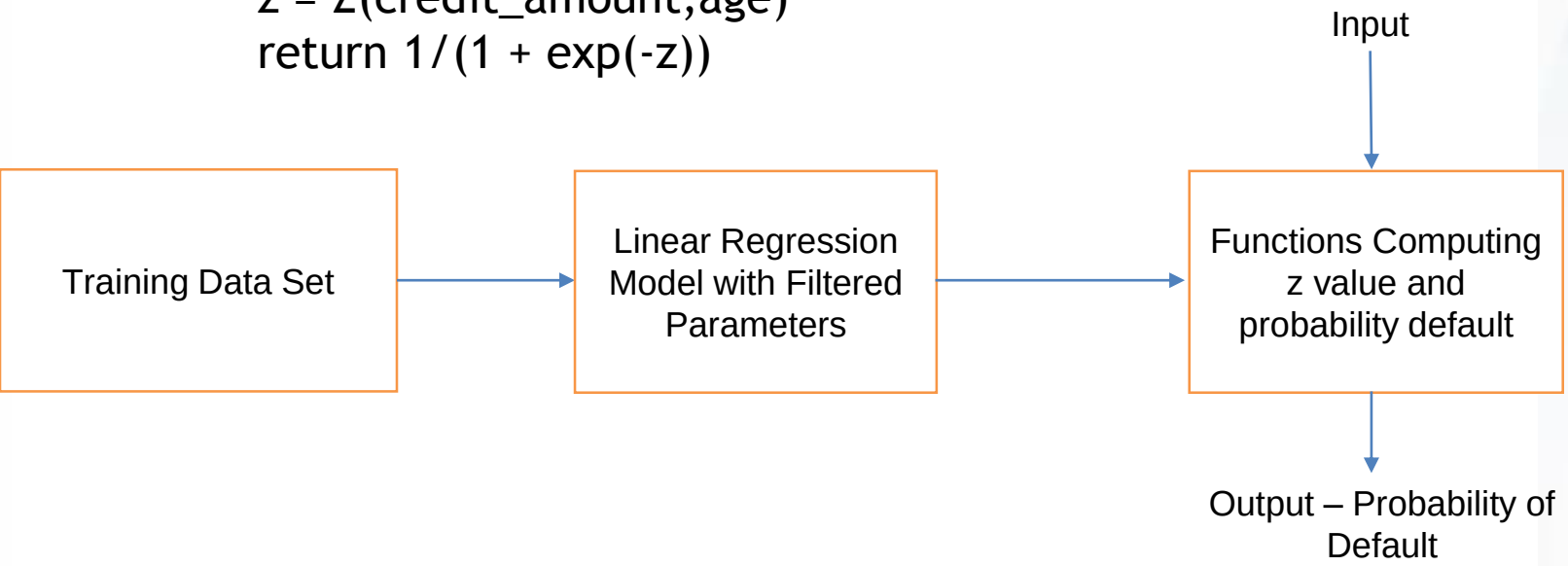
- The Code imports a training data set from Kaggle.
- It prepares the data by filtering out the required parameters.
- The required parameters are then applied to the logistic regression model estimator implemented from the class sklearn.linear_model.
- So the obtained values of the intercept and coefficients are fed to the function which computes the value of Z in the formula Pr(default = 1/X) = 1 / 1 + exp(-Z) where Z = w0 + w1.LimitBalance + w2.Age (where w0,w1 and w2 are constants obtained from the logistic regression model).
- This value of Z is then fed to the function that computes the probability default for any input value of the parameters of previous credit amount and the age.
- This also provides necessary statistical visual plots along with the probability of default which are convenient for banks to utilize in order to determine the customer's credibility.

- Interpretation with Algorithms & Pseudocode used.

Algorithm used is given as:-

```
func Z(pay_amount,age):
    return intercept[0] + coeff[0][0]*pay_amount + coeff[0][1]*age

func prob_default(credit_amount,age):
    z = Z(credit_amount,age)
    return 1/(1 + exp(-z))
```

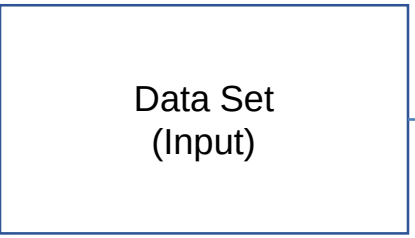


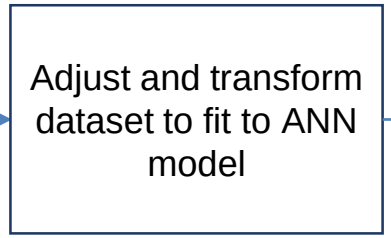
Implementation Details

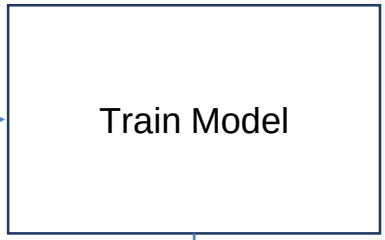
Algorithm for delinquency flagging - Model building

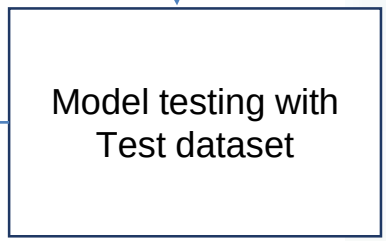
1. Import required libraries and Dataset needed
2. Data Pre-processing steps where required columns for the model is kept and rest is dropped
3. Encoding the Categorical data to numeric data using LabelEncoder from sklearn
4. Split data into training and test dataset(80-20)
5. Feature scaling using Standard scaler using sklearn
6. Building the ANN framework using keras- An input layer, 2 hidden layers and 1 output layer
7. Training the ANN model using adam as an optimizer and binary crossentropy as the last layer (output layer) loss function, with batch size 60 epochs
8. Testing and calculation of accuracy of model using accuracy_score()
9. Performance evaluation using confusion matrix

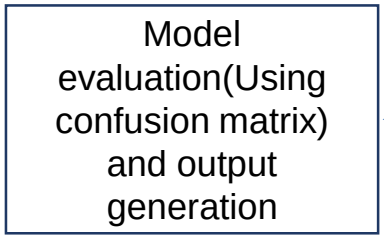

PES


Data Set
(Input)


Adjust and transform dataset to fit to ANN model


Train Model


Model testing with Test dataset

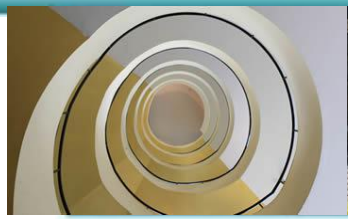

Model evaluation(Using confusion matrix) and output generation




References

- 1Charles Kwofie, 1Caleb Owusu-Ansah and 2Caleb Boadi “Predicting the Probability of Loan-Default: An Application of Binary Logistic Regression” 1Department of Statistics, University of Ghana, P.O. Box LG 25, 2 School of Business, University of Ghana, P.O. Box LG 25, Legon-Accra, Ghana- 2019
- Amir Ahmad Dar1* , N. Anuradha2 and Shahid Qadir3 “Estimation of pd of diffferent firms” by Journal of Global Enterpreneurship Research- 2019
- Apapan Pumsirirat, Liu Yan,School of Software Engineering, “Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine”, Tongji University, Shanghai, China (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 1, 2018
- UCI Machine Learning Repository. (2017, Nov. 29). Stalog (German credit data) dataset [Online]. Available: http://archive.ics.uci.edu/ml/datasets/Statlog+%28German+Credit+Data%29
- Kaggle. (2020, September). Credit Card Transactions Fraud Detection Dataset [Online]. Available: https://www.kaggle.com/kartik2112/fraud-detection?select=fraudTest.csv
- Single Hidden Layer Neural Network [Online]. Available: http://nicolamanzini.com/single-hidden-layer-neural-network/
- Charles Kwofie, 1Caleb Owusu-Ansah and 2Caleb Boadi Predicting the Probability of Loan-Default: An Application of Binary Logistic Regression 1Department of Statistics, University of Ghana, P.O. Box LG 25, 2 School of Business, University of Ghana, P.O. Box LG 25, Legon-Accra, Ghana- 2019
- Amir Ahmad Dar1* , N. Anuradha2 and Shahid Qadir3 Estimation of pd of diffferent firms by Journal of Global Enterpreneurship Research- 2019
- Apapan Pumsirirat, Liu Yan,School of Software Engineering, Credit Card Fraud Detection using Deep Learning based on Auto-Encoder and Restricted Boltzmann Machine, Tongji University, Shanghai, China (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 9, No. 1, 2018

PES UNIVERSITY






Thank You