Creditcard Fraud Prediction

Introduction

Overview

Credit Card Fraud Prediction is the process which is used to predict fraudranlent transcations. As this prediction is important that is related to finance section.
The credit card fraud detection features uses user behavior and location scanning to check for unusual patterns. These patterns include user characteristics such as user spending patterns as well as usual user geographic locations to verify his identity. If any unusual pattern is detected, the system requires reverification.

Linear Regression is used in this project to know whether the transactions are fraud or not.

LITERATURE SURVEY

Existing problem

Methods to solve this problem are:

- Logistic Regression
- Decision Tree
- Random Forest
- Naive Bayes
- ANN Model

Proposed solution

Solution that I suggest for the project is Linear Regression .
Linear Regression works with sigmoid function because the sigmoid function can be used to classify the output that is dependent feature and it uses the probability for classification of the dependent feature.
This algorithm works well with less amount of data set because of the use of sigmoid function. If the value of the sigmoid function is greature than 0.5 the output will be 1.If the output of the sigmoid function is less than 0.5 then the output is considered as 0.

Hardware/Software designing

Software Requirements:

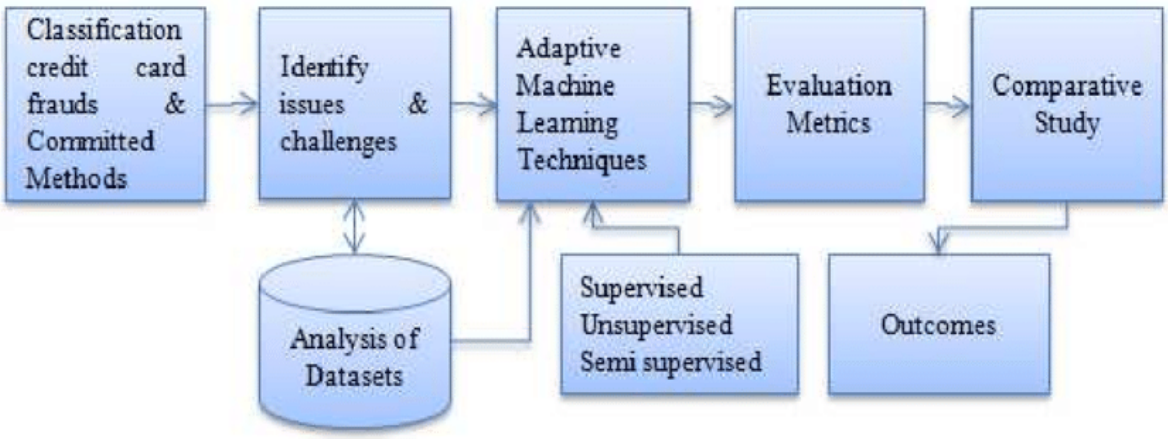
- Java JDK10
- Weka
- Eclipse IDE

Hardware Components:

- Processor – i3,i5
- Min Hard Disk – 4 GB
- Min Memory – 4GB RAM

THEORITICAL ANALYSIS

Block Diagram for the Credit Card fraud Detection

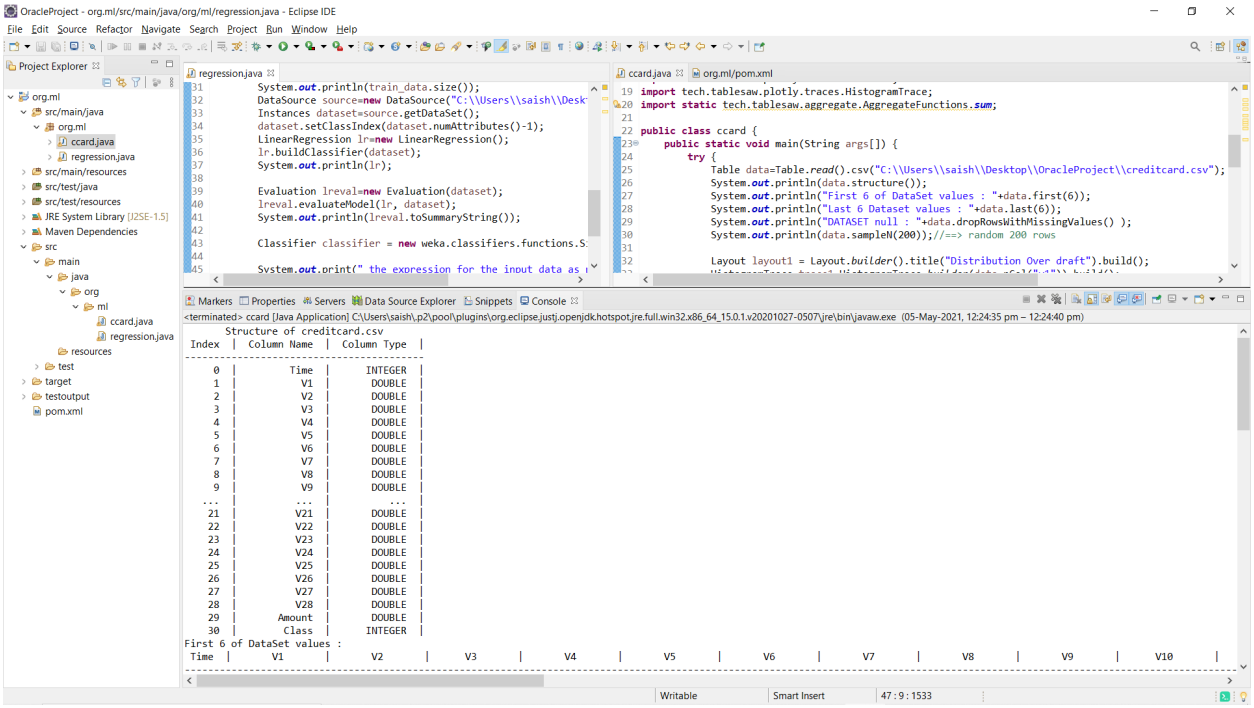


Experimental Investigations

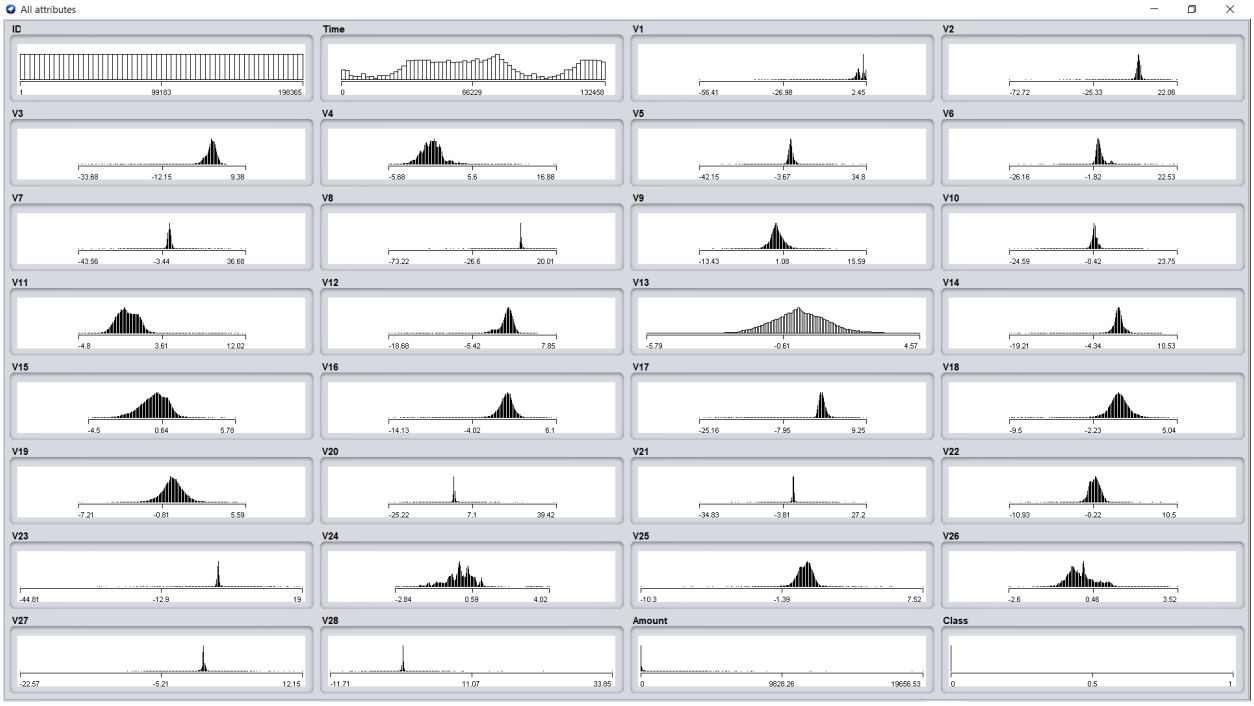
Analysis is using in this project to make sure which method is best to know or predict the fradulent transcations.

Result

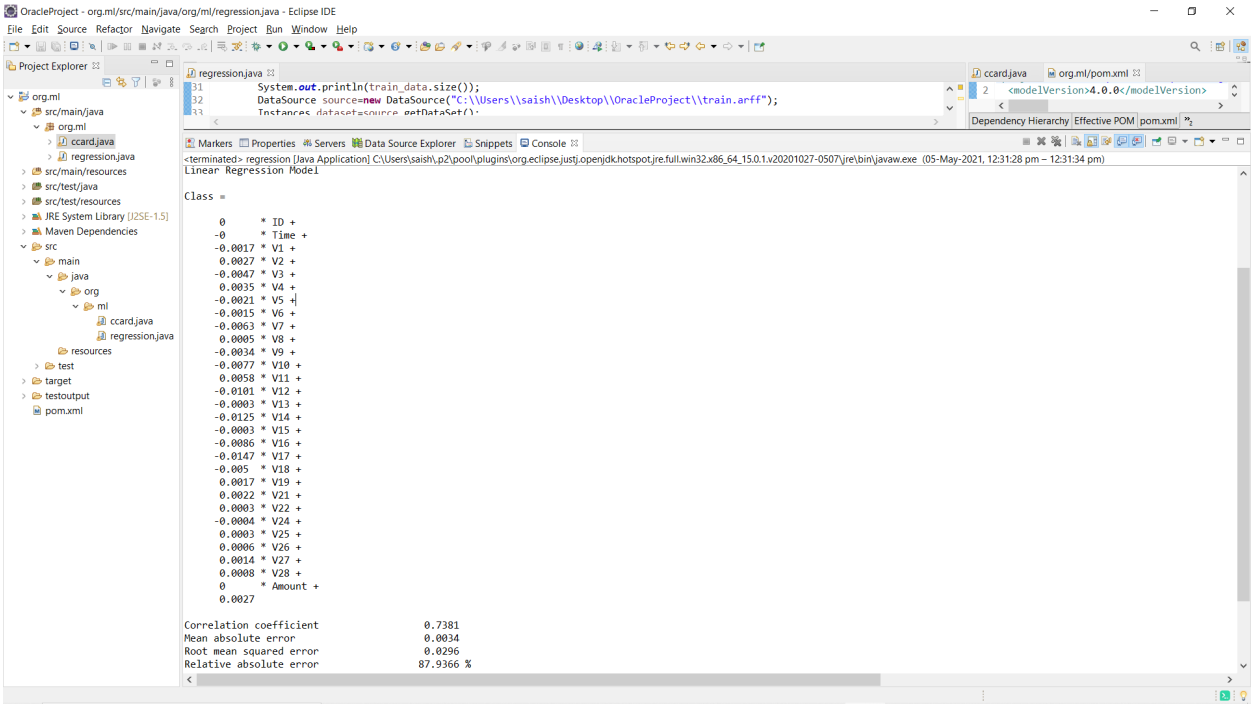
Structure of the Credit card DataSet



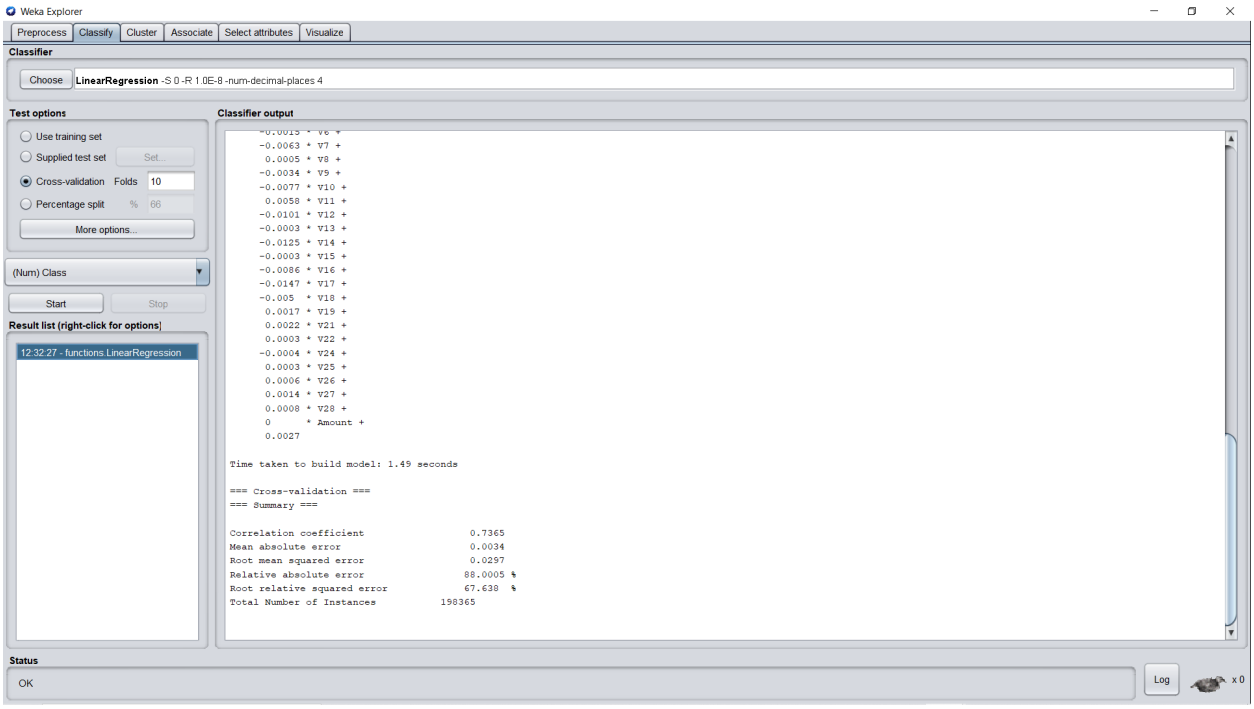
Visualization of Dataset(All attributes)



Linear Regression(The solution is used for this project)



The Linear Regression solution through WEKA



Advantages & Disadvantages:

Advantages:

- Linear Regression is simple to implement and easier to interpret the output coefficients.
- When we know the relationship between the independent and dependent variable have a linear relationship, this 'Linear Regression ' is the best to use because of it's less complexity to compared to the other algoritms.

Disadvantages:

- Linear regression technique outliers can have huge effects on the regression and boundaries are linear in this technique.
- Diversely, linear regression assumes a linear relationship between dependent and independent variables. That means it assumes that there is a straight-line relationship between them. It assumes independence between attributes.

Conclusion:

It is important for credit card companies to be able to recognize fraudulent credit card transactions so that customers may not face the issue regrading the transcations. So, this project helps to predit whether the transactions are fraud or not.

Refrences:

Dataset for Credit card fraud prediction:
https://www.kaggle.com/mlg-ulb/creditcardfraud

> Source code:
> https://github.com/smartinternz02/SPS-10710-Creditcard-Fraud-Prediction-