

Machine Learning Nanodegree Capstone Project

Credit Card Fraud Detection using Supervised learning

Proposed by : Ahmed Roshdy

Domain Background:

This project is proposed on the domain of supervised learning, supervised learning refers to a class of systems and algorithms that determine a predictive model using data points with known outcomes. The model is learned by training through an appropriate learning algorithm (such as [linear regression](#), [random forests](#), or [neural networks](#)) that typically works through some optimization routine to minimize a loss or error function

Problem Statement :

The main objective of this project is to detect the fraud credit cards , Credit card companies shall be able to recognize fraudulent credit card transactions so that customers are not charged for items that they did not purchase

Datasets and Inputs:

The dataset used in this project is produced by kaggle website and available to download from [here](#)

the dataset is labeled and consisting of 284807 rows and 31 columns , and our target feature that we should predict is the Class column , and if the value of the class is 0 that means it's transaction without fraud , and if it is equal 1 that means it's transaction with fraud.

Solution statement:

The proposed solution to this problem is to apply supervised learning algorithm to detect if the new transaction is fraud or not , and we will use the logistic regression model .

Benchmark Model :

For the benchmark model, we will use logistic regression model.

Evaluation Metrics:

The evaluation metric for this problem is simply precision, recall , and, f1-score

Project Design:

- Preprocessing

we will show the data and see the shape of it

- Data splitting

Split the data into a training set and validation set with an 80-20 split.

- Model training and evaluation

- I will start with the simple model architecture first before training and evaluating it. Then splitting the data , fit the model , and test the model by using predict function