[Credit Card Fraud Detection using Machine   
Learning Approach]

[Shanti Mutyala]

**Data Science Capstone Project   
Exploratory Data Analytics Report**

Date:

[09/02/2023]

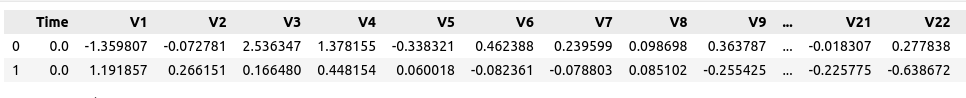
Team Members:

Name: Shanti Mutyala

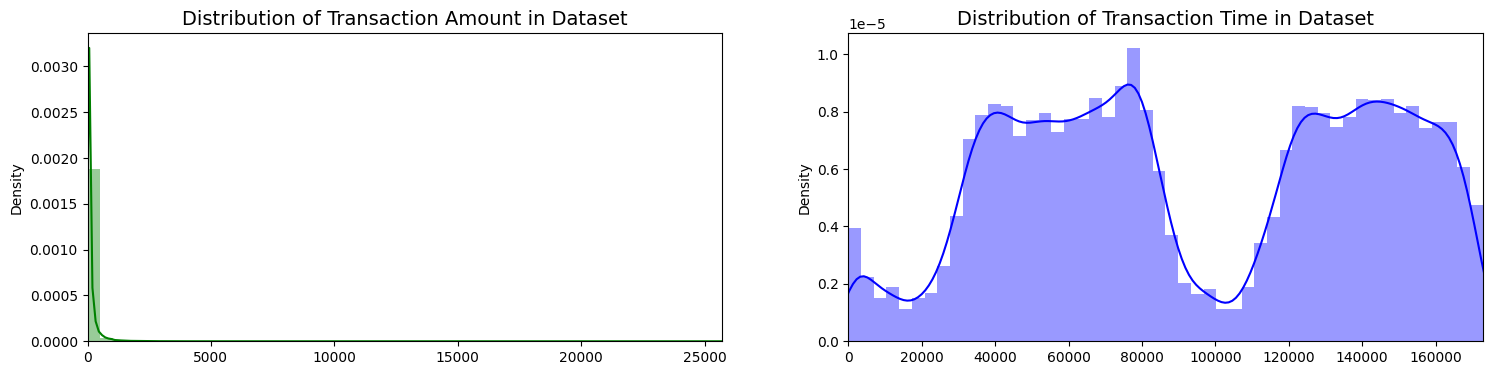
**Credit Card Dataset**

**Analysis the basic metrics of variables**

1. The dataset contains transactions made by credit cards in September 2013 by European cardholders.
2. This dataset presents transactions that occurred in two days, where I have 492 frauds out of 284,807 transactions. The dataset is highly unbalanced, the positive class (frauds) account for 0.172% of all transactions.
3. Below is sample dataset example



**Non-graphical and graphical univariate analysis**

1. I scaled the Amount and Time data using Robust Scaler as I wanted scaling to be less affected with outliers. Below is the result before scaling.

**Missing value analysis and outlier analysis**

1. I inspected the data and found that there was no null or missing value in it.

**Feature engineering and analysis**

1. I divided our data into train and test sets so that I can see the performance effectively on unseen data as well.
2. I also used some Oversampling and Undersampling as well on the train dataset so as to see if any performance improvement or not in using various models.

**Student Performance Dataset**

**Analysis the basic metrics of variables**

1. The dataset is related to performance of students in various exams and some information such as Level of Education, Race, Gender, etc.
2. The data consisted of 1000 rows and I chose Math Score as our target column and formulated the problem as a regression technique.

**Missing value analysis and outlier analysis**

1. I inspected the data and found that there was no null or missing value in it.

**Feature engineering and analysis**

1. I did the One Hot Encoding for the variables since most of the variables were categorical types.
2. I used Cross validation technique for evaluation of the model performance.

Table of Contributions

The table below identifies contributors to various sections of this document.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Section** | **Writing** | **Editing** |
| **1** | **Analysis the basic metrics of variables** | **Shanti** | **Shanti** |
| **2** | **Non-graphical and graphical univariate analysis** | **Shanti** | **Shanti** |
| **3** | **Feature engineering and analysis** | **Shanti** | **Shanti** |
| **4** | **Appendix** |  |  |

**Grading**

The grade is given on the basis of quality, clarity, presentation, completeness, and writing of each section in the report. This is the grade of the group. Individual grades will be assigned at the end of the term when peer reviews are collected.