VERZEO

Minor Project

- **Project Name:**

Machine Learning July Minor Project

- **Project Description:**

According to the Data Breach Index, more than 5 million records are being stolen on a daily basis, a concerning statistic that shows - fraud is still very common both for Card-Present and Card-not Present type of payments.

 In today's digital world where trillions of Card transaction happens per day, detection of fraud is challenging.As fraudster change tactics, learning algorithm can be improvised adding more analyzed features

As a data scientist, you are required to construct a ML model based on the available data and justify how mature your model is for industry (bank- payment gateways or VISA / Mastercard) in categorization & authorization of transaction based on efficiency in fraud detection.

- Explore other performance characteristics like accuracy, specificity, the area under the precision-recall curve, confusion matric etc. on the given dataset.

- Identify all possible features critical in the identification of card frauds.

Flow of your approach -

- Understanding -Problem statement

- Perform exploratory data analysis

- Shape of your data

- Display few rows of dataframe

- Data types of all columns

- Statistics summary of the features

- Check missing values for each feature
- Distribution of target variable
- Distribution of all features
- Boxplots
- Outliers
- Correlation Heatmaps
- Conclusions
- Preprocess the data
- Select Training data, test data

Following are some points for you to take note of, while doing the assignment:

- State all your assumptions clearly
- Provide clear explanations to justify your stand

Dataset Link:

https://drive.google.com/file/d/1P_UXMfF7fIShogzRLHXWyUKqVuRyzgEr/view?usp=sharing