Predicting Customer claim for insurance policy

PREPARED BY: Jose Baby (Emp. Id. 2211583)

Prepared as part of Kaggle projects given by Mr. Mrityunjoy Pandey.

**Data Description**

Data you will predict whether a customer made a claim upon an insurance policy. The ground truth claim is binary valued, but a prediction may be any number from 0.0 to 1.0, representing the probability of a claim. The features in this dataset have been anonymized and may contain missing values.

**Steps that I have followed throughout the project:**

1. Importing all the necessary libraries and the dataset.
2. Pre-processing the data, applying missing value imputation, scaling, etc
3. Creating 13 CatBoost Classifiers.
4. Then using the VotingClassifier class from sklearn to combine conceptually different machine learning classifiers and use a majority vote or the average predicted probabilities (soft vote) to predict the class labels. In majority voting, the predicted class label for a particular sample is the class label that represents the majority (mode) of the class labels predicted by each individual classifier.
5. Storing the results in a separate csv file.