* The training data consists of majority categorical features.
* The data does not seems to be having much of Outliers. I did not apply any anomaly detection.
* We have only three continuous features: **Age, Vintage, Avg\_Account\_Balance.**
* I have solved the problem with four different algorithms: **Catboost**, **Light Gradient Boosting**, **Random Forest** and **Gradient Boosting Classifier**.
* I have used **label** and **One-hot encoding** to transform all the features to numerical.
* The feature: Credit\_Product had missing values. In order to fill the missing values, I have used **forwardfill** for imputation. The reason for using this is to maintain the ration of the categoried(Yes:No) at a constant(same as before imputation)
* For using CatboostClassifier, we need to have only categorical features. So I have transformed the three continuous features into categorical feature using binning specifically **KBinsDiscretizer.**
* For getting more insights, I have made combinations of all column\_pairs and started taking the unique values from 1st col by using GROUPBY on 2nd column.
* I got my **highest score** with **LGBM Classifier with extensive hyperparameter tuning(0.80078).** 
  + For validation, I have used **RepeatedStratifiedKFold.**
  + In order to avoid **overfitting,** I have used **learning\_rate**=0.05 with **decaypower** of 0.99
  + Also I have used **early\_stopping\_rounds.** This is used, again for the same reason.(prevent overfitting)