Credit Card Lead Prediction

Happy Customer Bank

INDUSTRY TREND

Credit Cards are usually allotted based on Age Profile

- Credit Profile of Customer
- Salary , Account Balance
- Age of Customer in Bank
- Transactional History

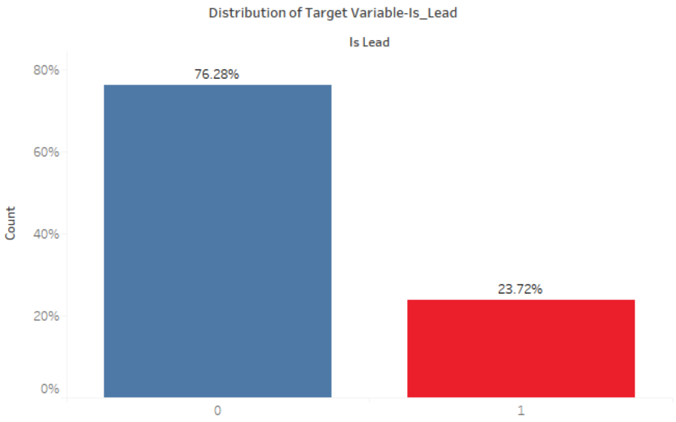
Approach

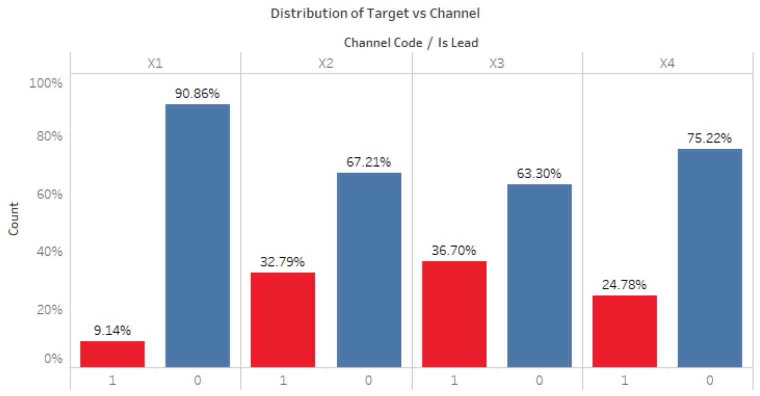
A brief on the approach, which you have used to solve the problem

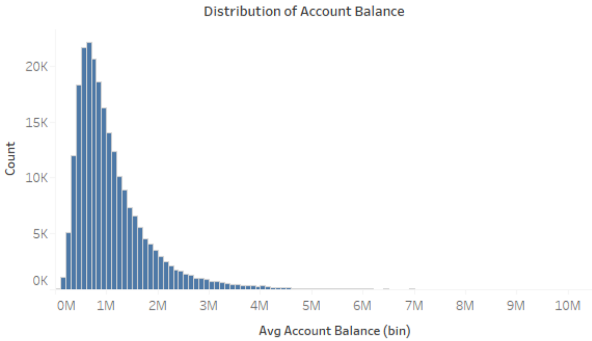
What data-preprocessing / feature engineering ideas really worked? How did you discover them?

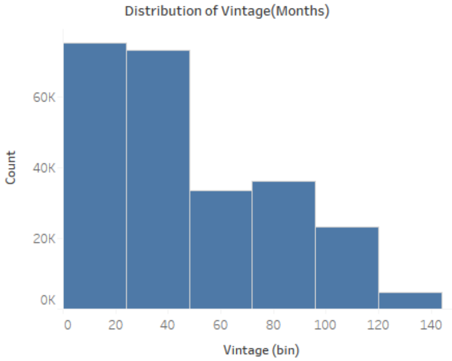
What does your final model look like? How did you reach it?

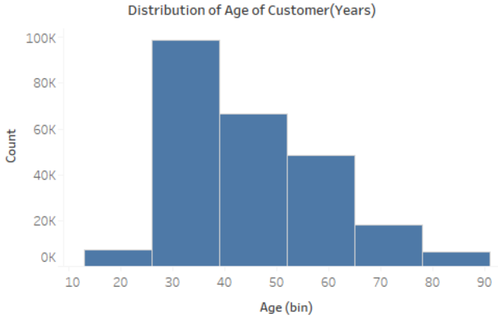
Exploratory Data Analysis





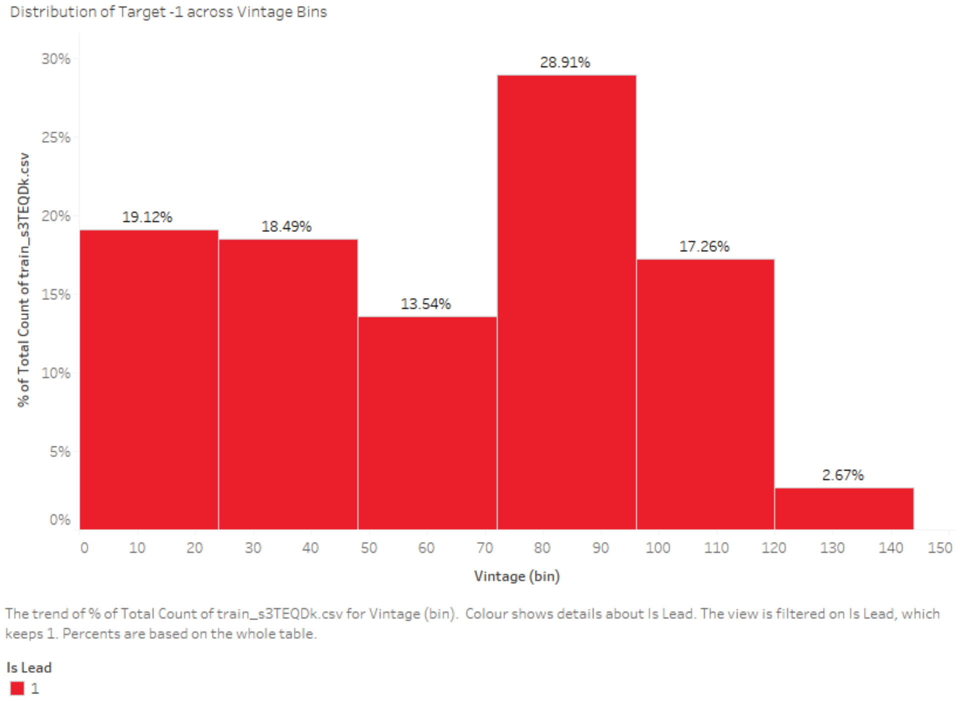


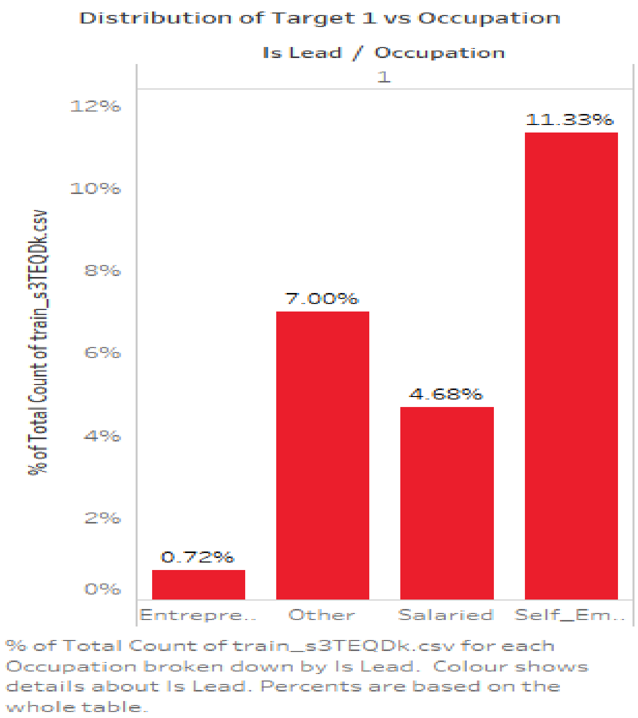


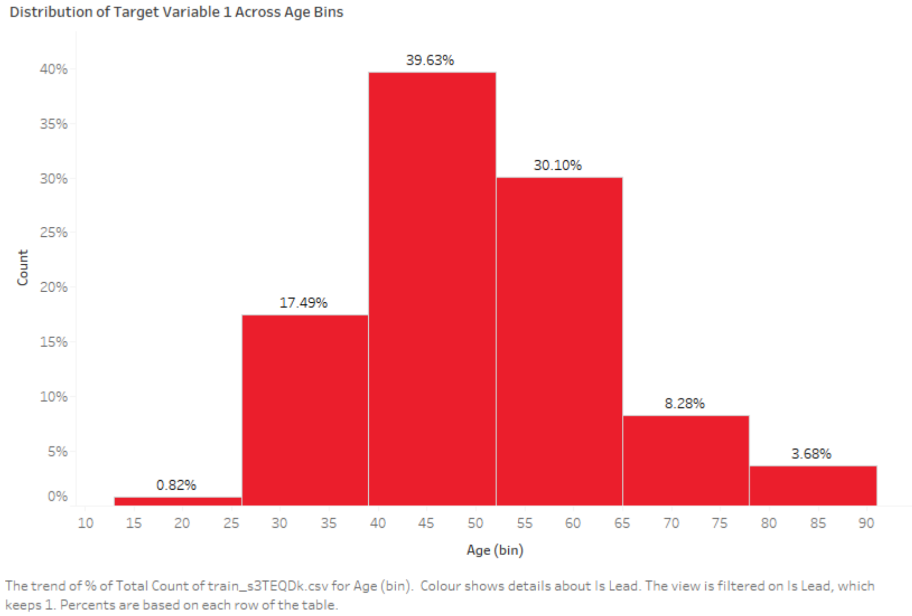


Is_Lead	23.72% conversions
Channel	X2,X3
Continuous Variables	Right skewed Account _Balance

Target 1 Distribution

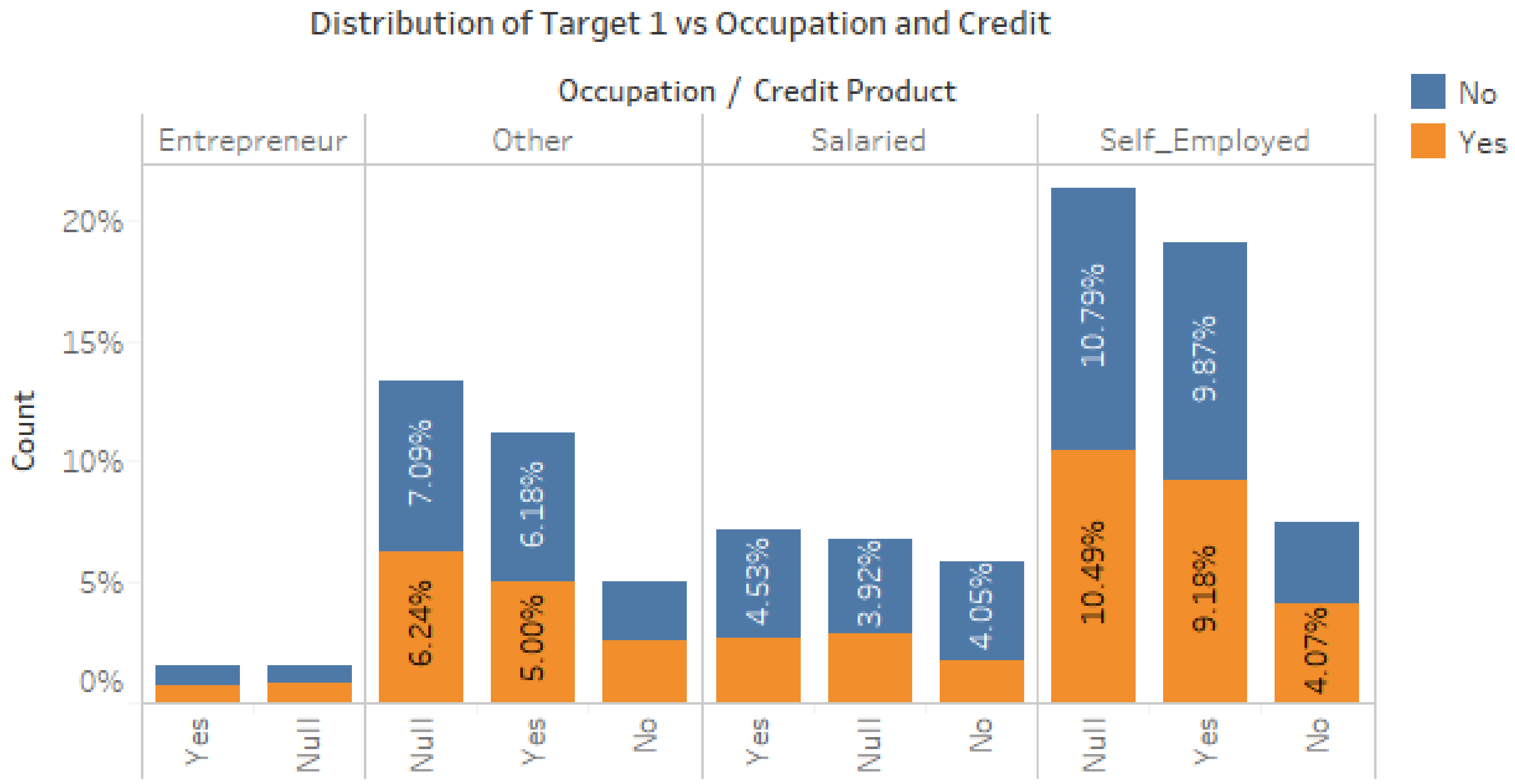






Age Bins	Max conversions from Age-Bins 40-60
Vintage Bins	Max conversions from Vintage Bins 46% from 6-8 years,32% from 0-4years
Occupation	Max Conversions from Occupation Self Employed

Target 1 vs Occupation|Credit



% of Total Count of train_s3TEQDk.csv for each Credit Product broken down by Occupation. Colour shows details about Is Active. The data is filtered on Is Lead, which keeps 1. Percents are based on the whole table.

Occupation and Credit have a higher impact

Building Interaction Features might help

Pre-Processing

Min-Max Scaler of Age, Vintage , Account Balance

Label Encoding of Categorical variables

Frequency Encoding of interaction features since distribution of data across categorical features is captured

Reducing levels of Region – Top > 5000 were identified and grouped

Feature Engineering

Total of 138 features were built

Dropped ID

Aggregate Features of interaction features – min, max,mean,sum ,standard deviation of all features

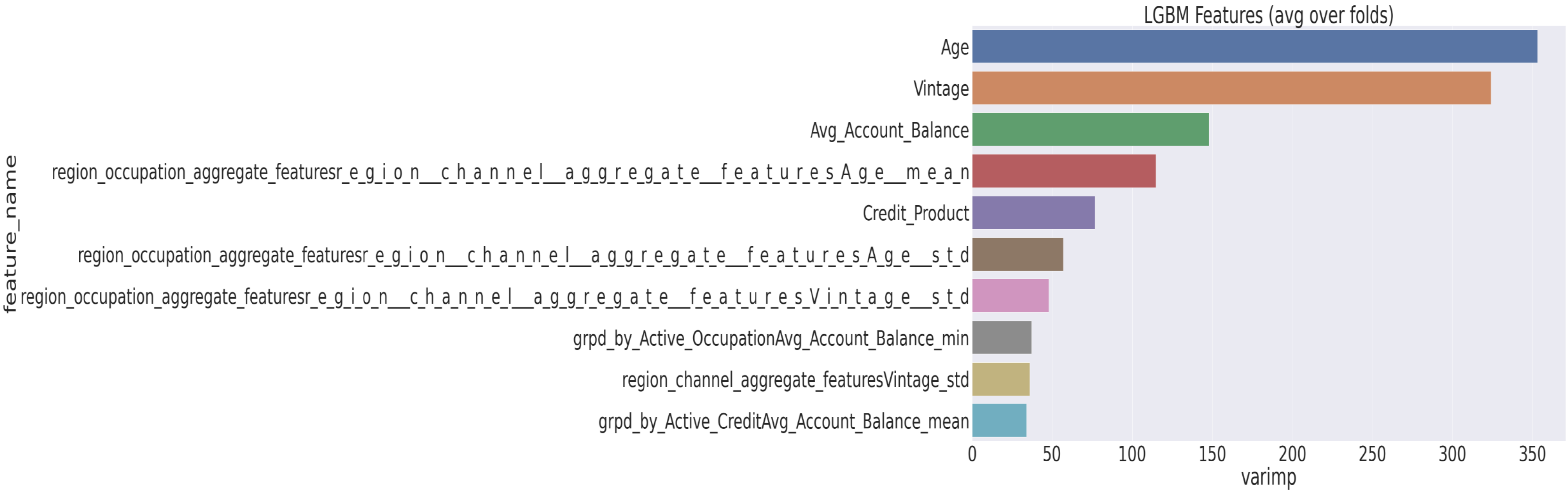
. Final shape

Train Predictors-(245725, 137)

Test Predictors-(105312,137)

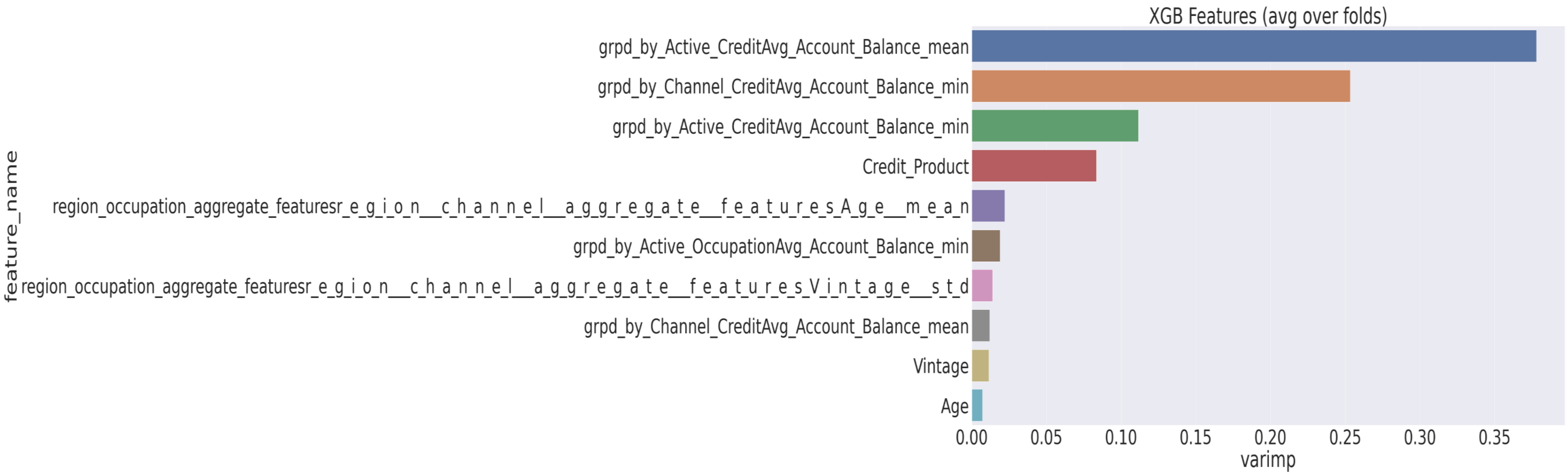
Models Used-LGBM

Average Stratified-KFold Score : 0.874109277726837



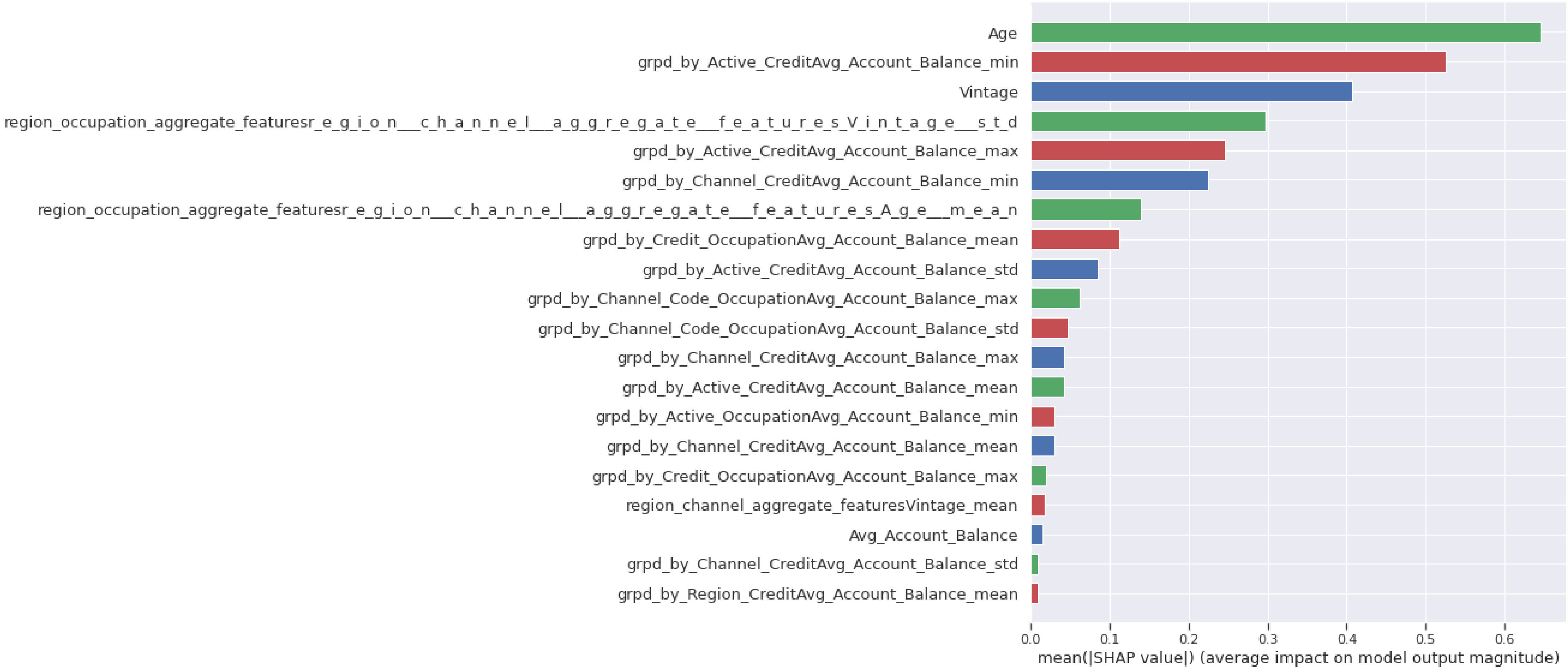
Models Used-XGBM

Average StratifiedKFold Score :
0.8731464590413118

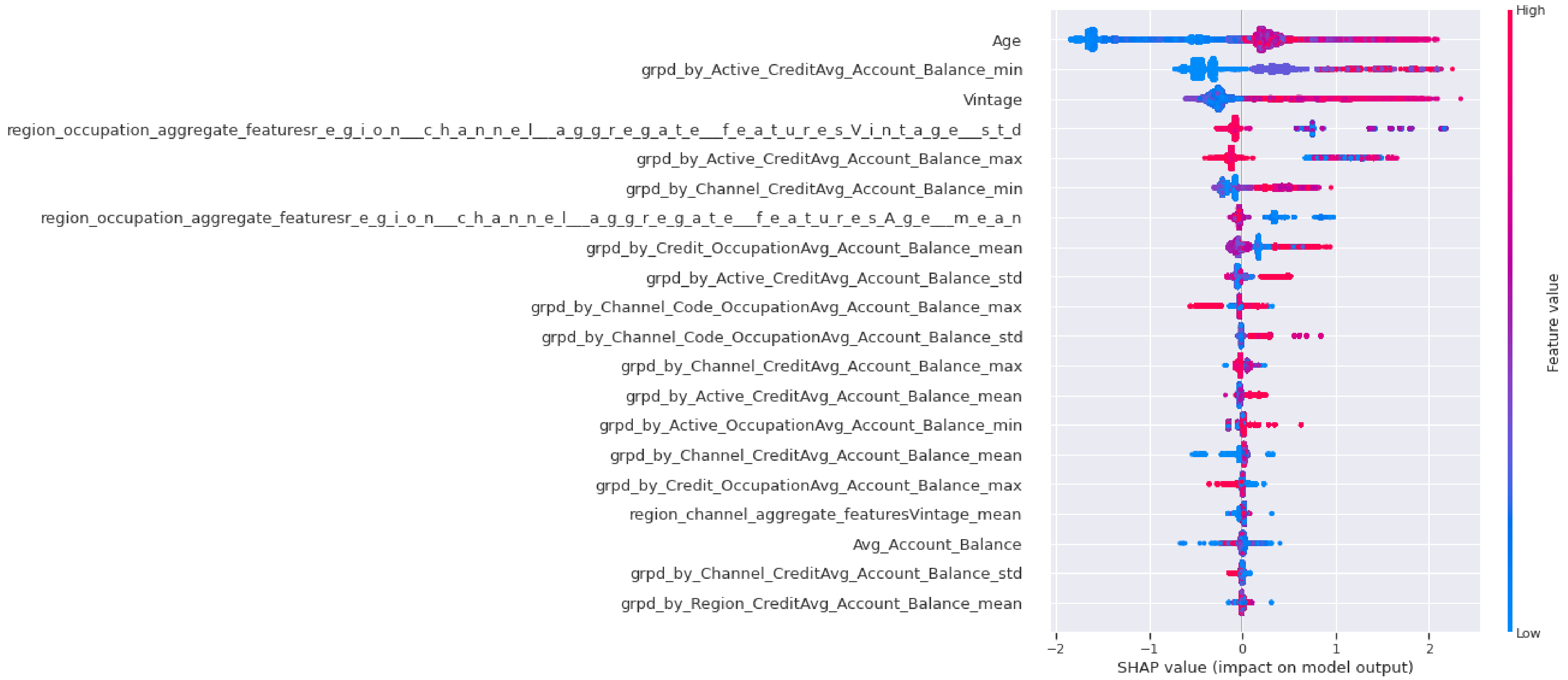


Models Used-CatBoost

Average Stratified-KFold Score : 0.8736018320655713



Models Used-CatBoost-SHAP values



Thank You