

Bank Churn Prediction

Project 4
(PGP-AIML-BA-UTA-OCT23-A)

Saurabh Mittal

April 5, 2024

Contents / Agenda

- Executive Summary
- Business Problem Overview and Solution Approach
- EDA Results
- Data Preprocessing
- Model Performance Summary
- Appendix

Executive Summary

- The objective of this study is to build a neural network-based classifier that can determine whether a customer will leave the bank or not in the next 6 months
- The data contains 10000 records: 14 attributes, with no missing values of the attributes. The data was of good quality.
- The data from the past 12 months was analyzed using Machine Learning methodology that revealed the following insights
 - About 20% of the members exited
 - 30% of German, and ~20% of Spain and France members exited
 - About 23% of non-active members exited. About ~8% of active members exited
 - Members between 40-50 age group exited
 - No significant positive or negative correlation between variables exist
- Insights and Recommendations
 - Six different neural networks were trained
 - Neural network incorporating SMOTE, Adam and Dropout provided the maximum Recall and the best validation

Business Problem Overview and Solution Approach

- Problem
 - Customers' leaving credit cards services would lead bank to loss, so the bank wants to analyze the data of customers and identify the customers who will leave their credit card services and reason for same
 - Develop a neural network-based classification model that will help the bank improve its to understand which customer will leave in the next 6 months
- Solution approach / methodology
 - Bank would want Recall to be maximized, greater the Recall higher the chances of minimizing false negatives. Hence, the focus should be on increasing Recall or minimizing the false negatives or in other words identifying the true positives(i.e. Class 1) so that the bank can retain their valuable customers by identifying the customers who are at risk of attrition.
 - Build a suite of neural network models employing: SGD, Adam, Dropout, and SMOTE, in various combinations
 - Select the optimally tuned model and verify with test set

EDA Results

EDA Summary

- The dataset has 10000 rows and 14 columns
- Only 3 variables are object types, rest all are numerical types.

Data Cleaning

- Drop "RowID", "CustomerID" and "Surname" as they are unique for each candidate and might not add value to modeling.

Data Preprocessing

- Duplicate value check: No duplicates
- Missing Value treatment: None
- Outlier check
 - none
- Data preparation for modeling
 - Training set has 6000 records, Validation set has 2000 records, Test set has records
 - Total columns: 11
 - Normalized attributes: "CreditScore","Age","Tenure","Balance","EstimatedSalary"

Model Performance Summary

Of the 6 neural network models, the best model incorporated SMOTE, Adam and Dropout

Training Set

Training performance comparison

	recall
NN with SGD	1.000000
NN with Adam	0.476697
NN with Adam & Dropout	0.463614
NN with SMOTE & SGD	0.757798
NN with SMOTE & Adam	0.853883
NN with SMOTE,Adam & Dropout	0.867071

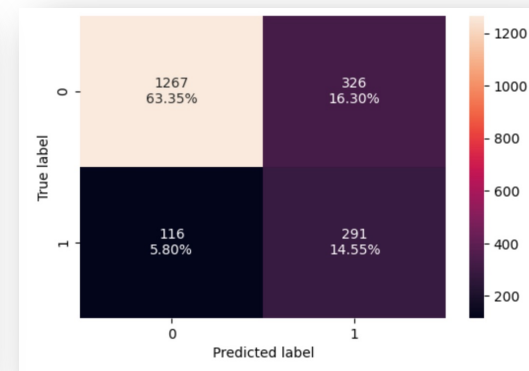
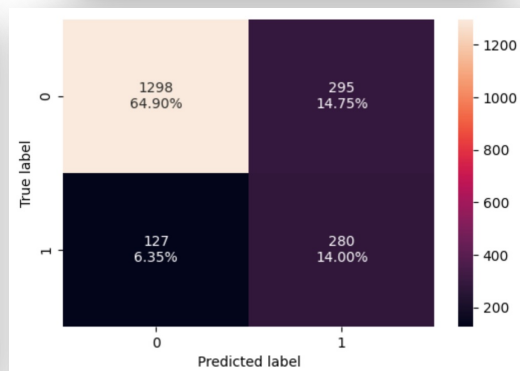
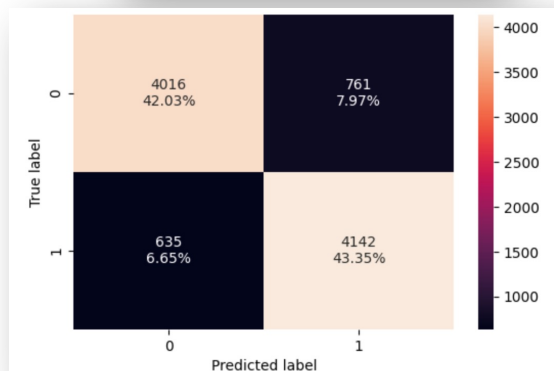
Validation Set

Validation set performance comparison

	recall
NN with SGD	1.000000
NN with Adam	0.425061
NN with Adam & Dropout	0.395577
NN with SMOTE & SGD	0.663391
NN with SMOTE & Adam	0.604423
NN with SMOTE,Adam & Dropout	0.687961

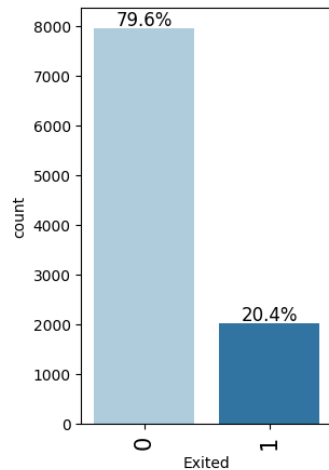
Test Set

	precision	recall	f1-score	support
0	0.92	0.80	0.85	1593
1	0.47	0.71	0.57	407
accuracy			0.78	2000
macro avg	0.69	0.76	0.71	2000
weighted avg	0.83	0.78	0.79	2000

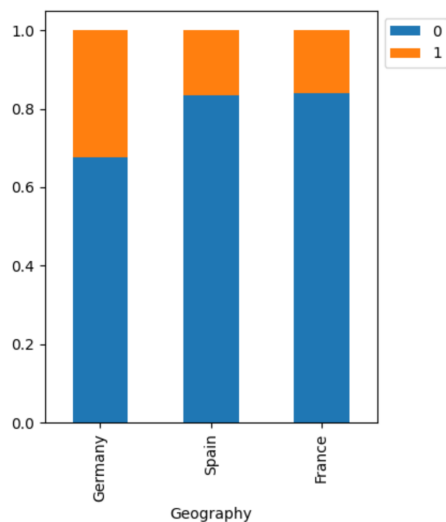


APPENDIX

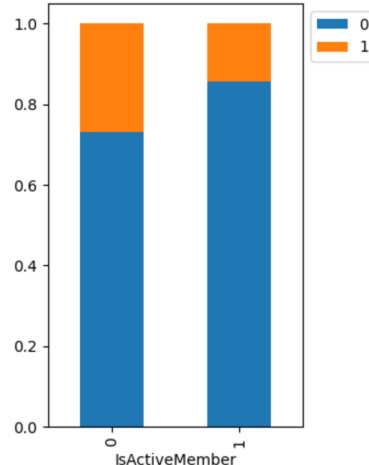
Variables and Distributions for Exited in the Past 12 Months



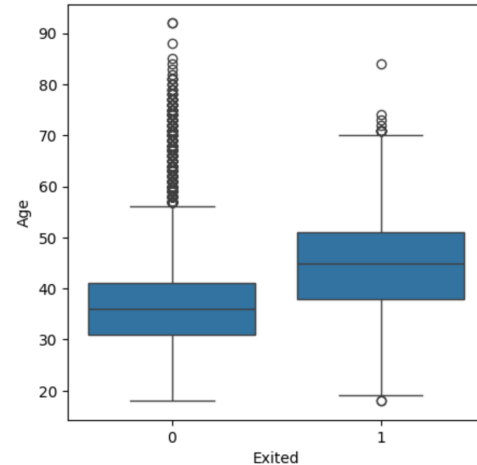
About 20% of members exited



30% of German, and ~20% of Spain and France members exited



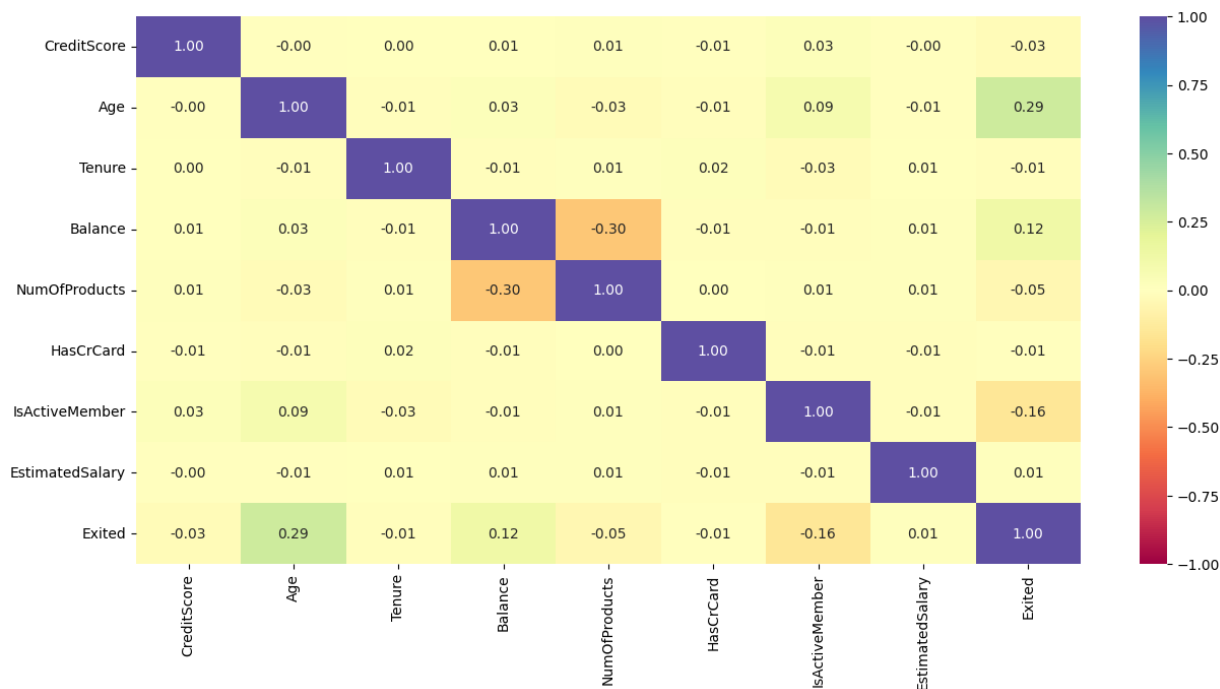
About 23% of non-active members exited. About ~8% of active members exited



Members between 40-50 age group exited

Correlation Findings

- No significant positive or negative correlation exists.
- Neural network is the preferred option for building a classifier.





Happy Learning !

