Customer churn in banking industry!

Introduction:

Customer churn: is the percentage at which customers stop doing business with an entity.
The competition in the banking industry is fast increasing, So the banks are required to implement customer retention strategies and try to increase their market share by acquiring new customers at the same time to stay in this industry. Improving customers retention rate by up to 5 % will increase your bank's profit by up to 85 %. Also, attracting new customers costs more than retaining the old customers who are more likely to produce more profit. Thus, banks may maintain their competitive advantage by using machine learning models to predict customer churn.

Data:

It consists of 10000 observations and 14 variables.
Independent variables contain information about customers. The dependent variable refers to customer abandonment.
Features:

• RowNumber: Row Number

• CustomerId: Customer ID

• Surname: Surname

• CreditScore: Credit score

• Geography: Country (Germany / France / Spain)

• Gender: Gender (Female / Male)

• Age: Age

• Tenure: How many years of customer

• Balance: Balance

• NumOfProducts: Bank product used

• HasCrCard: Credit card status (0 = No, 1 = Yes)

• IsActiveMember: Active membership status (0 = No, 1 = Yes)

• EstimatedSalary: Estimated salary

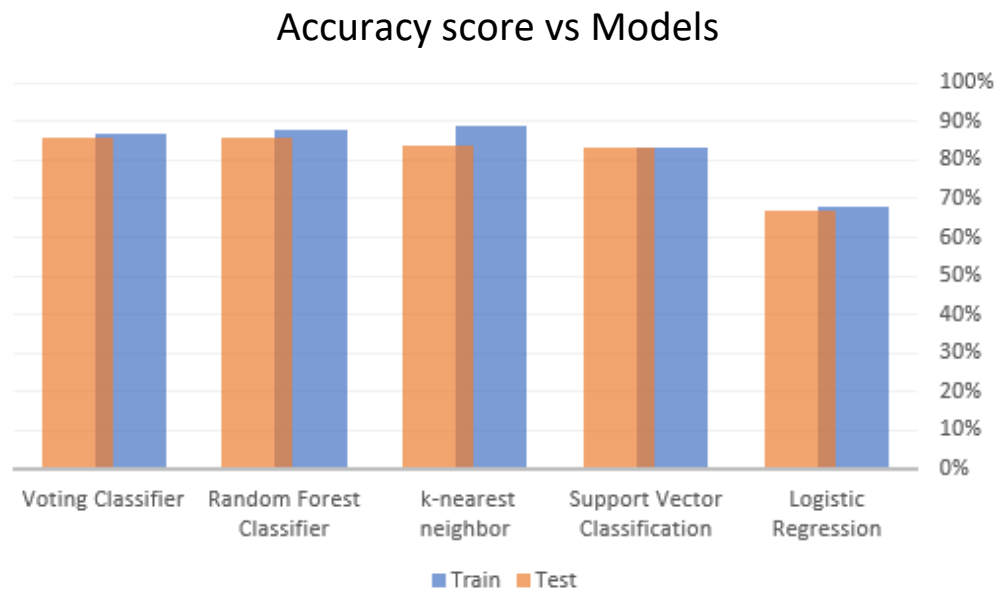• Exited: Abandoned or not? (0 = No, 1 = Yes)
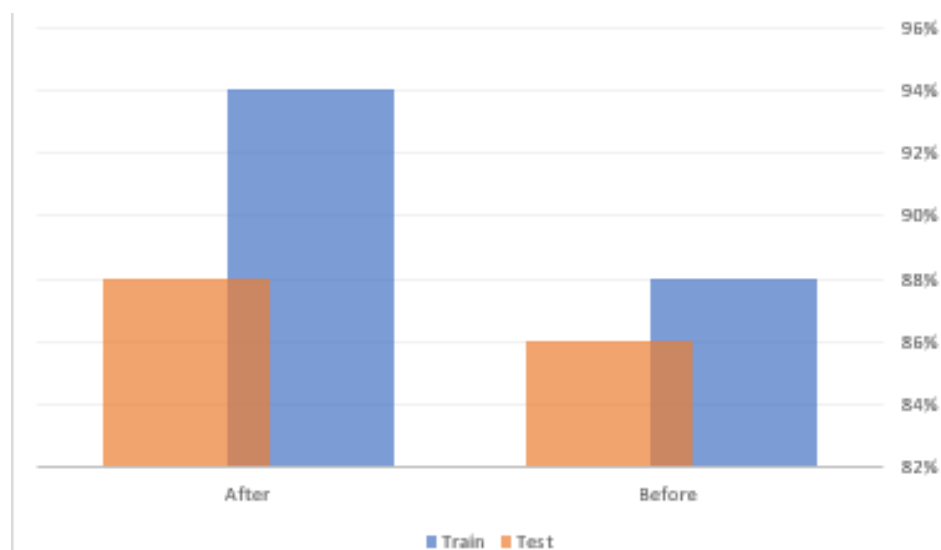
Tools:

Data processing: Pandas and NumPy.
Modeling: Scikit-learn.
Visualization: Matplotlib and Seaborn.

Algorithms

Accuracy score vs Models

Hyperparameter Tuning for random forest model

Conclusion

The accuracy of all algorithms was good, but after adjusting the random forest algorithm, it became better than others, so I recommend applying it.