APPROACH TO SOLVE THE BANK CHURN ANALYSIS PROBLEM:

- **A brief on the approach used to solve the problem.**

The target of the project was to analyze the past data and predict whether the customer will churn or not in the next 6 months. For this, I have tried to analyze the behavior of the customers who have churned in the past and also the patterns of those who have not churned in the past.

This analysis helped me to create a model that can differentiate the characteristics of the people who can churn from those who are less likely to churn.

As this is a supervised machine learning problem with categorical dependent variables, so I have used a few classification models to predict the target variable. The Random Forest Classifier turned out to be the best model.

- **Which Data-preprocessing / Feature Engineering ideas really worked? How did you discover them?**

1. IQR method to remove outliers worked as it helped in dealing with outliers present in the balance feature. Outliers were discovered using boxplot.
2. Applying standard scaler to all the features helped in reducing the skewness in the dataset. Skewness was discovered using histograms.
3. Applying label encoder helped in converting categorical data into numerical data. Came to know about all the categorical data with the help of data types information.

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

By using the visualization patterns, I have discovered the following characteristics of the individuals that could help in predicting whether a person will churn or not in the next 6 months:

Those who are more likely to churn:

1. Higher proportion of females churns as compared to that of males.
2. Those with poor credit category have high chances of churn.
3. Lower age group people are more likely to churn.
4. People with low balance have high probability of churning.

Those who are less likely to churn:

1. For those with income more than 15L are less likely to churn.
2. Those with 3+ product holdings are very less likely to churn.
3. Those who owns a credit card.

So, I have considered all these features as important one in predicting the churn behavior of the customers.