**Business Requirement Document** 

**Problem Statement:**

1. Despite the continuous efforts of RBC bank to attract and retain customers, the banking industry faces a persistent challenge in the form of customer churn, leading to financial losses and reduced customer satisfaction.

2. It is advantageous for RBC bank to know what leads a client towards the decision to leave the company.

**Project Objective: -**

1. The aim of this project is to analyse the customer churn rate for bank because it is useful to understand why the customers leave.

2. After Analysing we need to train a Machine Learning Model which can find the key factors that significantly influence the customer churn or attrition.

3. In the end will choose the most reliable model that will attach a probability to the churn to make it easier for customer service to target right customer in order to minimize their efforts to prevent customer’s churn.

**Project Overview:**

1. Churn refers to customers leaving a bank or discontinuing their banking services.

2. Banking Churn Analysis is a process of studying customer behaviour in the banking industry to predict and understand customer attrition or churn.

3. Banking Churn Modelling aims to identify patterns and factors that contribute to customer churn, enabling banks to take proactive measures to retain customers and improve customer satisfaction.

**Table of Contents:**

**1. Importing Libraries & Dataset.**

**2. Data Wrangling.**

\* Data Cleaning.

\* Handling Missing Values.

\* Handling Inconsistences.

**3. Exploratory Data Analysis (EDA)**

\* Visualizing Dependent Variable.

\* Visualizing Independent Variables.

\* Generating Insights.

**4. Data Preprocessing.**

\* Variable Selection and Importance.

\* Feature Transformation, Scaling and Encoding.

\* Splitting Data for Model Training.

\* Applying SMOTE to reduce class-imbalance.

**5. Model Creation, Training and Evaluation.**

\* Selection of Classification Algorithms.

\* Model Training and Tuning.

\* Model Evaluation and Performance.

\* Confusion Matrix Analysis.

\* Accuracy, Precision, Recall, and F1 Score.

\* Receiver Operating Characteristic (ROC) Curve and AUC.

\* Feature Importance and Contribution.

**Churn Analysis:**

Analyse the data and bring out few insights on the customer Churn.

It is advantageous for banks to know what leads a client towards the decision to leave the company.

**Churn prevention allows companies to develop loyalty programs and retention campaigns to keep as many customers as possible.**

**ABOUT DATA**

•Row Number—corresponds to the record (row) number and has no effect on the output.

•Customer Id—contains random values and has no effect on customer leaving the bank.

•Surname—the surname of a customer has no impact on their decision to leave the bank.

•Credit Score—can have an effect on customer churn, since a customer with a higher credit score is less likely to leave the bank.

**Credit score:**

•Excellent: 800–850

•Very Good: 740–799

•Good: 670–739

•Fair: 580–669

•Poor: 300–579

•Geography—a customer’s location can affect their decision to leave the bank.

•Gender—it’s interesting to explore whether gender plays a role in a customer leaving the bank.

•Age—this is certainly relevant, since older customers are less likely to leave their bank than younger ones.

•Tenure—refers to the number of years that the customer has been a client of the bank. Normally, older clients are more loyal and less likely to leave a bank.

* Balance—also a very good indicator of customer churn, as people with a higher balance in their accounts are less likely to leave the bank compared to those with lower balances.
* Number Of Products—refers to the number of products that a customer has purchased through the bank. Ex. FD, RD, Current Account, Insurance, SIP etc
* Has Credit Card—denotes whether or not a customer has a credit card. This column is also relevant, since people with a credit card are less likely to leave the bank.

•1 represents credit card holder

•0 represents non credit card holder

* Is Active Member—active customers are less likely to leave the bank.

•1 represents Active Member

•0 represents Inactive Member

* Estimated Salary—as with balance, people with lower salaries are more likely to leave the bank compared to those with higher salaries.
* Exited—whether or not the customer left the bank.

0 represents Retain

1 represents Exit

* Bank DOJ — date when the Customer associated/joined with the bank.

**Data Gathering:**

Please use the following data assets to pull the data related to Bank customer and associated details.

* Active Customer
* Bank Churn
* Credit Card
* Customer Info
* Exit Customer
* Gender
* Geography