**Customer Churn Analysis and Prediction**

**Author:**

**1. Mohamed Waleed - 41910017**

**2. Ahmed Nasser - 41920015**

**3. Omar Mahmoud Abdelkarim - 41910312**

**4. Omar Nasr Elden - 41910045**

**5. Ibrahim Emam Mahmoud - 41910270**

**6. Mostafa Mohamed Ibraheem – 41810276**

Executive Summary:

This report presents an in-depth analysis of customer churn in a banking dataset, aiming to identify key factors contributing to churn and build a predictive model for customer retention. The analysis includes data set review, exploratory data analysis, feature engineering, and model fitting.

Objectives:

1-Identify and Visualize Factors Contributing to Customer Churn

To understand the factors influencing customer churn, exploratory data analysis (EDA) was conducted. Key insights include:

Geographical Impact: Churn rate inversely related to population density, suggesting potential resource allocation issues.

Gender Influence: Female customers have a higher churn rate compared to males.

Credit Card Connection: Customers with credit cards show a higher likelihood of churn.

Inactive Members: Inactive members exhibit a greater churn rate, highlighting the need for targeted engagement programs.

2- Build a Predictive Model:

A predictive model was developed to classify customers as churned or retained and assign probabilities to aid customer service efforts. Model fitting included logistic regression, SVM with various kernels, and ensemble models.

Data Set Review & Preparation:

1-Data Structure Overview

- 1000 rows, 14 attributes.

- Questions raised about the snapshot nature of the data and the relevance of specific dates.

- Customers with a balance post-exit raise questions about product vs. bank exit.

2 Key Questions and Observations:

1.Temporal Snapshot: Clarify the date relevance; consider obtaining balances over time.

2.Exit with Balance: Investigate customers with balances post-exit for potential product exit.

3.Active Members: Define and explore degrees of activity; consider transaction counts.

4.Product Breakdown: Analyze product count for additional insights.

3. Exploratory Data Analysis (EDA):

1 Baseline Model and Customer Insights

- Baseline model predicts 20% churn due to the observed churn rate.

- Customer insights:

- Geographical disparities in churn rates.

- Gender and credit card influence on churn.

- Concerning proportion of inactive members.

Demographic and Behavioral Analysis:

- No significant credit score difference between retained and churned customers.

- Age impact on churn, suggesting a need to review retention strategies by age group.

- Tenure analysis reveals extremes more likely to churn.

- Customer churn impact on bank balances, posing a potential capital challenge.

- Product and salary insignificance on churn likelihood.

Feature Engineering:

- Derived features to enhance the model's predictive capabilities.

Data Preparation for Model Fitting:

- Split the dataset into training and testing sets.

Model Fitting and Selection:

- Explored logistic regression, SVM with various kernels, and ensemble models.

- Keen focus on predicting churned customers.

- Model fit accuracy reviewed for optimal selection.

Conclusion:

This report provides a comprehensive analysis of factors influencing customer churn and the development of a predictive model. Recommendations include addressing geographical disparities, targeting inactive members, and refining retention strategies based on age groups. The predictive model aims to empower customer service efforts to proactively retain customers and minimize churn impact.

Next Steps:

1.Address Data Questions: Obtain clarity on temporal snapshots and exit with balance scenarios.

2.Refine Model: Fine-tune the selected model based on ongoing data and feedback.

3.Implementation: Deploy the model in a real-world environment for continuous monitoring and improvement.