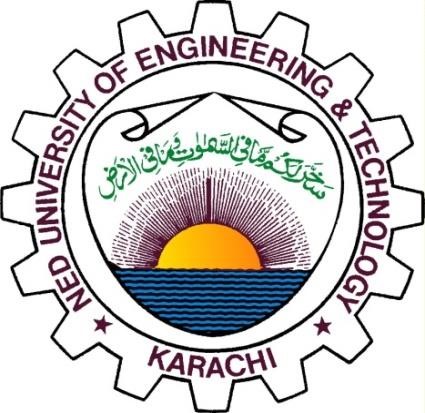
**FINAL PROJECT**



**NED ACADEMY**

**Centre for Multidisciplinary Postgraduate Programs**

**NED UNIVERSITY OF ENGINEERING AND TECHNOLOGY KARACHI, PAKISTAN**

**Batch: 4th**

**Name:Fazal Diyan Roll No.: May/2022**

Introduction

The Credit Card Churn Customer Dataset sourced from <https://leaps.analyttica.com/home> encompasses information about 10,000 customers, including attributes like age, salary, marital status, credit card limit, and category. With a mere 16.07% of customers identified as churning, the dataset presents a challenging scenario for predicting customer churn within the bank's credit card services. This study aims to delve into the dataset, conducting Feature Exploration, Exploratory Data Analysis (EDA), and applying Data Preprocessing techniques. Subsequently, three machine learning algorithms—Logistic Regression, Random Forest, and XG Boost—will be employed and compared to develop a robust predictive model for customer churn.

Objectives

The primary objective of this study is to build a predictive model capable of identifying customers at risk of churning from the bank's credit card services. Through Feature Exploration, EDA, and machine learning algorithms, the goal is to gain valuable insights into the dataset, optimize features, and develop a reliable model to enhance customer retention.

Scope of Work

The scope of this project involves a comprehensive exploration of the Credit Card Churn Customer Dataset, emphasizing Feature Exploration, EDA, and Data Preprocessing. The study extends to implementing machine learning algorithms, specifically Logistic Regression, Random Forest, and XG Boost, to predict and compare customer churn. Recommendations will be derived from the analysis to guide the bank in addressing potential churn risks.

Literature Review

Existing literature emphasizes the significance of predictive modeling in the banking sector to anticipate and manage customer attrition. Machine learning algorithms, including Logistic Regression, Random Forest, and XG Boost, have proven effective in addressing similar challenges across industries. Understanding features influencing customer churn is pivotal for developing targeted interventions and fostering customer loyalty.

Methodology

The methodology entails a thorough exploration of features within the dataset, followed by an in-depth EDA with various visualizations. Data preprocessing techniques will be applied to optimize the dataset for machine learning. Three machine learning algorithms—Logistic Regression, Random Forest, and XG Boost—will be employed and rigorously compared to identify the most effective model for predicting customer churn.

Findings

The study's findings will uncover key features influencing customer churn and provide insights into the performance of Logistic Regression, Random Forest, and XG Boost. Metrics such as accuracy, precision, and recall will be crucial in evaluating the models and determining their effectiveness in predicting churn.

Conclusions

Conclusions will be drawn based on the findings, highlighting the strengths and limitations of each machine learning algorithm. The study aims to offer actionable insights into developing a reliable predictive model for customer churn in the credit card services sector.

Recommendations

Recommendations will be provided to guide the bank in implementing strategies for customer retention based on the conclusions drawn from the machine learning models. These may include personalized retention offers, targeted communication, and enhanced customer service initiatives to minimize churn and foster long-term customer relationships.