

Algorithmic Machine Learning – Project Proposal

Group 7: Shreyash Mehta, Ronil Surve, Shubham Sharma, Namrata Rath

Topic: Bank Customer Churn Prediction

Abstract:

In the contemporary banking sector, the challenge of customer retention has become a paramount concern, given the fierce competition and dynamic market trends. This study focuses on the development of a robust Bank Customer Churn Prediction Model to identify and forecast potential customer churn, thereby aiding banks in implementing proactive retention strategies.

The proposed model employs machine learning techniques and algorithms to analyze historical customer data, including transaction patterns, demographics, customer interactions, and other relevant features. Leveraging a comprehensive dataset from a leading financial institution, the model undertakes extensive data preprocessing, feature engineering, and selection processes to ensure optimal performance.

The model aims to accurately anticipate customers who are likely to churn, enabling banks to implement targeted marketing campaigns, personalized offers, and proactive intervention measures. By integrating the predictive model into the existing customer relationship management (CRM) system, banks can proactively engage with at-risk customers, thereby fostering stronger relationships and enhancing customer loyalty.

The project's objectives include the development of an interpretable and scalable churn prediction model, the evaluation of various machine learning algorithms for performance comparison, and the implementation of an effective deployment strategy for seamless integration into the bank's operational infrastructure.

Ultimately, the proposed Bank Customer Churn Prediction Model is anticipated to empower banking institutions with actionable insights to minimize customer attrition, optimize customer retention strategies, and enhance overall business performance in a highly competitive financial landscape.

The ML Pipeline to be followed:

1. Importing the necessary libraries and the dataset
2. Performing Data Preprocessing (Exploratory Data Analysis and Data Manipulation)
3. Modelling using Logistic Regression, KNN, and Random Forest
4. Performing Prediction
5. Visualization between Actual and predicted Values