Topic:

Bank Customer Churn Prediction Model

Dataset:
https://www.kaggle.com/datasets/gauravtopre/bank-customer-churn-dataset

Summary:

The project aims to address the issue of bank customer churn, where customers cease their business with a bank or switch to another one. Churn has significant implications for a bank's customer base and revenue. Predicting churn is crucial for banks to proactively manage customer relationships and reduce revenue loss.

Central Thesis/Questions:

1. **Primary Question:** Can machine learning algorithms accurately predict bank customer churn based on historical customer data?
2. **Thesis:** By leveraging various customer attributes, it is possible to build a predictive model that helps banks identify customers at risk of churning.
3. **Sub-questions:**
    - Which customer attributes have the most significant impact on churn prediction?
    - Which machine learning algorithm(s) provide the most accurate predictions for customer churn?
    - How can these predictions inform targeted retention strategies for at-risk customers?

Methods Planned:

1. **Data Collection:** Utilize the provided dataset from Kaggle, which includes features like credit score, age, tenure, balance, product subscriptions, and churn status.
2. **Data Preprocessing:** Cleanse the data by handling missing values, encoding categorical variables, and scaling numerical features to prepare it for machine learning algorithms.
3. **Model Selection:** Experiment with various machine learning algorithms such as Logistic Regression, Random Forest, Gradient Boosting, and Neural Networks to determine the most suitable for this prediction task.
4. **Model Training:** Split the dataset into training and testing sets. Train the selected models on the training data to learn the patterns and relationships.
5. **Model Evaluation:** Evaluate model performance using metrics like accuracy, precision, recall, and F1-score. Choose the model with the best performance for predictions.

6. **Predictions and Interpretation:** Use the trained model to predict churn for new or unseen data. Interpret the results to identify at-risk customers.

By following these steps, the project aims to develop an effective machine learning solution for predicting bank customer churn and implementing proactive measures to retain valuable customers.

Submitted by:

Group 7

Shreyash Mehta

Ronil Sanjay Surve

Shubham Sharma

Namrata Rath