Project Title: Customer Churn Prediction

Project Overview:

The project aims to develop a predictive model to forecast customer churn, helping our company proactively address customer retention efforts. Customer churn, or attrition, is a critical metric for businesses, and predicting it can significantly impact customer retention strategies.

Project Phases:

1. Data Collection:

We began by obtaining a dataset that contains customer information and their churn status.

The dataset includes fields such as CustomerID, Age, Gender, Location, Subscription Length (Months), Monthly Bill, Total Usage (GB), and Churn (target variable).

2. Data Exploration and Analysis:

We conducted an initial analysis to understand the dataset's characteristics and distribution.

We identified potential features for building the predictive model.

3. Data Preprocessing:

We split the dataset into training and testing sets to facilitate model training and evaluation.

We recognized the need for data preprocessing, including data standardization and one-hot encoding.

Issues related to column names and categorical variables ('Gender' and 'Location') were identified.

4. Model Building:

We attempted to build a Random Forest Classifier to predict customer churn.

Challenges related to data preprocessing were encountered during model training.

5. Streamlit App Integration:

We provided a Streamlit app code for deploying the model for real-time predictions.

Due to data preprocessing issues, the app faced challenges.

6. Data Transformation:

We addressed the data preprocessing issues by encoding categorical variables ('Gender' and 'Location') appropriately.

We successfully transformed the data for model training.

7. Model Training (Again):

We reattempted model training with the improved data preprocessing.

The model was trained on the dataset to predict customer churn effectively.

8. Model Performance Metrics and Visualizations:

We calculated key evaluation metrics, including accuracy, precision, recall, F1-score, and ROC AUC, to assess model performance.

We generated visualizations such as confusion matrices, classification reports, and ROC curves for better understanding.

Future Steps:

The project has reached the stage where a predictive model for customer churn has been developed and successfully integrated into a Streamlit app. Further steps might include ongoing model monitoring, retraining, and additional feature engineering to enhance model accuracy.

Conclusion:

The project has made significant progress in addressing the challenge of customer churn prediction. The primary difficulties encountered during data preprocessing have been resolved, and the model now demonstrates promising performance metrics. It is recommended to continue refining the model and the Streamlit app for potential deployment in a production environment.

By B.Phanindra Sai