# Customer Churn Prediction Report

## Objective:

The objective of this project is to develop and deploy a predictive model that can accurately predict customer churn in a telecom company. Customer churn refers to the phenomenon where customers switch to competitors or terminate their services. Aimed to build a machine learning model that predicts whether a customer is likely to churn based on various features.

## Approach:

### 1. Data Preprocessing:

- Loaded the dataset containing customer information.
- Checked for missing values and data types to ensure data quality.
- Created new features like 'Age_Monthly_Bill_Interact', 'Avg_Bill_By_Location', and 'Usage_Per_Month' to capture potentially relevant patterns in the data.
- Converted categorical features like 'Gender' and 'Location' into numeric representations.
- Dropped the 'Name' column as it was unnecessary for modelling.

### 2. Model Selection and Training:

- Split the data into training and testing sets using a 80-20 split.
- Trained a Random Forest model and evaluated its accuracy on the test set.
- Utilized classification report and cross-validation to assess model performance.

### 3. Alternative Model:

- Explored using a XGBoost as an alternative to Random Forest Classifier.
- Trained an initial XGBoost model and evaluated its accuracy on the test set.
- Conducted hyperparameter tuning for the XGBoost model using RandomizedSearchCV to improve its performance.
- Trained an optimized XGBoost model using the best hyperparameters.

### 4. Hyperparameter Tuning:

- Conducted hyperparameter tuning for the RandomForest model using RandomizedSearchCV.
- Identified the best hyperparameters for the model.

**5. Model Persistence:**

Saved the optimized Random Forest model using the joblib library for future use.

**6. New Customer Prediction:**

Demonstrated how to load the saved model and make predictions on new customer data.

Provided a code snippet to preprocess new customer data and obtain churn predictions.

**Conclusion:**

Successfully built and optimized predictive models using XGBoost and Random Forest algorithms to predict customer churn. The models demonstrated the potential to aid in reducing customer churn rates by identifying customers at risk of churning. The report provides a comprehensive overview of our approach, from data preprocessing to model evaluation, highlighting the importance of feature engineering, hyperparameter tuning, and model persistence.