

Customer Churn Prediction Using Machine Learning Algorithm

1) Introduction

- The rise of subscription-based models in consumer-oriented companies has made customer retention a key factor for sustainable growth. predictive modeling offers companies the ability to foresee potential churn and take preventive measures, ensuring customer loyalty and reduced attrition.

2) Problem Statement

- Customer churn is one of the most significant challenges faced by companies. Losing customers not only reduces revenue but also increases the cost of acquiring new customers. Traditional methods of identifying churn are reactive and inefficient, leading to late interventions.

3) Dataset And Its Attributes

- Dataset: - European Credit Card User dataset from Microset Azure Sample Dataset also available on Kaggle.
- Attributes: - Surname, CreditScore, Geography, Gender, Age, Tenure, Balance, NumOfProducts, HasCrCard, IsActiveMember, EstimatedSalary, Exited

4) Machine Learning Models: -

- Random Forest, Gradient Boosting (e.g., XGBoost, LightGBM), Neural Networks.

5) Exploratory Data Analysis (EDA): -

- i) Handle missing data with imputation techniques (e.g., mean, median for numerical data, mode for categorical).
- ii) Normalize numerical features (e.g., MinMaxScaler or StandardScaler).
- iii) Encode categorical variables using one-hot encoding or label encoding.
- iv) Balance the dataset using techniques like SMOTE (Synthetic Minority Oversampling Technique).
- v) Split data into training, validation, and test sets (e.g., 70-15-15 split).

6) QAM (Quality Assurance Modality)

- Precision, Recall, F1-Score: To balance false positives and false negatives. AUC-ROC Curve: To assess the models discriminative ability. Log Loss: For probability-based evaluation. Validation: Use cross-validation techniques to ensure robustness.

Project Synopsis

Submitted By: - Ayushi Dadhich, Marmik Parashar, Shashwat Mishra Vansh Sachdeva

7) System Requirements

a) **Hardware:** -

- Processor: - Intel i5 8th Generation or AMD Ryzen 5
- Main Memory: - 8GB DDR4 Ram
- Secondary Memory: - 100GigaBytes Solid State Drive
- Graphical Processing Unit: - WhiskeyLake-U GT2 [UHD Graphics 620]

b) **Software:** -

- Operating System
 - (1) Linux
 - (2) Windows 11 Home
- Text Editor
 - (1) VsCode
 - (2) Anaconda Jupyter
- Data Visualization Tools
 - (1) PowerBi
 - (2) Spreadsheet Viewer
- Web Browser
 - (1) Google Chrome Or Web Browser