

CUSTOMER CHURN PREDICTION

Data Preparation

- **Data Loading and Exploration:** The notebook starts by loading the dataset and performing exploratory data analysis (EDA). It involves checking for missing values, understanding data distributions, and visualizing relationships between variables using plots.

Feature Engineering

- **Feature Selection and Encoding:** The dataset features are processed, and categorical variables are encoded. Features are selected based on their importance or correlation with the target variable (churn).

Model Building

- **Logistic Regression Model:** A logistic regression model is trained on the processed data to predict customer churn.
- **Model Evaluation:** The model's performance is evaluated using metrics like accuracy, precision, recall, F1-score, and ROC AUC.

Visualization

- **Actual vs. Predicted Churn:** The notebook includes a plot showing the relationship between actual and predicted churn values, helping visualize model performance.

Results

- **Model Metrics:**
 - **Accuracy:** 78.75%
 - **Precision:** 62.06%
 - **Recall:** 51.60%
 - **F1-Score:** 56.35%
 - **ROC AUC:** 83.21%

Conclusion

The logistic regression model demonstrates decent predictive performance with an ROC AUC of 83.21%, indicating a good balance between sensitivity and specificity in predicting customer churn.