

CUSTOMER CHURN PREDICTION FOR SYRIATEL

- NAME : BRENDA CHEMUTAI
- DATE: JUNE 2025

Problem Statement

- High customer churn leads to revenue loss and increased acquisition costs.
- Identifying potential churners allows the company to take preventive action

Objective

- Develop a machine learning model for Churn classification or not.
- Evaluate model performance using ROC curve and classification metrics
- Use Logistic Regression ,Decision Tree and Linear Regression

Dataset overview

- Source :Syriatel Telecom dataset
- The target variable is churn

Data Preprocessing

- Used Label Encoding to encode categorical variable
- Dropped non informative columns like phone number
- Split data to training and testing sets
- Standardized numerical features using StandardScaler

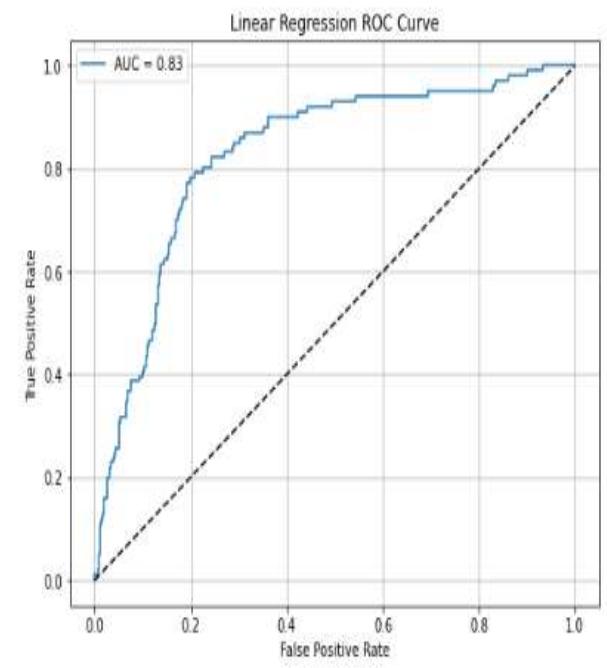
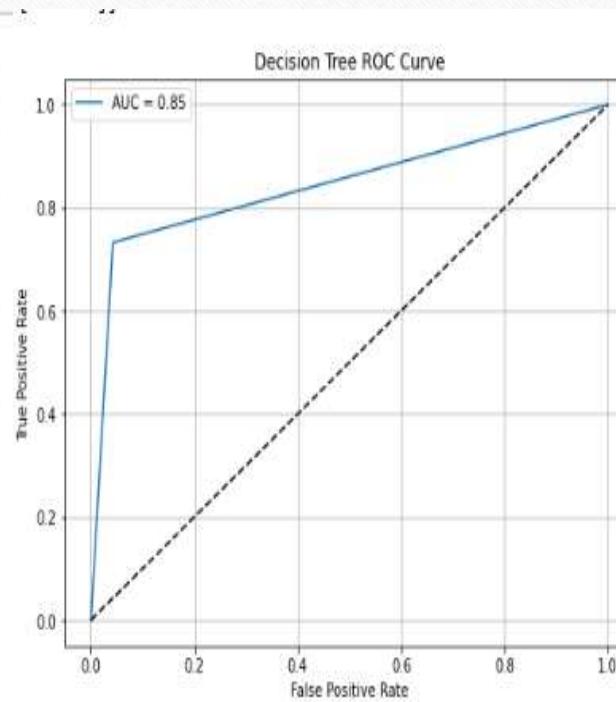
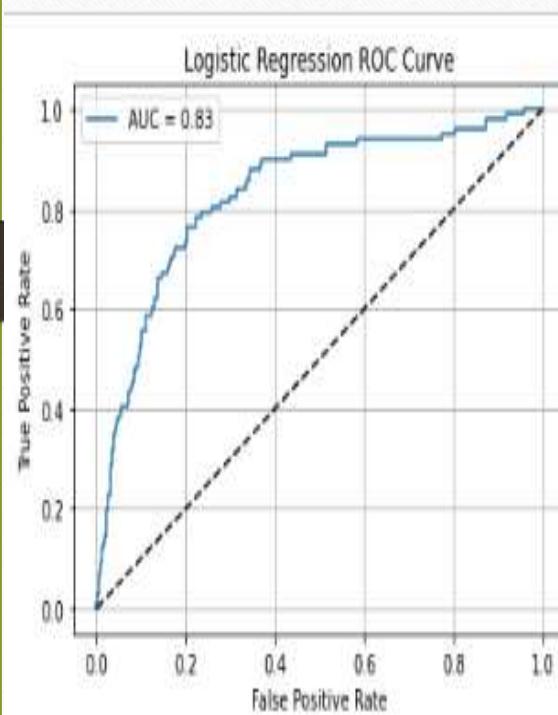
Model used

- Logistic Regression
- Decision Tree
- Linear Regression

Model Evaluation Metrics

- Recall: Measures how the model identifies all actual positive.
- Precision : It measures the correctness of positive predictions.
- Accuracy : Measures overall correctness.
- F1 Score : Harmonic mean of precision and recall.
- ROC Curve : discrimination ability across thresholds

ROC Curves



Graph interpretation

- Logistic regression ROC Curve – smooth and curved. It outputs probabilities that allow for fine grained threshold tuning to generate a smooth curve.
- Decision Tree ROC Curve – step-like and less smooth. Outputs labels or fewer probability levels leading to fewer distinct threshold.

Findings

- Logistic Regression and Decision Tree model has showed good prediction on churn's dataset making it suitable .also interprets feature importance via coefficients.
- It helps high risk customers and enables targeted retention strategies.
- ROC Curves validates model effectiveness
- Linear regression is not ideal for churn's dataset.

Conclusion & Recommendations

- Logistics Regression is suitable for Churn prediction
- I recommend these models for customer retention strategies
- Adopt the Logistic Regression model and Decision Tree for churn prediction due to its strong performance and interpretability.
- Monitor key features identified by the models and use them to actively engage high-risk customers

THANK YOU

- CONTACTS
- Email : brendachemuu19@gmail.com
- Phone number : 0704515760