TELECOM CUSTOMER CHURN

Project Overview: Predicting Telecom Customer Churn

Objective: The objective of this project was to develop and validate a predictive model for predicting customer churn in a telecom company. The project involved various stages including data preprocessing, feature engineering, exploratory data analysis, model selection, evaluation, and deployment using Python and key data science libraries. The goal was to build an accurate and efficient predictive model to identify customers at risk of churn.

Problem Statement: Customer churn is a critical issue for telecom companies, as losing customers can significantly impact revenue and profitability. The objective of this project was to predict which customers are likely to churn based on their historical behavior and interactions with the company.

Tools & Technologies:

- Python: Primary programming language used for data analysis and machine learning.
- Pandas & NumPy: Utilized for data manipulation and preprocessing.
- Scikit-Learn: Used for building and evaluating machine learning models.
- Matplotlib & Seaborn: Employed for data visualization and exploratory data analysis.
- Flask: Deployed the model as an interactive web application.

Data Preparation:

- Data Cleaning: Handled missing values and inconsistencies in the dataset, such as filling missing values and encoding categorical variables.
- Feature Engineering: Created new features to capture relevant information, such as customer tenure, contract type, and payment method.
- Encoding Categorical Variables: Converted categorical variables into numerical format using techniques like one-hot encoding.

Exploratory Data Analysis (EDA):

- Visualized distributions and relationships between features using histograms, box plots, and heatmaps.

- Identified key factors influencing customer churn, such as contract type, payment method, and monthly charges.

Model Development:

- Algorithm Selection: Evaluated multiple machine learning algorithms including logistic regression, decision trees, random forest, and gradient boosting.

- Model Tuning: Used cross-validation and grid search to optimize hyperparameters and improve model performance.

- Evaluation Metrics: Assessed models using accuracy, precision, recall, and F1-score to ensure balanced and robust performance.

Results:

- Performance: The final model achieved an accuracy of X% on the test set, indicating its effectiveness in predicting customer churn.

Key Insights:

- Contract Type: Customers with month-to-month contracts were more likely to churn compared to those with long-term contracts.

- Payment Method: Customers using electronic check as their payment method had higher churn rates.

- Monthly Charges: Higher monthly charges were associated with higher churn rates.

Deployment:

- Interactive Web Application: Deployed the model using Flask, allowing users to input customer details and get real-time churn predictions.

- User Interface: Designed a simple and intuitive interface for users to enter details such as contract type, payment method, and monthly charges, and obtain predictions.

Key Learnings:

- Data Preprocessing: Importance of handling missing data and encoding categorical variables for building accurate predictive models.

- Model Evaluation: Understanding the trade-offs between different performance metrics and selecting the best model based on balanced performance.

- Deployment: Practical experience in deploying a machine learning model as an interactive web application for real-world use.

****GitHub Repository****:

Explore the complete project and code here:

https://github.com/sravanthi224/Quanta-2.git