Customer Churn Prediction – Machine Learning Project

The **Customer Churn Prediction** project focuses on building a machine learning solution to identify customers who are likely to discontinue a company's services. Customer churn is a critical business challenge, as retaining existing customers is often more cost-effective than acquiring new ones. The objective of this project was to analyze historical customer data and predict churn behavior, enabling businesses to take proactive steps to improve customer retention.

The project began with collecting and understanding a structured customer dataset containing demographic information, service usage details, billing history, and customer support interactions. Data preprocessing was a key phase, where missing values were handled appropriately, categorical variables were encoded using suitable techniques, and numerical features were scaled to ensure consistency across the dataset. This step helped improve model performance and ensured data quality.

Exploratory Data Analysis (EDA) was performed to uncover patterns and relationships within the data. Visualizations and statistical analysis were used to identify factors influencing churn, such as customer tenure, monthly charges, contract type, service subscriptions, and frequency of support requests. These insights provided a strong foundation for model selection and feature engineering.

Multiple machine learning classification algorithms were trained and evaluated, including **Logistic Regression, Decision Tree, Random Forest, and K-Nearest Neighbors (KNN)**. Each model was assessed using standard evaluation metrics such as accuracy, precision, recall, F1-score, and ROC-AUC to ensure balanced performance and minimize false predictions. Comparative analysis helped in selecting the best-performing model based on predictive accuracy and business relevance.

Feature importance analysis was conducted to identify the most influential variables contributing to customer churn. The results highlighted key churn drivers and provided actionable insights that businesses can use to design targeted retention strategies, such as personalized offers or improved customer support.

Overall, the project demonstrates the practical application of machine learning in solving real-world business problems and showcases skills in data preprocessing, exploratory analysis, model building, evaluation, and interpretation of results to support data-driven decision-making.

Tools & Technologies:
Python, Pandas, NumPy, Scikit-learn, Matplotlib, Seaborn, VS Code