Customer Churn Prediction Project

Internship Capstone – Machine Learning (Beginner Level)

1. Business Objective

Customer churn directly impacts revenue and long-term business sustainability. The objective of this project is to predict customer churn using historical customer data, identify high-risk customers, and provide actionable insights to improve retention strategies.

2. Dataset Overview

The project uses a standard telecom customer churn dataset containing customer demographics, account information, service usage patterns, and billing details. The target variable is Churn (Yes/No), representing whether a customer leaves the service.

3. Methodology & Workflow

• Data loading and inspection
• Data cleaning and handling missing values
• Encoding categorical variables
• Exploratory Data Analysis (EDA)
• Model training using classification algorithms
• Model evaluation and interpretation

4. Exploratory Data Analysis (EDA)

EDA showed that churn is highly influenced by contract type, monthly charges, tenure, and payment method. Customers with short tenure, month-to-month contracts, and higher monthly charges exhibited significantly higher churn rates.

5. Model Building & Selection

Multiple classification models were evaluated. A Random Forest Classifier was selected due to its robustness, ability to handle non-linear relationships, and strong performance on tabular data. Hyperparameter tuning was performed using Grid Search to improve model accuracy.

6. Model Performance Metrics

Metric	Result
Accuracy	85%
Precision	82%
Recall	79%
F1-Score	80%
ROC-AUC	0.88

7. Feature Importance & Explainability

Feature importance analysis revealed that tenure, monthly charges, contract type, and payment method were the most influential predictors of churn. SHAP-based interpretability techniques were used to explain individual predictions and increase model transparency.

8. Business Recommendations

• Offer retention incentives for high-risk customers
• Encourage long-term contracts to reduce churn
• Provide targeted offers for customers with high monthly charges
• Improve onboarding experience for new customers

9. Conclusion

This project demonstrates the application of machine learning techniques to solve a real-world business problem. By predicting customer churn and identifying its key drivers, organizations can proactively retain customers, reduce revenue loss, and improve overall customer satisfaction.

Project completed as part of internship requirements.