Churn Prediction System
Machine Learning Task 2 – Future Interns

1. Problem Statement
Customer churn occurs when users stop using a service. Retaining customers is more cost-effective than acquiring new ones.
This project aims to build a machine learning-based system to predict customer churn and help businesses take proactive actions.

2. Dataset Description
Dataset Used: Telco Customer Churn Dataset (Kaggle).
The dataset includes customer demographics, services, contract details, and billing information.
The target variable is Churn (Yes/No).

3. Data Preprocessing & Feature Engineering
Data cleaning involved handling missing values, converting categorical variables into numerical format,
and encoding the target variable. Feature engineering was performed to improve model performance.

4. Machine Learning Models Used
Logistic Regression, Random Forest, and XGBoost classifiers were trained and evaluated.
XGBoost delivered the best performance.

5. Model Evaluation Metrics
Accuracy, Precision, Recall, F1-score, ROC-AUC score, and Confusion Matrix were used to evaluate the models.

6. Key Insights & Findings
Month-to-month contract customers have higher churn. High monthly charges increase churn probability.
Short-tenure customers are more likely to churn.

7. Business Recommendations
Target high-risk customers with retention offers, encourage long-term contracts,
and provide personalized plans to reduce churn.

8. Conclusion
This project demonstrates the effective use of machine learning for churn prediction and business decision support.