Customer Churn Analysis - Summary Report

1. Data Loading & Exploration

- Dataset loaded using pandas
- Used head(), info(), describe(), isnull().sum()
- Columns like customerID, gender, tenure, MonthlyCharges, Churn reviewed

2. Data Cleaning

- Converted TotalCharges to numeric
- Handled missing values using median
- Dropped unnecessary columns like customerID

3. Data Visualization

- Pie charts for churn distribution
- Histograms, countplots, boxplots
- Relationships between churn and features like Contract, PaymentMethod, etc.
- 10+ visualizations using seaborn and matplotlib

4. Feature Engineering

- Encoded categorical features using pd.get_dummies()
- Created train-test split using train_test_split

5. Model Training

- Trained models: Logistic Regression, Random Forest, XGBoost, SVM
- Accuracy scores printed before each classification report

6. Model Evaluation

- ROC curve, Precision-Recall curve
- Model accuracy comparison bar plot
- GridSearchCV for hyperparameter tuning

Customer Churn Analysis - Summary Report

- SHAP analysis for model explainability

7. Report Generation

- Final accuracy summary
- Report written to .txt and .pdf
- Model saved as .pkl using joblib