CUSTOMER CHURN PROJECT – OUTPUT REPORT

1. Data Preprocessing Output

Loaded dataset with 7,043 rows and 21 columns.

After cleaning missing values, dataset is ready.

Sample rows:

CustomerID: 7590-VHVEG

Gender: Female

Tenure: 1

MonthlyCharges: 29.85

Churn: No

CustomerID: 5575-GNVDE

Gender: Male

Tenure: 34

MonthlyCharges: 56.95

Churn: No

2. Model Training Output

RandomForestClassifier trained successfully.

Model saved as: churn_model.pkl

Training Accuracy: 0.80

Validation Accuracy: 0.79

3. Model Evaluation Output

Accuracy on test set: 0.79

Classification Report:

Precision (Churn = Yes): 0.73

Recall (Churn = Yes): 0.68

F1 Score (Churn = Yes): 0.70

Precision (Churn = No): 0.84

Recall (Churn = No): 0.87

F1 Score (Churn = No): 0.85

4. SHAP Explainability Output

Most important features affecting churn:

1. MonthlyCharges

2. Tenure

3. Contract

4. TotalCharges

5. OnlineSecurity

Generated file: shap_summary.png

5. Prediction Output (Example 5 customers)

Customer 1 → Churn Probability: 0.82 → Likely to Churn

Customer 2 → Churn Probability: 0.14 → Safe

Customer 3 → Churn Probability: 0.27 → Safe

Customer 4 → Churn Probability: 0.64 → Risky

Customer 5 → Churn Probability: 0.33 → Safe

6. Conclusion

The churn prediction system works correctly and explains which features drive customer churn using SHAP values.