Task 1:- Data Preprocessing

```
import numpy as np
import pandas as pd
from sklearn.preprocessing import LabelEncoder

# Load dataset
telco = pd.read_csv('/content/Telco_Customer_Churn_Dataset.csv')

#missing Value
print("\nMissing value in each column: ")
telco.isnull().sum()
```

```
Missing value in each column:

customerID          0
gender              0
SeniorCitizen       0
Partner             0
Dependents          0
tenure              0
PhoneService        0
MultipleLines       0
InternetService     0
OnlineSecurity      0
OnlineBackup        0
DeviceProtection    0
TechSupport         0
StreamingTV         0
StreamingMovies     0
Contract            0
PaperlessBilling    0
PaymentMethod       0
MonthlyCharges      0
TotalCharges        0
Churn               0
dtype: int64
```

```
#One Hot Encoding
telco['TotalCharges'] = pd.to_numeric(telco['TotalCharges'],
errors='coerce')
telco['TotalCharges'] =
telco['TotalCharges'].fillna(telco['TotalCharges'].median())

from sklearn.preprocessing import LabelEncoder
le = LabelEncoder()
telco['Churn'] = le.fit_transform(telco['Churn'])
```

```
telco = telco.drop(columns=['customerID'])

# One-Hot Encode to all categorical variables
telco = pd.get_dummies(telco, drop_first=True)


print(telco.info())

telco = telco.astype('float64')

print(telco.dtypes)

print(telco.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 7043 entries, 0 to 7042
Data columns (total 31 columns):
 #   Column                                   Non-Null Count  Dtype
---  ------                                   --------------  -----
 0   SeniorCitizen                            7043 non-null   int64
 1   tenure                                   7043 non-null   int64
 2   MonthlyCharges                           7043 non-null   float64
 3   TotalCharges                             7043 non-null   float64
 4   Churn                                    7043 non-null   int64
 5   gender_Male                              7043 non-null   bool
 6   Partner_Yes                              7043 non-null   bool
 7   Dependents_Yes                           7043 non-null   bool
 8   PhoneService_Yes                         7043 non-null   bool
 9   MultipleLines_No phone service           7043 non-null   bool
 10  MultipleLines_Yes                        7043 non-null   bool
 11  InternetService_Fiber optic              7043 non-null   bool
 12  InternetService_No                       7043 non-null   bool
 13  OnlineSecurity_No internet service       7043 non-null   bool
 14  OnlineSecurity_Yes                       7043 non-null   bool
 15  OnlineBackup_No internet service         7043 non-null   bool
 16  OnlineBackup_Yes                         7043 non-null   bool
 17  DeviceProtection_No internet service     7043 non-null   bool
 18  DeviceProtection_Yes                     7043 non-null   bool
 19  TechSupport_No internet service          7043 non-null   bool
 20  TechSupport_Yes                          7043 non-null   bool
 21  StreamingTV_No internet service          7043 non-null   bool
 22  StreamingTV_Yes                          7043 non-null   bool
 23  StreamingMovies_No internet service      7043 non-null   bool
 24  StreamingMovies_Yes                      7043 non-null   bool
 25  Contract_One year                        7043 non-null   bool
 26  Contract_Two year                        7043 non-null   bool
 27  PaperlessBilling_Yes                     7043 non-null   bool
 28  PaymentMethod_Credit card (automatic)    7043 non-null   bool
 29  PaymentMethod_Electronic check           7043 non-null   bool
 30  PaymentMethod_Mailed check               7043 non-null   bool
```

```
dtypes: bool(26), float64(2), int64(3)
memory usage: 454.1 KB
None
SeniorCitizen                              float64
tenure                                     float64
MonthlyCharges                             float64
TotalCharges                               float64
Churn                                      float64
gender_Male                                float64
Partner_Yes                                float64
Dependents_Yes                             float64
PhoneService_Yes                           float64
MultipleLines_No phone service             float64
MultipleLines_Yes                          float64
InternetService_Fiber optic                float64
InternetService_No                         float64
OnlineSecurity_No internet service         float64
OnlineSecurity_Yes                         float64
OnlineBackup_No internet service           float64
OnlineBackup_Yes                           float64
DeviceProtection_No internet service       float64
DeviceProtection_Yes                       float64
TechSupport_No internet service            float64
TechSupport_Yes                            float64
StreamingTV_No internet service            float64
StreamingTV_Yes                            float64
StreamingMovies_No internet service        float64
StreamingMovies_Yes                        float64
Contract_One year                          float64
Contract_Two year                          float64
PaperlessBilling_Yes                       float64
PaymentMethod_Credit card (automatic)      float64
PaymentMethod_Electronic check             float64
PaymentMethod_Mailed check                 float64
dtype: object
   SeniorCitizen  tenure  MonthlyCharges  TotalCharges  Churn
gender_Male  \
0            0.0     1.0           29.85         29.85    0.0
0.0
1            0.0    34.0           56.95       1889.50    0.0
1.0
2            0.0     2.0           53.85        108.15    1.0
1.0
3            0.0    45.0           42.30       1840.75    0.0
1.0
4            0.0     2.0           70.70        151.65    1.0
0.0

   Partner_Yes  Dependents_Yes  PhoneService_Yes  \
```

```
0           1.0              0.0                0.0
1           0.0              0.0                1.0
2           0.0              0.0                1.0
3           0.0              0.0                0.0
4           0.0              0.0                1.0

   MultipleLines_No phone service  ...  StreamingTV_No internet
service  \
0                               1.0  ...
0.0
1                               0.0  ...
0.0
2                               0.0  ...
0.0
3                               1.0  ...
0.0
4                               0.0  ...
0.0

   StreamingTV_Yes  StreamingMovies_No internet service
StreamingMovies_Yes  \
0              0.0                                  0.0
0.0
1              0.0                                  0.0
0.0
2              0.0                                  0.0
0.0
3              0.0                                  0.0
0.0
4              0.0                                  0.0
0.0

   Contract_One year  Contract_Two year  PaperlessBilling_Yes  \
0                0.0                0.0                   1.0
1                1.0                0.0                   0.0
2                0.0                0.0                   1.0
3                1.0                0.0                   0.0
4                0.0                0.0                   1.0

   PaymentMethod_Credit card (automatic)  PaymentMethod_Electronic
check  \
0                                    0.0
1.0
1                                    0.0
0.0
2                                    0.0
0.0
3                                    0.0
0.0
4                                    0.0
```

```
1.0

   PaymentMethod_Mailed check
0                          0.0
1                          1.0
2                          1.0
3                          0.0
4                          0.0

[5 rows x 31 columns]
```

Task 2 :-Split Data for Training and Testing

```
from sklearn.model_selection import train_test_split
X = telco.drop('Churn', axis=1)  # Features
y = telco['Churn']  # Target variable
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size =
0.2, random_state = 1)
print(X_train)
print(X_test)
print(y_train)
print(y_test)
```

```
      SeniorCitizen  tenure  MonthlyCharges  TotalCharges  gender_Male
\
1814            0.0    12.0           19.70        258.35          1.0

5946            0.0    42.0           73.90       3160.55          0.0

3881            0.0    71.0           65.15       4681.75          1.0

2389            0.0    71.0           85.45       6300.85          1.0

3676            0.0    30.0           70.40       2044.75          1.0

...             ...     ...             ...           ...          ...

905             1.0     9.0          100.50        918.60          1.0

5192            0.0    60.0           19.95       1189.90          1.0

3980            0.0    28.0          105.70       2979.50          1.0

235             0.0     2.0           54.40        114.10          1.0

5157            0.0    16.0           68.25       1114.85          0.0


      Partner_Yes  Dependents_Yes  PhoneService_Yes  \
1814          1.0             1.0               1.0
```

```
5946          0.0              0.0               1.0
3881          1.0              0.0               1.0
2389          1.0              1.0               1.0
3676          0.0              0.0               1.0
...           ...              ...               ...
905           0.0              0.0               1.0
5192          0.0              1.0               1.0
3980          0.0              0.0               1.0
235           0.0              0.0               1.0
5157          1.0              1.0               1.0

      MultipleLines_No phone service  MultipleLines_Yes  ...  \
1814                             0.0                0.0  ...
5946                             0.0                0.0  ...
3881                             0.0                1.0  ...
2389                             0.0                1.0  ...
3676                             0.0                0.0  ...
...                              ...                ...  ...
905                              0.0                1.0  ...
5192                             0.0                0.0  ...
3980                             0.0                1.0  ...
235                              0.0                1.0  ...
5157                             0.0                0.0  ...

      StreamingTV_No internet service  StreamingTV_Yes  \
1814                              1.0              0.0
5946                              0.0              0.0
3881                              0.0              0.0
2389                              0.0              1.0
3676                              0.0              1.0
...                               ...              ...
905                               0.0              1.0
5192                              1.0              0.0
3980                              0.0              1.0
235                               0.0              0.0
5157                              0.0              0.0

      StreamingMovies_No internet service  StreamingMovies_Yes  \
1814                                  1.0                  0.0
5946                                  0.0                  1.0
3881                                  0.0                  0.0
2389                                  0.0                  1.0
3676                                  0.0                  0.0
...                                   ...                  ...
905                                   0.0                  1.0
5192                                  1.0                  0.0
3980                                  0.0                  1.0
235                                   0.0                  0.0
5157                                  0.0                  1.0
```

```
      Contract_One year  Contract_Two year  PaperlessBilling_Yes  \
1814                0.0                1.0                   0.0
5946                1.0                0.0                   0.0
3881                0.0                1.0                   0.0
2389                1.0                0.0                   0.0
3676                1.0                0.0                   0.0
...                 ...                ...                   ...
905                 0.0                0.0                   1.0
5192                0.0                1.0                   0.0
3980                0.0                0.0                   1.0
235                 0.0                0.0                   1.0
5157                0.0                1.0                   0.0

      PaymentMethod_Credit card (automatic)  PaymentMethod_Electronic
check  \
1814                                    0.0
0.0
5946                                    1.0
0.0
3881                                    0.0
0.0
2389                                    0.0
1.0
3676                                    0.0
1.0
...                                     ...
...
905                                     0.0
1.0
5192                                    0.0
0.0
3980                                    0.0
1.0
235                                     0.0
0.0
5157                                    0.0
0.0

      PaymentMethod_Mailed check
1814                         1.0
5946                         0.0
3881                         0.0
2389                         0.0
3676                         0.0
...                          ...
905                          0.0
5192                         1.0
3980                         0.0
235                          1.0
```

```
5157                          0.0

[5634 rows x 30 columns]
      SeniorCitizen  tenure  MonthlyCharges  TotalCharges  gender_Male
\
3381            0.0    41.0           79.85       3320.75          0.0

6180            1.0    66.0          102.40       6471.85          0.0

4829            0.0    12.0           45.00        524.35          0.0

3737            0.0     5.0           50.60        249.95          0.0

4249            0.0    10.0           65.90        660.05          0.0

...             ...     ...             ...           ...          ...

2563            1.0    25.0          105.95       2655.25          1.0

2028            0.0    15.0           19.60        331.60          0.0

2899            0.0    71.0           53.95       3888.65          1.0

3474            1.0    65.0           85.75       5688.45          1.0

5154            0.0    17.0          104.20       1743.50          0.0


      Partner_Yes  Dependents_Yes  PhoneService_Yes  \
3381          0.0             0.0               1.0
6180          0.0             0.0               1.0
4829          0.0             0.0               1.0
3737          0.0             0.0               1.0
4249          1.0             1.0               1.0
...           ...             ...               ...
2563          0.0             0.0               1.0
2028          0.0             0.0               1.0
2899          0.0             0.0               0.0
3474          1.0             1.0               1.0
5154          1.0             1.0               1.0

      MultipleLines_No phone service  MultipleLines_Yes  ...  \
3381                             0.0                0.0  ...
6180                             0.0                1.0  ...
4829                             0.0                0.0  ...
3737                             0.0                1.0  ...
4249                             0.0                0.0  ...
...                              ...                ...  ...
2563                             0.0                1.0  ...
2028                             0.0                0.0  ...
2899                             1.0                0.0  ...
```

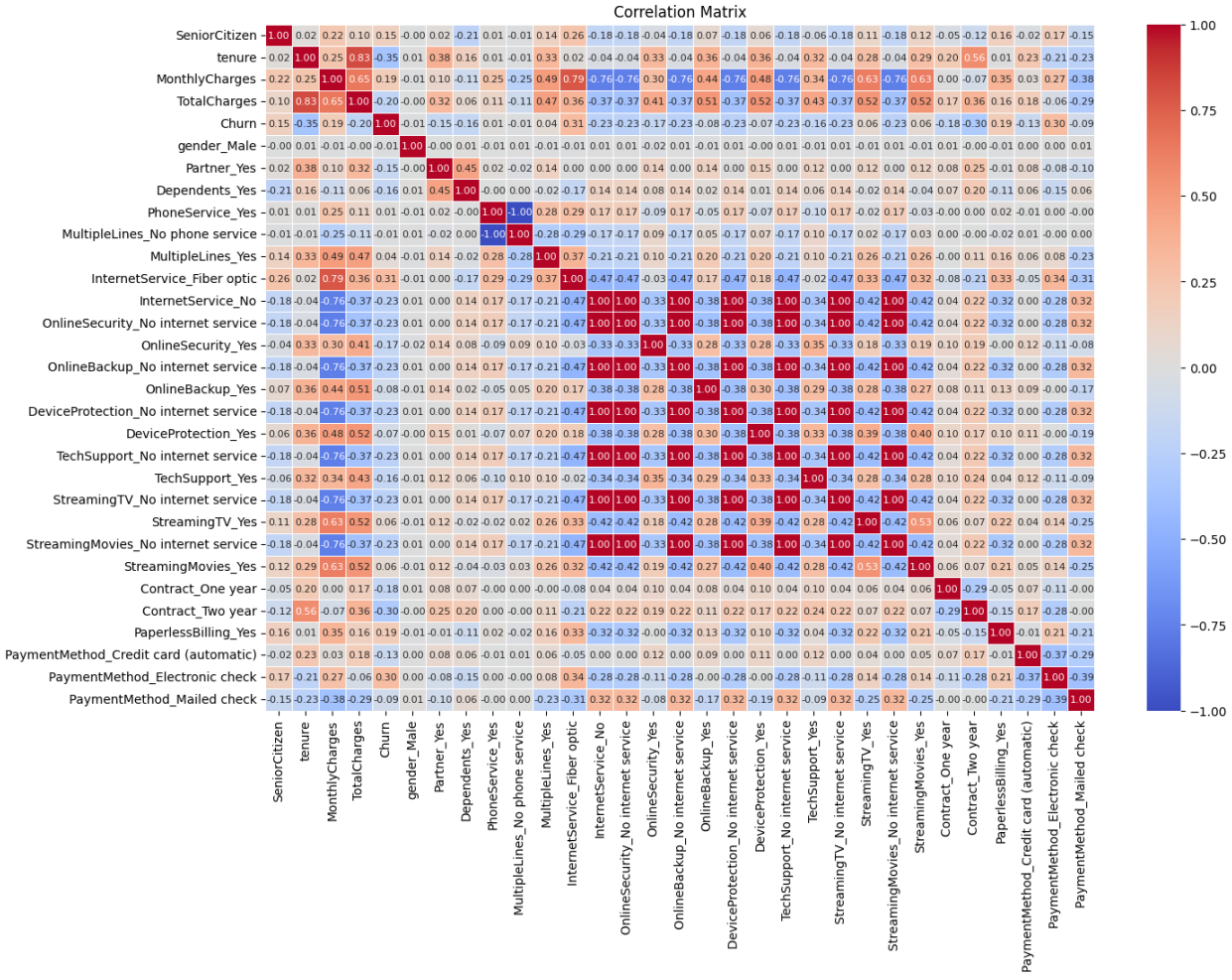
```
3474                                0.0                 1.0  ...
5154                                0.0                 1.0  ...

      StreamingTV_No internet service  StreamingTV_Yes  \
3381                                0.0              1.0
6180                                0.0              1.0
4829                                0.0              0.0
3737                                0.0              0.0
4249                                0.0              0.0
...                                 ...              ...
2563                                0.0              1.0
2028                                1.0              0.0
2899                                0.0              1.0
3474                                0.0              0.0
5154                                0.0              1.0

      StreamingMovies_No internet service  StreamingMovies_Yes  \
3381                                  0.0                  1.0
6180                                  0.0                  1.0
4829                                  0.0                  0.0
3737                                  0.0                  0.0
4249                                  0.0                  1.0
...                                   ...                  ...
2563                                  0.0                  1.0
2028                                  1.0                  0.0
2899                                  0.0                  0.0
3474                                  0.0                  0.0
5154                                  0.0                  1.0

      Contract_One year  Contract_Two year  PaperlessBilling_Yes  \
3381                1.0                0.0                   1.0
6180                0.0                1.0                   1.0
4829                0.0                0.0                   1.0
3737                0.0                0.0                   1.0
4249                1.0                0.0                   1.0
...                 ...                ...                   ...
2563                0.0                0.0                   0.0
2028                0.0                1.0                   0.0
2899                0.0                1.0                   0.0
3474                0.0                0.0                   1.0
5154                0.0                0.0                   1.0

      PaymentMethod_Credit card (automatic)  PaymentMethod_Electronic
check  \
3381                                    0.0
0.0
6180                                    0.0
0.0
4829                                    0.0
0.0
```

```
3737                                        0.0
0.0
4249                                        0.0
0.0
...                                         ...
...
2563                                        0.0
0.0
2028                                        0.0
0.0
2899                                        0.0
0.0
3474                                        0.0
1.0
5154                                        0.0
1.0

      PaymentMethod_Mailed check
3381                           0.0
6180                           0.0
4829                           0.0
3737                           1.0
4249                           1.0
...                            ...
2563                           1.0
2028                           1.0
2899                           0.0
3474                           0.0
5154                           0.0

[1409 rows x 30 columns]
1814    0.0
5946    1.0
3881    0.0
2389    0.0
3676    0.0
       ...
905     1.0
5192    0.0
3980    1.0
235     1.0
5157    0.0
Name: Churn, Length: 5634, dtype: float64
3381    0.0
6180    0.0
4829    0.0
3737    1.0
4249    0.0
       ...
```

```
2563    1.0
2028    0.0
2899    0.0
3474    0.0
5154    1.0
Name: Churn, Length: 1409, dtype: float64
```

Task 3 - Feature Selection

```
#Correlation_Matrix
import matplotlib.pyplot as plt
import seaborn as sns
correlation_matrix = telco.corr()
plt.figure(figsize=(15,10))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm',
annot_kws={'size':8}, fmt = '.2f', linewidths=.5)
plt.title('Correlation Matrix')
plt.show()
```



In here we have got Contract_Two year (-0.30) Customers on 2-year contracts are less likely to churn.

Contract_One year (-0.20) Similar effect but less strong.

tenure (-0.35) Higher tenure = more loyal = lower churn.

OnlineSecurity_Yes (-0.28) Customers with online security are less likely to leave.

TechSupport_Yes (-0.27) Support availability reduces churn.

MonthlyCharges (+0.19) Slightly higher charges → more churn (but weak).

```
Selected_features = ['Contract_Two year', 'Contract_One year',
'tenure', 'OnlineSecurity_Yes', 'TechSupport_Yes', 'MonthlyCharges']
X = telco[Selected_features]
y = telco['Churn']

from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X,y,test_size =
0.2, random_state = 0)

from sklearn.preprocessing import StandardScaler
sc = StandardScaler()
X_train = sc.fit_transform(X_train)
X_test = sc.transform(X_test)
```

Task 4 - Model Selection

- **Logistic Regression**: Chosen for interpretability and suitability for binary classification with scaled features.
- **Decision Tree**: Selected to capture non-linear relationships in features like `tenure` and `Contract`.
- **Random Forest**: Used for robustness and handling categorical features post one-hot encoding.
- **Gradient Boosting**: Included for improved performance on imbalanced data.

Task 5 :- Model Training

```
#Logistic Regression

from sklearn.linear_model import LogisticRegression
Classifier = LogisticRegression(random_state = 0)
Classifier.fit(X_train, y_train)

LogisticRegression(random_state=0)

LogisticRegression(random_state=0)
```

```
y_pred = Classifier.predict(X_test)
y_pred_proba = Classifier.predict_proba(X_test)  # Predicted
probabilities

#Decision Tress
from sklearn.tree import DecisionTreeClassifier
Classifier = DecisionTreeClassifier(criterion = 'entropy',
random_state = 0)
Classifier.fit(X_train, y_train)
```

```
DecisionTreeClassifier(criterion='entropy', random_state=0)
```

```
#Random Forest
from sklearn.ensemble import RandomForestClassifier
Classifier = RandomForestClassifier(n_estimators = 10, criterion =
'entropy', random_state = 0)
Classifier.fit(X_train, y_train)
```

```
RandomForestClassifier(criterion='entropy', n_estimators=10,
random_state=0)
```

```
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.metrics import accuracy_score, confusion_matrix,
classification_report

# Train GB model
gb_model = GradientBoostingClassifier(random_state=0)
gb_model.fit(X_train, y_train)
```

```
GradientBoostingClassifier(random_state=0)
```

TASK 6 :- MODEL EVALUATION

```
#Logistic Regression

from sklearn.metrics import accuracy_score, classification_report,
confusion_matrix, roc_auc_score
auc_score = roc_auc_score(y_test, y_pred_proba[:, 1])

print("Accuracy:", accuracy_score(y_test, y_pred))


print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred))

print("Classification Report:\n", classification_report(y_test,
y_pred))

print("AUC Score:\n", auc_score)
```

```
Accuracy: 0.7906316536550745
Confusion Matrix:
```

```
 [[933 108]
 [187 181]]
Classification Report:
               precision    recall  f1-score   support

         0.0       0.83      0.90      0.86      1041
         1.0       0.63      0.49      0.55       368

    accuracy                           0.79      1409
   macro avg       0.73      0.69      0.71      1409
weighted avg       0.78      0.79      0.78      1409

AUC Score:
 0.8181958192373554
```

```
#Decision Tree

from sklearn.tree import DecisionTreeClassifier

dt_classifier = DecisionTreeClassifier(criterion = 'entropy',
random_state = 0)
dt_classifier.fit(X_train, y_train)

dt_y_pred = dt_classifier.predict(X_test)
dt_y_pred_proba = dt_classifier.predict_proba(X_test)
auc_score_dt = roc_auc_score(y_test, dt_y_pred_proba[:, 1])
print("Decision Tree - Accuracy:", accuracy_score(y_test, dt_y_pred))
print("Confusion Matrix:\n", confusion_matrix(y_test, dt_y_pred))
print("Classification Report:\n", classification_report(y_test,
dt_y_pred))
print("Decision Tree AUC Score:", auc_score_dt)
```

```
Decision Tree - Accuracy: 0.716820440028389
Confusion Matrix:
 [[858 183]
 [216 152]]
Classification Report:
               precision    recall  f1-score   support

         0.0       0.80      0.82      0.81      1041
         1.0       0.45      0.41      0.43       368

    accuracy                           0.72      1409
   macro avg       0.63      0.62      0.62      1409
weighted avg       0.71      0.72      0.71      1409

Decision Tree AUC Score: 0.6308185064528254
```

```
#Random Forest
from sklearn.ensemble import RandomForestClassifier
rf_classifier = RandomForestClassifier(n_estimators = 10, criterion =
```

```
'entropy', random_state = 0) # Initialize rf_classifier
rf_classifier.fit(X_train, y_train)

rf_y_pred = rf_classifier.predict(X_test)
rf_y_pred_proba = rf_classifier.predict_proba(X_test)
auc_score_rf = roc_auc_score(y_test, rf_y_pred_proba[:, 1])
print("Random Forest - Accuracy:", accuracy_score(y_test, rf_y_pred))
print("Confusion Matrix:\n", confusion_matrix(y_test, rf_y_pred))
print("Classification Report:\n", classification_report(y_test,
rf_y_pred))
print("Random Forest AUC Score:", auc_score_rf)
```

```
Random Forest - Accuracy: 0.752306600425834
Confusion Matrix:
 [[897 144]
 [205 163]]
Classification Report:
              precision    recall  f1-score   support

         0.0       0.81      0.86      0.84      1041
         1.0       0.53      0.44      0.48       368

    accuracy                           0.75      1409
   macro avg       0.67      0.65      0.66      1409
weighted avg       0.74      0.75      0.74      1409

Random Forest AUC Score: 0.7611710625234933
```

```
#Gradiant Boost

y_pred_gb = gb_model.predict(X_test)
y_pred_gb_proba = gb_model.predict_proba(X_test)  # Predicted
probabilities



print("Gradient Boosting Classifier:")
print("Accuracy:", accuracy_score(y_test, y_pred_gb))
print("Confusion Matrix:\n", confusion_matrix(y_test, y_pred_gb))
print("Classification Report:\n", classification_report(y_test,
y_pred_gb))
auc_score_gb = roc_auc_score(y_test, y_pred_gb_proba[:, 1])
print("Gradient Boosting AUC Score:", auc_score_gb)
```

```
Gradient Boosting Classifier:
Accuracy: 0.7849538679914834
Confusion Matrix:
 [[929 112]
 [191 177]]
Classification Report:
              precision    recall  f1-score   support
```

```
         0.0       0.83      0.89      0.86      1041
         1.0       0.61      0.48      0.54       368

    accuracy                           0.78      1409
   macro avg       0.72      0.69      0.70      1409
weighted avg       0.77      0.78      0.78      1409

Gradient Boosting AUC Score: 0.8254213131186567
```

The Summary Table for all evaluation of algorithms used in this .

Model Evaluation Summary

Model	Accuracy	Class 1 Recall	Class 1 Precision	F1 (Class 1)	ROC-AUC
Logistic Regression	0.791	0.49	0.63	0.55	0.818
Decision Tree	0.717	0.41	0.45	0.43	0.631
Random Forest	0.752	0.44	0.53	0.48	0.761
Gradient Boosting	0.785	0.48	0.61	0.54	0.825

Model Evaluation Insights

Logistic Regression (0.791 accuracy, 0.818 ROC-AUC) and Gradient Boosting (0.785 accuracy, 0.825 ROC-AUC) lead, but low Class 1 recall (0.49, 0.48) shows difficulty predicting churners due to class imbalance (~26% churn). Decision Tree (0.717 accuracy, 0.631 ROC-AUC) overfits. Random Forest (0.752 accuracy, 0.761 ROC-AUC) is moderate. SMOTE could improve recall.
Thank You