

Syria Tel Customer Churn - A Classification Case

TASK: PHASE 3 PROJECT

BY: OCHIENG' DENNIS OUMA

Introduction

- ▶ Customer churn is a major challenge to many businesses, especially those dealing with customers who make subscriptions.
- ▶ The project seeks to solve this real-world problem through giving informed recommendations based on an effective model.
- ▶ The key stakeholder in the case of SyriaTel is the Customer Retention Department, who will need to adopt the recommendations so as to reduce the rate of customers leaving.

Business Problem

- ▶ Syriatel, a telecom provider, is experiencing customer churn.
- ▶ The retention team wants to use data to predict which customers are most likely to leave, so they can intervene with retention strategies.

Research Objectives

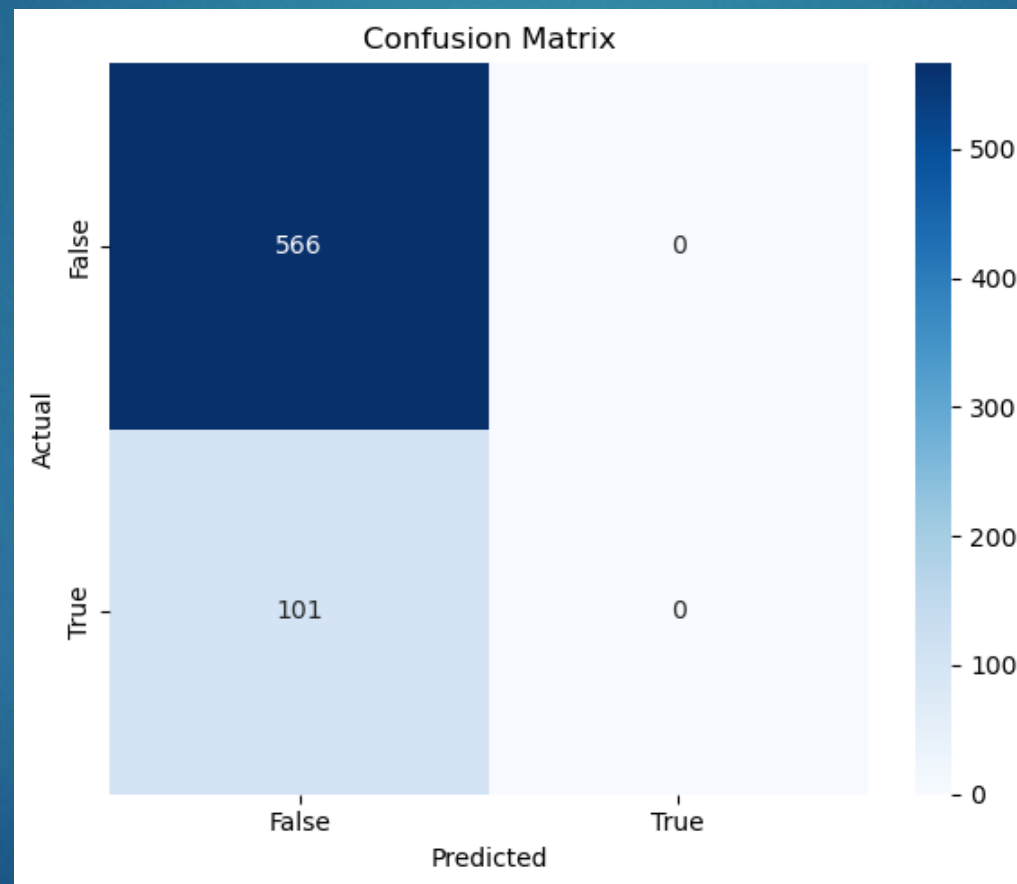
- ▶ To build a classification model to predict customer churn.
- ▶ To provide business recommendations based on model insights.

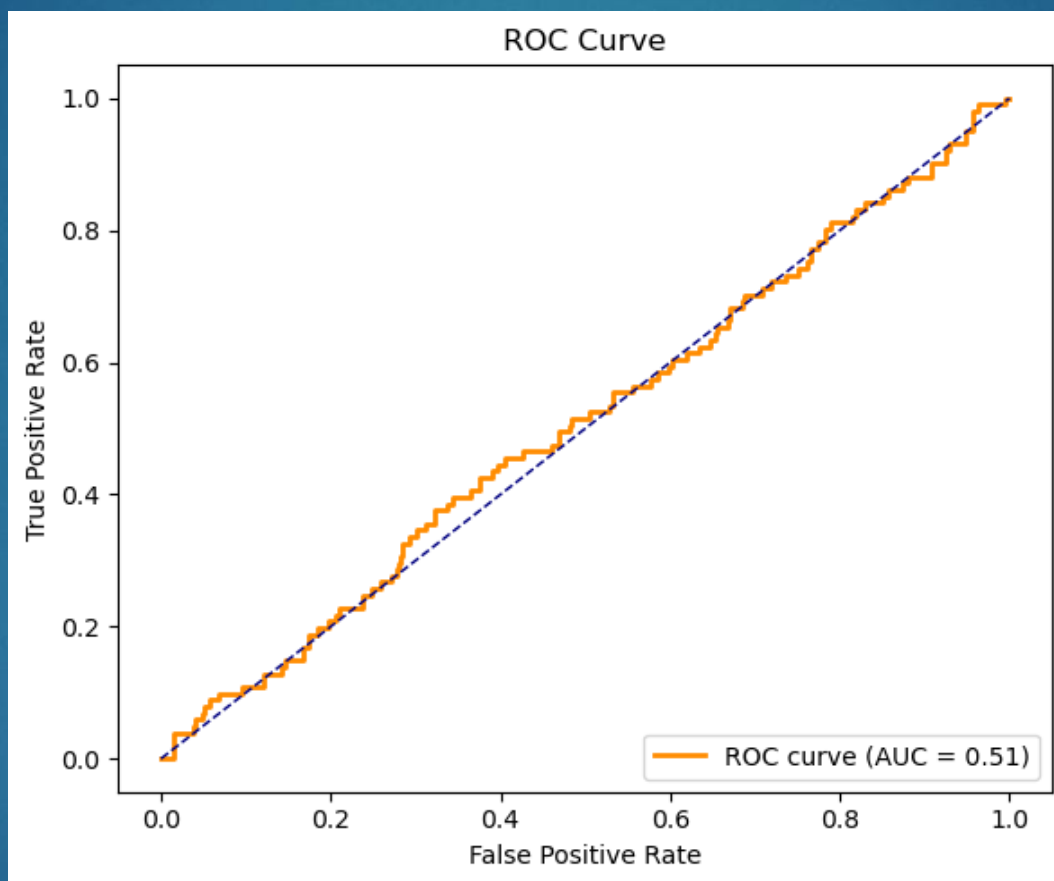
Baseline Modeling - Logistic Regression

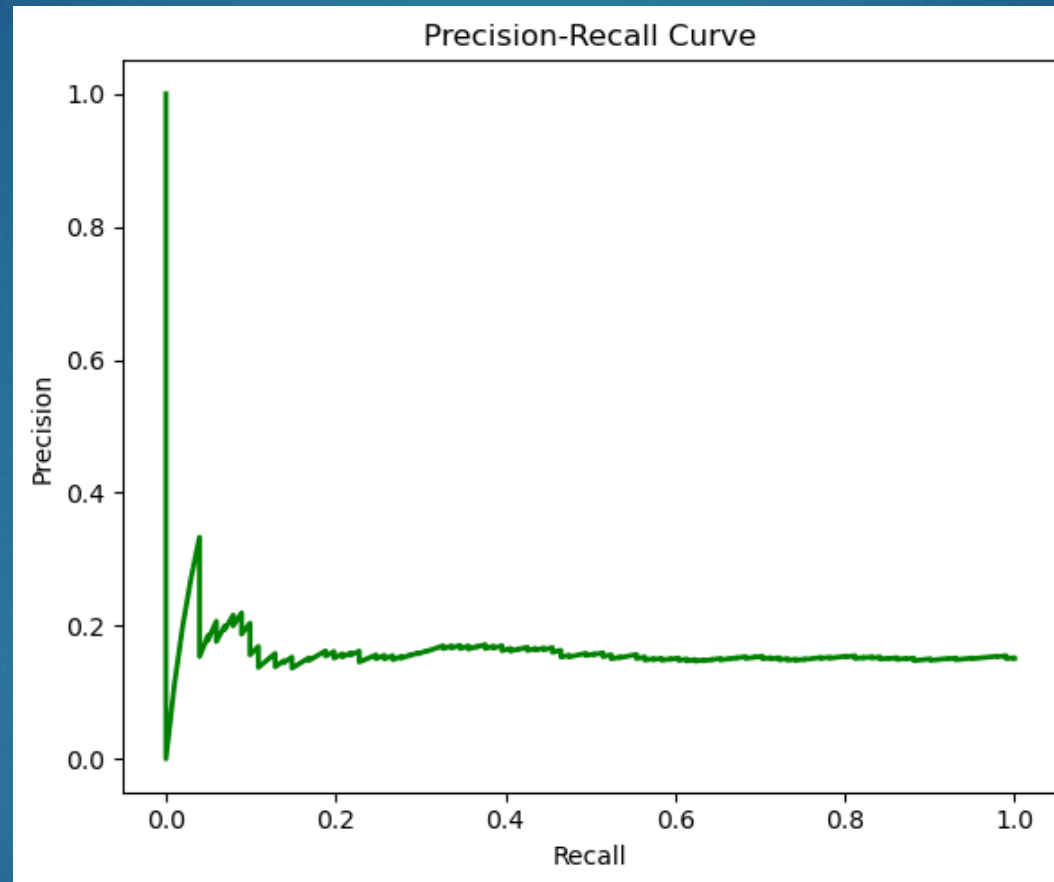
- The logistics regression returned:

	precision	recall	f1-score	support
False	0.85	1.00	0.92	566
True	0.00	0.00	0.00	101
accuracy			0.85	667
macro avg	0.42	0.50	0.46	667
weighted avg	0.72	0.85	0.78	667

ROC AUC: 0.8219046286254067



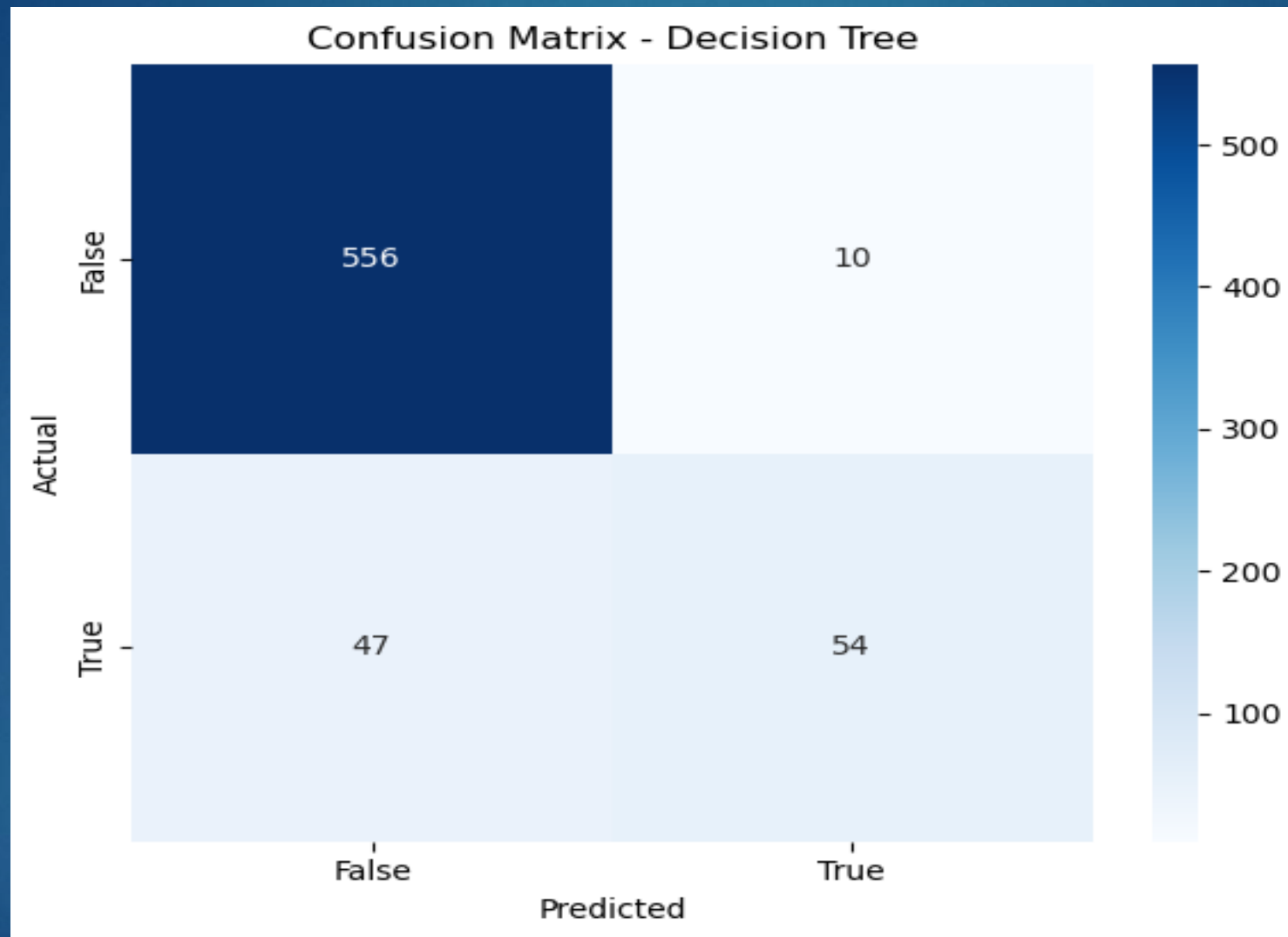




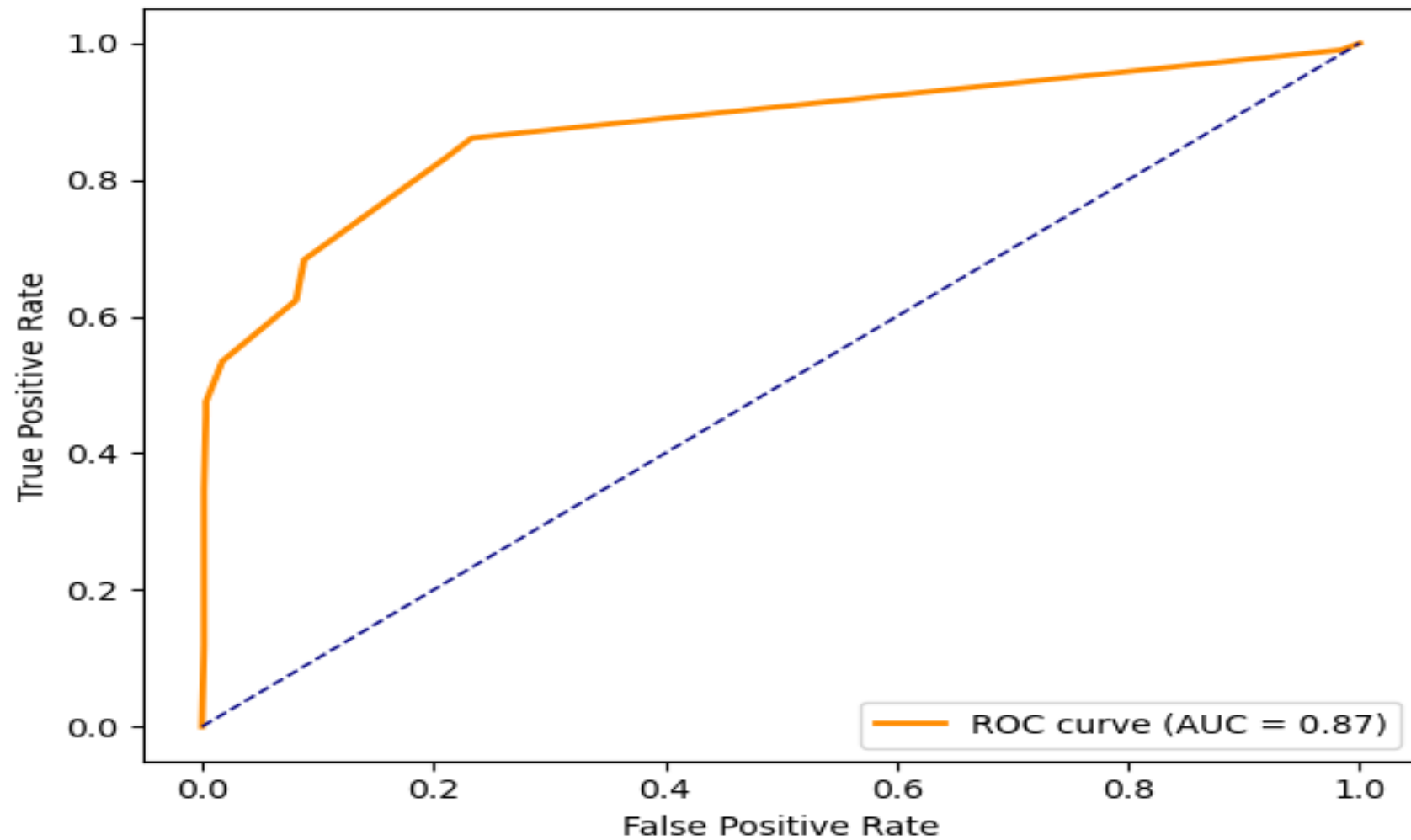
Non-Parametric Model - Decision Tree

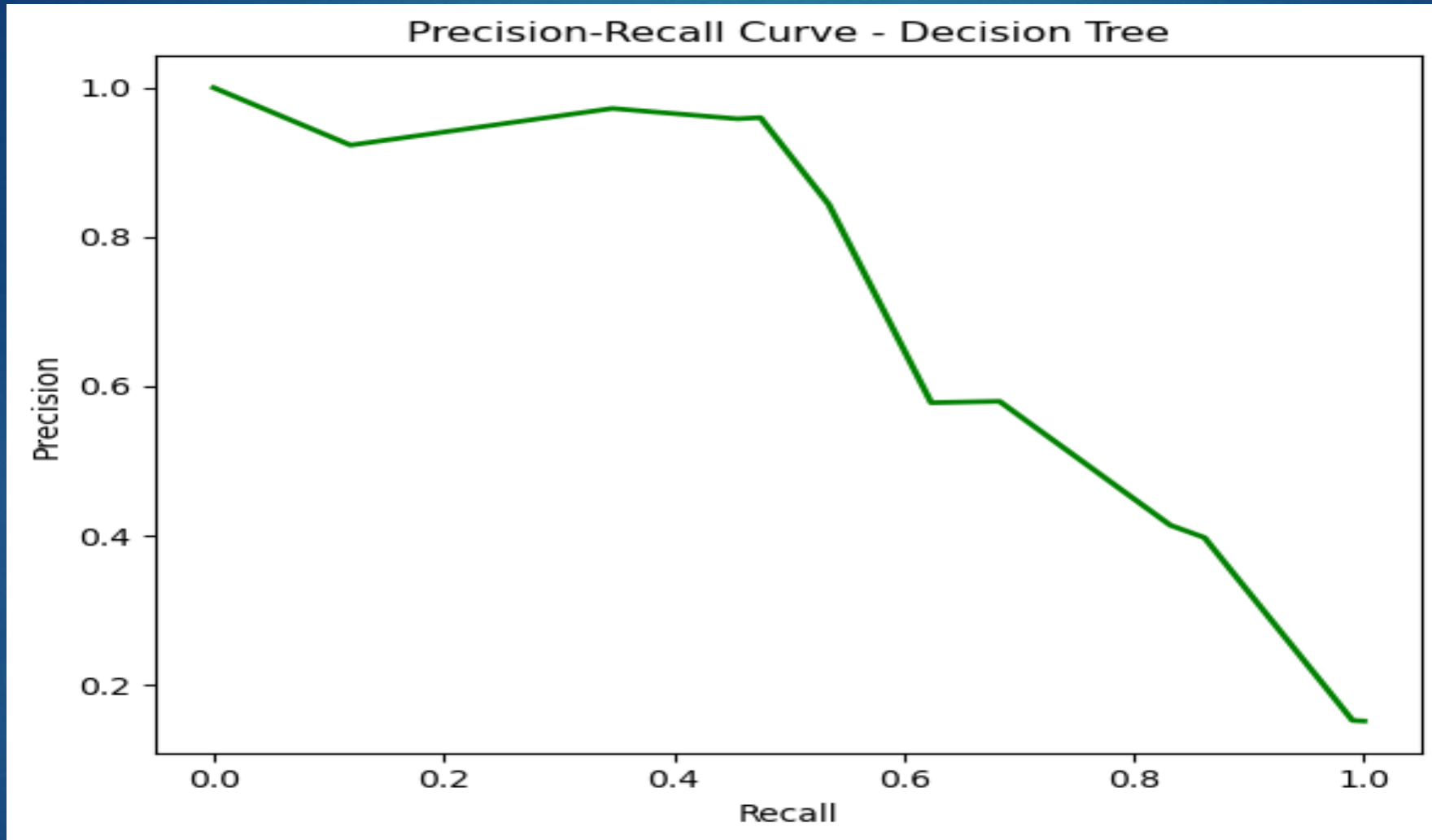
- The Decision tree model when run returned:

	precision	recall	f1-score	support
False	0.92	0.98	0.95	566
True	0.84	0.53	0.65	101
accuracy			0.91	667
macro avg	0.88	0.76	0.80	667
weighted avg	0.91	0.91	0.91	667

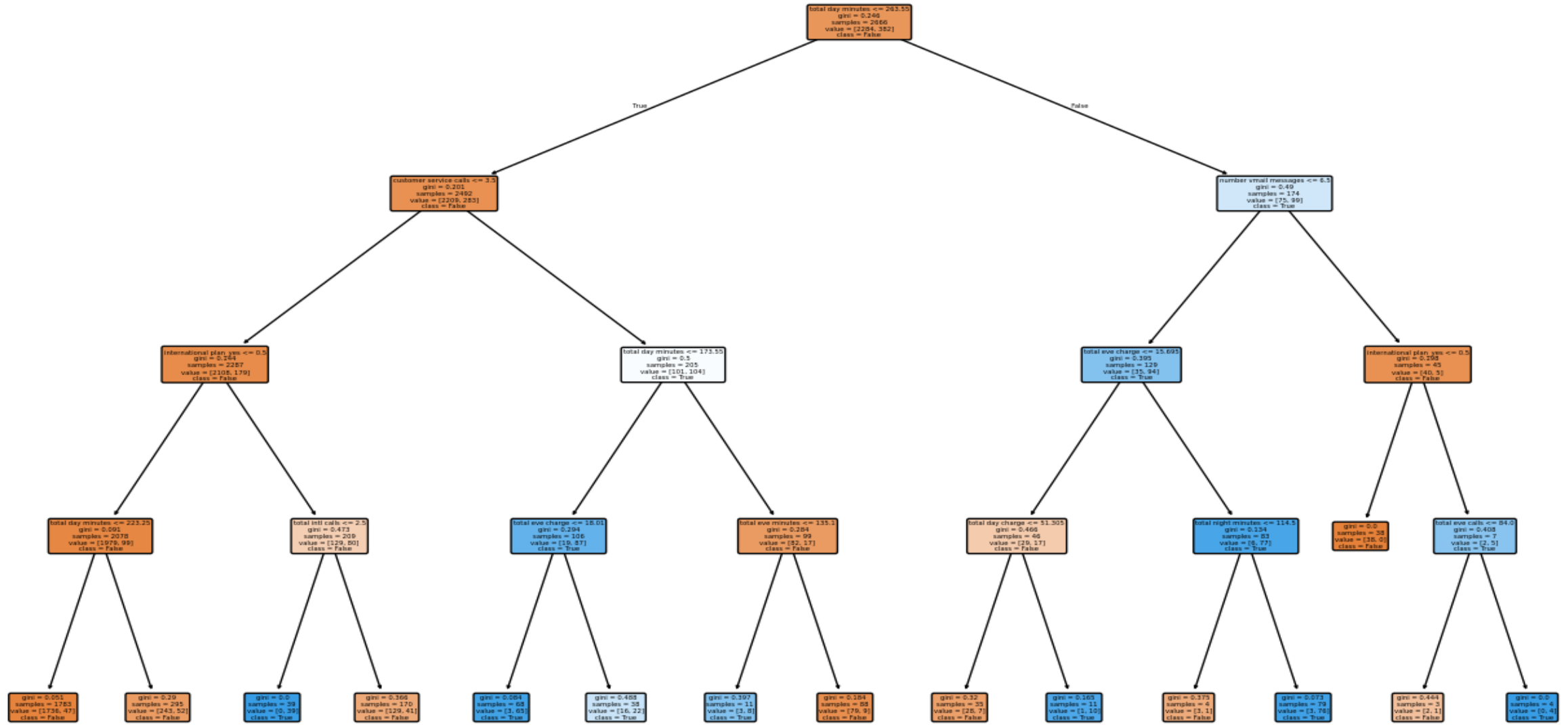


ROC Curve - Decision Tree





Decision Tree Visualization



Model Tuning - Random Forest

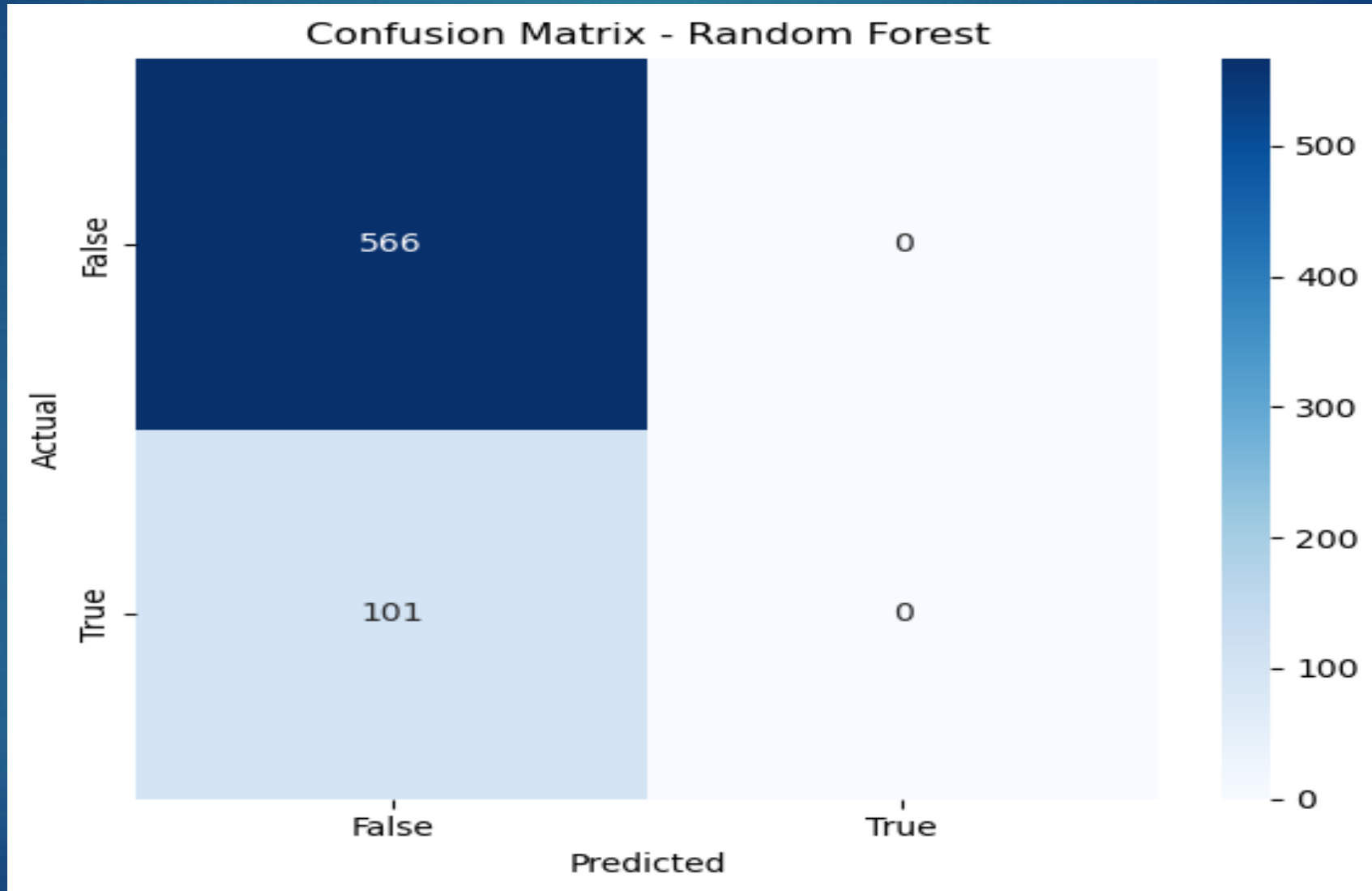
- After considering random forest and evaluating best model:

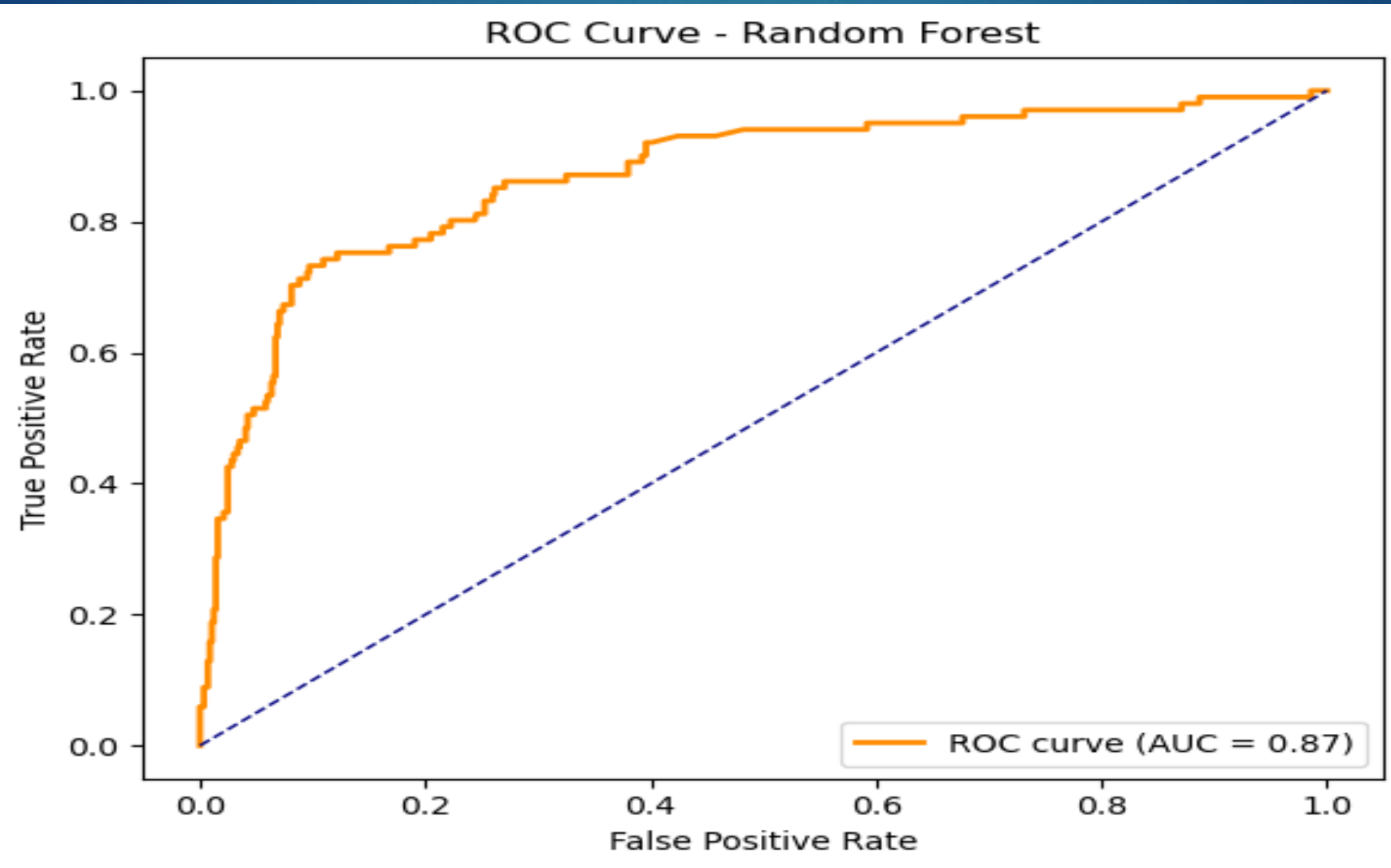
```
Best Parameters: {'max_depth': 4, 'min_samples_split': 2, 'n_estimators': 100}
```

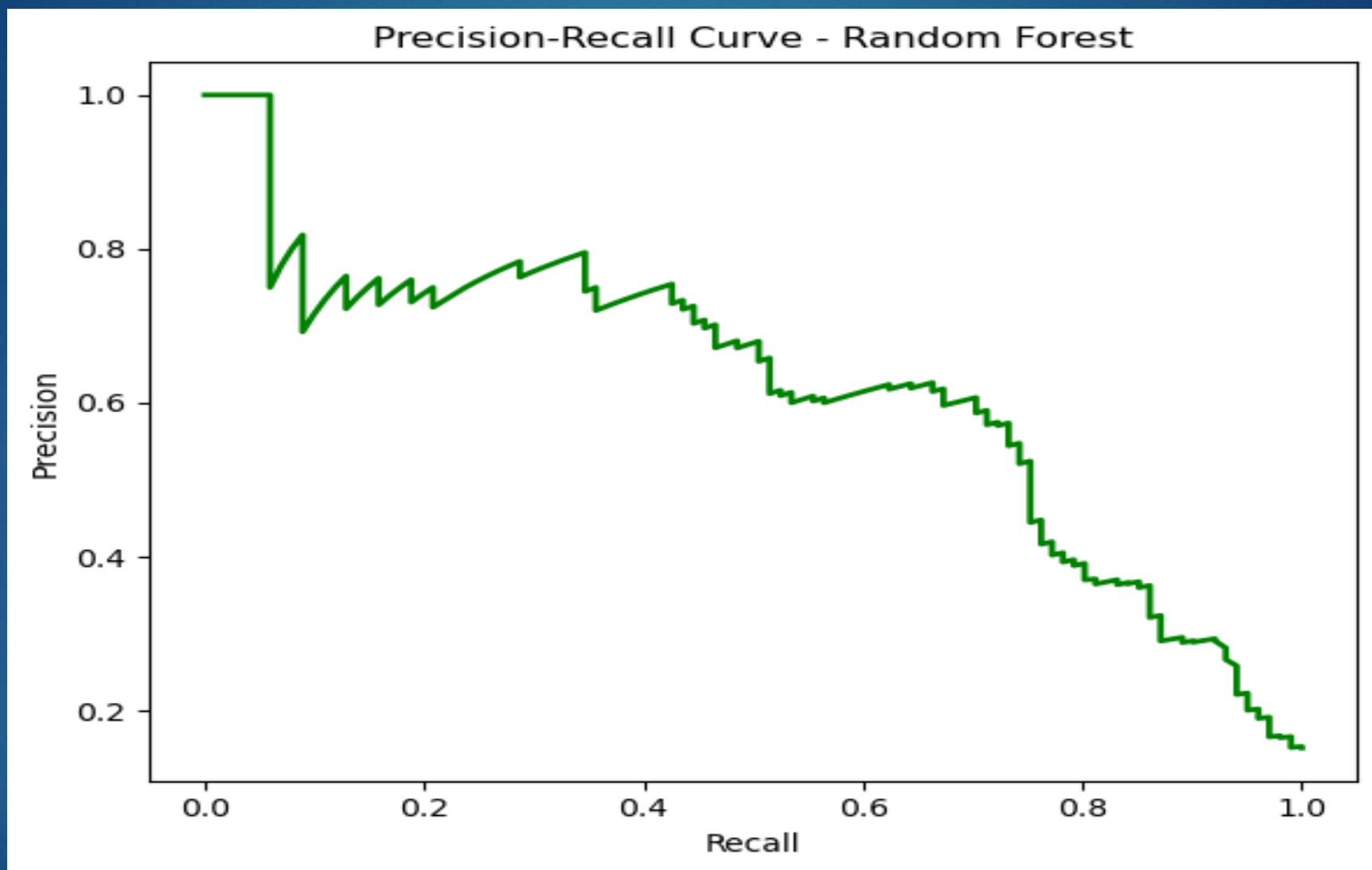
```
Best Recall Score: 0.0
```

```
Classification Report:
```

	precision	recall	f1-score	support
False	0.85	1.00	0.92	566
True	0.00	0.00	0.00	101
accuracy			0.85	667
macro avg	0.42	0.50	0.46	667
weighted avg	0.72	0.85	0.78	667







Evaluation

Final Model: Random Forest

Random Forest is picked because:-

- ▶ - Chosen for better recall (catching churners)
- ▶ - Best parameters: {...}
- ▶ - Test set recall
- ▶ - Top features: {plot feature importances}

Recommendations

- ▶ To offer loyalty plans to month-to-month customers
- ▶ To Improve customer service efficiency for those with high-tech support needs
- ▶ To explore additional features (e.g., customer satisfaction scores if available)
- ▶ To test model performance on new customer segments
- ▶ To deploy model for live churn monitoring