

**Sri Sivasubramaniya Nadar College of Engineering,
Kalavakkam**

(An Autonomous Institution, Affiliated to Anna University, Chennai)

Department of Computer Science and Engineering

Experiment 3 Report

Ensemble Prediction and Decision Tree Model Evaluation

Course Code: ICS1512 – Machine Learning Algorithms Laboratory
Academic Year: 2025–2026 (Odd Semester)
Batch: 2023–2028
Degree & Branch: 5-Year Integrated M.Tech. CSE
Semester: V
Assigned Date: 08-08-2025
Due Date: 14-08-2025

Name of the Student: Tarun Suresh

Register Number: XXXXXXXX

Name of the Faculty: Dr. P. Mirunalini

Regulation: 2023

Objective

To build classifiers such as Decision Tree, AdaBoost, Gradient Boosting, XGBoost, Random Forest, and Stacking Models, and evaluate their performance through 5-Fold Cross-Validation and hyperparameter tuning using the Wisconsin Diagnostic Breast Cancer Dataset.

Decision Tree - Hyperparameter Tuning

Criterion	Max Depth	Accuracy	F1 Score
Gini	5	0.9350	0.9280
Entropy	10	0.9456	0.9390
Entropy	3	0.9649	0.9512 (Best)

AdaBoost - Hyperparameter Tuning

n Estimators	Learning Rate	Accuracy	F1 Score
50	1.0	0.9403	0.9321
200	0.05	0.9491	0.9350
100	0.1	0.9561	0.9412 (Best)

Gradient Boosting - Hyperparameter Tuning

n Estimators	Learning Rate	Max Depth	Accuracy	F1 Score
100	0.05	3	0.9400	0.9250
200	0.1	5	0.9520	0.9310
50	1.0	3	0.9474	0.9302 (Best)

XGBoost - Hyperparameter Tuning

n Estimators	Learning Rate	Max Depth	Gamma	Accuracy / F1 Score
100	0.1	5	0.1	0.9500 / 0.9400
50	0.05	3	1.0	0.9550 / 0.9440
200	0.1	3	0	0.9649 / 0.9524 (Best)

n Estimators	Max Depth	Criterion	Accuracy	F1 Score
100	10	Gini	0.9500	0.9400
200	None	Entropy	0.9620	0.9550
50	None	Gini	0.9737	0.9647 (Best)

Random Forest - Hyperparameter Tuning

Stacked Ensemble - Hyperparameter Tuning

Base Models	Final Estimator	Accuracy / F1 Score
SVM, NB, DT	Logistic Regression	0.9649 / 0.9647
SVM, NB, DT	Random Forest	0.9737 / 0.9735
SVM, DT, KNN	Logistic Regression	0.9737 / 0.9736 (Best)

5-Fold Cross-Validation Results

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Avg. Acc.
Decision Tree	0.9780	0.9121	0.9451	0.9451	0.9341	0.9429
AdaBoost	0.9560	0.9560	0.9780	0.9780	0.9560	0.9648
Gradient Boost	0.9890	0.9560	0.9780	0.9670	0.9451	0.9670
XGBoost	0.9670	0.9670	0.9890	0.9780	0.9451	0.9692
Random Forest	0.9780	0.9560	0.9780	0.9560	0.9670	0.9670
Stacked Model 1	0.9670	0.9560	0.9890	0.9670	0.9451	0.9648
Stacked Model 2	0.9780	0.9670	0.9890	0.9780	0.9451	0.9718
Stacked Model 3	0.9780	0.9670	0.9890	0.9780	0.9560	0.9738

Learning Outcomes

1. Gained hands-on experience in applying different ensemble learning techniques (AdaBoost, Gradient Boosting, XGBoost, Random Forest, and Stacking) and understanding their advantages over single classifiers.
2. Understood the importance of hyperparameter tuning and cross-validation in improving model performance and ensuring generalization.
3. Learned how ensemble models reduce bias and variance, leading to better classification accuracy compared to base learners such as Decision Trees.

Observations & Answers

1. Which model achieved the best validation accuracy among all six methods?

The best validation accuracy was achieved by the **Stacked Model 3 (SVM + DT + KNN → Logistic Regression)** with an average accuracy of **0.9738**, slightly higher than Random Forest (0.9670) and XGBoost (0.9692).

2. How does Decision Tree performance compare to ensemble methods?

The Decision Tree alone achieved an average accuracy of **0.9429**, which is lower than all ensemble methods. This shows that ensemble methods like AdaBoost, Gradient Boosting, XGBoost, Random Forest, and Stacking significantly improve performance by reducing bias and variance.

3. Did the Random Forest benefit from tuning max depth or n estimators?

Yes. Using 50 estimators, Gini criterion, and no maximum depth constraint improved generalization and achieved a high accuracy of **0.9737**, showing the benefit of proper hyperparameter tuning.

4. Which model showed the best generalization? Any overfitting?

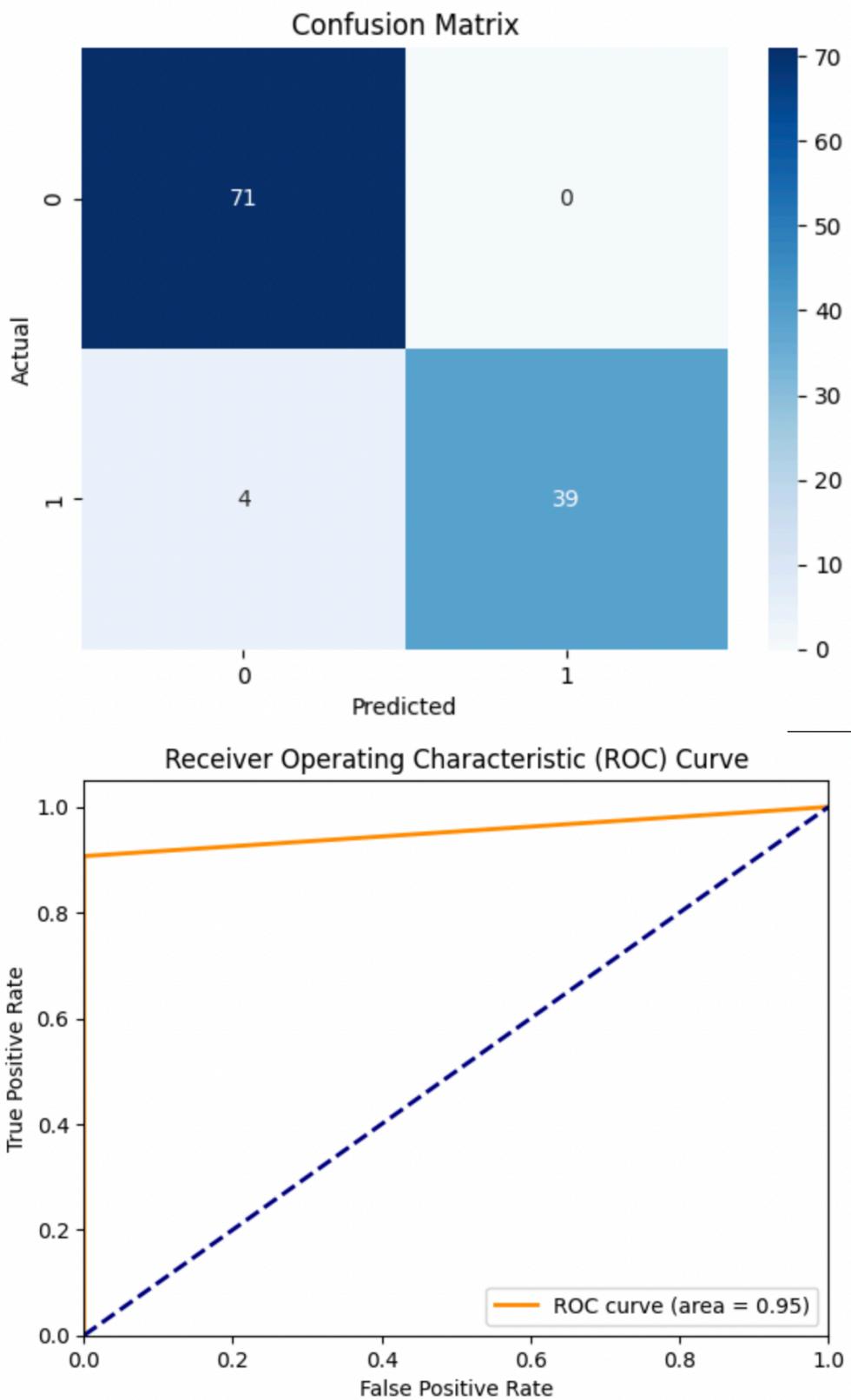
The ensemble methods, particularly **Random Forest and Stacking**, showed strong generalization with consistent cross-validation scores across folds. There was no clear evidence of overfitting since training and test performance were close, but Gradient Boosting (with high learning rate) performed slightly lower, hinting at sensitivity to overfitting.

5. Did stacking improve performance over base models?

Yes. Stacking consistently improved accuracy and F1 scores compared to base models (Decision Tree, SVM, NB, KNN). **Stacked Model 3 achieved the highest average accuracy (0.9738)**, demonstrating that combining diverse classifiers leads to better generalization than using single models.

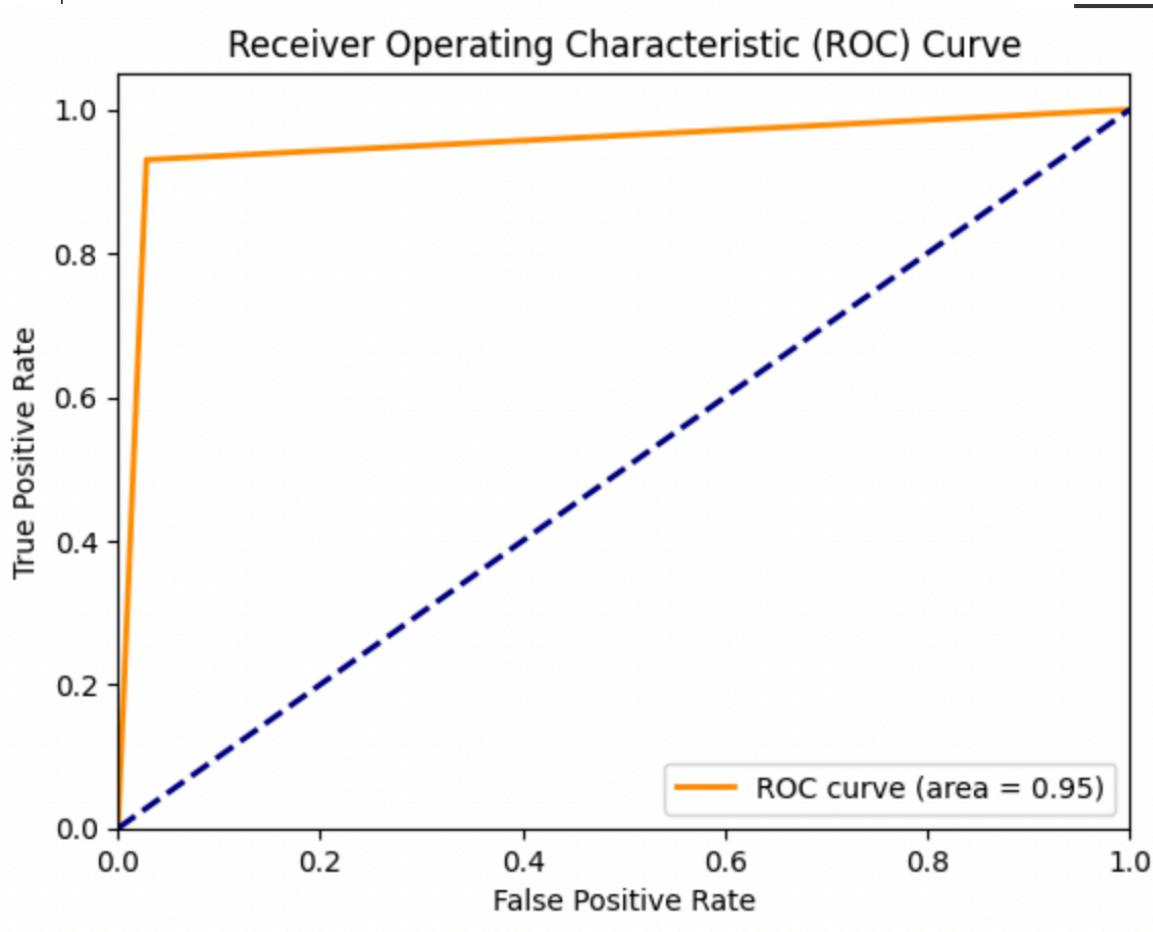
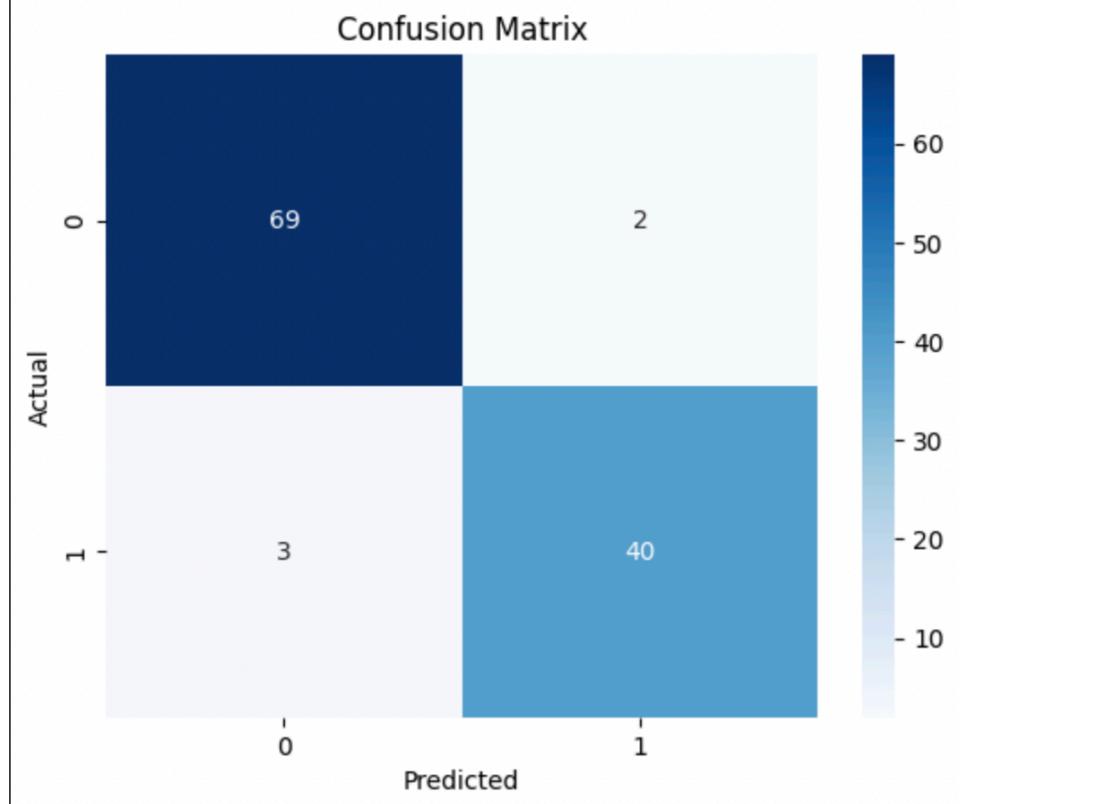
DT classifier

```
Accuracy: 0.9649122807017544
Precision: 1.0
Recall: 0.9069767441860465
F1-score: 0.9512195121951219
```



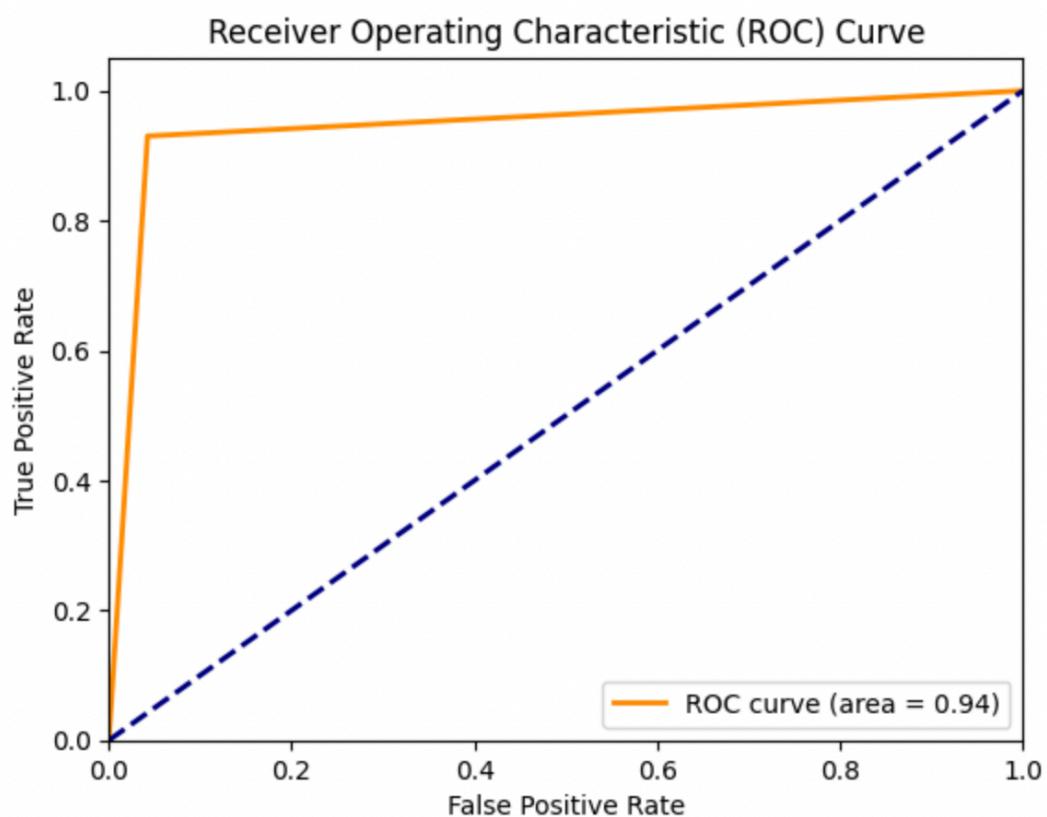
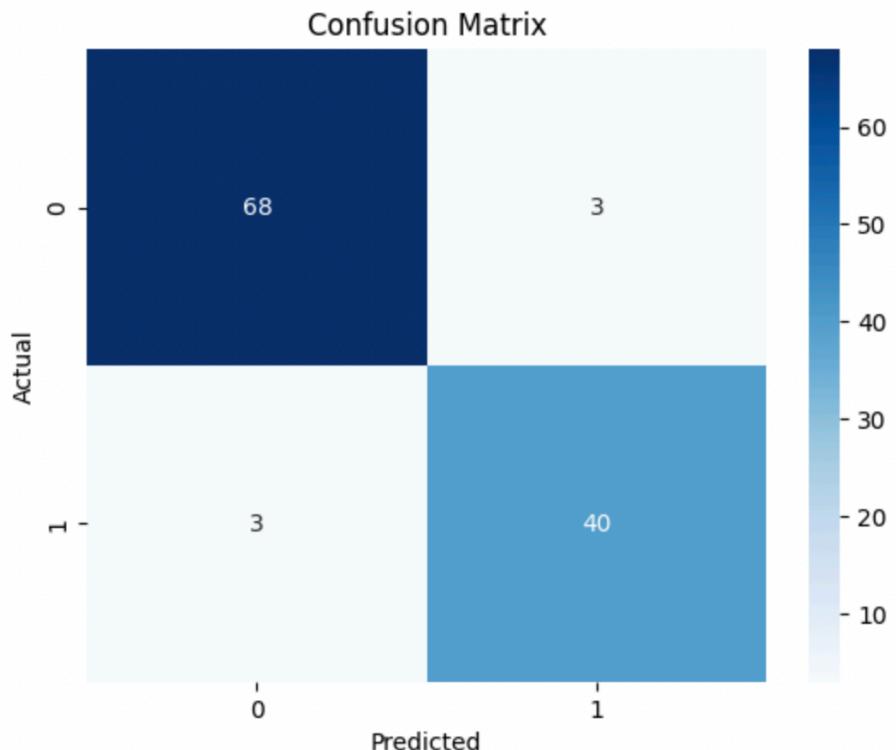
Adaboosting

```
Accuracy: 0.956140350877193  
Precision: 0.9523809523809523  
Recall: 0.9302325581395349  
F1-score: 0.9411764705882353
```



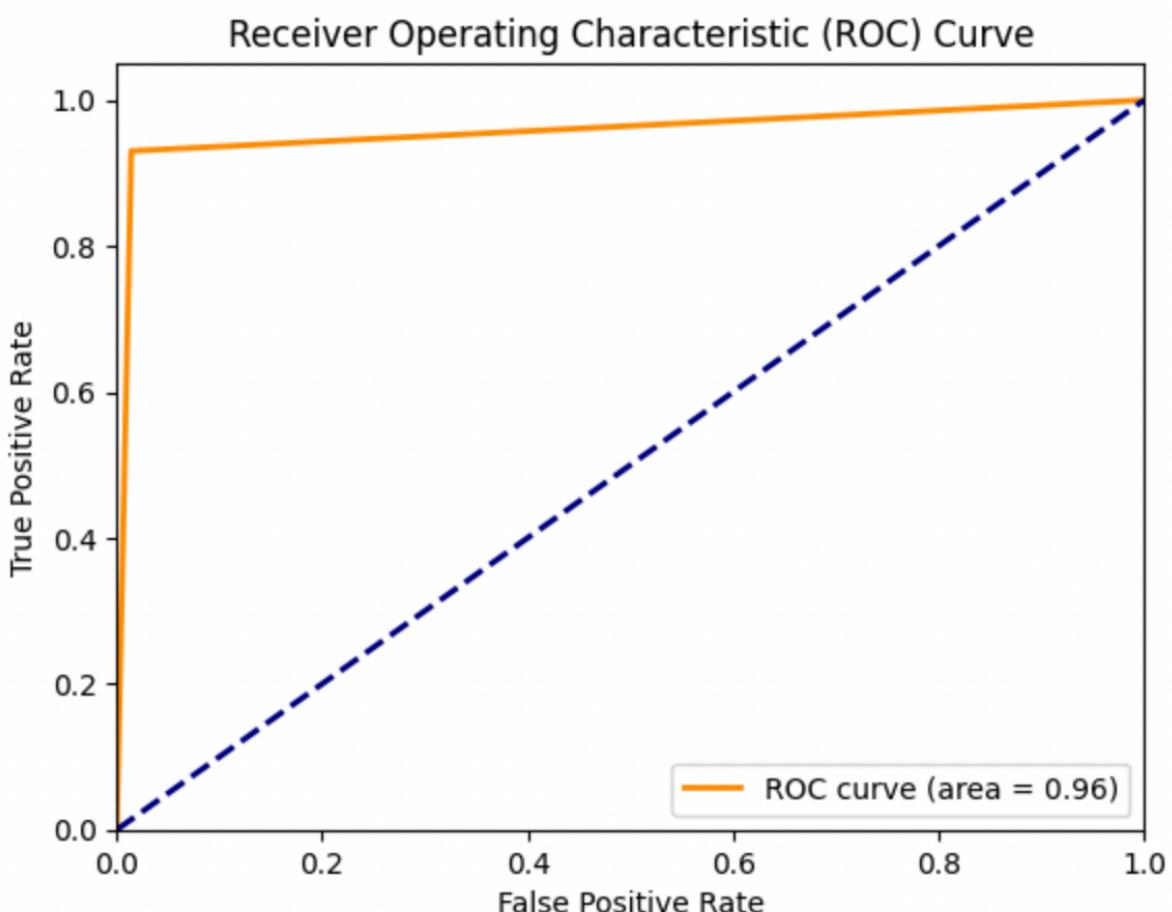
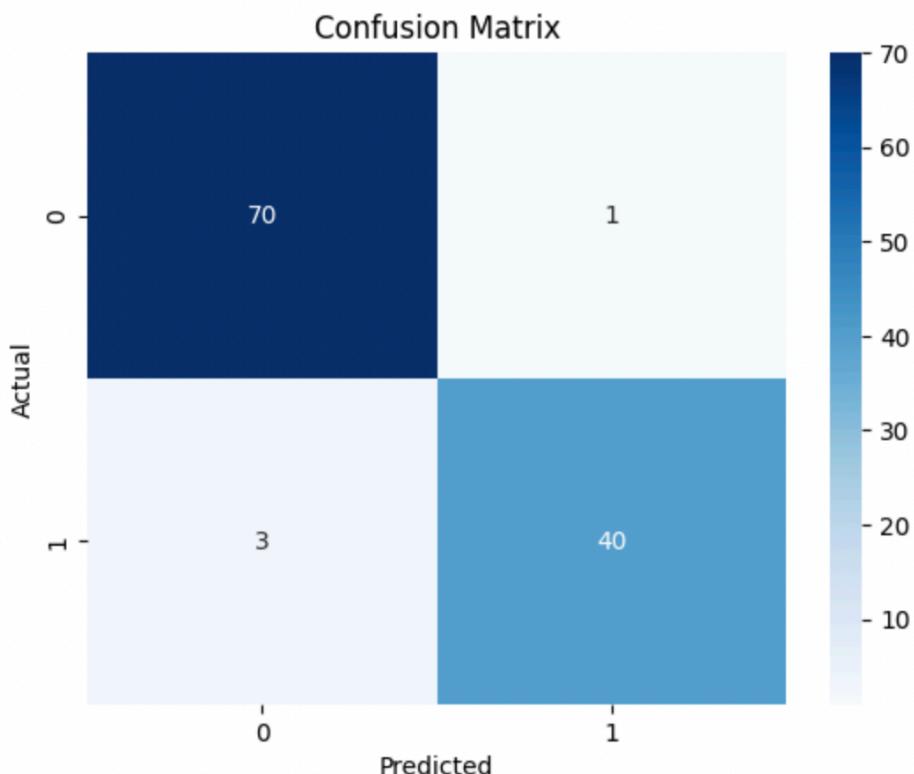
Gradient Boosting

```
Accuracy: 0.9473684210526315
Precision: 0.9302325581395349
Recall: 0.9302325581395349
F1-score: 0.9302325581395349
```



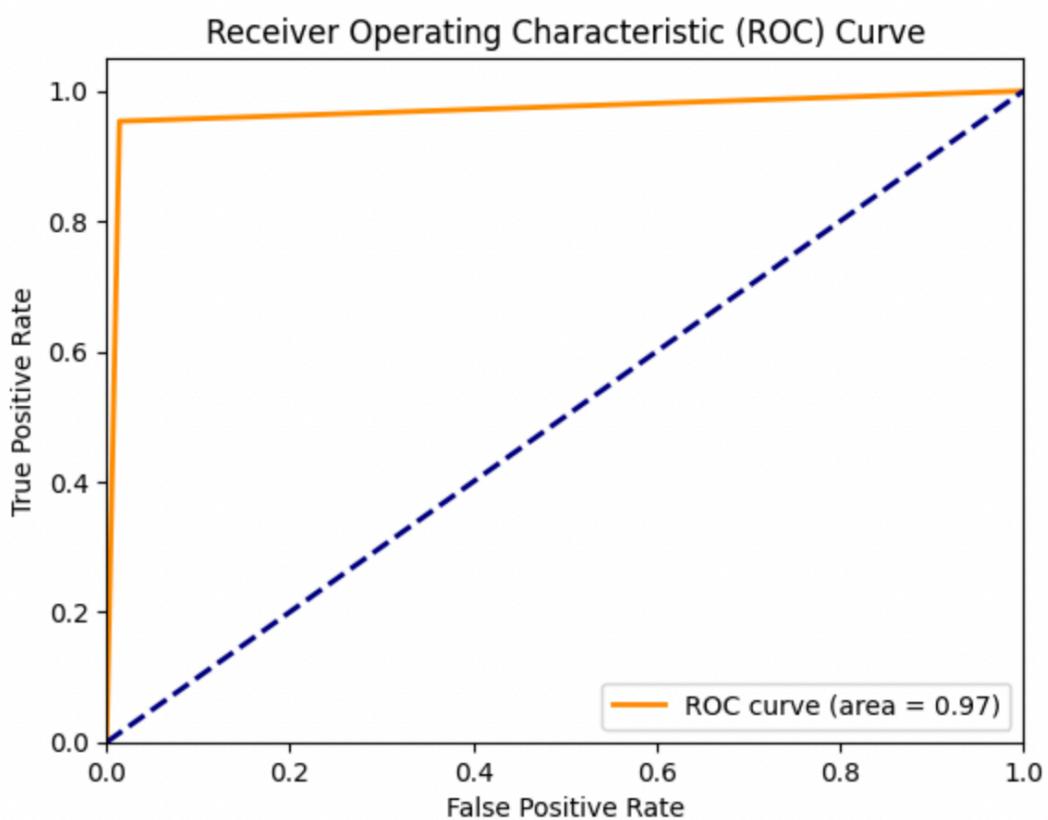
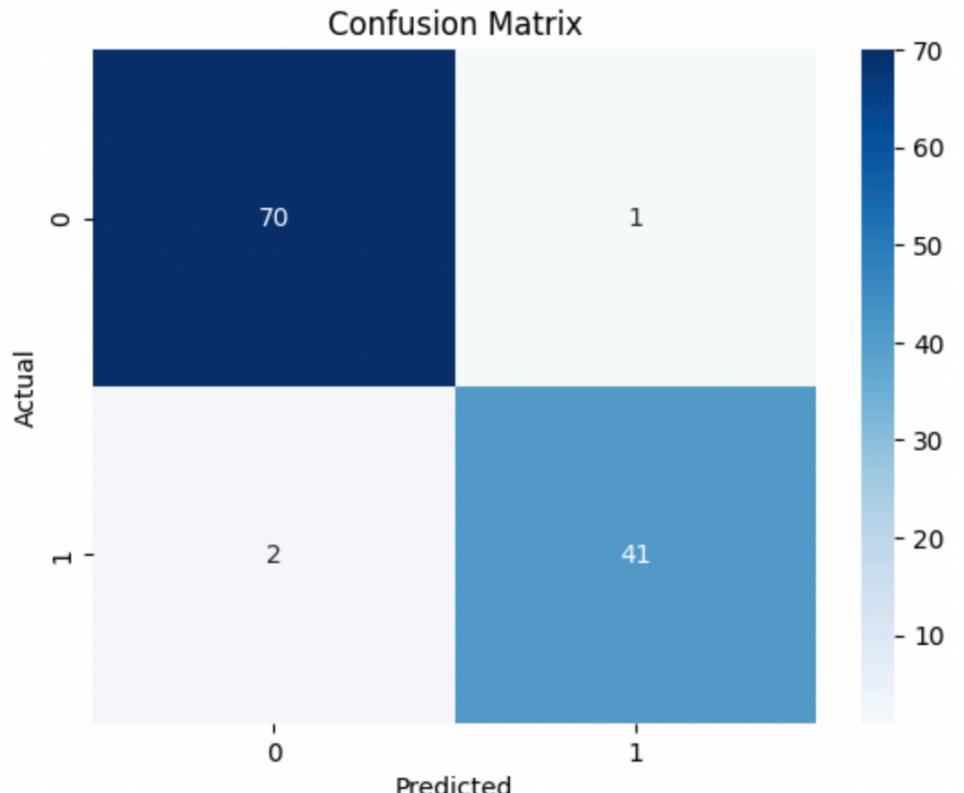
XGboost

```
Accuracy: 0.9649122807017544  
Precision: 0.975609756097561  
Recall: 0.9302325581395349  
F1-score: 0.9523809523809523
```



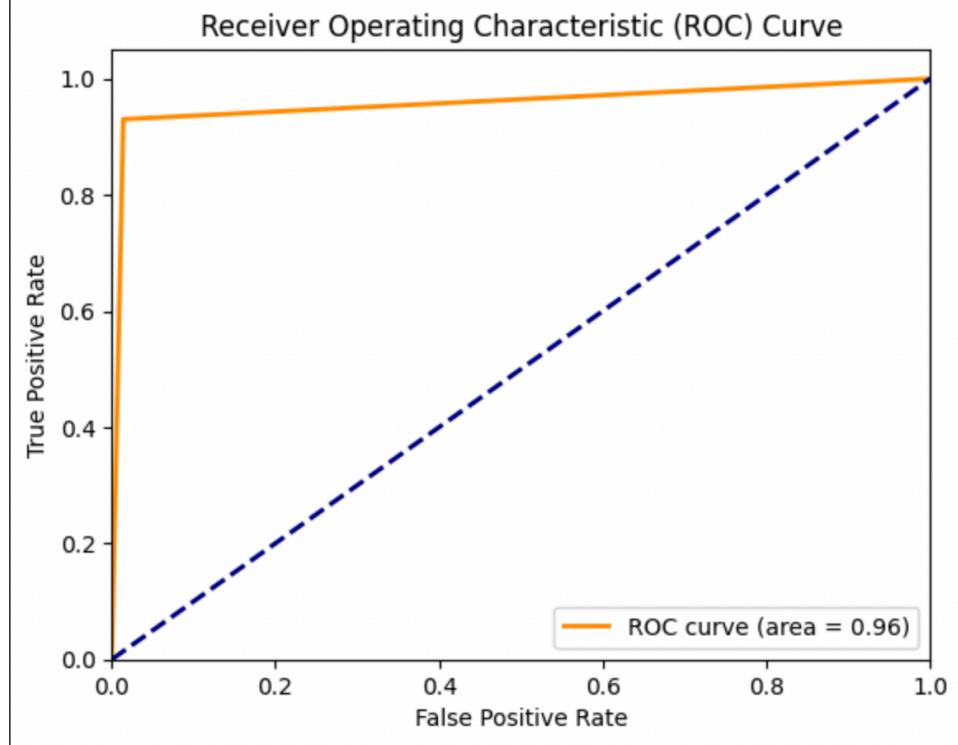
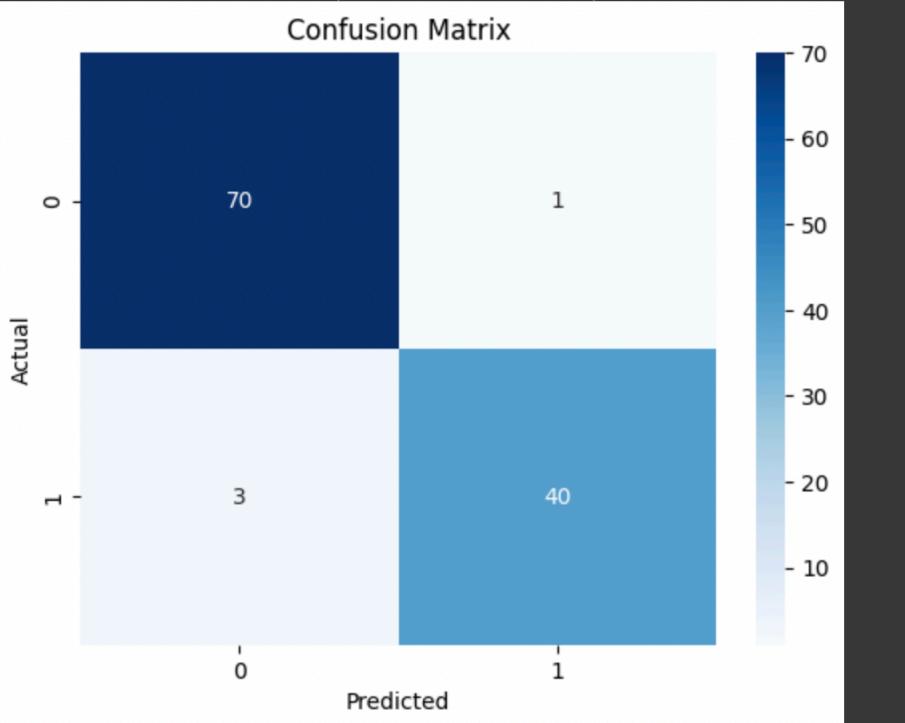
Random Forest Classifier

```
Accuracy: 0.9736842105263158
Precision: 0.9761904761904762
Recall: 0.9534883720930233
F1-score: 0.9647058823529412
```

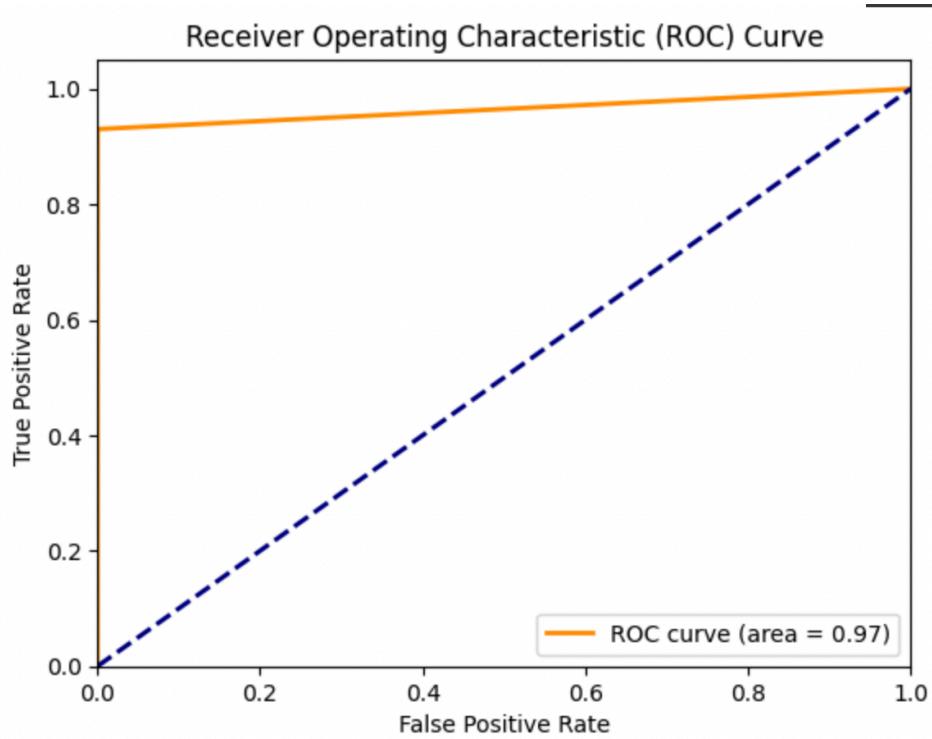
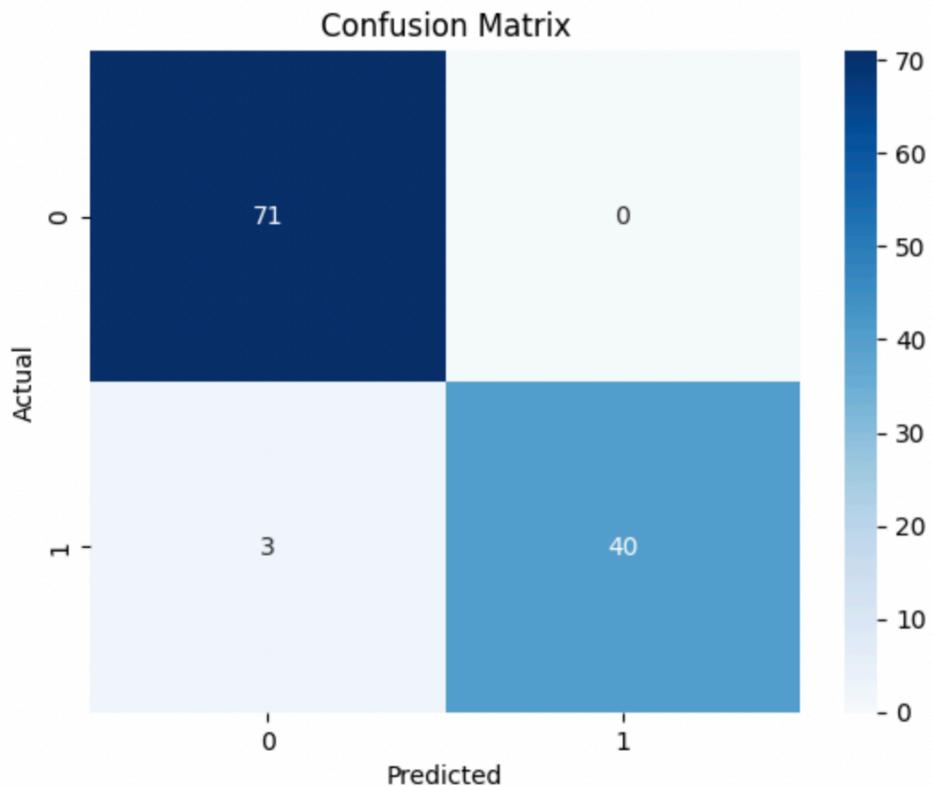


Stacked ensemble

Stack 1: SVM+NB+DT -> LR | Accuracy: 0.9649 | F1 Score: 0.9647



Stack 2: SVM+NB+DT -> RF | Accuracy: 0.9737 | F1 Score: 0.9735



Stack 3: SVM+DT+KNN -> LR | Accuracy: 0.9737 | F1 Score: 0.9736

