# Artificial Intelligence Model for Parkinson Disease Detection Using Machine Learning Algorithms

Shahid Ali FA24-RAI-015

Department of Computer Science
COMSATS University Islamabad

# Introduction

- **Parkinson's Disease (PD)** is a progressive neurological disorder that affects movement and speech.
- It impacts **motor functions**, **speech**, and **posture** due to dopamine deficiency.
- Symptoms include tremors, muscle rigidity, slow movements, and speech difficulties.
- Parkinson's disease mainly affects older people and slowly worsens over time.
- Early diagnosis is crucial to manage symptoms and improve quality of life.
- Early treatment can help patients a lot, but currently, diagnosis depends a lot on doctors' experience and observation, which can sometimes be wrong or inconsistent.
- Since PD affects speech, analyzing voice recordings can be a useful way to detect PD automatically.
- This Research leverages machine learning algorithms to detect PD using **voice features** to provide a fast, objective, and accurate decision support system.

# Problem Statement

- The current diagnostic techniques for Parkinson's Disease are subjective, rely on clinical observations, and often detect the disease at a later stage.
- The main challenge is the lack of quantitative, objective, and reliable methods for early diagnosis and staging of Parkinson’s disease.
- Existing diagnostic procedures are subjective and prone to variability. There is a need for automated and accurate models that can select important features from complex biomedical data and effectively classify patients.
- This **paper addresses** the problem by proposing machine learning models trained on carefully selected voice features to classify as Parkinson's patient versus Healthy individual.

# Related Work ( Literature review )

- Various AI methods have been proposed for PD diagnosis using speech ( Vocal Features ).
- Previous studies utilized feature selection techniques such as **LASSO**, **mRMR**, genetic algorithms, and different classifiers including SVM, KNN, and neural networks.
- While many achieved **good accuracy**, feature selection remains critical to improve classifier performance and reduce computational cost.

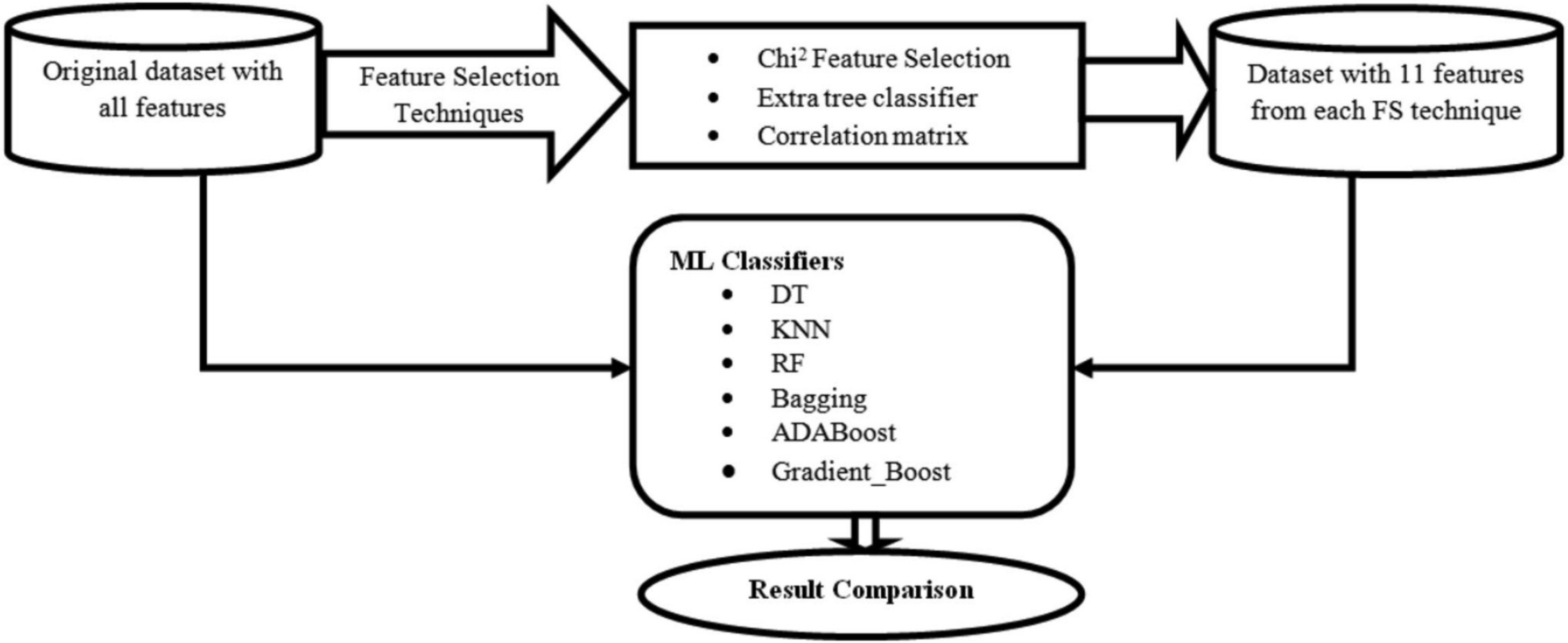
**Gaps in Previous Research:**

- Many studies use **black-box models** with limited explainability.
- **Overfitting** due to too many features is common.
- Lack of **comparison across multiple classifiers** using **consistent feature sets**.

The **current paper** **bridges these gaps** by:

- Using **feature selection** to minimize noise
- Applying **six classifiers** on the same dataset
- Evaluating both **full and reduced features**

# Methodology



# Methodology

**1. Dataset:**

- Sourced from **UCI Machine Learning Repository**
- **197 voice samples**, 23 features, and a binary label:
  - 1 = Parkinson's patient
  - 0 = Healthy individual

**Features include:**

- **Frequency-based**: MDVP:Fo(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz)
- **Jitter/Shimmer-based**: Indicators of vocal irregularity and amplitude variation
- **NHR / HNR**: Noise-to-Harmonic and Harmonic-to-Noise Ratios
- **DFA, PPE, Spread1, Spread2, D2:** Nonlinear dynamics and complexity measure
- **Data Split**: 80% training, 20% testing

# Methodology

**2. Feature Selection Techniques:**

- **Chi2 Test:** It checks which features are most related to whether a person has PD or not. Features with higher scores are more important.
- **Extra Trees Classifier:** An ensemble method that provides feature importance scores based on randomized decision trees. Features that help the most get higher importance scores.
- **Correlation Matrix:** Computes Pearson correlation between features and the target to select features with highest absolute correlation.

**3. Machine Learning Classifiers:**

- Decision Tree (DT)
- K-Nearest Neighbors (KNN)
- Random Forest (RF)
- Bagging
- AdaBoost
- Gradient Boosting **Our added models** ➡ **Logestic Regression, Naïve Bayes and SVM.**
- The methodology involves selecting top features from each method and training/testing classifiers on both full and reduced feature sets.

# Base Paper Results

| No. of features | % Accuracy | | | |
|---|---|---|---|---|
| | Without feature selection (23-Attributes) | Chi$^2$ feature selection Technique (11- Attributes) | Extra trees classifier feature selection Technique (11-Attributes) | Correlation matrix feature selection technique (11-Attributes) |
| DT | 94.87 | 94.87 | 94.87 | 94.87 |
| KNN | 82.05 | 82.05 | 82.05 | 79.48 |
| RF | 92.30 | 92.30 | 89.74 | 87.17 |
| Bagging | 92.30 | 89.74 | 89.74 | 87.17 |
| AdaBoosting | 87.17 | 84.61 | 87.17 | 89.74 |
| Gradient boosting | 92.30 | 94.87 | 92.30 | 94.87 |

# Our Results

| ML Classifier | Accuracy (Full Features) | Accuracy Chi2 | Accuracy Extra Tree Classifier | Accuracy Corelation Matrix | AUC |
|---|---|---|---|---|---|
| DT | 94.92% | 82% | 98.31% | 85% | 94.90% |
| KNN | 98% | 83% | 97% | 98% | 100% |
| RF | 100% | 100% | 100% | 97% | 100% |
| Bagging | 99.57% | 76.26% | 99.57% | 97% | 100% |
| AdaBoosting | 85% | 76.27% | 85% | 80% | 95% |
| Gradient Boosting | 98.31% | 92% | 100% | 96.61% | 100% |
| Logestic Regression | 80% | 70% | 87% | 84% | 95% |
| Naive Bayes | 79% | 65% | 83% | 80% | 97% |
| SVM | 90% | 76% | 89.83% | 83% | 95% |

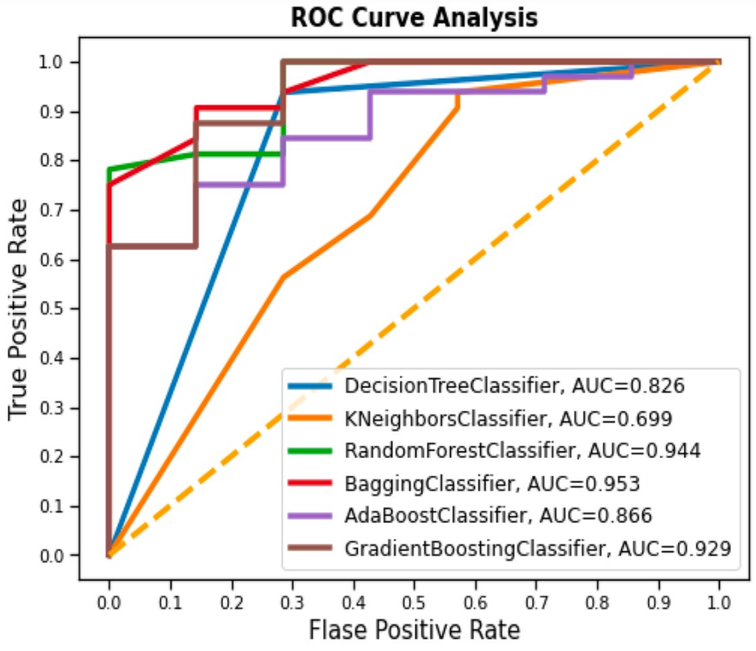
- **Random Forest, Bagging and Gradient Boosting** performed the best.
- Feature selection improved model performance and reduced complexity.

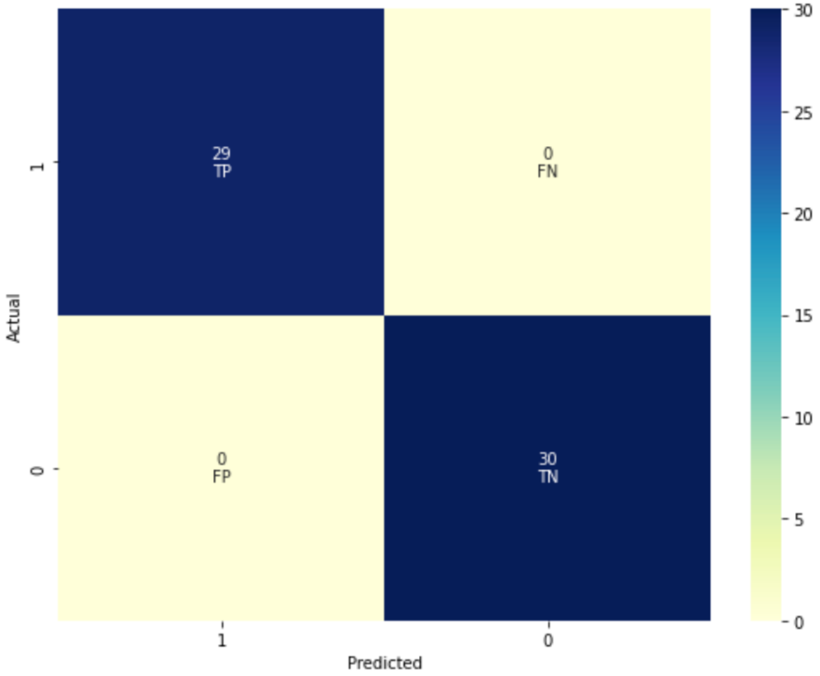
📈 **ROC Curve Analysis**

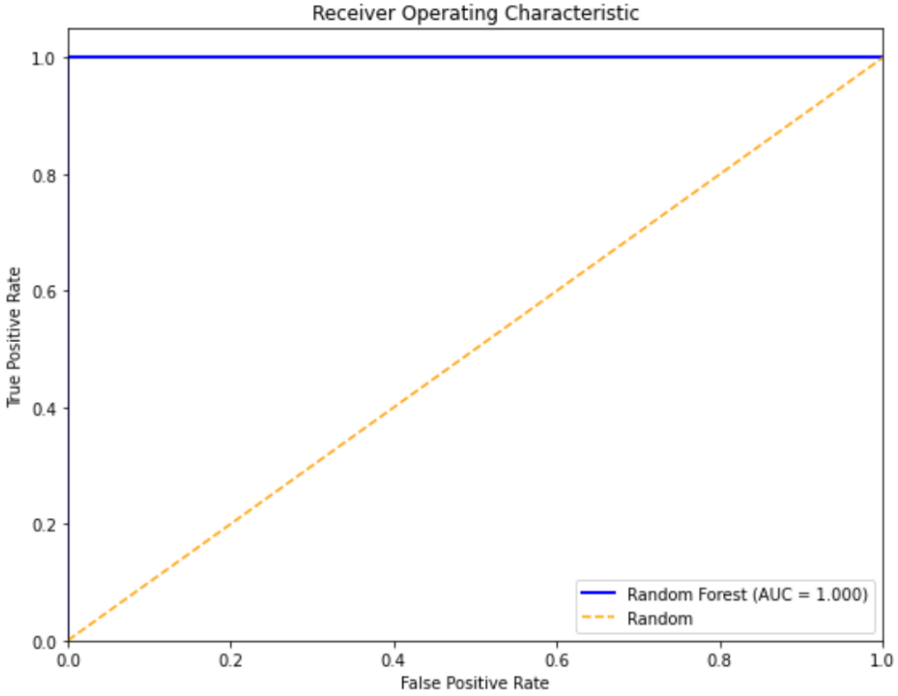
- ROC curves confirm high model performance.
- Best AUC = **100%** (RF and Gradient Boosting).

# Random Forest Results

```
=== Random Forest ===
Classification report (Test data):
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        30
           1       1.00      1.00      1.00        29

    accuracy                           1.00        59
   macro avg       1.00      1.00      1.00        59
weighted avg       1.00      1.00      1.00        59

Train Accuracy: 100.00%
Test Accuracy: 100.00%
Train RMSE: 0.0000
Test RMSE: 0.0000
Area under the curve: 1.000
```

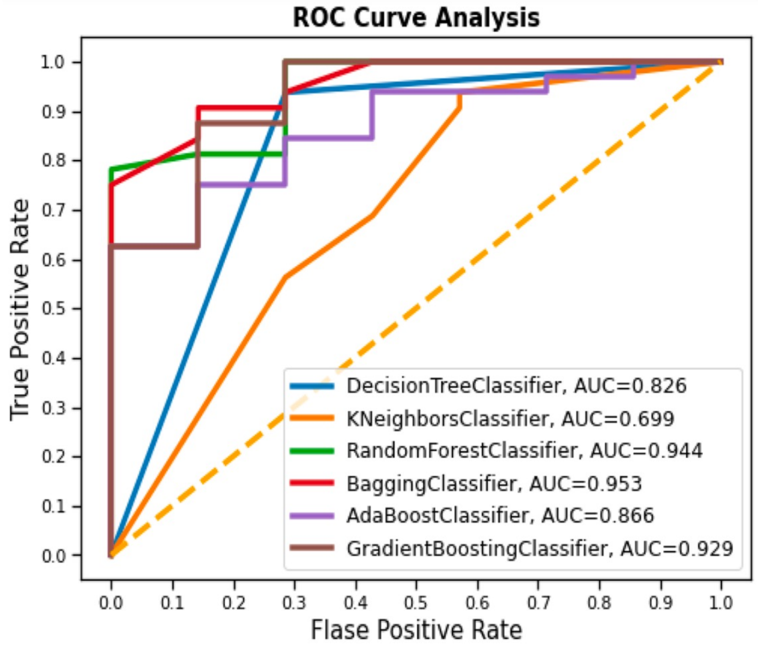


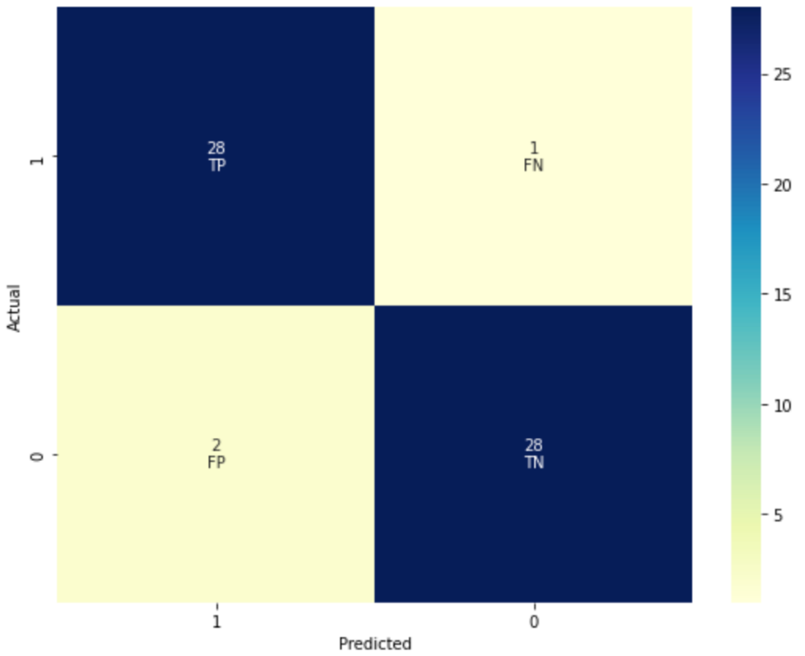



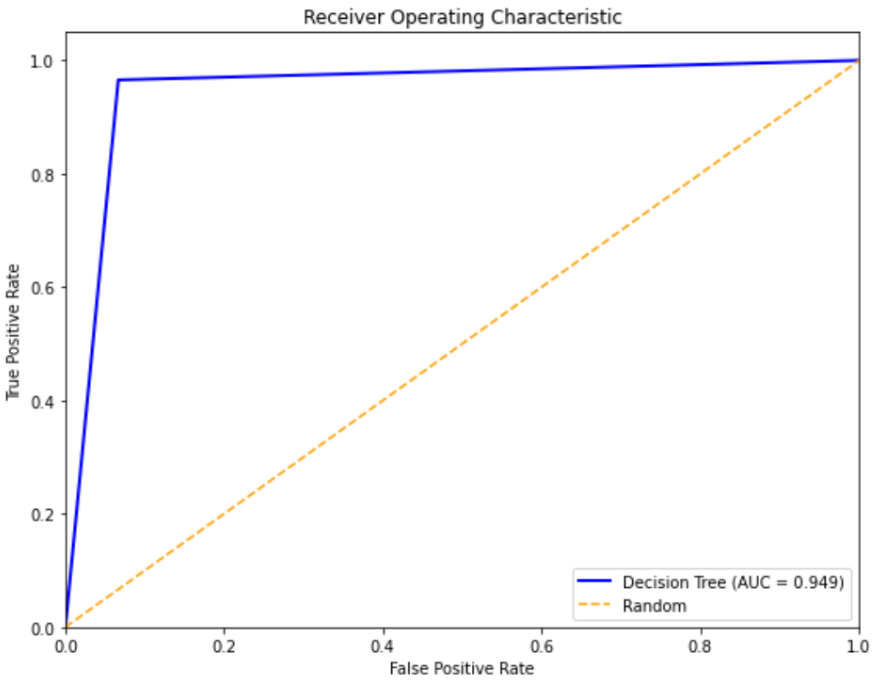



# Decision Trees Classifier Result

```
=== Decision Tree ===
Classification report (Test data):
              precision    recall  f1-score   support

           0       0.97      0.93      0.95        30
           1       0.93      0.97      0.95        29

    accuracy                           0.95        59
   macro avg       0.95      0.95      0.95        59
weighted avg       0.95      0.95      0.95        59

Train Accuracy: 100.00%
Test Accuracy: 94.92%
Train RMSE: 0.0000
Test RMSE: 0.2255
Area under the curve: 0.949
```

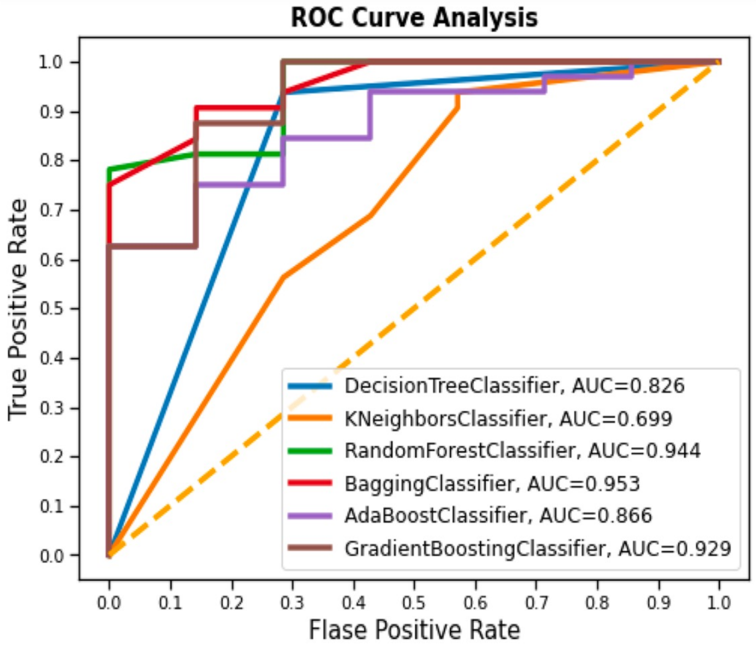


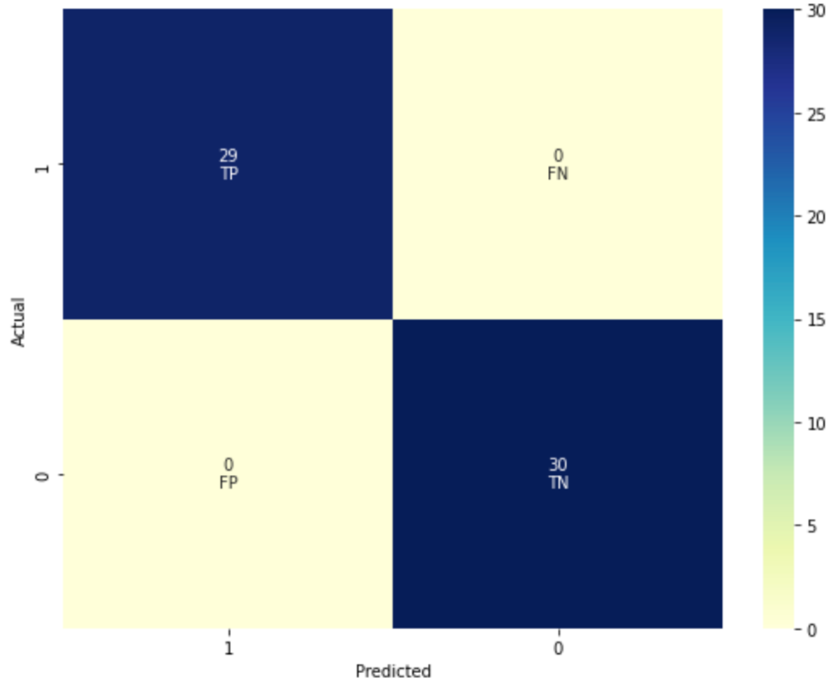



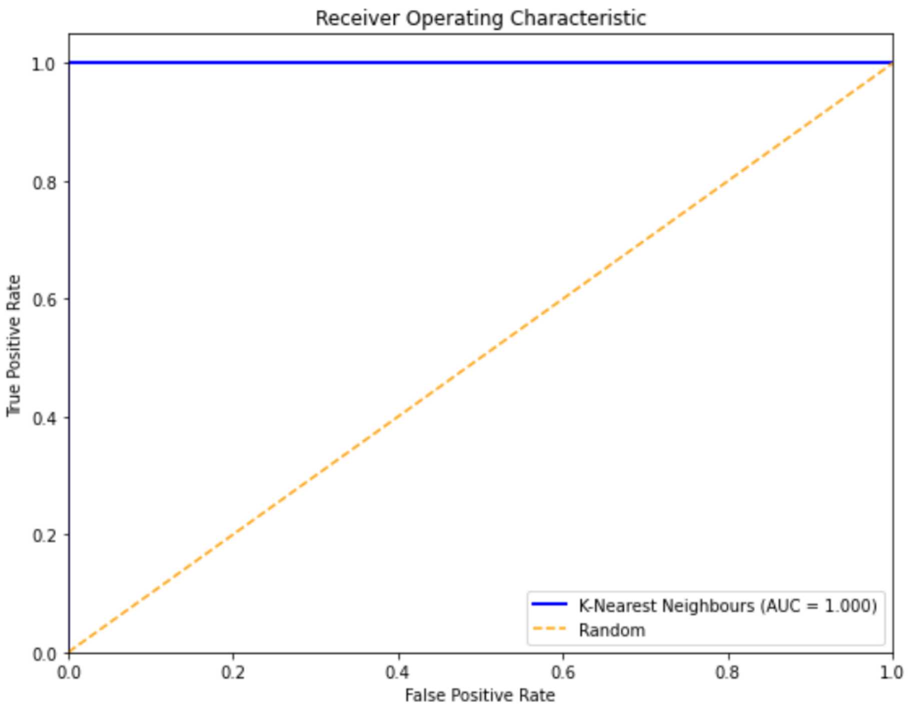



# KNN Results

```
=== K-Nearest Neighbours ===
Classification report (Test data):
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        30
           1       1.00      1.00      1.00        29

    accuracy                           1.00        59
   macro avg       1.00      1.00      1.00        59
weighted avg       1.00      1.00      1.00        59

Train Accuracy: 97.45%
Test Accuracy: 100.00%
Train RMSE: 0.1598
Test RMSE: 0.0000
Area under the curve: 1.000
```

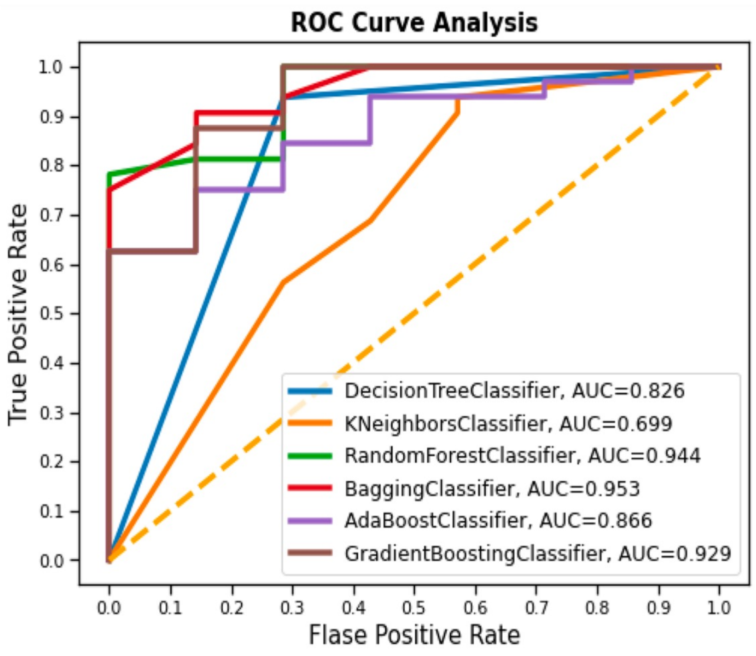


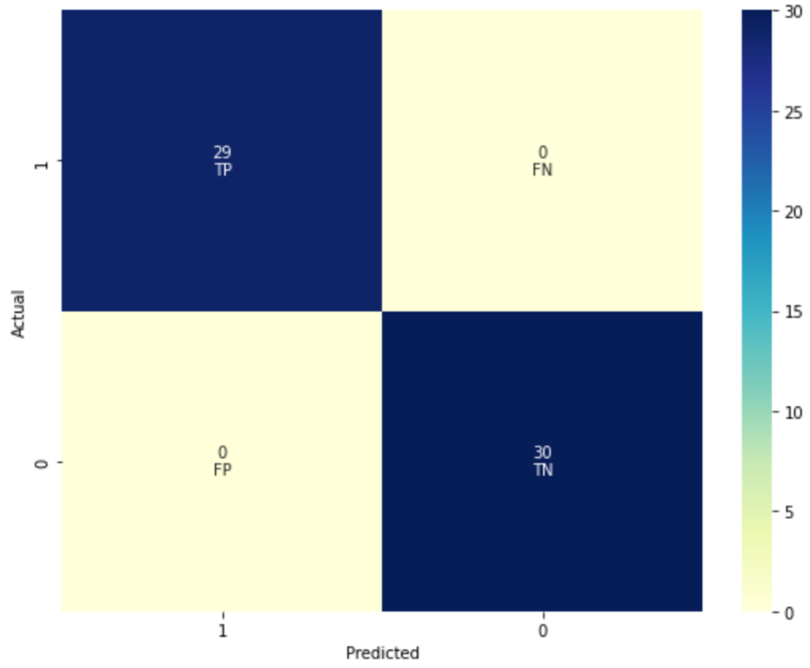



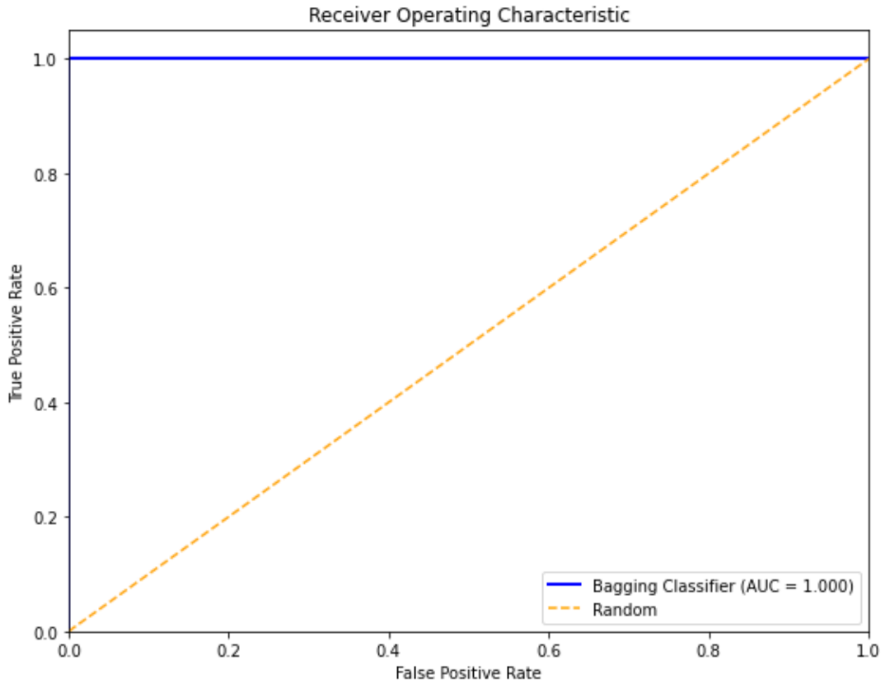



# Bagging Results

```
=== Bagging Classifier ===
Classification report (Test data):
              precision    recall  f1-score   support

           0       1.00      1.00      1.00        30
           1       1.00      1.00      1.00        29

    accuracy                           1.00        59
   macro avg       1.00      1.00      1.00        59
weighted avg       1.00      1.00      1.00        59

Train Accuracy: 100.00%
Test Accuracy: 100.00%
Train RMSE: 0.0000
Test RMSE: 0.0000
Area under the curve: 1.000
```

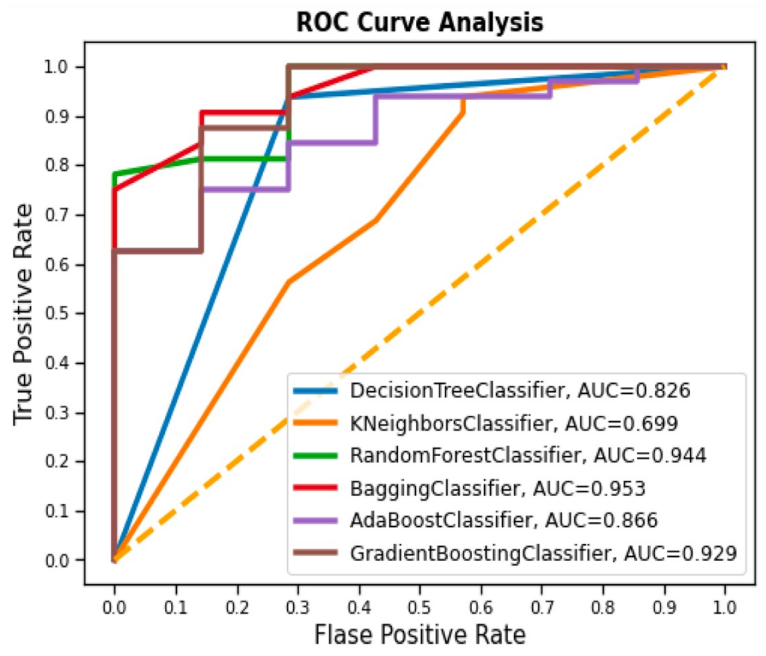


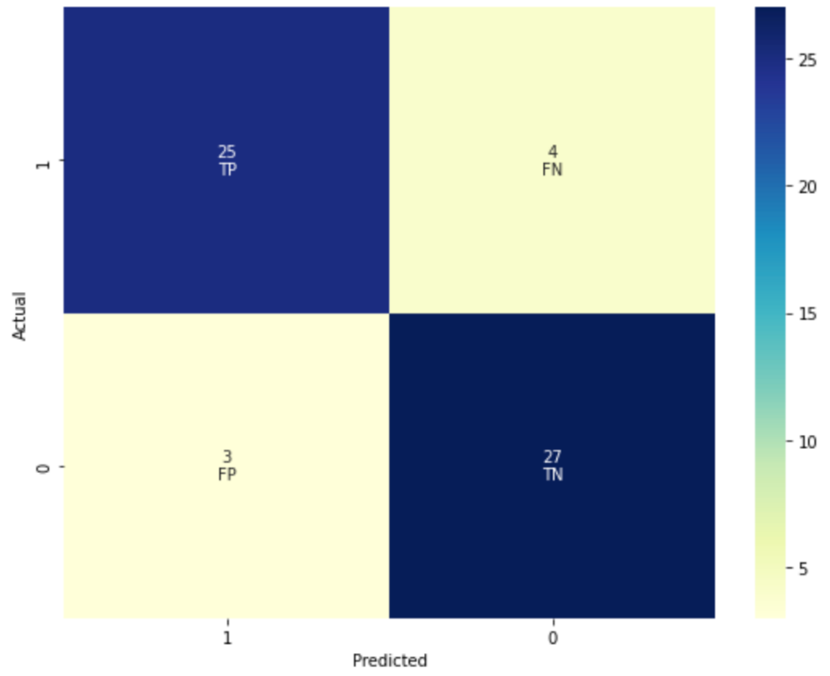



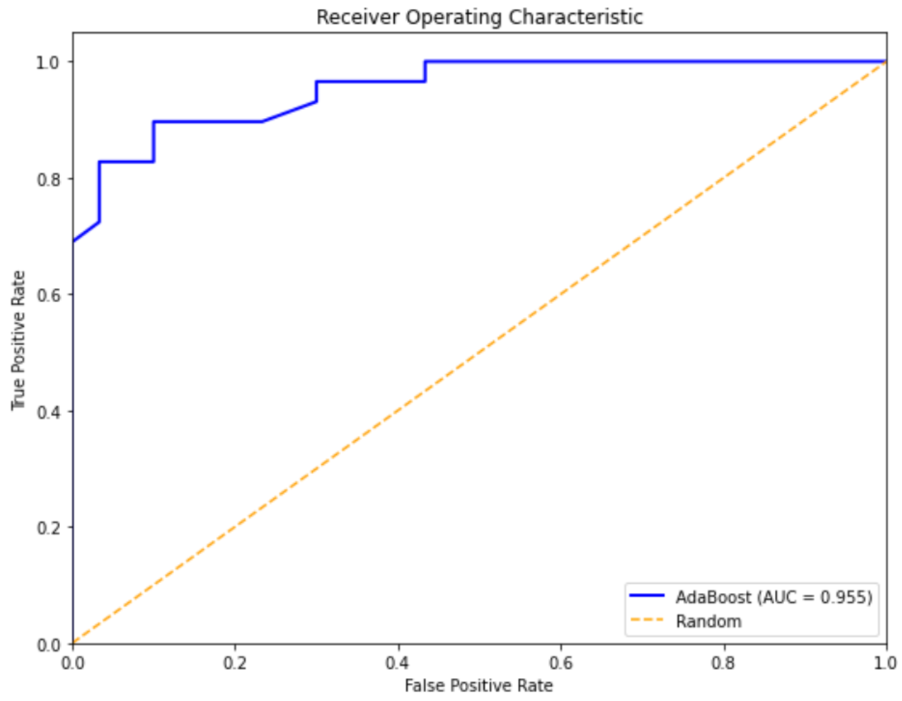



# AdaBoosting Results

```
=== AdaBoost ===
Classification report (Test data):
              precision    recall  f1-score   support

           0       0.87      0.90      0.89        30
           1       0.89      0.86      0.88        29

    accuracy                           0.88        59
   macro avg       0.88      0.88      0.88        59
weighted avg       0.88      0.88      0.88        59

Train Accuracy: 84.26%
Test Accuracy: 88.14%
Train RMSE: 0.3968
Test RMSE: 0.3444
Area under the curve: 0.955
```

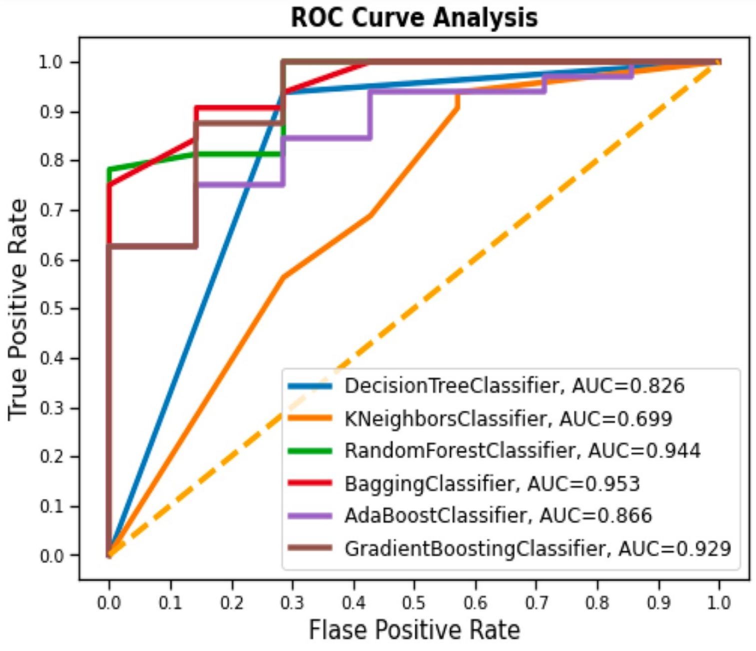


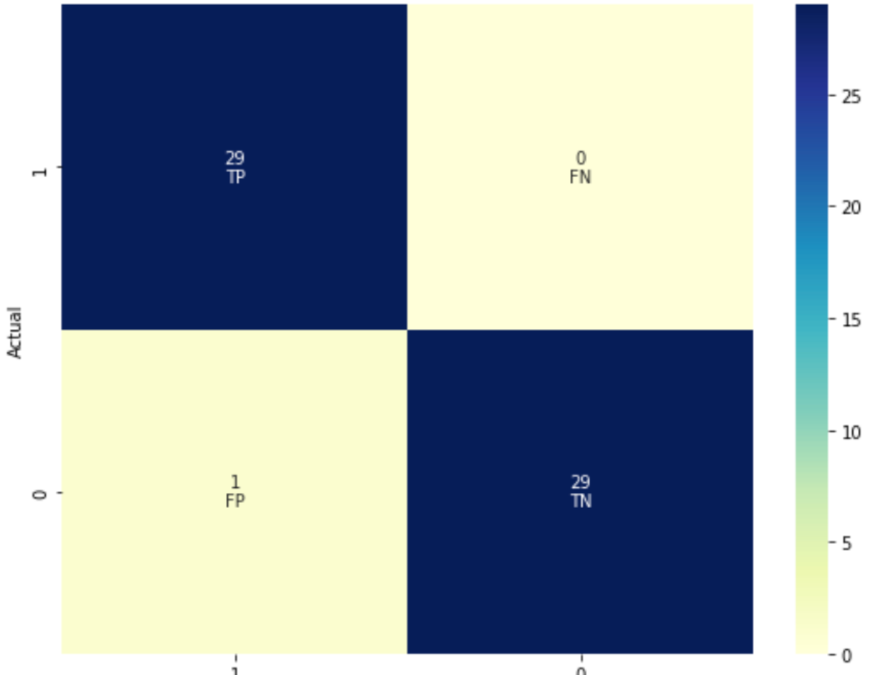



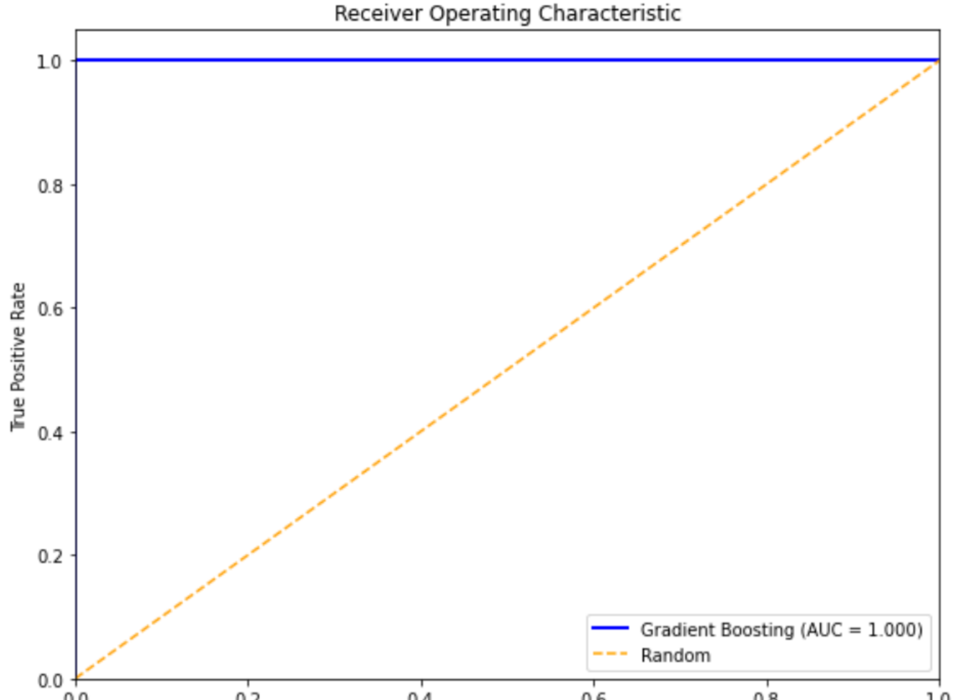



# Gradient Boosting Results

```
=== Gradient Boosting ===
Classification report (Test data):
              precision    recall  f1-score   support

           0       1.00      0.97      0.98        30
           1       0.97      1.00      0.98        29

    accuracy                           0.98        59
   macro avg       0.98      0.98      0.98        59
weighted avg       0.98      0.98      0.98        59

Train Accuracy: 100.00%
Test Accuracy: 98.31%
Train RMSE: 0.0000
Test RMSE: 0.1302
Area under the curve: 1.000
```

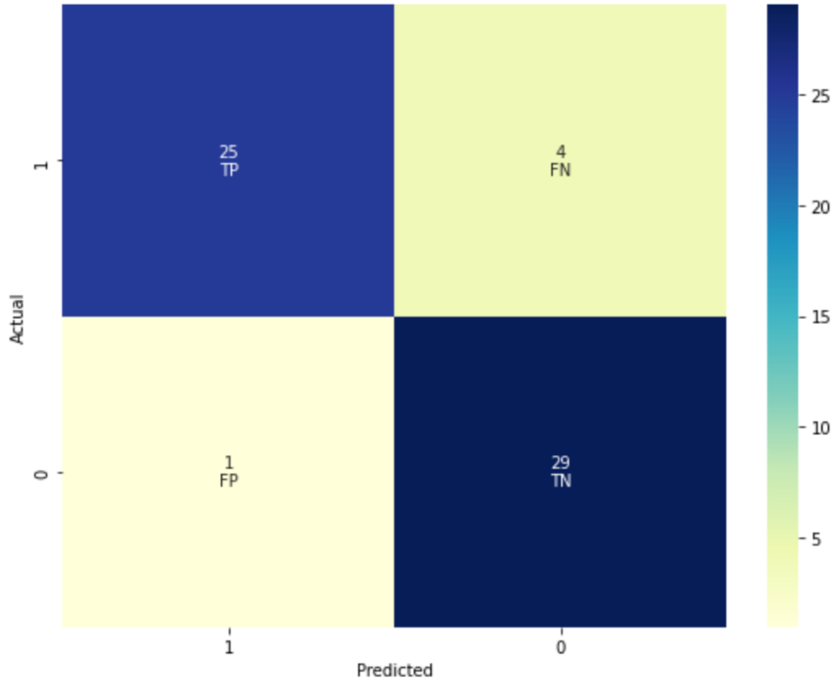


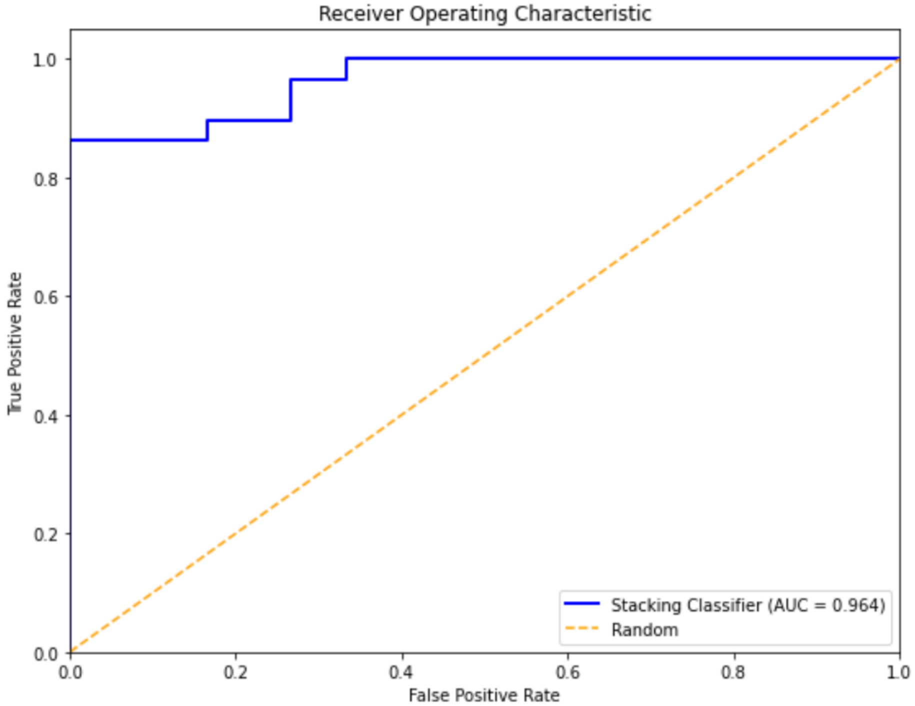



# Stacking Classifier Results

```
=== Stacking Classifier ===
Classification report (Test data):
              precision    recall  f1-score   support

           0       0.88      0.97      0.92        30
           1       0.96      0.86      0.91        29

    accuracy                           0.92        59
   macro avg       0.92      0.91      0.91        59
weighted avg       0.92      0.92      0.91        59

Train Accuracy: 90.21%
Test Accuracy: 91.53%
Train RMSE: 0.3128
Test RMSE: 0.2911
Area under the curve: 0.964
```



# Conclusion

- This study presents an effective AI-based approach to Parkinson’s disease detection using feature selection and classification algorithms.
- The combination of Chi2, Extra Trees, and Correlation Matrix methods helps identify the most relevant voice features for PD diagnosis.
- **Random Forest, Bagging, GB and SC** classifier emerged as the best performer.
- The method is suitable for clinical use due to its high accuracy and low computational complexity.
- Future research could explore deep learning techniques and multimodal data to further improve detection accuracy.

# THANKS