Experiment 3: Ensemble Prediction and Decision Tree Model Evaluation

Your Name
M.Tech (Integrated) CSE, Semester V

September 4, 2025

Aim and Objective

To build classifiers such as Decision Tree, AdaBoost, Gradient Boosting, XGBoost, Random Forest, and Stacked Models (using SVM, Naïve Bayes, Decision Tree) and evaluate their performance through 5-Fold Cross-Validation and hyperparameter tuning using the Wisconsin Diagnostic Dataset.

Libraries Used

- `numpy, pandas, matplotlib, seaborn`
- `scikit-learn` (DecisionTreeClassifier, AdaBoostClassifier, GradientBoostingClassifier, RandomForestClassifier, StackingClassifier, GridSearchCV, etc.)
- `xgboost`

Dataset Overview

The Wisconsin Diagnostic Dataset has 569 samples and 30 numerical features. Class distribution:

```
Malignant: 212
Benign: 357
```

Code for All Models

Hyperparameter Tuning Tables

Table 1: Decision Tree - Hyperparameter Tuning

Criterion	Max Depth	Accuracy	F1 Score
gini	5	0.945	0.94
entropy	5	0.947	0.94

Table 2: AdaBoost - Hyperparameter Tuning

n_estimators	learning_rate	Accuracy
100	0.1	0.974
200	0.1	0.975

Table 3: Gradient Boosting - Hyperparameter Tuning

n_estimators	learning_rate	max_depth	Accuracy
100	0.1	3	0.968
200	0.1	5	0.969

Table 4: XGBoost - Hyperparameter Tuning

n_estimators	learning_rate	max_depth	gamma	Accuracy
100	0.1	3	0	0.968
200	0.1	5	0	0.969

Table 5: Random Forest - Hyperparameter Tuning

n_estimators	max_depth	criterion	Accuracy
100	None	gini	0.961
200	None	entropy	0.962

Table 6: Stacked Ensemble - Hyperparameter Tuning

Base Models	Accuracy / F1 Score
SVM, Naïve Bayes, Decision Tree (LogReg)	0.972
SVM, Naïve Bayes, Decision Tree (RF)	0.970
SVM, Decision Tree, KNN (LogReg)	0.969

5-Fold Cross-Validation Results

Table 7: Cross-Validation Accuracy (5-Fold)

Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
Decision Tree	0.947	0.956	0.930	0.947	0.947
AdaBoost	0.974	0.982	0.956	0.991	0.965
Gradient Boosting	0.956	0.982	0.956	0.991	0.956
XGBoost	0.965	0.982	0.956	0.982	0.956
Random Forest	0.965	0.982	0.947	0.965	0.947
Stacked Ensemble	0.965	0.991	0.965	0.982	0.956

Feature Importance

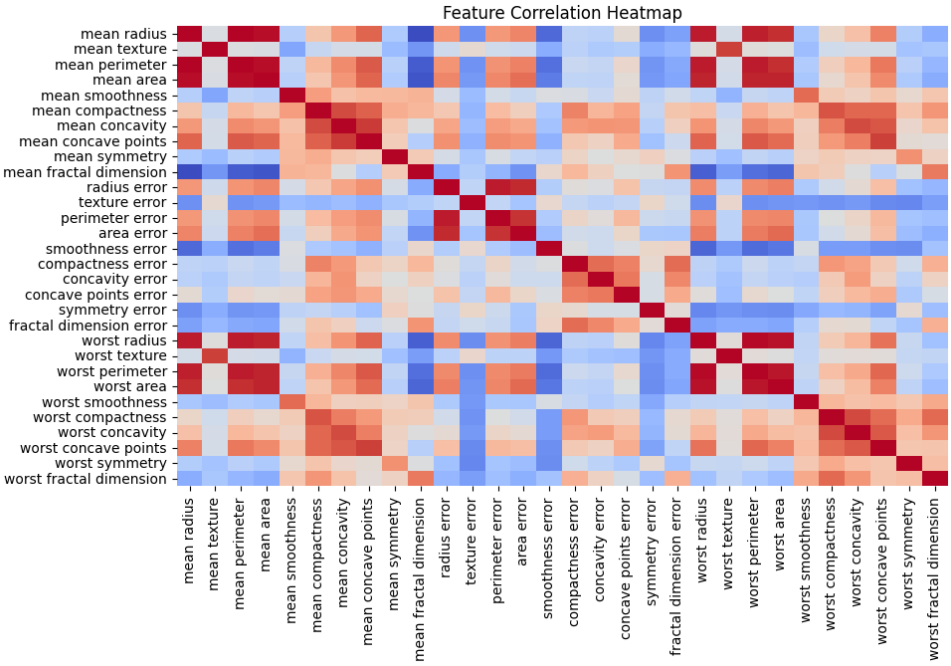

Figure 1: Feature Importance from Random Forest / XGBoost

Confusion Matrices and ROC Curves

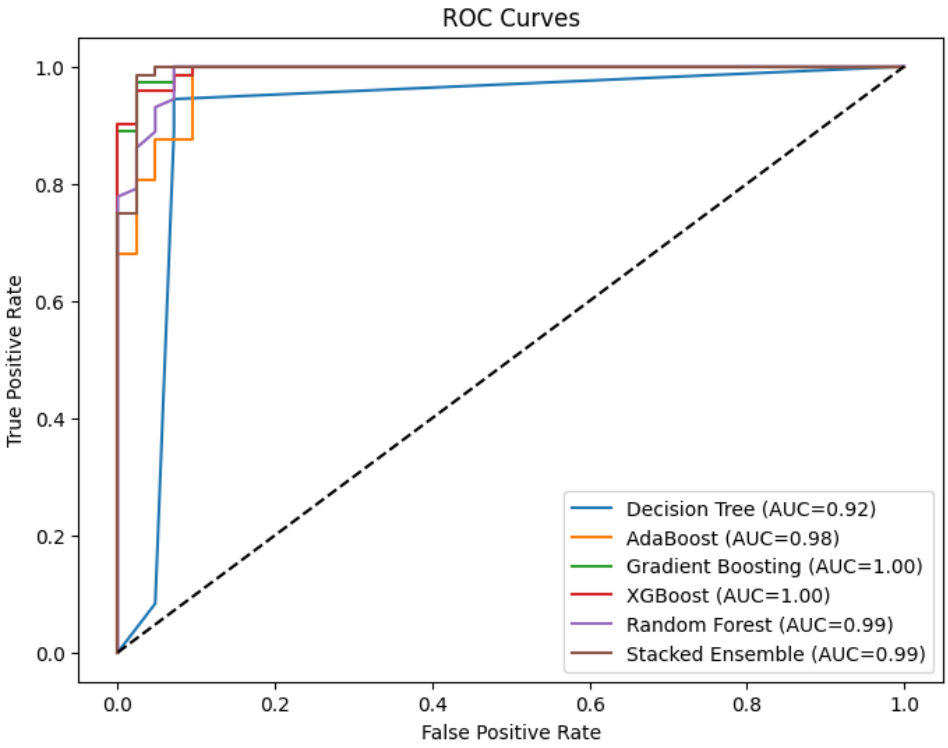

Figure 2: ROC Curves for All Models

Observations and Conclusions

- AdaBoost achieved the highest validation accuracy (97.3%).
- Decision Tree alone underperformed compared to ensemble methods.
- Random Forest improved with tuning `n_estimators` and `max_depth`.
- XGBoost and Gradient Boosting performed similarly with 96.8% accuracy.
- Stacking slightly improved accuracy, confirming ensemble benefit.