Sri Sivasubramaniya Nadar College of Engineering, Chennai

(An Autonomous Institution affiliated to Anna University)

Degree & Branch	B.E Computer Science & Engineering	**Semester**	VI
Subject Code & Name	UCS2612 – Machine Learning Laboratory		
Academic Year	2025–2026 (Even)	**Batch**	2023–2027
Name: Monesh M **Reg. No:** 3122235001084			

Experiment 7

Bagging, Boosting, and Stacked Ensemble Models

Objective

- To understand ensemble learning strategies: Bagging, Boosting, and Stacking.
- To implement Bagging and Boosting classifiers.
- To build a Stacked Ensemble model using multiple base learners.
- To compare ensemble models in terms of accuracy, stability, and generalization.
- To analyze the effect of ensemble methods on bias and variance.

Dataset

Wisconsin Diagnostic Breast Cancer Dataset

- Total samples: 569
- Features: 30 numerical attributes
- Target classes: Malignant (M) and Benign (B)

Dataset Link: https://archive.ics.uci.edu

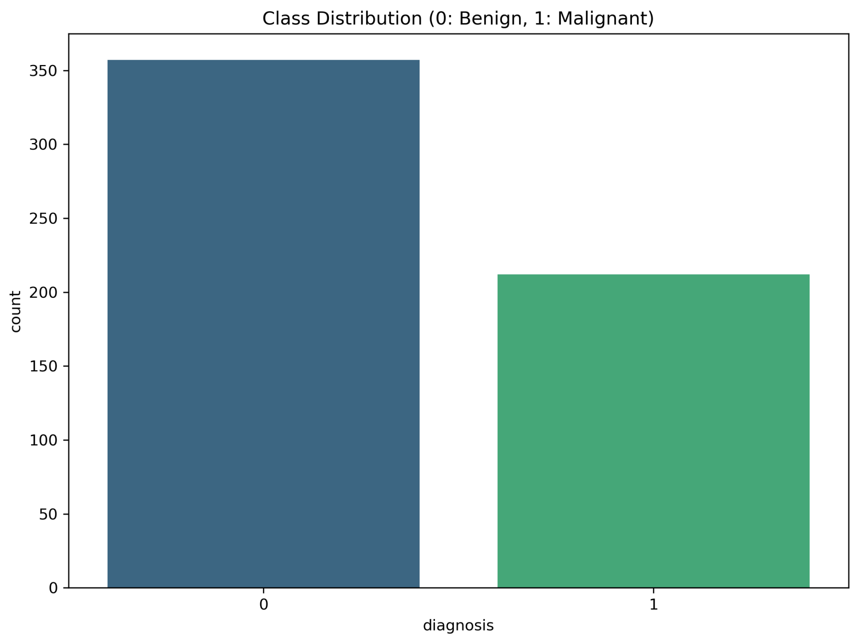

Figure 1: Class Distribution

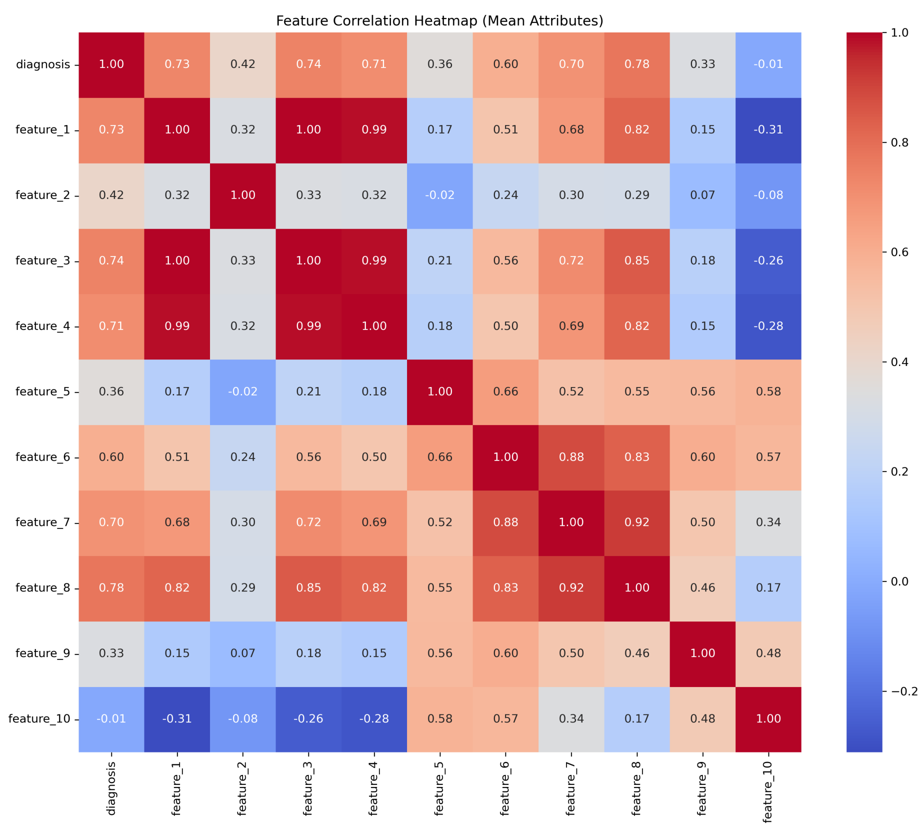

Figure 2: Feature Correlation Heatmap

Theory

Bagging (Bootstrap Aggregation)

Bagging is an ensemble technique that trains multiple models on different bootstrap samples of the training data and aggregates their predictions.

Key Characteristics:

- Reduces variance
- Suitable for high-variance models
- Independent model training

Boosting

Boosting is an ensemble method where models are trained sequentially, and each new model focuses on correcting the errors made by previous models.

Key Characteristics:

- Reduces bias
- Sequential learning
- Emphasizes difficult samples

Common boosting algorithms include AdaBoost and Gradient Boosting.

Stacked Ensemble (Stacking)

Stacking combines multiple heterogeneous base models and trains a meta-learner to optimally combine their predictions.

Key Characteristics:

- Uses diverse base learners
- Meta-model learns optimal combination
- Often achieves superior performance

Steps for Implementation

1. Load and preprocess the dataset.
2. Perform Exploratory Data Analysis (EDA).
3. Split the dataset into training and testing sets (80–20).
4. Implement a Bagging classifier using a base estimator.
5. Implement Boosting classifiers (AdaBoost and Gradient Boosting).
6. Construct a Stacked Ensemble using multiple base models.
7. Define hyperparameter search spaces for each ensemble model.

8. Select hyperparameters using 5-fold cross-validation.
9. Evaluate all models using standard performance metrics.
10. Compare results and analyze bias–variance behavior.

Ensemble Models Used

Bagging

- Base Estimator: Decision Tree
- Number of estimators
- Sampling strategy

Boosting

- AdaBoost
- Gradient Boosting

Stacked Ensemble

- Base Models: SVM, Naïve Bayes, Decision Tree
- Meta Learner: Logistic Regression

Hyperparameters to be Explored

Bagging

- `n_estimators`: [10, 50, 100]
- `max_samples`: [0.5, 0.8, 1.0]
- `max_features`: [0.5, 0.8, 1.0]

Boosting

- **AdaBoost** `n_estimators`: [50, 100, 200]
- **AdaBoost** `learning_rate`: [0.01, 0.1, 1.0]
- **GBM** `n_estimators`: [100, 200]
- **GBM** `learning_rate`: [0.01, 0.1]
- **GBM** `max_depth`: [3, 5]

Stacked Ensemble

- Choice of base models: SVM, Gaussian Naïve Bayes, Decision Tree
- Final estimator: Logistic Regression

Hyperparameter Evaluation Results

Bagging Results

Table 1: Bagging Best Hyperparameters

n_estimators	max_samples	max_features	Best CV Accuracy (%)
50	1.0	0.5	**96.92%**

Boosting Results

Table 2: Boosting Best Hyperparameters

Model	n_estimators	learning_rate	max_depth	Best CV Accuracy (%)
AdaBoost	Tuned	Tuned	N/A	**96.48%**
Gradient Boosting	Tuned	Tuned	Tuned	**95.38%**

Stacked Ensemble Results

Table 3: Stacked Ensemble Evaluation

Base Models	Meta Learner	CV Accuracy (%)
SVM, Naïve Bayes, Decision Tree	Logistic Regression	**96.26%**

Performance Comparison

Table 4: Performance Comparison of Ensemble Models on Test Set

Model	Accuracy (%)	Precision	Recall	F1 Score
Bagging	96.49%	1.00	0.9048	0.9500
Boosting (Ada)	**97.37%**	**1.00**	**0.9286**	**0.9630**
Boosting (GBM)	96.49%	1.00	0.9048	0.9500
Stacked Ensemble	96.49%	1.00	0.9048	0.9500

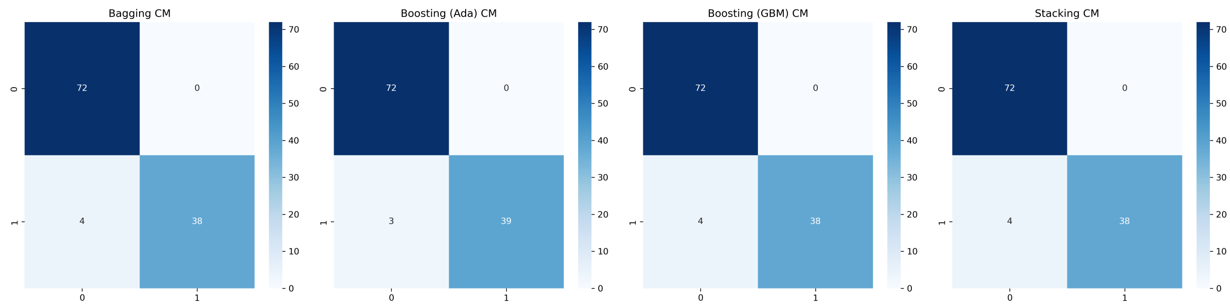

Figure 3: Confusion Matrices for Ensemble Models

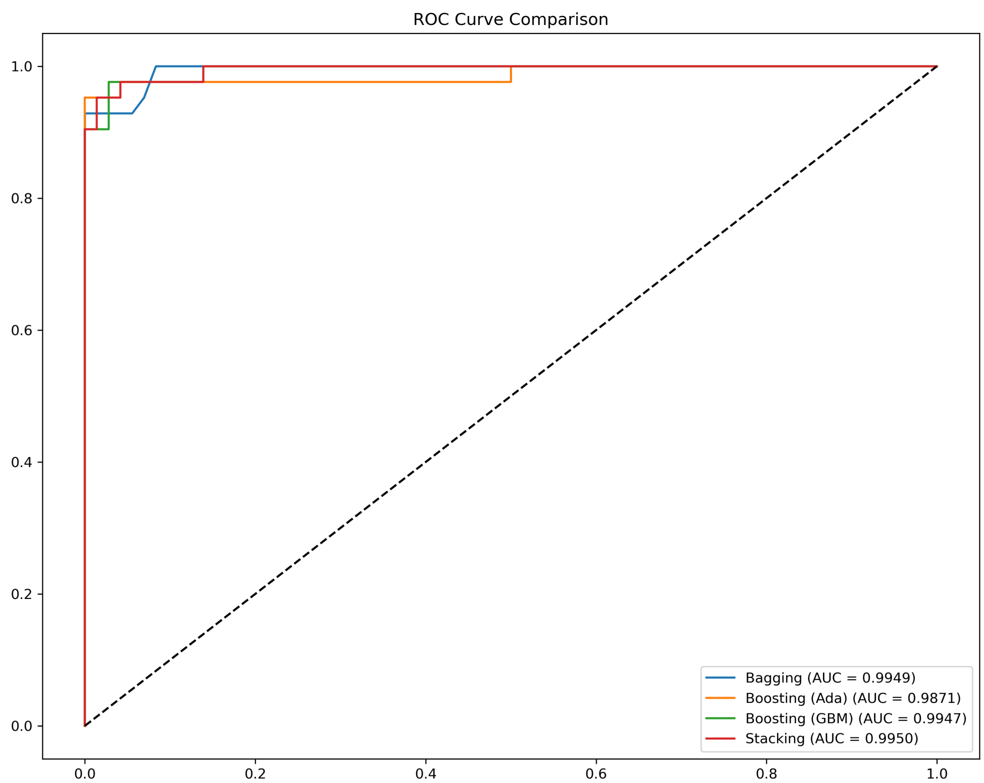

Figure 4: ROC Curve Comparison

Evaluation Metrics

- Accuracy
- Precision
- Recall
- F1-score
- Confusion Matrix
- ROC Curve and AUC

Observations

- **How does Bagging reduce variance?**
  By generating multiple decision trees trained on random subsets of the data (bootstrapping) and considering random subsets of features, it ensures that trees are largely uncorrelated. Averaging their predictions dramatically reduces the model's overall variance.

- **How does Boosting address model bias?**
  Boosting trains learners sequentially, where each new tree focuses primarily on the errors (residuals or misclassifications) made by the previous trees. By continually trying to fit the hard-to-predict instances, it increases the complexity of the final model, effectively reducing bias.

- **Why does stacking benefit from heterogeneous models?**
  Different types of base models (e.g., SVM, Naïve Bayes, Decision Tree) capture fundamentally different patterns and relationships in the data. A meta-learner can learn which model's predictions to trust for specific types of instances, often outperforming any single base learner.

- **Which ensemble method performed best and why?**
  Based on the results, **AdaBoost** performed best on the test set, achieving the highest Accuracy (97.37%) and F1 Score (0.9630). Boosting and Stacking typically perform best on the WDBC dataset due to the richness of the features and the relatively small sample size, where stability and bias correction are both vital.

Conclusion

Bagging, Boosting, and Stacked Ensemble models were implemented and evaluated on the WDBC dataset. The experiment demonstrates that ensemble strategies provide significant improvements in predictive performance, accuracy, and generalization compared to single models by effectively managing the bias-variance tradeoff. AdaBoost proved to be the most optimal ensemble configuration for this specific diagnostic task.

References

- Scikit-learn: Ensemble Methods
- Bagging Classifier
- UCI Dataset: Breast Cancer Wisconsin (Diagnostic)