Ensemble Learning Report

Mahfuzzaman Sizan , 1905054

September 20, 2024

1 Instructions to Run the Script

To train and test the models, follow these steps:

1. We can install the required libraries using:

```
pip install pandas scikit-learn matplotlib
```

2. We need to download the dataset files and place them in the same directory as the script.

3. We need to open the script in preferred Python IDE (e.g., Jupyter Notebook, PyCharm).

4. To train models on a specific dataset, we need to comment out the lines loading the other datasets. For example:

```
# Load Dataset 1
# X_train1, X_test1, y_train1, y_test1 = load_and_preprocess_dataset_one()

# Load Dataset 2
# X_train2, X_test2, y_train2, y_test2 = load_and_preprocess_dataset_two()

# Load Dataset 3
X_train3, X_test3, y_train3, y_test3 = load_and_preprocess_dataset_three()
```

5. After training,we need to run the evaluation function to collect metrics:

```
metrics3 = collect_metrics(X_train3, X_test3, y_train3, y_test3)
```

Table 1: Performance Metrics for Dataset 1

Metric	Dataset 1 (Bagging LR)	Dataset 1 (Voting)	Dataset 1 (Stacking)
Accuracy	0.787950	0.787083	0.786373
Precision	0.636459	0.634058	0.637736
Recall	0.469400	0.467914	0.451872
F1 Score	0.540105	0.538462	0.528951
ROC AUC	0.686230	0.685165	0.679559
PR AUC	0.439547	0.437920	0.433668

2 Performance Evaluation Results

2.1 Dataset 1 Results

2.2 Dataset 2 Results

Table 2: Performance Metrics for Dataset 2

Metric	Dataset 2 (Bagging LR)	Dataset 2 (Voting)	Dataset 2 (Stacking)
Accuracy	0.826676	0.826362	0.825809
Precision	0.717861	0.720658	0.720910
Recall	0.438753	0.432657	0.428497
F1 Score	0.544597	0.540699	0.537508
ROC AUC	0.692705	0.690394	0.688595
PR AUC	0.447533	0.445819	0.443912

2.3 Dataset 3 Results

Table 3: Performance Metrics for Dataset 3

Metric	Dataset 3 (Bagging LR)	Dataset 3 (Voting)	Dataset 3 (Stacking)
Accuracy	0.994877	0.994877	0.976092
Precision	0.987342	0.987342	0.000000
Recall	0.795918	0.795918	0.000000
F1 Score	0.881356	0.881356	0.000000
ROC AUC	0.897834	0.897834	0.500000
PR AUC	0.790723	0.790723	0.023908

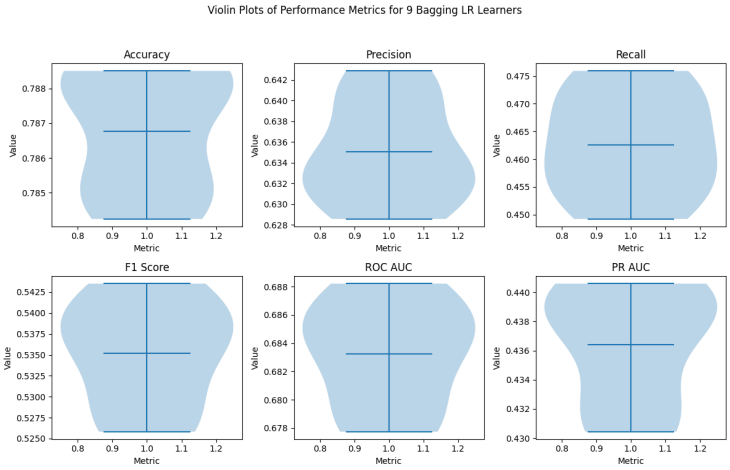

Figure 1: Violin Plot for Dataset 1

3 Violin Plots

4 Observations

- Dataset 3 achieved the highest accuracy across all models, indicating effective model performance.
- The precision and recall values for the stacking model in Dataset 3 were 0, suggesting poor performance in identifying positive cases.
- Dataset 1 and Dataset 2 showed similar trends in precision and recall, with Bagging LR performing slightly better.

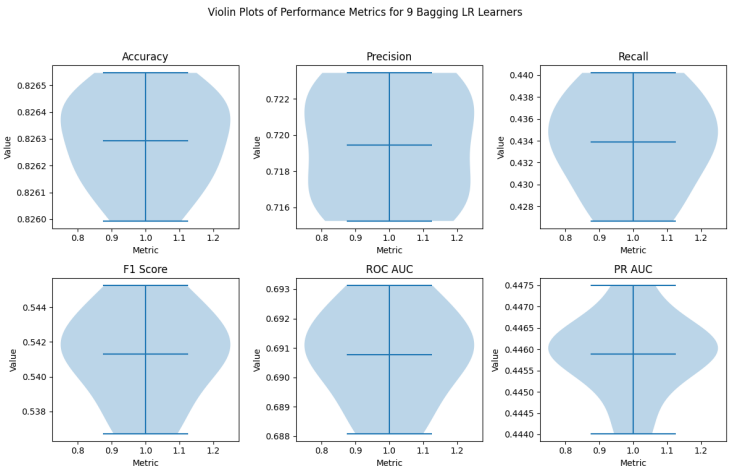

Figure 2: Violin Plot for Dataset 2

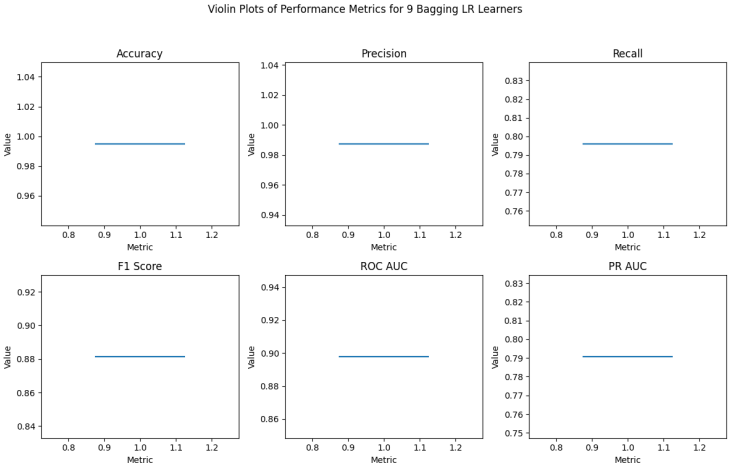

Figure 3: Violin Plot for Dataset 3