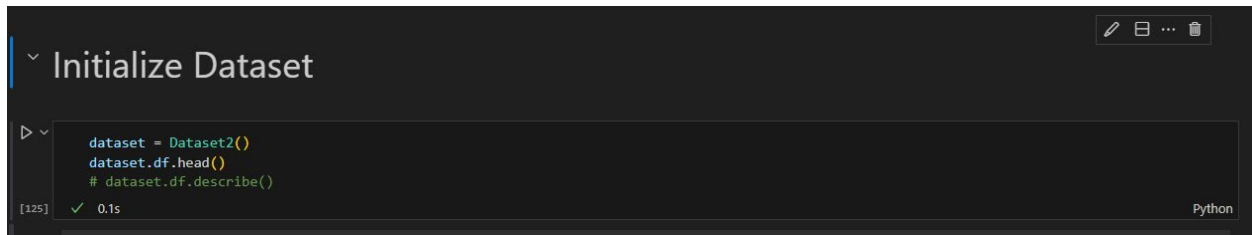# Offline2 Report

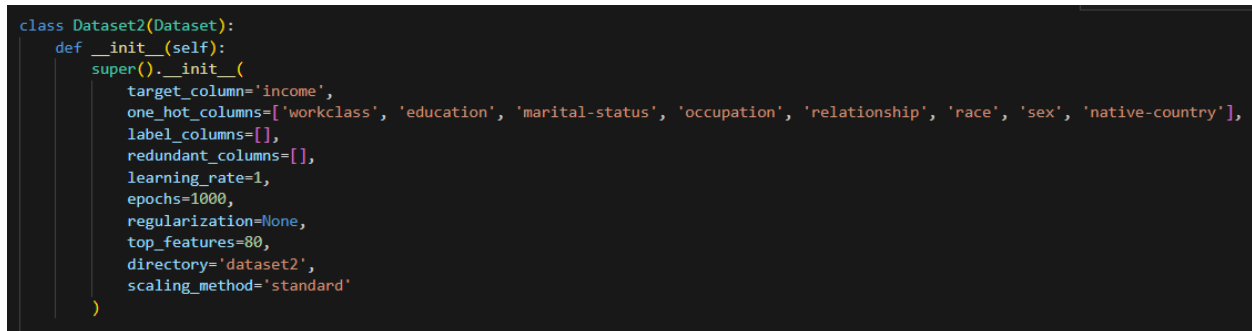1905072 – Mahir Labib Dihan

Running the Code

1. Install Packages: **pip install -r requirements.txt**
2. Choosing the desired dataset to run.
   - There are Classes defined for each dataset. So, we need to create object of our desired class.
   - For example, if we want to run dataset 2, just assign Dataset2 object to dataset variable like below. That's it.

   - One thing to make sure that the dataset directory in each Class is set correctly. For example, in my machine dataset2 is in **'dataset2'** folder of my working directory.

```python
class Dataset2(Dataset):
    def __init__(self):
        super().__init__(
            target_column='income',
            one_hot_columns=['workclass', 'education', 'marital-status', 'occupation', 'relationship', 'race', 'sex', 'native-country'],
            label_columns=[],
            redundant_columns=[],
            learning_rate=1,
            epochs=1000,
            regularization=None,
            top_features=80,
            directory='dataset2',
            scaling_method='standard'
        )
```

Dataset1 Evaluation

Scaling	Top Features	Learning Rate	Epochs
standard	20	1	1000

	Accuracy	Sensitivity	Specificity	Precision	F_1-score	AUROC	AUPR
LR*	0.8114 ± 0.0034	0.5623 ± 0.0180	0.8925 ± 0.0075	0.6303 ± 0.0113	0.5941 ± 0.0088	0.8575 ± 0.0014	0.6238 ± 0.0056
Voting ensemble	0.815658	0.568116	0.896226	0.640523	0.602151	0.858994	0.625409
Stacking ensemble	0.818505	0.602899	0.886792	0.638037	0.61997	0.857312	0.61891

```
dataset1 Performance Metric    Accuracy         Sensitivity      Specificity      Precision        F1-Score         AUROC            AUPR
LR                             0.814235         0.568116         0.89434          0.636364         0.600306         0.858884         0.628012
LR*               0.8114 ± 0.0034  0.5623 ± 0.0180  0.8925 ± 0.0075  0.6303 ± 0.0113  0.5941 ± 0.0088  0.8575 ± 0.0014  0.6238 ± 0.0056
Voting Ensemble                0.815658         0.568116         0.896226         0.640523         0.602151         0.858994         0.625409
Stacking Ensemble              0.818505         0.602899         0.888679         0.638037         0.61997          0.857312         0.61891
```
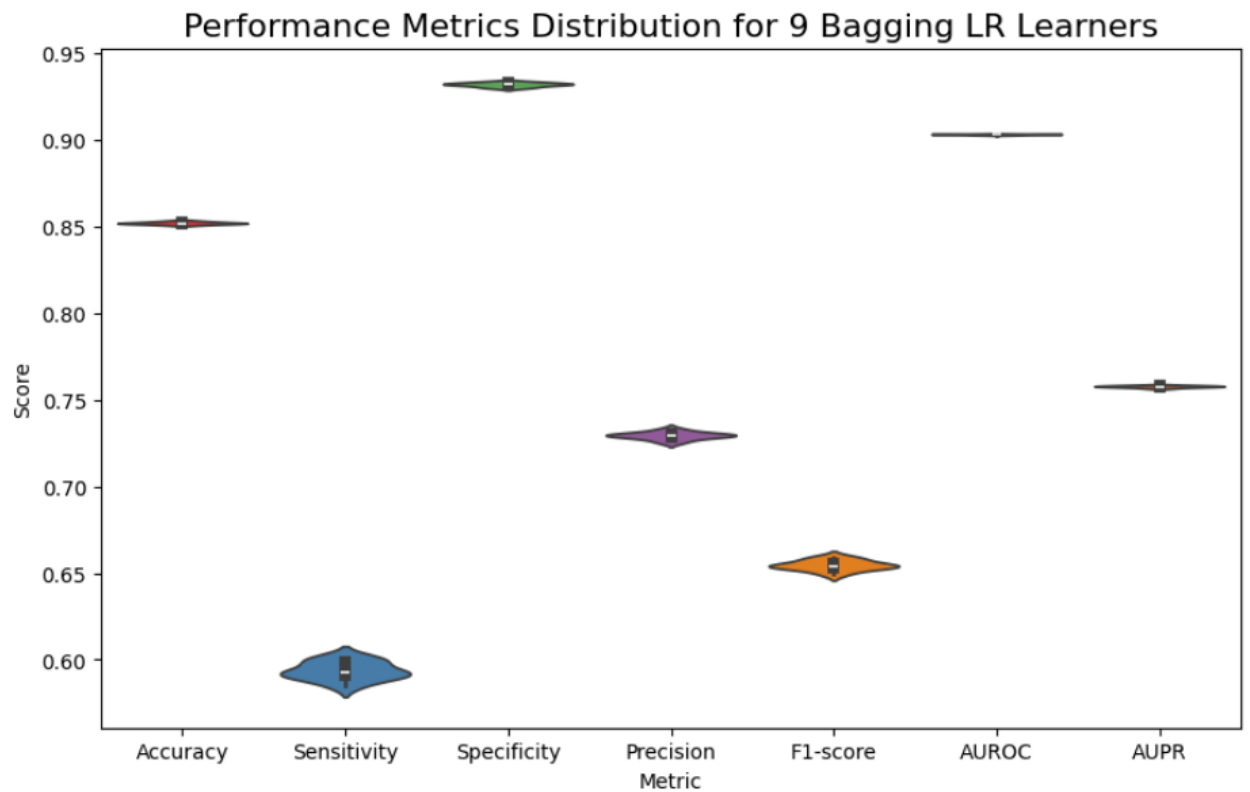

Performance Metrics Distribution for 9 Bagging LR Learners
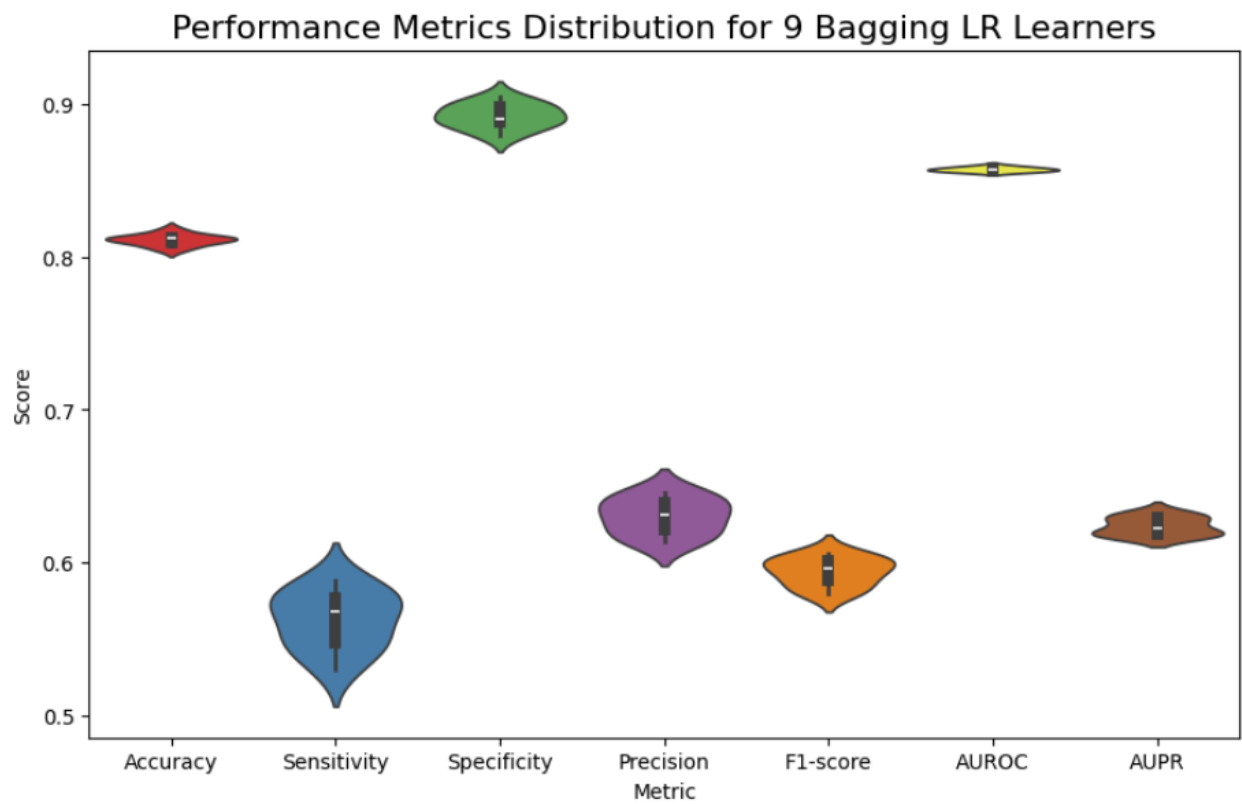
Dataset2 Evaluation

Scaling	Top Features	Learning Rate	Epochs
standard	50	1	1000

	Accuracy	Sensitivity	Specificity	Precision	F_1-score	AUROC	AUPR
LR*	0.8519 ± 0.0007	0.5938 ± 0.0049	0.9318 ± 0.0011	0.7293 ± 0.0019	0.6546 ± 0.0027	0.9031 ± 0.0002	0.7577 ± 0.0005
Voting ensemble	0.852298	0.593864	0.932261	0.730646	0.655192	0.903403	0.758316
Stacking ensemble	0.849226	0.572803	0.934755	0.730922	0.642274	0.902178	0.753568

```
dataset2 Performance Metric      Accuracy      Sensitivity      Specificity        Precision         F1-Score            AUROC            AUPR
LR                              0.852359        0.591004         0.933226         0.732517         0.654195         0.903572         0.75842
LR*                 0.8519 ± 0.0007  0.5938 ± 0.0049  0.9318 ± 0.0011  0.7293 ± 0.0019  0.6546 ± 0.0027  0.9031 ± 0.0002  0.7577 ± 0.0005
Voting Ensemble                 0.852298        0.593864         0.932261         0.730646         0.655192         0.903403         0.758316
Stacking Ensemble               0.849226        0.572803         0.934755         0.730922         0.642274         0.902178         0.753568
```
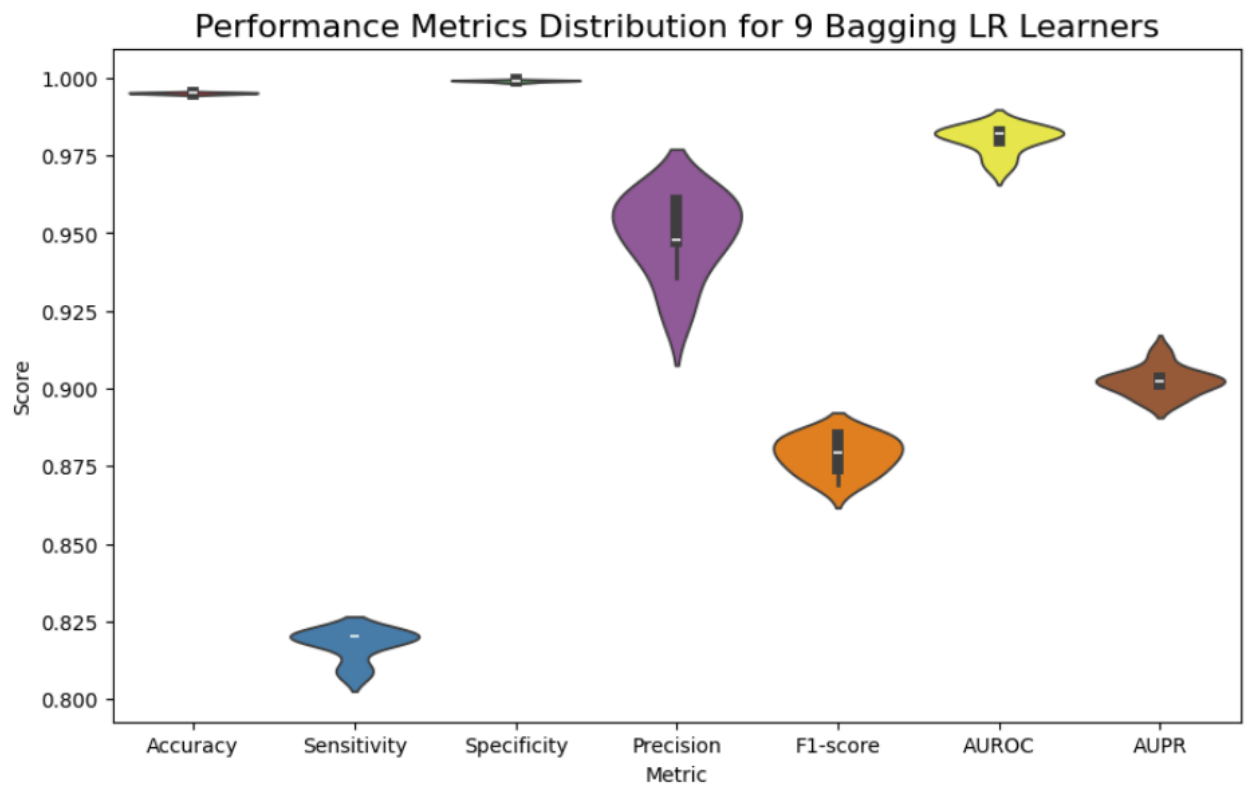

Performance Metrics Distribution for 9 Bagging LR Learners

Dataset3 Evaluation

Scaling	Top Features	Learning Rate	Epochs
standard	20	1	1000

	Accuracy	Sensitivity	Specificity	Precision	F_1-score	AUROC	AUPR
LR*	0.9951 ± 0.0002	0.8177 ± 0.0047	0.9990 ± 0.0002	0.9494 ± 0.0121	0.8786 ± 0.0054	0.9802 ± 0.0041	0.9024 ± 0.0042
Voting ensemble	0.995357	0.820225	0.999251	0.960526	0.884848	0.981932	0.90616
Stacking ensemble	0.994624	0.808989	0.998751	0.935065	0.86747	0.965888	0.863091

```
dataset3 Performance Metric     Accuracy        Sensitivity      Specificity       Precision          F1-Score            AUROC              AUPR
LR                              0.995112        0.820225         0.999001          0.948052           0.879518           0.983967           0.908289
LR*                 0.9951 ± 0.0002  0.8177 ± 0.0047  0.9990 ± 0.0002  0.9494 ± 0.0121  0.8786 ± 0.0054  0.9802 ± 0.0041  0.9024 ± 0.0042
Voting Ensemble                 0.995357        0.820225         0.999251          0.960526           0.884848           0.981932           0.90616
Stacking Ensemble               0.994624        0.808989         0.998751          0.935065           0.86747            0.965888           0.863091
```


Performance Metrics Distribution for 9 Bagging LR Learners

Observation

For all the datasets we can see that Sensitivity is lowest. One solution for that I observed is to decrease the threshold of logistic regression. In Logistic regression we classify based on the output of sigmoid function. If the value ≥ 0.5, then it is in Class 1, otherwise Class 0. If we decrease it to be less than 0.5 like 0.32 then Sensitivity increases. But this increases the false positive rate and decreases precision.
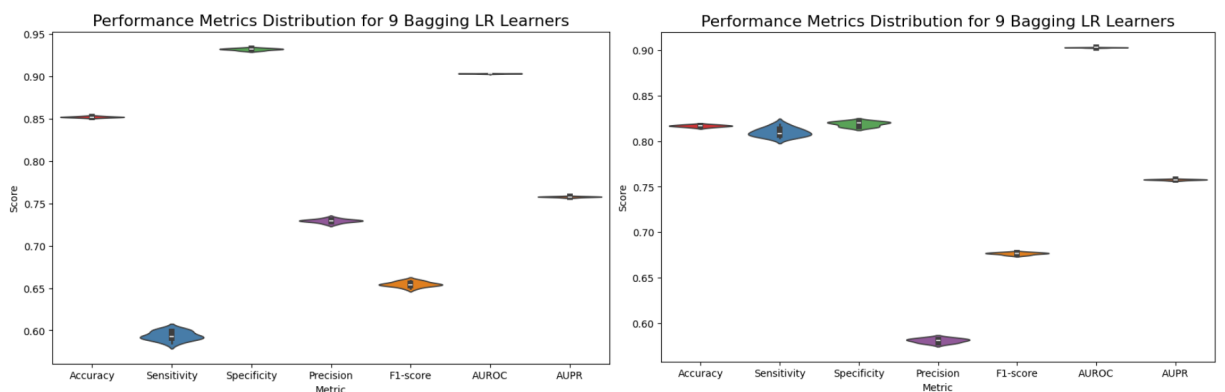
Dataset 1: Threshold 0.5 vs 0.32

dataset1 Performance Metric	Accuracy	Sensitivity	Specificity	Precision	F1-Score	AUROC	AUPR
LR	0.77153	0.794203	0.764151	0.522901	0.63061	0.858884	0.628012
LR*	0.7730 ± 0.0061	0.7945 ± 0.0152	0.7659 ± 0.0115	0.5252 ± 0.0091	0.6322 ± 0.0053	0.8575 ± 0.0014	0.6238 ± 0.0056
Voting Ensemble	0.772242	0.8	0.763208	0.523719	0.633028	0.858994	0.625409
Stacking Ensemble	0.769395	0.802899	0.758491	0.5197	0.630979	0.857104	0.618945
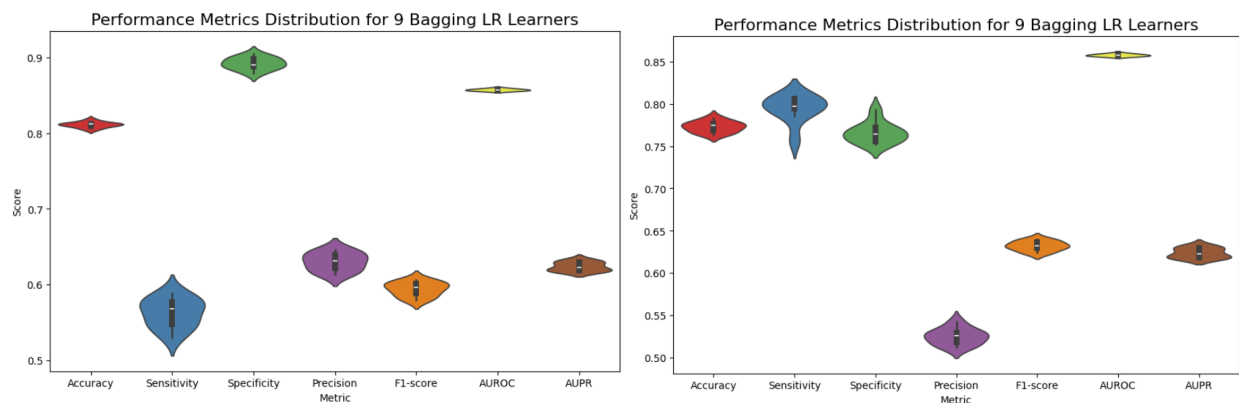
Dataset 2: Threshold 0.5 vs 0.28

dataset2 Performance Metric	Accuracy	Sensitivity	Specificity	Precision	F1-Score	AUROC	AUPR
LR	0.818506	0.807072	0.822043	0.583898	0.677581	0.903572	0.75842
LR*	0.8169 ± 0.0010	0.8101 ± 0.0044	0.8190 ± 0.0024	0.5807 ± 0.0022	0.6765 ± 0.0011	0.9031 ± 0.0002	0.7577 ± 0.0005
Voting Ensemble	0.817523	0.812533	0.819067	0.581504	0.677874	0.903403	0.758316
Stacking Ensemble	0.817707	0.802392	0.822446	0.583034	0.675347	0.901983	0.753536

Dataset 3: Threshold 0.5 vs 0.2

dataset3 Performance Metric	Accuracy	Sensitivity	Specificity	Precision	F1-Score	AUROC	AUPR
LR	0.995357	0.865169	0.998251	0.916667	0.890173	0.983967	0.908289
LR*	0.9950 ± 0.0004	0.8539 ± 0.0130	0.9981 ± 0.0005	0.9102 ± 0.0196	0.8809 ± 0.0082	0.9802 ± 0.0041	0.9024 ± 0.0042
Voting Ensemble	0.995601	0.865169	0.998501	0.927711	0.895349	0.981932	0.90616
Stacking Ensemble	0.994379	0.853933	0.997502	0.883721	0.868571	0.967305	0.860423

Performance Metrics Distribution for 9 Bagging LR Learners