HOW TO RUN:

1) Install these libraries:

```
requirements.txt
1    pandas
2    numpy
3    matplotlib
4    scikit-learn
5    ucimlrepo
     ...
6    seaborn
7
8
```

2) Uncomment for the corresponding datasets or ensembling

```python
#categorical_df, numerical_df,target_column_name = dataset_3_preprocessing()
categorical_df, numerical_df,target_column_name = dataset_1_preprocessing()
#categorical_df, numerical_df,target_column_name = dataset_2_preprocessing()
categorical_df.info()
numerical_df.info()
```
✓ 1.1s

```python
# # #predict with stack ensembling--------------------------------------------------------
# y_test,y_pred, y_pred_prob = predict_with_stack_ensembling(scaled_and_encoded_df, target_column_name, n_models=9)

# # get the accuracy score

# evaluate_model('StackEnsemble',y_test, y_pred, y_pred_prob,save_scores=True)


                                                                          (parameter) soft_voting:
# predict with voting ensembling---------------------------------------------------------
y_test,y_pred, y_pred_prob = predict_with_voting_ensembling(scaled_and_encoded_df, target_column_name, n_models=9,soft_voting=True)

# get the accuracy score

evaluate_model('VotingEnsemble',y_test, y_pred, y_pred_prob,save_scores=True)

violin_plot()

# delete the scores file
import os
#os.remove('LR_scores.csv')
```
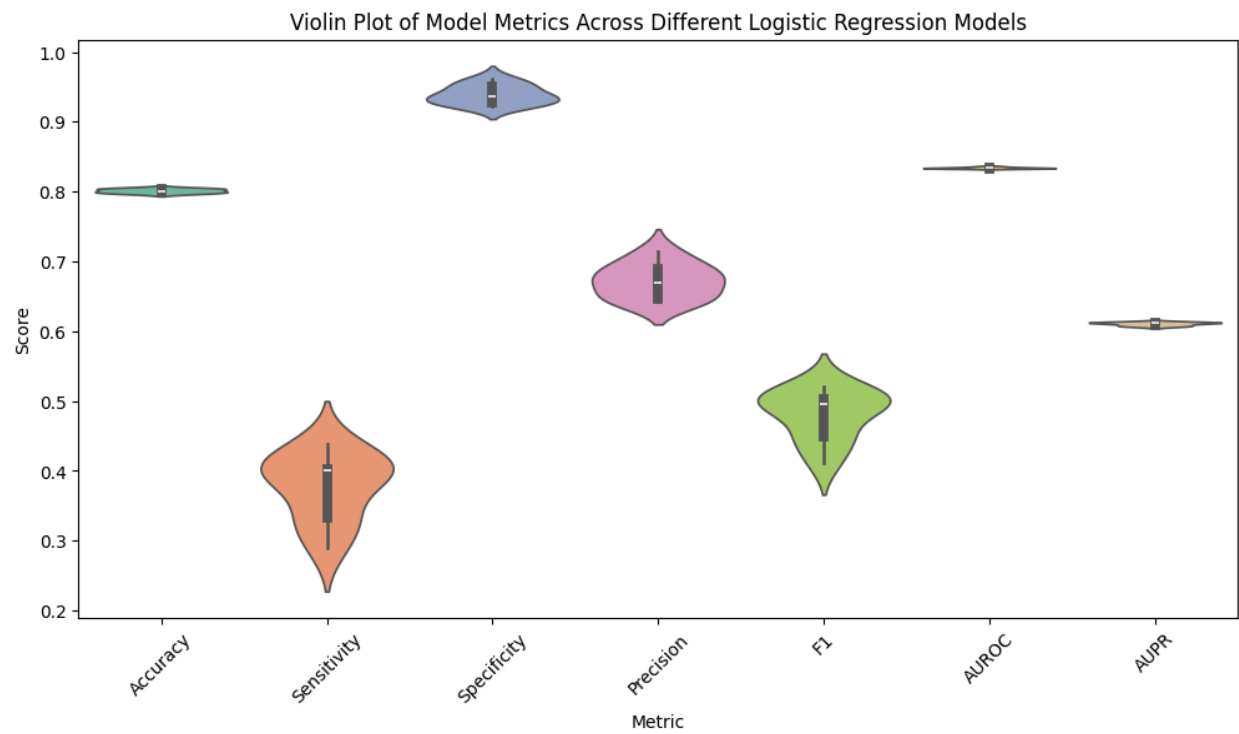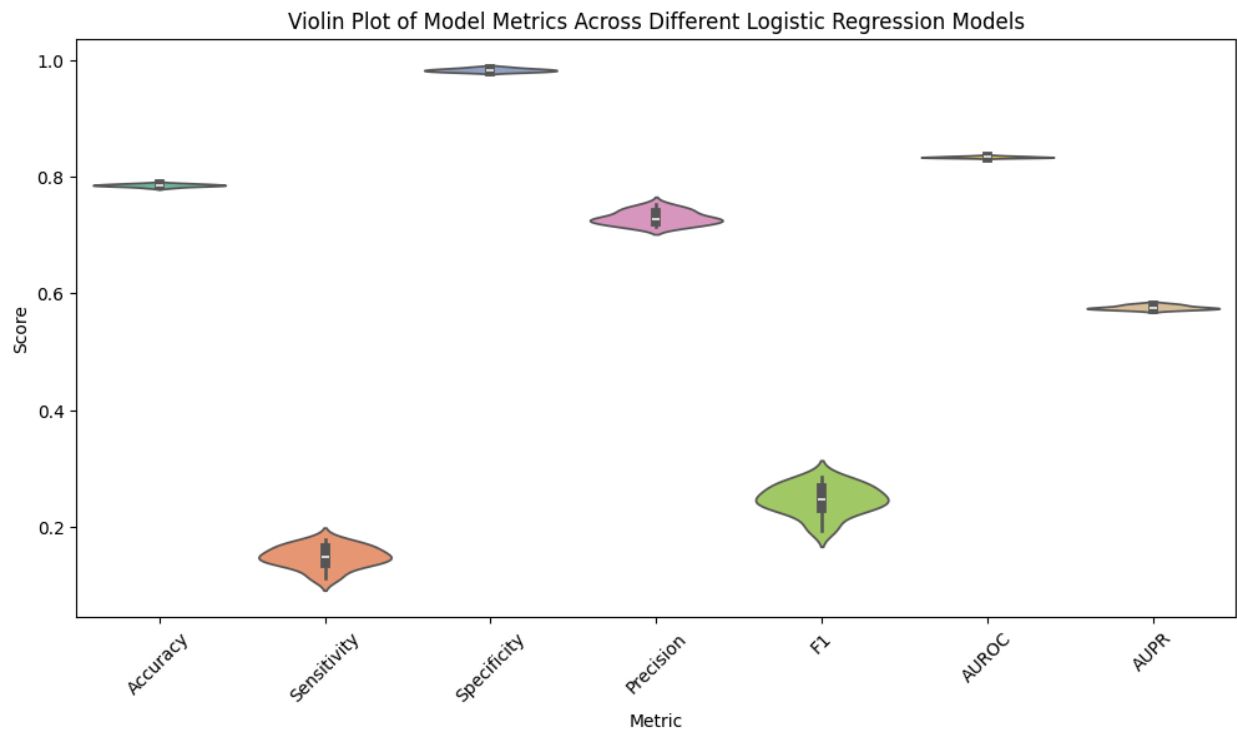
DATASET-1

Model	Accuracy	Sensitivity	Specificity	Precision	F1	AUROC	AUPR
LR	0.80 ± 0.00	0.38 ± 0.05	0.94 ± 0.01	0.67 ± 0.02	0.48 ± 0.04	0.83 ± 0.00	0.61 ± 0.00
StackEnsemble	0.7914590747	0.4144927536	0.9141509434	0.6111111111	0.493955095	0.8402351654	0.6020901973
VotingEnsemble	0.8042704626	0.3855072464	0.9405660377	0.6785714286	0.4916820702	0.8342083675	0.6103738052

Violin Plot of Model Metrics Across Different Logistic Regression Models

DATASET-2

Model	Accuracy	Sensitivity	Specificity	Precision	F1	AUROC	AUPR
LR	0.78 ± 0.01	0.13 ± 0.03	0.98 ± 0.02	0.76 ± 0.05	0.22 ± 0.06	0.81 ± 0.01	0.58 ± 0.01
StackEnsemble	0.7766960443	0.1225108225	0.9795918367	0.6505747126	0.2061930783	0.8134347639	0.5271785048
VotingEnsemble	0.7773109244	0.1255411255	0.9794575725	0.6546275395	0.210679259	0.8142744094	0.5277588407

Violin Plot of Model Metrics Across Different Logistic Regression Models

DATASET-3

Model	Accuracy	Sensitivity	Specificity	Precision	F1	AUROC	AUPR
LR	0.98 ± 0.00	0.00 ± 0.00	1.00 ± 0.00	nan ± nan	nan ± nan	0.90 ± 0.00	0.83 ± 0.00
StackEnsemble	0.9777723498	0	1			0.9021832527	0.8272443622
VotingEnsemble	0.9777723498	0	1			0.9020487382	0.8273617452

Violin Plot of Model Metrics Across Different Logistic Regression Models