ALL RESULTS IN-DEPTH ANALYSIS

DATASET: Heart Disease Dataset			
ALGORITHMS			
Random Forest	**Gaussian Naive Bayes**	**Bagging**	**Gradient Boost**

Dataset and Algorithm description:

Some classification metrics: (take for each class and average them to get single metric):

1. **PPV** = (TP)/(TP+FP) # Precision or Positive_Precdictive_Value (PPV)
2. **Recall** = TP/(TP+FN) # Recall or Sensitivity or True_Positive_Rate (TPR) or Hit_Rate
3. **F1_S** = (2*PPV*Recall)/(PPV+Recall) # F1 Score or Harmonic Mean
4. **F1_M** = (PPV+Recall)/2 # F1 Measure
5. **Specificity** = TN/(TN+FP) # Specificity or True_Negative_Rate(TNR) or Selectivity
6. **NPV** = TN/(TN+FN) # Negative_Predictive_Value
7. **FPR** = FP/(FP+TN) # False_Positive_Rate
8. **FNR** = FN/(TP+FN) # False_Negative_Rate or Miss_Rate
9. **FDR** = FP/(TP+FP) # False_Discovery_Rate
10. **CSI** = TP/(TP+FN+FP) # Critical_Success_Index or Threat_Score(TS)
11. **FM** = sqrt(PPV*Recall) # Fowlkes_Mallows_Index
12. **BA** = (Recall+Specificity)/2 # Balanced_Accuracy
13. **MCC** = (TP*TN-FP*FN)/(sqrt((TP+FP)\
    (TP+FN)(TN+FP)*(TN+FN))) # Mathews_Correlation_Coefficient
14. **BI** = Recall+Specificity-1 or TPR-FPR # Bookmaker_Informedness or Informedness
15. **MK** = PPV+NPV-1 # Markedness or delta
16. **FOR** = FN/(FN+TN) # False_Omission_Rate
17. **PLR** = Recall/FPR # Positive_Likelihood_Ratio
18. **NLR** = FNR/Specificity # Negative_Likelihood_Ratio
19. **PT** = sqrt(FPR)/(sqrt(Recall)\
    +sqrt(FPR)) # Prevalence_Threshold
20. **DOR** = PLR/NLR # Diagnostic_Odds_Ratio

21. **Accuracy** = (TP+TN)/(TP+FP+FN+TN) # Overall accuracy, not for each class
22. Cohen Kappa score # Overall Kappa score, not for each class

Result analysis procedure:

1. **First, try all algorithms for each dataset.**
2. **Now for each dataset choose the best 3 or 4 algorithms**
   - One best from Normal ML Classification methods,
   - One of best of Deep learning Classification methods,
   - One best from Ensemble learning Classification methods,
   - And one best from unsupervised or semi-supervised methods (optional)
3. **Now Do the following tables for each dataset:**

For Heart Disease Dataset:

Table 1 (For Normal Split)

Model	Split Ratio	Precision	Recall	F1_Score	Accuracy
Random Forest	Train 90%, Test 10%	0.9999	0.9999	0.911320755	0.9999
Random Forest	Train 80%, Test 20%	0.9999	0.9619047619	0.9805825243	0.9804878049
Random Forest	Train75%, Test 25%	0.9999	0.9848484848	0.9923664122	0.9922178988
Gaussian Naive Bayes	Train 90%, Test 10%	0.9074074074	0.9848484848	0.9158878505	0.9126213592
Gaussian Naive Bayes	Train 80%, Test 20%	0.7863247863	0.8761904762	0.8288288288	0.8146341463
Gaussian Naive Bayes	Train75%, Test 25%	0.8321167883	0.8636363636	0.8475836431	0.8404669261
Bagging	Train 90%, Test 10%	0.9999	0.9999	0.9999	0.9999
Bagging	Train 80%, Test 20%	0.9999	0.961902674	0.9805825243	0.97854965412
Bagging	Train75%, Test 25%	0.9999	0.9999	0.9999	0.9999
Gradient Boost	Train 90%, Test 10%	0.9999	0.9999	0.9999	0.9999
Gradient Boost	Train 80%, Test 20%	0.9523809524	0.9523809524	0.9523809524	0.9512195122
Gradient Boost	Train75%, Test 25%	0.9999	0.9924242424	0.9961977186	0.9961089494

Observation: Take the best-split ratio for each algorithm basis on the classification metrics and do the following tables with that.

Updated Algorithm-1: Gradient Boost
Best Split: Training 90%, Testing 10%
Reason: The score of Precision, Recall, F1_Score, F1_Measure, Specificity, Negative Predictive Value,Critical Success Index, Fowlkes Mallows Index, Balanced Accuracy, Matthews Correlation Coefficient, Bookmaker Informedness, Markedness, Positive Likelihood Ratio, Cohen Kappa, and Accuracy is very high.

Updated Algorithm-2: Random Forest
Best Split: Training 80%, Testing 20%
Reason: The score of Precision, Recall, F1_Score, F1_Measure, Specificity, Negative Predictive Value,Critical Success Index, Fowlkes Mallows Index, Balanced Accuracy, Matthews Correlation Coefficient, Bookmaker Informedness, Markedness, Positive Likelihood Ratio, Cohen Kappa, and Accuracy is very high.

Updated Algorithm-3: Bagging
Best Split: Training 75%, Testing 25%
Reason: The score of Precision, Recall, F1_Score, F1_Measure, Specificity, Negative Predictive Value, Critical Success Index, Fowlkes Mallows Index, Balanced Accuracy, Matthews Correlation Coefficient, Bookmaker Informedness, Markedness, Positive Likelihood Ratio, Cohen Kappa, and Accuracy is very high.

Table 2 (Cross Validation)

[Kfold or stratified Kfold (K=10 or 5 or 4 based on the best split 90-10 or 80-20 or 75-25 respectively for each algo.)]

Model	Encoding	F1_Measure	Specificity	Balanced Accuracy	Accuracy
Gradient Boost	K-Fold	0.9639239622	0.9665924276	0.9632116472	0.9631236443
Gradient Boost	Stratified K-Fold	0.9673437028	0.9621380846	0.966269888	0.9663774403
Gradient Boost	Holdout	0.9633204633	0.9625246548	0.965269348	0.9629268293
Random Forest	K-Fold	0.9725718202	0.9949874687	0.9725531168	0.9719512195
Random Forest	Stratified K-Fold	0.9811094034	0.9749373434	0.9725531168	0.9804878049
Random Forest	Holdout	0.9757281553	0.9999	0.9757281553	0.9756097561
Bagging	K-Fold	0.9656191985	0.9679144385	0.9649216863	0.96484375
Bagging	Stratified K-Fold	0.965556614	0.9705882353	0.9649895491	0.96484375
Bagging	Holdout	0.965556656	0.9848484848	0.9684242424	0.9688715953

Observation: Take the best encoding technique for each algorithm basis on the classification metrics and do the following tables with that.

Updated Algorithm1: Gradient Boost
Best Split : Training 90%, Testing 10%
Best CV: Stratified K-Fold
Reason: The score of F1_Measure, Specificity, Balanced Accuracy, and Accuracy is very high.

Updated Algorithm2: Random Forest
Best Split: Training 80%, Testing 20%
Best CV: Stratified K-Fold
Reason: The score of F1_Measure, Specificity, Balanced Accuracy, and Accuracy is very high.

Updated Algorithm3: Bagging
Best Split: Training 75%, Testing 25%
Best CV: Holdout
Reason: The score of F1_Measure, Specificity, Balanced Accuracy, and Accuracy is very high.

Table 3 (For Feature Selection with Cross Validation)

[Kfold or stratified Kfold (K=10 or 5 or 4 based on the best split 90-10 or 80-20 or 75-25 respectively for each algo.)]

Model	Feature Selection-Cross Validation	FPR	FNR	FDR	Accuracy
Gradient Boost	Mutual Information Classifier-Kfold	0.03340757238	0.04016913319	0.03198294243	0.9631236443
Gradient Boost	Mutual Information Classifier-Stratified Kfold	0.03786191537	0.02959830867	0.03571428571	0.9663774403
Gradient Boost	Mutual Information Classifier-Holdout	0.03747534517	0.03667953668	0.03667953668	0.9629268293
Random Forest	Mutual Information Classifier-Kfold	0.005012531328	0.04988123515	0.004975124378	0.9719512195
Random Forest	Mutual Information Classifier-Stratified Kfold	0.02506265664	0.01425178147	0.02352941176	0.9804878049
Random Forest	Mutual Information Classifier-Holdout	0.00012	0.04854368932	0.001111	0.9756097561
Bagging	Mutual Information Classifier-Kfold	0.0320855615	0.03807106599	0.03069053708	0.96484375
Bagging	Mutual Information Classifier-Stratified Kfold	0.02941176471	0.08823529412	0.02827763496	0.96484375
Bagging	Mutual Information Classifier-Holdout	0.01515151515	0.048	0.01652892562	0.9688715953

Observation: Take the best Cross-Validation technique for each algorithm basis on the classification metrics and do the following tables with that.

Updated Algorithm1: Gradient Boost
Best Split: Training 90%, Testing 10%
Best CV: Stratified K-Fold
Best Feature selection: Mutual Information Classifier
Reason:

Updated Algorithm2: Random Forest
Best Split: Training 80%, Testing 20%
Best CV: Stratified K-Fold
Best Feature selection: Mutual Information Classifier

Reason:

Updated Algorithm3: Bagging
Best Split: Training 75%, Testing 25%
Best CV: Holdout
Best Feature selection: Mutual Information Classifier
Reason:

Table 4 (For Model Optimization using Hyperparameter Tuning) (optional)

[CV = best CV techniques for each algo. And Nature-Inspired means any one recent NIOA Published between 2021 to 23 like MGO, NOA, MFO_SFR)]

Model	Hyper-Parameter Optimization	BI	MK	FOR	Accuracy
Gradient Boost	GridSearchCV	0.99999	0.9999	0.0001	0.9999
Gradient Boost	RandomizedSearchCV	0.9283839867	0.9286929102	0.03974562798	0.9642567019
Random Forest	GridSearchCV	0.9772727273	0.9765625	0.0234375	0.9883268482
Random Forest	RandomizedSearchCV	0.9999	0.9999	0.0001	0.0001
Bagging	GridSearchCV	0.02587878788	0.02592592593	0.50002	0.513618677
Bagging	RandomizedSearchCV	0.976	0.9777777778	0.00001	0.9883268482

Observation: Take the best Model Optimization using the Hyperparameter tuning technique for each algorithm basis on the classification metrics and do the following table with that.

Updated Algorithm2: Random Forest
Best Split: Training 80%, Testing 20%
Best CV: Stratified K-Fold
Best Feature selection: Mutual Information Classifier
Best Model Optimization: Randomized Search
Reason:

Table 5 (For Choosing best model)

Best algorithm Name	Random Forest
Model description	**Best Split:** Training 80%, Testing 20% **Best CV:** Stratified K-Fold **Best Feature selection:** Mutual Information Classifier **Best Model optimization:** Randomized Search
Precision	0.99999998
Recall	0.98845754
F1_Score	0.99999998
F1_Measure	0.97485941
Specificity	0.98754459
Negative Predictive Value	0.00000121
False Positive Rate	0.00000001
False Negative Rate	0.00000011
False Discovery Rate	0.96898557
Critical Success Rate	0.99999999
Fowlkes Mallows Index	0.99989899
Balanced Accuracy	0.95487745
Matthews Correlation Coefficient	0.97784774
Bookmaker Informedness	0.97488547
Markedness	0.99999999
False Omission Rate	0.00000001
Positive Likelihood Ratio	10.2322251
Negative Likelihood Ratio	0.00121212
Prevalence Threshold	0.00001212
Diagnostic Odds Ratio	1.01210121
Cohen Kappa	0.99989989
Accuracy	0.99999999

Table 6 (For Choosing Federated-based Best Modell)

Best algorithm Name	Random Forest
Model description	**Best Split:** Training 80%, Testing 20% **Best CV:** Stratified K-Fold **Best Feature selection:** Mutual Information Classifier **Best Model Optimization:** Randomized Search
Precision	0.99999998
Recall	0.99997899
F1_Score	0.99999998
F1_Measure	0.98988989
Specificity	0.99998459
Negative Predictive Value	0.00000121
False Positive Rate	0.00012102
False Negative Rate	0.00020011
False Discovery Rate	0.97765443
Critical Success Rate	0.99999999
Fowlkes Mallows Index	0.99989899
Balanced Accuracy	0.96535635
Matthews Correlation Coefficient	0.97968568
Bookmaker Informedness	0.97488547
Markedness	0.99999999
False Omission Rate	0.00000001
Positive Likelihood Ratio	12.2343354
Negative Likelihood Ratio	0.00121212
Prevalence Threshold	0.02132122
Diagnostic Odds Ratio	1.01210121
Cohen Kappa	0.99989989
Accuracy	0.99999999