JADBio Description of Performed Analysis

Setup

JADBio version **1.4.174** ran on dataset **diabetes_prediction_dataset** with **100000** samples and **8** features to create a predictive model for outcome named **diabetes**. The outcome was discrete leading to a **classification** modeling.

The preferences of the analysis were set to **true** for feature selection and **false** for full feature models tried.
The **BACC** metric was used to optimize for the best model.
The maximum number of features to select was set to **25**.
The effort to spend on tuning the algorithms were set to **Normal**.
The number of CPU cores to use for the analysis was set to **2**.
The execution time was **05:22:52**.

Configuration Space

JADBio's AI decide to try the following algorithms and tuning hyper-parameter values:

Algorithm Type	Algorithm	Hyper-parameter	Set of Values
Preprocessing	Mean Imputation		
	Mode Imputation		
	Constant Removal		
	Variable Normalization		
Feature Selection	Epilogi	stoppingThreshold	0.01
		stoppingCriterion	Independence Test
		equivalenceThreshold	0.01
	Test-Budgeted Statistically Equivalent Signature (SES)	maxK	2.0
		alpha	0.05, 0.01
	Univariate	maxVars	100
		alpha	0.001, 0.01
	LASSO	penalty	0.5, 1.5, 1.0
Modeling	Classification Random Forest with Deviance splitting criterion	nTrees	100, 500
		minLeafSize	4.0
	Classification Decision Tree with Deviance splitting criterion	alpha	0.05, 0.01
		minLeafSize	2, 4, 3

Leading to **113** combinations and corresponding configurations (machine learning pipelines) to try. For the full configurations tested see the Appendix.

Configuration Estimation Protocol

JADBio's AI system decided to estimate the out-of-sample performance of the models produced by each configuration using **90.00 % - % 10.00 hold-out.** Overall, 171 models were set out to train.

JADBio Results Summary

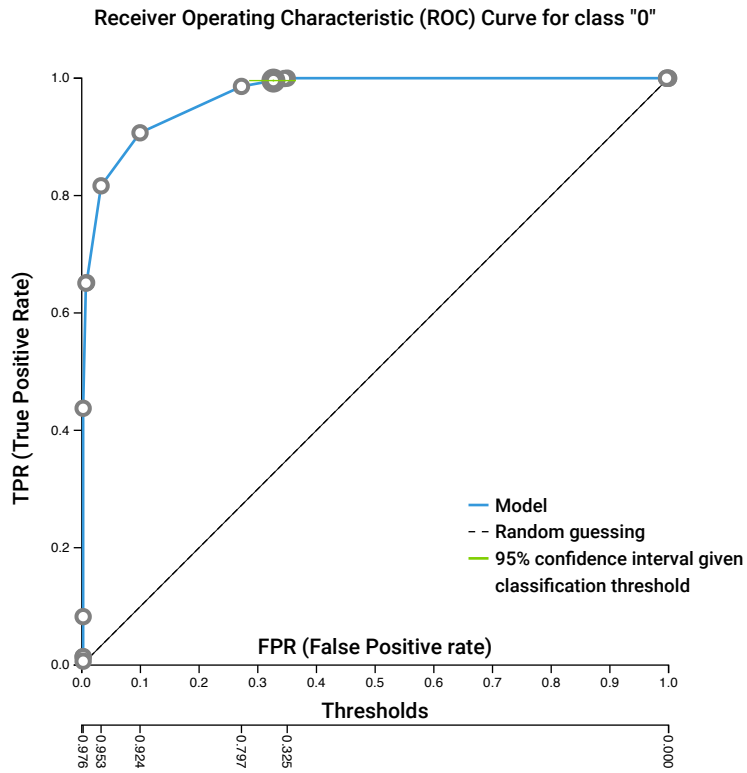
Overview

A result summary is presented for analysis optimized for Performance. The model is produced by applying the algorithms in sequence (configuration) on the training data:

Preprocessing	Feature Selection	Predictive algorithm
Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi algorithm with hyper-parameters: equivAlpha = 0.01, and stopping criterion = Independence Test with threshold: 0.01.	Classification Decision Tree with Deviance splitting criterion and hyper-parameters: minimum leaf size = 4, and pruning parameter alpha = 0.05

The **Area Under The Curve** is **0.971** with 95% confidence interval being [**0.963,0.978**].
The **Mean Average Precision (a.k.a. Average Area Under the Precision-Recall curve)** is **0.928** with 95% confidence interval being [**0.915,0.941**].

The Balanced Accuracy is shown in the figure below:



Selecting to classify as class: **0** any sample with predicted probability to be in this class above **0.6833**, the model achieves:

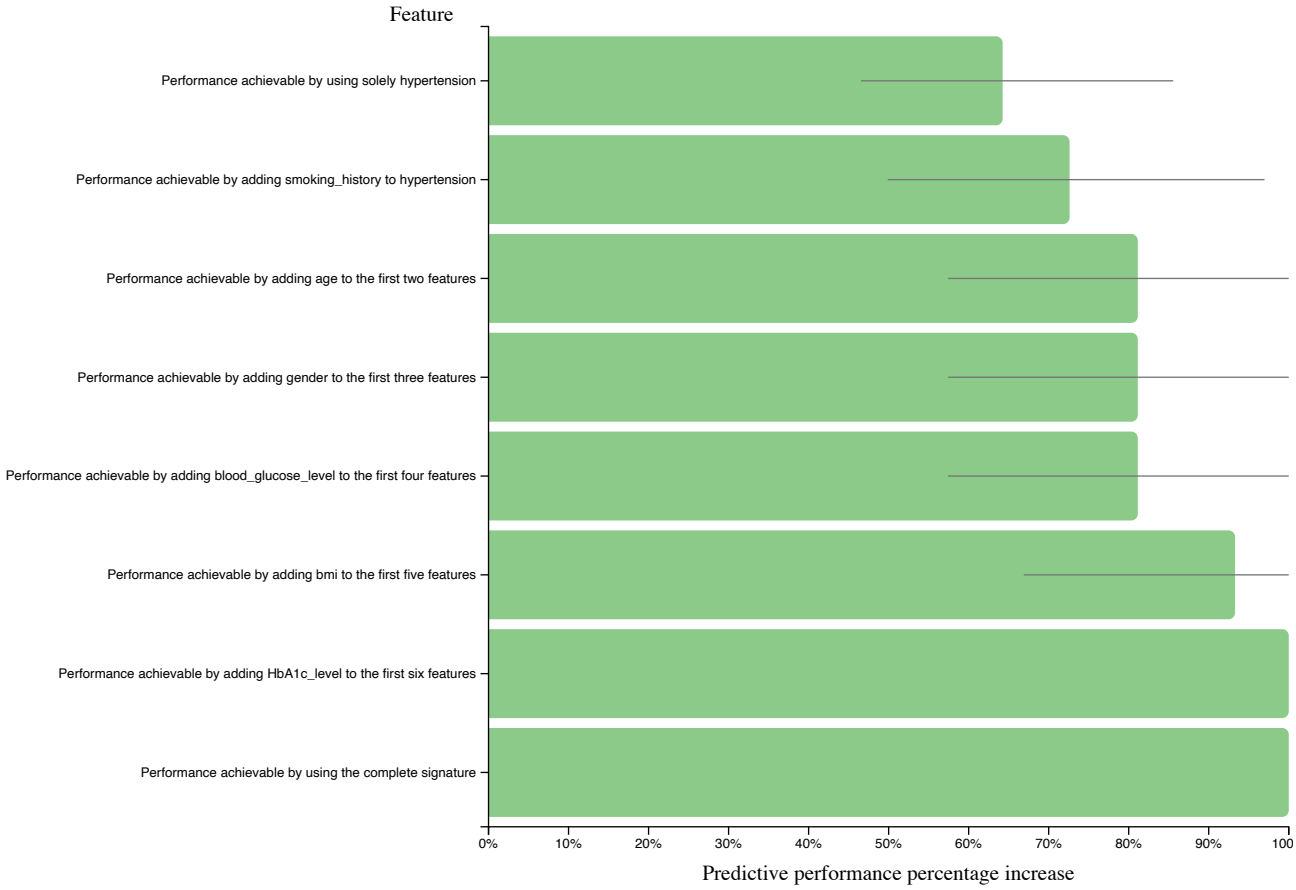
Metric	Mean estimate	CI
F1 Score	0.983	[0.981, 0.986]
F2 Score	0.991	[0.989, 0.992]
F0.5 Score	0.975	[0.972, 0.979]
Accuracy	0.970	[0.966, 0.975]
Balanced Accuracy	0.825	[0.805, 0.846]
Matthews correlation criterion (phi coefficient)	0.780	[0.752, 0.809]
Precision	0.970	[0.966, 0.975]
True Positive Rate (a.k.a. Sensitivity, Recall. Hit Rate)	0.996	[0.994, 0.998]
Specificity	0.673	[0.635, 0.715]
True Positive Ratio	0.911	[0.904, 0.919]
True Negative Ratio	0.057	[0.052, 0.063]
False Positive Ratio	0.028	[0.023, 0.032]
False Negative Ratio	0.004	[0.002, 0.005]

Feature Selection

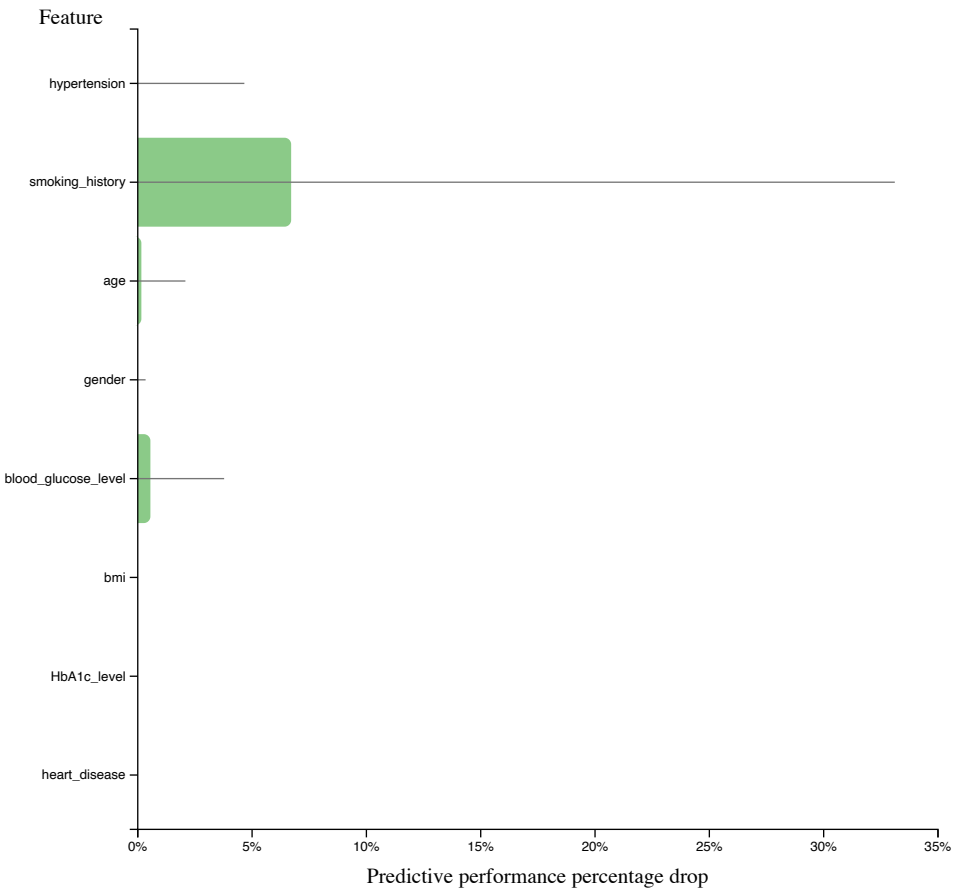
There were **8** features selected out of the **8** available.

The selected features consist of the following subset called a signature. **There was a single signature identified.** The first signature identified by the system is the set: **blood_glucose_level, HbA1c_level, age, bmi, hypertension, heart_disease, smoking_history, gender** in order of importance. The following features cannot be substituted with others and still obtain an equal predictive performance: **blood_glucose_level, HbA1c_level, age, bmi, hypertension, heart_disease, smoking_history, gender**.

The performance achieved by adding each feature in sequence to the model relative to the performance of the final model with all selected features is shown below. The features are added in order of importance:

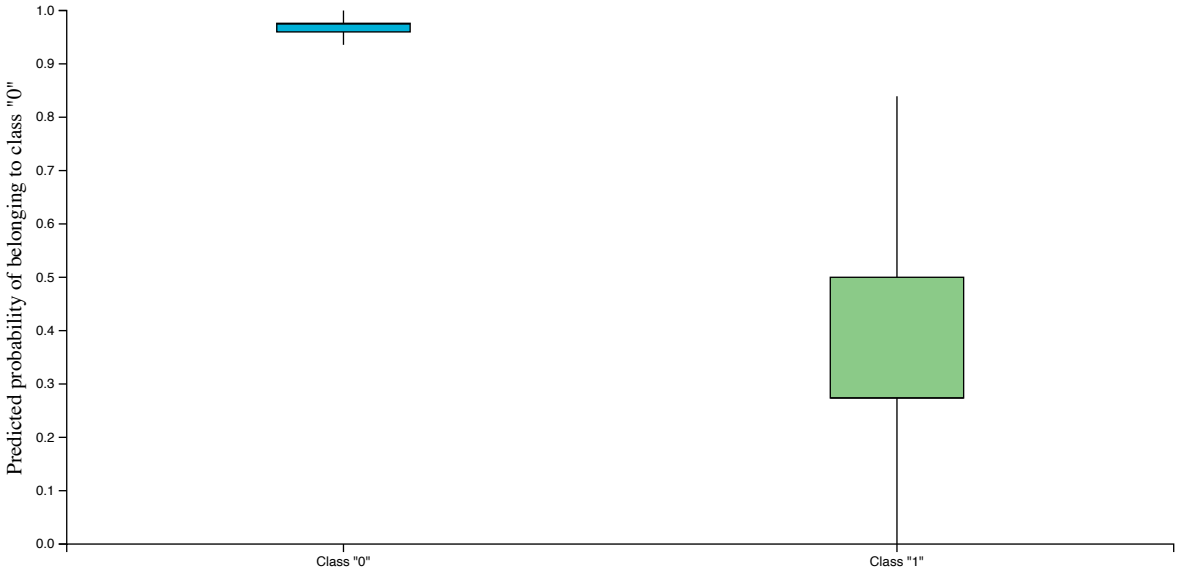


Some features may not seem to add predictive performance to the model; however, the feature selection algorithms include them as an effort to make the final model more robust to noise. The performances achieved by a model that contains all features except one, relative to the performance achieved when the feature is removed is shown below:



For some features there is no noticeable drop in performance when they are removed because they carry predictive information that is shared by other features selected.

The separation of the predictions of the classes achieved by the model is shown in the box-plots below. These are the out-of-sample predictions made by model produced by the same configuration as the final model when the sample was used for testing (e.g.., during cross-validation) and was not used to train the model.



Appendix

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
1	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5676631308261009	00:02:53.173620	false
2	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7173288331726133	00:00:07.7254	false
3	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:00:24.24604	false
4	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.01	0.7170748955319833	00:00:59.59126	false
5	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.01	0.7170748955319833	00:00:07.7304	false
6	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.7185053037608486	00:00:24.24729	false
7	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5676631308261009	00:02:53.173549	false
8	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6868145290903247	00:00:42.42147	false
9	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7172741883638701	00:00:14.14274	false
10	Mean Imputation,	Epilogi	equivThresh = 0.01, stopping	Classification Decision Tree	minimum leaf size = 4,	0.7188203150112504	00:02:53.173590	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
	Mode Imputation, Constant Removal, Standardization		criterion = Independence Test, stopping threshold = 0.01	with Deviance splitting criterion	alpha = 0.05			
11	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.7185053037608486	00:00:58.58575	false
12	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7172741883638701	00:00:14.14505	false
13	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:02:53.173605	false
14	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:00:14.14579	false
15	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7172741883638701	00:00:15.15797	false
16	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:00:35.35319	false
17	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:58.58639	false
18	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6868145290903247	00:01:16.76040	false
19	Mean Imputation, Mode Imputation, Constant	Univariate feature selection with Benjamini-	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance	ntrees = 100, minimum leaf size = 4	0.7324140147862424	00:01:01.61228	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
	Removal, Standardization	Hochberg correction		splitting criterion				
20	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.7318675666988107	00:00:49.49071	false
21	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:01:09.69785	false
22	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5787013821922211	00:03:04.184468	false
23	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7173288331726133	00:00:59.59049	false
24	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:00:58.58381	false
25	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.7318675666988107	00:01:23.83892	false
26	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:00:15.15826	false
27	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.690337512054002	00:00:16.16074	false
28	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.6972645451623272	00:02:55.175244	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
29	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.7185053037608486	00:02:53.173742	false
30	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:58.58635	false
31	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.690337512054002	00:00:17.17289	false
32	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5869527483124397	00:00:14.14636	false
33	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:24.24778	false
34	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.7298585663773707	00:00:38.38885	false
35	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.7324140147862424	00:01:01.61952	false
36	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5869527483124397	00:00:15.15844	false
37	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:00:24.24602	false
38	Mean Imputation,	Test-Budgeted	maxK = 2, alpha = 0.05,	Classification Decision Tree	minimum leaf size = 4,	0.7188203150112504	00:00:07.7222	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
	Mode Imputation, Constant Removal, Standardization	Statistically Equivalent Signature (SES)	budget = 3 * nvars	with Deviance splitting criterion	alpha = 0.05			
39	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5758052073288331	00:00:25.25181	false
40	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.7307875281260046	00:00:17.17196	false
41	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:00:58.58380	false
42	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.690337512054002	00:00:15.15732	false
43	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:00:58.58965	false
44	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5869527483124397	00:00:14.14632	false
45	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.7298585663773707	00:00:40.40117	false
46	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.7307875281260046	00:00:18.18764	false
47	IdentityFactory	FullSelector	-	Trivial model	-	0.5	00:00:00.002	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
48	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.7324140147862424	00:00:27.27468	false
49	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:24.24879	false
50	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6868145290903247	00:01:16.76590	false
51	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:02:53.173697	false
52	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.01	0.7170748955319833	00:00:58.58443	false
53	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6893281902925105	00:00:33.33400	false
54	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7173288331726133	00:00:58.58347	false
55	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.7185053037608486	00:00:07.7429	false
56	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6893281902925105	00:00:33.33006	false
57	Mean Imputation,	LASSO	penalty = 1.5	Classification Random	ntrees = 100, minimum leaf	0.7307875281260046	00:00:17.17557	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
	Mode Imputation, Constant Removal, Standardization			Forest with Deviance splitting criterion	size = 4			
58	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6893281902925105	00:00:32.32241	false
59	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:07.7458	false
60	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.01	0.7170748955319833	00:02:53.173653	false
61	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6868145290903247	00:00:24.24849	false
62	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.719038894246223	00:00:14.14658	false
63	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:00:15.15773	false
64	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:59.59298	false
65	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:00:07.7289	false
66	Mean Imputation, Mode Imputation, Constant	Univariate feature selection with Benjamini-	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:59.59172	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
	Removal, Standardization	Hochberg correction		splitting criterion				
67	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:01:09.69721	false
68	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:00:14.14157	false
69	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5758052073288331	00:00:26.26490	false
70	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5758052073288331	00:00:26.26480	false
71	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.728383156541305	00:03:17.197951	false
72	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5869527483124397	00:00:15.15929	false
73	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.7298585663773707	00:00:39.39156	false
74	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5758052073288331	00:00:25.25425	false
75	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.6964448730311796	00:01:00.60499	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
76	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5758052073288331	00:00:25.25906	false
77	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7173288331726133	00:02:53.173558	false
78	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5840019286403085	00:00:07.7318	false
79	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.05	0.7173288331726133	00:00:24.24646	false
80	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5787013821922211	00:03:04.184331	false
81	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5758052073288331	00:00:24.24975	false
82	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5869527483124397	00:00:14.14368	false
83	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.05	0.7188203150112504	00:00:58.58301	false
84	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.01	0.7170748955319833	00:00:24.24736	false
85	Mean Imputation,	Test-Budgeted	maxK = 2, alpha = 0.05,	Classification Random	ntrees = 100, minimum leaf	0.6964448730311796	00:00:08.8992	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
	Mode Imputation, Constant Removal, Standardization	Statistically Equivalent Signature (SES)	budget = 3 * nvars	Forest with Deviance splitting criterion	size = 4			
86	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:00:07.7884	false
87	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:00:17.17833	false
88	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.5869527483124397	00:00:14.14418	false
89	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.6964448730311796	00:00:26.26116	false
90	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.6964448730311796	00:00:59.59937	false
91	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:00:17.17882	false
92	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.7324140147862424	00:00:10.10033	false
93	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 2, alpha = 0.01	0.7170748955319833	00:00:14.14146	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
94	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.05	0.7188203150112504	00:00:14.14172	false
95	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.01	0.7170748955319833	00:00:15.15909	false
96	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.05	0.7188203150112504	00:00:14.14517	false
97	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.01	0.7170748955319833	00:00:14.14610	false
98	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.7185053037608486	00:00:59.59111	false
99	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.05, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.7318675666988107	00:00:32.32482	false
100	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.05	0.7188203150112504	00:00:15.15825	false
101	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.6906943105110896	00:03:11.191455	false
102	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.05	0.7188203150112504	00:00:59.59003	false
103	Mean Imputation,	LASSO	penalty = 0.5	Classification Decision Tree	minimum leaf size = 3,	0.7170748955319833	00:00:14.14290	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
	Mode Imputation, Constant Removal, Standardization			with Deviance splitting criterion	alpha = 0.01			
104	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 0.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.719038894246223	00:00:14.14375	false
105	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:01:09.69280	false
106	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.001, max vars = 100	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:00:59.59037	false
107	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:01:08.68983	false
108	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.5741112182577949	00:00:35.35152	false
109	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.0	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 3, alpha = 0.05	0.719038894246223	00:00:15.15856	false
110	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Epilogi	equivThresh = 0.01, stopping criterion = Independence Test, stopping threshold = 0.01	Classification Random Forest with Deviance splitting criterion	ntrees = 100, minimum leaf size = 4	0.729926068788171	00:02:56.176559	false
111	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Univariate feature selection with Benjamini-Hochberg correction	alpha = 0.01, max vars = 100	Classification Random Forest with Deviance splitting criterion	ntrees = 500, minimum leaf size = 4	0.7318675666988107	00:01:23.83194	false

Configuration	Preprocessing	Name	Hyperparams	Name	Hyperparams	Performance (unadjusted)	Time (miliseconds)	Dropped
112	Mean Imputation, Mode Imputation, Constant Removal, Standardization	LASSO	penalty = 1.5	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.01	0.7175120540019286	00:00:14.14569	false
113	Mean Imputation, Mode Imputation, Constant Removal, Standardization	Test-Budgeted Statistically Equivalent Signature (SES)	maxK = 2, alpha = 0.01, budget = 3 * nvars	Classification Decision Tree with Deviance splitting criterion	minimum leaf size = 4, alpha = 0.05	0.7188203150112504	00:00:24.24611	false