## STREAMLINE Training Summary Report: 2022-06-21 20:57:04.879407

## **General Pipeline Settings:**

Data Path: /home/ryanurb/idata/datasets/multiplexer\_set

Output Path: /home/ryanurb/idata/output Experiment Name: MultiplexerSet\_noSelect\_Sub

Class Label: Class Instance Label: None Ignored Features: None

Specified Categorical Features: None

CV Partitions: 10 Partition Method: S Match Label: None Categorical Cutoff: 10

Statistical Significance Cutoff: 0.05 Export Feature Correlations: True Export Univariate Plots: False

Random Seed: 42

TURF Cutoff: 0.5

Run From Jupyter Notebook: False

Use Data Scaling: True
Use Data Imputation: True
Use Multivariate Imputation: True
Use Mutual Information: True
Use MultiSURF: True
Use TURF: False

MultiSURF Instance Subset: 2000 Max Features to Keep: 2000 Filter Poor Features: False Top Features to Display: 40

Export Feature Importance Plot: True
Overwrite CV Datasets: True

Overwrite CV Datasets: True
Primary Metric: balanced\_accuracy

Training Subsample for KNN,ANN,SVM,and XGB: 5000 Uniform Feature Importance Estimation (Models): True Hyperparameter Sweep Number of Trials: 200

Hyperparameter Timeout: 900

Export Hyperparameter Sweep Plots: True

Export ROC Plot: True
Export PRC Plot: True
Export Metric Boxplots: True

Export Feature Importance Boxplots: True

Metric Weighting Composite FI Plots: balanced\_accuracy

Top Model Features To Display: 40

## ML Modeling Algorithms:

Naive Bayes: True

Logistic Regression: True
Decision Tree: True
Random Forest: True
Gradient Boosting: True
Extreme Gradient Boosting: True
Light Gradient Boosting: True
Category Gradient Boosting: True
Support Vector Machine: True
Artificial Neural Network: True
K-Nearest Neightbors: True
Genetic Programming: True
eLCS: False
XCS: False
ExSTraCS: True

## LCS Settings (eLCS,XCS,ExSTraCS):

Do LCS Hyperparameter Sweep: False

nu: 10

Training Iterations: 500000 N (Rule Population Size): 5000

LCS Hyperparameter Sweep Timeout: 1200

## Datasets:

 $D1 = A_6_bit_mutliplexer_500_01$ 

 $D2 = B\_11\_bit\_mutliplexer\_1000\_01$ 

 $D3 = C_20_bit_mutliplexer_2000_01$ 

 $D4 = D\_37\_bit\_mutliplexer\_5000\_01$ 

 $D5 = E\_70\_bit\_mutliplexer\_10000\_01$ 

 $D6 = F_135_bit_mutliplexer_20000_01$ 

## Univariate Analysis of Each Dataset (Top 10 Features for Each): Page 1

## $D1 = A_6_{bit_mutliplexer_500_01}$

### Feature: P-Value

R\_0: 1.26939024268301e-10 R\_3: 3.509298244030297e-07 R\_1: 9.059261024221464e-07

R\_2: 1.7448416343049778e-05 A\_0: 0.5694973919380477

## $D2 = B_11_bit_mutliplexer_1000_01$

### Feature: P-Value

R\_4: 2.2248801885218974e-05

R\_2: 0.0001766997611545

A\_1: 0.7971447941453101

R\_6: 0.000193302557659

R\_0: 0.0003044847642591

R\_1: 0.0012360481186965

K\_1. 0.0012300461160903

R\_7: 0.0052238913440066

R\_3: 0.0224712913608342

R\_5: 0.0225536002419627 A\_2: 0.2035317994195949

A\_0: 0.4098568324810853

## D3 = C\_20\_bit\_mutliplexer\_2000\_01

## Feature: P-Value

R\_2: 8.852163558112362e-05

R\_3: 0.0001193452640023

R\_6: 0.000384686668138

R\_5: 0.0008311970831204

 $R_1: \ 0.0008925709376199$ 

R\_4: 0.0016002295505668

 $R\_14:\ 0.0023725177489401$ 

 $R\_13\colon\ 0.0033974532166476$ 

R\_15: 0.0048580643893158

 $R_7: \ 0.0050257374945953$ 

## $D4 = D_37_bit_mutliplexer_5000_01$

### Feature: P-Value

R\_31: 0.0007736754593663

R\_30: 0.0008832966181209

R\_13: 0.0009918366525363

R\_2: 0.0016474655174039

R\_19: 0.0023476917358282

R\_1: 0.0024614843515886 R\_3: 0.0027526782036786

R\_15: 0.0095181037929174

R\_29: 0.0106617018265068

R\_5: 0.0110527366236116

## **D5** = **E\_70\_bit\_mutliplexer\_10000\_01**

## Feature: P-Value

R\_22: 2.225314574216959e-07

R\_34: 3.5048304422665855e-05

R\_25: 0.0004675117666111

A\_5: 0.0026933624905259

R\_38: 0.0030718690386894

 $R\_13:\ 0.0035117346967141$ 

R\_29: 0.0039853662019388 R\_12: 0.0045245747368835

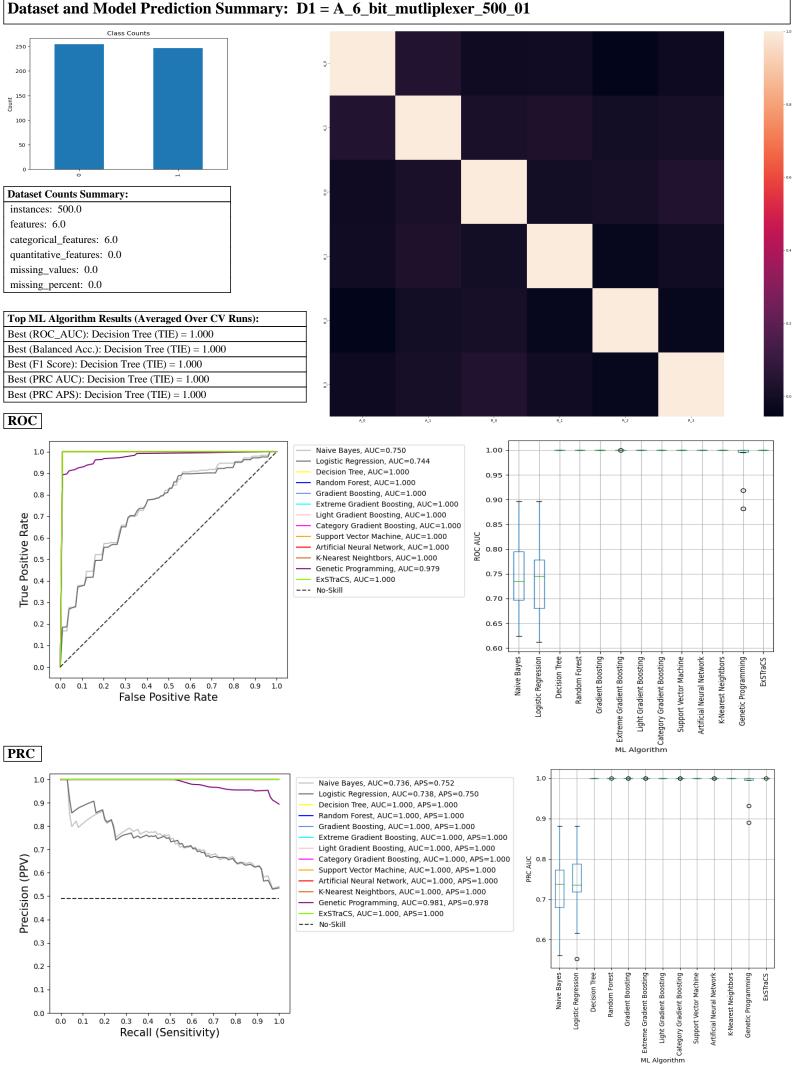
R\_12: 0.0043243747308833 R\_52: 0.0057657049009223

R\_59: 0.0060974802599191

# Univariate Analysis of Each Dataset (Top 10 Features for Each): Page 2

## $\boxed{D6 = F\_135\_bit\_mutliplexer\_20000\_01}$

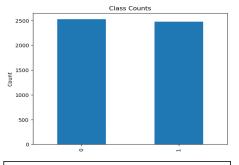
Feature: P-Value
R_23: 0.0008774724917323
R_65: 0.0028169569692886
R_49: 0.0033365077024448
R_74: 0.0055759947488155
R_105: 0.0058873767124799
R_116: 0.0075815765786996
R_94: 0.0090281431801193
R_26: 0.0124580448268327
R_59: 0.0168492271574137
R_35: 0.0240939875657274



#### Dataset and Model Prediction Summary: D2 = B\_11\_bit\_mutliplexer\_1000\_01 400 200 100 **Dataset Counts Summary:** instances: 1000.0 features: 11.0 categorical\_features: 11.0 quantitative\_features: 0.0 missing\_values: 0.0 missing\_percent: 0.0 Top ML Algorithm Results (Averaged Over CV Runs): Best (ROC\_AUC): Random Forest (TIE) = 1.000 Best (Balanced Acc.): Random Forest (TIE) = 1.000 Best (F1 Score): Random Forest (TIE) = 1.000 Best (PRC AUC): Random Forest (TIE) = 1.000 Best (PRC APS): Random Forest (TIE) = 1.000 ROC Naive Bayes, AUC=0.659 1.0 Logistic Regression, AUC=0.663 Decision Tree, AUC=0.981 0.9 Random Forest, AUC=1,000 0.8 Gradient Boosting, AUC=1.000 Extreme Gradient Boosting, AUC=1.000 Light Gradient Boosting, AUC=1.000 True Positive Rate 0.7 Category Gradient Boosting, AUC=1.000 Support Vector Machine, AUC=1.000 Artificial Neural Network, AUC=1.000 K-Nearest Neightbors, AUC=0.968 0.7 Genetic Programming, AUC=0.513 ExSTraCS, AUC=1.000 0.4 --- No-Skill 0.3 0.2 0.5 0.1 Genetic Programming Logistic Regression Light Gradient Boosting Category Gradient Boosting Support Vector Machine K-Nearest Neightbors Decision Tree Random Forest Gradient Boosting Extreme Gradient Boosting Artificial Neural Network 0.4 0.5 0.6 0.7 0.8 0.9 False Positive Rate PRC Naive Bayes, AUC=0.664, APS=0.671 Logistic Regression, AUC=0.668, APS=0.675 0.9 Decision Tree, AUC=0.986, APS=0.973 Random Forest, AUC=1.000, APS=1.000 0.9 0.8 Gradient Boosting, AUC=1.000, APS=1.000 Extreme Gradient Boosting, AUC=1.000, APS=1.000 Light Gradient Boosting, AUC=1.000, APS=1.000 0.7 Precision (PPV) Category Gradient Boosting, AUC=1.000, APS=1.000 PRC AUC Support Vector Machine, AUC=1.000, APS=1.000 Artificial Neural Network, AUC=1.000, APS=1.000 K-Nearest Neightbors, AUC=0.971, APS=0.972 Genetic Programming, AUC=0.628, APS=0.518 ExSTraCS, AUC=1.000, APS=1.000 0.4 --- No-Skill ∄ 0.3 ogistic Regression Random Forest Gradient Boosting Extreme Gradient Boosting Light Gradient Boosting Category Gradient Boosting Support Vector Machine Artificial Neural Network 0.1 0.0 0.3 0.4 0.5 0.6 0.7 0.0 0.1 0.2 0.8 0.9 Recall (Sensitivity)

#### Dataset and Model Prediction Summary: D3 = C\_20\_bit\_mutliplexer\_2000\_01 200 **Dataset Counts Summary:** instances: 2000.0 features: 20.0 categorical\_features: 20.0 quantitative\_features: 0.0 missing\_values: 0.0 missing\_percent: 0.0 Top ML Algorithm Results (Averaged Over CV Runs): Best (ROC\_AUC): Gradient Boosting (TIE) = 1.000 Best (Balanced Acc.): Category Gradient Boosting (TIE) = 1.000 Best (F1 Score): Category Gradient Boosting (TIE) = 1.000 £13 Best (PRC AUC): Gradient Boosting (TIE) = 1.000 Best (PRC APS): Gradient Boosting (TIE) = 1.000 ROC Naive Bayes, AUC=0.627 1.0 Logistic Regression, AUC=0.625 Decision Tree, AUC=0.772 0.9 Random Forest, AUC=0.987 0.9 0.8 Gradient Boosting, AUC=1.000 Extreme Gradient Boosting, AUC=1.000 Light Gradient Boosting, AUC=1.000 True Positive Rate 0.7 Category Gradient Boosting, AUC=1.000 0.8 Support Vector Machine, AUC=0.910 Artificial Neural Network, AUC=0.942 K-Nearest Neightbors, AUC=0.827 0.7 Genetic Programming, AUC=0.537 ExSTraCS, AUC=1.000 0.4 --- No-Skill 0.3 0.2 0.5 Genetic Programming Naive Bayes Logistic Regression Gradient Boosting Extreme Gradient Boosting Light Gradient Boosting S Category Gradient Boosting Support Vector Machine Artificial Neural Network K-Nearest Neightbors Decision Tree Random Forest 0.4 0.5 0.6 0.7 0.8 0.9 False Positive Rate PRC Naive Bayes, AUC=0.617, APS=0.622 Logistic Regression, AUC=0.615, APS=0.620 0.9 Decision Tree, AUC=0.805, APS=0.723 Random Forest, AUC=0.987, APS=0.987 0.9 0.8 Gradient Boosting, AUC=1.000, APS=1.000 Extreme Gradient Boosting, AUC=1.000, APS=1.000 Light Gradient Boosting, AUC=1.000, APS=1.000 0.7 Precision (PPV) Category Gradient Boosting, AUC=1.000, APS=1.000 PRC AUC 0.8 Support Vector Machine, AUC=0.907, APS=0.907 Artificial Neural Network, AUC=0.941, APS=0.941 K-Nearest Neightbors, AUC=0.827, APS=0.828 Genetic Programming, AUC=0.645, APS=0.512 ExSTraCS, AUC=1.000, APS=1.000 0.4 --- No-Skill 0.3 0.2 Naive Bayes ogistic Regression Random Forest Gradient Boosting Extreme Gradient Boosting Light Gradient Boosting Category Gradient Boosting Artificial Neural Network 0.1 0.0 0.3 0.4 0.5 0.6 0.7 0.0 0.1 0.2 0.8 0.9 Recall (Sensitivity)

## Dataset and Model Prediction Summary: D4 = D\_37\_bit\_mutliplexer\_5000\_01



## **Dataset Counts Summary:**

instances: 5000.0 features: 37.0

categorical\_features: 37.0 quantitative\_features: 0.0 missing\_values: 0.0 missing\_percent: 0.0

Best (PRC APS): ExSTraCS = 1.000

## ROC

0.9

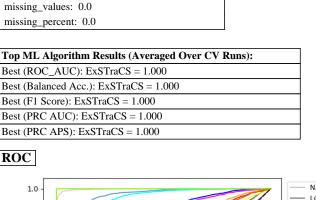
0.8

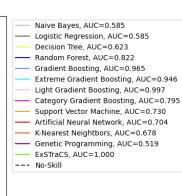
0.7

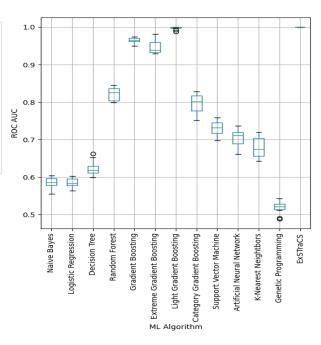
0.4

0.3 0.2

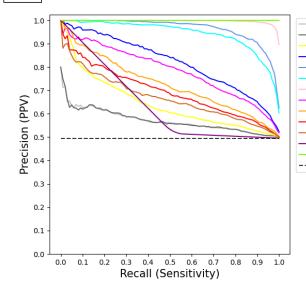
True Positive Rate







# PRC



0.3 0.4 0.5 0.6 0.7

False Positive Rate

0.8 0.9

Naive Bayes, AUC=0.571, APS=0.573 Logistic Regression, AUC=0.571, APS=0.573 Decision Tree, AUC=0.643, APS=0.601 Random Forest, AUC=0.820, APS=0.821 Gradient Boosting, AUC=0.966, APS=0.966 Extreme Gradient Boosting, AUC=0.945, APS=0.945 Light Gradient Boosting, AUC=0.997, APS=0.997 Category Gradient Boosting, AUC=0.785, APS=0.786 Support Vector Machine, AUC=0.727, APS=0.728 Artificial Neural Network, AUC=0.696, APS=0.697 K-Nearest Neightbors, AUC=0.669, APS=0.670 Genetic Programming, AUC=0.636, APS=0.506 ExSTraCS, AUC=1.000, APS=1.000 --- No-Skill

0.9 ŧ PRC AUC 0.7 # P Naive Bayes ogistic Regression न् नु Category Gradient Boosting Decision Tree Random Forest Gradient Boosting Extreme Gradient Boosting Light Gradient Boosting Support Vector Machine Artificial Neural Network

#### Dataset and Model Prediction Summary: D5 = E\_70\_bit\_mutliplexer\_10000\_01 2000 **Dataset Counts Summary:** instances: 10000.0 features: 70.0 categorical\_features: 70.0 quantitative\_features: 0.0 missing\_values: 0.0 missing\_percent: 0.0 Top ML Algorithm Results (Averaged Over CV Runs): Best (ROC\_AUC): ExSTraCS = 0.992 Best (Balanced Acc.): ExSTraCS = 0.983 Best (F1 Score): ExSTraCS = 0.983 Best (PRC AUC): ExSTraCS = 0.989 Best (PRC APS): ExSTraCS = 0.988 ROC Naive Bayes, AUC=0.570 1.0 Logistic Regression, AUC=0.570 Decision Tree, AUC=0.536 0.9 Random Forest, AUC=0.613 0.9 0.8 Gradient Boosting, AUC=0.672 Extreme Gradient Boosting, AUC=0.597 Light Gradient Boosting, AUC=0.684 True Positive Rate 0.7 Category Gradient Boosting, AUC=0.590 0.8 Support Vector Machine, AUC=0.591 Artificial Neural Network, AUC=0.562 K-Nearest Neightbors, AUC=0.591 0.7 Genetic Programming, AUC=0.514 ExSTraCS. AUC=0.992 0.4 --- No-Skill 0.3 0.6 $\downarrow$ 0.2 中 Genetic Programming Naive Bayes Logistic Regression Decision Tree Extreme Gradient Boosting Category Gradient Boosting Random Forest Gradient Boosting Light Gradient Boosting Support Vector Machine Artificial Neural Network K-Nearest Neightbors 0.4 0.5 0.6 0.7 0.8 0.9 False Positive Rate PRC Naive Bayes, AUC=0.566, APS=0.568 Logistic Regression, AUC=0.566, APS=0.567 0.9 Decision Tree, AUC=0.559, APS=0.543 Random Forest, AUC=0.609, APS=0.610 0.9 0.8 Gradient Boosting, AUC=0.660, APS=0.661 Extreme Gradient Boosting, AUC=0.589, APS=0.589 Light Gradient Boosting, AUC=0.673, APS=0.674 0.7 Precision (PPV) Category Gradient Boosting, AUC=0.586, APS=0.587 PRC AUC Support Vector Machine, AUC=0.582, APS=0.583 Artificial Neural Network, AUC=0.556, APS=0.557 0.7 K-Nearest Neightbors, AUC=0.578, APS=0.577 Genetic Programming, AUC=0.633, APS=0.506 ExSTraCS, AUC=0.989, APS=0.988 ₽ 0.4 --- No-Skill Ŧ 0.3 0.2 Naive Bayes ogistic Regression Random Forest Extreme Gradient Boosting Light Gradient Boosting Category Gradient Boosting Artificial Neural Network 0.1

0.0

0.0

0.1 0.2

0.3 0.4 0.5 0.6 0.7

Recall (Sensitivity)

0.8 0.9

#### Dataset and Model Prediction Summary: D6 = F\_135\_bit\_mutliplexer\_20000\_01 8000 4000 2000 **Dataset Counts Summary:** instances: 20000.0 features: 135.0 categorical\_features: 135.0 quantitative\_features: 0.0 missing\_values: 0.0 missing\_percent: 0.0 Top ML Algorithm Results (Averaged Over CV Runs): Best (ROC\_AUC): Light Gradient Boosting = 0.559 Best (Balanced Acc.): Light Gradient Boosting = 0.543 Best (F1 Score): ExSTraCS = 0.612 Best (PRC AUC): Genetic Programming = 0.625 Best (PRC APS): Light Gradient Boosting = 0.549 ROC Naive Bayes, AUC=0.525 1.0 0.58 Logistic Regression, AUC=0.523 Decision Tree, AUC=0.501 0.9 Random Forest, AUC=0.542 0.8 Gradient Boosting, AUC=0.542 0.56 Extreme Gradient Boosting, AUC=0.529 Light Gradient Boosting, AUC=0.559 True Positive Rate 0.7 Category Gradient Boosting, AUC=0.533 Support Vector Machine, AUC=0.534 Artificial Neural Network, AUC=0.528 K-Nearest Neightbors, AUC=0.535 Genetic Programming, AUC=0.502 0.52 ExSTraCS, AUC=0.512 0.4 --- No-Skill 0.3 0.50 0.2 0.1 De Light Gradient Boosting Light Gradient Boosting Category Gradient Boosting Naive Bayes Logistic Regression Decision Tree Gradient Boosting Extreme Gradient Boosting Support Vector Machine K-Nearest Neightbors Genetic Programming Random Forest Artificial Neural Network 0.4 0.5 0.6 0.7 0.8 0.9 False Positive Rate PRC 0.64 1.0 Naive Bayes, AUC=0.517, APS=0.518 Logistic Regression, AUC=0.515, APS=0.516 0.9 Decision Tree, AUC=0.505, APS=0.500 Random Forest, AUC=0.536, APS=0.536 0.8 Gradient Boosting, AUC=0.532, APS=0.532 Extreme Gradient Boosting, AUC=0.517, APS=0.518 0.58 Light Gradient Boosting, AUC=0.548, APS=0.549 0.7 Precision (PPV) Category Gradient Boosting, AUC=0.524, APS=0.525 PRC AUC Support Vector Machine, AUC=0.525, APS=0.525 0.56 Artificial Neural Network, AUC=0.521, APS=0.522 K-Nearest Neightbors, AUC=0.525, APS=0.524 Genetic Programming, AUC=0.625, APS=0.497 ExSTraCS, AUC=0.509, APS=0.507 --- No-Skill 0.50 0.3 ogistic Regression Decision Tree **Extreme Gradient Boosting** Light Gradient Boosting Category Gradient Boosting Artificial Neural Network 0.1 0.0 0.4 0.5 0.6 0.7 0.0 0.1 0.2 0.3 0.8 0.9 Recall (Sensitivity)

#### Average Model Prediction Statistics (Rounded to 3 Decimal Points): Page 1 $D1 = A_6_{bit_mutliplexer_500_01}$ TN FP ML Algorithm Balanced Accuracy TP FN NPV LR+ LR-ROC PRC PRC Sensitivity Specificity Precision Score (Recall) | 0.653 | 0.654 | 0.637 | 0.63 AUC AUC APS 0.75 0.736 0.752 0.744 0.738 0.75 Accuracy 0.664 Naive Bayes Logistic Regression Decision Tree Random Forest Gradient Boosting 16.1 15.5 24.6 24.6 24.6 8.3 8.3 0.0 0.0 0.0 8.5 9.1 0.0 0.0 0.0 0.664 0.674 2.194 0.661 3.242 1.0 0.0 1.0 0.0 1.0 0.0 0.549 0.0 0.0 0.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 Gradient Boosting Extreme Gradient Boosting Light Gradient Boosting Category Gradient Boosting Support Vector Machine Artificial Neural Network K-Nearest Neightbors Genetic Programming ExSTraCS 1.0 1.0 1.0 1.0 1.0 1.0 24.6 24.6 24.6 24.6 24.6 24.6 24.6 0.0 0.0 0.0 0.0 0.0 0.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 0.0 1.0 1.0 0.0 0.0 0.0 1.0 1.0 1.0 1.0 1.0 1.0 0.0 1.0 1.0 1.0 0.0 0.0 0.0 1.0 1.0

0.949

0.948

0.976

1.0

0.974 6.708 1.0 0.0

0.6

1.0

0.978

0.979 0.981 1.0 1.0

$D2 = B_11_bit_muth$	iplexer_	1000_0	1													
ML Algorithm	Balanced	Accuracy	F1		Specificity		TP	TN	FP	FN	NPV	LR+	LR-	ROC	PRC	PRC
	Accuracy		Score	(Recall)		(PPV)								AUC	AUC	APS
Naive Bayes	0.613	0.613	0.624	0.638	0.587	0.613	32.2	29.1	20.4	18.3	0.614	1.596	0.63	0.659	0.664	0.671
Logistic Regression	0.605	0.606	0.618	0.636	0.575	0.607	32.1	28.5	21.0	18.4	0.607	1.557	0.648	0.663	0.668	0.675
Decision Tree	0.979	0.979	0.979	0.976	0.982	0.982	49.3	48.6	0.9	1.2	0.976	23.409	0.024	0.981	0.986	0.973
Random Forest	1.0	1.0	1.0	1.0	1.0	1.0	50.5	49.5		0.0	1.0	0.0	0.0	1.0	1.0	1.0
Gradient Boosting	1.0	1.0	1.0	1.0	1.0	1.0	50.5	49.5	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0
Extreme Gradient Boosting	1.0	1.0	1.0	1.0	1.0	1.0	50.5	49.5	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0
Light Gradient Boosting	1.0	1.0	1.0	1.0	1.0	1.0	50.5	49.5		0.0	1.0	0.0	0.0	1.0	1.0	1.0
Category Gradient Boosting	1.0	1.0	1.0	1.0	1.0	1.0	50.5	49.5	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0
Support Vector Machine	0.991	0.991	0.991	0.994	0.988	0.988	50.2	48.9	0.6	0.3	0.994	22.102	0.006	1.0	1.0	1.0
Artificial Neural Network	0.991	0.991	0.991	0.996	0.986	0.986	50.3	48.8	0.7	0.2	0.996	19.602	0.004	1.0	1.0	1.0
K-Nearest Neightbors	0.905	0.905	0.906	0.911	0.899	0.904	46.0	44.5	5.0	4.5	0.909	11.462	0.1	0.968	0.971	0.972
Genetic Programming	0.503	0.503	0.508	0.511	0.495	0.509	25.8	24.5	25.0	24.7	0.496	1.025	1.011	0.513	0.628	0.518
ExSTraCS	1.0	1.0	1.0	1.0	1.0	1.0	50.5	49.5	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0

$D3 = C_20_bit_mutl$	iplexer_	_2000_0	)1													
ML Algorithm	Balanced	Accuracy	F1	Sensitivity	Specificity	Precision	TP	TN	FP	FN	NPV	LR+	LR-	ROC	PRC	PRC
	Accuracy		Score	(Recall)		(PPV)								AUC	AUC	APS
Naive Bayes	0.589		0.574		0.615		55.3	62.6	39.2	42.9	0.592	1.517	0.722	0.627		0.622
Logistic Regression	0.582	0.582	0.555	0.534	0.63	0.583	52.4	64.1	37.7	45.8	0.584	1.495	0.747	0.625	0.615	0.62
Decision Tree	0.751	0.751	0.743	0.751	0.751	0.741	73.7	76.4	25.4	24.5	0.766	3.449	0.345	0.772	0.805	0.723
Random Forest	0.938	0.938	0.936	0.928	0.948		91.1	96.5	5.3	7.1	0.932	26.245	0.076	0.987	0.987	0.987
Gradient Boosting	1.0		0.999		0.999		98.2	101.7		0.0	1.0	10.1	0.0	1.0	1.0	1.0
Extreme Gradient Boosting	0.993	0.993	0.993	0.993	0.993		97.5	101.1	0.7	0.7	0.993	40.592	0.007	1.0	1.0	1.0
Light Gradient Boosting	0.999	0.999	0.999	1.0	0.998	0.998	98.2	101.6	0.2	0.0	1.0	20.3	0.0	1.0	1.0	1.0
Category Gradient Boosting	1.0	1.0	1.0	1.0	1.0	1.0	98.2	101.8	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0
Support Vector Machine	0.829		0.825	0.824	0.834		80.9	84.9	16.9	17.3	0.832	5.206	0.211	0.91	0.907	0.907
Artificial Neural Network	0.878	0.878	0.876	0.881	0.876	0.872	86.5	89.2	12.6	11.7	0.886	8.043	0.139		0.941	0.941
K-Nearest Neightbors	0.751	0.751	0.744		0.763	0.75	72.5	77.7	24.1	25.7	0.753	3.19	0.344	0.827	0.827	0.828
Genetic Programming	0.537	0.537	0.531	0.538	0.535	0.526	52.8	54.5	47.3	45.4	0.548	1.16	0.865	0.537	0.645	0.512
ExSTraCS	1.0	1.0	1.0	1.0	1.0	1.0	98.2	101.8	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0

$D4 = D_37_bit_mutl$	iplexer_	_5000_0	1													
ML Algorithm		Accuracy	F1		Specificity		TP	TN	FP	FN	NPV	LR+	LR-	ROC	PRC	PRC
	Accuracy		Score	(Recall)		(PPV)								AUC	AUC	APS
Naive Bayes	0.559	0.559		0.534	0.584	0.558	132.2	147.3	104.9	115.6	0.56	1.285	0.799	0.585	0.571	0.573
Logistic Regression	0.556	0.556	0.544	0.535	0.577	0.554	132.6	145.4	106.8	115.2	0.558	1.268	0.807	0.585	0.571	0.573
Decision Tree	0.59	0.59	0.594	0.606	0.575	0.584	150.1	145.0	107.2	97.7	0.598	1.434	0.686	0.623	0.643	0.601
Random Forest	0.735	0.735	0.728	0.718	0.753	0.741	177.8	189.8	62.4	70.0	0.731	2.941	0.375	0.822	0.82	0.821
Gradient Boosting	0.899	0.899	0.898	0.897	0.901	0.9	222.2	227.3	24.9	25.6	0.899	9.404	0.115	0.965	0.966	0.966
Extreme Gradient Boosting	0.874	0.874	0.872	0.869	0.879	0.877	215.3	221.7	30.5	32.5	0.873	8.41	0.15	0.946	0.945	0.945
Light Gradient Boosting	0.98	0.98	0.98	0.981	0.978	0.978	243.1	246.7	5.5	4.7	0.981	67.068	0.02	0.997	0.997	0.997
Category Gradient Boosting	0.717	0.717	0.706	0.685	0.748	0.729	169.7	188.7	63.5	78.1	0.708	2.788	0.422	0.795	0.785	0.786
Support Vector Machine	0.669	0.669	0.665	0.665	0.673	0.667	164.7	169.8		83.1	0.672	2.044	0.498	0.73		0.728
Artificial Neural Network	0.646	0.646	0.642	0.641	0.651	0.644	158.8	164.3	87.9	89.0	0.649	1.849	0.552	0.704		0.697
K-Nearest Neightbors	0.624	0.624	0.611	0.595	0.653	0.628	147.5	164.6	87.6	100.3	0.622	1.738	0.622	0.678		0.67
Genetic Programming	0.519	0.519	0.516	0.518	0.52	0.515	128.3	131.1	121.1	119.5	0.523	1.082	0.93	0.519	0.636	0.506
ExSTraCS	1.0	1.0	1.0	1.0	1.0	1.0	247.8	252.2	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0

													-			
$D5 = E_70_bit_mutb$	iplexer_	_10000_	01													
ML Algorithm	Balanced Accuracy	Accuracy	F1 Score	Sensitivity (Recall)	Specificity	Precision (PPV)	TP	TN	FP	FN	NPV	LR+	LR-	ROC AUC	PRC AUC	PRC APS
Naive Bayes	0.549	0.549	0.546	0.544	0.555	0.549	271.3	277.9	223.2	227.6	0.55	1.225	0.824	0.57	0.566	0.568
Logistic Regression	0.549	0.55	0.547	0.545	0.554	0.549	271.8	277.7	223.4	227.1	0.55	1.226	0.823	0.57	0.566	0.567
Decision Tree	0.522	0.522	0.515	0.51	0.534	0.521	254.4	267.5	233.6	244.5	0.522	1.096	0.919	0.536	0.559	0.543
Random Forest	0.577	0.577	0.572	0.568	0.586	0.577	283.3	293.6	207.5	215.6	0.577	1.374	0.738	0.613	0.609	0.61
Gradient Boosting	0.622	0.622	0.621	0.622	0.623	0.621	310.1	312.0	189.1	188.8	0.623	1.651	0.608	0.672	0.66	0.661
Extreme Gradient Boosting	0.568	0.568	0.559	0.551	0.584	0.569	274.9	292.7	208.4	224.0	0.567	1.326	0.768	0.597	0.589	0.589
Light Gradient Boosting	0.633	0.633	0.632	0.633	0.632	0.632	315.9	316.9	184.2	183.0	0.634	1.733	0.581	0.684	0.673	0.674
Category Gradient Boosting	0.557	0.557	0.549	0.54	0.574	0.558	269.3	287.7	213.4	229.6	0.556	1.269	0.802	0.59	0.586	0.587
Support Vector Machine	0.563	0.564	0.56	0.557	0.57	0.563	277.7	285.8	215.3	221.2	0.564	1.296	0.777	0.591	0.582	0.583
Artificial Neural Network	0.542	0.542	0.547	0.556	0.527	0.54	277.6	264.0	237.1	221.3	0.544	1.183	0.845	0.562	0.556	0.557
K-Nearest Neightbors	0.562	0.562	0.555	0.549	0.576	0.563	273.7	288.4	212.7	225.2	0.562	1.295	0.784	0.591	0.578	0.577
Genetic Programming	0.514	0.514	0.51	0.507	0.521	0.513	253.1	261.1	240.0	245.8	0.515	1.061	0.947	0.514	0.633	0.506
ExSTraCS	0.983	0.983	0.983	0.982	0.984	0.984	489.8	493.2	7.9	9.1	0.983	129,236	0.019	0.992	0.989	0.988

# Average Model Prediction Statistics (Rounded to 3 Decimal Points): Page 2

$D6 = F_135_bit_mut$	liplexer	_20000	_01													
ML Algorithm	Balanced	Accuracy	F1		Specificity		TP	TN	FP	FN	NPV	LR+	LR-	ROC	PRC	PRC
	Accuracy		Score	(Recall)		(PPV)								AUC	AUC	APS
Naive Bayes	0.515	0.515	0.503	0.494	0.536	0.512	490.5	540.2	467.1	502.2	0.518	1.066	0.944	0.525	0.517	0.518
Logistic Regression	0.512	0.512	0.485	0.469	0.555	0.51	466.0	558.7	448.6	526.7	0.515	1.058	0.957	0.523	0.515	0.516
Decision Tree	0.502	0.502	0.493	0.489	0.515	0.498	485.3	518.7	488.6	507.4	0.506	1.008	0.993	0.501	0.505	0.5
Random Forest	0.529	0.529	0.519	0.512	0.545	0.526	508.4	549.3	458.0	484.3	0.532	1.127	0.895	0.542	0.536	0.536
Gradient Boosting	0.531	0.531	0.524	0.52	0.541	0.528	516.7	545.4	461.9	476.0	0.534	1.136	0.886	0.542	0.532	0.532
Extreme Gradient Boosting	0.519	0.519			0.56	0.517	475.2	563.6	443.7	517.5	0.521	1.088	0.932	0.529	0.517	0.518
Light Gradient Boosting	0.543	0.543	0.536	0.531	0.554	0.54	527.6	558.1	449.2	465.1	0.546	1.193	0.846	0.559	0.548	0.549
Category Gradient Boosting	0.52	0.521	0.453	0.4	0.639	0.523	397.5	643.9	363.4	595.2	0.52	1.111	0.938	0.533	0.524	0.525
Support Vector Machine	0.524	0.524	0.518	0.517	0.53	0.52	512.8	534.3	473.0	479.9	0.527	1.101	0.912	0.534	0.525	0.525
Artificial Neural Network	0.522	0.522	0.515	0.511	0.533	0.519	507.3	537.1	470.2	485.4	0.525	1.096	0.917	0.528	0.521	0.522
K-Nearest Neightbors	0.525	0.525	0.509	0.497	0.553	0.523	493.5	556.8	450.5	499.2	0.527	1.114	0.911	0.535	0.525	0.524
Genetic Programming	0.502	0.502	0.502		0.496	0.498	503.4	499.9	507.4	489.3	0.505	1.008	0.994	0.502	0.625	0.497
ExSTraCS	0.505	0.503	0.612	0.799	0.212	0.5	792.9	213.0	794.3	199.8	0.508	1.015	0.991	0.512	0.509	0.507

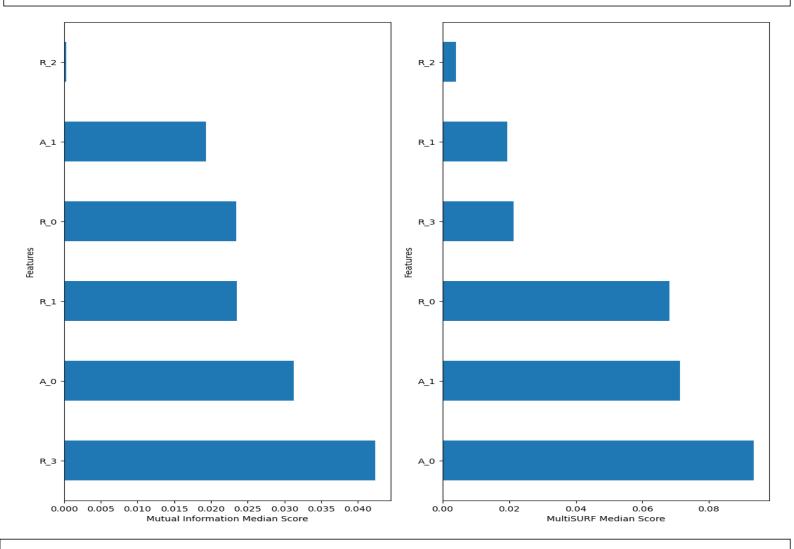
#### Median Model Prediction Statistics (Rounded to 3 Decimal Points): Page 1 $D1 = A_6$ bit\_mutliplexer 500 01 ML Algorithm TP TN FP FN NPV Balanced Accuracy LR+ LR-ROC PRC PRC Sensitivity Specificity Precision AUC AUC 0.735 0.738 0.745 0.736 Score (Recall) 0.667 | 0.673 Accuracy 16.0 15.0 25.0 25.0 25.0 Naive Baves 15.0 25.0 25.0 25.0 10.0 Logistic Regression 0.646 0.648 0.648 1.885 0.746 0.0 Decision Tree Random Forest 1.0 1.0 1.0 1.0 1.0 1.0 0.0 1.0 0.0 0.0 0.0 Gradient Boosting 1.0 1.0 0.0 1.0 1.0 1.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 25.0 0.0 Extreme Gradient Boosting Light Gradient Boosting 1.0 1.0 1.0 1.0 0.0 1.0 1.0 0.0 1.0 1.0 Category Gradient Boosting Support Vector Machine Artificial Neural Network K-Nearest Neightbors 0.0 0.0 0.0 0.0 1.0 1.0 1.0 1.0 0.0 0.0 0.0 1.0 1.0 0.0 0.0 0.0 1.0 1.0 1.0 0.0 0.0 1.0 1.0 1.0 1.0 1.0 Genetic Programming ExSTraCS 0.99 1.0 0.981 0.981 24.0 25.0 0.5 0.0 0.0 1.0 1.0 $D2 = B_11_bit_mutliplexer_1000_01$ ML Algorithm Balanced Accuracy Sensitivity Specificity Precision TP TNFP FN NPV LR+ LR-ROC PRC PRC Score (Recall) 0.647 | 0.654 0.63 | 0.644 (PPV) 0.625 0.614 APS 0.654 0.659 AUC 0.658 AUC 0.645 Accuracy 0.633 Naive Bayes Logistic Regression Decision Tree 0.639 1.634 0.618 1.56 0.6 0.586 17.5 18.0 0.609 0.61 0.607 50.5 50.5 50.5 50.5 50.5 48.5 1.5 0.985 0.97 0.98 0.971 20.016 0.03 Random Forest Gradient Boosting Extreme Gradient Boosting Light Gradient Boosting Category Gradient Boosting 0.0 0.0 1.0 1.0 1.0 1.0 1.0 1.0 49.5 0.0 0.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 1.0 49.5 0.0 0.0 1.0 1.0 0.0 0.0 1.0 1.0 1.0 0.0 0.0 0.0 4.5 25.0 0.0 0.5 0.5 4.5 23.5 0.99 1.0 0.995 1.0 0.908 0.91 Support Vector Machine Artificial Neural Network K-Nearest Neightbors 0.99 0.995 0.905 0.99 0.99 0.909 0.525 50.0 50.5 46.5 49.0 49.0 45.0 0.99 0.99 0.90 0.909 0.09 Genetic Programming 0.5 0.5 0.508 | 0.5 0.505 26.0 0.495 1.001 0.998 0.504 | 0.631 50.5 49.5 0.0 0.0 $D3 = C_20\_bit\_mutliplexer\_2000\_01$ Balanced Accuracy TP TN FP FN NPV PRC LR+ LR-PRC ML Algorithm F1 Sensitivity Specificity Precision ROC Accuracy 0.589 0.573 0.781 Score (Recall) 0.572 | 0.556 APS 0.603 0.591 1.468 0.714 0.574 1.451 0.769 0.784 3.253 0.287 0.942 17.867 0.064 Naive Bayes Logistic Regression 43.5 44.5 22.5 62.5 64.0 39.5 38.0 0.616 0.555 0.546 0.784 0.772 0.769 75.0 96.5 102.0 Decision Tree 26.5 0.804 | 0.841 | 0.747 Random Forest Gradient Boosting 0.939 0.939 0.946 0.99 0.0 98.0 1.0 1.0 1.0 1.0 1.0 1.0 0.0 1.0 0.0 0.0 1.0 1.0 1.0 Extreme Gradient Boosting 0.5 25.5 0.01 1.0 Light Gradient Boosting Category Gradient Boosting 0.0 Support Vector Machine Artificial Neural Network K-Nearest Neightbors Genetic Programming ExSTraCS 82.5 88.5 14.0 12.5 23.0 0.842 5.473 0.902 7.489 0.759 3.131 0.836 0.835 0.835 | 0.84 0.863 0.841 88.0 89.5 15.5 10.0 0.92 0.912 0.912 0.961 0.96 0.96 0.823 0.822 0.823 0.531 0.637 0.508 0.112 0.893 0.877 0.755 0.531 0.751 0.755 79.0 0.889 0.539 55.0 102.0 46.0 98.0 0.0 1.0 0.0 1.0 1.0 1.0 1.0 1.0 1.0 0.0 0.0 1.0 1.0 1.0 $D4 = D_37_bit_mutliplexer_5000_01$ ML Algorithm Balanced Accuracy ΤP TN FN NPV LR+ PRC Sensitivity Specificity Precision ROC Score (Recall) | 0.545 | 0.533 | 0.548 | 0.53 | 0.592 | 0.599 | 0.727 | 0.723 (PPV) 0.557 0.554 0.579 0.742 AUC AUC APS 0.586 0.568 0.571 0.584 0.565 0.567 0.618 0.65 0.592 0.826 0.83 0.83 Accuracy 0.56 0.557 Naive Bayes 148.0 0.505 1.262 0.791 0.559 1.263 0.803 0.595 1.406 0.692 0.729 2.924 0.377 0.906 9.892 0.106 0.876 6.455 0.145 0.577 0.563 0.75 0.909 0.865 Logistic Regression Decision Tree Random Forest 0.557 0.588 0.734 131.5 148.5 179.0 145.5 142.0 189.0 229.0 218.0 248.5 189.0 170.5 165.0 106.5 110.0 116. 99.5 0.557 0.588 0.734 0.904 0.868 68.5 24.0 32.0 4.5 80.5 63.0 23.0 34.0 0.904 0.904 0.903 0.867 0.87 0.907 Gradient Boosting Extreme Gradient Boosting Light Gradient Boosting 0.864 0.984 0.713 0.667 0.647 0.870 0.435 0.143 0.999 0.999 0.999 0.999 0.982 61.857 0.018 0.999 0.999 0.999 0.999 0.707 2.717 0.421 0.801 0.784 0.786 0.673 1.998 0.494 0.732 0.731 0.732 0.651 1.842 0.545 0.712 0.703 0.705 243.5 167.5 4.0 63.0 81.5 88.0 0.984 0.984 Category Gradient Boosting Support Vector Machine Artificial Neural Network K-Nearest Neightbors Genetic Programming 0.707 0.67 0.677 0.667 0.667 0.673 0.645 0.642 81.5 0.644 0.619 0.623 99.0 0.639

Genetic Programming	0.521	0.521	0.513		0.518	0.516	127.5	130.5	121.5	120.5			0.919		0.638	
ExSTraCS	1.0	1.0	1.0	1.0	1.0	1.0	248.0	252.0	0.0	0.0	1.0	0.0	0.0	1.0	1.0	1.0
$D5 = E_70_bit_mutl$	iplexer	10000	01													
ML Algorithm	_	Accuracy	F1 Score	Sensitivity (Recall)	Specificity	Precision (PPV)	TP	TN	FP	FN	NPV	LR+	LR-	ROC AUC	PRC AUC	PRC APS
Naive Bayes	0.544	0.544	0.541	0.547	0.555	0.544	273.0	278.0	223.0	226.0	0.544	1.198	0.843	0.564	0.564	0.565
Logistic Regression	0.545	0.545	0.541	0.546	0.562	0.546	272.5	281.5	219.5	226.5	0.544	1.207	0.843	0.566	0.563	0.564
Decision Tree	0.516	0.516	0.519	0.513	0.53	0.514	256.0	265.5	235.5	243.0	0.517	1.064	0.94	0.531	0.56	0.535
Random Forest	0.572	0.572	0.57	0.573	0.579	0.574	286.0	290.0	211.0	213.0	0.573	1.355	0.75	0.614	0.614	0.615
Gradient Boosting	0.624	0.624	0.622	0.62	0.625	0.624	309.5	313.0	188.0	189.5	0.624	1.666	0.605	0.666	0.657	0.658
Extreme Gradient Boosting	0.566	0.566	0.56	0.552	0.585	0.568	275.5	293.0	208.0	223.5	0.566	1.32	0.771	0.595	0.586	0.587
Light Gradient Boosting	0.627	0.627	0.633	0.637	0.632	0.624	318.0	316.5	184.5	181.0	0.634	1.669	0.581	0.679	0.67	0.671
Category Gradient Boosting	0.56	0.56	0.546	0.536	0.578	0.563	267.5	289.5	211.5	231.5	0.557	1.292	0.799	0.597	0.594	0.595
Support Vector Machine	0.56	0.56	0.554	0.551	0.57	0.561	275.0	285.5	215.5	224.0	0.56	1.286	0.79	0.588	0.585	0.586
Artificial Neural Network	0.536	0.536	0.554	0.559	0.532	0.534	279.0	266.5	234.5	220.0	0.537	1.15	0.865	0.556	0.561	0.562
K-Nearest Neightbors	0.564	0.564	0.555	0.551	0.58	0.567	275.0	290.5	210.5	224.0	0.565	1.316	0.774	0.591	0.581	0.58
Genetic Programming	0.513	0.513	0.513	0.509	0.519	0.512	254.0	260.0	241.0	245.0	0.514	1.053	0.949	0.513	0.634	0.506
ExSTraCS	0.99	0.99	0.99	0.995	0.988	0.988	496.0	495.0	6.0	2.5	0.995	40.594	0.005	0.996	0.995	0.994
		-									-					

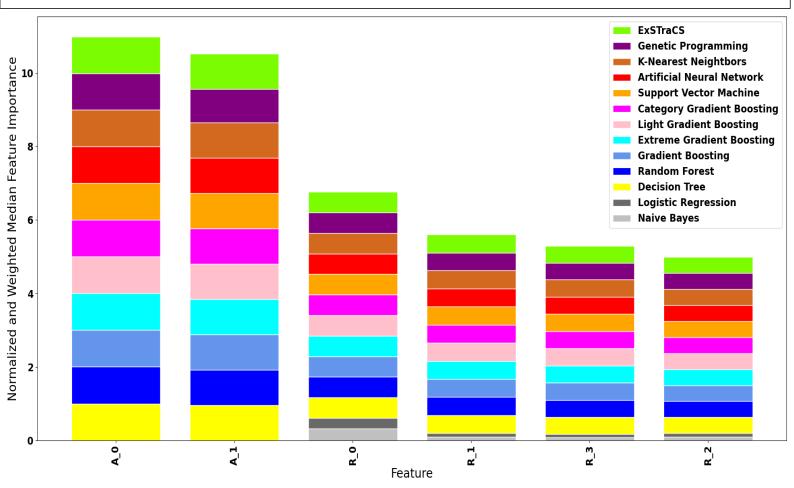
# Median Model Prediction Statistics (Rounded to 3 Decimal Points): Page 2

$D6 = F_135_bit_mut$	tliplexer	_20000	_01													
ML Algorithm	Balanced Accuracy	Accuracy	F1	Sensitivity (Recall)	Specificity	Precision (PPV)	TP	TN	FP	FN	NPV	LR+	LR-	ROC AUC	PRC AUC	PRC APS
Naive Baves	0.517	0.517	Score 0.504	0.495	0.535	0.514	492.0	538.5	468.5	501.0	0.521	1.074	0.934	0.523		0.516
Logistic Regression	0.513	0.513	0.504	0.503	0.523	0.509	499.0	526.5	480.5	494.0	0.515	1.053	0.956	0.522		0.515
Decision Tree	0.502	0.502	0.492	0.489	0.517	0.498	485.5	521.0	486.5	507.0	0.505	1.007	0.993	0.502	0.503	0.499
Random Forest	0.529	0.529	0.52	0.516	0.54	0.528	512.0	544.0	463.0	481.0	0.532	1.137	0.891	0.545	0.537	0.537
Gradient Boosting	0.528	0.528	0.523	0.522	0.547	0.525	518.5	550.5	456.5	474.0	0.531	1.122	0.896	0.54	0.532	0.533
Extreme Gradient Boosting	0.519	0.519	0.499	0.479	0.553	0.518	475.5	557.0	450.0	517.5	0.522	1.089	0.931	0.529	0.518	0.519
Light Gradient Boosting	0.545	0.545	0.537	0.532	0.558	0.542	528.0	562.0	446.0	465.0	0.548	1.2	0.838	0.564	0.551	0.552
Category Gradient Boosting	0.52	0.521	0.451	0.398	0.644	0.524	395.5	649.0	358.0	597.5	0.52	1.116	0.938	0.534	0.526	0.527
Support Vector Machine	0.522	0.522	0.518	0.519	0.53	0.519	515.0	534.0	473.0	478.0	0.526	1.094	0.915	0.534	0.524	0.525
Artificial Neural Network	0.527	0.527	0.518	0.514	0.533	0.524	510.0	536.5	470.5	482.5	0.529	1.119	0.904	0.527	0.523	0.524
K-Nearest Neightbors	0.528	0.529	0.508	0.499	0.545	0.526	495.5	549.0	458.5	497.0	0.53	1.128	0.902	0.535	0.521	0.522
Genetic Programming	0.502	0.502	0.507	0.511	0.498	0.498	507.0	501.5	505.5	486.0	0.505	1.007	0.993		0.627	0.497
EvCTroCC	0.501	0.409	0.607	0.772	0.240	0.407	766.5	251.0	756.0	226.0	0.511	1.002	0.072	0.515	0.506	0.506

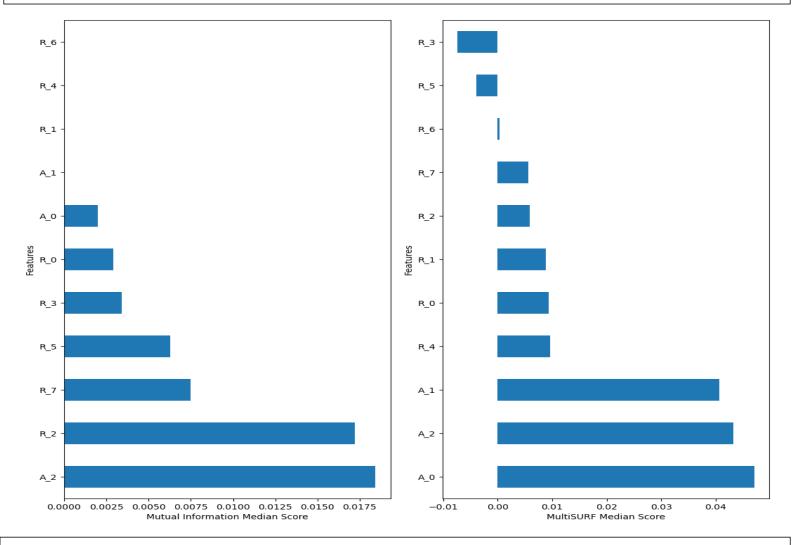
Feature Importance Summary: D1 = A\_6\_bit\_mutliplexer\_500\_01



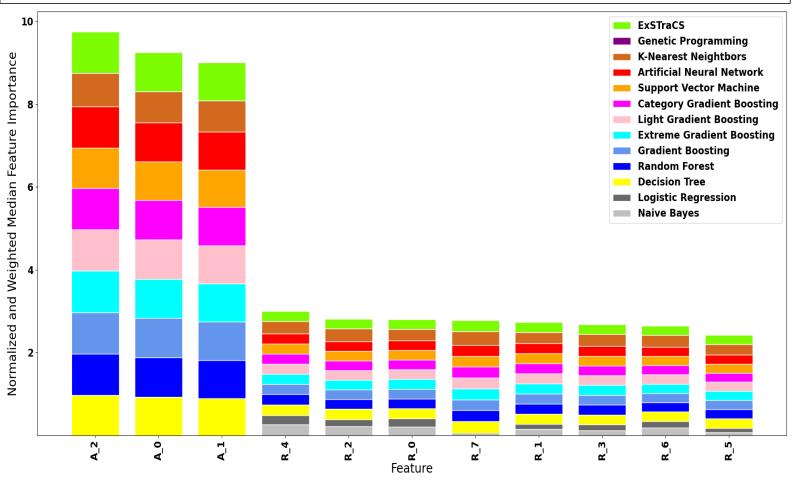
Composite Feature Importance Plot (Normalized and Performance Weighted)



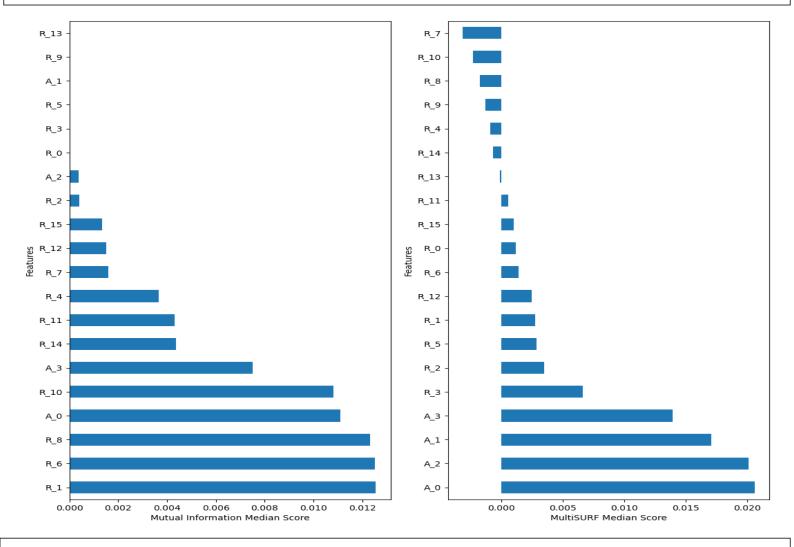
Feature Importance Summary: D2 = B\_11\_bit\_mutliplexer\_1000\_01



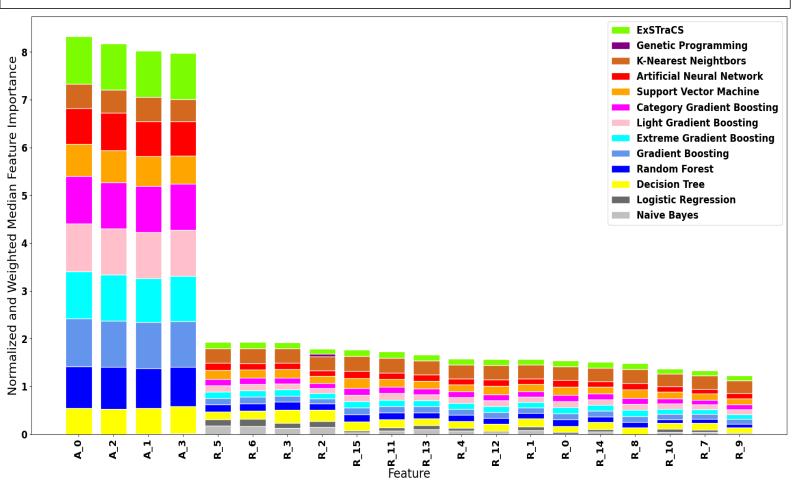
Composite Feature Importance Plot (Normalized and Performance Weighted)



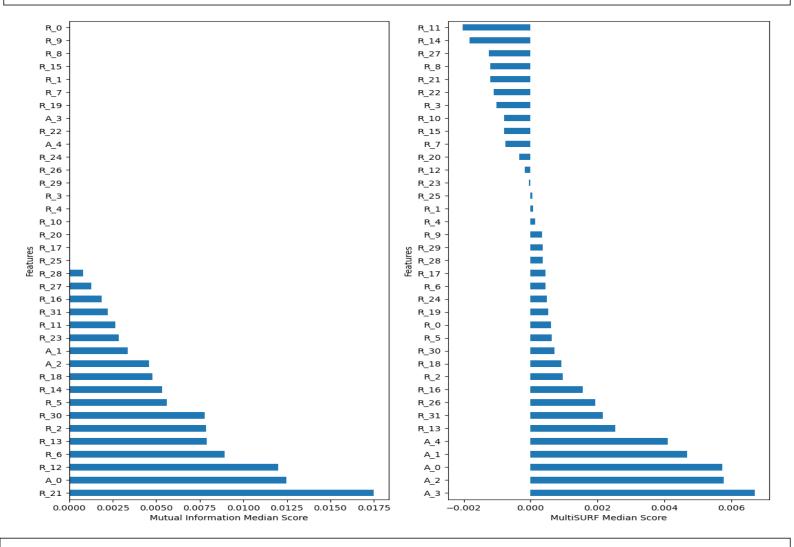
Feature Importance Summary: D3 = C\_20\_bit\_mutliplexer\_2000\_01



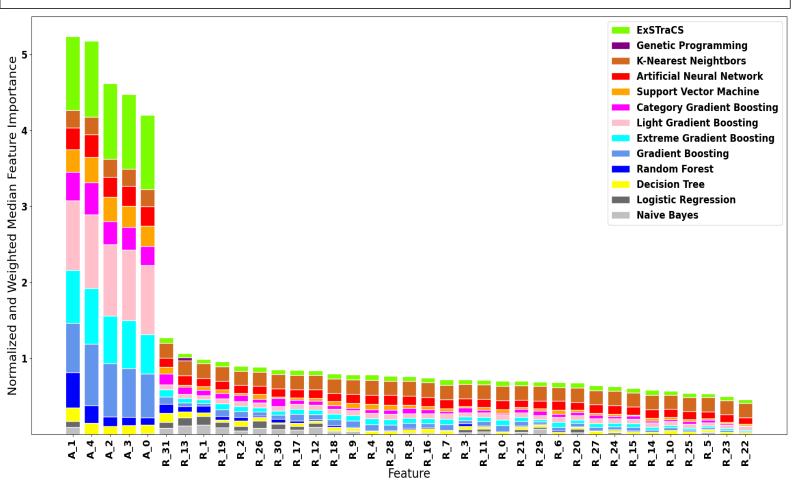
Composite Feature Importance Plot (Normalized and Performance Weighted)



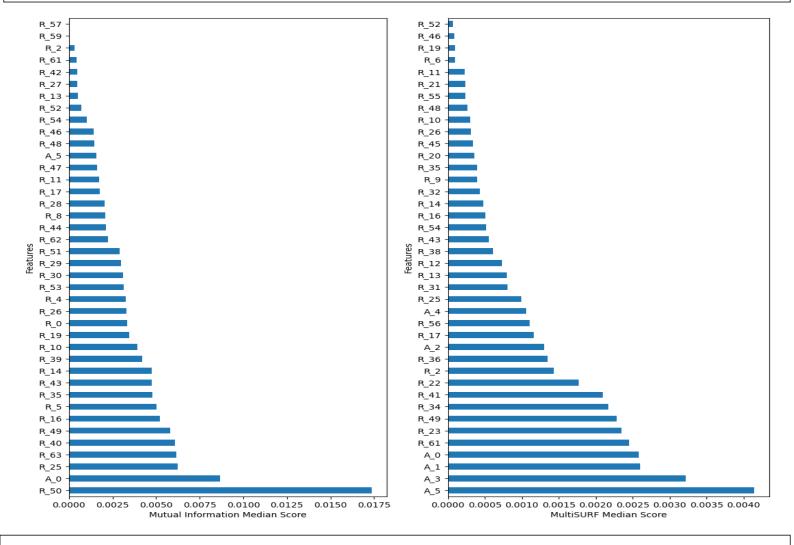
Feature Importance Summary: D4 = D\_37\_bit\_mutliplexer\_5000\_01



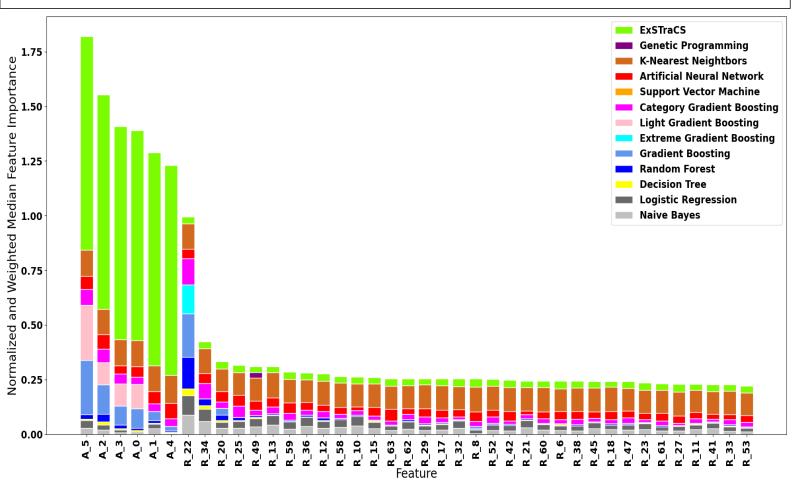
Composite Feature Importance Plot (Normalized and Performance Weighted)



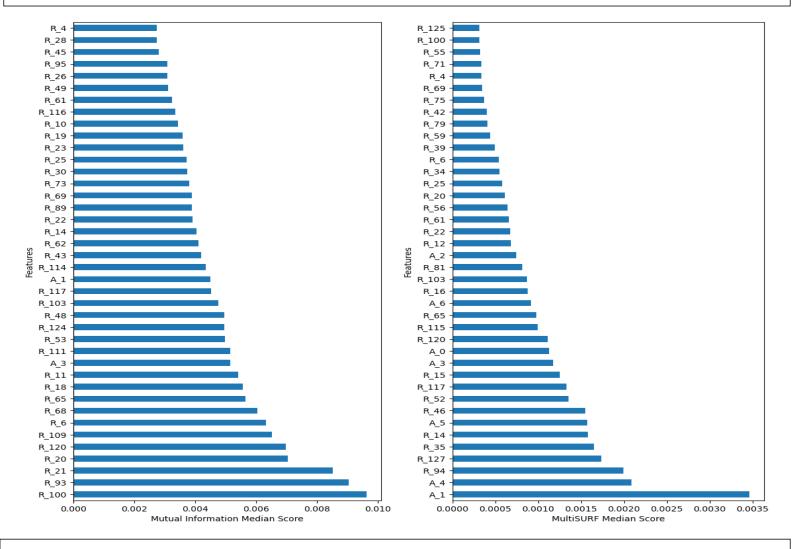
**Feature Importance Summary: D5 = E\_70\_bit\_mutliplexer\_10000\_01** 



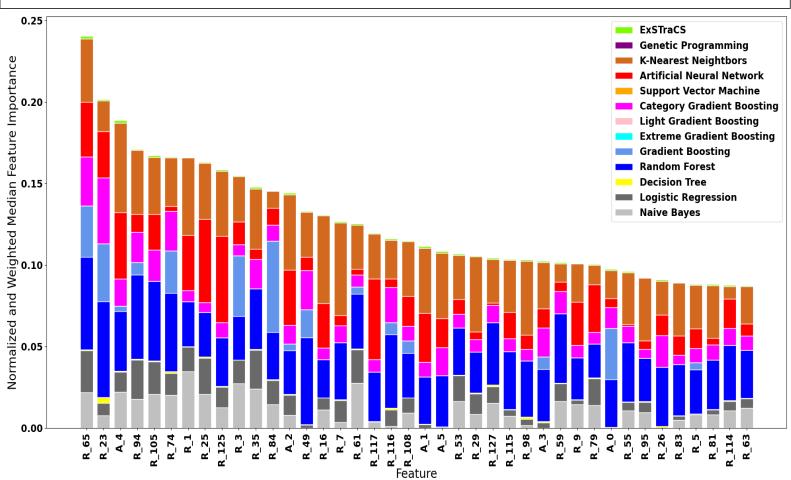
Composite Feature Importance Plot (Normalized and Performance Weighted)



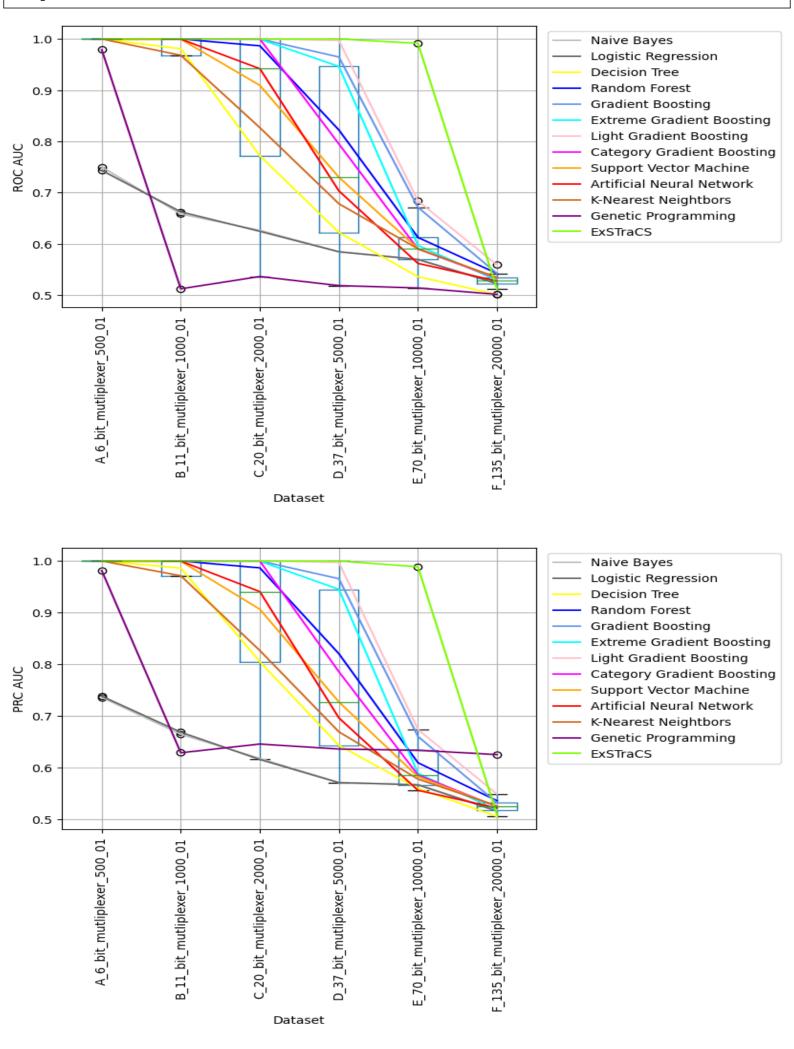
**Feature Importance Summary: D6 = F\_135\_bit\_mutliplexer\_20000\_01** 



Composite Feature Importance Plot (Normalized and Performance Weighted)



## **Compare ML Performance Across Datasets**



# Using Best Performing Algorithms (Kruskall Wallis Compare Datasets)

## Datasets:

D1 = A\_6\_bit\_mutliplexer\_500\_01

 $D2 = B_11_bit_mutliplexer_1000_01$ 

D3 = C\_20\_bit\_mutliplexer\_2000\_01

 $D4 = D_37_bit_mutliplexer_5000_01$ 

D5 = E\_70\_bit\_mutliplexer\_10000\_01

 $D6 = F\_135\_bit\_mutliplexer\_20000\_01$ 

index	P-Value	Best_Alg_D1	Median_D1	Best_Alg_D2	Median_D2	Best_Alg_D3	Median_D3
Balanced Accuracy	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
Accuracy	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
F1 Score	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
Sensitivity (Recall)	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
Specificity	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
Precision (PPV)	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
TP	0.0	Decision Tree	25.0	Random Forest	50.5	Gradient Boosting	98.0
TN	0.0	Decision Tree	25.0	Random Forest	49.5	Gradient Boosting	102.0
FP	0.0	Logistic Regression	10.0	Genetic Programming	23.5	Genetic Programming	47.0
FN	0.0	Logistic Regression	9.5	Genetic Programming	25.0	Genetic Programming	46.0
NPV	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
LR+	0.0196	Naive Bayes	1.9383	Decision Tree	20.0163	Extreme Gradient Boosting	25.5
LR-	0.0	Logistic Regression	0.556	Genetic Programming	0.9984	Genetic Programming	0.8894
ROC AUC	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
PRC AUC	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0
PRC APS	0.0	Decision Tree	1.0	Random Forest	1.0	Gradient Boosting	1.0

index	P-Value	Best_Alg_D4	Median_D4	Best_Alg_D5	Median_D5	Best_Alg_D6	Median_D6
Balanced Accuracy	0.0	ExSTraCS	1.0	ExSTraCS	0.9905		0.5451
Accuracy	0.0	ExSTraCS	1.0	ExSTraCS	0.9905		0.5452
F1 Score	0.0	ExSTraCS	1.0	ExSTraCS	0.9904		0.6066
Sensitivity (Recall)	0.0	ExSTraCS	1.0	ExSTraCS	0.995	ExSTraCS	0.7723
Specificity	0.0	ExSTraCS	1.0	ExSTraCS	0.988	Category Gradient Boosting	0.6445
Precision (PPV)	0.0	ExSTraCS	1.0	ExSTraCS	0.9881	Light Gradient Boosting	0.5418
TP	0.0	ExSTraCS	248.0	ExSTraCS	496.0	ExSTraCS	766.5
TN	0.0	ExSTraCS	252.0	ExSTraCS	495.0	Category Gradient Boosting	649.0
FP	0.0	Genetic Programming	121.5	Genetic Programming	241.0	ExSTraCS	756.0
FN	0.0	Genetic Programming	120.5	Genetic Programming	245.0	Category Gradient Boosting	597.5
NPV	0.0	ExSTraCS	1.0	ExSTraCS	0.9949	Light Gradient Boosting	0.5477
LR+	0.0196	Light Gradient Boosting	61.8569	ExSTraCS	40.5937	Light Gradient Boosting	1.2003
LR-	0.0	Genetic Programming	0.9189	Genetic Programming	0.9492	Genetic Programming	0.9931
ROC AUC	0.0	ExSTraCS	1.0	ExSTraCS	0.9958	Light Gradient Boosting	0.5642
PRC AUC	0.0	ExSTraCS	1.0	ExSTraCS	0.9949		0.6272
PRC APS	0.0	ExSTraCS	1.0	ExSTraCS	0.9944	Light Gradient Boosting	0.5519

## **Pipeline Runtime Summary**

A_6_bit_mutliplexer_500	_01	B_11_bit_mutliplexer_1000_01						
Pipeline Component	Time (sec)	Pipeline Component	Time (sec)					
Exploratory Analysis	0.67	Exploratory Analysis	0.81					
Preprocessing	0.02	Preprocessing	0.03					
Mutual Information	0.22	Mutual Information	0.47					
MultiSURF	7.04	MultiSURF	36.71					
Feature Selection	0.24	Feature Selection	0.27					
Naive Bayes	0.65	Naive Bayes	1.47					
Logistic Regression	44.39	Logistic Regression	58.18					
Decision Tree	48.91	Decision Tree	52.11					
Random Forest	3980.24	Random Forest	3012.63					
Gradient Boosting	1488.16	Gradient Boosting	3614.84					
Extreme Gradient Boosting	6447.96	Extreme Gradient Boosting	8154.53					
Light Gradient Boosting	416.71	Light Gradient Boosting	1127.0					
Category Gradient Boosting	9420.65	Category Gradient Boosting	9320.95					
Support Vector Machine	1874.1	Support Vector Machine	1010.71					
Artificial Neural Network	2364.51	Artificial Neural Network	5922.24					
K-Nearest Neightbors	92.67	K-Nearest Neightbors	123.02					
Genetic Programming	14222.49	Genetic Programming	12499.19					
ExSTraCS	5436.65	ExSTraCS	34330.41					
Stats Summary	16.61	Stats Summary	16.06					

C_20_bit_mutliplexer_20	00_01	D_37_bit_mutliplexer_5000_01						
Pipeline Component	Time (sec)	Pipeline Component	Time (sec)					
Exploratory Analysis	1.17	Exploratory Analysis	2.55					
Preprocessing	0.05	Preprocessing	0.17					
Mutual Information	1.35	Mutual Information	5.44					
MultiSURF	220.78	MultiSURF	552.11					
Feature Selection	0.33	Feature Selection	0.48					
Naive Bayes	4.18	Naive Bayes	15.91					
Logistic Regression	85.61	Logistic Regression	194.49					
Decision Tree	68.9	Decision Tree	124.38					
Random Forest	8546.24	Random Forest	11109.55					
Gradient Boosting	8547.08	Gradient Boosting	9739.42					
Extreme Gradient Boosting	9170.34	Extreme Gradient Boosting	9643.67					
Light Gradient Boosting	2363.66	Light Gradient Boosting	6106.74					
Category Gradient Boosting	10655.57	Category Gradient Boosting	11145.88					
Support Vector Machine	7783.43	Support Vector Machine	16191.87					
Artificial Neural Network	8910.17	Artificial Neural Network	9194.73					
K-Nearest Neightbors	917.35	K-Nearest Neightbors	39670.32					
Genetic Programming	10558.32	Genetic Programming	11294.3					
ExSTraCS	56322.18	ExSTraCS	96453.65					
Stats Summary	17.7	Stats Summary	22.13					

E_70_bit_mutliplexer_10000_01		F_135_bit_mutliplexe	F_135_bit_mutliplexer_20000_01	
Pipeline Component	Time (sec)	Pipeline Component	Time (sec)	
Exploratory Analysis	7.1	Exploratory Analysis	22.12	
Preprocessing	0.45	Preprocessing	1.62	
Mutual Information	21.09	Mutual Information	87.74	
MultiSURF	967.6	MultiSURF	1920.68	
Feature Selection	0.56	Feature Selection	0.53	
Naive Bayes	88.41	Naive Bayes	491.1	
Logistic Regression	479.16	Logistic Regression	2953.84	
Decision Tree	299.64	Decision Tree	834.76	
Random Forest	14742.59	Random Forest	31430.96	
Gradient Boosting	11423.64	Gradient Boosting	18523.58	
Extreme Gradient Boosting	11716.09	Extreme Gradient Boosting	15423.16	
Light Gradient Boosting	13558.08	Light Gradient Boosting	28213.6	
Category Gradient Boosting	14275.26	Category Gradient Boosting	20532.51	
Support Vector Machine	53373.8	Support Vector Machine	54617.27	
Artificial Neural Network	8158.81	Artificial Neural Network	8524.78	
K-Nearest Neightbors	184243.87	K-Nearest Neightbors	332107.3	
Genetic Programming	9675.91	Genetic Programming	10696.83	
ExSTraCS	242323.1	ExSTraCS	876241.06	
Stats Summary	25.06	Stats Summary	25.7	