**Results for Solver and PC as separate attributes for ML**

4648 data points, all features

Filename: solver\_pc\_separate1.arff

BayesNet

Steps: 1. Make class as last column in arff

2. Remove diagonal-dominance

|  |  |  |
| --- | --- | --- |
| **Combination** | **All feat. Accuracy with 10cv**  **(%)** | **All feat. Accuracy with 66-34% split (%)** |
| Solver+solver\_name+pc\_name | 87.7 | 88.6 |
| solver | 87.8 | 89.2 |
| Solver separate, PC separate | 87.7 | 88.6 |
| Solver\_name | 87.7 | 88.4 |
| Pc\_name | 87.5 | 88.4 |
| Solver + solver\_name | 87.8 | 89.0 |
| Solver + pc\_name | 87.7 | 88.6 |

With Reduced Set 1 : (8 features)

BayesNet

|  |  |  |
| --- | --- | --- |
| **Combination** | **RS1 Accuracy with 10cv**  **(%)** | **All feat. Accuracy with 66-34% split (%)** |
| Solver+solver\_name+pc\_name | 86.7 | 83.8 |
| solver | 87.3 | 88.1 |
| Solver separate, PC separate | 86.4 | 86.2 |
| Solver\_name | 87.4 (improved) | 89.5(improved) |
| Pc\_name | 86.5 | 85.7 |
| Solver + solver\_name | 86.9 | 88.1 |
| Solver + pc\_name | 86.2 | 84.6 |

With Reduced Set 1 : (6 features)

BayesNet

|  |  |  |
| --- | --- | --- |
| **Combination** | **RS2 Accuracy with 10cv**  **(%)** | **All feat. Accuracy with 66-34% split (%)** |
| Solver+solver\_name+pc\_name | 84.1 | 80.0 |
| solver | 86.9 | 81.4 |
| Solver separate, PC separate | 86.0 | 81.8 |
| Solver\_name | 87.5 (improved) | 86.4 (improved) |
| Pc\_name | 85.4 | 81.4 |
| Solver + solver\_name | 85.7 | 83.8 |
| Solver + pc\_name | 84.9 | 80.5 |

With all features

Solver\_name + pc\_name solver\_name solver

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ML mtd** | **With solver separate + pc separate (Accuracy)** | | **With only solver separate (Accuracy)** | | **With solver and pc together (Accuracy)** | |
|  | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** |
| LibSVM | 68.3 | 64.5 | 68.3 | 64.5 | 45.8 | 37.9 |
| RF | 87.7 | 86.2 | 78.5 | 77.2 | 74.7 | 71.9 |
| BN | 87.7 | 88.6 | 87.7 | 88.4 | 87.7 | 89.2 |
| knn | 55.2 | 52.2 | 42.4 | 43.0 | 29.2 | 30.9 |
| ADT | 66.8 | 69.8 | 47.0 | 66.2 | 65.3 | 66.5 |
| J48 | 84.1 | 83.0 | 78.7 | 74.8 | 76.8 | 78.3 |

With RS1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ML mtd** | **With solver separate + pc separate (Accuracy)** | | **With only solver separate (Accuracy)** | | **With solver and pc together (Accuracy) single** | |
|  | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** |
| LibSVM | 64.5 | 63.8 | 61.7 | 59.7 | 43.8 | 37.7 |
| RF | 87.0 | 84.4 | 77.8 | 75.6 | 79.6 | 76.5 |
| BN | 86.4 | 86.2 | 87.4 | 89.5 | 87.3 | 88.1 |
| knn | 68.5 | 62.1 | 67.7 | 61.9 | 68.4 | 66.0 |
| ADT | 61.6 | 55.5 | 44.8 | 44.9 | 44.7 | 57.4 |
| J48 | 82.3 | 78.5 | 77.2 | 76.3 | 78.3 | 76.5 |

With RS2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ML mtd** | **With solver separate + pc separate (Accuracy)** | | **With only solver separate (Accuracy)** | | **With solver and pc together (Accuracy)** | |
|  | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** |
| LibSVM | 65.4 | 64.3 | 62.3 | 60.7 | 44.9 | 38.8 |
| RF | 84.7 | 82.0 | 77.4 | 75.6 | 78.6 | 76.8 |
| BN | 86.0 | 81.8 | 87.5 | 86.4 | 86.9 | 81.4 |
| knn | 69.7 | 62.5 | 64.4 | 62.9 | 69.0 | 64.5 |
| ADT | 66.0 | 70.2 | 57.3 | 61 | 63.4 | 71.3 |
| J48 | 77.6 | 75.7 | 76.7 | 71.7 | 75.1 | 76.8 |

On aciss data:

File: solver\_pc\_separate\_aciss\_1.arff

3049 instances

2250 bad, 799 good

All features

68 🡪 56 feat (removeuseless)

change class numeric to -1,1 in arff file

With all features

Solver\_name + pc\_name solver\_name solver

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **ML mtd** | **With solver separate + pc separate (Accuracy)** | | **With only solver separate (Accuracy)** | | **With solver and pc together (Accuracy)** | |
|  | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** |
| LibSVM | 29.9 | 99.5 | 29.9 | 99.5 (not good as predicts good for almost all pts) | 29.9 | 99.5 |
| RF | 93.6 | 94.9 | 93.8 | 94.7 | 85.6 | 85.3 |
| BN | 84.3 | 84.8 | 84.3 | 84.8 | 84.3 | 84.8 |
| knn | 92.5 | 93.5 | 92.1 | 92.5 | 94.7 | 95.3 |
| ADT | 90.0 | 89.0 | 89.6 | 85.6 | 91.1 | 85.6 |
| J48 | 91.0 | 92.3 | 91.0 | 92.3 | 91.0 | 92.3 |
| VFI | Bad | Bad | Bad | bad | 36.5 | 56.4 |

For solver alone: (for last 2 columns)

ADT 10cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.911 0.355 0.878 0.911 0.894 0.892 -1

0.645 0.089 0.719 0.645 0.68 0.892 1

Weighted Avg. 0.841 0.286 0.837 0.841 0.838 0.892

=== Confusion Matrix ===

a b <-- classified as

2049 201 | a = -1

284 515 | b = 1

Knn 10cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.947 0.745 0.782 0.947 0.857 0.795 -1

0.255 0.053 0.632 0.255 0.364 0.795 1

Weighted Avg. 0.766 0.563 0.742 0.766 0.727 0.795

=== Confusion Matrix ===

a b <-- classified as

2131 119 | a = -1

595 204 | b = 1

For solver\_name+pc\_name

BayesNet 10cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.843 0.467 0.836 0.843 0.839 0.766 -1

0.533 0.157 0.546 0.533 0.54 0.766 1

Weighted Avg. 0.762 0.386 0.76 0.762 0.761 0.766

=== Confusion Matrix ===

a b <-- classified as

1896 354 | a = -1

373 426 | b = 1

Knn k = 10, 10cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.925 0.329 0.888 0.925 0.906 0.917 -1

0.671 0.075 0.76 0.671 0.713 0.917 1

Weighted Avg. 0.858 0.263 0.854 0.858 0.855 0.917

=== Confusion Matrix ===

a b <-- classified as

2081 169 | a = -1

263 536 | b = 1

ADT 10cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.9 0.38 0.87 0.9 0.885 0.894 -1

0.62 0.1 0.688 0.62 0.652 0.894 1

Weighted Avg. 0.827 0.307 0.822 0.827 0.824 0.894

=== Confusion Matrix ===

a b <-- classified as

2026 224 | a = -1

304 495 | b = 1

J48 10cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.91 0.269 0.905 0.91 0.907 0.83 -1

0.731 0.09 0.742 0.731 0.736 0.83 1

Weighted Avg. 0.863 0.222 0.862 0.863 0.863 0.83

=== Confusion Matrix ===

a b <-- classified as

2047 203 | a = -1

215 584 | b = 1

RF 10 cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.936 0.339 0.886 0.936 0.91 0.926 -1

0.661 0.064 0.786 0.661 0.718 0.926 1

Weighted Avg. 0.864 0.267 0.86 0.864 0.86 0.926

=== Confusion Matrix ===

a b <-- classified as

2106 144 | a = -1

271 528 | b = 1

LibSVM 10CV

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.299 0.288 0.745 0.299 0.427 0.506 -1

0.712 0.701 0.265 0.712 0.386 0.506 1

Weighted Avg. 0.407 0.396 0.619 0.407 0.416 0.506

=== Confusion Matrix ===

a b <-- classified as

673 1577 | a = -1

230 569 | b = 1

With RS1

Solver\_name + pc\_name solver

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ML mtd** | **With solver separate + pc separate (Accuracy)** | | **With solver and pc together (Accuracy) single** | |
|  | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** |
| LibSVM | 49.9 | 96.5 | 29.9 | 99.0 (13.5 bad accuracy) |
| RF | 92.5 | 93.3 | 91.5 | 92.9 |
| BN | 84.2 | 84.8 | 84.3 | 84.8 |
| knn | 91.1 | 90.5 | 91.5 | 91.3 |
| ADT | 84.1 | 78.9 | 89.2 | 90.7 |
| J48 | 89.7 | 88.5 | 90.2 | 88.6 |

RF 10cv:

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.925 0.314 0.892 0.925 0.909 0.929 -1

0.686 0.075 0.765 0.686 0.723 0.929 1

Weighted Avg. 0.863 0.251 0.859 0.863 0.86 0.929

=== Confusion Matrix ===

a b <-- classified as

2082 168 | a = -1

251 548 | b = 1

With RS2

Solver\_name + pc\_name solver

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ML mtd** | **With solver separate + pc separate (Accuracy)** | | **With solver and pc together (Accuracy)** | |
|  | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** |
| LibSVM | Too long |  | Too long |  |
| RF | 91.9 | 92.3 | 91.0 | 92.6 |
| BN | 84.2 | 84.8 | 84.3 | 84.8 |
| knn | 90.7 | 90.3 | 90.7 | 91.3 |
| ADT | 93.7 | 79.2 | 92.6 | 86.4 |
| J48 | 89.0 | 91.4 | 88.8 | 91.7 |
| VFI | 36.4 | 36.6 | 8.6 | 41.4 |

On aciss data:

File: solver\_pc\_separate\_aciss\_30.arff

Petsc\_anamod\_aciis\_30.arff

3049 instances

bad: 2037, good: 1012

All features

68🡪 57 feat (removeuseless)

With all features

Solver\_name + pc\_name solver

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ML mtd** | **With solver separate + pc separate (Accuracy)** | | **With solver and pc together (Accuracy)** | |
|  | **10-cv** | **66-34%split** | **10-cv** | **66-34%split** |
| LibSVM | 45.4 | 80.6 | 39.9 | 99.7 |
| RF | 92.4 | 93.4 | 79.8 | 83.5 |
| BN | 56.0 | 78.8 | 84.6 | 85.6 |
| knn | 91.4 | 92.4 | 89.4 | 91.4 |
| ADT | 91.8 | 93.1 | 86.5 | 90.2 |
| J48 | 91.5 | 91.1 | 91.2 | 91.8 |
| VFI | 40.3 | 40.6 | 63.5 | 54.1 |

RF 10cv

=== Detailed Accuracy By Class ===

TP Rate FP Rate Precision Recall F-Measure ROC Area Class

0.924 0.214 0.897 0.924 0.91 0.94 -1

0.786 0.076 0.837 0.786 0.81 0.94 1

Weighted Avg. 0.878 0.169 0.877 0.878 0.877 0.94

=== Confusion Matrix ===

a b <-- classified as

1882 155 | a = -1

217 795 | b = 1