

Avaliando algoritmos de classificação em Fluxos de dados aplicados na detecção de ataques à rede usando o MOA (*Massive Online Analysis*)

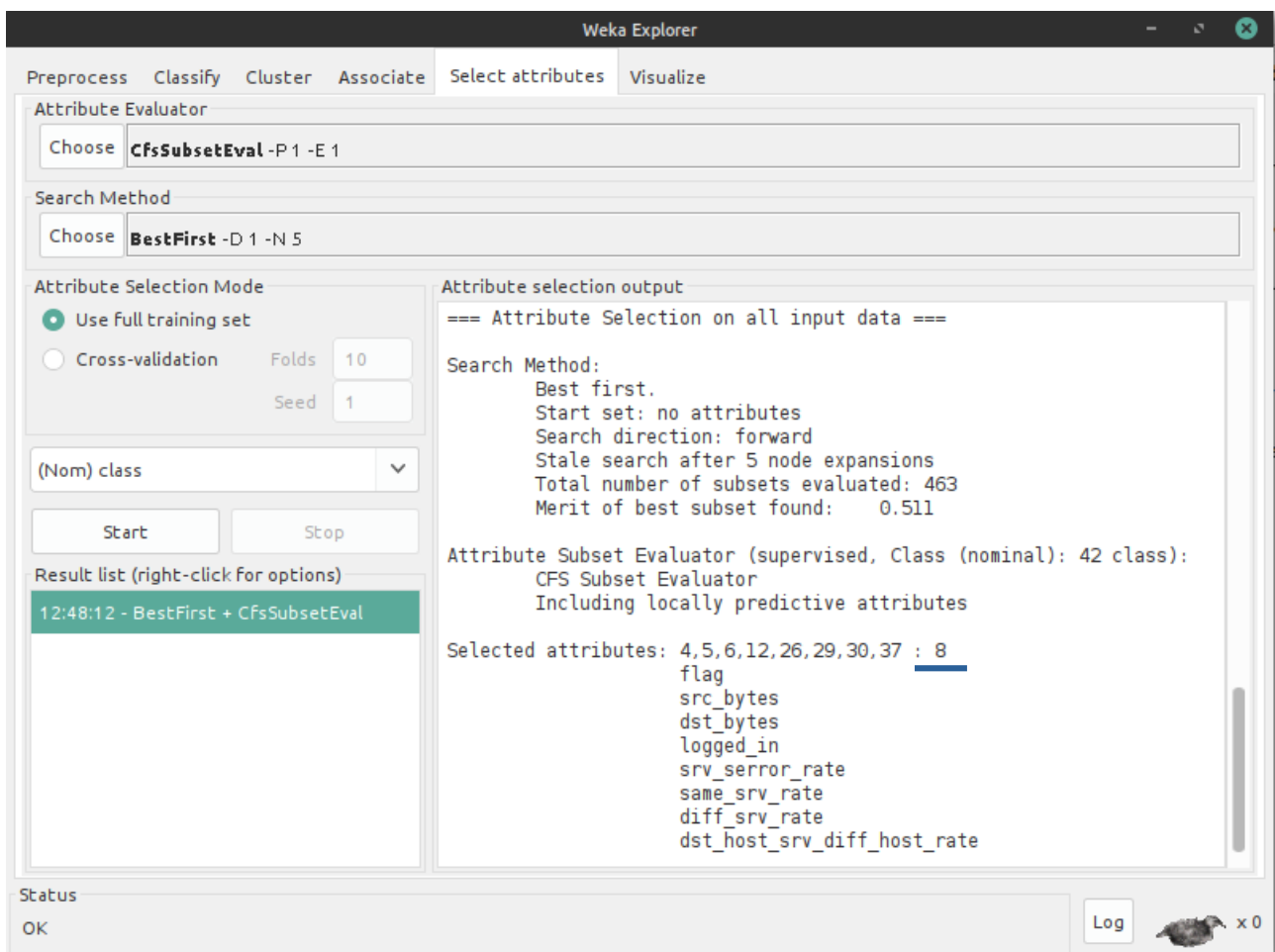
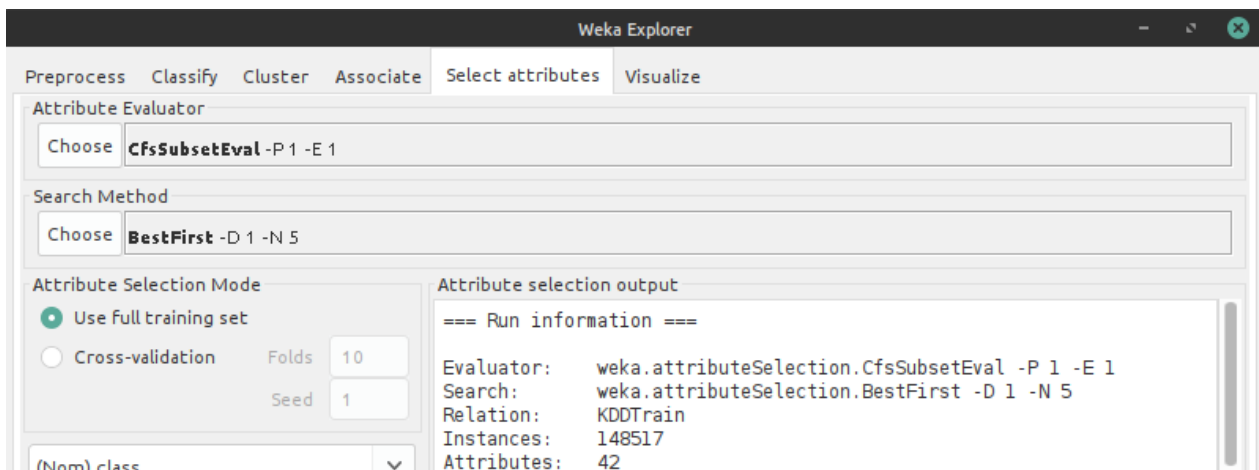
(Documentação dos Experimentos)

Alunos:

Augusto Parisot de Gusmão Neto

Marcelo Marques da Rocha

Tela do ambiente Weka Explorer, acessado através da interface do MOA. Redução do número de atributos, dos 42 (41 atributos mais o atributo *class*) contidos no dataset original, foram selecionados apenas 8.



A partir daqui, um novo dataset, contendo apenas os atributos selecionados pelo algoritmo, será salvo. Será um dataset com o mesmo número de instâncias (148.517), porém, com 9 atributos (8 selecionados mais o atributo *class*).

Treinamento dos modelos NSL-KDD-Train42

```
LearnModel -l bayes.NaiveBayes -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-naivebayes-42.moa
```

```
LearnModel -l trees.DecisionStump -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-decisionstump-42
```

```
LearnModel -l functions.MajorityClass -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-majority-42.moa
```

```
LearnModel -l functions.NoChange -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-nochange-42.moa
```

```
LearnModel -l functions.Perceptron -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-perceptron-42.moa
```

```
LearnModel -l lazy.kNN -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-lazyknn-42.moa
```

```
LearnModel -l trees.HoeffdingAdaptiveTree -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-hoefadatree-42.moa
```

```
LearnModel -l trees.HoeffdingOptionTree -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-hoefopttree-42.moa
```

```
LearnModel -l trees.HoeffdingTree -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN42.arff) -O /home/marcelo/KDD_final/mod-hoefmtree-42.moa
```

Avaliação Prequential dos modelos NSL-KDD-Test42 (Janela = 1000)

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-naivebayes-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-multinbayes-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-hoefadatree-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-hoefopttree-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-hoefmtree-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-lazyknn-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-majority-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-nochange-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-perceptron-42.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST42.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

Resultados (Métricas de Avaliação) – Dataset 1 [42 atributos]

Acurácia dos Classificadores (Média e Desvio padrão)

Summary Viewer

Algorithm	mod-naivebayes-42	mod-majority-42	mod-hoeftree-42	mod-decisionstump-42	mod-nochange-42	mod-perceptron-42	mod-hoeftadtree-42	mod-hoeftree-42	mod-lazyknn-42
NSL-KDD42_atributos	78,88±1,80	43,03±1,79	94,09±0,90	81,63±2,48	50,60±1,93	43,03±1,79	94,97±1,33	94,05±0,92	95,07±1,31

Summary classifications correct (percent) Export Summaries

Tempo de Avaliação (CPU seconds)

Summary Viewer									
Algorithm	mod-naivebayes-42	mod-majority-42	mod-hoeftpttree-42	mod-decisionstump-42	mod-nochange-42	mod-perceptron-42	mod-hoeftadtree-42	mod-hoeftree-42	mod-lazyknn-42
NSL-KDD42_atributos	0,14±0,07	0,11±0,06	0,93±0,55	2,62±1,77	0,16±0,06	0,16±0,07	0,79±0,43	0,68±0,41	5,99±3,36

Precisão para a Classe 0 (normal)

Summary Viewer									
Algorithm	mod-naivebayes-42	mod-majority-42	mod-hoefopttree-42	mod-decisionstump-42	mod-nochange-42	mod-perceptron-42	mod-hoefadatree-42	mod-hoefetree-42	mod-lazyknn-42
NL-KDD42_atributos	69,10±2,76	43,03±1,79	91,22±1,87	73,94±3,53	42,54±2,42	43,03±1,79	91,88±2,75	91,15±1,91	95,01±2,54
<div> Summary Precision for class 0 (percent) ▼ Export Summaries </div>									

Precisão para a Classe 1 (ataque)

Summary Viewer									
Algorithm	mod-naivebayes-42	mod-majority-42	mod-hoefopttree-42	mod-decisionstump-42	mod-nochange-42	mod-perceptron-42	mod-hoefadatree-42	mod-hoefetree-42	mod-lazyknn-42
NSL-KDD42_atributos	92,11±1,08	0,00±0,00	96,43±0,52	89,95±1,28	56,58±2,67	4,35±20,39	97,59±0,82	96,42±0,53	95,19±0,84

Janela da Sessão de Análise dos Resultados (Dataset 1)

Nesta janela, temos os classificadores que serão “ranqueados”. No caso da comparação de mais de dois modelos, em Demsar (2006) há a recomendação da utilização do teste de Friedman. Esse teste é baseado na comparação de ranqueamentos de desempenhos. Como todos os algoritmos estão sendo comparados entre si em pares, os valores de q_0 serão fornecidos pela estatística de Nemenyi (Nemenyi, 1963).

MOA Graphical User Interface

Experiments Summary Plot **Analyze**

Configuration

Result Folder:

Algorithm	Algorithm ID
mod-naivebayes-42.txt	mod-naivebayes-42
mod-majority-42.txt	mod-majority-42
mod-hoeffpttree-42.txt	mod-hoeffpttree-42
mod-decisionstump-42.txt	mod-decisionstump-42
mod-nochange-42.txt	mod-nochange-42
mod-perceptron-42.txt	mod-perceptron-42
mod-hoeffadtree-42.txt	mod-hoeffadtree-42

Stream	Stream ID
NSL-KDD42_atributos	NSL-KDD42_atributos

Statistical Test

Test:

pValue:

measure:

type:

Output

```
mod-hoeffadtree-42 vs mod-hoefftree-42: 21.80075818807247
mod-naivebayes-42 vs mod-decisionstump-42: 28.66512293055501
mod-naivebayes-42 vs mod-nochange-42: 28.66512293055501
mod-majority-42 vs mod-perceptron-42: 28.66512293055501
mod-hoeffpttree-42 vs mod-hoeffadtree-42: 28.66512293055501
mod-hoeffpttree-42 vs mod-hoefftree-42: 28.66512293055501
mod-decisionstump-42 vs mod-hoefftree-42: 28.66512293055501
mod-nochange-42 vs mod-perceptron-42: 28.66512293055501
mod-hoeffadtree-42 vs mod-lazyknn-42: 28.66512293055501
```

Ranqueamento dos Algoritmos (Dataset 1 [42 atributos])

P-values involving all algorithms

P-value computed by Friedman Test: 0.4334701203765876

P-value computed by Iman and Daveport Test: NaN

Ranking of the algorithms

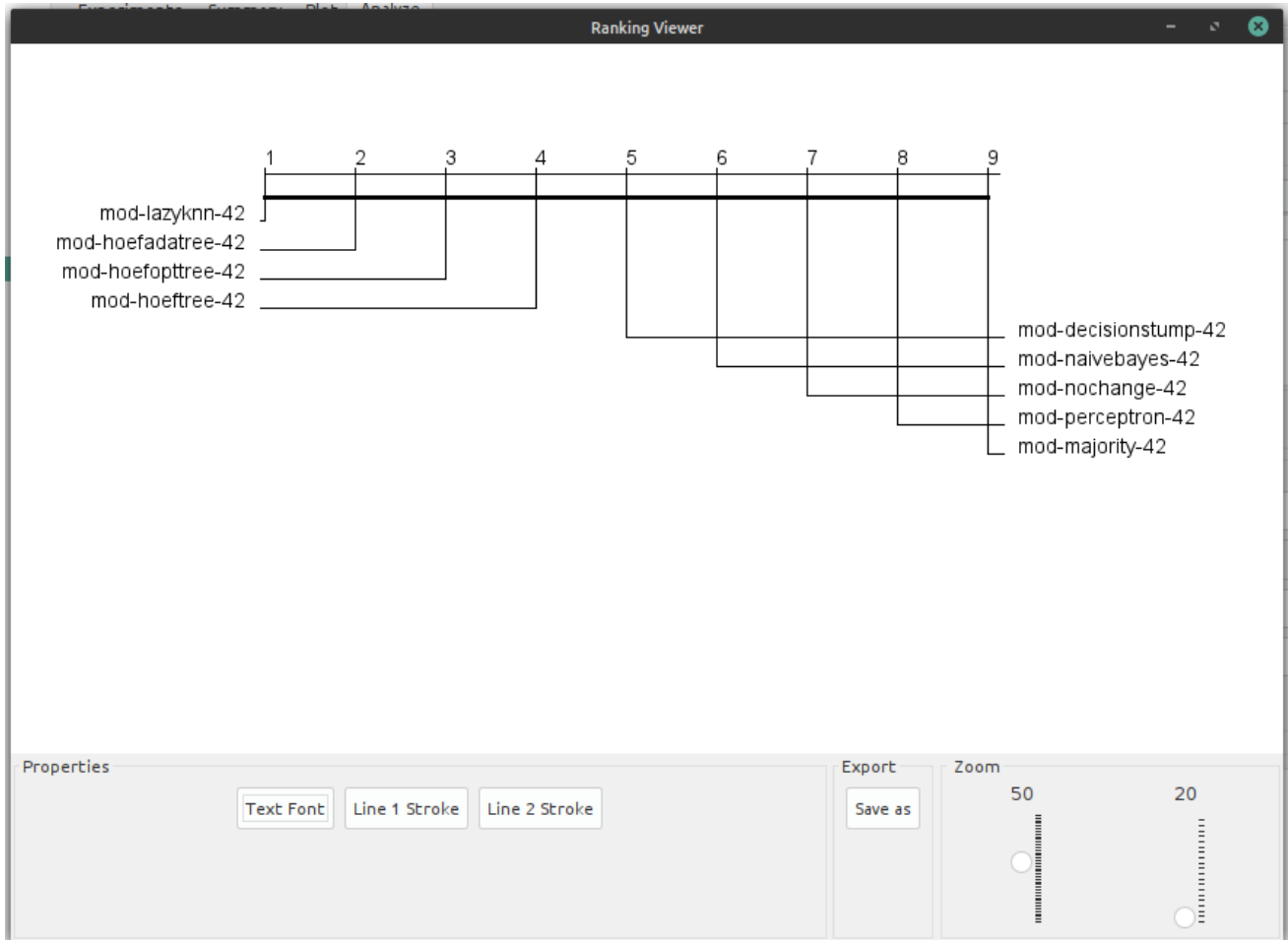
mod-lazyknn-42: 1.0
mod-hoefadatree-42: 2.0
mod-hoefopttree-42: 3.0
mod-hoefmtree-42: 4.0
mod-decisionstump-42: 5.0
mod-naivebayes-42: 6.0
mod-nochange-42: 7.0
mod-perceptron-42: 8.0
mod-majority-42: 9.0

P-values of classifiers against each other

mod-majority-42 vs mod-lazyknn-42: 1.3992157372470226
mod-majority-42 vs mod-hoefadatree-42: 2.5452412151753903
mod-perceptron-42 vs mod-lazyknn-42: 2.5452412151753903
mod-majority-42 vs mod-hoefopttree-42: 4.368069012905356
mod-nochange-42 vs mod-lazyknn-42: 4.368069012905356
mod-perceptron-42 vs mod-hoefadatree-42: 4.368069012905356
mod-naivebayes-42 vs mod-lazyknn-42: 7.081401688522092
mod-majority-42 vs mod-hoefmtree-42: 7.081401688522092
mod-hoefopttree-42 vs mod-perceptron-42: 7.081401688522092
mod-nochange-42 vs mod-hoefadatree-42: 7.081401688522092
mod-naivebayes-42 vs mod-hoefadatree-42: 10.861184969220528
mod-majority-42 vs mod-decisionstump-42: 10.861184969220528
mod-hoefopttree-42 vs mod-nochange-42: 10.861184969220528
mod-decisionstump-42 vs mod-lazyknn-42: 10.861184969220528
mod-perceptron-42 vs mod-hoefmtree-42: 10.861184969220528
mod-naivebayes-42 vs mod-majority-42: 15.788808938915995
mod-naivebayes-42 vs mod-hoefopttree-42: 15.788808938915995
mod-decisionstump-42 vs mod-perceptron-42: 15.788808938915995
mod-decisionstump-42 vs mod-hoefadatree-42: 15.788808938915995
mod-nochange-42 vs mod-hoefmtree-42: 15.788808938915995
mod-hoefmtree-42 vs mod-lazyknn-42: 15.788808938915995
mod-naivebayes-42 vs mod-perceptron-42: 21.80075818807247
mod-naivebayes-42 vs mod-hoefmtree-42: 21.80075818807247
mod-majority-42 vs mod-nochange-42: 21.80075818807247
mod-hoefopttree-42 vs mod-decisionstump-42: 21.80075818807247
mod-hoefopttree-42 vs mod-lazyknn-42: 21.80075818807247
mod-decisionstump-42 vs mod-nochange-42: 21.80075818807247
mod-hoefadatree-42 vs mod-hoefmtree-42: 21.80075818807247
mod-naivebayes-42 vs mod-decisionstump-42: 28.66512293055501
mod-naivebayes-42 vs mod-nochange-42: 28.66512293055501
mod-majority-42 vs mod-perceptron-42: 28.66512293055501
mod-hoefopttree-42 vs mod-hoefadatree-42: 28.66512293055501
mod-hoefopttree-42 vs mod-hoefmtree-42: 28.66512293055501
mod-decisionstump-42 vs mod-hoefmtree-42: 28.66512293055501
mod-nochange-42 vs mod-perceptron-42: 28.66512293055501
mod-hoefadatree-42 vs mod-lazyknn-42: 28.66512293055501

Representação Gráfica de Pós-testes (Dataset 1)

Em Demsar (2006) é apresentada uma maneira gráfica de dispor os resultados dos pós-testes.



Treinamento dos modelos NSL-KDD-Train9

```
LearnModel -l bayes.NaiveBayes -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-naivebayes-9.moa
```

```
LearnModel -l trees.DecisionStump -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-decisionstump-9.moa
```

```
LearnModel -l functions.MajorityClass -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-majority-9.moa
```

```
LearnModel -l functions.NoChange -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-nochange-9.moa
```

```
LearnModel -l functions.Perceptron -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-perceptron-9.moa
```

```
LearnModel -l lazy.kNN -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-lazyknn-9.moa
```

```
LearnModel -l trees.HoeffdingAdaptiveTree -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-hoefadatree-9.moa
```

```
LearnModel -l trees.HoeffdingOptionTree -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-hoefopttree-9.moa
```

```
LearnModel -l trees.HoeffdingTree -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TRAIN9.arff) -O /home/marcelo/KDD_final/mod-hoefmtree-9.moa
```

Avaliação Prequential dos modelos NSL-KDD-Test9 (Janela = 1000)

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-naivebayes-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-decisionstump-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-hoefadatree-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-hoefopttree-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-hoefmtree-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-lazyknn-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-majority-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-nochange-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```

```
EvaluatePrequential -l file:/home/marcelo/KDD_final/mod-perceptron-9.moa -s (ArffFileStream -f /home/marcelo/KDD_final/NSL-KDD-TEST9.arff) -e (WindowClassificationPerformanceEvaluator -p) -f 1000
```


Resultados (Métricas de Avaliação) – Dataset 2 [9 atributos]

Acurácia dos Classificadores (Média e Desvio padrão)

Summary Viewer									
Algorithm	mod-hoefmtree-9	mod-lazyknn-9	mod-majority-9	mod-decisionstump-9	mod-perceptron-9	mod-hoefadatree-9	mod-nochange-9	mod-hoefopttree-9	mod-naivebayes-9
NSL-KDD9_atributos	90,09±1,82	92,33±1,45	43,03±1,79	69,92±1,51	71,20±1,74	82,38±1,81	50,60±1,93	89,72±3,48	73,67±1,50

Tempo de Avaliação (CPU seconds)

[illegible]

Precisão para a Classe 0 (normal)

Summary Viewer									
Algorithm	mod-hoefmtree-9	mod-lazyknn-9	mod-majority-9	mod-decisionstump-9	mod-perceptron-9	mod-hoefadatree-9	mod-nochange-9	mod-hoefopttree-9	mod-naivebayes-9
NSL-KDD9_atributos	84,17±3,33	92,60±2,21	43,03±1,79	59,08±2,18	59,99±2,32	83,68±8,39	42,54±2,42	84,75±6,36	62,40±2,26

Precisão para a Classe 1 (ataque)

Summary Viewer									
Algorithm	mod-hoefree-9	mod-lazyknn-9	mod-majority-9	mod-decisionstump-9	mod-perceptron-9	mod-hoefadtree-9	mod-nochange-9	mod-hoefopttree-9	mod-naivebayes-9
NSL-KDD9_atributos	95,79±0,81	92,16±1,65	0,00±0,00	96,74±0,88	98,85±0,63	83,92±5,44	56,58±2,67	94,70±1,09	96,91±0,83

Janela da Sessão de Análise dos Resultados (Dataset 2)

MOA Graphical User Interface

ExperimentsSummaryPlotAnalyze

Configuration

Result FolderBrowse

Algorithm	Algorithm ID
mod-hoefftree-9.txt	mod-hoefftree-9
mod-lazyknn-9.txt	mod-lazyknn-9
mod-majority-9.txt	mod-majority-9
mod-decisionstump-9.txt	mod-decisionstump-9
mod-perceptron-9.txt	mod-perceptron-9
mod-hoeffadatree-9.txt	mod-hoeffadatree-9
mod-nochange-9.txt	mod-nochange-9

Delete Algorithm

Stream	Stream ID
NSL-KDD9_atributos	NSL-KDD9_atributos

Delete Stream

Statistical Test

TestNemenyi

pValue

measureclassifications correct (percent)

typeMean

Run Test

Image

Reset to Default

Output

mod-majority-9: 9.0

P-values of classifiers against each other

mod-lazyknn-9 vs mod-majority-9: 1.3992157372470226
mod-hoefftree-9 vs mod-majority-9: 2.5452412151753903
mod-lazyknn-9 vs mod-nochange-9: 2.5452412151753903
mod-hoefftree-9 vs mod-nochange-9: 4.368069012905356
mod-lazyknn-9 vs mod-decisionstump-9: 4.368069012905356
mod-majority-9 vs mod-hoeffpttree-9: 4.368069012905356

Ranqueamento dos Algoritmos (Dataset 2 [9 atributos])

P-values involving all algorithms

P-value computed by Friedman Test: 0.4334701203765876

P-value computed by Iman and Davenport Test: NaN

Ranking of the algorithms

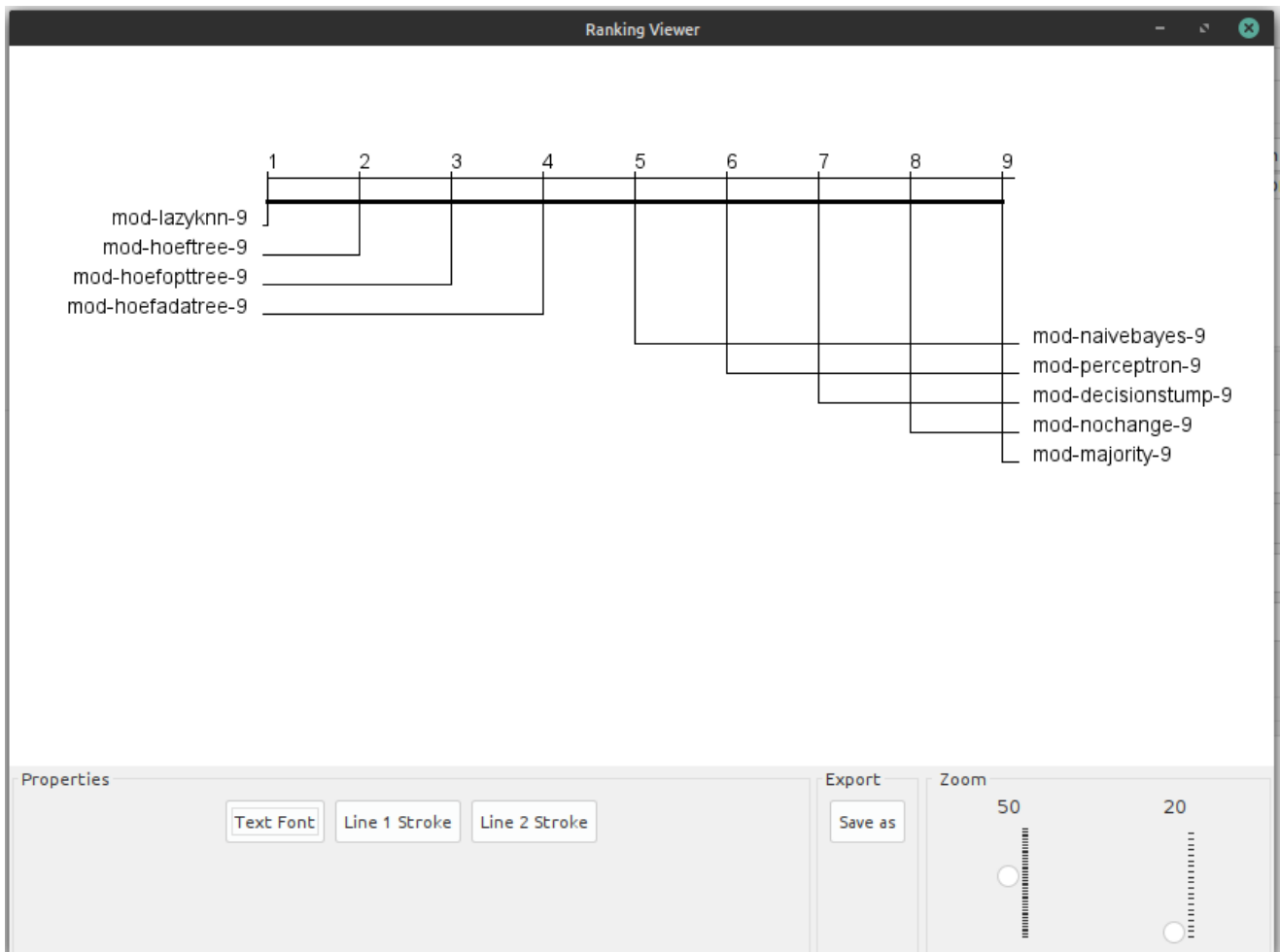
mod-lazyknn-9: 1.0
mod-hoefmtree-9: 2.0
mod-hoefopttree-9: 3.0
mod-hoefadatree-9: 4.0
mod-naivebayes-9: 5.0
mod-perceptron-9: 6.0
mod-decisionstump-9: 7.0
mod-nochange-9: 8.0
mod-majority-9: 9.0

P-values of classifiers against each other

mod-lazyknn-9 vs mod-majority-9: 1.3992157372470226
mod-hoefmtree-9 vs mod-majority-9: 2.5452412151753903
mod-lazyknn-9 vs mod-nochange-9: 2.5452412151753903
mod-hoefmtree-9 vs mod-nochange-9: 4.368069012905356
mod-lazyknn-9 vs mod-decisionstump-9: 4.368069012905356
mod-majority-9 vs mod-hoefopttree-9: 4.368069012905356
mod-hoefmtree-9 vs mod-decisionstump-9: 7.081401688522092
mod-lazyknn-9 vs mod-perceptron-9: 7.081401688522092
mod-majority-9 vs mod-hoefadatree-9: 7.081401688522092
mod-nochange-9 vs mod-hoefopttree-9: 7.081401688522092
mod-hoefmtree-9 vs mod-perceptron-9: 10.861184969220528
mod-lazyknn-9 vs mod-naivebayes-9: 10.861184969220528
mod-majority-9 vs mod-naivebayes-9: 10.861184969220528
mod-decisionstump-9 vs mod-hoefopttree-9: 10.861184969220528
mod-hoefadatree-9 vs mod-nochange-9: 10.861184969220528
mod-hoefmtree-9 vs mod-naivebayes-9: 15.788808938915995
mod-lazyknn-9 vs mod-hoefadatree-9: 15.788808938915995
mod-majority-9 vs mod-perceptron-9: 15.788808938915995
mod-decisionstump-9 vs mod-hoefadatree-9: 15.788808938915995
mod-perceptron-9 vs mod-hoefopttree-9: 15.788808938915995
mod-nochange-9 vs mod-naivebayes-9: 15.788808938915995
mod-hoefmtree-9 vs mod-hoefadatree-9: 21.80075818807247
mod-lazyknn-9 vs mod-hoefopttree-9: 21.80075818807247
mod-majority-9 vs mod-decisionstump-9: 21.80075818807247
mod-decisionstump-9 vs mod-naivebayes-9: 21.80075818807247
mod-perceptron-9 vs mod-hoefadatree-9: 21.80075818807247
mod-perceptron-9 vs mod-nochange-9: 21.80075818807247
mod-hoefopttree-9 vs mod-naivebayes-9: 21.80075818807247
mod-hoefmtree-9 vs mod-lazyknn-9: 28.66512293055501
mod-hoefmtree-9 vs mod-hoefopttree-9: 28.66512293055501
mod-majority-9 vs mod-nochange-9: 28.66512293055501
mod-decisionstump-9 vs mod-perceptron-9: 28.66512293055501
mod-decisionstump-9 vs mod-nochange-9: 28.66512293055501
mod-perceptron-9 vs mod-naivebayes-9: 28.66512293055501
mod-hoefadatree-9 vs mod-hoefopttree-9: 28.66512293055501
mod-hoefadatree-9 vs mod-naivebayes-9: 28.66512293055501

Representação Gráfica de Pós-testes (Dataset 2)

Em Demsar (2006) é apresentada uma maneira gráfica de dispor os resultados dos pós-testes.



Referências

Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine Learning Research*, 7, 1-30.

Nemenyi, P. B. (1963). *Distribution-free multiple comparisons*. Princeton University.