Rapport des resulats de l'apprentissage automatique pour la prédiction du degré de sévérité de la maladie de Tuberculose pulmonaire

Benkedadra Mohamed
Youcefi Mohemmed Yassine
Benkorreche Mohamed El Amine

Groupe 1 - ISI

A - J48 (MÉTHODE 1) :
- Y_features_opt2_300:
  - Y_features_opt2_300 - OPTIONS:

-P 57 -O -S 54 -I 15
-store-out-of-bag-predictions -output-out-of-bag-complexity-statistics
-represent-copies-using-weights -W weka.classifiers.trees.J48 -- -U -Q 90

- Y_features_opt2_300 - VAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances         22               64.7059 %
Incorrectly Classified Instances       12               35.2941 %
Kappa statistic                         0.2867
Mean absolute error                     0.4435
Root mean squared error                 0.4674
Relative absolute error                88.9415 %
Root relative squared error            93.6397 %
Total Number of Instances              34


=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
               0.722    0.438    0.650      0.722   0.684      0.289  0.710     0.743     LOW
               0.563    0.278    0.643      0.563   0.600      0.289  0.710     0.702     HIGH
Weighted Avg.  0.647    0.362    0.647      0.647   0.645      0.289  0.710     0.723


=== Confusion Matrix ===

  a  b   <-- classified as
 13  5 |  a = LOW
  7  9 |  b = HIGH
```

- Y_features_opt2_300 - TRAINVAL MESURE:

```
=== Error on test data ===
  : 0.581
Correctly Classified Instances        143               87.1951 %
Incorrectly Classified Instances       21               12.8049 %
Kappa statistic                         0.7428
Mean absolute error                     0.3278
Root mean squared error                 0.3585
Relative absolute error                65.7046 %
Root relative squared error            71.7869 %
Total Number of Instances             164


=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
               0.895    0.154    0.865      0.895   0.880      0.743  0.944     0.948     LOW
               0.846    0.105    0.880      0.846   0.863      0.743  0.944     0.944     HIGH
Weighted Avg.  0.872    0.130    0.872      0.872   0.872      0.743  0.944     0.946


=== Confusion Matrix ===

  a  b   <-- classified as
 77  9 |  a = LOW
 12 66 |  b = HIGH
```

- Z_features_opt2_300 :
- Z_features_opt2_300 - OPTIONS:
-O -S 89 -I 18

-W weka.classifiers.trees.J48 -- -C 0.1355511376705627 -Q 64 -J

- Z_features_opt2_300 - VAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances          17              50      %
Incorrectly Classified Instances        17              50      %
Kappa statistic                         -0.007
Mean absolute error                      0.4872
Root mean squared error                  0.5094
Relative absolute error                 97.6962 %
Root relative squared error            102.0379 %
Total Number of Instances               34


=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
              0.556    0.563    0.526      0.556   0.541      -0.007  0.528     0.651     LOW
              0.438    0.444    0.467      0.438   0.452      -0.007  0.528     0.507     HIGH
Weighted Avg. 0.500    0.507    0.498      0.500   0.499      -0.007  0.528     0.583


=== Confusion Matrix ===

  a  b   <-- classified as
 10  8 |  a = LOW
  9  7 |  b = HIGH
```

- Z_features_opt2_300 - TRAINVAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances         146             89.0244 %
Incorrectly Classified Instances        18             10.9756 %
Kappa statistic                          0.78
Mean absolute error                      0.2838
Root mean squared error                  0.3223
Relative absolute error                 56.8913 %
Root relative squared error             64.5401 %
Total Number of Instances              164


=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
              0.895    0.115    0.895      0.895   0.895      0.780  0.962     0.966     LOW
              0.885    0.105    0.885      0.885   0.885      0.780  0.962     0.963     HIGH
Weighted Avg. 0.890    0.110    0.890      0.890   0.890      0.780  0.962     0.965


=== Confusion Matrix ===

  a  b   <-- classified as
 77  9 |  a = LOW
  9 69 |  b = HIGH
```

- YZ_features_opt2_600:

-YZ_features_opt2_600 – OPTIONS:

-O -S 70 -W weka.classifiers.trees.J48 -- -N 104 -R -Q 105

- YZ_features_opt2_600 - VAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances         18              52.9412 %
Incorrectly Classified Instances       16              47.0588 %
Kappa statistic                         0.0423
Mean absolute error                     0.4882
Root mean squared error                 0.5032
Relative absolute error                97.898  %
Root relative squared error           100.8013 %
Total Number of Instances              34


=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
              0.667    0.625    0.545      0.667   0.600      0.044   0.580     0.675     LOW
              0.375    0.333    0.500      0.375   0.429      0.044   0.580     0.540     HIGH
Weighted Avg. 0.529    0.488    0.524      0.529   0.519      0.044   0.580     0.611


=== Confusion Matrix ===

  a  b   <-- classified as
 12  6 |  a = LOW
 10  6 |  b = HIGH
```

- YZ_features_opt2_600 - TRAINVAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances        127              77.439  %
Incorrectly Classified Instances       37              22.561  %
Kappa statistic                         0.5436
Mean absolute error                     0.4077
Root mean squared error                 0.4236
Relative absolute error                81.7157 %
Root relative squared error            84.8109 %
Total Number of Instances             164


=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
              0.872    0.333    0.743      0.872   0.802      0.553   0.867     0.864     LOW
              0.667    0.128    0.825      0.667   0.738      0.553   0.867     0.850     HIGH
Weighted Avg. 0.774    0.236    0.782      0.774   0.771      0.553   0.867     0.857


=== Confusion Matrix ===

  a  b   <-- classified as
 75 11 |  a = LOW
 26 52 |  b = HIGH
```

A - HoeffdingTree (MÉTHODE 2) :
- Y_features_opt2_300:
- Y_features_opt2_300 - OPTIONS:
-P 29 -O -S 55 -num-slots 17 -output-out-of-bag-complexity-statistics
-W weka.classifiers.trees.HoeffdingTree --
-H 0.41623 -L 1 -E 0.41813 -M 0.21027 -P

- Y_features_opt2_300 - VAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances          16              47.0588 %
Incorrectly Classified Instances        18              52.9412 %
Kappa statistic                         -0.0851
Mean absolute error                      0.4927
Root mean squared error                  0.5409
Relative absolute error                 98.8107 %
Root relative squared error            108.3633 %
Total Number of Instances               34
```

```
=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.667	0.750	0.500	0.667	0.571	-0.091	0.545	0.669	LOW
	0.250	0.333	0.400	0.250	0.308	-0.091	0.545	0.487	HIGH
Weighted Avg.	0.471	0.554	0.453	0.471	0.447	-0.091	0.545	0.583	

```
=== Confusion Matrix ===

  a  b   <-- classified as
 12  6 |  a = LOW
 12  4 |  b = HIGH
```

- Y_features_opt2_300 - TRAINVAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances          97              59.1463 %
Incorrectly Classified Instances        67              40.8537 %
Kappa statistic                          0.1666
Mean absolute error                      0.4366
Root mean squared error                  0.4987
Relative absolute error                 87.5073 %
Root relative squared error             99.8629 %
Total Number of Instances              164
```

```
=== Detailed Accuracy By Class ===
```

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.779	0.615	0.583	0.779	0.667	0.179	0.646	0.695	LOW
	0.385	0.221	0.612	0.385	0.472	0.179	0.646	0.633	HIGH
Weighted Avg.	0.591	0.428	0.597	0.591	0.574	0.179	0.646	0.666	

```
=== Confusion Matrix ===

  a  b   <-- classified as
 67 19 |  a = LOW
 48 30 |  b = HIGH
```

- Z_features_opt2_300 :
- Z_features_opt2_300 - OPTIONS:
-P 53 -S 77 -W weka.classifiers.trees.HoeffdingTree -- -H 0.65928 -L 1 -S 1 -G 0 -N 57

- Z_features_opt2_300 - VAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances          14              41.1765 %
Incorrectly Classified Instances        20              58.8235 %
Kappa statistic                         -0.2143
Mean absolute error                      0.5055
Root mean squared error                  0.5088
Relative absolute error                101.3803 %
Root relative squared error            101.9297 %
Total Number of Instances               34


=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
               0.667    0.875    0.462      0.667   0.545      -0.245  0.451     0.574     LOW
               0.125    0.333    0.250      0.125   0.167      -0.245  0.451     0.431     HIGH
Weighted Avg.  0.412    0.620    0.362      0.412   0.367      -0.245  0.451     0.507

=== Confusion Matrix ===

  a  b   <-- classified as
 12  6 |   a = LOW
 14  2 |   b = HIGH
```

- Z_features_opt2_300 - TRAINVAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances          97              59.1463 %
Incorrectly Classified Instances        67              40.8537 %
Kappa statistic                          0.1686
Mean absolute error                      0.4841
Root mean squared error                  0.4871
Relative absolute error                 97.0391 %
Root relative squared error             97.5359 %
Total Number of Instances              164


=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
               0.756    0.590    0.586      0.756   0.660      0.177  0.661     0.716     LOW
               0.410    0.244    0.604      0.410   0.489      0.177  0.661     0.610     HIGH
Weighted Avg.  0.591    0.425    0.594      0.591   0.578      0.177  0.661     0.666

=== Confusion Matrix ===

  a  b   <-- classified as
 65 21 |   a = LOW
 46 32 |   b = HIGH
```

- YZ_features_opt2_600:
-YZ_features_opt2_600 – OPTIONS:
-I 8 -output-out-of-bag-complexity-statistics -represent-copies-using-weights
-W weka.classifiers.trees.HoeffdingTree -- -H 0.77991 -S 0 -M 0.85064 -G 50
-N 41 -P

- YZ_features_opt2_600 - VAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances          18                52.9412 %
Incorrectly Classified Instances        16                47.0588 %
Kappa statistic                          0.0556
Mean absolute error                      0.4909
Root mean squared error                  0.5015
Relative absolute error                 98.4358 %
Root relative squared error            100.4713 %
Total Number of Instances               34


=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
              0.556    0.500    0.556      0.556   0.556      0.056  0.502     0.536     LOW
              0.500    0.444    0.500      0.500   0.500      0.056  0.502     0.510     HIGH
Weighted Avg. 0.529    0.474    0.529      0.529   0.529      0.056  0.502     0.524


=== Confusion Matrix ===

  a  b   <-- classified as
 10  8 |  a = LOW
  8  8 |  b = HIGH
```

- YZ_features_opt2_600 - TRAINVAL MESURE:

```
=== Error on test data ===

Correctly Classified Instances          90                54.878  %
Incorrectly Classified Instances        74                45.122  %
Kappa statistic                          0.0932
Mean absolute error                      0.4826
Root mean squared error                  0.4921
Relative absolute error                 96.7442 %
Root relative squared error             98.5391 %
Total Number of Instances              164


=== Detailed Accuracy By Class ===

              TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
              0.593    0.500    0.567      0.593   0.580      0.093  0.568     0.569     LOW
              0.500    0.407    0.527      0.500   0.513      0.093  0.568     0.566     HIGH
Weighted Avg. 0.549    0.456    0.548      0.549   0.548      0.093  0.568     0.568


=== Confusion Matrix ===

  a  b   <-- classified as
 51 35 |  a = LOW
 39 39 |  b = HIGH
```