**Assignment 7: Logic-based approaches**

**Group members:**

**Yumeng Lu**

**Zhaoxuan Ren**

**Tao Li**

## Iris

We use trees to classify this dataset, and the validation approach is cross validation. In addition, we use different folders to test the results, and show the accuracy.

(1) BF trees

Test option: Cross Validation (10 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 142 | 94.6667 % |
| Incorrectly Classified Instances | 8 | 5.3333 % |
| Kappa statistic | 0.92 | |
| Mean absolute error | 0.041 | |
| Root mean squared error | 0.1754 | |

| | |
|---|---|
| Relative absolute error | 9.2315 % |
| Root relative squared error | 37.2061 % |
| Total Number of Instances | 150 |

(2) Decision Stump

Test option: Cross Validation (10 folders)

=== Evaluation on test split ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 32 | 62.7451 % |
| Incorrectly Classified Instances | 19 | 37.2549 % |
| Kappa statistic | 0.4507 | |
| Mean absolute error | 0.2357 | |
| Root mean squared error | 0.3438 | |
| Relative absolute error | 52.911 % | |
| Root relative squared error | 72.713 % | |
| Total Number of Instances | 51 | |

(3) FT

Test Option: Cross Validation (5 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 144 | 96 % |
| Incorrectly Classified Instances | 6 | 4 % |
| Kappa statistic | 0.94 | |
| Mean absolute error | 0.0332 | |
| Root mean squared error | 0.1364 | |
| Relative absolute error | 7.4671 % | |
| Root relative squared error | 28.9385 % | |
| Total Number of Instances | 150 | |

(4) J48

Test Option: Cross Validation (4 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 143 | 95.3333 % |
| Incorrectly Classified Instances | 7 | 4.6667 % |
| Kappa statistic | 0.93 | |

| | |
|---|---|
| Mean absolute error | 0.041 |
| Root mean squared error | 0.1761 |
| Relative absolute error | 9.2165 % |
| Root relative squared error | 37.3534 % |
| Total Number of Instances | 150 |

(5) J48graft

Test Option: Cross Validation (4 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 142 | 94.6667 % |
| Incorrectly Classified Instances | 8 | 5.3333 % |
| Kappa statistic | 0.92 | |
| Mean absolute error | 0.0447 | |
| Root mean squared error | 0.188 | |
| Relative absolute error | 10.0497 % | |
| Root relative squared error | 39.87 % | |
| Total Number of Instances | 150 | |

(6) LADTress

Test Option: Cross Validation (8 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | | |
|---|---|---|---|
| Correctly Classified Instances | 141 | 94 | % |
| Incorrectly Classified Instances | 9 | 6 | % |
| Kappa statistic | 0.91 | | |
| Mean absolute error | 0.0424 | | |
| Root mean squared error | 0.1866 | | |
| Relative absolute error | 9.5268 % | | |
| Root relative squared error | 39.5677 % | | |
| Total Number of Instances | 150 | | |

(7) LMT

Test Option: Cross Validation (5 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 143 | 95.3333 % |
| Incorrectly Classified Instances | 7 | 4.6667 % |
| Kappa statistic | 0.93 | |
| Mean absolute error | 0.0426 | |
| Root mean squared error | 0.1402 | |
| Relative absolute error | 9.5738 % | |
| Root relative squared error | 29.7515 % | |
| Total Number of Instances | 150 | |

(8) NBTree

Test Option: Cross Validation (5 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 137 | 91.3333 % |
| Incorrectly Classified Instances | 13 | 8.6667 % |
| Kappa statistic | 0.87 | |
| Mean absolute error | 0.0639 | |
| Root mean squared error | 0.2036 | |

| | |
|---|---|
| Relative absolute error | 14.3665 % |
| Root relative squared error | 43.1822 % |
| Total Number of Instances | 150 |

(9) RandomForest

Test Option: Cross Validation (5 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 144 | 96 % |
| Incorrectly Classified Instances | 6 | 4 % |
| Kappa statistic | 0.94 | |
| Mean absolute error | 0.0383 | |
| Root mean squared error | 0.1516 | |
| Relative absolute error | 8.61 % | |
| Root relative squared error | 32.1537 % | |
| Total Number of Instances | 150 | |

(10) RandomTree

Test Option: Cross Validation (8 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 140 | 93.3333 % |
| Incorrectly Classified Instances | 10 | 6.6667 % |
| Kappa statistic | 0.9 | |
| Mean absolute error | 0.0444 | |
| Root mean squared error | 0.2108 | |
| Relative absolute error | 9.997 % | |
| Root relative squared error | 44.7071 % | |
| Total Number of Instances | 150 | |

(11) REPTree

Test Option: Cross Validation (8 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 142 | 94.6667 % |
| Incorrectly Classified Instances | 8 | 5.3333 % |

| | |
|---|---|
| Kappa statistic | 0.92 |
| Mean absolute error | 0.05 |
| Root mean squared error | 0.1817 |
| Relative absolute error | 11.2563 % |
| Root relative squared error | 38.526 % |
| Total Number of Instances | 150 |

(12) SimpleCart

Test Option: Cross Validation (4 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 142 | 94.6667 % |
| Incorrectly Classified Instances | 8 | 5.3333 % |
| Kappa statistic | 0.92 | |
| Mean absolute error | 0.0498 | |
| Root mean squared error | 0.1821 | |
| Relative absolute error | 11.2113 % | |
| Root relative squared error | 38.6187 % | |
| Total Number of Instances | 150 | |

The best  accuracy from all the tree models weka provide is 96%.

In our assignment, we can get the same value by using FT and RandomForest.

Weka also provides some rules. In our tests, some rules can get highest 96% accuracy

(1) Rule: JRip

Test Option: Cross Validation (9 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 144 | 96 % |
| Incorrectly Classified Instances | 6 | 4 % |
| Kappa statistic | 0.94 | |
| Mean absolute error | 0.041 | |
| Root mean squared error | 0.1648 | |
| Relative absolute error | 9.2182 % | |
| Root relative squared error | 34.9407 % | |
| Total Number of Instances | 150 | |

(2) Rule: NNge

Test Option: Cross Validation (10 folders)

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 144 | 96 % |
| Incorrectly Classified Instances | 6 | 4 % |
| Kappa statistic | 0.94 | |
| Mean absolute error | 0.0267 | |
| Root mean squared error | 0.1633 | |
| Relative absolute error | 6 % | |
| Root relative squared error | 34.641 % | |
| Total Number of Instances | 150 | |

## House and vote

After trying to improve the accuracy by using different using different learning approaches, different values of parameters, best classification accuracy we achieved is 85.7471 % by using LMT (10-fold cross validation).

Time taken to build model: 0.53 seconds

=== Stratified cross-validation ===

=== Summary ===

| | | |
|---|---|---|
| Correctly Classified Instances | 373 | 85.7471 % |
| Incorrectly Classified Instances | 62 | 14.2529 % |
| Kappa statistic | 0.7302 | |
| Mean absolute error | 0.1618 | |
| Root mean squared error | 0.2923 | |
| Relative absolute error | 44.2905 % | |
| Root relative squared error | 68.4773 % | |
| Total Number of Instances | 435 | |

=== Detailed Accuracy by Class ===

| TP Rate | FP Rate | Precision | Recall | F-Measure | ROC Area | Class |
|---|---|---|---|---|---|---|
| 0.903 | 0.122 | 0.869 | 0.903 | 0.886 | 0.905 | n |
| 0.894 | 0.145 | 0.849 | 0.894 | 0.871 | | |

| | | | | | | |
|---|---|---|---|---|---|---|
| 0.888 | y | | | | | |
| | | 0.091 | 0.002 | 0.667 | 0.091 | 0.16 |
| 0.706 | w | | | | | |
| Weighted | Avg. | 0.857 | 0.127 | 0.849 | 0.857 | 0.842 |
| 0.887 | | | | | | |