Na przykładzie pliku *heart_2020_cleaned.csv*

OverFitting – wyłączono przycinanie drzewa
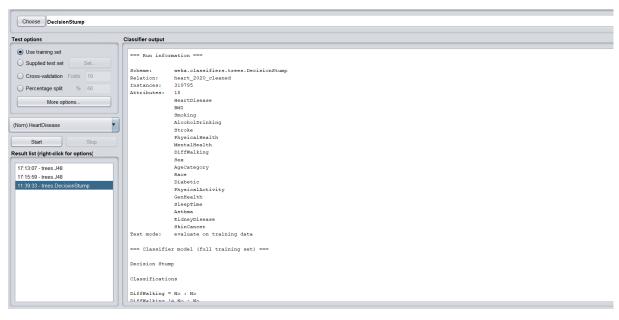
```
|   |   |   |   |   |   |   |   |   |   |   |   MentalHealth <= 4: No  (3.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   MentalHealth > 4: Yes  (3.0)
|   |   |   |   |   |   |   |   |   |   Asthma = No
|   |   |   |   |   |   |   |   |   |   |   MentalHealth <= 12
|   |   |   |   |   |   |   |   |   |   |   |   MentalHealth <= 1: No  (17.0/7.0)
|   |   |   |   |   |   |   |   |   |   |   |   MentalHealth > 1: Yes  (3.0)
|   |   |   |   |   |   |   |   |   |   |   MentalHealth > 12: No  (4.0)
|   |   |   |   |   |   |   |   AgeCategory = 65-69
|   |   |   |   |   |   |   |   |   AlcoholDrinking = No
|   |   |   |   |   |   |   |   |   |   Asthma = Yes: Yes  (17.0/5.0)
|   |   |   |   |   |   |   |   |   |   Asthma = No
|   |   |   |   |   |   |   |   |   |   |   PhysicalActivity = Yes: No  (8.0/2.0)
|   |   |   |   |   |   |   |   |   |   |   PhysicalActivity = No
|   |   |   |   |   |   |   |   |   |   |   |   PhysicalHealth <= 18: Yes  (8.0)
|   |   |   |   |   |   |   |   |   |   |   |   PhysicalHealth > 18
|   |   |   |   |   |   |   |   |   |   |   |   |   MentalHealth <= 3
|   |   |   |   |   |   |   |   |   |   |   |   |   |   PhysicalHealth <= 29: Yes  (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   PhysicalHealth > 29
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   SleepTime <= 9
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   SleepTime <= 5: Yes  (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   SleepTime > 5: No  (5.0/1.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   SleepTime > 9: Yes  (3.0)
|   |   |   |   |   |   |   |   |   |   |   |   |   MentalHealth > 3: No  (11.0/3.0)
|   |   |   |   |   |   |   |   |   AlcoholDrinking = Yes: Yes  (2.0)
|   |   |   |   |   |   |   |   AgeCategory = 75-79
|   |   |   |   |   |   |   |   |   MentalHealth <= 27
|   |   |   |   |   |   |   |   |   |   Asthma = Yes
|   |   |   |   |   |   |   |   |   |   |   BMI <= 25.01
|   |   |   |   |   |   |   |   |   |   |   |   BMI <= 22.1: No  (2.0)
|   |   |   |   |   |   |   |   |   |   |   |   BMI > 22.1: Yes  (3.0)
|   |   |   |   |   |   |   |   |   |   |   BMI > 25.01: No  (5.0)
|   |   |   |   |   |   |   |   |   |   Asthma = No
|   |   |   |   |   |   |   |   |   |   |   SleepTime <= 8
|   |   |   |   |   |   |   |   |   |   |   |   BMI <= 32.32: Yes  (11.0/1.0)
```

Classifier

Choose   J48 -S -C 0.25 -M 2

Test options

- (•) Use training set
- ( ) Supplied test set   Set...
- ( ) Cross-validation Folds 10
- ( ) Percentage split   % 66

More options...

(Nom) HeartDisease ▾

Start   Stop

Result list (right-click for options)

17:13:07 - trees.J48
17:15:59 - trees.J48

Classifier output

```
|   |   |   |   |   |   |   |   Asthma = Yes: No  (3.0)
|   |   |   |   |   |   |   |   Asthma = No: Yes  (3.0/1.0)
|   |   |   |   |   |   BMI > 26.6: Yes  (32.0/2.0)
|   |   |   |   AgeCategory = 50-54: Yes  (11.0/1.0)
|   |   |   |   AgeCategory = 45-49
|   |   |   |   |   Diabetic = Yes: No  (2.0)
|   |   |   |   |   Diabetic = No: Yes  (2.0)
|   |   |   |   |   Diabetic = No, borderline diabetes: No  (0.0)
|   |   |   |   |   Diabetic = Yes (during pregnancy): No  (0.0)
|   |   |   |   AgeCategory = 18-24: Yes  (1.0)
|   |   |   |   AgeCategory = 35-39: No  (2.0)
|   |   |   |   AgeCategory = 30-34: Yes  (0.0)
|   |   |   |   AgeCategory = 25-29: Yes  (0.0)
|   |   AlcoholDrinking = Yes: Yes  (9.0)
|   GenHealth = Excellent
|   |   Diabetic = Yes: Yes  (14.0)
|   |   Diabetic = No
|   |   |   Race = White: No  (15.0/6.0)
|   |   |   Race = Black: No  (0.0)
|   |   |   Race = Asian: No  (0.0)
|   |   |   Race = American Indian/Alaskan Native: No  (0.0)
|   |   |   Race = Other: Yes  (2.0)
|   |   |   Race = Hispanic: No  (0.0)
|   |   Diabetic = No, borderline diabetes: No  (3.0/1.0)
|   |   Diabetic = Yes (during pregnancy): Yes  (1.0)

Number of Leaves  :    1792

Size of the tree :    2819


Time taken to build model: 61.96 seconds
```

Under Fitting

Choose | DecisionStump

Test options
- ○ Use training set
- ○ Supplied test set   Set...
- ○ Cross-validation   Folds   10
- ○ Percentage split   %   66

More options...

(Nom) HeartDisease

Start | Stop

Result list (right-click for options)
- 17:13:07 - trees.J48
- 17:15:59 - trees.J48
- 11:39:33 - trees.DecisionStump

Classifier output

```
=== Run information ===

Scheme:       weka.classifiers.trees.DecisionStump
Relation:     heart_2020_cleaned
Instances:    319795
Attributes:   18
              HeartDisease
              BMI
              Smoking
              AlcoholDrinking
              Stroke
              PhysicalHealth
              MentalHealth
              DiffWalking
              Sex
              AgeCategory
              Race
              Diabetic
              PhysicalActivity
              GenHealth
              SleepTime
              Asthma
              KidneyDisease
              SkinCancer
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

Decision Stump

Classifications

DiffWalking = No : No
DiffWalking != No : No
```


Classifier output

```
              KidneyDisease
              SkinCancer
Test mode:    evaluate on training data

=== Classifier model (full training set) ===

Decision Stump

Classifications

DiffWalking = No : No
DiffWalking != No : No
DiffWalking is missing : No

Class distributions

DiffWalking = No
No      Yes
0.9370154510957387      0.0629845489042613
DiffWalking != No
No      Yes
0.7741950011258726      0.22580499887412744
DiffWalking is missing
No      Yes
0.9144045404086993      0.08559545959130067


Time taken to build model: 1.52 seconds

=== Evaluation on training set ===

Time taken to test model on training data: 0.35 seconds


=== Summary ===
```

```
=== Summary ===

Correctly Classified Instances        292422               91.4405 %
Incorrectly Classified Instances      27373                 8.5595 %
Kappa statistic                           0
Mean absolute error                       0.1502
Root mean squared error                   0.274
Relative absolute error                  95.9482 %
Root relative squared error              97.9538 %
Total Number of Instances            319795

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                 1,000    1,000    0,914      1,000   0,955      ?        0,624     0,934     No
                 0,000    0,000    ?          0,000   ?          ?        0,624     0,137     Yes
Weighted Avg.    0,914    0,914    ?          0,914   ?          ?        0,624     0,866

=== Confusion Matrix ===

      a       b   <-- classified as
 292422       0 |     a = No
  27373       0 |     b = Yes
```