

Numeric Prediction

For this homework, use the [world happiness dataset](#).

We want to predict the SWL-index attribute (prediction target) from the other predicting attributes: continent, life-expectancy, GDP-per-capita, access-to-education-score. Note that the attribute country identifies each data instance uniquely and as such will be disregarded in our analysis. It is provided just for context.

[5 points] Build a regression tree for this dataset in Weka using M5P with the following parameters:

build regression tree: True unpruned: True useUnsmoothed: True default values
for the remaining parameters

1. Record the tree in your report.

=== Run information ===

Scheme: weka.classifiers.trees.M5P -N -U -R -M 4.0 -num-decimal-places 4

Relation: world_happiness-weka.filters.unsupervised.attribute.Remove-R2

Instances: 178

Attributes: 6

continent

SWL-ranking

SWL-index

life-expectancy

GDP-per-capita

access-to-education-score

Test mode: 2-fold cross-validation

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

M5 unpruned regression tree:

SWL-ranking <= 95.5 :

| SWL-ranking <= 55.5 :

| | SWL-ranking <= 41.5 :

| | | SWL-ranking <= 15.5 :

| | | | SWL-ranking <= 4.5 :

| | | | | SWL-ranking <= 2.5 : LM1 (2/0%)

| | | | | SWL-ranking > 2.5 : LM2 (2/0%)

| | | | SWL-ranking > 4.5 :

| | | | | SWL-ranking <= 7.5 : LM3 (3/0%)

| | | | | SWL-ranking > 7.5 : LM4 (8/4.224%)

| | | SWL-ranking > 15.5 :

| | | | SWL-ranking <= 25.5 : LM5 (10/0%)

| | | | SWL-ranking > 25.5 :

| | | | | SWL-ranking <= 32.5 : LM6 (7/0%)

| | | | | SWL-ranking > 32.5 : LM7 (9/2.742%)

| | SWL-ranking > 41.5 : LM8 (14/4.318%)

| SWL-ranking > 55.5 :

| | SWL-ranking <= 76.5 :

| | | SWL-ranking <= 70.5 :

| | | | SWL-ranking <= 60.5 : LM9 (5/3.501%)

| | | | SWL-ranking > 60.5 : LM10 (10/0%)

| | | SWL-ranking > 70.5 : LM11 (6/0%)

| | SWL-ranking > 76.5 :

| | | SWL-ranking <= 91.5 :

| | | | SWL-ranking <= 80.5 : LM12 (4/0%)

| | | | SWL-ranking > 80.5 : LM13 (11/4.198%)

| | | SWL-ranking > 91.5 : LM14 (4/0%)

SWL-ranking > 95.5 :

| SWL-ranking <= 142.5 :

| | SWL-ranking <= 121.5 :

| | | SWL-ranking <= 109.5 :

| | | | SWL-ranking <= 103.5 :

| | | | | SWL-ranking <= 99.5 : LM15 (4/3.778%)

| | | | | SWL-ranking > 99.5 : LM16 (4/0%)

| | | | SWL-ranking > 103.5 : LM17 (6/0%)

| | | SWL-ranking > 109.5 : LM18 (12/4.315%)

| | SWL-ranking > 121.5 :

| | | SWL-ranking <= 129.5 : LM19 (8/0%)

| | | SWL-ranking > 129.5 :

| | | | SWL-ranking <= 136.5 : LM20 (7/4.331%)

| | | | SWL-ranking > 136.5 : LM21 (6/0%)

| SWL-ranking > 142.5 :

| | SWL-ranking <= 168.5 :

| | | SWL-ranking <= 158.5 :

| | | | SWL-ranking <= 151.5 :
 | | | | | SWL-ranking <= 148.5 : LM22 (6/3.262%)
 | | | | | SWL-ranking > 148.5 : LM23 (3/0%)
 | | | | SWL-ranking > 151.5 : LM24 (7/3.954%)
 | | | SWL-ranking > 158.5 :
 | | | | SWL-ranking <= 161.5 : LM25 (3/0%)
 | | | | SWL-ranking > 161.5 :
 | | | | | SWL-ranking <= 163.5 : LM26 (2/0%)
 | | | | | SWL-ranking > 163.5 : LM27 (5/3.49%)
 | | SWL-ranking > 168.5 :
 | | | SWL-ranking <= 171.5 : LM28 (3/4.126%)
 | | | SWL-ranking > 171.5 :
 | | | | SWL-ranking <= 175.5 :
 | | | | | life-expectancy <= 66.9 : LM29 (2/0%)
 | | | | | life-expectancy > 66.9 : LM30 (2/8.726%)
 | | | | SWL-ranking > 175.5 : LM31 (3/12.353%)

LM num: 1

SWL-index =

+ 273.33

LM num: 2

SWL-index =

+ 260

LM num: 3

SWL-index =

+ 256.67

LM num: 4

SWL-index =

+ 252.0812

LM num: 5

SWL-index =

+ 246.67

LM num: 6

SWL-index =

+ 243.33

LM num: 7

SWL-index =

+ 239.63

LM num: 8

SWL-index =

+ 231.4271

LM num: 9

SWL-index =

+ 223.998

LM num: 10

SWL-index =

+ 220

LM num: 11

SWL-index =

+ 216.67

LM num: 12

SWL-index =

+ 213.33

LM num: 13

SWL-index =

+ 208.7891

LM num: 14

SWL-index =

+ 203.33

LM num: 15

SWL-index =

+ 197.5025

LM num: 16

SWL-index =

+ 193.33

LM num: 17

SWL-index =

+ 190

LM num: 18

SWL-index =

+ 185.2783

LM num: 19

SWL-index =

+ 180

LM num: 20

SWL-index =

+ 175.2386

LM num: 21

SWL-index =

+ 170

LM num: 22

SWL-index =

+ 163.8867

LM num: 23

SWL-index =

+ 160

LM num: 24

SWL-index =

+ 155.7157

LM num: 25

SWL-index =

+ 150

LM num: 26

SWL-index =

+ 146.67

LM num: 27

SWL-index =

+ 142.664

LM num: 28

SWL-index =

+ 134.4433

LM num: 29

SWL-index =

+ 120

LM num: 30

SWL-index =

+ 120

LM num: 31

SWL-index =

+ 106.6667

Number of Rules : 31

Time taken to build model: 0.01 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient	0.9948
Kendall's tau	0.9722
Spearman's rho	0.9968
Mean absolute error	2.4737
Root mean squared error	3.8921
Relative absolute error	7.6425 %
Root relative squared error	10.1985 %
Total Number of Instances	178

[5 points] Build a regression tree for this dataset in Weka using M5P with the following parameters:

build regression tree: True unpruned: False useUnsmoothed: True default values
for the remaining parameters

2.

Record the tree in your report [5 points].

=== Run information ===

Scheme: weka.classifiers.trees.M5P -U -R -M 4.0 -num-decimal-places 4

Relation: world_happiness-weka.filters.unsupervised.attribute.Remove-R2

Instances: 178

Attributes: 6

continent

SWL-ranking

SWL-index

life-expectancy

GDP-per-capita

access-to-education-score

Test mode: 2-fold cross-validation

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

M5 pruned regression tree:

SWL-ranking <= 95.5 :

| SWL-ranking <= 55.5 :

| | SWL-ranking <= 41.5 :

| | | SWL-ranking <= 15.5 :

| | | | SWL-ranking <= 4.5 :

| | | | | SWL-ranking <= 2.5 : LM1 (2/0%)

| | | | | SWL-ranking > 2.5 : LM2 (2/0%)

| | | | SWL-ranking > 4.5 :

| | | | | SWL-ranking <= 7.5 : LM3 (3/0%)

| | | | | SWL-ranking > 7.5 : LM4 (8/4.224%)

| | | SWL-ranking > 15.5 :

| | | | SWL-ranking <= 25.5 : LM5 (10/0%)

| | | | SWL-ranking > 25.5 :

| | | | | SWL-ranking <= 32.5 : LM6 (7/0%)

| | | | | SWL-ranking > 32.5 : LM7 (9/2.742%)

| | SWL-ranking > 41.5 : LM8 (14/4.318%)

| SWL-ranking > 55.5 :

| | SWL-ranking <= 76.5 :

| | | SWL-ranking <= 70.5 :

| | | | SWL-ranking <= 60.5 : LM9 (5/3.501%)

| | | | SWL-ranking > 60.5 : LM10 (10/0%)

| | | SWL-ranking > 70.5 : LM11 (6/0%)

| | SWL-ranking > 76.5 :

| | | SWL-ranking <= 91.5 :

| | | | SWL-ranking <= 80.5 : LM12 (4/0%)

| | | | SWL-ranking > 80.5 : LM13 (11/4.198%)

| | | SWL-ranking > 91.5 : LM14 (4/0%)

SWL-ranking > 95.5 :

| SWL-ranking <= 142.5 :

| | SWL-ranking <= 121.5 :

| | | SWL-ranking <= 109.5 :

| | | | SWL-ranking <= 103.5 :

| | | | | SWL-ranking <= 99.5 : LM15 (4/3.778%)

| | | | | SWL-ranking > 99.5 : LM16 (4/0%)

| | | | SWL-ranking > 103.5 : LM17 (6/0%)

| | | SWL-ranking > 109.5 : LM18 (12/4.315%)

- | | SWL-ranking > 121.5 :
- | | | SWL-ranking <= 129.5 : LM19 (8/0%)
- | | | SWL-ranking > 129.5 :
- | | | | SWL-ranking <= 136.5 : LM20 (7/4.331%)
- | | | | SWL-ranking > 136.5 : LM21 (6/0%)
- | SWL-ranking > 142.5 :
- | | SWL-ranking <= 168.5 :
- | | | SWL-ranking <= 158.5 :
- | | | | SWL-ranking <= 151.5 :
- | | | | | SWL-ranking <= 148.5 : LM22 (6/3.262%)
- | | | | | SWL-ranking > 148.5 : LM23 (3/0%)
- | | | | SWL-ranking > 151.5 : LM24 (7/3.954%)
- | | | SWL-ranking > 158.5 :
- | | | | SWL-ranking <= 161.5 : LM25 (3/0%)
- | | | | SWL-ranking > 161.5 : LM26 (7/5.585%)
- | | SWL-ranking > 168.5 :
- | | | SWL-ranking <= 171.5 : LM27 (3/4.126%)
- | | | SWL-ranking > 171.5 : LM28 (7/19.65%)

LM num: 1

SWL-index =

+ 273.33

LM num: 2

SWL-index =

+ 260

LM num: 3

SWL-index =

+ 256.67

LM num: 4

SWL-index =

+ 252.0812

LM num: 5

SWL-index =

+ 246.67

LM num: 6

SWL-index =

+ 243.33

LM num: 7

SWL-index =

+ 239.63

LM num: 8

SWL-index =

+ 231.4271

LM num: 9

SWL-index =

+ 223.998

LM num: 10

SWL-index =

+ 220

LM num: 11

SWL-index =

+ 216.67

LM num: 12

SWL-index =

+ 213.33

LM num: 13

SWL-index =

+ 208.7891

LM num: 14

SWL-index =

+ 203.33

LM num: 15

SWL-index =

+ 197.5025

LM num: 16

SWL-index =

+ 193.33

LM num: 17

SWL-index =

+ 190

LM num: 18

SWL-index =

+ 185.2783

LM num: 19

SWL-index =

+ 180

LM num: 20

SWL-index =

+ 175.2386

LM num: 21

SWL-index =

+ 170

LM num: 22

SWL-index =

+ 163.8867

LM num: 23

SWL-index =

+ 160

LM num: 24

SWL-index =

+ 155.7157

LM num: 25

SWL-index =

+ 150

LM num: 26

SWL-index =

+ 143.8086

LM num: 27

SWL-index =

+ 134.4433

LM num: 28

SWL-index =

+ 114.2857

Number of Rules : 28

Time taken to build model: 0.01 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient	0.979
Kendall's tau	0.936
Spearman's rho	0.9887
Mean absolute error	4.8182
Root mean squared error	7.8153
Relative absolute error	14.8857 %

Root relative squared error 20.4787 %

Total Number of Instances 178

[5 points] Build a model tree for this dataset in Weka using M5P with the following parameters:

build regression tree: False unpruned: True useUnsmoothed: True default values
for the remaining parameters

3.

Record the tree in your report.

=== Run information ===

Scheme: weka.classifiers.trees.M5P -N -U -M 4.0 -num-decimal-places 4

Relation: world_happiness-weka.filters.unsupervised.attribute.Remove-R2

Instances: 178

Attributes: 6

continent

SWL-ranking

SWL-index

life-expectancy

GDP-per-capita

access-to-education-score

Test mode: 2-fold cross-validation

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

M5 unpruned model tree:

SWL-ranking ≤ 95.5 :

| SWL-ranking ≤ 55.5 :

| | SWL-ranking ≤ 41.5 :

| | | SWL-ranking ≤ 15.5 :

| | | | SWL-ranking ≤ 4.5 :

| | | | | SWL-ranking ≤ 2.5 : LM1 (2/0%)

| | | | | SWL-ranking > 2.5 : LM2 (2/0%)

| | | | SWL-ranking > 4.5 :

| | | | | SWL-ranking ≤ 7.5 : LM3 (3/0%)

| | | | | SWL-ranking > 7.5 : LM4 (8/4.224%)

| | | SWL-ranking > 15.5 :

| | | | SWL-ranking ≤ 25.5 : LM5 (10/0%)

| | | | SWL-ranking > 25.5 :

| | | | | SWL-ranking ≤ 32.5 : LM6 (7/0%)

| | | | | SWL-ranking > 32.5 : LM7 (9/2.742%)

| | SWL-ranking > 41.5 : LM8 (14/4.318%)

| SWL-ranking > 55.5 :

| | SWL-ranking ≤ 76.5 :

| | | SWL-ranking ≤ 70.5 :

| | | | SWL-ranking ≤ 60.5 : LM9 (5/3.501%)

| | | | SWL-ranking > 60.5 : LM10 (10/0%)

| | | SWL-ranking > 70.5 : LM11 (6/0%)

| | SWL-ranking > 76.5 :

- | | | SWL-ranking <= 91.5 :
- | | | | SWL-ranking <= 80.5 : LM12 (4/0%)
- | | | | SWL-ranking > 80.5 : LM13 (11/4.198%)
- | | | SWL-ranking > 91.5 : LM14 (4/0%)
- SWL-ranking > 95.5 :
- | SWL-ranking <= 142.5 :
- | | SWL-ranking <= 121.5 :
- | | | SWL-ranking <= 109.5 :
- | | | | SWL-ranking <= 103.5 :
- | | | | | SWL-ranking <= 99.5 : LM15 (4/3.778%)
- | | | | | SWL-ranking > 99.5 : LM16 (4/0%)
- | | | | SWL-ranking > 103.5 : LM17 (6/0%)
- | | | SWL-ranking > 109.5 : LM18 (12/4.315%)
- | | SWL-ranking > 121.5 :
- | | | SWL-ranking <= 129.5 : LM19 (8/0%)
- | | | SWL-ranking > 129.5 :
- | | | | SWL-ranking <= 136.5 : LM20 (7/4.331%)
- | | | | SWL-ranking > 136.5 : LM21 (6/0%)
- | SWL-ranking > 142.5 :
- | | SWL-ranking <= 168.5 :
- | | | SWL-ranking <= 158.5 :
- | | | | SWL-ranking <= 151.5 :
- | | | | | SWL-ranking <= 148.5 : LM22 (6/3.262%)
- | | | | | SWL-ranking > 148.5 : LM23 (3/0%)

| | | | SWL-ranking > 151.5 : LM24 (7/3.954%)
 | | | SWL-ranking > 158.5 :
 | | | | SWL-ranking <= 161.5 : LM25 (3/0%)
 | | | | SWL-ranking > 161.5 :
 | | | | | SWL-ranking <= 163.5 : LM26 (2/0%)
 | | | | | SWL-ranking > 163.5 : LM27 (5/3.49%)
 | | SWL-ranking > 168.5 :
 | | | SWL-ranking <= 171.5 : LM28 (3/4.126%)
 | | | SWL-ranking > 171.5 :
 | | | | SWL-ranking <= 175.5 :
 | | | | | life-expectancy <= 66.9 : LM29 (2/0%)
 | | | | | life-expectancy > 66.9 : LM30 (2/8.726%)
 | | | | SWL-ranking > 175.5 : LM31 (3/12.353%)

LM num: 1

SWL-index =

+ 273.33

LM num: 2

SWL-index =

+ 260

LM num: 3

SWL-index =

+ 256.67

LM num: 4

SWL-index =

+ 252.0812

LM num: 5

SWL-index =

+ 246.67

LM num: 6

SWL-index =

+ 243.33

LM num: 7

SWL-index =

+ 239.63

LM num: 8

SWL-index =

+ 231.4271

LM num: 9

SWL-index =

+ 223.998

LM num: 10

SWL-index =

+ 220

LM num: 11

SWL-index =

+ 216.67

LM num: 12

SWL-index =

+ 213.33

LM num: 13

SWL-index =

+ 208.7891

LM num: 14

SWL-index =

+ 203.33

LM num: 15

SWL-index =

+ 197.5025

LM num: 16

SWL-index =

+ 193.33

LM num: 17

SWL-index =

+ 190

LM num: 18

SWL-index =

+ 185.2783

LM num: 19

SWL-index =

+ 180

LM num: 20

SWL-index =

+ 175.2386

LM num: 21

SWL-index =

+ 170

LM num: 22

SWL-index =

+ 163.8867

LM num: 23

SWL-index =

+ 160

LM num: 24

SWL-index =

+ 155.7157

LM num: 25

SWL-index =

+ 150

LM num: 26

SWL-index =

+ 146.67

LM num: 27

SWL-index =

+ 142.664

LM num: 28

SWL-index =

+ 134.4433

LM num: 29

SWL-index =

+ 120

LM num: 30

SWL-index =

+ 120

LM num: 31

SWL-index =

+ 106.6667

Number of Rules : 31

Time taken to build model: 0.01 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient	0.9948
Kendall's tau	0.9722
Spearman's rho	0.9968
Mean absolute error	2.4737
Root mean squared error	3.8921
Relative absolute error	7.6425 %
Root relative squared error	10.1985 %
Total Number of Instances	178

[5 points] Build a model tree for this dataset in Weka using M5P with the following parameters:

build regression tree: False unpruned: False useUnsmoothed: True default values
for the remaining parameters

4.

Record the tree in your report [5 points].

=== Run information ===

Scheme: weka.classifiers.trees.M5P -U -M 4.0 -num-decimal-places 4

Relation: world_happiness-weka.filters.unsupervised.attribute.Remove-R2

Instances: 178

Attributes: 6

continent

SWL-ranking

SWL-index

life-expectancy

GDP-per-capita

access-to-education-score

Test mode: 2-fold cross-validation

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

M5 pruned model tree:

SWL-ranking <= 95.5 :

| SWL-ranking <= 55.5 :

| | SWL-ranking <= 41.5 :

| | | SWL-ranking <= 15.5 :

| | | | SWL-ranking <= 4.5 :

| | | | | SWL-ranking <= 2.5 : LM1 (2/0%)

| | | | | SWL-ranking > 2.5 : LM2 (2/0%)

| | | | SWL-ranking > 4.5 : LM3 (11/2.303%)

| | | SWL-ranking > 15.5 :

| | | | SWL-ranking <= 25.5 : LM4 (10/0%)

| | | | SWL-ranking > 25.5 :

| | | | | SWL-ranking <= 32.5 : LM5 (7/0%)

| | | | | SWL-ranking > 32.5 : LM6 (9/2.742%)

| | SWL-ranking > 41.5 : LM7 (14/4.318%)

| SWL-ranking > 55.5 : LM8 (40/3.267%)

SWL-ranking > 95.5 :

| SWL-ranking <= 142.5 : LM9 (47/3.288%)

| SWL-ranking > 142.5 :

| | SWL-ranking <= 168.5 : LM10 (26/3.435%)

| | SWL-ranking > 168.5 : LM11 (10/5.96%)

LM num: 1

SWL-index =

+ 273.33

LM num: 2

SWL-index =

+ 260

LM num: 3

SWL-index =

-0.7276 * SWL-ranking

+ 260.6091

LM num: 4

SWL-index =

+ 246.67

LM num: 5

SWL-index =

+ 243.33

LM num: 6

SWL-index =

+ 239.63

LM num: 7

SWL-index =

+ 231.4271

LM num: 8

SWL-index =

-0.5513 * SWL-ranking

+ 256.2046

LM num: 9

SWL-index =

-0.6002 * SWL-ranking

+ 254.6806

LM num: 10

SWL-index =

-1.0131 * SWL-ranking

+ 311.7692

LM num: 11

SWL-index =

-3.818 * SWL-ranking

+ 782.756

Number of Rules : 11

Time taken to build model: 0.01 seconds

=== Cross-validation ===

=== Summary ===

Correlation coefficient	0.9976
Kendall's tau	0.9854
Spearman's rho	0.9993
Mean absolute error	1.6844
Root mean squared error	2.6748
Relative absolute error	5.2039 %
Root relative squared error	7.0089 %
Total Number of Instances	178

Raking: 8/10 took some background research but overall easy to interpret

