HOMEWORK-6

### Numeric Prediction

For this homework, use the [world happiness dataset](http://csis.pace.edu/~benjamin/teaching/cs619/webfiles/world_happiness_w_continents.arff).

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.

1. [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

Record the tree in your report.

**ANSWER: -**

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%)

1. [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

Record the tree in your report [5 points].

**ANSWER: -**

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%)

1. [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

Record the tree in your report.

**ANSWER: -**

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%)

1. [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

Record the tree in your report [5 points].

**ANSWER: -**

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%)