Project 2.2

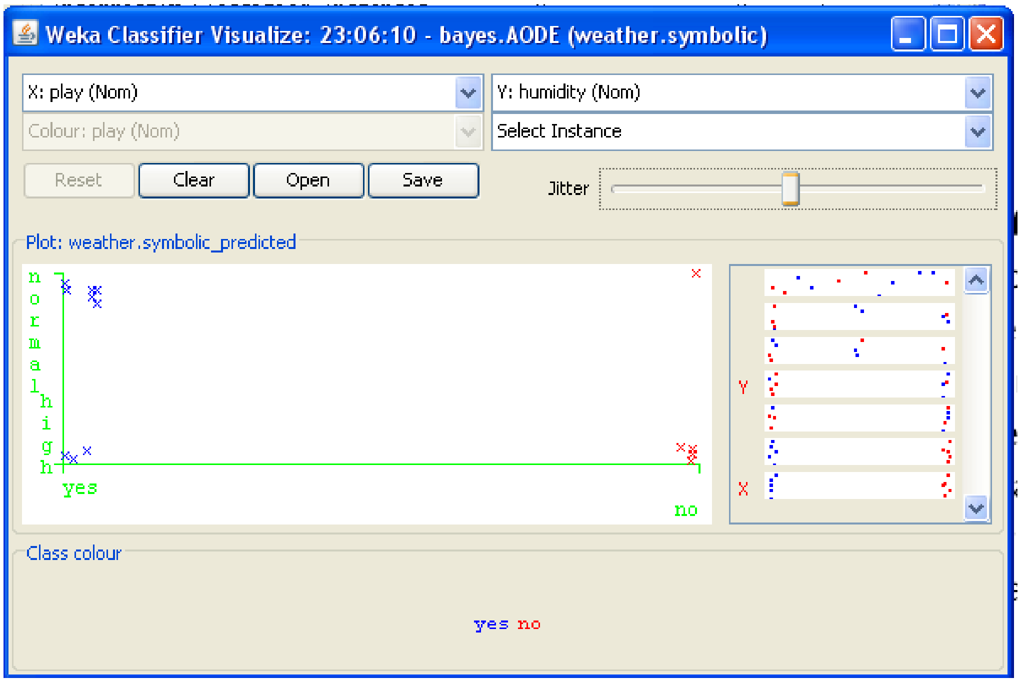
Experiment with Bayes Classifiers

AOED Classifier:

Averaged One-Dependence Estimators (AODE) is a probabilistic classification learning technique. It was developed to address the attribute-independence problem of the popular Naive Bayes classifier. It frequently develops substantially more accurate classifiers than naive Bayes at the cost of a modest increase in the amount of computation. Like naive Bayes, AODE does not perform model selection and hence has low variance. It supports incremental learning whereby the classifier can be updated efficiently with information from new examples as they become available. It predicts class probabilities rather than simply predicting a single class, allowing the user to determine the confidence with which each classification can be made. Its probabilistic model can directly handle situations where some data are missing.

As you can see the classification error has been %50, which is noticeable.

I was very interested to see the role of humidity in our classifier. So I visualized the error vs. the humidity. As you can see, people generally want to play when the humidity is normal (and hence the error is small), but when the humidity is high, then the result depends on other factors (like wind).



Weka Classifier Visualize: 23:06:10 - bayes.AODE (weather.symbolic)

X: play (Nom)

Y: humidity (Nom)

Colour: play (Nom)

Plot: weather.symbolic_predicted

```
=== Run information ===

Scheme:       weka.classifiers.bayes.AODE -F 1

Relation:     weather.symbolic

Instances:    14

Attributes:   5

              outlook

              temperature

              humidity

              windy

              play

Test mode:    10-fold cross-validation


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


The AODE Classifier

Class yes: Prior probability = 0.63



Class no: Prior probability = 0.38


Dataset: weather.symbolic

Instances: 14

Attributes: 5

Frequency limit for superParents: 1

Correction: laplace



Time taken to build model: 0 seconds
```

```
=== Stratified cross-validation ===

=== Summary ===


Correctly Classified Instances          7               50      %

Incorrectly Classified Instances        7               50      %

Kappa statistic                       -0.0426

Mean absolute error                    0.4706

Root mean squared error                0.498

Relative absolute error               98.8198 %

Root relative squared error          100.941  %

Total Number of Instances             14


=== Detailed Accuracy By Class ===


               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class

                 0.556     0.6       0.625     0.556     0.588      0.533    yes

                 0.4       0.444     0.333     0.4       0.364      0.533    no

Weighted Avg.    0.5       0.544     0.521     0.5       0.508      0.533


=== Confusion Matrix ===


 a b   <-- classified as

 5 4 | a = yes

 3 2 | b = no
```

BayesNet

Bayes Network learning using various search algorithms and quality measures.

Base class for a Bayes Network classifier. Provides datastructures (network structure, conditional probability distributions, etc.) and facilities common to Bayes Network learning algorithms like K2 and B.Application of BayesNet gives ~57% of accuracy. Below is the run information:

```
=== Run information ===


Scheme:       weka.classifiers.bayes.BayesNet -D -Q weka.classifiers.bayes.net.search.local.K2 -- -P 1 -S BAYES -E weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5

Relation:     weather.symbolic

Instances:    14

Attributes:   5

              outlook

              temperature

              humidity

              windy

              play

Test mode:    10-fold cross-validation


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


Bayes Network Classifier

not using ADTree

#attributes=5 #classindex=4

Network structure (nodes followed by parents)

outlook(3): play

temperature(3): play

humidity(2): play

windy(2): play
```

```
play(2):

LogScore Bayes: -69.07317135664013

LogScore BDeu: -83.46880542273105

LogScore MDL: -82.71568504897063

LogScore ENTROPY: -65.56181240647145

LogScore AIC: -78.56181240647145



Time taken to build model: 0 seconds


=== Stratified cross-validation ===

=== Summary ===


Correctly Classified Instances          8               57.1429 %

Incorrectly Classified Instances        6               42.8571 %

Kappa statistic                        -0.0244

Mean absolute error                     0.415

Root mean squared error                 0.4909

Relative absolute error                87.1501 %

Root relative squared error            99.5104 %

Total Number of Instances              14


=== Detailed Accuracy By Class ===


               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
```

```
              0.778    0.8      0.636    0.778    0.7      0.622    yes

              0.2      0.222    0.333    0.2      0.25     0.622    no

Weighted Avg.    0.571    0.594    0.528    0.571    0.539    0.622


=== Confusion Matrix ===


 a b   <-- classified as

 7 2 | a = yes

 4 1 | b = no
```

Naïve Bayes

This classifier is expected to give a worse accuracy because of the assumption of independence. Class for a Naive Bayes classifier using estimator classes. Numeric estimator precision values are chosen based on analysis of the training data. For this reason, the classifier is not an UpdateableClassifier (which in typical usage are initialized with zero training instances) -- if you need the UpdateableClassifier functionality, use the NaiveBayesUpdateable classifier. The NaiveBayesUpdateable classifier will use a default precision of 0.1 for numeric attributes when buildClassifier is called with zero training instances.

```
=== Run information ===


Scheme:       weka.classifiers.bayes.NaiveBayes

Relation:     weather.symbolic

Instances:    14

Attributes:   5

              outlook

              temperature

              humidity
              windy
              play
Test mode:    10-fold cross-validation

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

Naive Bayes Classifier

                Class
Attribute         yes     no
               (0.63) (0.38)
============================
outlook
  sunny           3.0    4.0
  overcast         5.0    1.0
  rainy           4.0    3.0
  [total]        12.0    8.0

temperature
  hot             3.0    3.0
  mild            5.0    3.0
  cool            4.0    2.0
  [total]        12.0    8.0

humidity
```

```
  high          4.0    5.0
  normal          7.0    2.0
  [total]        11.0    7.0

windy
  TRUE           4.0    4.0
  FALSE          7.0    3.0
  [total]        11.0    7.0



Time taken to build model: 0 seconds


=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances           8              57.1429 %
Incorrectly Classified Instances         6              42.8571 %
Kappa statistic                        -0.0244
Mean absolute error                      0.4374
Root mean squared error                  0.4916
Relative absolute error                 91.8631 %
Root relative squared error             99.6492 %
Total Number of Instances               14
```

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

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.778     0.8       0.636     0.778     0.7        0.578    yes
                 0.2       0.222     0.333     0.2       0.25       0.578    no
Weighted Avg.    0.571     0.594     0.528     0.571     0.539      0.578

=== Confusion Matrix ===

 a b   <-- classified as
 7 2 | a = yes
 4 1 | b = no
```

Experiment with ANN

Multi-layer Perceptron

Experiment with ANN gives ~75 accuracy at best with variations of number of hidden layers.

```
=== Run information ===

Scheme:       weka.classifiers.functions.MultilayerPerceptron -L 0.3 -M 0.2 -N 500 -V 0 -S 0 -E 20 -H a
Relation:     weather.symbolic
Instances:    14
Attributes:   5
              outlook
              temperature

              humidity

              windy

              play

Test mode:    10-fold cross-validation


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


Sigmoid Node 0

    Inputs    Weights

    Threshold    -4.597967080790813

    Node 2    2.433270074007239

    Node 3    2.0546443732203774

    Node 4    1.364159803860347

    Node 5    2.6974766889493536

    Node 6    3.908322709064356

Sigmoid Node 1

    Inputs    Weights

    Threshold    4.601251960011152

    Node 2    -2.4045226373071156

    Node 3    -2.0532744956144127

    Node 4    -1.379986429753948

    Node 5    -2.756274547604192

    Node 6    -3.877948258791871

Sigmoid Node 2
```

```
    Inputs    Weights
    Threshold    -0.1550798021501342
    Attrib outlook=sunny    -1.323464477913686
    Attrib outlook=overcast    1.6602675280399888
    Attrib outlook=rainy    -0.3207802552865604
    Attrib temperature=hot    -0.2873122456981835
    Attrib temperature=mild    1.181190360097958
    Attrib temperature=cool    -0.7853150475848826
    Attrib humidity    2.808930687905
    Attrib windy    1.9190213581350706
Sigmoid Node 3
    Inputs    Weights
    Threshold    -0.18031675012278034
    Attrib outlook=sunny    -1.1524514010228344
    Attrib outlook=overcast    1.5760227701429683
    Attrib outlook=rainy    -0.32578400279223824
    Attrib temperature=hot    -0.2760307631136823
    Attrib temperature=mild    1.0450876279343007
    Attrib temperature=cool    -0.6318819517738498
    Attrib humidity    2.4504774603875408
    Attrib windy    1.678251292646871
Sigmoid Node 4
    Inputs    Weights
    Threshold    -0.3554146745674961
    Attrib outlook=sunny    -0.46574052680925143
    Attrib outlook=overcast    1.4382073898080827
    Attrib outlook=rainy    -0.6194183985830608
    Attrib temperature=hot    -0.0670794406887232
    Attrib temperature=mild    0.6337484752708613
    Attrib temperature=cool    -0.20814280117719502
    Attrib humidity    1.982466584793048
    Attrib windy    0.9946423645131915
Sigmoid Node 5
    Inputs    Weights
    Threshold    -0.06888405078498452
    Attrib outlook=sunny    -1.3982064219096493
    Attrib outlook=overcast    1.8084944112736516
    Attrib outlook=rainy    -0.31997269602762973
    Attrib temperature=hot    -0.3035821635771427
    Attrib temperature=mild    1.2908528760310662
    Attrib temperature=cool    -0.8921466424329777
    Attrib humidity    3.1090049574873424
    Attrib windy    2.0747113212966872
Sigmoid Node 6
    Inputs    Weights
    Threshold    0.04399369934901554
    Attrib outlook=sunny    -1.80182134279014
    Attrib outlook=overcast    2.2544547024444554
    Attrib outlook=rainy    -0.40095717506501327
    Attrib temperature=hot    -0.41558677311306397
    Attrib temperature=mild    1.589170285947685
    Attrib temperature=cool    -1.2545441906677217
    Attrib humidity    4.119310666164331
    Attrib windy    2.740851006387263
Class yes
    Input
    Node 0
Class no
    Input
    Node 1


Time taken to build model: 0.45 seconds


=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances          10               71.4286 %
Incorrectly Classified Instances         4               28.5714 %
Kappa statistic                          0.3778
Mean absolute error                      0.287
Root mean squared error                  0.5268
Relative absolute error                 60.2616 %
Root relative squared error            106.7798 %
Total Number of Instances               14
```

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

             TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class

              0.778     0.4       0.778     0.778     0.778      0.778    yes

              0.6       0.222      0.6       0.6       0.6        0.778    no

Weighted Avg.    0.714     0.337     0.714     0.714     0.714      0.778

=== Confusion Matrix ===

 a b   <-- classified as

 7 2 | a = yes

 2 3 | b = no
```

Voted Perceptron

This network works much worse and gives ~57% accuracy at best. Below are the run information.

```
=== Run information ===

Scheme:       weka.classifiers.functions.VotedPerceptron -I 1 -E 1.0 -S 1 -M 10000

Relation:     weather.symbolic

Instances:    14

Attributes:   5

              outlook

              temperature
              humidity

              windy

              play

Test mode:    10-fold cross-validation


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


VotedPerceptron: Number of perceptrons=4


Time taken to build model: 0 seconds


=== Stratified cross-validation ===

=== Summary ===


Correctly Classified Instances           8               57.1429 %

Incorrectly Classified Instances         6               42.8571 %

Kappa statistic                       -0.1351

Mean absolute error                      0.3918

Root mean squared error                  0.5825

Relative absolute error                82.2829 %

Root relative squared error            118.0664 %

Total Number of Instances               14


=== Detailed Accuracy By Class ===
          TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class

           0.889     1         0.615     0.889     0.727      0.622     yes

           0        0.111      0        0         0          0.622     no

Weighted Avg.    0.571     0.683      0.396     0.571     0.468      0.622


=== Confusion Matrix ===


 a b   <-- classified as

 8 1  | a = yes

 5 0  | b = no
```

Winnow Network

Winnow network works based on sparse connections and hence can't beat the multi-layer perceptron. Here it gives only 50% of accuracy. Below is the run information.

```
=== Run information ===


Scheme:       weka.classifiers.functions.Winnow -I 1 -A 2.0 -B 0.5 -H -1.0 -W 2.0 -S 1

Relation:     weather.symbolic

Instances:    14

Attributes:   5

              outlook

              temperature

              humidity

              windy

              play

Test mode:    10-fold cross-validation
```

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

Winnow

Attribute weights

w0 8.0
w1 1.0
w2 2.0
w3 4.0
w4 2.0
w5 2.0
w6 1.0
w7 1.0

Cumulated mistake count: 7



Time taken to build model: 0 seconds

=== Stratified cross-validation ===
=== Summary ===
Correctly Classified Instances          7             50      %

Incorrectly Classified Instances        7             50      %

Kappa statistic                      -0.2564

Mean absolute error                    0.5

Root mean squared error                0.7071

Relative absolute error              105      %

Root relative squared error          143.3236 %

Total Number of Instances              14


=== Detailed Accuracy By Class ===


            TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class

             0.778     1         0.583     0.778     0.667      0.389    yes

             0        0.222      0        0         0          0.389    no

Weighted Avg.    0.5       0.722      0.375     0.5       0.429      0.389


=== Confusion Matrix ===


 a b   <-- classified as

 7 2 | a = yes

 5 0 | b = no
```

Discuss Your Results

As mentioned before, due to dependence on sparse connections, Winnow is the worst choice for the network when we have few attributes and a small database. Winnow is good when we are dealing with numerous attributes and large dataset which contain redundancy. With our dataset, Multi-layer Perceptron gives the best results.