John M. Warlop

2018-05-22

Lab #3: Bayesian Network

UCSD Extension - Data Mining II Adv. Concepts & Techniques

1	What is Bayes Rule Formula? $P(A|B) = [P(B|A) * P(A)]/P(B)$, which reads: The probability of A being true given that B is true is equal to the probability of B being true given that A is true times the probability of A being true, all divided by the probability of B being true. Sometimes, this is written as: $P(H|E) = [P(E|H)*P(E)]/P(H)$ or the probability of hypothesis being true given evidence E is equal to the probability of the evidence is true given the hypothesis times the probability of hypothesis being true all divided by the probability of the evidence being true.
2	What does "maximum likelihood" mean? Given a distribution, "maximum likelihood" finding the optimal to fit a distribution to some data. In other words, you are looking for the optimal value for mean and standard deviation for a distribution given observed data -- assuming normal distribution.
3	What is Bayesian Learning? The process of forming a bayesian network by using Bayes theorem, total probability and conditional probability. The learning process involves two parts: #1 propogating the effect of independent variables(parent nodes) on dependent variables(child nodes) and #2: infer the effect of observations on dependent variables while updating the knowledge of independent variables(parent nodes). #1 is a probabilistic problem while #2 is "reverse probability" or liklihood problem.

4	What is a prior? What is posterior? In statistics, prior refers to the probability before a test is run. Posterior probability is a new probability attained after new evidence is gathered.
5	LNK In Weka open the "BayesNetwork_Weather_data_Set.csv" and under the "Classify" Tab click on the Bayes Folder and then BayesNet method.

Train the Bayesian Network on the data set using several different parameters. First Train the netowrk by appying the Simple K2 search algorithm. You can visualize the model as well as the graph. You can also visualize probabilities by clicking on node.
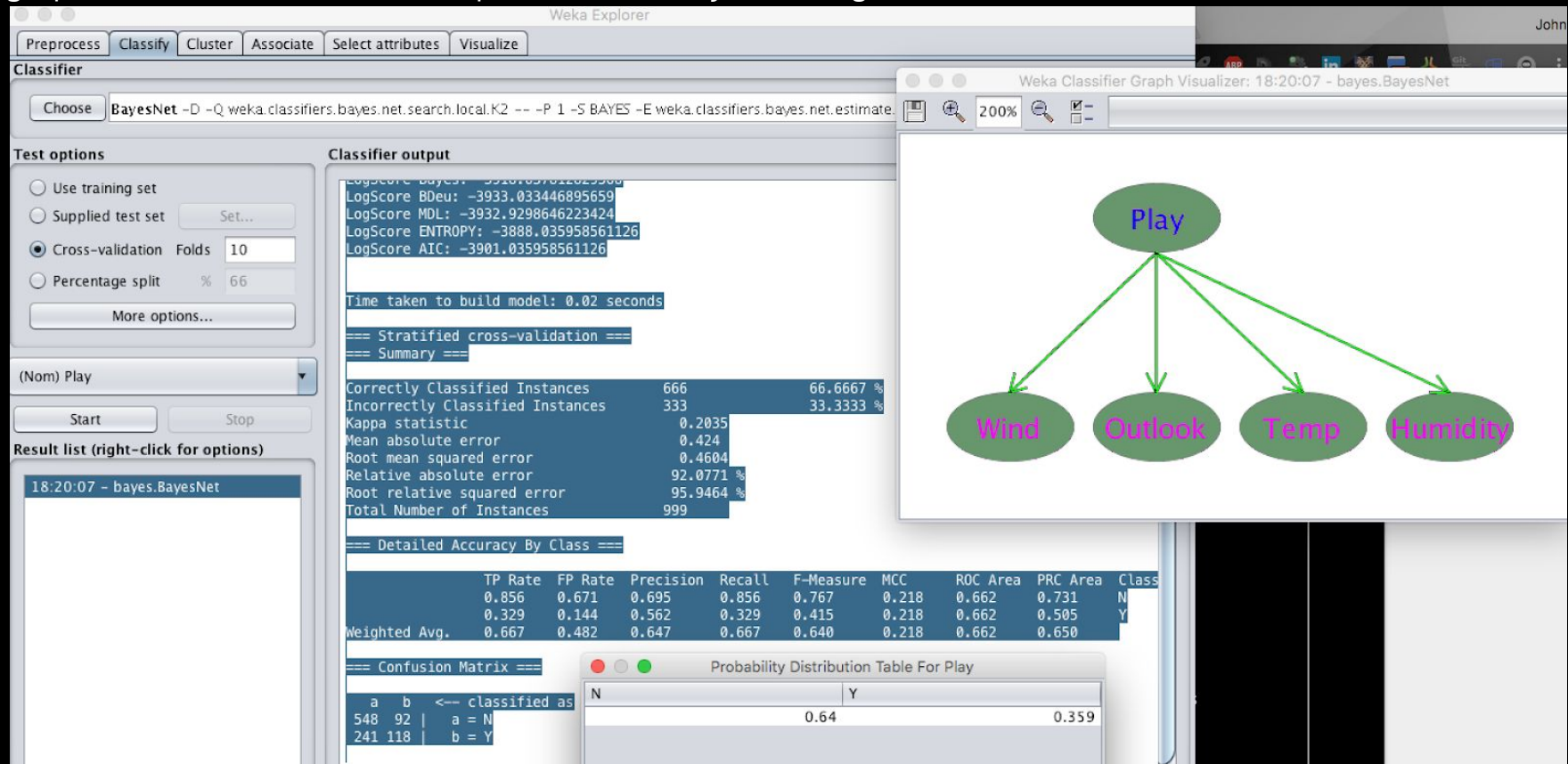
I ran BayesNet w/o Markov Blanket and with Markov Blanket and performance of model was unchanged(66.67% and 33.33%). I also switched on random node order, this too had no effect on model accuracy. I also tried ENTROPY switch with K2, no effect.

NAME
weka.classifiers.bayes.net.search.local.K2

SYNOPSIS
This Bayes Network learning algorithm uses a hill climbing algorithm restricted by an order on the variables.

For more information see:

G.F. Cooper, E. Herskovits (1990). A Bayesian method for constructing Bayesian belief networks from databases.

G. Cooper, E. Herskovits (1992). A Bayesian method for the induction of probabilistic networks from data. Machine Learning. 9(4):309-347.

Works with nominal variables and no missing values only.

OPTIONS
markovBlanketClassifier -- When set to true (default is false), after a network structure is learned a Markov Blanket correction is applied to the network structure. This ensures that all nodes in the network are part of the Markov blanket of the classifier node.

randomOrder -- When set to true, the order of the nodes in the network is random. Default random order is false and the order of the nodes in the dataset is used. In any case, when the network was initialized as Naive Bayes Network, the class variable is first in the ordering though.

scoreType -- The score type determines the measure used to judge the quality of a network structure. It can be one of Bayes, BDeu, Minimum Description Length (MDL), Akaike Information Criterion (AIC), and Entropy.

initAsNaiveBayes -- When set to true (default), the initial network used for structure learning is a Naive Bayes Network, that is, a network with an arrow from the classifier node to each other node. When set to false, an empty network is used as initial network structure

maxNrOfParents -- Set the maximum number of parents a node in the Bayes net can have. When initialized as Naive Bayes, setting this parameter to 1 results in a Naive Bayes classifier. When set to 2, a Tree Augmented Bayes Network (TAN) is learned, and when set >2, a Bayes Net Augmented Bayes Network (BAN) is learned. By setting it to a value much larger than the number of nodes in the network (the default of 100000 pretty much guarantees this), no restriction on the number of parents is enforced

Repeat the training process by utilizing different serach algoirithms (ReapededHillClimber, SimulatedAnealing, etc under BayesianNet paremeters options). Save each graph, compare and contrast the confusion matrix, output classification errors and predictive power of each model. Describe and discuss in details the differences, pros/cons between the models.

Each of the seraching algorthms will have a set of their own parameters you can adjust:
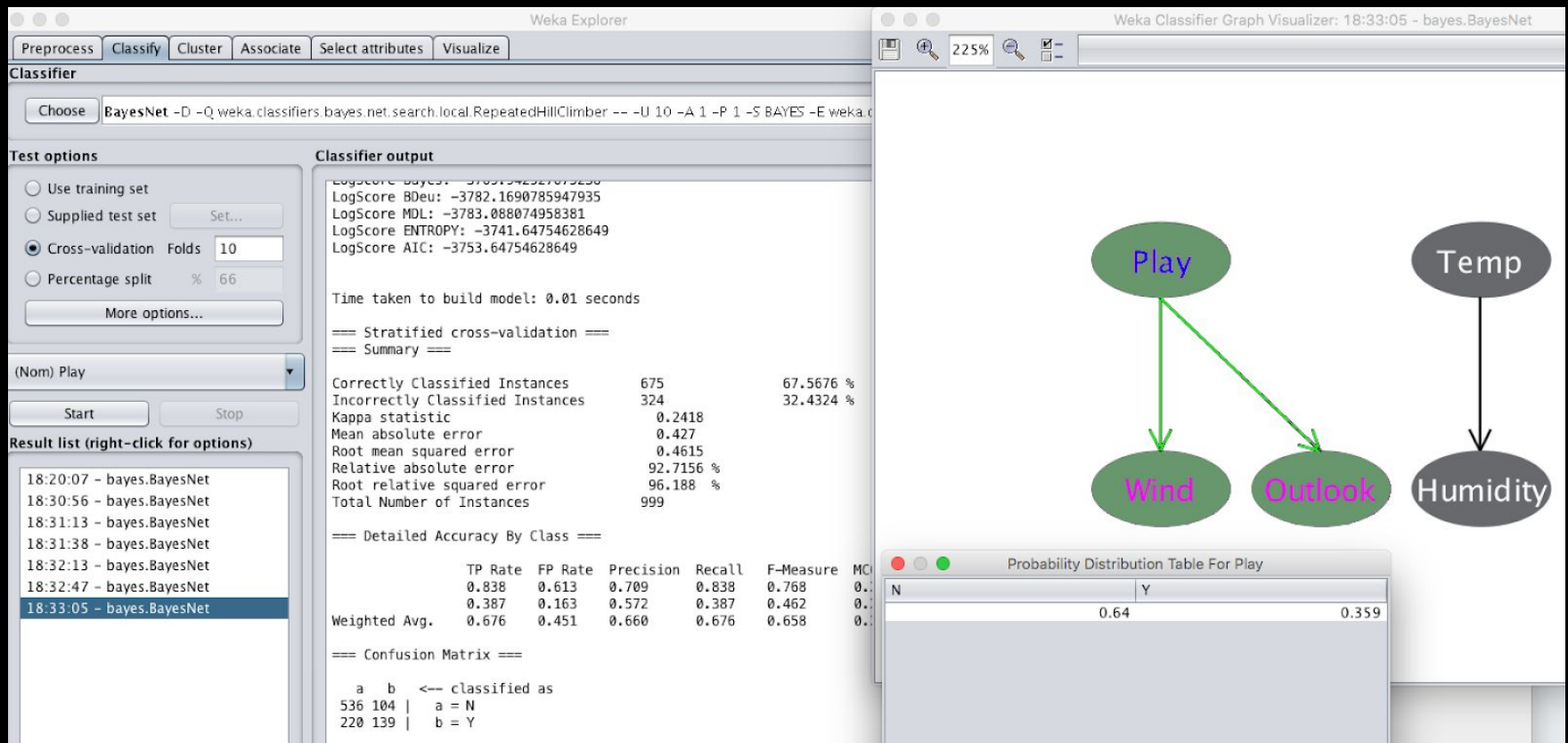
NAME
weka.classifiers.bayes.net.search.local.RepeatedHillClimber

SYNOPSIS
This Bayes Network learning algorithm repeatedly uses hill climbing starting with a randomly generated network structure and return the best structure of the various runs.

OPTIONS
runs -- Sets the number of times hill climbing is performed.

seed -- Initialization value for random number generator. Setting the seed allows replicability of experiments.

useArcReversal -- When set to true, the arc reversal operation is used in the search.

markovBlanketClassifier -- When set to true (default is false), after a network structure is learned a Markov Blanket correction is applied to the network structure. This ensures that all nodes in the network are part of the Markov blanket of the classifier node.

scoreType -- The score type determines the measure used to judge the quality of a network structure. It can be one of Bayes, BDeu, Minimum Description Length (MDL), Akaike Information Criterion (AIC), and Entropy.

initAsNaiveBayes -- When set to true (default), the initial network used for structure learning is a Naive Bayes Network, that is, a network with an arrow from the classifier node to each other node. When set to false, an empty network is used as initial network structure

maxNrOfParents -- Set the maximum number of parents a node in the Bayes net can have. When initialized as Naive Bayes, setting this parameter to 1 results in a Naive Bayes classifier. When set to 2, a Tree Augmented Bayes Network (TAN) is learned, and when set >2, a Bayes Net Augmented Bayes Network (BAN) is learned. By setting it to a value much larger than the number of nodes in the network (the default of 100000 pretty much guarantees this), no restriction on the number of parents is enforced.

Running the RepeatedHillCllimber as the search algorithm returned a slightly better model. I tried several different parameters as I did on K2(ScoreType => Entropy, initAsNaiveBayes, maxNumParents) none of these parameter changes had a significant impact on model accuracy(see below).
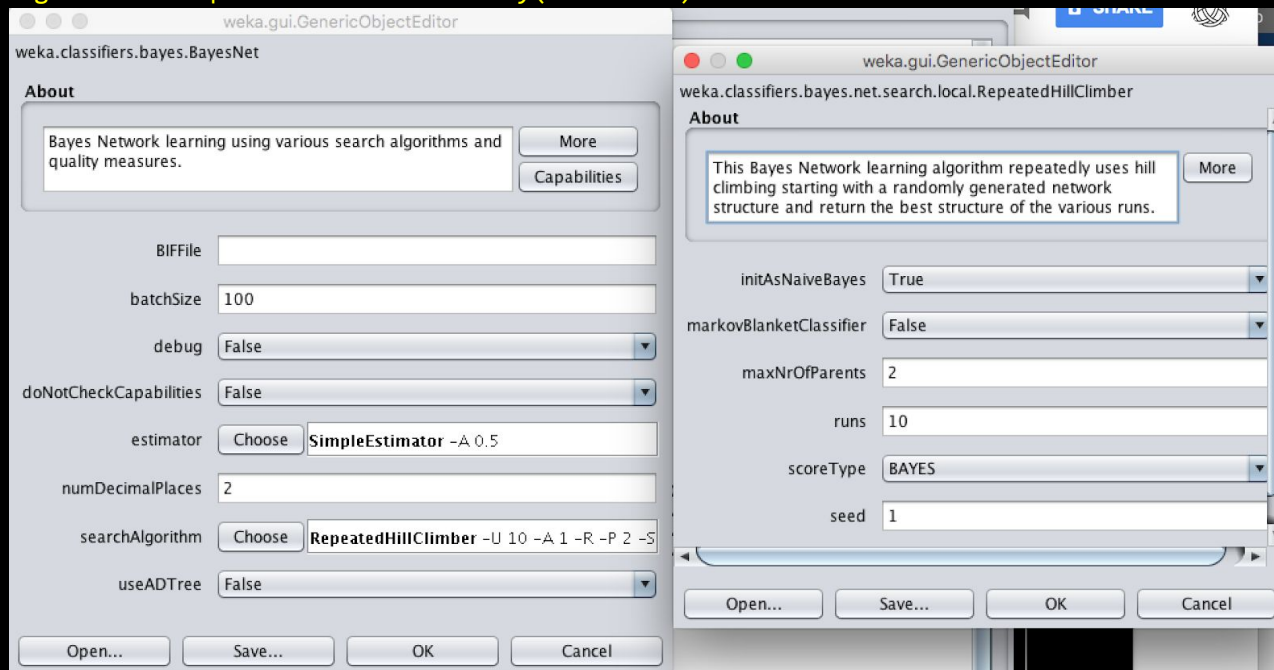
Part 2:

From Weka's Start Menu click on Tools and then Bayes net Editor
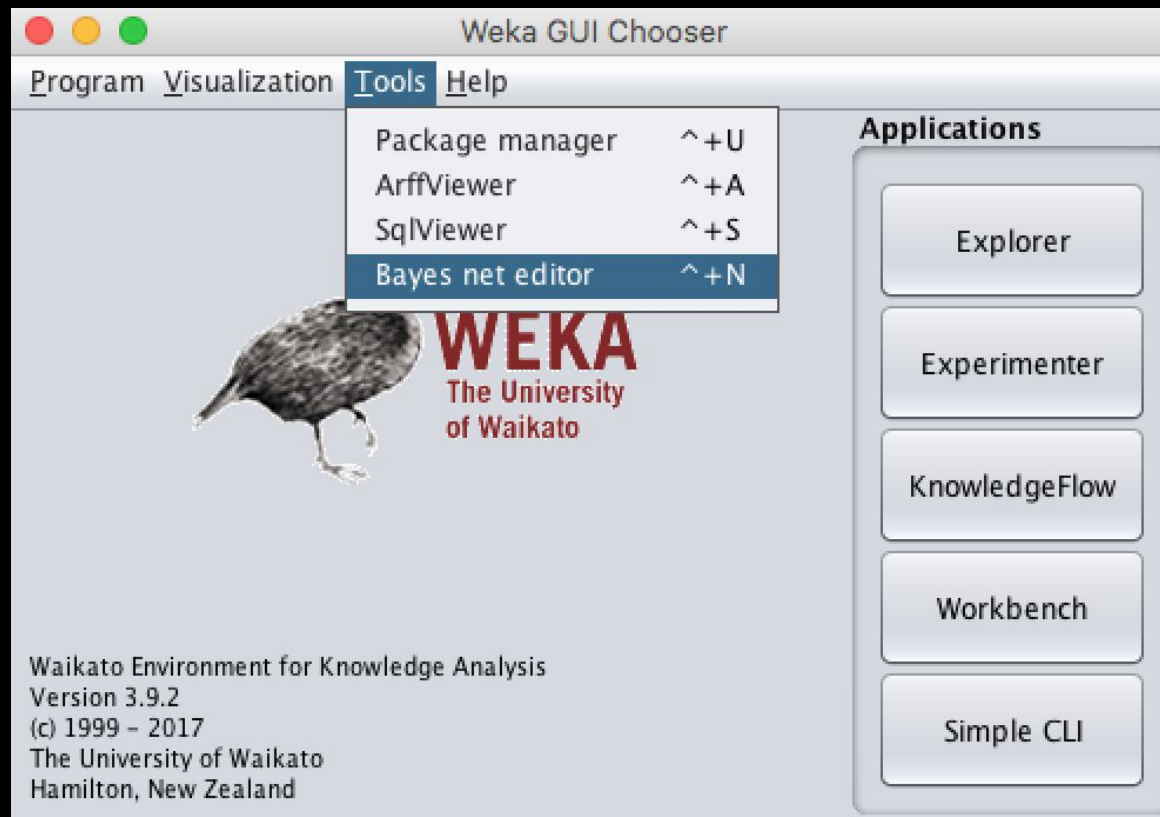
Now you can load in the graph structure you have previously saved in the XML BIF format.

In this Editor, you can now click on Tools – Show Margins to see the probabilities associated with probability table for each node.
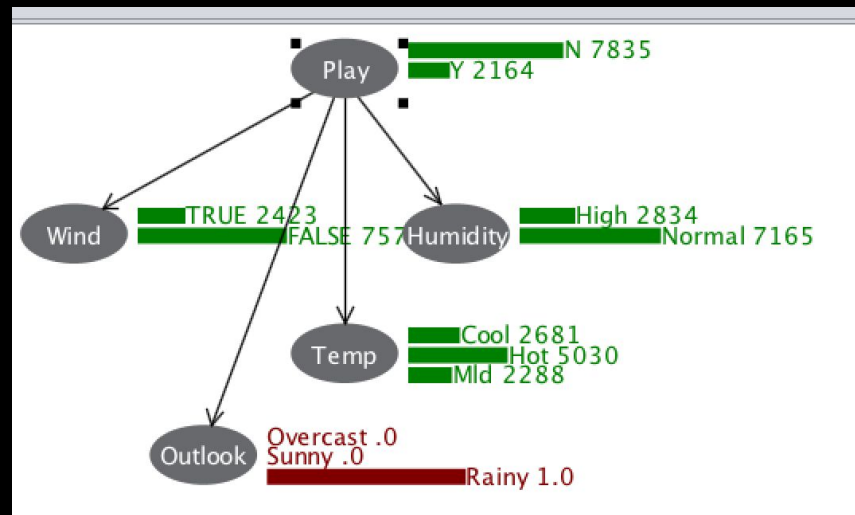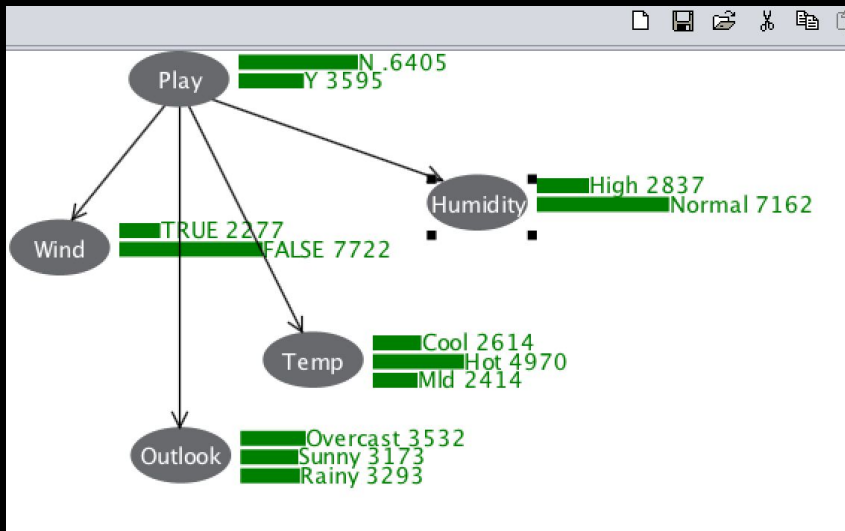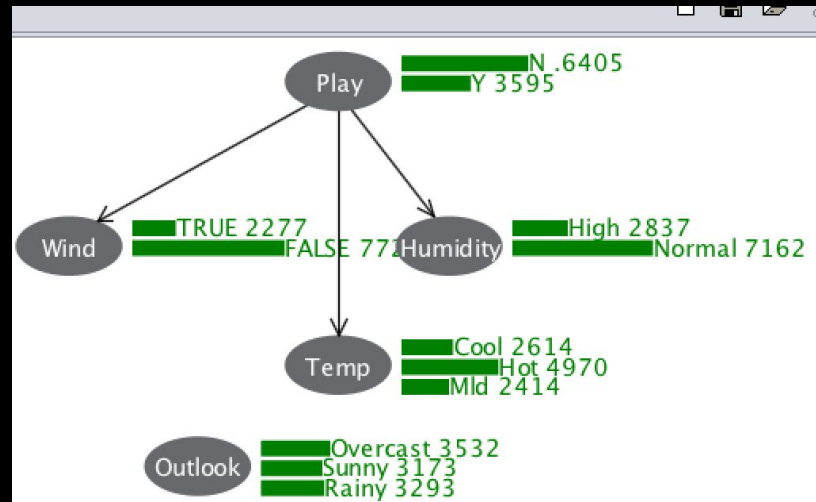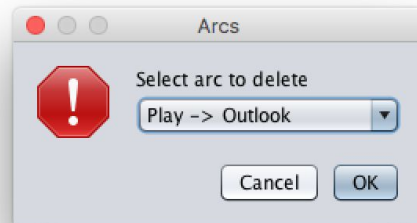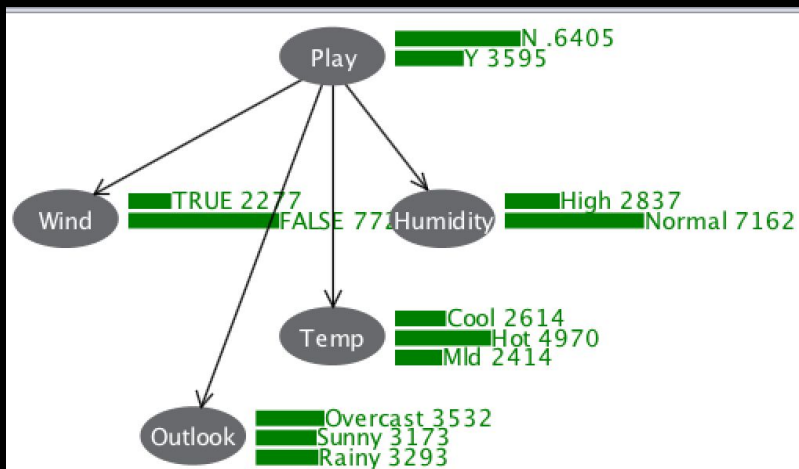
Describe and discuss the apparent changes. Observe how the evidence has propagated, what else changed? Why? When the evidence is changed, the probabilities adjust as well. The adjustments are necessary because of Bayes Theorem. If you change the P(E|H) and each evidence piece is independent, this will change the overall graph.

In this Editor, you can delete or add arcs and/or nodes from the Edit tab. Once you have deleted the arc of your choice, you can retrain by set Data and Learn CPT commands. You can now set evidence on one of the nodes and explain how these changes influenced the posterior probabilities in the graph.
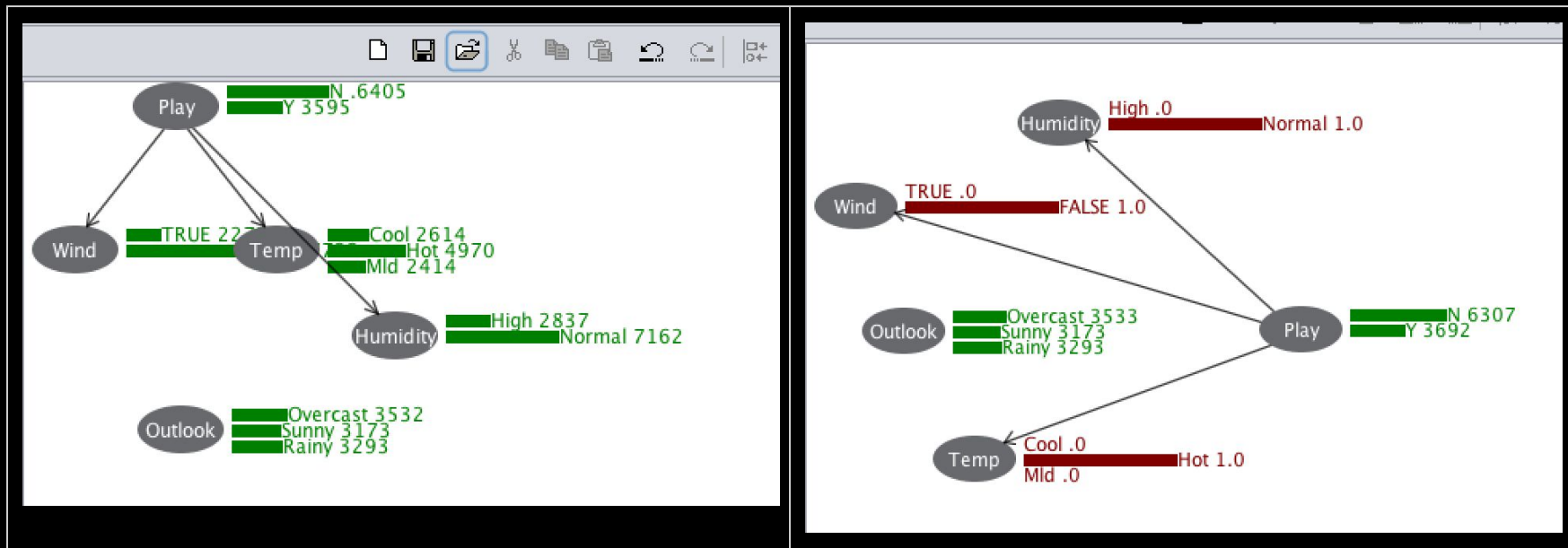
When I deleted an arc and then changed the evidence, the posterior probabilities changed. The outlook node did not change because it is not connected to network after deleting the arc.

Below as you can see, I merged data values that were more suitable for playing a game outside. I used the WEKA mergeattributes filter. I Merged Outlook(overcast,sunny) and Temp(Cool,Mild). I then deleted arc Humidity->Outlook and the node probabilities changed. The probabilities must change because the bayesian equation must be valid for each node.

Left window:
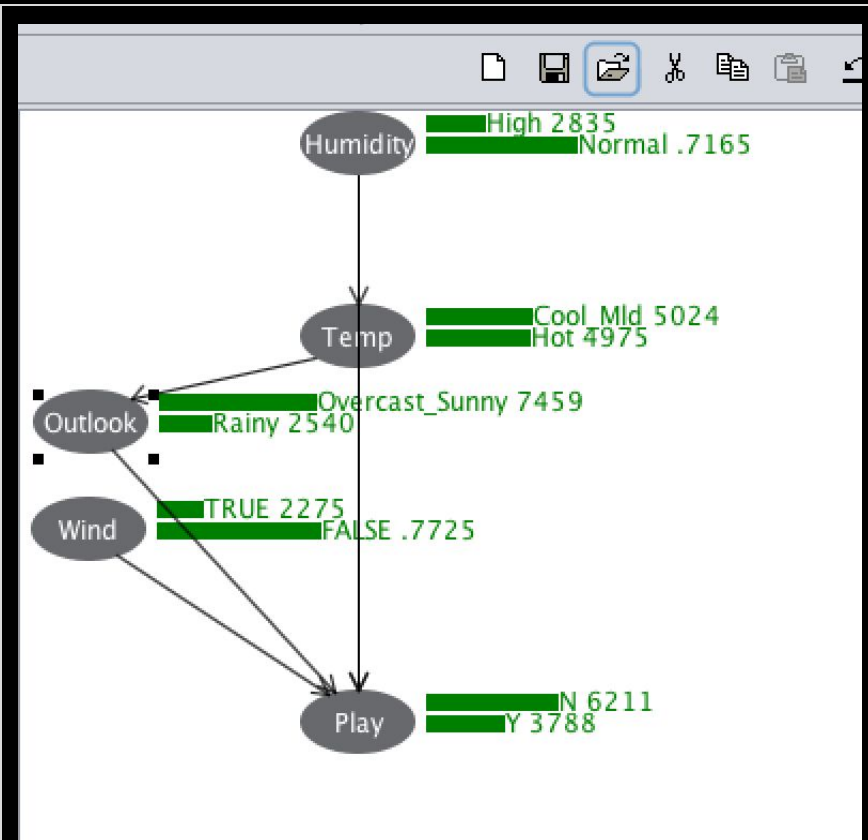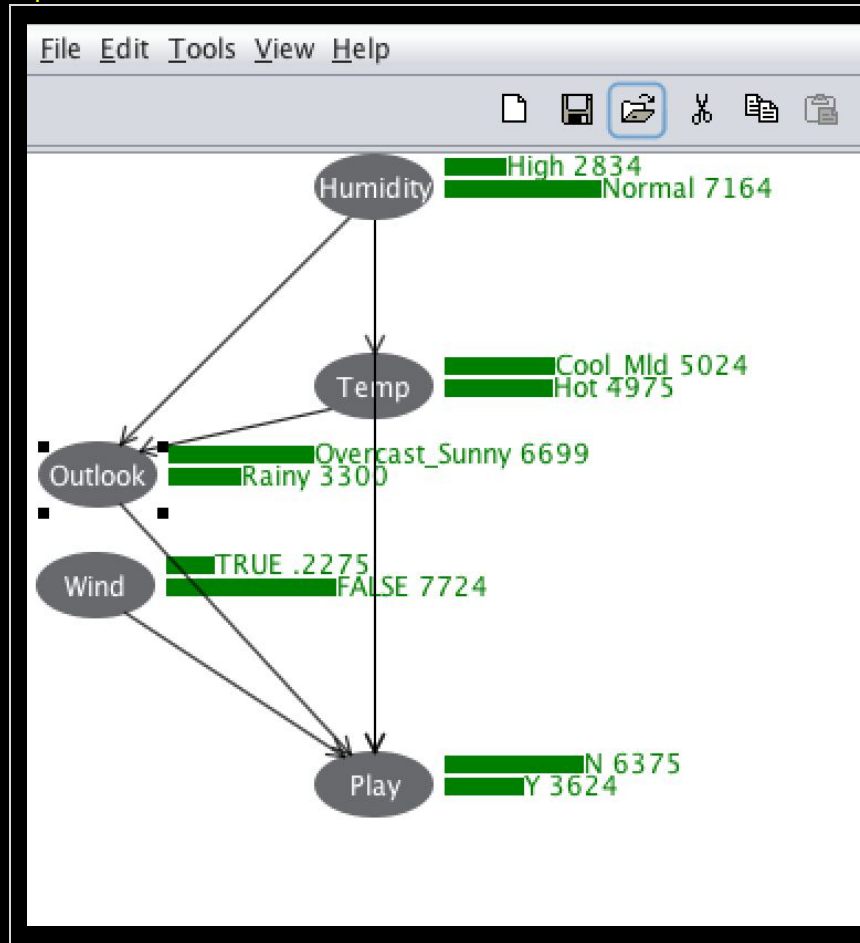
File   Edit   Tools   View   Help

Humidity — High 2835 / Normal .7165

Temp — Cool_Mld 5024 / Hot 4975

Outlook — Overcast_Sunny 7459 / Rainy 2540

Wind — TRUE 2275 / FALSE .7725

Play — N 6211 / Y 3788
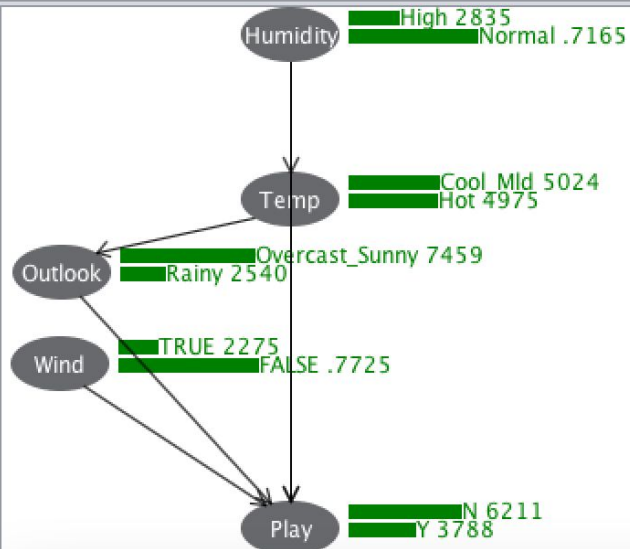
Probability Distribution Table For Outlook

Temp	Overcast_Sunny	Rainy
Cool_Mld	0.737	0.263
Hot	0.755	0.245

Randomize   Ok   Cancel

Right window:

Humidity — High 2835 / Normal .7165

Temp — Cool_Mld 5024 / Hot 4975

Outlook — Overcast_Sunny 7459 / Rainy 2540

Wind — TRUE 2275 / FALSE .7725

Play — N 6211 / Y 3788
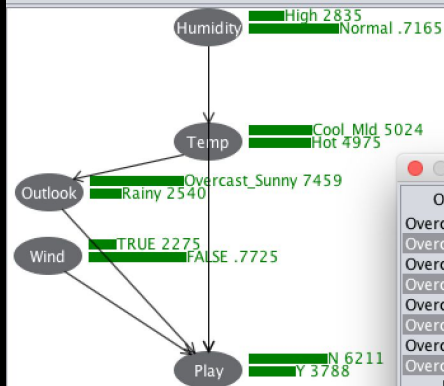
Probability Distribution Table For Play

Outlook	Wind	Temp	Humidity	N	Y
Overcast_Sunny	TRUE	Cool_Mld	High	0.855	0.145
Overcast_Sunny	TRUE	Cool_Mld	Normal	0.717	0.283
Overcast_Sunny	TRUE	Hot	High	0.792	0.208
Overcast_Sunny	TRUE	Hot	Normal	0.743	0.257
Overcast_Sunny	FALSE	Cool_Mld	High	0.518	0.482
Overcast_Sunny	FALSE	Cool_Mld	Normal	0.476	0.524
Overcast_Sunny	FALSE	Hot	High	0.482	0.518
Overcast_Sunny	FALSE	Hot	Normal	0.526	0.474
Rainy	TRUE	Cool_Mld	High	0.656	0.344
Rainy	TRUE	Cool_Mld	Normal	0.596	0.404
Rainy	TRUE	Hot	High	0.9	0.1
Rainy	TRUE	Hot	Normal	0.703	0.297
Rainy	FALSE	Cool_Mld	High	0.777	0.223
Rainy	FALSE	Cool_Mld	Normal	0.755	0.245
Rainy	FALSE	Hot	High	0.833	0.167
Rainy	FALSE	Hot	Normal	0.868	0.132

Randomize   Ok   Cancel
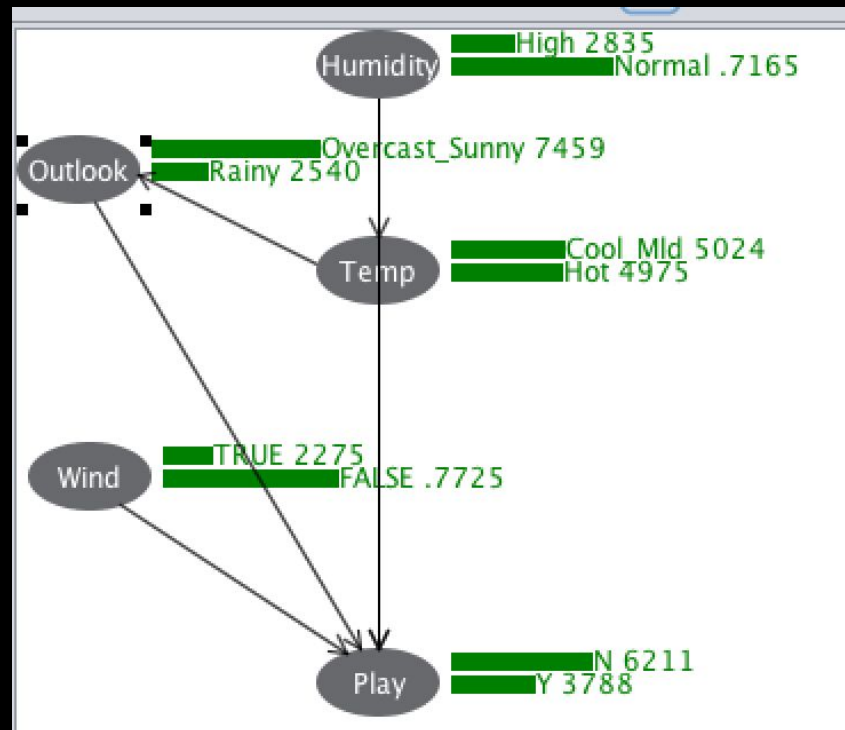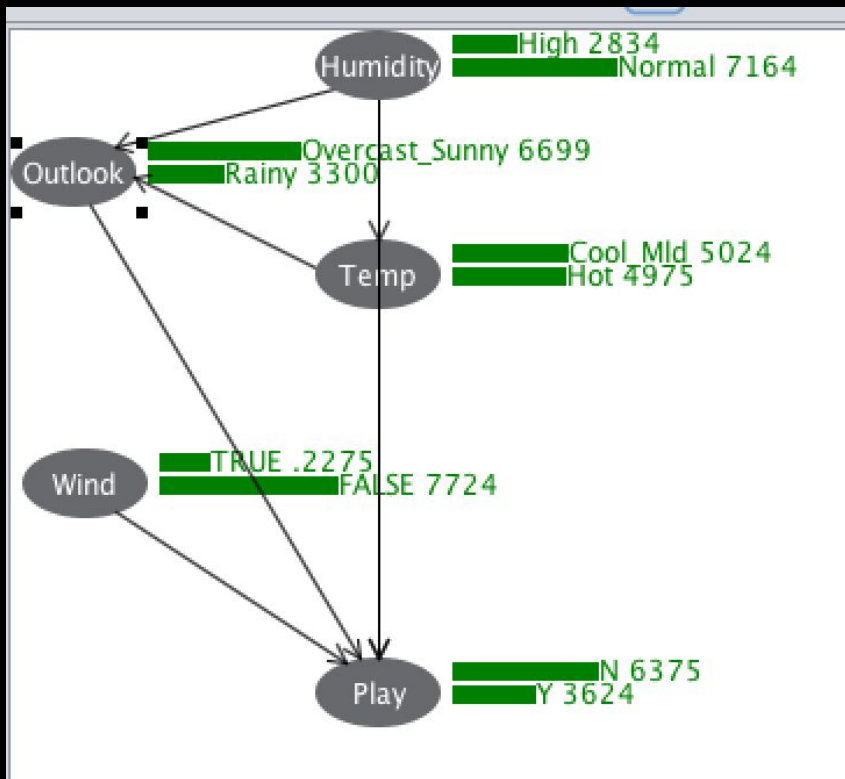
| 6 | Follow the process for training and visualizing the BN graph from the Question #5. Train the Bayesian Network and Decision Tree model on Hypothyroid data set. Describe the process and evaluation of the models. What are pros and cons of each? Which one performed better? Why? Which model is more suitable for particular applications? I loaded the hypothyroid data set into each classifier. I chose BayesNet and J48. J48 is easier to understand because it is a simple decision |

tree. It does not use probabilities. The J48 decision tree performed better than the BayesNet(99% to 98%). The J48 model does not use probabilities whereas the Bayesian network does use probabilities. The Bayesian network assumes each variable is independent. J48(decision tree) is easier to understand.

Classifier

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

Test options

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

More options...

(Nom) Class

Start   Stop

Result list (right-click for options)

14:01:42 – bayes.BayesNet

Classifier output

```
LogScore Aic:   23333.20323020033


Time taken to build model: 0.08 seconds

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

Correctly Classified Instances        3719               98.5949 %
Incorrectly Classified Instances        53                1.4051 %
Kappa statistic                          0.9028
Mean absolute error                      0.011
Root mean squared error                  0.075
Relative absolute error                 15.135  %
Root relative squared error             39.3888 %
Total Number of Instances             3772

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall  F-Measure  MCC    ROC Area  PRC Area  Class
                 0.993    0.086    0.993      0.993   0.993      0.910  0.997     1.000     negative
                 0.892    0.004    0.930      0.892   0.911      0.906  0.999     0.975     compensated_hypothyroid
                 0.926    0.004    0.854      0.926   0.889      0.887  0.996     0.921     primary_hypothyroid
                 0.000    0.000    ?          0.000   ?          ?      0.749     0.002     secondary_hypothyroid
Weighted Avg.    0.986    0.080    ?          0.986   ?          ?      0.997     0.996

=== Confusion Matrix ===

    a    b    c    d    <-- classified as
 3458    9   14    0 |    a = negative
   20  173    1    0 |    b = compensated_hypothyroid
    3    4   88    0 |    c = primary_hypothyroid
    2    0    0    0 |    d = secondary_hypothyroid
```

Classifier

Choose   J48 -C 0.25 -M 2

Test options

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

More options...

(Nom) Class

Start   Stop

Result list (right-click for options)

14:01:42 - bayes.BayesNet
14:06:12 - trees.J48

Classifier output

```
|   |   |   | thyroid surgery = t: negative (6.74)
|   |   |   TSH measured = f: negative (30.75)
|   |   on thyroxine = t: negative (56.17)

Number of Leaves  :     15

Size of the tree :      29


Time taken to build model: 0.08 seconds

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

Correctly Classified Instances        3756               99.5758 %
Incorrectly Classified Instances        16                0.4242 %
Kappa statistic                          0.9707
Mean absolute error                      0.003
Root mean squared error                  0.0414
Relative absolute error                  4.1612 %
Root relative squared error             21.7445 %
Total Number of Instances             3772

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
                 0.999    0.021    0.998      0.999    0.998      0.979   0.993     0.999     negative
                 0.985    0.002    0.970      0.985    0.977      0.976   0.999     0.964     compensated_hypothyroid
                 0.937    0.001    0.957      0.937    0.947      0.946   1.000     0.988     primary_hypothyroid
                 0.000    0.000    ?          0.000    ?          ?       0.197     0.000     secondary_hypothyroid
Weighted Avg.    0.996    0.019    ?          0.996    ?          ?       0.993     0.996

=== Confusion Matrix ===

    a    b    c    d    <-- classified as
 3476    3    2    0 |   a = negative
    1  191    2    0 |   b = compensated_hypothyroid
    3    3   89    0 |   c = primary_hypothyroid
    2    0    0    0 |   d = secondary_hypothyroid
```

Appendix

1	This intentially left blank
2	This intentially left blank
3	This intentially left blank
4	

5

```
******************************************* K2 *******************************************

=== 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:     BayesNetwork_Weather_data_Set
Instances:    999
Attributes:   5
              Wind
              Outlook
              Temp
              Humidity
              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)
Wind(2): Play
Outlook(3): Play
Temp(3): Play
Humidity(2): Play
Play(2):
```

```
LogScore Bayes: -3918.637812829568
LogScore BDeu: -3933.033446895659
LogScore MDL: -3932.9298646223424
LogScore ENTROPY: -3888.035958561126
LogScore AIC: -3901.035958561126


Time taken to build model: 0.02 seconds

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

Correctly Classified Instances         666                  66.6667 %
Incorrectly Classified Instances       333                  33.3333 %
Kappa statistic                          0.2035
Mean absolute error                      0.424
Root mean squared error                  0.4604
Relative absolute error                 92.0771 %
Root relative squared error             95.9464 %
Total Number of Instances              999

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area  PRC Area  Class
                 0.856    0.671    0.695      0.856   0.767      0.218    0.662     0.731     N
                 0.329    0.144    0.562      0.329   0.415      0.218    0.662     0.505     Y
Weighted Avg.    0.667    0.482    0.647      0.667   0.640      0.218    0.662     0.650

=== Confusion Matrix ===

   a    b   <-- classified as
 548   92 |   a = N
 241  118 |   b = Y

K2 w/Markov Blanket
=== Run information ===

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

```
Relation:        BayesNetwork_Weather_data_Set
Instances:    999
Attributes:   5
              Wind
              Outlook
              Temp
              Humidity
              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)
Wind(2): Play
Outlook(3): Play
Temp(3): Play
Humidity(2): Play
Play(2):
LogScore Bayes: -3918.637812829568
LogScore BDeu: -3933.033446895659
LogScore MDL: -3932.9298646223424
LogScore ENTROPY: -3888.035958561126
LogScore AIC: -3901.035958561126



Time taken to build model: 0 seconds

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

Correctly Classified Instances          666               66.6667 %
Incorrectly Classified Instances        333               33.3333 %
Kappa statistic                           0.2035
Mean absolute error                       0.424
Root mean squared error                   0.4604
Relative absolute error                  92.0771 %
```

```
Root relative squared error           95.9464 %
Total Number of Instances             999


=== Detailed Accuracy By Class ===

                TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
                0.856    0.671    0.695      0.856   0.767      0.218   0.662     0.731     N
                0.329    0.144    0.562      0.329   0.415      0.218   0.662     0.505     Y
Weighted Avg.   0.667    0.482    0.647      0.667   0.640      0.218   0.662     0.650

=== Confusion Matrix ===

   a    b    <-- classified as
 548   92 |   a = N
 241  118 |   b = Y
```

```
******************************* RepeatedHillClimber ********************************

=== Run information ===

Scheme:       weka.classifiers.bayes.BayesNet -D -Q
weka.classifiers.bayes.net.search.local.RepeatedHillClimber -- -U 10 -A 1 -P 1 -S BAYES -E
weka.classifiers.bayes.net.estimate.SimpleEstimator -- -A 0.5
Relation:     BayesNetwork_Weather_data_Set
Instances:    999
Attributes:   5
              Wind
              Outlook
              Temp
              Humidity
              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)
Wind(2): Play
Outlook(3): Play
Temp(3):
Humidity(2): Temp
Play(2):
LogScore Bayes: -3769.942927675236
LogScore BDeu: -3782.1690785947935
LogScore MDL: -3783.088074958381
LogScore ENTROPY: -3741.64754628649
LogScore AIC: -3753.64754628649


Time taken to build model: 0.01 seconds

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

Correctly Classified Instances         675               67.5676 %
Incorrectly Classified Instances       324               32.4324 %
Kappa statistic                          0.2418
Mean absolute error                      0.427
Root mean squared error                  0.4615
Relative absolute error                 92.7156 %
Root relative squared error             96.188  %
Total Number of Instances              999

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall  F-Measure  MCC     ROC Area  PRC Area  Class
                 0.838    0.613    0.709      0.838   0.768      0.251   0.650     0.718     N
                 0.387    0.163    0.572      0.387   0.462      0.251   0.650     0.493     Y
Weighted Avg.    0.676    0.451    0.660      0.676   0.658      0.251   0.650     0.637
```

```
=== Confusion Matrix ===

   a   b   <-- classified as
 536 104 |   a = N
 220 139 |   b = Y


=== 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:      hypothyroid
Instances:     3772
Attributes:    30
               age
               sex
               on thyroxine
               query on thyroxine
               on antithyroid medication
               sick
               pregnant
               thyroid surgery
               I131 treatment
               query hypothyroid
               query hyperthyroid
               lithium
               goitre
               tumor
               hypopituitary
               psych
               TSH measured
               TSH
               T3 measured
               T3
               TT4 measured
               TT4
```

```
                T4U measured
                T4U
                FTI measured
                FTI
                TBG measured
                TBG
                referral source
                Class
Test mode:      10-fold cross-validation

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

Bayes Network Classifier
not using ADTree
#attributes=30 #classindex=29
Network structure (nodes followed by parents)
age(1): Class
sex(2): Class
on thyroxine(2): Class
query on thyroxine(2): Class
on antithyroid medication(2): Class
sick(2): Class
pregnant(2): Class
thyroid surgery(2): Class
I131 treatment(2): Class
query hypothyroid(2): Class
query hyperthyroid(2): Class
lithium(2): Class
goitre(2): Class
tumor(2): Class
hypopituitary(2): Class
psych(2): Class
TSH measured(2): Class
TSH(4): Class
T3 measured(2): Class
T3(3): Class
TT4 measured(2): Class
TT4(6): Class
T4U measured(2): Class
```

```
T4U(1): Class
FTI measured(2): Class
FTI(5): Class
TBG measured(1): Class
TBG(1): Class
referral source(5): Class
Class(4):
LogScore Bayes: -29286.076203253557
LogScore BDeu: -29611.200986222455
LogScore MDL: -29816.503686177814
LogScore ENTROPY: -29178.26323628693
LogScore AIC: -29333.26323628693


Time taken to build model: 0.08 seconds

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

Correctly Classified Instances        3719               98.5949 %
Incorrectly Classified Instances        53                1.4051 %
Kappa statistic                          0.9028
Mean absolute error                      0.011
Root mean squared error                  0.075
Relative absolute error                 15.135  %
Root relative squared error             39.3888 %
Total Number of Instances             3772


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

	TP Rate	FP Rate	Precision	Recall	F-Measure	MCC	ROC Area	PRC Area	Class
	0.993	0.086	0.993	0.993	0.993	0.910	0.997	1.000	negative
	0.892	0.004	0.930	0.892	0.911	0.906	0.999	0.975	compensated_hypothyroid
	0.926	0.004	0.854	0.926	0.889	0.887	0.996	0.921	primary_hypothyroid
	0.000	0.000	?	0.000	?	?	0.749	0.002	secondary_hypothyroid
Weighted Avg.	0.986	0.080	?	0.986	?	?	0.997	0.996	

```
=== Confusion Matrix ===

    a    b    c    d   <-- classified as
 3458    9   14    0 |    a = negative
   20  173    1    0 |    b = compensated_hypothyroid
    3    4   88    0 |    c = primary_hypothyroid
    2    0    0    0 |    d = secondary_hypothyroid


*********************************************** J48 ***********************************************
=== Run information ===

Scheme:       weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:     hypothyroid
Instances:    3772
Attributes:   30
              age
              sex
              on thyroxine
              query on thyroxine
              on antithyroid medication
              sick
              pregnant
              thyroid surgery
              I131 treatment
              query hypothyroid
              query hyperthyroid
              lithium
              goitre
              tumor
              hypopituitary
              psych
              TSH measured
              TSH
              T3 measured
              T3
              TT4 measured
              TT4
              T4U measured
```

```
              T4U
              FTI measured
              FTI
              TBG measured
              TBG
              referral source
              Class
Test mode:    10-fold cross-validation

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

J48 pruned tree
------------------

TSH <= 6: negative (3366.31/2.0)
TSH > 6
|   FTI <= 64
|   |   TSH measured = t
|   |   |   T4U measured = t
|   |   |   |   thyroid surgery = f
|   |   |   |   |   T3 <= 2.3: primary_hypothyroid (82.7)
|   |   |   |   |   T3 > 2.3
|   |   |   |   |   |   TSH <= 15: negative (2.06/0.06)
|   |   |   |   |   |   TSH > 15: primary_hypothyroid (3.24)
|   |   |   |   thyroid surgery = t
|   |   |   |   |   TT4 <= 49: negative (3.0)
|   |   |   |   |   TT4 > 49: primary_hypothyroid (2.0)
|   |   |   T4U measured = f: compensated_hypothyroid (7.08/2.62)
|   |   TSH measured = f: negative (6.24)
|   FTI > 64
|   |   on thyroxine = f
|   |   |   TSH measured = t
|   |   |   |   thyroid surgery = f
|   |   |   |   |   TT4 <= 150
|   |   |   |   |   |   TT4 <= 48
|   |   |   |   |   |   |   T4U measured = t: negative (2.0/1.0)
|   |   |   |   |   |   |   T4U measured = f: primary_hypothyroid (3.04/0.04)
|   |   |   |   |   |   TT4 > 48: compensated_hypothyroid (191.5/3.06)
|   |   |   |   |   TT4 > 150: negative (9.16/0.16)
```

```
|   |   |   |   thyroid surgery = t: negative (6.74)
|   |   |   TSH measured = f: negative (30.75)
|   |   on thyroxine = t: negative (56.17)

Number of Leaves  :      15

Size of the tree :      29


Time taken to build model: 0.08 seconds

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

Correctly Classified Instances        3756               99.5758 %
Incorrectly Classified Instances        16                0.4242 %
Kappa statistic                          0.9707
Mean absolute error                      0.003
Root mean squared error                  0.0414
Relative absolute error                  4.1612 %
Root relative squared error             21.7445 %
Total Number of Instances             3772

=== Detailed Accuracy By Class ===

        TP Rate  FP Rate  Precision  Recall   F-Measure  MCC     ROC Area  PRC Area  Class
        0.999    0.021    0.998      0.999    0.998      0.979   0.993     0.999     negative
        0.985    0.002    0.970      0.985    0.977      0.976   0.999     0.964     compensated_hypothyroid
        0.937    0.001    0.957      0.937    0.947      0.946   1.000     0.988     primary_hypothyroid
        0.000    0.000    ?          0.000    ?          ?       0.197     0.000     secondary_hypothyroid
WAvg    0.996    0.019    ?          0.996    ?          ?       0.993     0.996

=== Confusion Matrix ===

     a     b    c    d   <-- classified as
  3476     3    2    0 |    a = negative
     1   191    2    0 |    b = compensated_hypothyroid
     3     3   89    0 |    c = primary_hypothyroid
     2     0    0    0 |    d = secondary_hypothyroid
```