User Input on CLIPS and Bayesian on Weka

Lecture #5 – Expert System
Department of Computer Science
Bogor Agricultural University (IPB)
2013

Input dari user

- Untuk membaca inputan dari user gunakan perintah (read)
- Untuk membaca banyak inputan dari user gunakan perintah (readline)

Contoh read

```
CLIPS>
(defrule read-input
=>
(printout t "Name a primary color" crlf)
(assert (color (read))))
CLIPS>
(defrule check-input
?color <- (color ?color-read&red|yellow|blue)
=>
(retract ?color)
(printout t "Correct" crlf))
```

Contoh readline

```
CLIPS>
(defrule test-readline
=>
(printout t "Enter input" crlf)
(bind ?string (readline))
(assert-string (str-cat "(" ?string ")")))
CLIPS> (reset)
CLIPS> (run)
```

Practice of Bayesian Networks

Bayesian Networks. Two tasks

- Infer the structure of the network from the data (in practice, the structure of the network is identified by data experts, not by machine)
- Fill in conditional probabilities tables

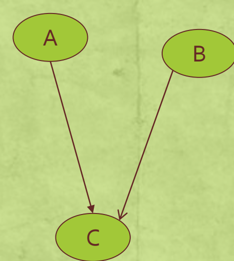
The elementary inference types in Bayes nets

Causal

Increased probability of A makes B more likely.

A can cause B

Intercausal



A and B can each cause C. B explains C and so is evidence against A

Diagnostic

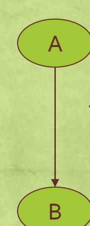
Increased probability of B makes A more likely.

B is evidence for A

Bayesian Net for Weather Data

Outlook	Temp.	Humidity	Windy	Play
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No

Diagnostic



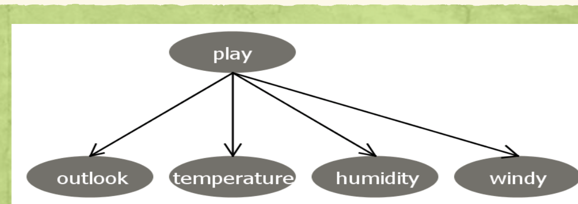
Increased probability of B makes A more likely.

B is evidence for A

Naïve Bayes

Assuming the attributes are independent of each other, we have a Naïve Bayesian Network:

Outlook	Temp.	Humidity	Windy	Play
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No



P(play=yes)=9/14,

with Laplace correction:
P(play=yes)=9+1/14+2=0.625

In general, to make Laplace correction, we add an initial count (1) to the total of all instances with a given attribute value, and we add the number of distinct values of the same attribute to the total number of instances in the group.

Naïve Bayes

And to fill the Conditional Probability Tables we compute conditional probabilities for each node in form: Pr (attribute=value | parents values) for each combinations of attributes values in parent nodes

Outlook	Temp.	Humidity	Windy	Play
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No

P(outlook=sunny|play=yes)
=(2+1)/(9+3)=3/12

P(outlook=rainy|play=yes)
=(3+1)/(9+3)=4/12

P(outlook=overcast|play=yes)
=(4+1)/(9+3)=5/12

Sum is1

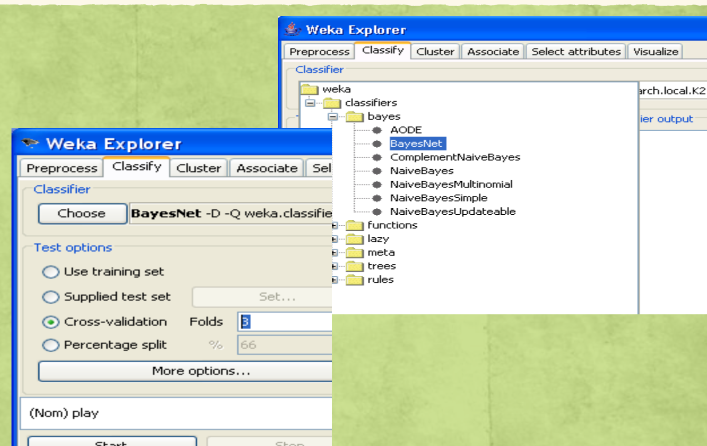
P(outlook=sunny|play=yes)
=(2+1)/(9+3)=3/12

P(outlook=sunny|play=no)
=(3+1)/(5+3)=4/8

Sum is NOT 1

WEKA Exercise 1. Bayesian network for weather data with default parameters

- **Preprocess** tab
  - ***Open file*** *weather.nominal.arff*
  - Perform ***filters*** (if needed): discretize or replace missing values
- **Classify** tab
  - ***Classifier->choose->classifiers->bayes->BayesNet***
  - Click on row with selected classifier, change Laplace correction (initial count) to 1 (instead of 0.5) in the ***option row for the estimator***
  - ***Cross-validation*** change to ***3 folds*** (since we have only 14 instances, with 10 folds cross validation we will have test groups of size less than 2, which makes the classifier less reliable). Press ***Start***

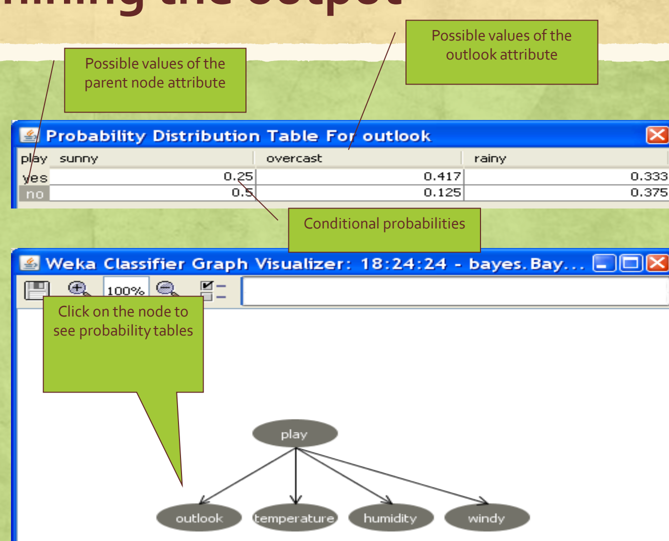


WEKA Exercise 2. Examining the output

- In the history box, right-click and choose ***visualize graph***. Check that probabilities in CPT correspond to what we calculated before (clicking on the graph node brings table of conditional probabilities)
- Naïve Bayes? Study parameters of the program. Click on ***choose*** line again.
- Save this model in file weather.xml for later use
- Click on ***searchAlgorithm*** row. Default parameters are:

  initAsNaiveBayes=true

  maxNrOfParents=1
- Change ***maxNrOfParents=2***. Run. Visualize graph
- Change to ***initAsNaiveBayes=false***. Run. Visualize graph. Change back to true.



Conditional Probabilities Tables in WEKA

- After the structure is learned, the CPT for each node are computed.
- Simple estimator computes the relative frequencies of the associated combinations of the attribute values in the training data (just like we do in our excercises).

How it was computed

Probability Distribution Table For outlook

play	sunny	overcast	rainy
yes	0.25	0.417	0.333
no	0.5	0.125	0.375

Outlook	Temp.	Humidity	Windy	Play
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No

Total number of different values for outlook

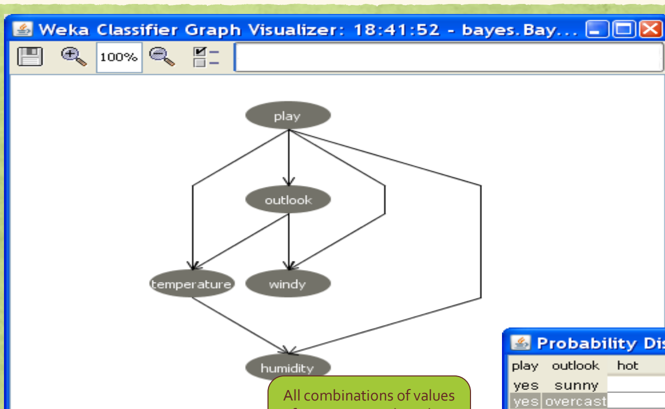
P(outlook=sunny|play=yes)
=(2+1)/(9+3)=3/12=1/4=0.25

Number of instances with play=yes

Initial count for attribute value=sunny

Number of instances with outlook=sunny and play=yes

More complex Bayesian Network for weather data (with ***maxNrOfParents=2)***



Possible values of the attribute temperature

Probability Distribution Table For temperature

play	outlook	hot	mild	cool
yes	sunny	0.2	0.4	0.4
yes	overcast	0.429	0.286	0.286
yes	rainy	0.167	0.5	0.333
no	sunny	0.5	0.333	0.167
no	overcast	0.333	0.333	0.333
no	rainy	0.2	0.4	0.4

All combinations of values for 2 parent nodes: play and outlook

Conditional probabilities

How it was computed

Probability Distribution Table For temperature

play	outlook	hot	mild	cool
yes	sunny	0.2	0.4	0.4
yes	overcast	0.429	0.286	0.286
yes	rainy	0.167	0.5	0.333
no	sunny	0.5	0.333	0.167
no	overcast	0.333	0.333	0.333
no	rainy	0.2	0.4	0.4

Outlook	Temp.	Humidity	Windy	Play
Sunny	Hot	High	False	No
Sunny	Hot	High	True	No
Overcast	Hot	High	False	Yes
Rainy	Mild	High	False	Yes
Rainy	Cool	Normal	False	Yes
Rainy	Cool	Normal	True	No
Overcast	Cool	Normal	True	Yes
Sunny	Mild	High	False	No
Sunny	Cool	Normal	False	Yes
Rainy	Mild	Normal	False	Yes
Sunny	Mild	Normal	True	Yes
Overcast	Mild	High	True	Yes
Overcast	Hot	Normal	False	Yes
Rainy	Mild	High	True	No

P(temperature=hot|play=yes,outlook=sunny) =(0+1)/(2+3)=1/5=0.2

Number of instances with play=yes and outlook=sunny

Number of instances with temperature=hot, outlook=sunny and play=yes

How WEKA infers a structure of the network

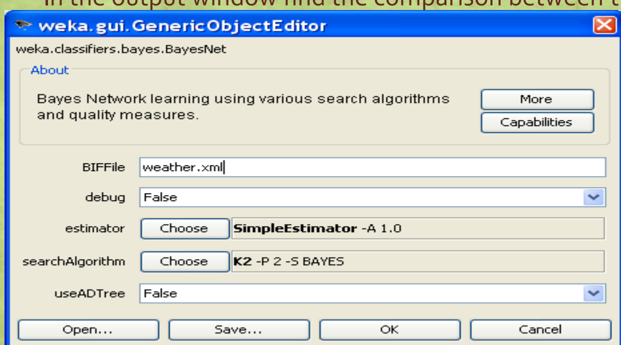
- The nodes correspond to the attributes
- Learning the structure is to find edges
  - Searching through the possible edges sets
  - For each set estimate the conditional probability tables from the data
  - Estimate quality of the network as the probability of obtaining the set of data given this network

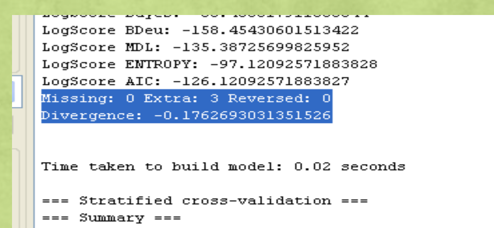
WEKA Search Algorithms. Example

- By default: K2.
- Starts with a given ordering of attributes.
- Adds one node in order and considers adding edges from each previously added node to a new node.
- Then it adds the edge which maximizes the network score.
- The number of parents is restricted to a predefined maximum.
- *The Markov blanket* of a node includes all its parents, children and children parents. It is proven, that a given node is conditionally dependent only on nodes in its Markov blanket. So the edge is added from the class node to the node which is not in its Markov blanket. Otherwise the value of this attribute would be irrelevant for the class.

WEKA Exercise 3. Improving the network supplied as a file

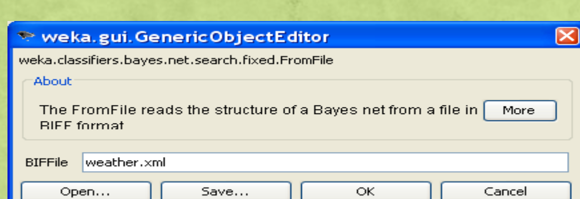
- Bring in the window of ***classifier's options***
  - Type in the ***BIFF file*** box: *weather.xml*
- ***Run***
- In the output window find the comparison between two networks:





WEKA Exercise 4. Structure supplied by the user

- Bring in the window of ***classifier's options***
- In ***searchAlgorithm*** row press ***Choose*** button
- Choose search->fixed->FromFile. OK
- Press ***searchAlgorithm*** row to define parameters
- Type in the ***BIFF file*** box: *weather.xml (Do NOT use the button Open…)*
- ***Run***
- Check that WEKA has produced the Naïve Bayes, as it was supplied in your file



LKP 5

- Suppose you are working for a financial institution and you are asked to build a fraud detection system. You plan to use the following information:

When the card holder is traveling abroad, fraudulent transaction are more likely since tourists are prime targets for thieves. More precisely, 1% of transactions are fraudulent when the card holder is traveling, whereas only 0.2% of the transactions are fraudulent when he is not traveling. On average, 5% of all transactions happen while card holder is traveling. If a transaction is fraudulent, then the likelihood of a foreign purchase increases, unless the card holder happens to be traveling. More precisely, when the card holder is not traveling, 10% of the fraudulent transactions are foreign purchases, whereas only 1% of the legitimate transactions are foreign purchases. On the other hand, when the card holder is traveling, 90% of the transactions are foreign purchases regardless of the legitimacy of the transactions.

1) Build a Bayesian Network Model for the system, based on the given information above and write also the conditional probability for each node within.

2) System has detected the foreign purchase. What is the probability of a fraud if we don't know whether the card holder is traveling or not?