

Errors in the course handouts

Errors in the handouts will be noted here so you can correct your copies. The versions on the web will be corrected by the time the errors are recorded here. Please let me know if you spot any further errors (dfg@doc.ic.ac.uk).

Coursework handout:

Task 1.4: 4 marks

**** $P(A|B) = P(A \& B) / P(B)$. ****

Task 1.5: 6 marks (harder)

Reworded:

Finally complete the function Query which calculates the probability distribution over the root node of a naive Bayesian network. To represent a naive network in Python we will use a list containing an entry for each node (in numeric order) giving the associated probability table: [prior, cpt1, cpt2, cpt3, cpt4, cpt5]. You can calculate the prior of the root and the conditional probability tables between each child variable and the root using your solutions to Tasks 1 and 2. A query is a list of the instantiated states of the child nodes, for example [1,0,3,2,0]. The returned value is a list (or vector) giving the posterior probability distribution over the states of the root node, for example [0.1,0.3,0.4,0.2].

Results File (Coursework 1)

- The prior probability distribution of ****node 0**** in the data set
- The results of queries ****[4,0,0,0,5] and [6, 5, 2, 5, 5]**** on the naive network

IDAPICourseworkSkeleton.py

The function DependencyMatrix needs a third parameter:

```
def DependencyMatrix(theData, noVariables, noStates):
```

Results File (Coursework 3)

4. The MDLScore of the your network for Hepatitis C data set

Task 3.6: 4 marks (hard)

Write a function to find the **best** scoring network formed by deleting one arc from the spanning tree.