University of Pittsburgh
CS 2750 Machine Learning
Professor Milos Hauskrecht

Homework 7
Apr 9, 2007

Solutions to Problem set 7

1 Problem 1

(a) **A** and **B** are independent not given **D**;

A and **C** are independent not given **D**;

B and **C** are independent not given **D**;

A and **F** are independent given **D**;

B and **F** are independent given **D**;

C and **F** are independent given **D**;

D and **E** are independent given **C**;

(b) The number of parameters to the full joint distribution:

$2 \times 2 \times 2 \times 3 \times 2 \times 2 - 1 = 96 - 1 = 95$

(c)

$$\begin{aligned}
&P(A = T, B = T, C = F, D = F, E = F, F = T) = \\
&\quad = \quad P(A = T) \cdot P(B = T) \cdot P(C = T) \cdot P(D = F|A = T, B = T, C = F) \cdot P(E = F|C = F) \\
&\qquad\quad \cdot P(F = T|D = F)
\end{aligned}$$

(d) Number of parameters to define the BBN: $1 + 1 + 1 + 16 + 2 + 3 = 24$

(e) **B** and **E** are independent if we do not know **D** and **F**.. Thus

$$\begin{aligned}
P(B = T, E = T) &= P(B = T) \times P(E = T) \\
&= P(B = T) \times \sum_C P(E = T, C) \\
&= P(B = T) \times \sum_C P(E = T|C) \times P(C)
\end{aligned}$$

2 Problem 2

(a)

Pneumonia=T	**Pneumonia**=F
0.02	0.98

Pneumonia	**Fever**=T	**Fever**=F
T	0.9	0.1
F	0.6	0.4

Pneumonia	**Paleness**=T	**Paleness**=F
T	0.7	0.3
F	0.5	0.5

Pneumonia	**Cough**=T	**Cough**=F
T	0.9	0.1
F	0.1	0.9

Pneumonia	**HWBC**=T	**HWBC**=F
T	0.8	0.2
F	0.5	0.5

(b)

$$
\begin{aligned}
&P(Pn = T|F = T, P = F, C = T, H = F) \\
&\quad = \frac{P(Pn = T, F = T, P = F, C = T, H = F)}{P(F = T, P = F, C = T, H = F)} \\
&\quad = \frac{P(Pn = T, F = T, P = F, C = T, H = F)}{P(Pn = T, F = T, P = F, C = T, H = F) + P(Pn = F, F = T, P = F, C = T, H = F)} \\
&\quad = \frac{P(Pn = T) \cdot P(F = T|Pn = T) \cdot P(P = F|Pn = T) \cdot P(C = T|Pn = T) \cdot P(H = f|Pn = T)}{\sum_{Pn} P(Pn) \cdot P(F = T|Pn) \cdot P(P = F|Pn) \cdot P(C = T|Pn) \cdot P(H = f|Pn)} \\
&\quad = \frac{0.02 \times 0.9 \times 0.3 \times 0.9 \times 0.2}{0.02 \times 0.9 \times 0.3 \times 0.9 \times 0.2 + 0.98 \times 0.6 \times 0.5 \times 0.1 \times 0.5} \\
&\quad = 0.062
\end{aligned}
$$

(c)

$$
\begin{aligned}
&P(Pn = T|F = T, C = T) \\
&\quad = \frac{P(Pn = T, F = T, C = T)}{P(F = T, C = T)} \\
&\quad = \frac{P(Pn = T, F = T, C = T)}{P(Pn = T, F = T, C = T, H = F) + P(Pn = F, F = T, C = T)} \\
&\quad = \frac{P(Pn = T) \cdot P(F = T|Pn = T) \cdot P(C = T|Pn = T)}{\sum_{Pn} P(Pn) \cdot P(F = T|Pn) \cdot P(C = T|Pn)} \\
&\quad = \frac{0.02 \times 0.9 \times 0.9}{0.02 \times 0.9 \times 0.9 + 0.98 \times 0.6 \times 0.1} = 0.216
\end{aligned}
$$

Note, that if some of the symptoms (e.g. P) in the Naive Bayes model are missing, the terms corresponding to them (e.g. $P(P|Pn)$) are not used in the formula.

(d)

Symptoms	**P(Pn=T | Symptoms)**
1 0 1 0	0.062
1 0 1 -1	0.1419
0 1 -1 0	0.0028