

CIS 520, Machine Learning, Fall 2021

Homework 9

Due: Tuesday, November 16th, 11:59pm

Submit to Gradescope

Yixuan Meng, Zhouyang Fang

Problem 1

1. $p(X_1, X_2, X_3, X_4, X_5, X_6) = p(X_1)p(X_2)p(X_3 | X_1)p(X_4 | X_1, X_2)p(X_5 | X_3, X_4)p(X_6 | X_4)$
2. No. $p(X_3 | X_1)p(X_4 | X_1, X_2)p(X_5 | X_3, X_4)p(X_6 | X_4) \neq p(X_3)p(X_4)p(X_5 | X_3)p(X_6 | X_3)$. So this distribution is not included in the Bayesian network above.

3. Will be smaller.

Removing the edge means cutting one possible way for the information flow. Before removing it represent the model that direct causality exists or doesn't exist between X_3 and X_5 . However, after the removing, it can only represent that direct causality doesn't exist between X_3 and X_5 .

4. (a) True.

Both $X_1 \rightarrow X_4 \leftarrow X_2$ and $X_3 \rightarrow X_5 \leftarrow X_4$ are not active path. So they are independent.

- (b) True.

Both $X_1 \rightarrow X_4 \rightarrow X_2$ and $X_5 \leftarrow X_4 \rightarrow X_6$ are not active path. So they are conditionally independent given X_4 .

- (c) False.

X_6 is the descendant of X_4 . According to the last rules of active trails, $X_{i-1} \rightarrow X_i \leftarrow X_{i+1}$ and X_i or its descendant is observed, $X_1 \rightarrow X_4 \leftarrow X_2$ is an active trail. So not conditionally independent given X_6 .

- (d) False.

$X_1 \rightarrow X_4 \leftarrow X_2$ is an active trail given X_4 . Also $X_1 \rightarrow X_3 \rightarrow X_5$ is an active path. So they are not conditionally independent given X_4 .

Problem 2

1. Nothing to Report. Please skip.
2. Do you add the link(yes/no)? **No**
Steps:

B	P(B)%
T	70.6
F	29.4

A	B	$P(B A)\%$
T	T	72.7
	F	27.3
F	T	66.7
	F	33.3

$|P(B) - P(B | A)| = 0.02 < 0.05, |P(B) - P(B | \sim A)| = 0.04 < 0.05$. So independent. No link.



3. Do you add the link(yes/no)? **Yes**
Steps:

A	$P(C A)\%$
T	55
F	67

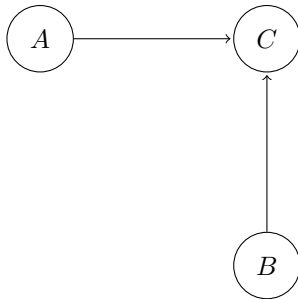
$P(C) = 0.59 \implies |P(C) - P(C | \sim A)| = 0.08$. Dependent. Link needed.



4. Do you add the link(yes/no)? **Yes**
Steps:

A	B	$P(C A, B)\%$
T	T	50
	F	67
F	T	75
	F	50

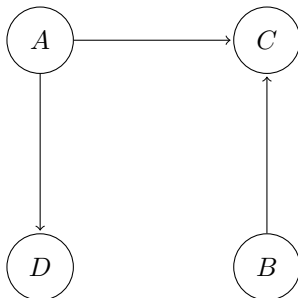
$|P(C | A) - P(C | A, \sim B)| = 0.12$. Dependent. Link needed.



5. Do you add the link(yes/no)? **Yes**
Steps:

$$P(D) = 0.41 \quad \left| \begin{array}{c|c} A & P(D|A)\% \\ \hline T & 45 \\ F & 33 \end{array} \right|$$

$\Rightarrow |P(D) - P(D | A)| = 8\%$. Dependent. Link needed.

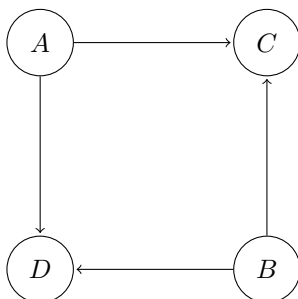


6. Do you add the link(yes/no)? **Yes**
Steps:

$$P(D | A) = 45\%$$

A	B	$P(D A, B)\%$
T	T	38
	F	67
F	T	50
	F	25

$\Rightarrow |P(D | A) - P(D | A, B)| = 7\%$ Dependent. Link needed.

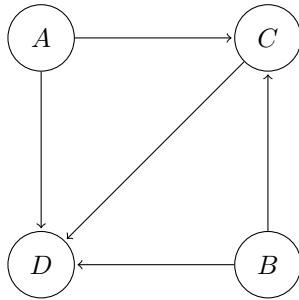


7. Do you add the link(yes/no)? **Yes**
Steps:

$$P(D | A, B) = 0.375$$

A	B	C	$P(D A, B, C)\%$
T	T	T	75
		F	0
	F	T	100
		F	0
F	T	T	33
		F	0
	F	T	100
		F	0

$\Rightarrow |P(D | A, B) - P(D | A, B, C)| = 0.375$. Dependent. Link needed.



8. The final Bayes net is as follow:

