CS 486/686 Spring 2023
Assignment 2 Solutions

June 20, 2023

Question 2a

Acc	$Pr(Acc)$
t	0.8
f	0.2

$Trav$	$Pr(Trav)$
t	0.05
f	0.95

$Trav$	$Fraud$	$Pr(Fraud\|Trav)$
t	t	0.01
t	f	0.99
f	t	0.004
f	f	0.996

$Trav$	$Fraud$	FP	$Pr(FP\|Trav, Fraud)$
t	t	t	0.9
t	t	f	0.1
t	f	t	0.9
t	f	f	0.1
f	t	t	0.1
f	t	f	0.9
f	f	t	0.01
f	f	f	0.99

Acc	$Fraud$	OP	$Pr(OP\|Acc, Fraud)$
t	t	t	0.8
t	t	f	0.2
t	f	t	0.6
t	f	f	0.4
f	t	t	0.3
f	t	f	0.7
f	f	t	0.1
f	f	f	0.9

Acc	PT	$Pr(PT\|Acc)$
t	t	0.1
t	f	0.9
t	t	0.01
t	f	0.99

Question 2b

Part I - Prior Probability

$$\begin{aligned} Pr(fraud) &= Pr(fraud|trav)Pr(trav) + Pr(fraud|\neg trav)Pr(\neg trav) \\ &= 0.01 * 0.05 + 0.004 * 0.95 \\ &= 0.0043 \end{aligned}$$

Part II - Posterior Probability

OC	$f_1(Acc) = Pr(Acc)$
t	0.8
f	0.2

$Fraud$	$Trav$	$f_2(Fraud, Trav) = Pr(Fraud\|Trav)$
t	t	0.01
t	f	0.004
f	t	0.99
f	f	0.996

$Trav$	$f_3(Trav) = Pr(Trav)$
t	0.05
f	0.95

OC	$f_4(Acc) = Pr(PT = t\|Acc)$
t	0.1
f	0.01

OC	$Fraud$	$f_5(Acc, Fraud) = Pr(OP = f \mid Acc, Fraud)$
t	t	0.2
t	f	0.4
f	t	0.7
f	f	0.9

$Trav$	$Fraud$	$f_6(Trav, Fraud) = Pr(FP = t \mid Trav, Fraud)$
t	t	0.9
t	f	0.9
f	t	0.1
f	f	0.01

Eliminate variable $Trav$:

$$f_7(Fraud) = sumout_{Trav}[f_2(Fraud, Trav) * f_3(Trav) * f_6(Trav, Fraud)]$$

$Fraud$	$f_7(Fraud)$
t	0.00083
f	0.054012

Eliminate variable Acc:

$$f_8(Fraud) = sumout_{Acc}[f_1(Acc) * f_4(Acc) * f_5(Acc, Fraud)]$$

$Fraud$	$f_8(Fraud)$
t	0.0174
f	0.0338

$$Pr(Fraud = t \mid FP = t, OP = f, PT = t) = k * f_7(Fraud = t) * f_8(Fraud = t) = 0.00784871$$

where k is a normalizing constant:

$$k = \frac{1}{f_7(Fraud = t) * f_8(Fraud = t) + f_7(Fraud = f) * f_8(Fraud = f)}$$

Question 2c

Acc	$f_1(Acc) = Pr(Acc)$
t	0.8
f	0.2

$Fraud$	$f_2(Fraud) = Pr(Fraud \vert Trav = t)$
t	0.01
f	0.99

$Trav$	$f_3() = Pr(Trav = t)$
	0.05

Acc	$f_4(Acc) = Pr(PT = t \vert Acc)$
t	0.1
f	0.01

Acc	$Fraud$	$f_5(Acc, Fraud) = Pr(OP = f \vert Acc, Fraud)$
t	t	0.2
t	f	0.4
f	t	0.7
f	f	0.9

$Fraud$	$f_6(Fraud) = Pr(FP = t \vert Trav = t, Fraud)$
t	0.9
f	0.9

Eliminate variable Acc:

$$f_7(Fraud) = sumout_{Acc}[f_1(Acc) * f_4(Acc) * f_5(Acc, Fraud)]$$

$Fraud$	$f_7(Fraud)$
t	0.0174
f	0.0338

$Pr(Fraud = t|FP = t, OP = f, PT = t, Trav = t)$
$= k * f_2(Fraud = t) * f_3() * f_6(Fraud = t) * f_7(Fraud = t) = 0.00517303$

where k is a normalizing constant:

$$k = \frac{1}{\sum_{Fraud} f_2(Fraud) * f_3() * f_6(Fraud) * f_7(Fraud)}$$

Question 2d

When an online purchase is made, the fraud detection system is likely to believe that the transaction is fraudulent unless it has reasons to believe that the card holder already has an account with the merchant. Therefore, an ingenious thief could simply make a transaction with the same merchant in the previous week to fool the fraud detection system into believing the card holder already has account with the merchant. After that, the thief can make the intended online purchase with a reduced risk of being rejected.

One can verify that the probability of a fraudulent transaction decreases when a previous transaction with the same merchant is observed in the prior week since $Pr(Fraud = t|OP = t) = 0.00600966$ whereas $Pr(Fraud = t|OP = t, PT = t) = 0.00575465$.

Question 3a

Each question 2 points, 1 for correct answer, 1 for justification.

i Dependent. Apply rule 1. Path open: $D \to G$, $D \to F \to G$.

ii Dependent. Apply rule 1. Path open: $D \to G$.

iii Independent. Apply rule 3. Path blocked: $A \to B \leftarrow\sim G$

iv Dependent. Apply rule 3. Path open: $A \to B \leftarrow\sim G$

v Independent. Apply rule 1. Path blocked: $A \sim\leftarrow C \leftarrow\sim G$. Apply rule 2 or rule 3. Path blocked: $A \sim\leftarrow C \to\sim G$. Path blocked: $A \sim\leftarrow E \to\sim G$.

vi Independent. Apply rule 2. Path blocked: $A \sim\leftarrow D \to\sim G$. Apply rule 3. Path blocked: $A \sim\to E \leftarrow\sim G$

vii Dependent. Apply rule 3. Path open: $A \sim\to B \leftarrow\sim\to E \leftarrow\sim G$

Question 3b

2 points will be subtracted if one relevant variable is wrong. 1 points will be subtracted if justification is wrong.

- C is relevant because it is query.
- D is relevant because it is a parent of relevant variable(C).
- E is relevant because it is in evidence and a descendent of relevant variable(C,D)
- F is relevant because it is a parent of relevant variable(E).