Portfolio Tasks

For AIML429

Document compiled on: March 26, 2024

Preface

These tasks. There are things to say about them.
The due date for this first set (Part I) is Friday 19th April (end of week 6).

Weighting of questions... some are obviously going to be easier than others. Does that matter? I like the simplicity of "flat" marking - all worth the same.

Contents

The Tasks (Part I)

1: Prosecutor

♠ Do you think "Meadows' Law" is an instance of the "Prosecutor's Fallacy"? Why?

2: Monty Hall favours one door

♠ You're a contestant on Monty Hall's game show, so you expect to opt for switching. When it's Monty's turn and he has a choice to open doors 2 or 3, suppose he doesn't decide evenly between them. Instead, 75 percent of the time he opts for door 2, and you know this. Let's say you've chosen door 1, and Monty has revealed a goat sitting behind door 2. What is the probability that switching wins?

3: prey-pet-temp

♠ At https://ecs.wgtn.ac.nz/Courses/AIML429_2024T1/Assignments you'll find a data set called "prey-pet-temp.csv", having three variables (columns). What 3-node belief net structure(s) would best match the dependencies exhibited by this data? (Justify your answer, *i.e.* are there independencies? conditional or otherwise? Does observation of one induce dependencies among the others?)

(Nb. if you use *GeNie* to look into this, you might want to assume a reasonably large link probability (and max computation time), not the 0.001 (and zero) that seems to be the default!)

4: Meal-Shirt-Wine

♠ At https://ecs.wgtn.ac.nz/Courses/AIML429_2024T1/Assignments you'll find a data set called "Meal-Shirt-Wine.csv", having three variables (columns). What 3-node belief net structure(s) would best match the dependencies exhibited by this data? (Justify your answer, *i.e.* are there independencies? conditional or otherwise? Does observation of one induce dependencies among the others?)

Y_1	Y_2	Y_3	X
1	1	0	ape
0	1	1	cat
0	1	0	cat
1	0	1	ape
1	0	1	ape
1	1	0	bat

5: values for factors

Suppose you're given the rather tiny data set $\mathcal{D}$ at left, and are prepared to assume that F_1 causes F_2 causes F_3, and F_2 also causes A. What values would you use in the factor parameters for

- $P(F_2|F_1)$?
- $P(A|F_2)$?

Provide a justification[a] your how you arrived at these values.

[a] Shorten to A options to 'a,b,c' if it helps

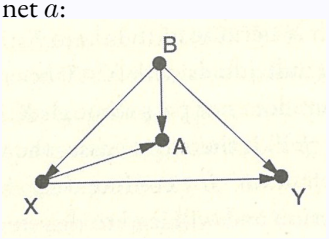
6: building in a structural assumption

Say I have 4 variables A,B,C,D, and I want to build in a "structural" assumption that C indep D, given knowledge of A (but don't want to build in any *other* assumptions besides this one). Which Belief Net structures would do this?

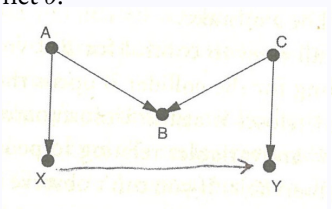
7: causality

Discuss confounding in the following 3 causal networks:

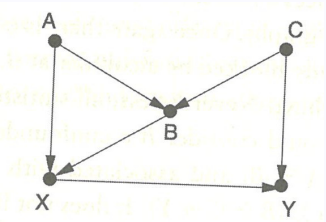
net a:



net b:



net c:



For each one, how could we deconfound the causal relationship $X \longrightarrow Y$?
i.e. for what choice(s)[a] of 'control' @ would $P(Y|X, @ \ldots) = P(Y|do(X), @)$?

[a] if several, list them

8: XOR

Exclusive-OR is the classic minimal failure case for a single layer Perceptron (neural net with a standard non-linear activation function like the sigmoid, or even a hard threshold).

- Can a belief net structured $X1 \longrightarrow Y \longleftarrow X2$ successfully implement XOR? (i.e. map $\mathbf{X}$ to target Y).
- How does the number of learnable parameters (degrees of freedom) of this Bayes Net compare to the number of learnable parameters in the simplest Neural Net capable of doing XOR? You can assume it's a standard Bayes Net (not the Dirichlet version)

X_1	X_2	Y
1	1	cat
0	0	cat
0	1	dog
0	1	dog
1	1	cat
1	1	dog
1	0	dog

9: prediction via Beta

Given the XOR-like data shown at left, what would the "fully Bayesian" prediction be for $Y =$ cat given $X_1, X_2 = (1, 1)$?

10: Dirichlet

It seems to me that Bayesian **prediction** for a Categorical taking K possible values is

$$\Pr(\text{outcome} = i \mid \mathcal{D} = \{n_1, n_2, \ldots n_i \ldots n_K\}) = \frac{\alpha_i + n_i}{\sum_{j=1}^{K} \alpha_j + n_j}$$

♠ where n_i is a count based on the data, and α_i is a 'pseudocount', often set to 1.
By refering to Dirichlet and giving an argument, justify the use (or else rejection!) of this prediction. As part of the argument, you might mention why $\Pr(\text{outcome} = i$ wouldn't just be the most plausible value of p_i (the MODE) under the Dirichlet posterior.
(Hint: we actually did this prediction for the bent coin case in lecture and got $P(\text{heads}|\mathcal{D} = \{h, h\}) = 3/4$ under a flat prior) .