HW 5 - CSE 473 Spring 2021

Due Date: June 9th, 2021 at *11:59 pm PDT*

- ***Please note that this homework does not allow for late submission.***

Total Points: 20 points

- This homework can be done individually or in groups of two.
  - Groups of two should submit their assignment as a group on gradescope.

True/False Questions [2 points]

1.(True / False) – The forward algorithm for HMMs has linear time complexity in the number of states.(0.5pt)
2.(True / False) – P(B|A, C) ∗ P(C|A, B)/P(B, C) = P(A|B, C), given no independence assumptions.(0.5pt)
3. (True / False) – Inference by enumeration can produce incorrect results if the Bayes network is dense (has many edges). (0.5pt)
4. (True/ False) – In order to achieve generalizability in the model's prediction, the algorithm should choose the best hyperparameter set according to the test set. (0.5pt)

Short Answer Questions [4 points]

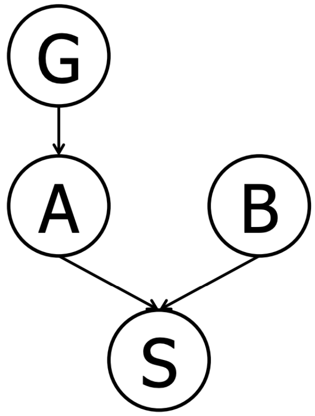
1. (1pt) HMMs can be seen as a special type of Bayesian network. Briefly describe one way in which they differ from the more general case.
2. (1pt) Briefly describe the pros and cons of using the forward algorithm vs. a particle filter for HMMs. When would you use each and why?
3. (1pt) Briefly describe why overfitting occurs and how we can reduce it.
4. (1pt) If you want to design a classification algorithm that given a written review, it finds whether the review is positive or negative, what will be your features. Please mention 2 types of features.

Bayes' Nets Representation (5 points)

Suppose that a patient can have a symptom (S) that can be caused by two different diseases (A and B). It is known that the variation of gene G plays a big role in the manifestation of disease A. The Bayes' Net and corresponding conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic expression.

$\mathbb{P}(G)$	
$+g$	0.1
$-g$	0.9

$\mathbb{P}(A \mid G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



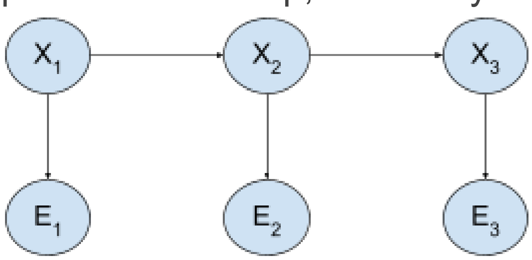
$\mathbb{P}(B)$	
$+b$	0.4
$-b$	0.6

$\mathbb{P}(S \mid A, B)$			
$+a$	$+b$	$+s$	1.0
$+a$	$+b$	$-s$	0.0
$+a$	$-b$	$+s$	0.9
$+a$	$-b$	$-s$	0.1
$-a$	$+b$	$+s$	0.8
$-a$	$+b$	$-s$	0.2
$-a$	$-b$	$+s$	0.1
$-a$	$-b$	$-s$	0.9

a) (2pt) Compute the following entry from the joint distribution: P(+g, +a, +b, +s)=?
b) (1pt) What is the probability that a patient has disease A? P(+a) = ?
c) (1pt) What is the probability that a patient has disease A given that they have disease B? P(+a|+b) = ?
d) (1pt) What is the probability that a patient has disease A given that they have symptom S and disease B? P(+a |+s, +b) = ?

Modified HMM (6 points)

A standard First-Order HMM is graphically represented like this, where the hidden variable at time t only depends on the hidden variable at the previous time step, and every state only includes one evidence variable.



The forward algorithm for a standard HMM is shown below:

$$P(X_t|e_{1:t-1}) = \sum_{x_{t-1}} P(X_t|x_{t-1})P(x_{t-1}|e_{1:t-1})$$

$$P(X_t|e_{1:t}) \propto P(X_t|e_{1:t-1})P(e_t|x_t)$$

Where the first equation is the Elapsed Time and the second equation is the Observation update.

Now imagine a modified HMM, where we receive two independent pieces of evidence each time step. Let A and B be random variables representing those pieces of evidence at time t.

A. (2pts) Draw three steps of a modified HMM chain for this setting. Your drawing should include the variables X, A, and B:
B. (1pt) Write down the formula for Observation for this setting:

$P(X_t | a_{1:t}, b_{1:t}) \propto$

Now consider a different setting where there is only one evidence variable, E, but the next state depends on the evidence from our previous time step as well as the previous hidden state. You could imagine a robot that changes its behavior based on what its sensors are reporting.

C. (2pt) Draw three steps of a modified HMM chain for this setting. Your drawing should include the variables X and E.
D. (1pt) Write down the formula for the Elapsed Time update for this setting:
   $P(X_t | e_{1:t-1}) =$

<u>Markov Chain</u> [3 points]

Consider the following process: You can be employed or unemployed. At each time step, you have a 5% chance of losing your job and a 60% chance of moving from unemployment to employment (finding a new job!). Unless otherwise specified, assume that you have a 50% chance of being employed at time 1.

A. (1pt) Describe the Markov chain that models this employment process. What are the states, initial distribution, and transition/emission probabilities?
B. (1pt) If you start out employed, what is the probability of being employed at step three? Show all of your work.
C. (1pt) What is the stationary distribution of the chain you defined?