CIS 520, Machine Learning, Fall 2021
Homework 9
Due: Tuesday, November 16th, 11:59pm
Submit to Gradescope

Instructions. Please write up your responses to the following problems clearly and concisely. We require you to write up your responses using LaTeX; we have provided a LaTeX template, available on Canvas, to make this easier. **Submit your answers in PDF form to Gradescope. We will not accept paper copies of the homework.**

Collaboration. You are allowed and encouraged to work together. You may discuss the **written homework** to understand the problem and reach a solution in groups. However, **it is recommended that each student also write down the solution independently and without referring to written notes from the joint session.** You must understand the solution well enough to reconstruct it by yourself. (This is for your own benefit: you have to take the exams alone.)

Learning Objectives

After completing this assignment, you will be able to:

- understand how Bayesian nets represent probability distributions
- be able to learn belief net stuctures

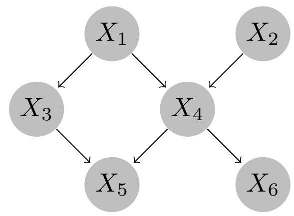
Deliverable

This homework can be completed individually or in groups of 2. You need to make one submission per group. Make sure to add your team member's name on Gradescope when submitting the homework's written part.

1. **A PDF compilation of** `hw9_template.tex`

1 Bayesian Networks [40 points]

Consider the Bayesian network over 6 random variables $X_1, X_2, X_3, X_4, X_5, X_6$ shown below (assume for simplicity that each random variable takes 2 possible values):



1. **[4 points]** Write an expression for the joint probability mass function $p(X_1, X_2, X_3, X_4, X_5, X_6)$ that makes the same (conditional) independence assumptions as the Bayesian network above.

2. **[6 points]** Consider a joint probability distribution satisfying the following factorization:

$$p(X_1, X_2, X_3, X_4, X_5, X_6) = p(X_1)p(X_2)p(X_3)p(X_4)p(X_5 \mid X_3)p(X_6 \mid X_3)\,.$$

   Is this distribution included in the class of joint probability distributions that can be represented by the Bayesian network above? Briefly explain your answer.

3. **[6 points]** If the edge from X_3 to X_5 is removed from the above network, will the class of joint probability distributions that can be represented by the resulting Bayesian network be smaller or larger than that associated with the original network? Briefly explain your answer.

4. **[24 points]** Given the above figure, determine whether each of the following is true or false. Briefly justify your answer.

   (a) $p(X_1, X_2) = p(X_1)p(X_2)$

   (b) $p(X_3, X_6|X_4) = p(X_3|X_4)p(X_6|X_4)$

   (c) $p(X_1, X_2 \mid X_6) = p(X_1 \mid X_6)p(X_2 \mid X_6)$

   (d) $p(X_2, X_5 \mid X_4) = p(X_2 \mid X_4)p(X_5 \mid X_4)$

2 Belief Net Construction [60 points]

Given the following observed counts for all different combinations of the binary random variables A, B, C and D (each variable can be true (T) or false (F)), construct a belief net using the algorithm described in class, where variables are added sequentially to the network. Note that there are 3400 observations in total. *Consider the variables in the order A, B, C, D*, and make sure to give *both* the graph and the conditional probability tables.

A	B	C	D	Count
T	T	T	T	600
T	T	T	F	200
T	T	F	T	0
T	T	F	F	800
T	F	T	T	400
T	F	T	F	0
T	F	F	T	0
T	F	F	F	200
F	T	T	T	200
F	T	T	F	400
F	T	F	T	0
F	T	F	F	200
F	F	T	T	200
F	F	T	F	0
F	F	F	T	0
F	F	F	F	200

In this problem, you will be comparing many conditional probabilities to assess whether some variables are independent of others. In modeling situations like this we need some criteria for deciding whether one variable depends on the other based on the observed conditional probabilities. For this problem, let's suppose if the range of conditional probabilities for a set of variables is larger than 0.05, then the variables are dependent. (In reality, we might evaluate dependence by either i) doing a statistical significance test, or ii) regularizing the Bayes net by adding a cost for each new connection).

Remember to check for both single and joint dependencies between variables. The key question is always to ask for each possible link that one might add to the graph: can it be removed? In building your belief net, follow the algorithm discussed in class:

1. Add A
2. **[9 points]** Add B and decide whether to add a link from A to B. Report if you add the link (3 point) and show steps (6 points).
3. **[9 points]** Add C and decide whether to add a link from A to C. Report if you add the link (3 point) and show steps (6 points).
4. **[9 points]** Decide whether to add a link from B to C. Report if you add the link (3 point) and show steps (6 points).
5. **[9 points]** Add D and decide whether to add a link from A to D. Report if you add the link (3 point) and show steps (6 points).
6. **[9 points]** Decide whether to add a link from B to D. Report if you add the link (3 point) and show steps (6 points).
7. **[9 points]** Decide whether to add a link from C to D. Report if you add the link (3 point) and show steps (6 points).
8. **[6 points]** Draw the final constructed network.

Again, be sure to give the conditional probability tables for downstream variables when appropriate.