CS410: Artificial Intelligence 2021 Fall
Homework 5: Regression & Neural Networks & Bayes Nets
Due date: 23:59:59 (GMT +08:00), January 3 2022

1. **Cross entropy loss.** Recall the statement in Lecture 8, Slide 67 that the cross entropy loss function is convex in θ. Prove this statement in this exercise.

2. **Backpropagation.**

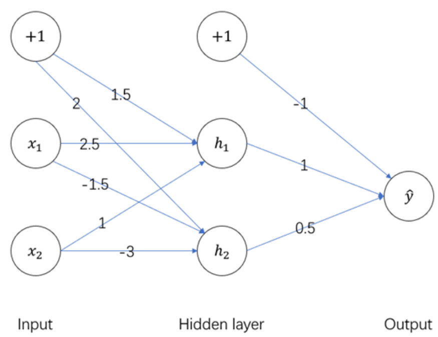


Figure 1: Problem 2.

   (a) Calculate the output values at nodes h_1, h_2 and $\hat{y}$ of this network for input $x_1 = 0, x_2 = 1$. Show all steps in your calculation. Assume that the neurons have sigmoid activation function.

   (b) Compute one step of the backpropagation algorithm with $\eta = 1$ for a given example with input $x_1 = 0, x_2 = 1$ and target output $y = 1$, using new weights and old weights respectively. Compute the updated weights for both the hidden layer and the output layer. Comment on whether a further forward pass gives a lower error. Show all steps in your calculation. The error on the given example is defined as $E = 1/2(y - O)^2$ where O is the real-valued network output of that example at the output node, and y is the integer-valued target output for that example.

3. **Bayes Nets.** Consider the following Bayes net. Calculate the marginal and conditional probabilities $\mathbb{P}(\neg P_3)$, $\mathbb{P}(P_2 \mid \neg P_3)$, $\mathbb{P}(P_1 \mid P_2, \neg P_3)$, and

$\mathbb{P}(P_1 \mid \neg P_3, P_4)$ using **inference by enumeration**. Show all steps in your calculation.

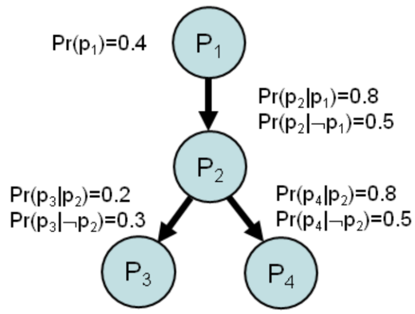


Figure 2: Problem 3.

4. **Bayes Nets.** Consider the same Bayes net in Exercise 3. Compute $\mathbb{P}(\neg P_3)$ and $\mathbb{P}(P_2 \mid \neg P_3)$ using **variable elimination**. Compare the computational complexity of inference by enumeration and variable elimination and discuss your findings. Show all steps in your calculation.

5. **Independence.** Answer the following questions by explicitly showing all steps in your calculation.

   (a) Is D independent from A given B in Figure 3?

   (b) Is D independent from C given E in Figure 4?

   (c) Is D independent from A given E in Figure 5?

6. **Likelihood Weighting.** Consider the following Bayesian network and the corresponding probabilities. Assume we generate the following six samples given the evidence $I_1 = T$ and $I_2 = F$: $(W_1, I_1, W_2, I_2) = \{(S,T,R,F),(R,T,R,F),(S,T,R,F),(S,T,S,F),(S,T,S,F),(R,T,S,F)\}$

   (a) What is the weight of the first sample (S,T,R,F) above?

   (b) Use likelihood weighting to estimate $P(W_2|I_1 = T, I_2 = F)$.

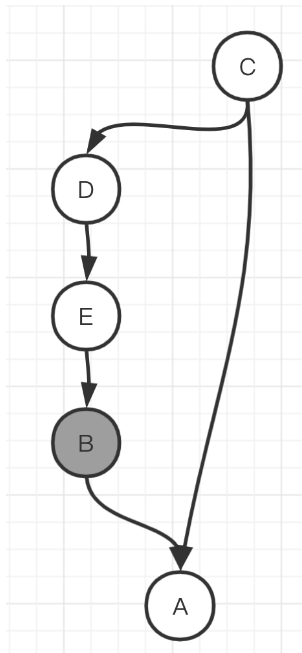


Figure 3: Problem 5.1.

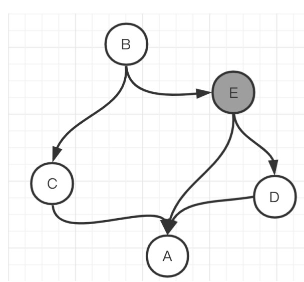


Figure 4: Problem 5.2.

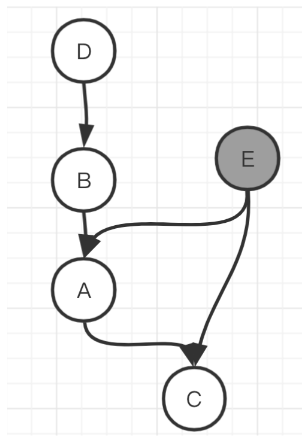
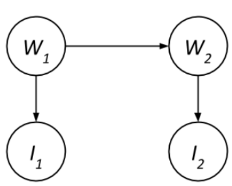


Figure 5: Problem 5.3.



W_1	$P(W_1)$
S	0.6
R	0.4

W_1	W_2	$P(W_2 \mid W_1)$
S	S	0.7
S	R	0.3
R	S	0.5
R	R	0.5

W	I	$P(I \mid W)$
S	T	0.9
S	F	0.1
R	T	0.2
R	F	0.8

Figure 6: Problem 6.