

Preparation for Third Midterm

Practice Questions

1.

- Is the following function P a valid probability function? If you answer "no", explain why not.

$P(\text{day} == \text{Monday}) = 0.4$
 $P(\text{day} == \text{Tuesday}) = 0.4$
 $P(\text{day} == \text{Wednesday}) = 0.2$

- Is the following function P a valid probability function? If you answer "no", explain why not.

$P(\text{day} == \text{Monday}) = 0.4$
 $P(\text{day} == \text{Tuesday}) = 0.4$
 $P(\text{day} == \text{Wednesday}) = 0.2$
 $P(\text{day} == \text{Thursday}) = 0.1$

- Is the following function P a valid probability density function? If you answer "no", explain why not.

$P(x) = 0.1, \text{ if } 10 \leq x \leq 20.$
 $P(x) = 0 \text{ otherwise.}$

- Is the following function P a valid probability density function? If you answer "no", explain why not.

$P(x) = 0.01, \text{ if } 10 \leq x \leq 20.$
 $P(x) = 0 \text{ otherwise.}$

2. Compute $P(\text{fire} \mid \text{alarm})$, given the following information:

$P(\text{alarm} \mid \text{fire}) = A$
 $P(\text{alarm} \mid \text{not fire}) = B$
 $P(\text{fire}) = C$

3. We are given the following information:

$P(\text{fire}) = 0.1$
 $P(\text{earthquake}) = 0.2$
 $P(\text{flood}) = 0.4$

1. Suppose that we do not know whether fire, earthquake, and flood, are independent events. Can we compute the probability $P(\text{fire and earthquake and flood})$? If yes, what is $P(\text{fire and earthquake and flood})$?
2. Suppose that we know that fire, earthquake, and flood, are independent events. Can we compute the probability $P(\text{fire and earthquake and flood})$? If yes, what is $P(\text{fire and earthquake and flood})$?
3. Suppose that we know that fire, earthquake, and flood, are not independent events. Can we compute the probability $P(\text{fire and earthquake and flood})$? If yes, what is $P(\text{fire and earthquake and flood})$?

4. Compute $P(\text{commute time} < 20 \text{ min} \mid \text{temperature} > 80)$, given the following joint probability distribution:

| Commute time | 40–60 Fahrenheit | 60–80 Fahrenheit | above 80 Fahrenheit |
|--------------|------------------|------------------|---------------------|
| < 20 min | 0.1 | 0.05 | 0.1 |
| 20–40 min | 0.2 | 0.1 | 0.1 |
| > 40 min | 0.05 | 0.1 | 0.2 |

5. For the Bayesian network shown in textbook figure 14.2: is $P(\text{Earthquake} \mid \text{Alarm})$ larger, equal to, or smaller than $P(\text{Earthquake} \mid \text{Alarm and Burglary})$? You can either (not recommended) compute both probabilities, or (recommended) provide an intuitive (but correct) justification for your answer.

6. For the Bayesian network shown in textbook figure 14.2: is $P(\text{Earthquake} \mid \text{Alarm})$ larger, equal to, or smaller than $P(\text{Earthquake} \mid \text{Alarm and MaryCalls})$? You can either (not recommended) compute both probabilities, or (recommended) provide an intuitive (but correct) justification for your answer.
7. We are building a decision tree to determine if the next car of a person will be a regular car or a minivan. We have 100 cases as examples. The following is true for those cases:
 - 40 people bought minivans. Out of those 40 people, 30 people were over 35 years of age, and 10 people were under 35 years of age.
 - 60 people bought regular cars. Out of those 60 people, 12 people were over 35 years of age, and 48 people were under 35 years of age.
 What is the entropy gain of selecting the "over 35 years of age" attribute as a test for the root node of the decision tree?
8. Given a set of training examples, is there always a decision tree that perfectly classifies all training examples in that set? If yes, prove your answer. If no, provide a counter example.
9. There are two types of candy bags, type A and type B. Both types of bags contain an infinite number of candies. A bag of type A contains 80% chocolate candies and 20% vanilla candies. A bag of type B contains 40% chocolate candies and 60% vanilla candies. The prior probability $P(A)$ of having a bag of type A is 0.99, and the prior probability $P(B)$ of having a bag of type B is 0.01. What is the posterior probability that we have a bag of type A if the first candy that we pick is a vanilla candy?
10. Design a perceptron takes two inputs X_1 and X_2 , and that outputs +1 if $X_1 > X_2 + 5$, and that outputs 0 if $X_1 \leq X_2 + 5$.
11. Consider a function F that takes three Boolean inputs and gives a +1 response when exactly two (no more, no fewer) of those inputs are set to true (for the inputs, true is encoded by value 1, false is encoded by value 0). Can we construct a perceptron that models function F perfectly? Why, or why not?
12. Design a neural network that implements the XOR function. You can use any number and any type of perceptrons you like. You do not have to specify the weights inside each perceptron, but you need to specify what function each perceptron implements (and, of course, the function should be a function that a perceptron can indeed model).