UNIVERSITY OF Nebraska Lincoln

Introduction to Machine Learning
CSCE 478/878

Written Assignment 1

Spring 2019

Last Name:

First Name:

NUID:

Grader:

Instructions: follow instructions carefully, failure to do so may result in points being deducted.

- Use US Letter (8.5 x 11.0 inches) size papers to prepare your solutions. Pages ripped out from notebooks are not acceptable.
- You MUST handwrite your answers.
- Print out a copy of this cover sheet and score table page; and staple it to the front of your assignment.
- Be sure to show sufficient work and detail calculation to justify your answer(s).
- The CSE academic dishonesty policy is in effect (see http://cse.unl.edu/academic-integrity-policy).

Questions	**Points**	**Score**
1	3	
2	10 (4 + 4 + 2)	
3	8 (4 + 4)	
4	6	
5	5	
6	6	
7	8 (4 + 4)	
8	5	
9	7 (3 + 1 + 3)	
10	9 (4 + 3 + 2)	
11	6	
12	9 (4 + 4 + 1)	
13	8 (2 + 6)	
14	6	
15	4	
16	5	
17	5	
Total	100	

Probability Theory

1) There are many lotteries now that award enormous prizes to people who correctly choose a set of six numbers out of the first n positive integers, where n is usually between 30 and 60. What is the probability that a person picks the correct six numbers out of 40?

2) This problem was posed by the Chevalier de Méré and was solved by Blaise Pascal and Pierre de Fermat.
   a) Find the probability of rolling at least one six when a fair die is rolled four times.
   b) Find the probability that a double six comes up at least once when a pair of dice is rolled 24 times. Answer the query the Chevalier de Méré made to Pascal asking whether this probability was greater than 1/2.
   c) Is it more likely that a six comes up at least once when a fair die is rolled four times or that a double six comes up at least once when a pair of dice is rolled 24 times?

3) Your neighbor has two children. Before getting any information from the neighbor, the probability of your neighbor having one boy is 1/2 and similarly having one girl is 1/2. Answer the following two questions. **You cannot use Bayes' rule**.

   a) Suppose you ask your neighbor whether he has any boys, and he says yes. What is the probability that one child is a girl?
   b) Suppose instead that you happen to see one of his children run by, and it is a boy. What is the probability that the other child is a girl?

4) Show that the statement of conditional independence $p(x, y \mid z) = p(x \mid z)p(y \mid z)$ is equivalent to each of the statements $p(x \mid y, z) = p(x \mid z)$ and $p(y \mid x, z) = p(y \mid z)$.

Expectation, Variance & Covariance

5) The variance of $f(x)$ is defined by: $var[f] = \mathbb{E}[(f(x) - \mathbb{E}[f(x)])^2]$

   Using this formula derive the following:

$$var[f] = \mathbb{E}[f(x)^2] - \mathbb{E}[f(x)]^2$$

6) Show that if two variables x and y are independent, then their covariance is zero.

Bayes' Rule

7) Feynman can get to work in three different ways: by bicycle, by car, or by bus. Because of commuter traffic, there is a 50% chance that he will be late when he drives his car. When he takes the bus, which uses a special lane reserved for buses, there is a 20% chance that he will be late. The probability that he is late when he rides his bicycle is only 5%. Feynman arrives late one day. His boss wants to estimate the probability that he drove his car to work that day.

    a) Suppose the boss assumes that there is a 1/3 chance that Feynman takes each of the three ways he can get to work. What estimate for the probability that Feynman drove his car does the boss obtain from Bayes' rule under this assumption?
    b) Suppose the boss knows that Feynman drives 30% of the time, takes the bus only 10% of the time, and takes his bicycle 60% of the time. What estimate for the probability that Feynman drove his car does the boss obtain from Bayes' rule using this information?

8) After your yearly checkup, the doctor has bad news and good news. The bad news is that you tested positive for a serious disease, and that the test is 99% accurate (i.e., the probability of testing positive given that you have the disease is 0.99, as is the probability of testing negative given that you don't have the disease). The good news is that this is a rare disease, striking only one in 10,000 people. What are the chances that you actually have the disease? (Show your calculations as well as giving the final result.)

Probability Density & Normalization

9) Answer the following three questions about normalization of an arbitrary function.
    a) Why do we need to normalize an arbitrary function?
    b) What happens if an arbitrary function is not normalized?
    c) Normalize the following function f(x).

$$f(x) = \exp\left({}^{\pi x}/_{10}\right) \quad for\ 0 \leq x \leq 10$$

Linear Algebra

10) Consider the matrix M and answer the questions.
    a) Find the rank of M (you MUST reduce it to its **row-echelon** form to compute the rank).
    b) Find the determinant of M.
    c) Does M have inverse? Justify your answer.

$$M = \begin{bmatrix} 6 & 18 & 3 \\ 3 & 9 & 6 \\ 4 & 12 & 9 \end{bmatrix}$$

11) Find the inverse of the matrix N by using the **Gauss-Jordan** elimination method.

$$N = \begin{bmatrix} -4 & 0 & 0 \\ 0 & 8 & 13 \\ 0 & 3 & 5 \end{bmatrix}$$

12) Consider the matrix P and answer the questions.
    a) Find the eigenvalues of the matrix P.
    b) Find the eigenvalues of the transpose of matrix P.
    c) What can you conclude about the two sets of eigenvalues?

$$P = \begin{bmatrix} 3 & 5 & 3 \\ 0 & 4 & 6 \\ 0 & 0 & 1 \end{bmatrix}$$

13) Consider the matrix P and answer the following questions.
    a) Does an ***eigenbasis*** exist for the matrix P? Justify your answer.
    b) Compute the ***eigenbasis*** of P.

$$P = \begin{bmatrix} 3 & 5 & 3 \\ 0 & 4 & 6 \\ 0 & 0 & 1 \end{bmatrix}$$

14) Diagonalize the matrix P. You MUST show detail calculation for computing inverse (**Gauss-Jordan** elimination method) and matrix multiplication.

$$P = \begin{bmatrix} 3 & 5 & 3 \\ 0 & 4 & 6 \\ 0 & 0 & 1 \end{bmatrix}$$

15) A positive definite matrix A satisfies $x^T A x > 0 \quad \forall x \neq 0$, where x is a vector. Using this definition of positive definite matrix, show that a 2 x 2 Identity matrix is positive definite.

Extra Credit Problems (Mandatory for 878 students):

16) **You MUST use the Bayes' rule** to solve the following problem (Monty Hall).

On a game show, a contestant is told the rules as follows:

There are three doors, labelled 1, 2, 3. A single prize has been hidden behind one of them. You get to select one door. Initially your chosen door will not be opened. Instead, the gameshow host will open one of the other two doors, and he will do so in such a way as not to reveal the prize. For example, if you first choose door 1, he will then open one of doors 2 and 3, and it is guaranteed that he will choose which one to open so that the prize will not be revealed.
At this point, you will be given a fresh choice of door: you can either stick with your first choice, or you can switch to the other closed door. All the doors will then be opened and you will receive whatever is behind your final choice of door.
Imagine that the contestant chooses door 1 first; then the gameshow host opens door 2, revealing nothing behind the door, as promised.

Should the contestant (a) stick with door 1, or (b) switch to door 3, or (c) does it make no difference? You may assume that initially, the prize is equally likely to be behind any of the 3 doors.

17) Your neighbor has two children. Before getting any information from the neighbor, the probability of your neighbor having one boy is 1/2 and similarly having one girl is 1/2. Answer the following two questions. **You MUST use Bayes' rule to solve this problem**.

a) Suppose you ask your neighbor whether he has any boys, and he says yes. What is the probability that one child is a girl?
b) Suppose instead that you happen to see one of his children run by, and it is a boy. What is the probability that the other child is a girl?