CS177, Homework 4
Due Date: Wednesday, April 30th

Reading

- Olofsson Chapter 7, Sections 7.1-7.2.3 (p.407-424) : Markov chains, Chapman-Kolmogorov equation, classification of states

Problems

Please document clearly how you arrived at your answer step-by-step.

Problem 1: Correlated errors

Your new startup company has a telephone hotline that PC owners can call for technical support. Mr. John Doe calls your hotline to report a problem with his PC running Windows. Let E be the event that his machine has a virus. Let F be the event that the registry file on his PC is corrupted. The following table represents the joint probabilities of various combinations of these two events.

	E=0	E=1
F=0	0.01	0.19
F=1	0.02	0.78

1. What is $P(E = 1)$?
2. What is $P(F = 1)$?
3. What is $P(E = 1|F = 1)$?
4. What is $P(F = 1|E = 1)$?
5. If John reports that his hard drive has a virus, the probability that his registry file is corrupted has changed. Has it increased or decreased? By what multiplicative factor?
6. If John reports that his registry file is corrupted, the probability that he has a virus has changed. Has it increased or decreased? By what multiplicative factor?

Problem 2: Diagnostic Features

Consider building a model for the following fault diagnosis problem. The class variable C represents the health of a disk drive: C = 1 means it is operating normally, and C = 0 means it is in a failed state. When the drive is running, it continuously monitors itself using a temperature and shock sensor. It records two binary features, X and Y, where each takes values 0 or 1. X is whether the drive has been subject to shock (e.g. dropped), and Y is whether the drive has ever been above 70 degrees C.

The following is the joint pmf of all three variables:

x	y	c	p(x,y,c)
0	0	0	0.1
0	1	0	0.15
1	0	0	0.15
1	1	0	0.1
0	0	1	0.35
0	1	1	0.1
1	0	1	0.05
1	1	1	0.0

Please answer the following questions, showing the equations used to compute all numerical values.

1. What is $P(C = 1)$?
2. What is $P(C = 0|X = 1, Y = 0)$?
3. What is $P(X = 0, Y = 0)$?
4. What is $P(C = 0|X = 0)$?
5. Are X and Y independent? Are X and Y conditional independent given C? Justify your answers

Problem 3: Entropy

Recall from class that we defined the entropy (measured in bits) of a discrete random variable X by the formula:

$$H(X) = \sum_{x_i} p(x_i) \log_2 \frac{1}{p(x_i)} = E\left[\log_2 \frac{1}{p(x)}\right]$$

We also defined the mutual information between two discrete random variables X, Y as the change in entropy of X once Y is observed and derived the formula:

$$I(X,Y) = \sum_{x_i, y_i} p(x_i, y_i) \log_2 \frac{p(x_i, y_i)}{p(x_i)p(y_i)} = E\left[\log_2 \frac{p(x,y)}{p(x)p(y)}\right]$$

Using the same joint pmf from problem 2, please answer the following

1. Compute $H(X)$ and $H(Y)$, the entropy of the sensor variables. Which variable has more “uncertainty”?
2. Compute the mutual information between each feature and the drive health, $I(C,X)$ and $I(C,Y)$
3. Which single diagnostic feature, temperature or shock, would you measure to get the most information about the health of the drive C?
4. Is X independent of C? What would $I(Y,C)$ be if Y and C were independent variables? (Hint: you should be able to answer this question based on the formula for I without having to compute any probabilities.)

Problem 4: Coding

Suppose now we wanted to transmit the state of the drive over the network. The drive can be in one of 8 different states $S = s_1, s_2, \ldots, s_8$ given by the combination of possible values of X, Y and C. For example, $S = s_1 = (X = 0, Y = 0, C = 0)$. These states are listed in the table in problem 2 along with their probabilities.

1. If we use a fixed length code for S, what is the expected codeword length?
2. Construct a Huffman code for S based on the probabilities given in problem 2. Derive the code graphically using the technique described in class of building a tree by recursively merging the lowest probability symbols.

3. What is the expected codeword length for your Huffman code?

4. Calculate the entropy of the hard drive state $H(S)$. Briefly comment on the differences between the entropy, the length of a fixed length code and that of the Huffman code.

Problem 5: Weather

Consider a very simple model for the weather each day in Irvine, where if it rains the weather is recorded as ”rainy” and otherwise the weather is recorded as ”sunny” for that day. If it is rainy on a given day, then the probability is 0.5 that it is rainy again the next day. If it is sunny on a given day then the probability is 0.9 that it is sunny the next day.

1. Draw a state diagram for this Markov chain. Label the two states and draw arrows for the state transitions labeled with their corresponding probabilities.

2. Define the transition matrix for this Markov chain.

3. If it rains on Friday, what is the probability of it being sunny on Saturday?

4. If it rains on Friday, what is the probability of it being sunny the next three days?

5. If it rains on Friday, what is the probability of it being sunny on Sunday?

Problem 6: Web Navigation

Consider a Markov chain with 4 states, named 1,2,3,4 with the following transition probability matrix:

0.3	0.1	0.2	0.4
0.0	0.7	0.1	0.2
0.0	0.0	0.4	0.6
0.0	0.0	0.0	1.0

where rows/columns 1 thru 4 correspond to states 1 thru 4. The initial state probability vector is $(1, 0, 0, 0)$, i.e. the system always starts in state 1.

This could be a very simple model to represent how a user navigates a simple Web site, where each state represents a set of similar pages on the site. For example state 1 could be the set of introductory pages, state 2 represents specific product information pages, state 3 could be the product purchase and checkout pages, and state 4 could represent the state where the person leaves the Web site. State 4 is known as an ”absorbing” state: once the system enters this state it never leaves. Here we use state 4 to represent the end of an individual session on the Web site.

For simplicity in this problem to keep the calculations simple there are no transitions ”backwards”, e.g., from state 3 to state 2 – in a better model would certainly allow such transitions since real users typically go back and forth between different parts of a Web site.

1. Draw the state transition diagram corresponding to this Markov Chain. Only draw arrows for those transitions with non-zero probability.

2. What is the probability that a person will leave the site at time step 2? In other words, what is the probability that the second state in a sequence is state 4 given that the first state (at time step 1) is state 1.

3. What is the probability that a person will leave the site exactly at time step 3? (Hint: there is more than one way for this to happen)

4. What is the probability that a person exists the site without visiting the product purchase page (i.e. they get to state 4 without visiting state 3). Make sure to take into account the ways this can happen

5. Is this Markov chain *irreducible*? (see p. 414 in the text for the definition)