Name _____Jihui.Sheng_____

CPTS 437 - Introduction to Machine Learning

Spring 2021

Exam #1

Full Name: Jihui.Sheng

Duration: 50 minutes

Instructions: Clearly write your name at the top of each page of this exam. Complete all problems on this exam. Show all of your work on this exam. You may use your own calculator. You may *not* use a computer, notes, or other external help. Failure to turn in your exam at the end of 50 minutes will result in deduction of points. Anyone cheating on the exam will receive a zero.

Problem	Points Possible	Your Score
1	26	
2	16	
3	8	
4	12	
5	12	
6	12	
7	8	
8	6	
Total	100	

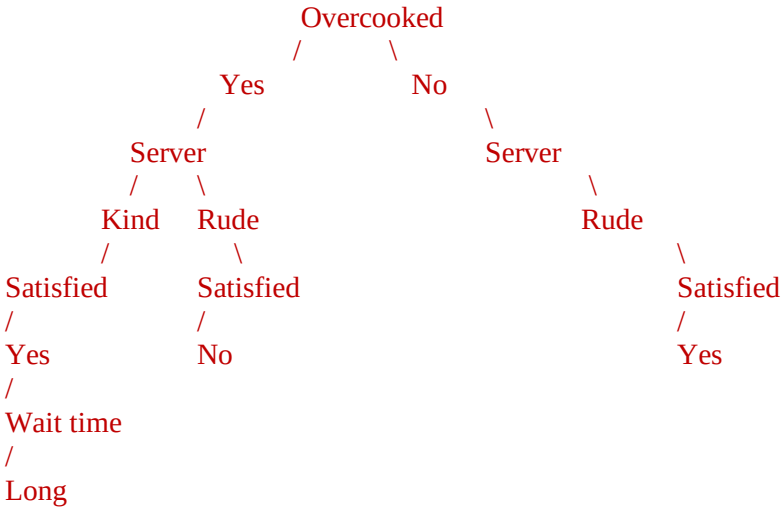
Name _____Jihui.Sheng_____

1. (26 points) This question focuses on the decision tree classifier.
   a. (16 points) Using the training data below, construct a decision tree for the classification of customers at "Pullman Pasta" into *Satisfied* or *Unsatisfied*. Use Information Gain as the decision criterion to select features for the tree nodes. Show your Gain calculations and build the tree such that every leaf node is pure (all of the training examples within a leaf node fit in one class).

Overcooked?	Wait time	Server	Satisfied
Yes	Long	Kind	Yes
No	Short	Rude	Yes
Yes	Long	Rude	No
No	Long	Rude	Yes
Yes	Short	Rude	No

Entropy Values
E(x,x)=1
E(x,0)=E(0,x)=0
E(1,2)=E(2,1)=.92
E(1,3)=E(3,1)=.81
E(2,3)=E(3,2)=.97

Solution:
Let + represent Yes, and – represent No
E(P+, P-) = Entropy(S) = -(P+)log(P+)-(P-)log(P-) = 0.97
Gain(S, Wait time) =0.97 – (E(1,1)(Short) + E(2,1)(Long)) = 0.97 – (1*(2/5) +0.92*(3/5)) = 0.018
Gain(S, Server) = 0.97 – (E(1,0)(Kind) + E(2,2)(Rude)) = 0.97 – (0 + 1*(4/5))=0.17
Gain(S, Overcooked) = 0.97 – (E(1,2)(No) + E(2,0)(Yes)) = 0.97 – (0.92*(2/5)+0*(3/5)) = 0.602
Overcooked is root



   b. (4 points) Use the decision tree you built to predict whether the two test cases will be satisfied.

Overcooked?	Wait time	Server	Satisfied
No	Short	Kind	No
Yes	Long	Rude	No

Name _____Jihui.Sheng_____

c. (6 points) Write the if-then rules that correspond to your tree predicting Satisfied=Yes.

(Overcooked=No&Server=Rude)or(Overcooked=Yes&Server=Kind&Wait tim =Long)

2. (16 points) Simulate one iteration of the k-means algorithm on the data below. Each data point is represented by its <X,Y> location in a 2-dimensional space. Use Euclidean distance as the distance measure for this process. The current cluster centers are C1 = <1.00, 1.00>; C2=<2.50, 2.50>.

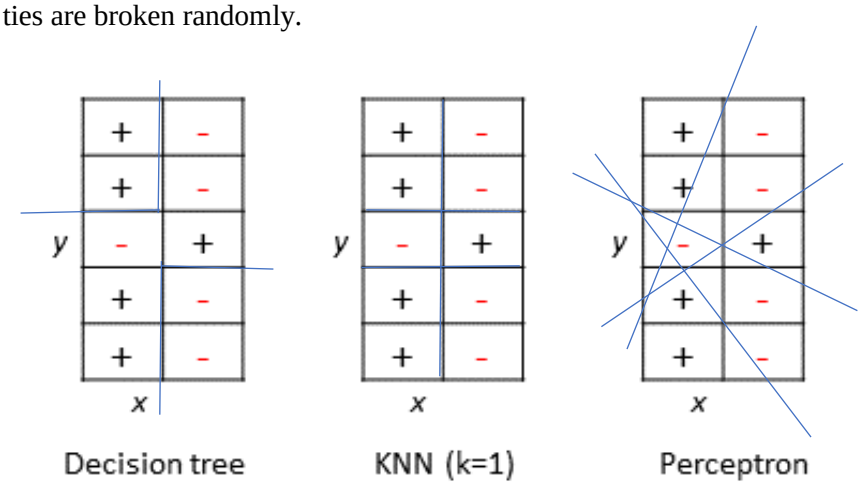
X	1.90	1.75	2.32	2.31	5.02	5.74
Y	0.97	0.84	1.63	2.09	3.02	3.84

Name _____Jihui.Sheng_____

3. (8 points) Does k nearest neighbors spend more computation time during training or testing? Explain your answer.

Yes, it will calculate the average of the training set every time。

4. (12 points) On the 2D dataset below, draw the decision boundaries learned by the following algorithms based on features x and y. Mark which decision regions are labeled positive or negative, and assume ties are broken randomly.



Name _____Jihui.Sheng_____

5. (12 points) What is the curse of dimensionality and what are two ways it can affect a geometric algorithm such as k nearest neighbors?

   [x1,x2,y]

6. (12 points) Write the formula for computing activation in a perceptron classifier. Define each of the variables used in your formula. Under what conditions would a perceptron increase the value of the bias term?

Name _____Jihui.Sheng_____

7. (8 points) Why is inductive bias important for a machine learning algorithm? Give two examples of ML algorithms that we discussed in class and their corresponding inductive biases.

8. (6 points) The graph below shows the classification accuracy of a decision tree as a function of the size of the tree. Explain why the accuracy on test data may be lower than for training data and give the name of this phenomenon.