CMSC 478, Spring 2022, Midterm Exam
**April 4**

Name:

Directions (please read carefully)

1. Be sure to write your name above
2. This exam is closed book and closed note, and you are not to communicate with anyone but the proctors during the exam
3. There is no need for a calculator of any kind, so the use of electronics is prohibited
4. Show your work, where appropriate, for partial credit

Question 1: Suppose you have a function of two variables $f(x, y) = (x - y)^2$ that you are trying to minimize. If the initial values are x = 2 and y = 1 and the learning rate is 0.1, what are the values of x and y after one step of gradient **descent**.

Question 2: Suppose the perceptron algorithm is running and the current weight vector is w = (2, -1) and b = 0. What are w and b after the algorithm processes the next instances which is x = (0, 1) and y = 1?

Question 3: Suppose the perceptron algorithm terminates with w = (w1, w2) = (1, 1) and b = -0.5. Draw the boundary between the positive and negative instances in the plot below (note that x1 is the horizontal axis and x2 is the vertical axis). And show which half plane is the positive half by putting a + in it.

x2

0.9	
0.8	
0.7	
0.6	
0.5	
0.4	
0.3	
0.2	
0.1	
0	

0   0.1   0.2   0.3   0.4   0.5   0.6   0.7   0.8   0.9   1   x1
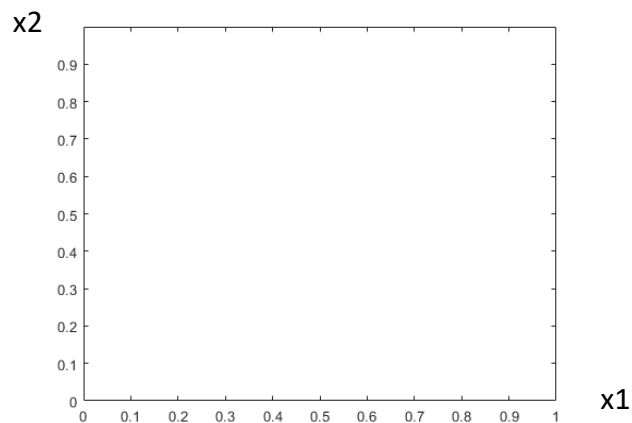
Question 4: Given the dataset below which has 3 features and a class label which is either 0 or 1, fill in values for the class label that would ensure that the root split of a decision tree will be x2. That is, you should put either a 0 or 1 in each row of the y column in the table below.

x1	x2	x3	y
0	0	0	
0	0	1	
0	1	0	
0	1	1	
1	0	0	
1	0	1	
1	1	0	
1	1	1	

Question 5: Explain why logistic regression is trained using **gradient ascent** instead of gradient descent.

Question 6: Both k-means and kNN have a k parameter. Say for each algorithm whether overfitting is more likely with larger or smaller values of k and why.
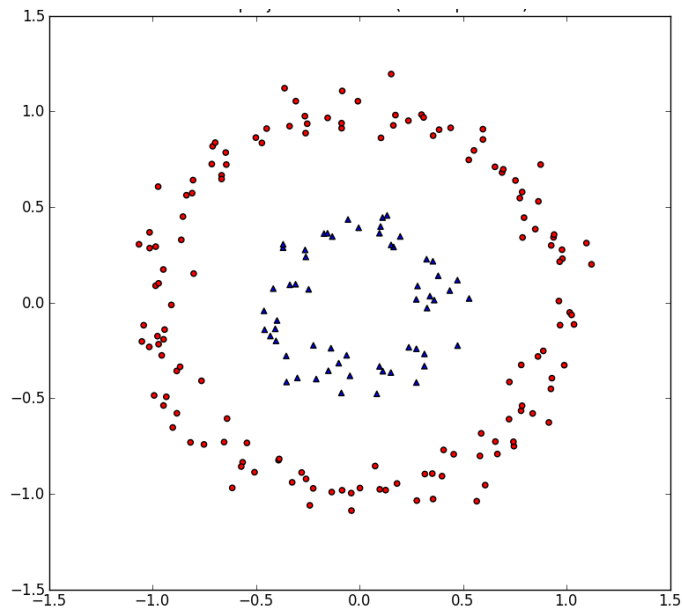
Question 7: Given the 2x2 contingency table below, what are recall and precision if y = 1 is the positive class? Given each of them as a ratio of whole numbers. Note that y is the true label and y' is the predicted label.

	y' = 1	y' = 0
y = 1	45	7
y = 0	4	22

Question 8: There is an MDP with 4 states (s1, s2, s3, s4) and two actions (A, B). Consider the Q-table below. What cell changes and to what value if the agent does a Q update after taking action A in state s3, gets a reward of 4 and winds up in state s1? The learning rate and discount factor are both 0.5. Give your final answer as a single number.

	A	B
S1	4	8
S2	8	7
S3	16	8
S4	5	5

Question 9: Consider the dataset in the figure which has two features (x1, x2) and two classes (denoted with red and blue). Given the polynomial kernel $K(a, b) = (a \bullet b + 1)^d$, what is a value of d for which the data will be linearly separable in the higher dimensional space? Explain why that value of d works.

Scratch Paper