CS 422 Exam 1 Name _____

Short Answer (10 Points)

1. In the following optimization problem, what values of λ would result in overfitting? What about underfitting? Briefly explain.

$$\min L(w, b) = e^{y\hat{y}} + \lambda ||w||^2$$

2. True/False: False negatives are only important for calculating Recall. Briefly explain.

Short Answer (10 Points)

3. True/False: There is no value of K for which K-NN will achieve 100% accuracy on the training data on every dataset. Briefly explain.

4. True/False: The perceptron will always converge to the same w and b. Briefly explain.

Decision Trees (20 Points)

5. Assume you have the following training data. Using the Information Gain algorithm from class, build the best depth-1 decision tree for this data. Show your work for full credit.

Sample	Junk Food	Exercise	Label
s_1	0	0	0
s_2	0	1	1
s_3	1	0	0
s_4	1	1	0
s_5	0	0	1
s_6	0	1	1
s_7	1	0	1
s_8	1	1	0

K-Means (10 Points)

6. How might K-means cluster the following data? Indicate the cluster centers and the rough clusters on the graph. Explain your choices briefly.
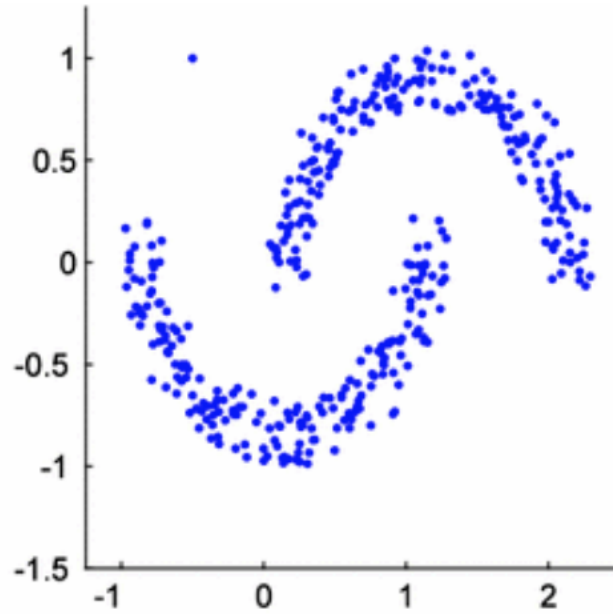
Linear Classifiers (20 Points)

7. Give the gradient descent update rules for the following regularized loss function. Show your work for partial credit!

$$L(w, b) = \sum_n (y_n - (wx_n + b))^2 + \lambda ||w||^2$$

Perceptron (10 Points)

8. Run the perceptron algorithm on the following data in the order provided for two epochs. Give the final w and b produced by the algorithm at the end of the first epoch.

| Sample	x_1	x_2	y
s_1	0	1	1
s_2	1	0	1
s_3	1	1	1
s_4	2	2	-1
s_5	2	1	-1
s_6	1	2	-1

Gradient Descent (10 Points)

9. Run gradient descent on the following function for two steps. Use the starting point $x_0 = 5$ and $\eta = 0.1$.

$$f(x) = x^2 + 3$$

KNN (10 Points)

10. Using K=1, 3 and 5, classify the following test data using the plotted training data (x = negative, filled circle = positive). Use the table below to record your final answers. Indicate ties with the answer $+/-$.
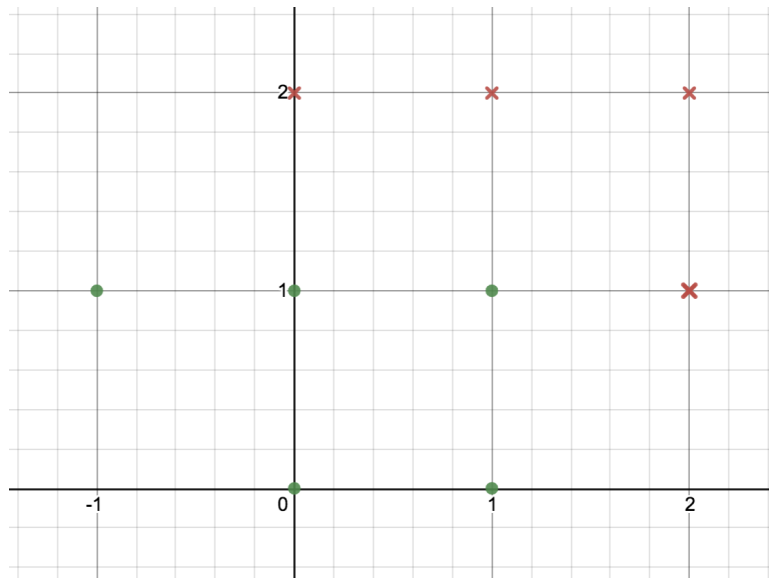
Sample	$K = 1$	$K = 3$	$K = 5$
(1.1,1)			
(0.5,0.5)			
(0,1.5)			
(2,0)			

Equations

Entropy and Information Gain

$$H = \sum_{c \in C} -p(c)log_2\big(p(c)\big)$$

$$IG = H - \sum_{t \in T} p(t)H(t)$$

p	$plog(p)$
$\frac{1}{8}$	-0.375
$\frac{1}{4}$	-0.5
$\frac{3}{8}$	-0.53
$\frac{1}{2}$	-0.5
$\frac{5}{8}$	-0.423
$\frac{3}{4}$	-0.311
$\frac{7}{8}$	-0.168
1	0

Perceptron

$$a = w \cdot x + b$$

$$w = w + xy$$

$$b = b + y$$

CS 422 Exam 1