UNIVERSITY OF WISCONSIN-LA CROSSE
Department of Computer Science

CS 419/519 **Machine Learning** **Fall 2018**
Midterm Exam (Practice) **24 October 2018**

- **Do not turn the page until instructed to do so.**
- **This is an open-book, open-notes exam.**
- **This booklet contains 6 pages including the cover page.**
- **You have 55 minutes.**
- **The maximum possible is 40 points.**
- **Write all your answers on the exam sheets.**

PROBLEM	SCORE
1	
2	
3	
4	
TOTAL	

NAME: _____

A note as you begin: the last page of this examination contains a number of data-sets that are used to answer various questions. I suggest you detach that pages now, for ease of reference, and then re-attach it after the exam.

1. **(8 pts.) SHORT ANSWER PROBLEMS.**

   Circle the appropriate answer and fill in the blanks where required.

   a. (*4 pts.*) Data-set **A** (end of exam) **(IS/ IS NOT)** linearly separable, because _____

   _____

   _____.

   Data-set **B** (end of exam) **(IS/ IS NOT)** linearly separable, because _____

   _____

   _____.

   b. (*4 pts.*) If we use a radial basis function kernel on 2-dimensional data, the result is a translation of that data to _____ dimensions.

   Such a function is of the form: $k(\mathbf{x}, \mathbf{z}) = e^{-\frac{||\mathbf{x}-\mathbf{z}||^2}{2\sigma^2}}$.

   It has highest value at _____ and that highest value is equal to _____ . The parameter σ controls _____ .

2. **(10 pts.) DECISION TREES.**

(a) (5 pts.) Compute the information-theoretic entropy of data-set **C** (end of exam). Show all necessary work; results should be accurate to no less than 3 decimal places.

(b) (5 pts.) Draw the decision tree that the algorithm covered in class would produce (using information gain as the principle for choosing attributes on which to split data).

3. **(10 pts.) CLASSIFICATION METHODS.**

Suppose we have a 2-dimensional data point, belonging to the class 0: $(\mathbf{x}, y) = (0.1, 0.2, 0)$. Suppose further that we have a weight vector: $(1, 1, 1)$

(a) (4 pts.) Compute the output of the logistic function upon this data-point, using that weight vector. Show all necessary work; results should be accurate to no less than 3 decimal places.

(b) (6 pts.) Compute the new weight vector we will get after doing a single iteration of logistic regression updating based upon the error made on the data-point given. Again, show all work, and use no less than 3 digits of decimal accuracy.

4. **(12 pts.) NEURAL NETWORKS.**

(a) (4 pts.) Draw a single-layer perceptron network for data-set **D** (end of exam). Assume the two layers are fully inter-connected. Label the output neurons with the relevant classes for which they stand in your network.

(b) (8 pts.) Assuming all weights (including bias weights) are set initially to 0.1, compute the output of each output neuron in the network, and the error they make, on the single data-point: $(\mathbf{x}, \mathbf{y}) = (0.5, 0.4, 0.2, (0, 1))$. Show all necessary work; results should use no less than 3 digits of decimal accuracy.

Data-set A (question 1): 1-dimensional numerical data-points belonging to two classes, X and Y:

$$\{(0.1, X),\ (0.2, X),\ (0.3, X),\ (0.4, Y),\ (0.5, Y),\ (0.6, Y)\}$$

Data-set B (question 1): 1-dimensional numerical data-points belonging to two classes, X and Y:

$$\{(0.1, X),\ (0.2, Y),\ (0.3, X),\ (0.4, X),\ (0.5, Y),\ (0.6, Y)\}$$

Data-set C (question 2): 2-dimensional data-points, each with two features, Color and Size, belonging to two classes, *Hippo* and *Not-Hippo*:

$$\{\ (Grey, Large, Hippo),\ (Grey, Medium, Hippo),\ (Grey, Small, Hippo),$$
$$(Pink, Small, \textit{Not-Hippo}),\ (Green, Large, \textit{Not-Hippo})\ \}$$

Data-set D (question 4): 3-dimensional numerical data-points belonging to two classes, X and Y; in this data-set an output vector $\mathbf{y} = (1, 0)$ means the data is of type X, while vector $\mathbf{y} = (0, 1)$ means the data is of type Y:

$$\{\ (0.5, 0.4, 0.2, (0, 1)),\ (0.1, 0.2, 0.3, (0, 1)),\ (0.3, 0.4, 0.5, (0, 1)),$$
$$(0.9, 0.8, 0.5, (1, 0)),\ (0.8, 0.9, 0.5, (1, 0)),\ (0.9, 0.9, 0.4, (1, 0))\ \}$$