CSCE 633: Machine Learning

EXAM # 2

Fall 2018

Total Time: 75 minutes

Name: _____

UID: _____

Question	*Point*	*Grade*
1	*20*	
2	*35*	
3	*45*	
Total	*100*	

Full Name of Person Sitting to Your Left:

Full Name of Person Sitting to Your Right:

1. (20 points) Concepts.

a) What are the differences between Bagging, Random Forest, and Boosting? What are some pros and cons of each?

b) Is there a limit to the kinds of methods that can be gradient boosted?

2. (35 points) Clustering

a) (5 points) Principal Component Analysis. Assume you have a matrix X that is n samples by p features. How do you prepare this data for use in PCA?

b) (10 points) Assume you have a Matrix A

$$A = \begin{pmatrix} 2 & 2 \\ 5 & -1 \end{pmatrix}$$

Find the Eigenvalues λ and associated eigenvectors.

This page intentionally left blank

c) (5 points) Assume you have 3 clusters in your data. Please explain how K-Means Clustering finds these clusters. How do you assign a new member to one of the existing clusters?

d) (5 points) Now assume you wanted to repeat the process using Gaussian Mixture Models, explain how Expectation Maximization works. Be sure to describe the differences between Diagonal and Full Covariance. How do you assign a new member to one of the existing clusters and how does this differ from K-Means?

3. (25 Points) Neural Networks

a) (5 points) Assume you have two inputs X_1 and X_2 and want to calculate an Output Y, using one hidden layer of two nodes. Please draw this network fully connected network.

b) (15 points) Assuming you have a sigmoid activation functions, X1 = 0.10, X2 = 0.05, initial weights between X1 and hidden node 1 = 0.2, between X1 and hidden node 2 = 0.3, X2 and hidden node 1 = 0.3, and X2 and hidden node 2 = 0.2, the weights between hidden node 1 and the output is 0.5, and the weight between hidden node 2 and the output is 0.1. Assume all biases are 0.1. Please write the formulation for the feed forward output and backward propagation for one iteration of updating weights. Note: We are not asking you to compute the values but do expect to see the numeric values in the equations.

This page intentionally left blank

c) (10 points) Autoencoder. What is an autoencoder, what is it used for, and why does it generally use a smaller number of dimensions for hidden layer compared to its input and output layers?

d) (15 points) Convolutional Neural Network. Consider one **convolution layer** with one 2 x 2 kernel and stride 1, succeeded by a **ReLU activation** and then a 2 x 2 **max pooling layer** with stride 2. The input and the kernel are shown as below. Please calculate the final output. Be sure to show your work.

Note: The convolution takes the formula

$$y_{i,j} = \sum_{k=0}^{K-1} \sum_{l=0}^{L-1} x_{i+k,j+l} \cdot f_{k+1,l+1}$$

Where K, L are the width and height of the kernel f. i, j iterate over the height and width of the input. So when the convolution starts the first element of the kernel will overlap with the first element of the input. Please use zero-padding to make sure the output of the convolution will have the same shape as the input.

Input:

0.3	**0.2**	**-0.1**	**0.7**
-0.7	**0.6**	**0.8**	**0.3**
0.9	**-0.3**	**-0.1**	**0.4**
-0.4	**0.2**	**1.0**	**-0.7**

Kernel:

0.1	**1.0**
0.8	**-0.2**

This page intentionally left blank