

EE 258 - HOMEWORK #2 FALL 2019

Due: Sep 18 3:00PM

PROBLEM 1

$$\varphi_1(z) = 1/(1+e^{-z})$$

$$\varphi_2(z) = \tanh(z)$$

$$\varphi_3(z) = \text{RELU}(z)$$

- Sketch the above activations functions on the same figure (manually and using Python).
- Find the derivative of the above activations. Express each derivative in terms of the activation function itself (if possible).

PROBLEM 2

Consider a feed-forward fully-connected neural network containing one input layer with 10 inputs, followed by one hidden layer with 20 neurons, and finally one output layer with 2 neurons. All neurons use the ReLU activation function. The output can be expressed as

$$\mathbf{y} = \text{ReLU}(\text{ReLU}(\mathbf{x}^T \cdot \mathbf{W}_h + \mathbf{b}_h^T) \cdot \mathbf{W}_o + \mathbf{b}_o^T)$$

where \mathbf{x} is the input matrix.

- Draw a representative figure for the described neural network.
- Explain what do \mathbf{x} , \mathbf{W}_h , \mathbf{b}_h , \mathbf{W}_o , and \mathbf{b}_o represent?
- What is the size of each matrix and vector in the expression? Explain your reasoning.
- How many neurons do you need in the output layer if you want to classify email into spam or ham? Is RELU an appropriate activation function at the output layer in this case?
- If after learning the coefficients and testing the performance of the above neural network, you observe that training error is much more lower than the test error. Which of the following will you change (and how?) in your network in order to design a better performing model?
 - Number of neurons in the hidden layer
 - Number of neurons in the output layer
 - Number of hidden layers
 - Number of inputs
 - Activation functions
 - Weights
 - Biases

- f) If after learning the coefficients and testing the performance of the above neural network, you observe that training error is still high. Which of the following will you change (and how?) in your network in order to design a better performing model?
- a. Number of neurons in the hidden layer
 - b. Number of neurons in the output layer
 - c. Number of hidden layers
 - d. Number of inputs
 - e. Activation functions
 - f. Weights
 - g. Biases

PROBLEM 3

1. **Get the data:** For this homework, use the iris dataset under Files-> Datasets on Canvas. (submit the code)

Hint: Pandas knows to treat rows with 'NA' as missing values

2. **Take a look at the data structure:** Write a brief paragraph about the data set related to each of the following observations (submit the code, the result, and the paragraphs):
- a. Look at the top five rows of the data set
 - b. Get a quick description of the data: Notice if there are any missing values or categorical features
 - c. Get a summary of the numerical features
 - d. Plot the histogram of the numerical features
3. **Discover and visualize the data to gain insights:** (submit the plots, code, and your observations of each plot)
- a. Obtain scatter matrix
 - b. Obtain the correlations among features and comment
4. **Data Cleaning** (submit the code, results and plots showing the changes in data)
- a. Drop the data points with NA in it
 - b. Tidy up the data by renaming the “class” data point correctly.
 - c. Remove the outliers: drop the 'Iris-setosa' rows with a sepal width less than 2.5 cm.
 - d. One of the data collectors forgot to convert the sepal length for “Iris-versicolor” to cm, instead added the data as meters. Find those and convert them to cm.
 - e. Handle the categorical variables
 - f. Save the clean data into a new file.