Neural Network and Deep Learning

Mid-term Exam
Total Points: 100
Deadline: March 1, 2022

1. [**Points 30**] In the logistic regression classifier, we are fitting an s-shape curve to fit the data. We are given 10 sample points with corresponding odds value as follows, 0.22, 0.33, 4, 5, 15.81, 19.01, 7.89, 3.95, 0.05, 0.15.

   (i) Compute log-odds values for those given data points. [**Points 5**]

   (ii) Compute the probability of those points. [**Points 8**]

   (iii) Compute log-likelihood for these given data points. [**Points 7**]

   (iv) Compute total cost for these data points. [**Points 10**]

2. [**Points 20**] Suppose we are given the task of detecting a cat present in a given image or not. To finish this task, you create two networks - a multilayer perceptron network and a neural network with similar configurations. Both have two layers – the first hidden layer has 4 units, and the second layer considers as the output layer. The input dimension is 3.

   (i) What is the difference between a perceptron and a neuron? [**Points 5**]

   (ii) Draw a diagram of both networks and show details weights and bias. [**Points 10**]

   (iii) Which network would be best to detect the cat in an image and why? [**Points 5**]

3. [**Points 25**] We are given equations as below.
   $y_1 = w_1x_1 + w_2x_2 + w_2x_3$
   $y_2 = w_4x_1 + 4x_1y_1$
   $y_3 = 2y_2 + 5x_1$
   $L = 1/y_3$

   (i) Draw the computational graph of these given equations. [**Points 5**]

   (ii) Compute derivative with respect to w_i and show detailed derivatives/loss propagation in your computational graph. [**Points 10**]

   (iii) Given w_1, w_2, w_3, w_4 are 1.1, 2, 1.5, 4.2 and given x_1, x_2, x_3 are 10, 15, 9, respectively. Compute forward pass values for each node in your graph and show the derivative loss in each node in the graph. [**Points 10**]

4. [**Points 25**] Given a five-layer neural network to classify cats, dogs, and goats. The first layer, the second layer, third layer, fourth layer comprise 24, 32, 24, and 8 neurons. Assume that each layer has the bias terms. Assume that you are given 100,000 images containing a cat, dog, and goat.

   (i) Compute the total number of parameters in this network. What are your hyper-parameters in this network? [**Points 10**]

   (ii) Which loss function will you use to train this network and why? [**Points 5**]

   (iii) How will you train this network and which data splitting mechanism you will use? How will you measure the performance of your model? Show detailed rationale behind your answer. [**Points 10**]