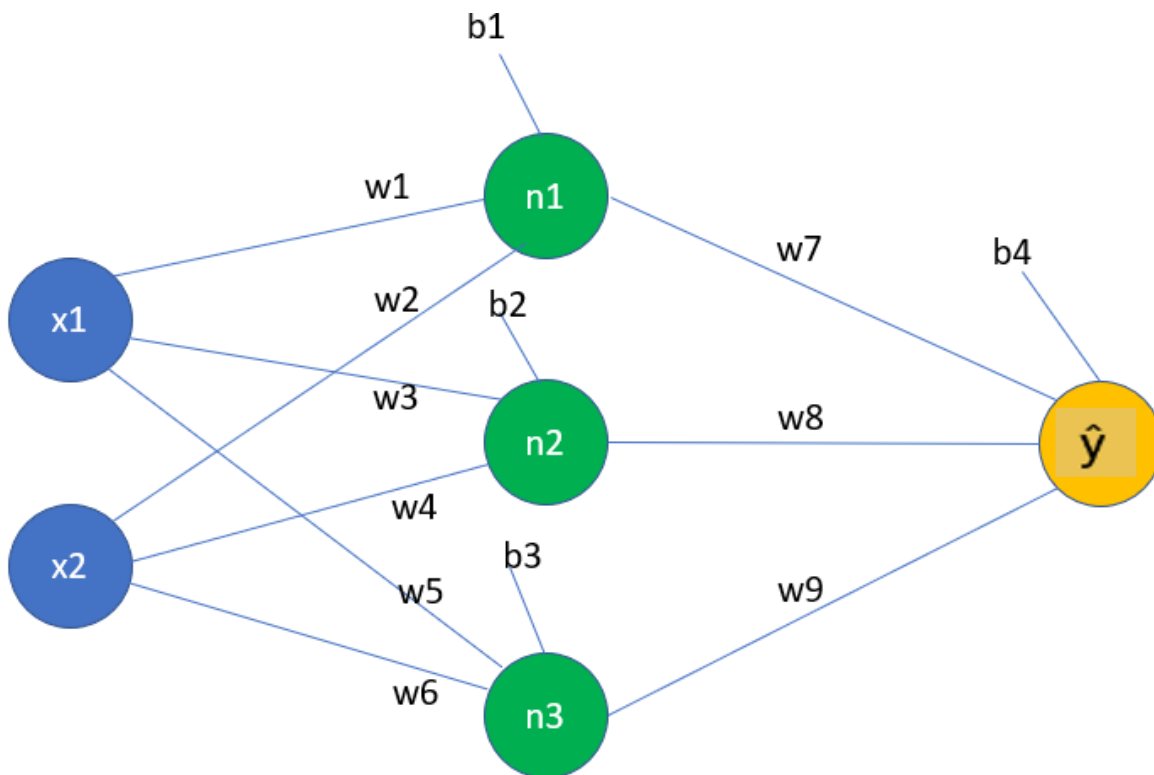


DATA 255 Deep Learning Technologies – Homework -1

Deadline – 11.59 PM – 9/20/2023

Problem 1 (8 pts): The figure below shows a 2-layer, feed-forward neural network with three hidden-layer nodes and one output node. x_1 and x_2 are the two inputs. For the following questions, assume the learning rate is $\alpha = 0.1$; activation function = sigmoid; loss function, $MSE = \frac{1}{2} (y - \hat{y})^2$; target output $y = 1$; For instance, the output of n_1 equal $\text{sigmoid}(w_1 * x_1 + w_2 * x_2 + b_1)$. Compute one step of the backpropagation. Here, a = last digit of your student ID

- Assume $x_1 = a$, $x_2 = 1$; all weights and biases equal 1. Compute the updated weights for both the hidden layer and output layer. Show all steps in your calculations. (4 pts)
- Assume $x_1 = 0$, $x_2 = a$; $w_1 = w_3 = w_5 = 0.5$; $w_2 = w_4 = w_6 = -0.5$; $w_7 = 1$, $w_8 = -1$, $w_9 = 0$ and biases weights equal 0.1. Compute the updated weights for both the hidden layer and output layer. Show all steps in your calculations. (4 pts)



Coding!!!

Problem 2: We will develop an Artificial Neural Networks using MNIST digit data, where we will train an ANN model and then classify new instances. You can directly download the data using <https://keras.io/api/datasets/mnist/> dataset currently contains 10 classes where each of the image sizes is (28×28) . You should split the data into train and test data, where train data should be used for only training the model and test data should be kept separated and used only for evaluation purpose. You should select last two digits of your student ID – meaning that if your student id is 006000104, then you should select 0 and 4 for developing the binary classification model.

- A. Build a NN for binary classification with **early stopping** criteria based on validation loss. Evaluate your model on the test data. Construct a confusion matrix. Present learning curve and include some examples of your prediction **(3pts)**
- B. Build three NNs for binary classification using three **different weight initializers** held other hyperparameters constant.
 - a. Construct three different confusion matrices **(3pts)**
 - b. Show three different learning curves and explain any differences **(2pts)**
 - c. Show accuracy using bar plots and explain – if there any difference in results for using three different initializers **(1pts)**

Problem 3 **(3 pts)**: Build a NN for multi-class classification considering all the classes (10 classes) in the MNIST digit dataset: consider **early stopping criteria based on the validation loss** and finally construct confusion matrix and discuss the results.

You are required to submit:

1. An MS/PDF/Scanned document:
 - a. Include all the steps of your calculations.
 - b. Attach screenshots of the code output.
 - c. Include the summary of the model
 - d. Include a Table - Mention all the hyperparameters you selected: activation function in hidden layer and output layer, weight initializer, number of hidden layers, neurons in hidden layers, loss function, optimizer, number of epochs, batch size, learning rate, evaluation metric
 2. Source code:
 - a. Python (Jupyter Notebook)
 - b. Ensure it is well-organized with comments and proper indentation.
- Failure to submit the source code will result in a deduction of 5 points.
 - Format your filenames as follows: "your_last_name_HW1.pdf" for the document and "your_last_name_HW1_source_code.ipynb" for the source code.
 - Before submitting the source code, please double-check that it runs without any errors.
 - Must submit the files separately.
 - Do not compress into a zip file.
 - HW submitted more than 24 hours late will not be accepted for credit.