MANASVI AGGARWAL
MTECH(Res)
DEEP LEARNING ASSIGNMENT-1
SR NO: 16223

PART 1:

In this part we have to make a python program to test if the numebr is divisible by 3 then the output should be Fizz and if number is divisible by 5 then the output is Buzz and if the number is divisible by 15 then the output should be FizzBuzz and if not divisible by 3 and 5 the the number should be in the output. As there are no hyperparameters to be learned in this program so no need for training.

PART 2:

In this I trained a pytorch neural network model which is tained on some inputs and tested on some other samples. My current model has 2 layers (excluding the input layer) and activation function used is Relu. Number of hidden units are 150. On training set the accuracy is 99-100% and test set also the saved model shows 100% accuracy. That means all are correctly labeled either Fizz, Buzz, FizzBuzz.

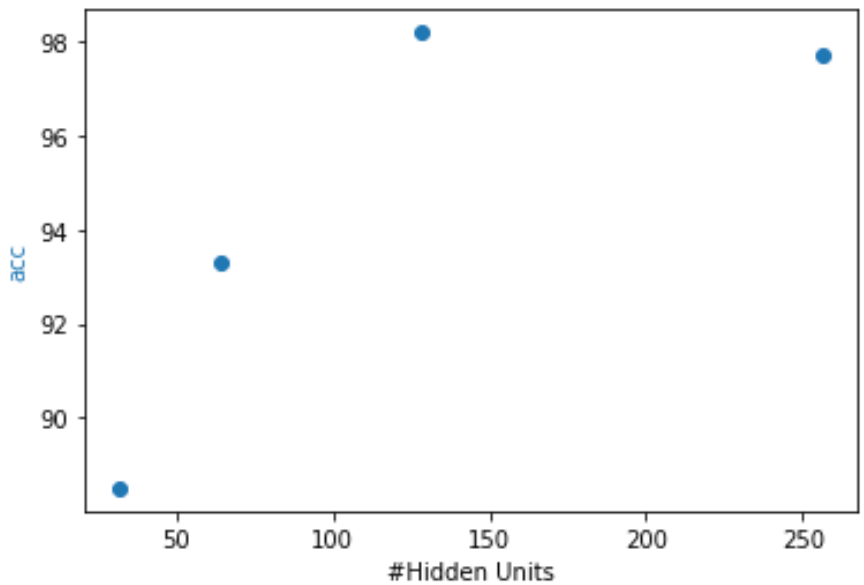
During training my model has different hyperparameters:

1) Learning rate
2) Number of hidden units
3) Epochs
4) Number of hidden layers
5) Activation Function
6) Optimizer

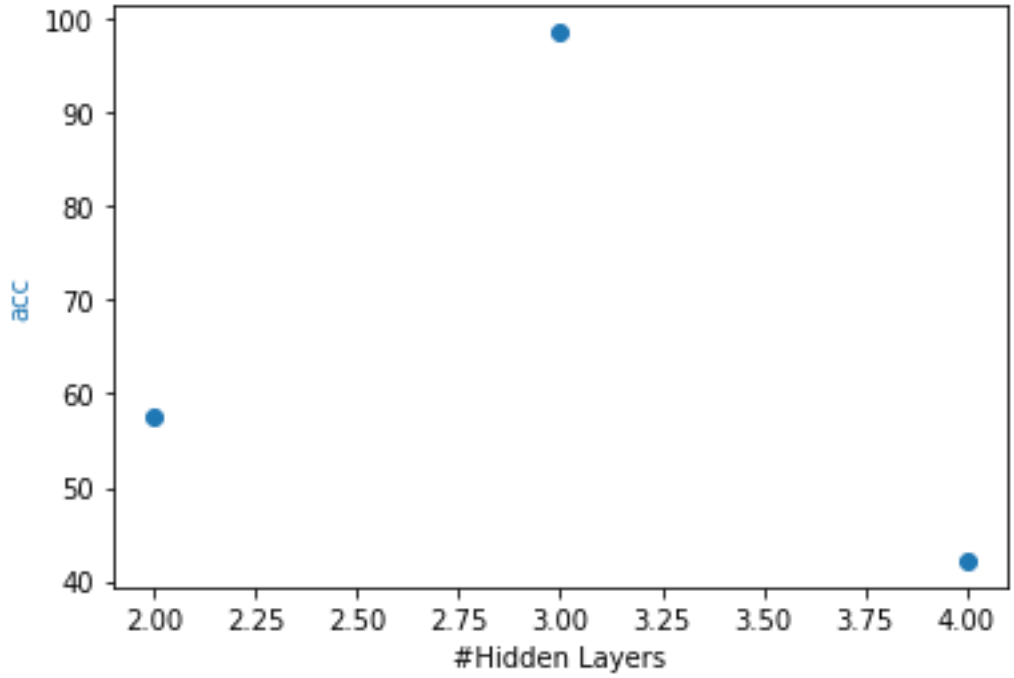
For activation function, I tried Relu and Sigmoid activation functions. For optimizer I tried SGD and Adam. Performance on test set was optimal when activation function set to Relu and Optimizer set to Adam. Learning rate of Ada is set to 0.001.

Given below are some plots for Number of hidden units, epochs, Number of hidden layers.

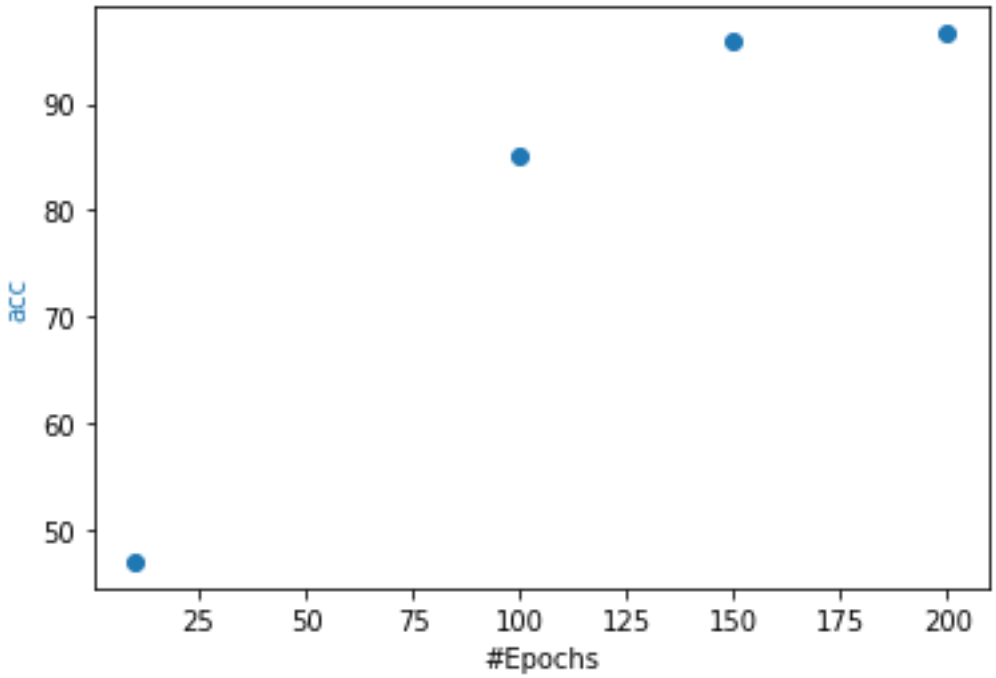
1) #Hidden Units:



2) #Hidden Layers:



3) #Epochs:



1) For #Hideen layers I tried 2,3,4. As shown in the plot also, the optimal performance is when the number of hidden layres are 2. So, my model has 2 hidden layers.

2) For #Hiddenunits I tried 32,64,128,256. As plot depicts the best performance is when hidden units are 128. I set it 150 as it was performing better than 128. So, my model has 150 hidden units.

3) For #epochs I tried on 10,100,150,200 as the no. of epochs is increasing the accuracy of the model is improving.

I used batch size of 128 and made batches using pytorch dataloader. The model is trained for 350 epochs. Loss function used is the cross entropy. #Epochs are 350.