Ass3, Q3:Multilayer Perceptron Regression

In this part of assignment, I had to implement a multi-layer Neural network for a regression problem. Then had to vary hyperparameter's and compare the loss value's.

Arya Marda

Problem description:

- Create a Multilayer Perceptron Regression class, and train it over the given Boston housing data.
- Also, Tune the Hyperparameter and find the combination that gives's minimum loss
- Also, have to log Several different loss metrics and finally report the best model.

Hyperparameter's tunning:

Hyperparameters to be varied:

- learningRate in [0.005, 0.001]
- HiddenLayers in [[6], [6,3], [6,5,7]]
- Number of epochs in [1500, 10000, 20000]:
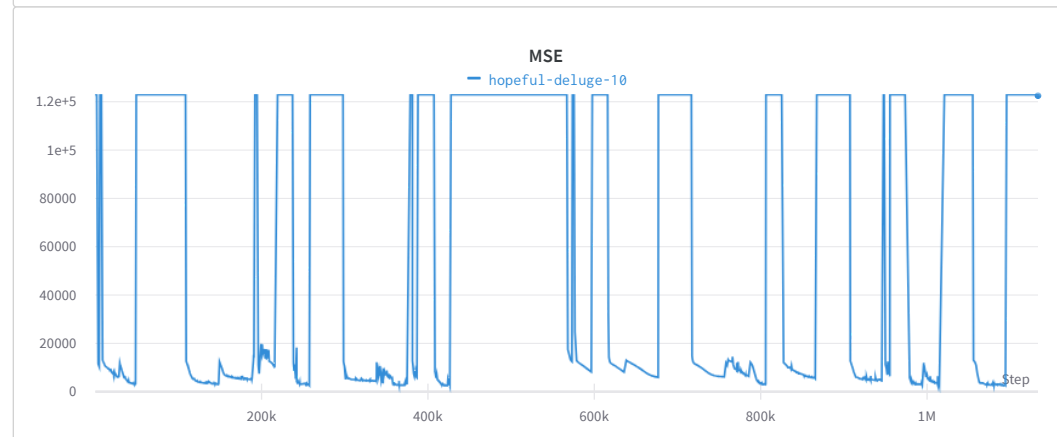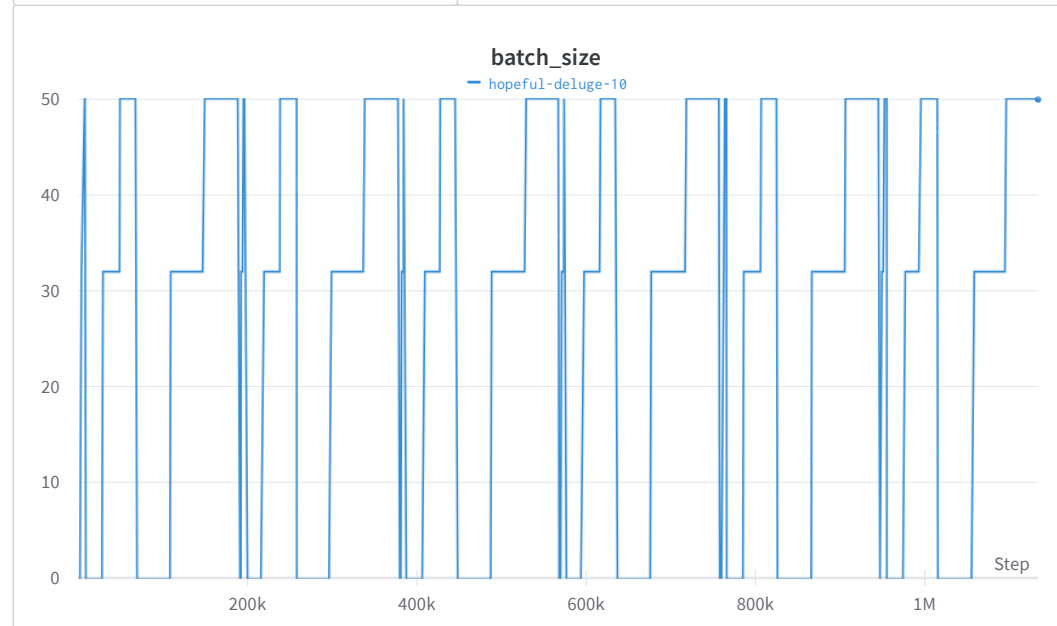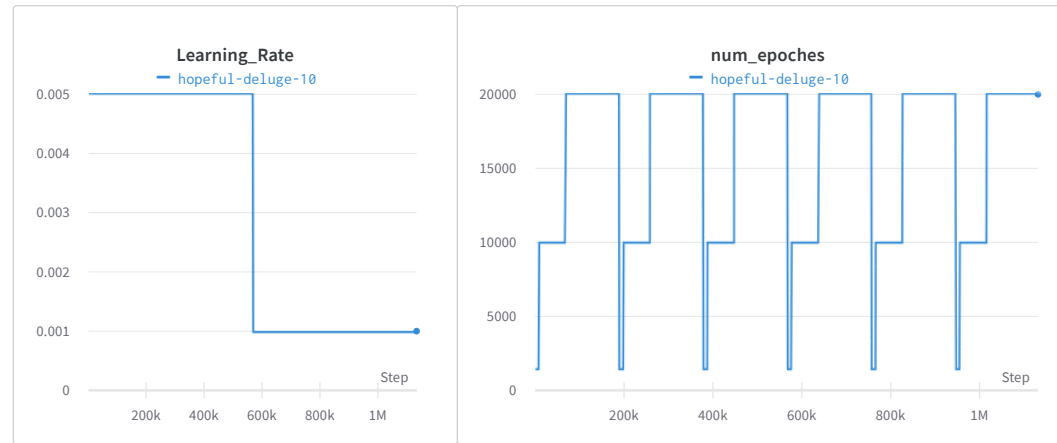- Batch_size in [0, 32, 50]:
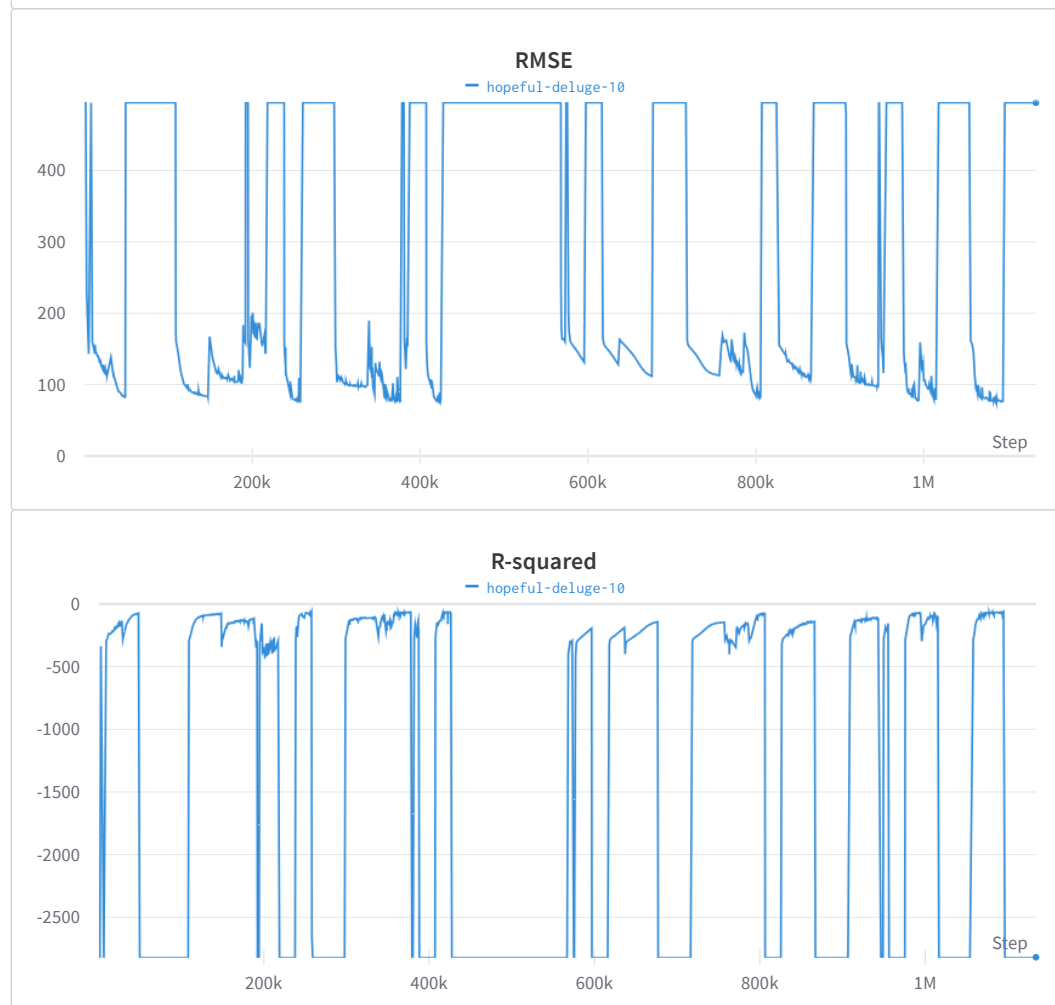
Points to understand before studying graphs:

- The activation function used is Relu, sigmoid and tanh cannot work as activation function because the range of value's of label's is not within -1 to1.
- Batch_Size = 0: Indicate's full batch gradient descent.
- In some run the Relu function can encounter all initial random output of the network as all negative values, in which case the Relu function make's them all zero, because of which the loss function will never converge.
- So basically, there can be some cases were the loss is non decreasing, constant.
- The variation of Hidden Layers couldn't be displayed in graph, hence for clarity I am displaying the variation's as for loop given below:

```
for learningRate in [0.005 , 0.001]:
    for hiddenLayers in [[6], [6,3], [6,5,7] ]:
        for num_epochs in [1500, 10000, 20000]:
            for batch_size in [0, 32, 50]:
                obj = Multi_layer_perceptron_wb(data = X_train , labels=y_train, hiddenLayers=hid
                obj.fit(num_epochs=num_epochs, batch_size=batch_size)
```

Graphs:

1. Below is a graph of variation of hyper-parameters
2. Following them are loss metric's valuese's.

Learning_Rate
— hopeful-deluge-10

num_epoches
— hopeful-deluge-10

batch_size
— hopeful-deluge-10

MSE
— hopeful-deluge-10

RMSE

R-squared

Finally the best model hyperparameters are :

- Epoche's: 10,000

- Learning Rate: 0.005

- Batch Size: 32

- Hidden Layer's: [6,5,7]