**DEEP LEARNING PROJECT**

**Working of script**

A neural network class with techniques for training, predicting, and visualizing learning progress is defined by the script that is given.   
The fit technique iterates over mini-batches and epochs to train the network. It monitors training and validation losses and updates parameters through forward and backward runs.   
The trained model is used by the `predict' technique to produce predictions.   
  
Plotting training and validation losses over epochs allows the plot learning approach to visually represent the learning process.   
The weights and biases of the network are initialized using the \_\_init\_layers function.   
Through the network levels, forward passes are computed using the \_\_forward method.   
While `\_\_calculate\_mse computes Mean Squared Error, \_\_calculate\_loss computes the loss between anticipated and actual values.   
Lastly, using computed gradients as a basis, the backward technique updates weights and biases using backpropagation.

Results for the configuration

For configuration 1 and 2

A graph of a graph

Description automatically generated with medium confidence

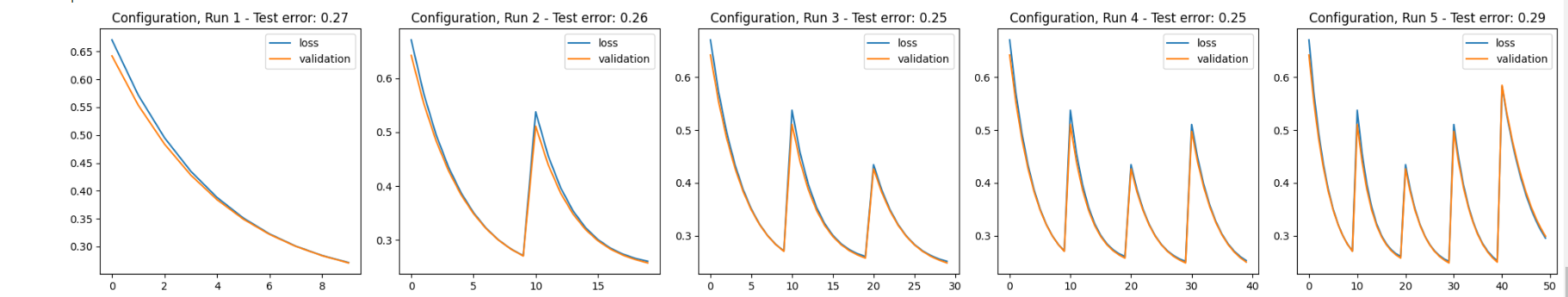
Configuration 1

A screenshot of a graph

Description automatically generated

Configuration 2

Here the configuration 2 performs well as compared with the configuration 1 in the sence of learning rate



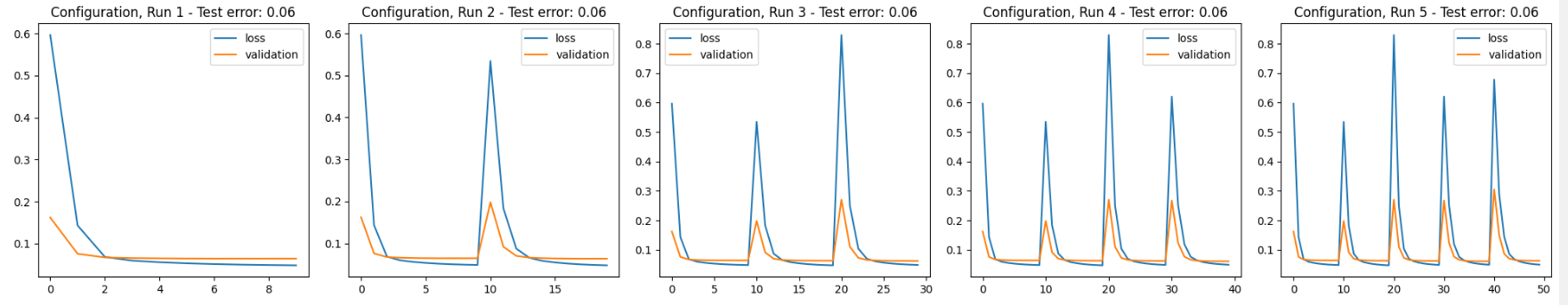
Configuration 3

A graph of a function

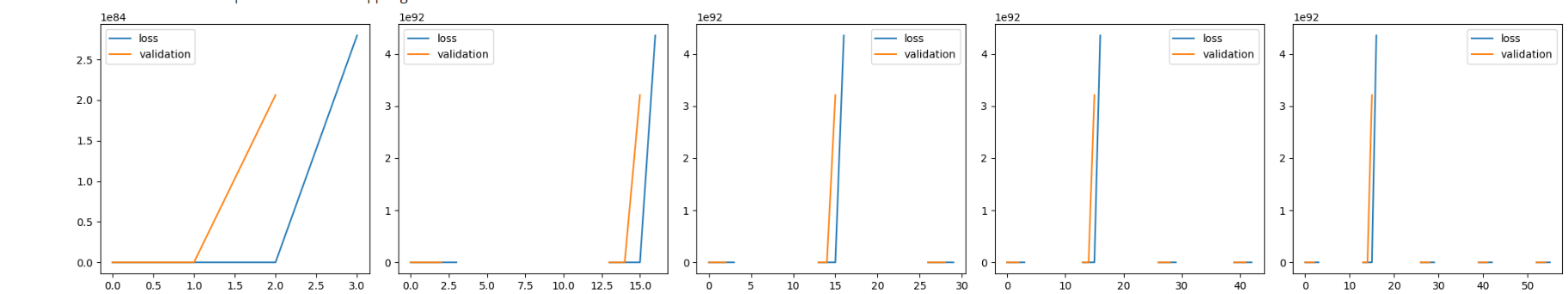
Description automatically generated with medium confidence

Configuration 4

In this I just added 3 neurons to the first and second layers I don’t find any significant difference after adding neurons to the given one but for the next configurations I wanted to change the learning rate



Configuration 5



Configuration 6

As i incressed learning rateit doesnot perform well but high learning rate the model will overshoot the minimum and result in divergence as a result it shows nan (not a number )

For the next the next step we need to decerese the learning rate

A graph of a graph

Description automatically generated with medium confidence

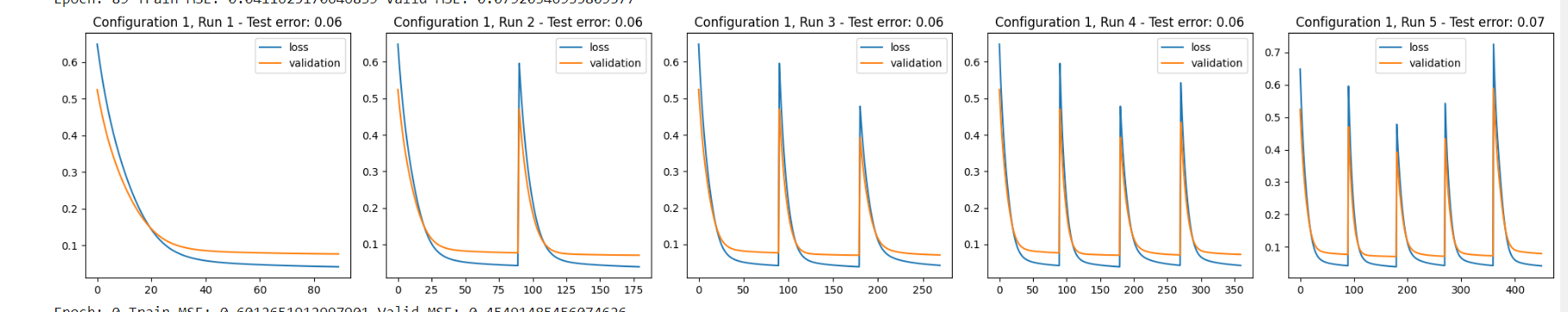
Configuration 7

A graph on a white background

Description automatically generated

Configuration 8

For configuration 7 and 8 I just increased the bandth width so that there is no significant differace but this is far better than the incressing learning rate .



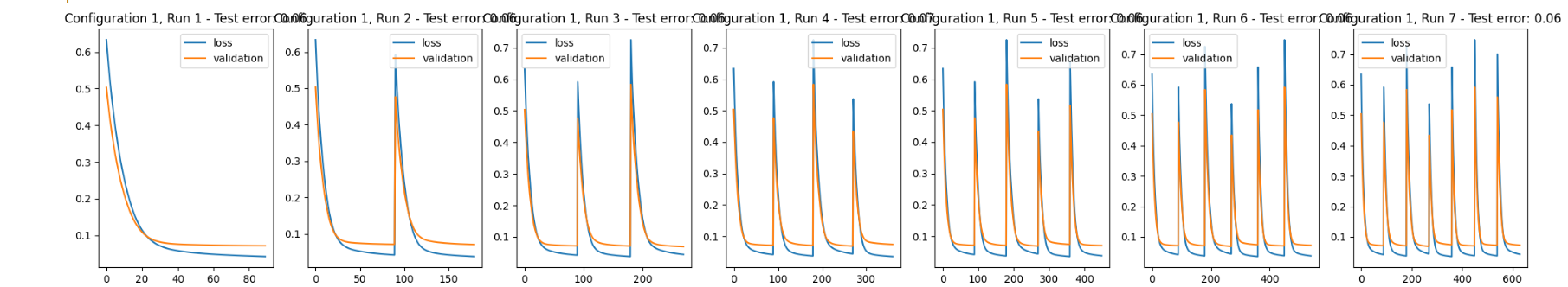
Configuration 9

A screenshot of a graph

Description automatically generated

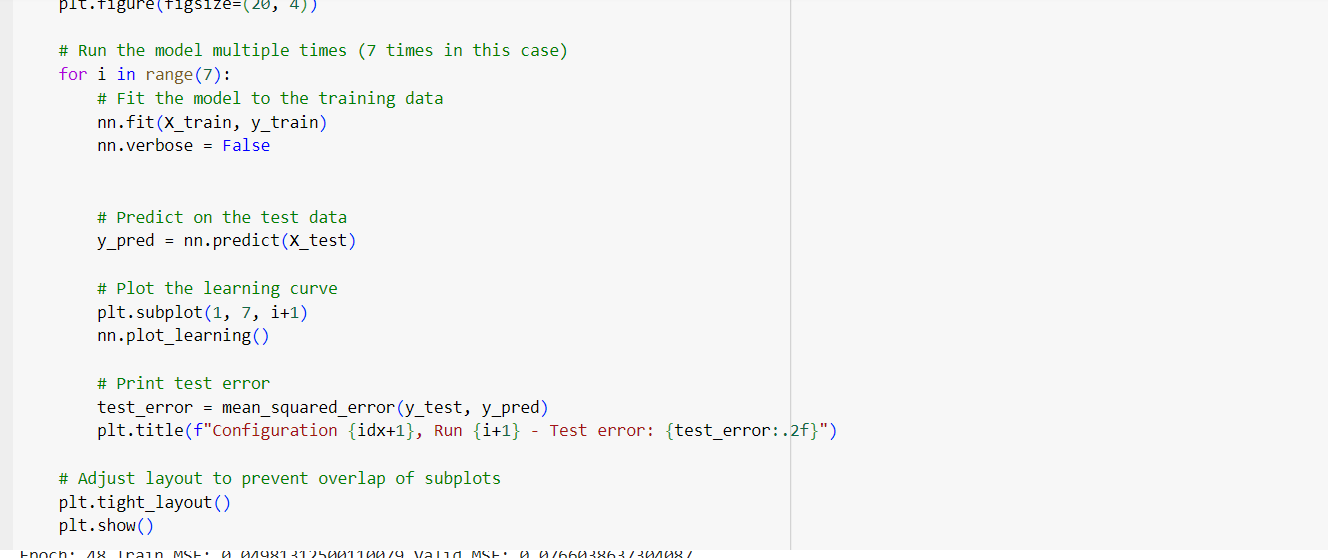
Configuration 10

For the configuration 9 and 10 in increased number of epochs this is really far better configuration so that this helps to give good test error as compared with all the configurations



Configuration 11

I incressed the running multiple times so that by running the code multiple times the configuration which gives the better configuration than compared to all



I turned off the vebose to get redused clutter

**Improvements**

To prevent overfitting, regularization techniques such as dropout or L2 regularization might be used.   
Monitoring and early termination: Put in place systems to keep an eye on the validation loss that occurs during training and to halt it as soon as it begins to rise.

**Limitations**

To prevent overfitting, regularization techniques such as dropout or L2 regularization might be used.   
Monitoring and early termination: Put in place systems to keep an eye on the validation loss that occurs during training and to halt it as soon as it begins to rise.