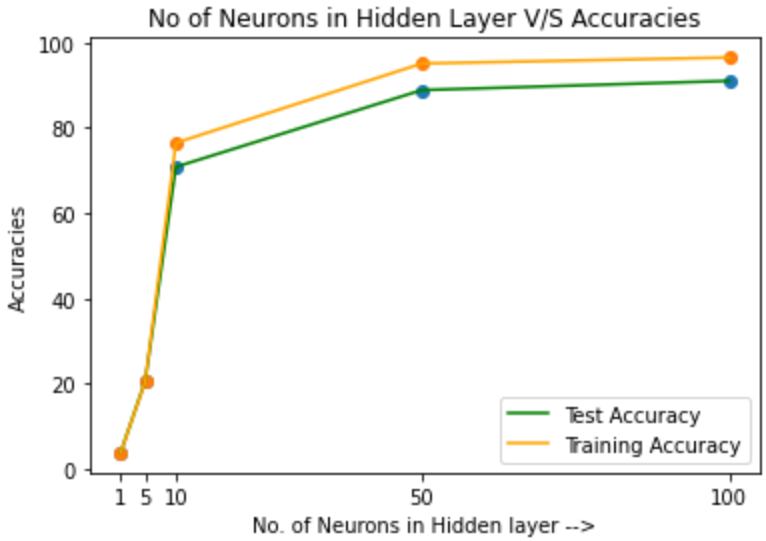
Que2 (A):-

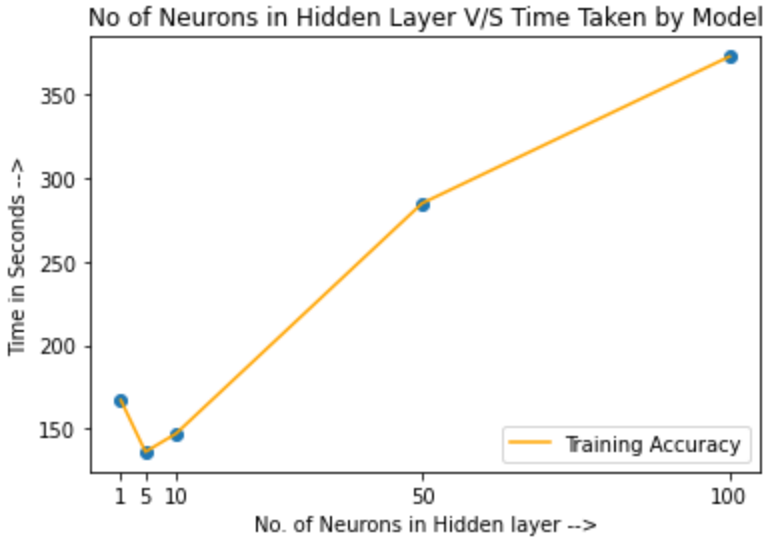
Neural network implemented.

Input Parameters:-

Hidden layer List : A list of number of neurons in each hidden layer
Batch Size: Mini batch size
Training Features : X
Train Labels : Y
Number of epochs
Activation Function : sigmoid or relu

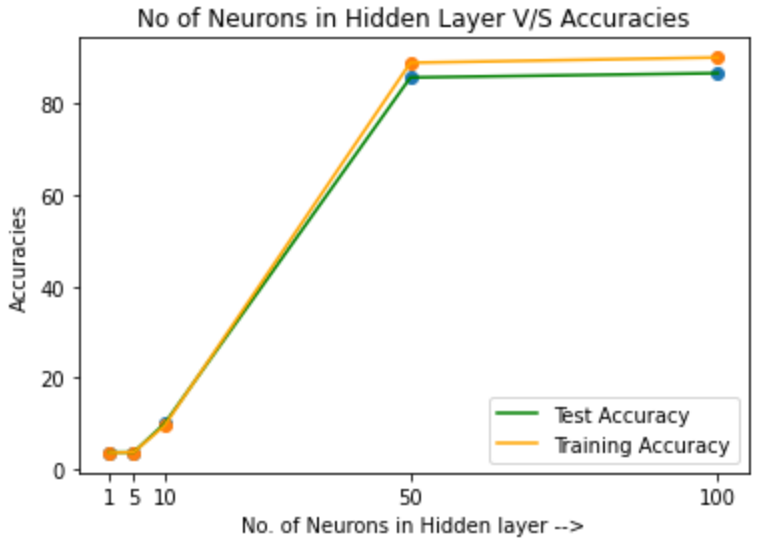
Que2 (B):-
Hidden Layer : [1,5,10,50,100]
Stopping Criteria: 1500 Epochs
Training Accuracies : [3.85, 20.48, 76.49, 95.09, 96.5]
Testing Accuracies : [3.85, 20.65, 70.86, 88.86, 91.02]
Time Taken : [167, 136, 147, 285, 373] seconds

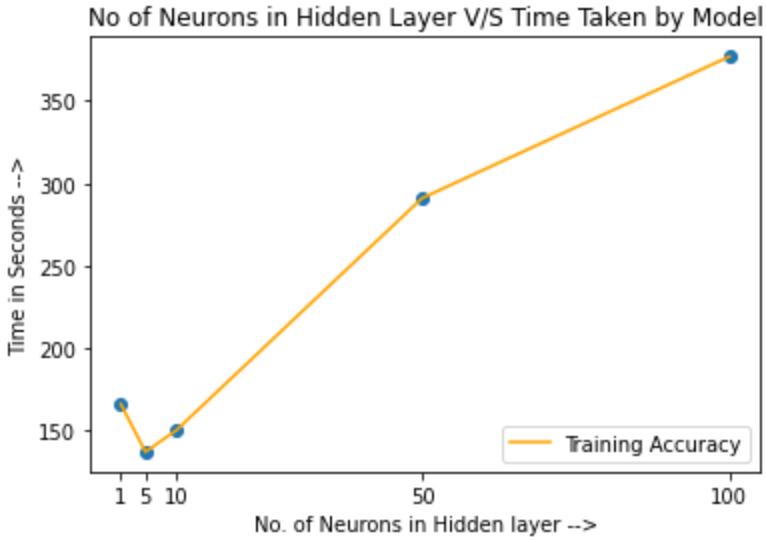




Observations:- If we increase the number of neurons in the hidden layer, time to train the model increases, accuracy on both test and train increases.It has been observed that hidden layers with 5 and 10 neurons took less time than 1 neuron.

Que_2 (C)
Learning Type = Adaptive
Hidden Layer : [1,5,10,50,100]
Stopping Criteria: 1500 Epochs
Training Accuracies : [3.64, 3.78, 9.56, 88.9, 90.05]
Testing Accuracies : [3.78, 3.77, 9.97, 85.68, 86.62]
Time Taken : [166, 137, 150, 291, 377] seconds





Observations:
Accuracies slightly have been reduced in comparison to fixed learning rate, and time taken is more to train the network.

Que_2(d)

a.

Hidden Layer = [100,100]
Activation: relu
Solver = sgd
Batch Size = 100
Epochs = 1500
Learning Type = Adaptive
Train Accuracy = 91.67%
Test Accuracy = 87.24%

b.

Hidden Layer = [100,100]
Activation: logistic
Solver = sgd
Batch Size = 100
Epochs = 1500
Learning Type = Adaptive
Train Accuracy = 80.41%
Test Accuracy = 78.76%

Observations:- Relu has performed better than sigmoid with [100,100] hidden layers, while a neural network with single layer with 100 neurons performed better as shown above.

Que_2(e):

a.

MLPClassifier
Activation = logistic
Time to train the model : 519seconds
Training data Accuracy : 87.15 %
Test Data accuracy : 88.30 %

Observations:- scikit MLPClassifier performed 8-10% on both the test and training data in comparison to part d.

b.

MLPClassifier
Activation = relu
Time to train the model : 1050 seconds
Training data Accuracy : 84.37 %
Test Data accuracy : 81.67 %

Observations:- With **relu** as activation function, the implementation in the part d performed better than the scikit MLPclassifier.