

## AI Assignment – 2 Report

**OUTLINE:** We need to train a multilayer perceptron with one hidden layer where the number of neurons in hidden layer varies between 5 to 8. There are two loss functions 'Sum of squared deviation' and 'Cross entropy loss function'. Sigmoid function is used as the activation function in both codes.

There are two stopping criterion -

1. Number of epochs = 100
2.  $||\Delta W|| < \epsilon = 0.001$

There are 16 attributes are given in the training set which are classified into 10 classes.

**OBSERVATIONS:** There are two codes. Both code initially have random generated weights b/w -0.5 to 0.5 . The first one has loss function as 'Sum of squared deviation' and the stopping criterion is 100 epochs.

The second one has the loss function as 'Cross entropy loss function' and the stopping condition as ' $||\Delta W|| < \epsilon$ '.

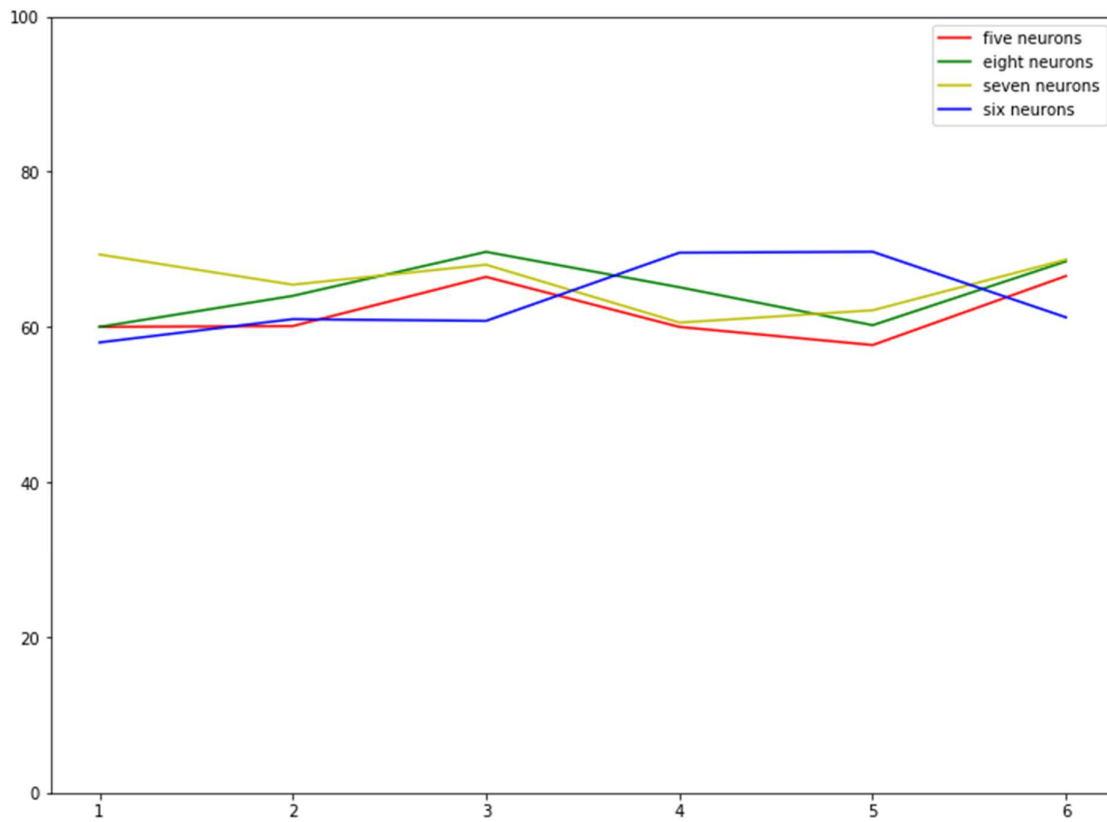
cross\_entropy.c :

- Loss function = Cross entropy function
- Stopping Condition = 100 epochs
- Learning rate = 0.001

At learning rate 0.001 and 100 epochs, we get a maximum accuracy of 50%-80%. A higher accuracy can be achieved by increasing the learning rate or number of epochs.

If we increasing learning rate by 10 times then the maximum accuracy obtained is 70% to 85% depending on the random weights generated.

We get different results of accuracy from 20% to 80% as initial weights are randomly generated. In the code, the number of neurons in the hidden layer is 7.



Graph b/w accuracy and number of neurons in hidden layer for code 1.

sum\_of\_sq\_dev.c :

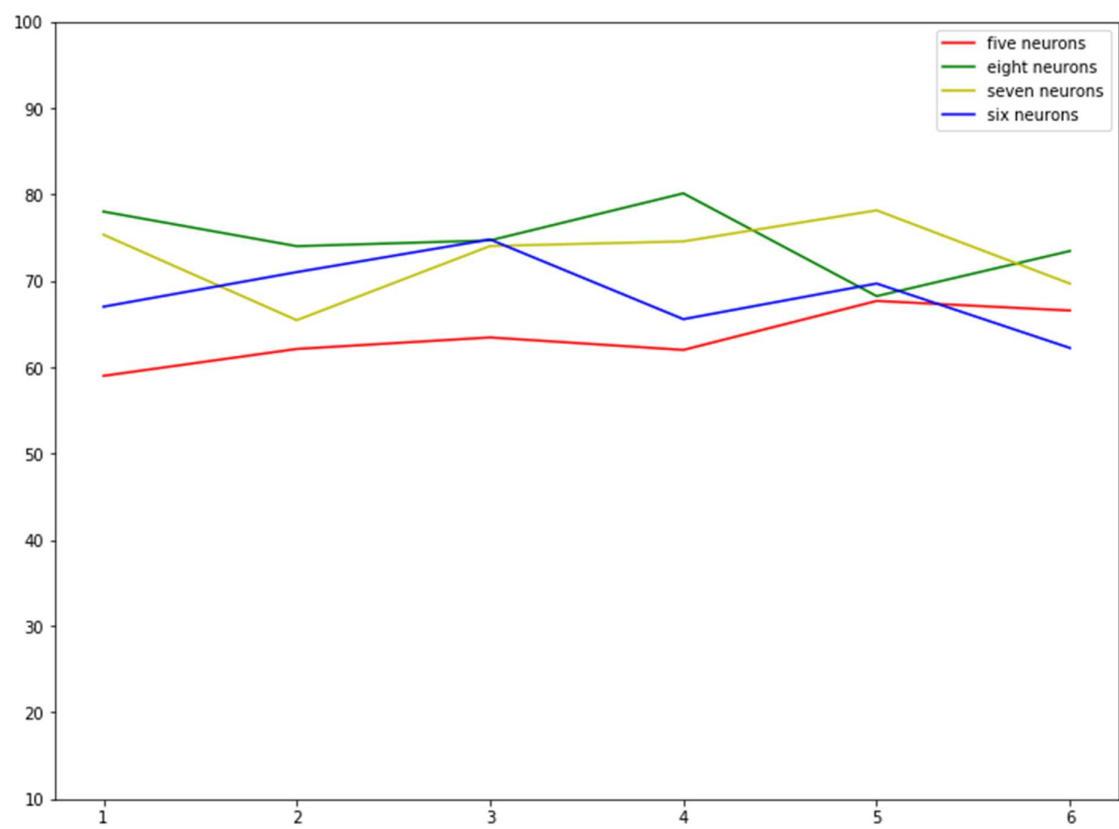
- Loss function = Sum of squared deviation
- Stopping condition -  $|| \Delta W || < \epsilon = 0.001$
- Learning rate = 0.01

In this code, the accuracy can be somewhat less as compared to previous one because of the problem of underfitting, i.e, inadequate training since the value of  $|| \Delta W ||$  was already too small.

The accuracy ranges between 40% to 75%.

However, the accuracy can be increased by decreasing the value of epsilon, which is the stopping condition and increasing the learning rate. This is because it increases the number of times the model is trained. The accuracy

ranges from 50% to 80%, for the given conditions. In the code, the number of neurons in the hidden layer is 7.



Graph b/w accuracy and number of neurons in hidden layer for code 2.