

I. When using 2 hidden layers with 16 neurons each, the accuracy on the training set and validation set increases continuously for the first 6 epochs, after which the training accuracy keeps rising but the validation accuracy keeps falling, indicating that the model is overfitting. In order to achieve the highest accuracy of 87.47%, we trained the model six more times before applying it to the test set.

II. Without changing any other hyperparameters, when  3 hidden layers are employed instead of 2 hidden layers, there is a very slight difference in validation and test accuracy, but no particular pattern is noticed. As a result, the data set used to develop the model reached saturation in accuracy at the first layer and no further improvement in accuracy is attainable for the present data set.

III. While keeping the number of hidden layers 2 constant and kept increasing the number of neurons in each layer to 32, 64, 128, 256 without making changes to any of the hyperparameters. When we did this the validation accuracy kept increasing with the addition of neurons in the hidden layers whereas no significance can be observed in the Test Accuracy.

IV. When binary-crossentropy is employed in place of the MSE loss function, there is a very slight decrease (approximately 0.5%) in both validation and test accuracy, indicating that binary cross-entropy is better to the MSE loss function for our IMDB database. However, the epoch with highest accuracy is 4 here.

V. There is a very slight decrease in both test and validation accuracy when the tanh activation function is employed in place of relu.

Concluding with the following by using different hyperparameters such as changing the number of hidden layers, number of neurons in each layers, changing optimizers and changing the loss function the maximum Test accuracy is almost the same.