

ELL409 REPORT

ASSIGNMENT 3

SARANSH AGARWAL 2019MT60763

NEURAL NETWORKS

To experiment with the use of Neural Networks for a multiclass classification problem

PART 1 A)

Standard backpropagation Technique

Here, we are going to learn a Neural Network classifier for these given images, starting with the raw pixels as input features, and thereby also to assess the usefulness of the different representations that Neural Network constructs. For this we have given dataset of 3000 images.

For this, we will take 75% of the set for training and remaining 25 % of the set for validation. Now we will tune hyperparameters and find out the best model

ACTIVATION FUNCTION- SIGMOID

NO HIDDEN LAYER

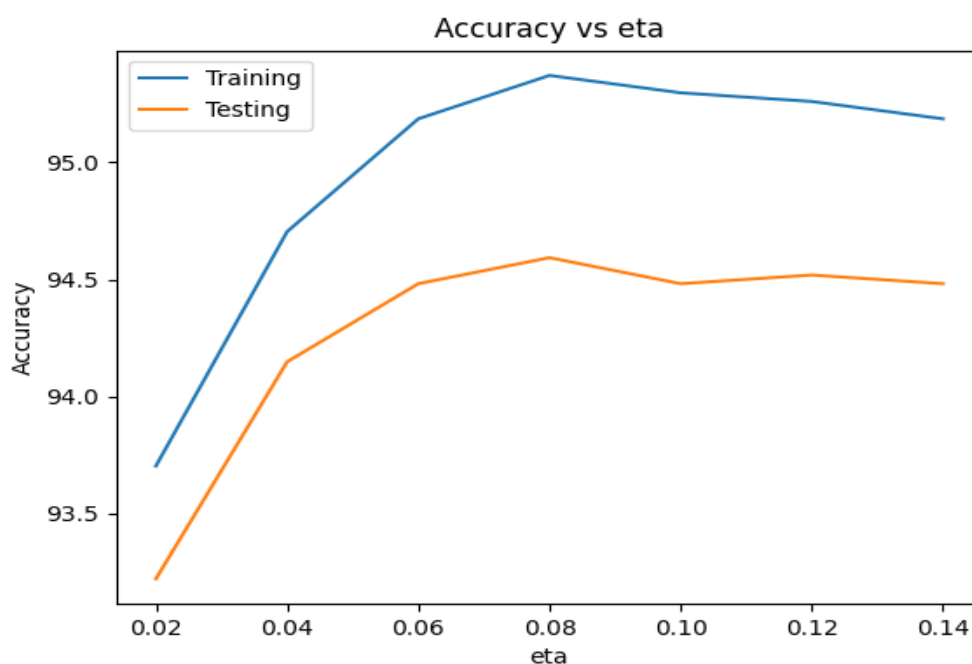


Figure 1: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 30. In the plot we can observe that at eta = 0.08 we get maximum testing accuracy = 94.5% after that model starts overfitting .

CONCLUSION

Optimal eta (learning rate) = 0.08

ONE HIDDEN LAYER

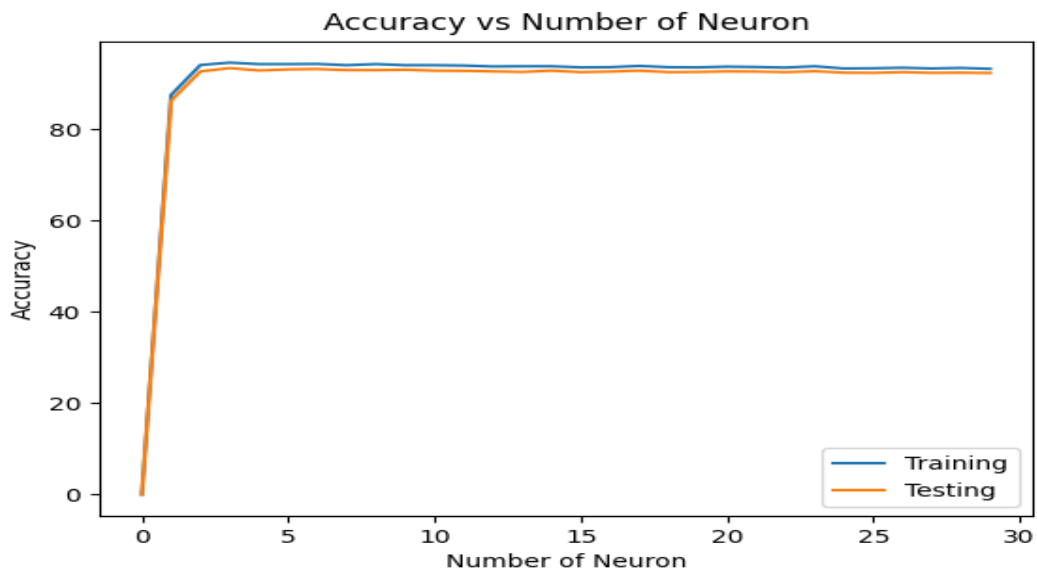


Figure 2: Accuracy vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs accuracy keeping number of iterations as 10 and $\eta = 0.08$. In the plot we can observe that at Number of Neuron in the Hidden layer = 30 we get maximum testing accuracy = 94% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the Hidden layer = 30

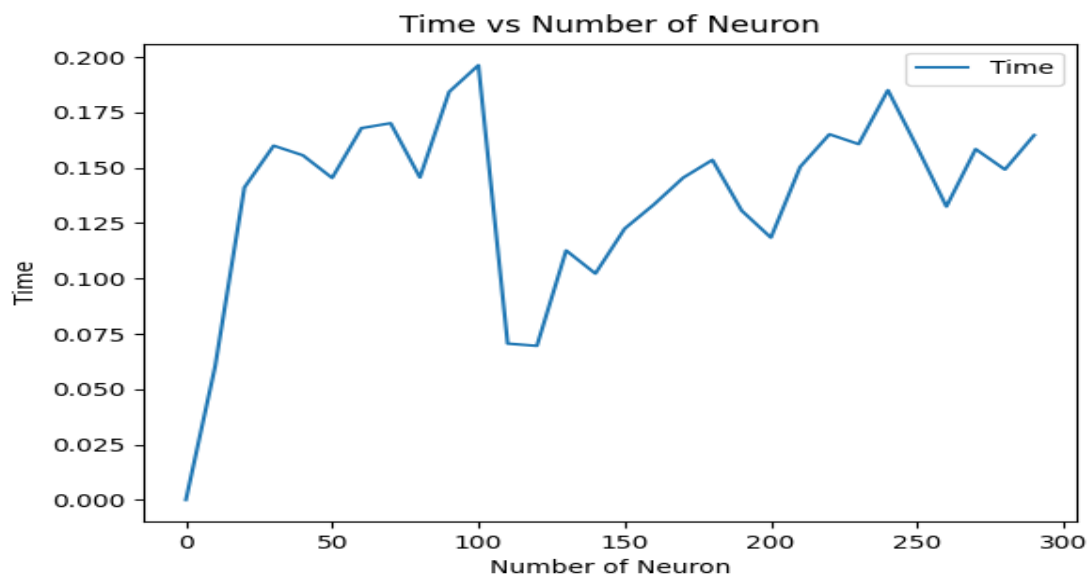


Figure 3: Time vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs Time keeping number of iterations as 10 and $\eta = 0.08$.

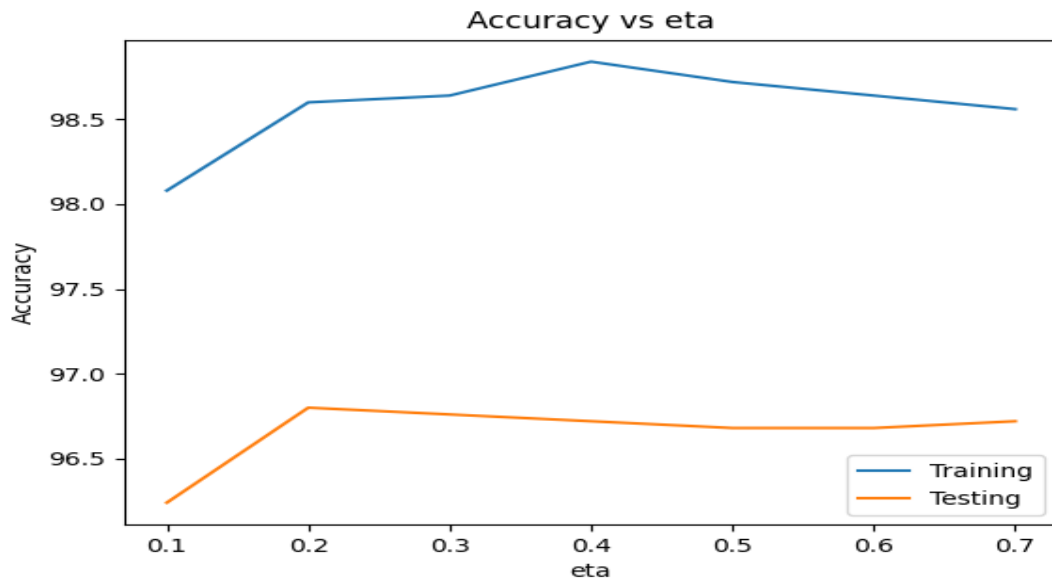


Figure 4: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 30 and Number of neurons in the hidden layer = 30. In the plot we can observe that at eta = 0.2 we get maximum testing accuracy = 96.8% after that model starts overfitting.

CONCLUSION

Optimal eta (learning rate) = 0.2

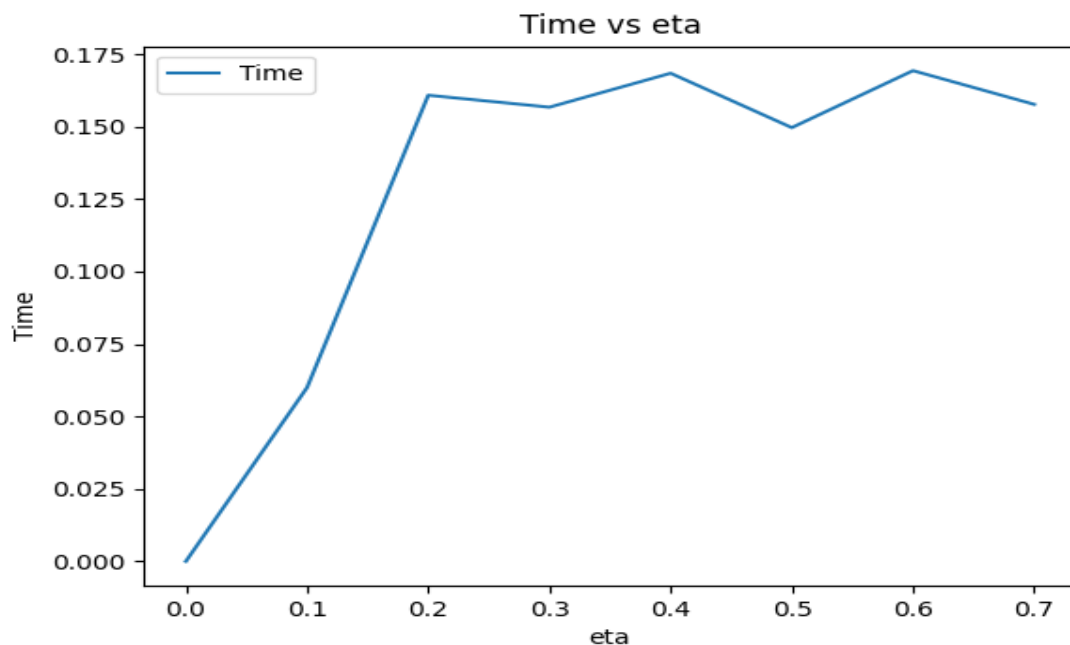


Figure 5: Time vs eta (learning rate) plot

The above graph shows variation of eta (learning rate) vs Time keeping number of iterations as 10 and Number of neurons in the hidden layer = 30.

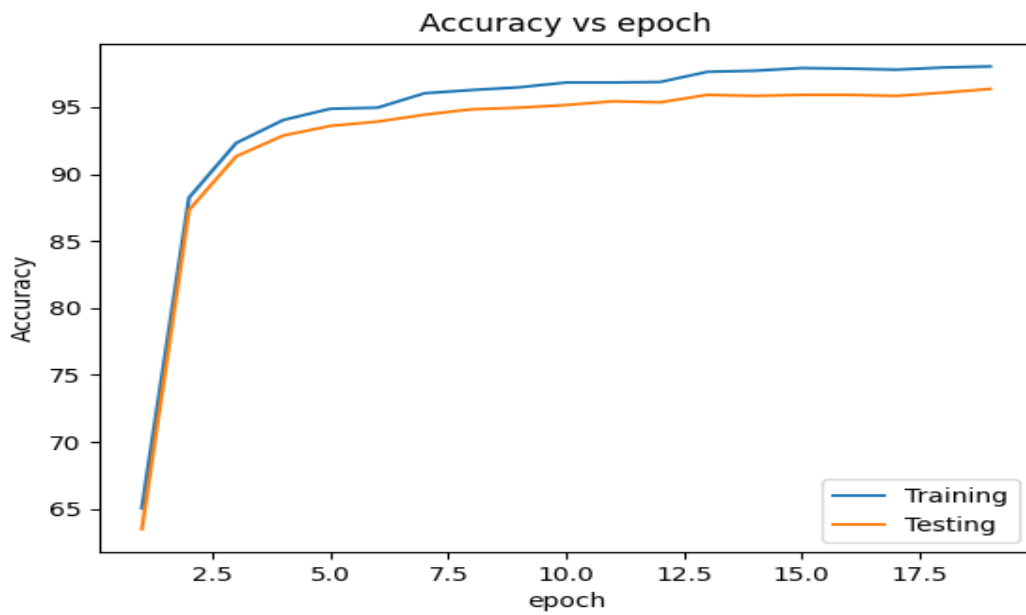


Figure 6: Accuracy vs epoch plot

The above graph shows variation of epoch (number of iterations) vs accuracy keeping Number of neurons in the hidden layer = 30 and eta = 0.2.

CONCLUSION

We can observe that as the epochs increases accuracy increases, it is because we are training more and more to tune the graph.

TWO HIDDEN LAYERS

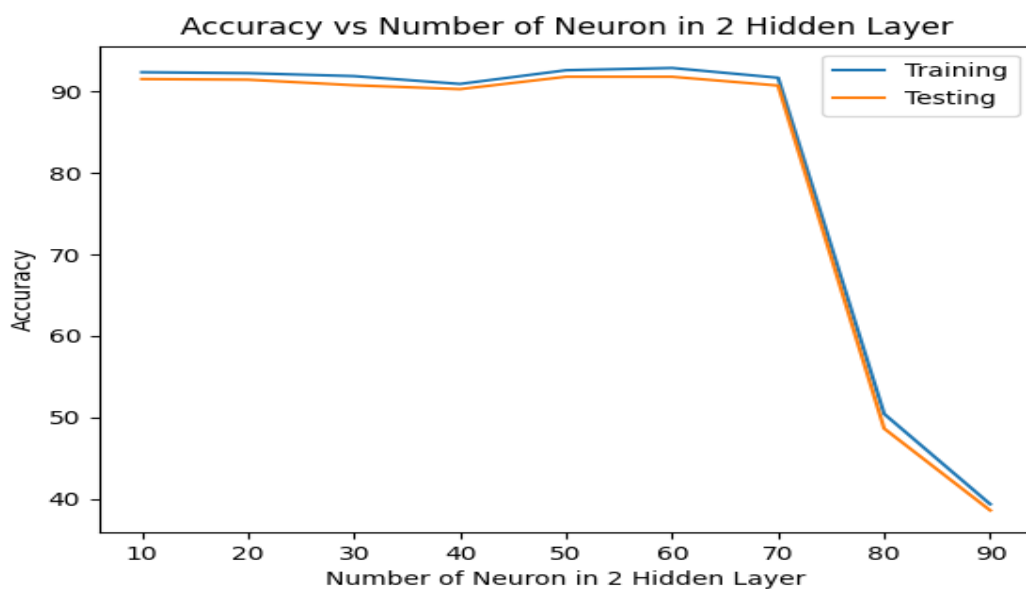


Figure 7: Accuracy vs Number of Neuron in second Hidden layer plot

The above graph shows variation of Number of Neuron in the second Hidden layer vs accuracy keeping number of neurons in the second Hidden layer = 30, number of iterations as 10 and eta = 0.08. In the plot we can observe that at Number of Neuron in the second Hidden layer = 60 we get maximum testing accuracy = 91.8% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the second Hidden layer = 60

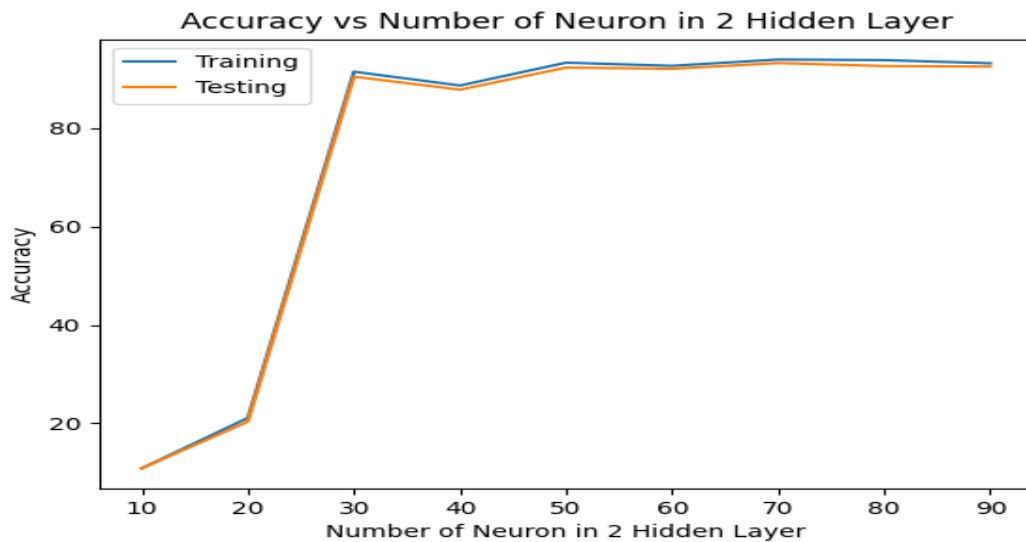


Figure 8: Accuracy vs Number of Neuron in first Hidden layer plot

The above graph shows variation of Number of Neuron in the first Hidden layer vs accuracy keeping number of neurons in the second Hidden layer = 60, number of iterations as 10 and eta = 0.8. In the plot we can observe that at Number of Neuron in the first Hidden layer = 50 we get maximum testing accuracy = 92.28% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the first Hidden layer = 50

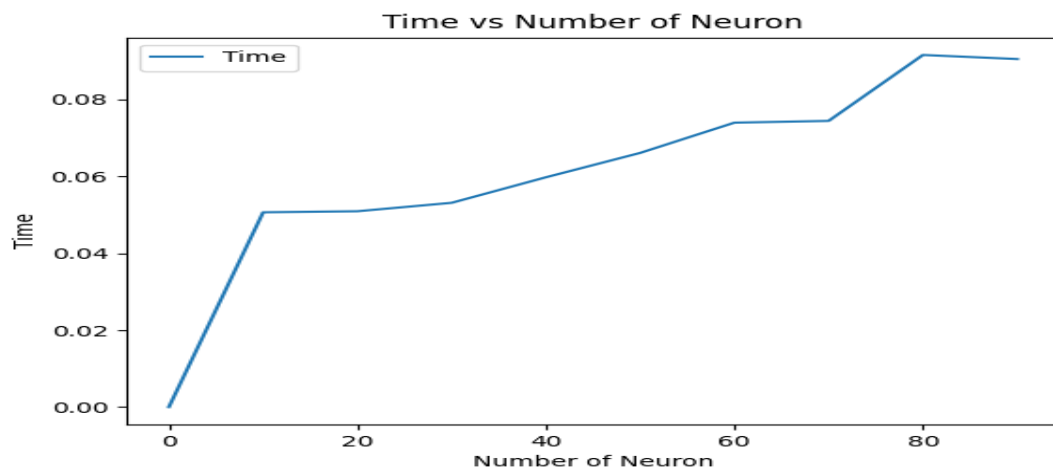


Figure 9: Time vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs Time keeping number of iterations as 10 and $\eta = 0.7$.

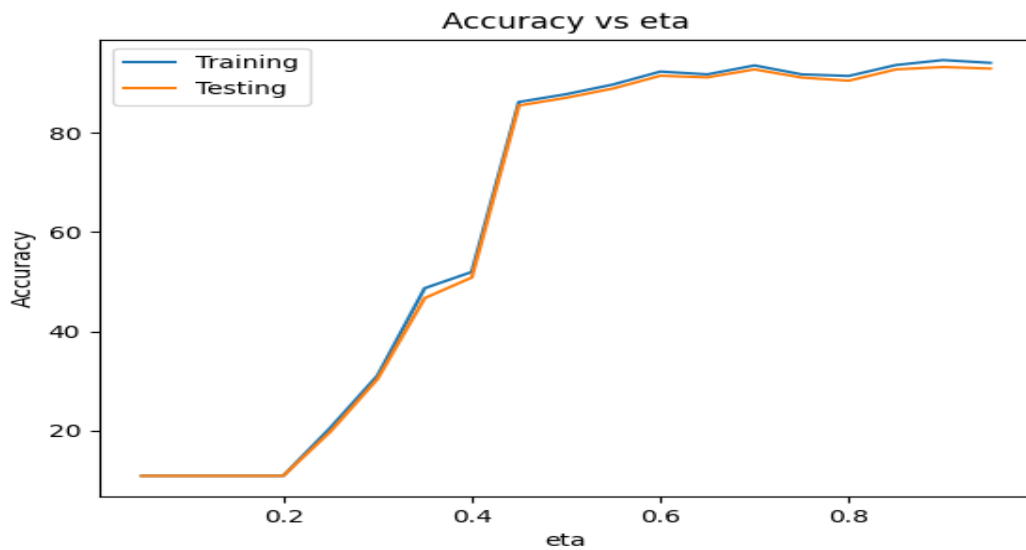


Figure 10: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 10 and Number of neurons in the first hidden layer = 50 and in second hidden layer = 60. In the plot we can observe that at $\eta = 0.7$ we get maximum testing accuracy = 94% after that model starts overfitting.

CONCLUSION

Optimal η (learning rate) = 0.7

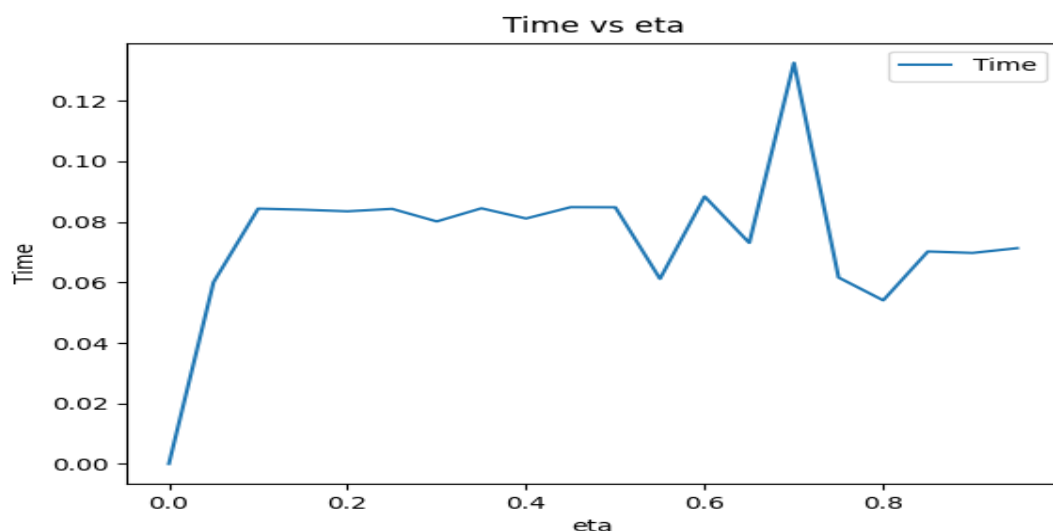


Figure 11: Time vs eta (learning rate) plot

The above graph shows variation of eta (learning rate) vs Time keeping accuracy keeping number of iterations as 10 and Number of neurons in the first hidden layer = 50 and in second hidden layer = 60.

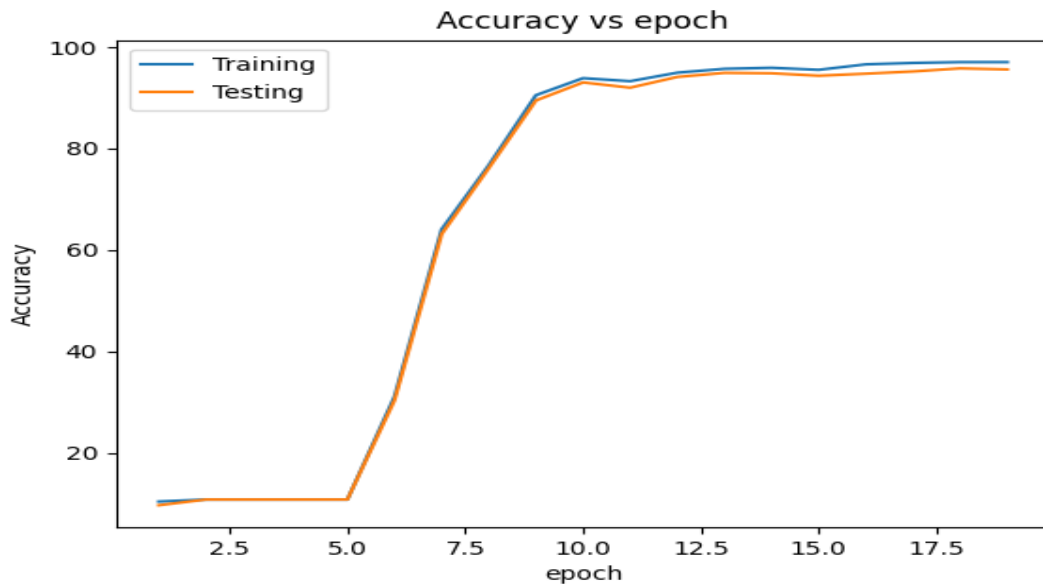


Figure 12: Accuracy vs epoch plot

The above graph shows variation of epoch (number of iterations) vs accuracy keeping Number of neurons in the first hidden layer = 50 and in second = 60 and eta = 0.7.

CONCLUSION

We can observe that as the epochs increases accuracy increases, it is because we are training more and more to tune the graph.

COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

NUMBER OF HIDDEN LAYERS	TRAINING SCORE	TESTING SCORE	BEST ETA	OPTIMAL NUMBER OF NEURONS
0	94.08	92.16	0.08	2
1	96.72	95.32	0.2	30
2	92.88	92.12	0.7	50,60

Figure 13: COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

CONCLUSION

We can observe from the above table that accuracy first increases from 92.16 with zero hidden layer to 95.32 with one hidden layer but then it decreases to 92.12 with two hidden layers as the model starts overfitting with increasing the number of hidden layers.

ACTIVATION FUNCTION- TANH

NO HIDDEN LAYER

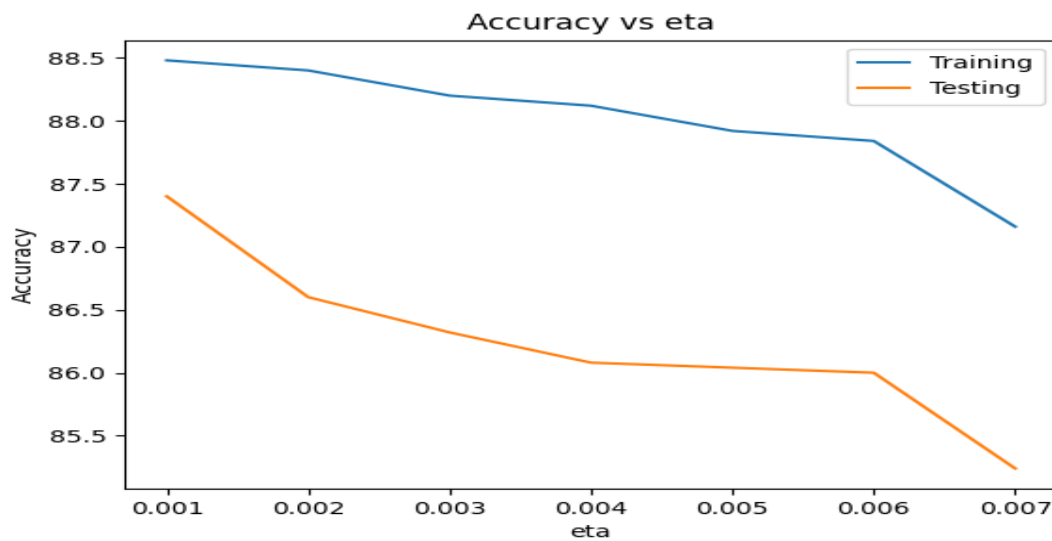


Figure 14: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 10. In the plot we can observe that at eta = 0.006 we get maximum testing accuracy = 86% after that model starts overfitting.

CONCLUSION

Optimal eta (learning rate) = 0.006

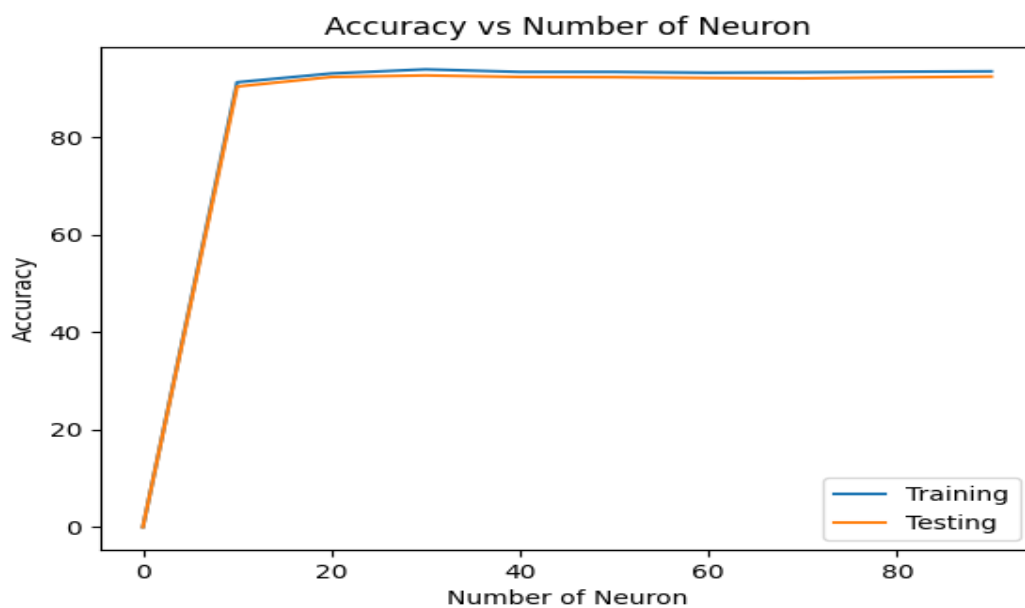


Figure 15: Accuracy vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs accuracy keeping number of iterations as 10 and $\eta = 0.006$. In the plot we can observe that at Number of Neuron in the Hidden layer = 30 we get maximum testing accuracy = 92.64% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the Hidden layer = 30

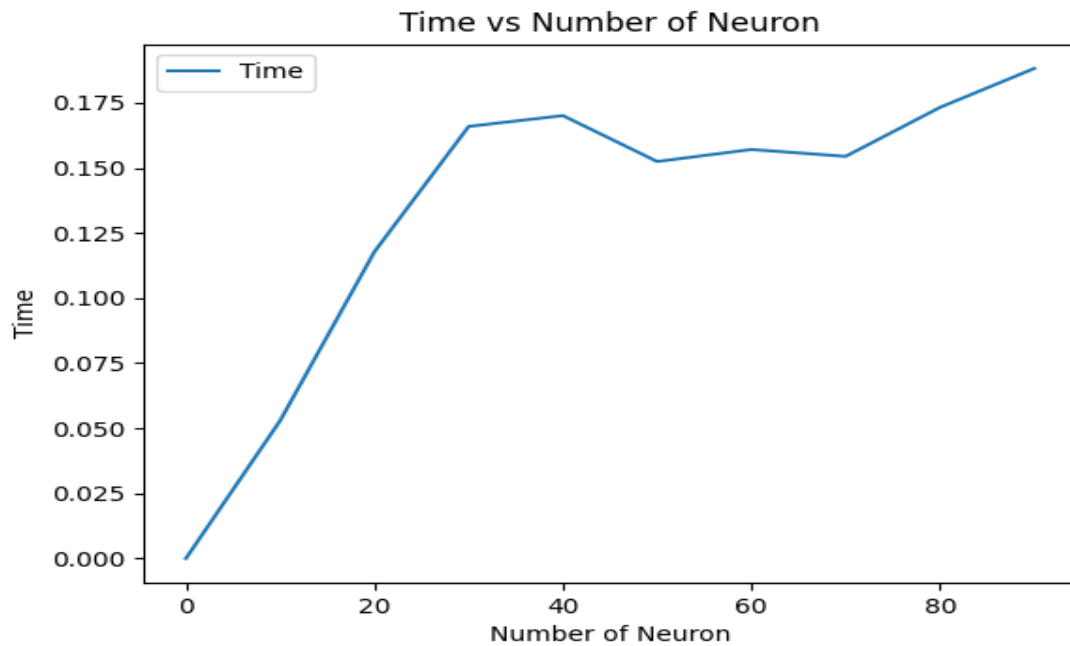


Figure 16: Time vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs Time keeping number of iterations as 10 and $\eta = 0.006$.

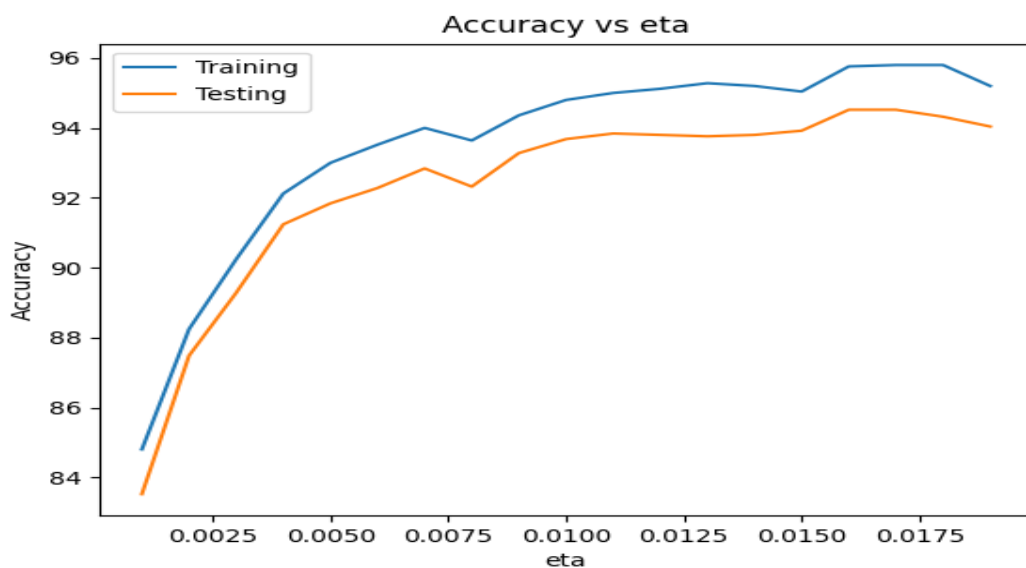


Figure 17: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 10 and Number of neurons in the hidden layer = 30. In the plot we can observe that at eta = 0.016 we get maximum testing accuracy = 94.54% after that model starts overfitting.

CONCLUSION

Optimal eta (learning rate) = 0.016

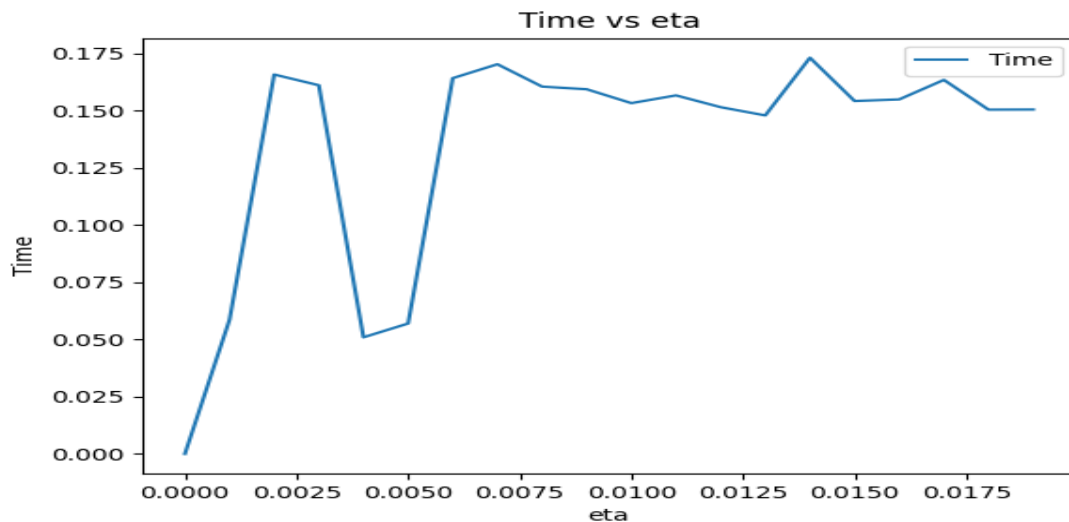


Figure 18: Time vs eta (learning rate) plot

The above graph shows variation of eta (learning rate) vs Time keeping number of iterations as 10 and Number of neurons in the hidden layer = 30.

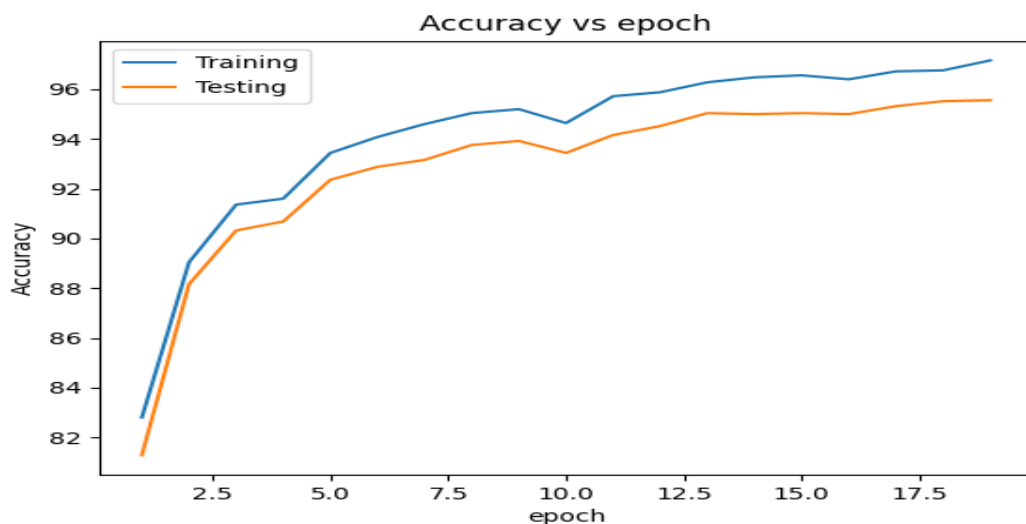


Figure 19: Accuracy vs epoch plot

The above graph shows variation of epoch (number of iterations) vs accuracy keeping Number of neurons in the hidden layer = 30 and eta = 0.016.

CONCLUSION

We can observe that as the epochs increases accuracy increases, it is because we are training more and more to tune the graph.

TWO HIDDEN LAYERS

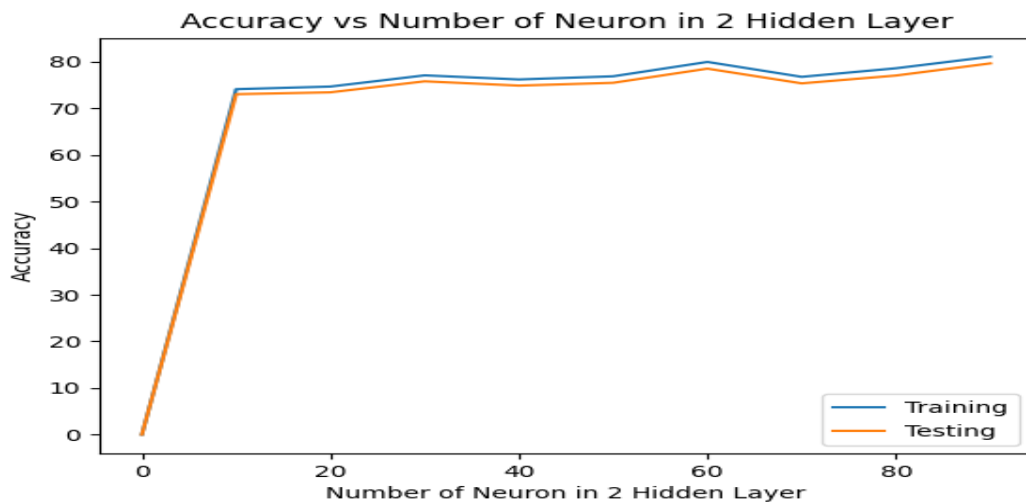


Figure 20: Accuracy vs Number of Neuron in second Hidden layer plot

The above graph shows variation of Number of Neuron in the second Hidden layer vs accuracy keeping number of neurons in the first Hidden layer = 30, number of iterations as 10 and eta = 0.006. In the plot we can observe that at Number of Neuron in the second Hidden layer = 60 we get maximum testing accuracy = 78.48% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the second Hidden layer = 60

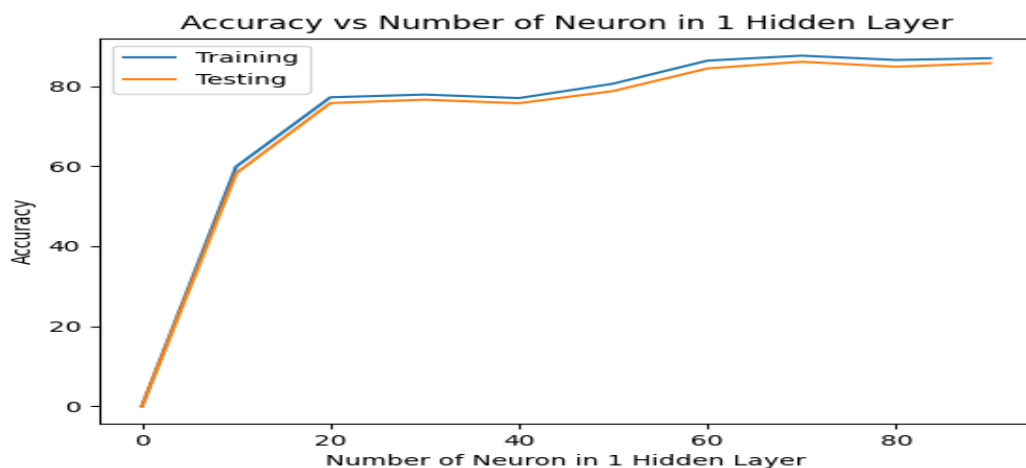


Figure 21: Accuracy vs Number of Neuron in first Hidden layer plot

The above graph shows variation of Number of Neuron in the first Hidden layer vs accuracy keeping number of neurons in the second Hidden layer = 60, number of iterations as 10 and eta = 0.006. In the plot we can observe that at Number of Neuron in the first Hidden layer = 30 we get maximum testing accuracy = 76.6% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the first Hidden layer = 30

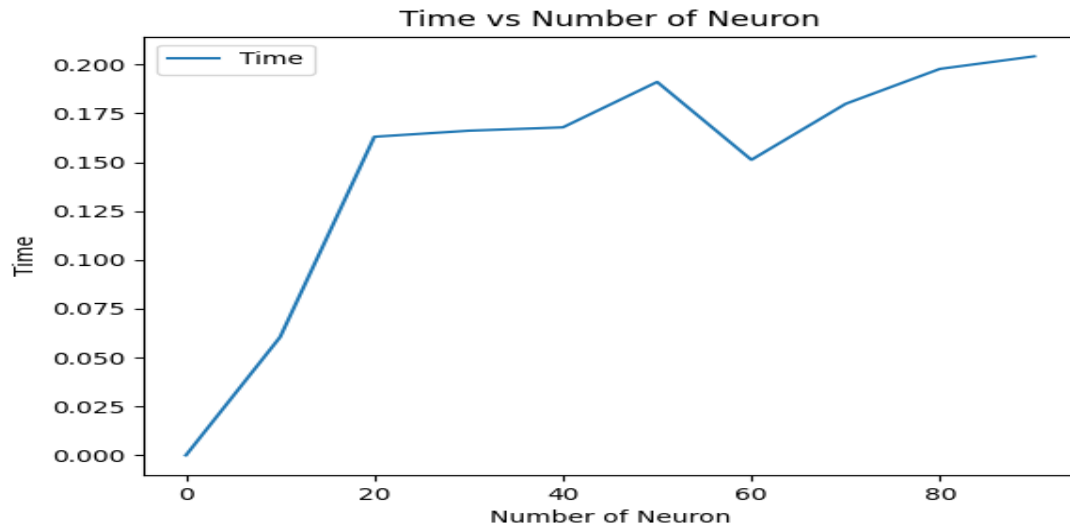


Figure 22: Time vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs Time keeping number of iterations as 10 and eta = 0.006.

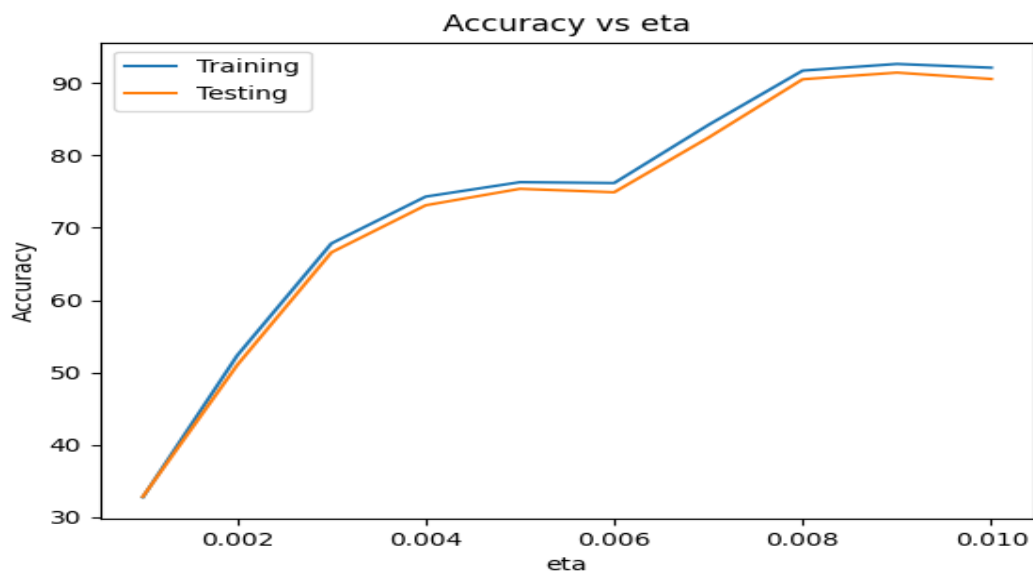


Figure 23: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 10 and Number of neurons in the first hidden layer = 30 and in second hidden layer = 60. In the plot we can observe that at eta = 0.005 we get maximum testing accuracy = 75.4% after that model starts overfitting.

CONCLUSION

Optimal eta (learning rate) = 0.005

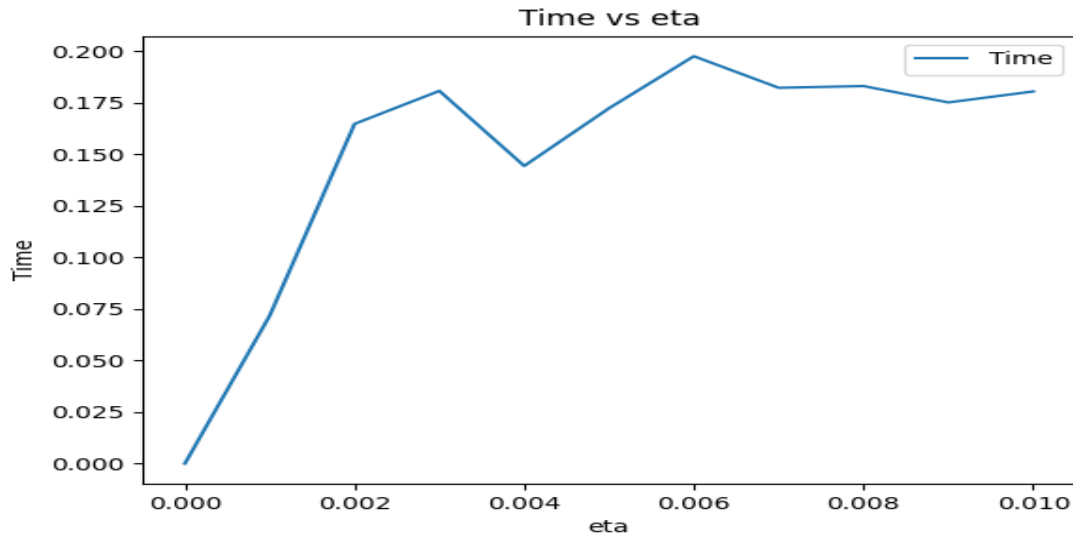


Figure 24: Time vs eta (learning rate) plot

The above graph shows variation of eta (learning rate) vs Time keeping accuracy keeping number of iterations as 10 and Number of neurons in the first hidden layer = 30 and in second hidden layer = 60.

COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

NUMBER OF HIDDEN LAYERS	TRAINING SCORE	TESTING SCORE	BEST ETA	OPTIMAL NUMBER OF NEURONS
0	87.52	85.64	0.006	-
1	95	93.67	0.016	30
2	75.64	74.64	0.005	30,60

Figure 25: COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

CONCLUSION

We can observe from the above table that accuracy first increases from 85.64 with zero hidden layer to 93.67 with one hidden layer but then it decreases to 74.64 with two hidden layers as the model starts overfitting with increasing the number of hidden layers.

ACTIVATION FUNCTIONN - SIGMOID

NUMBER OF HIDDEN LAYERS	TRAINING SCORE	TESTING SCORE	BEST ETA	OPTIMAL NUMBER OF NEURONS
0	94.08	92.16	0.08	-
1	96.72	95.32	0.2	30
2	92.88	92.12	0.7	50,60

Figure 26: COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

ACTIVATION FUNCTIONN - TANH

NUMBER OF HIDDEN LAYERS	TRAINING SCORE	TESTING SCORE	BEST ETA	OPTIMAL NUMBER OF NEURONS
0	87.52	85.64	0.006	-
1	95	93.67	0.016	30
2	75.64	74.64	0.005	30,60

Figure 27: COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

COMPARISION BETWEEN TANH LAYER AND SIGMOID LAYER:

The Above Table shows the training and testing score using sigmoid and tanh as activation functions. We can observe that accuracy is higher when we used sigmoid activation function.

USING TENSORFLOW

Now we train our model using the inbuild library TensorFlow by using the tune hyperparameters.

Using our own Implementation

ACTIVATION FUNCTIONN - SIGMOID

NUMBER OF HIDDEN LAYERS	TRAINING SCORE	TESTING SCORE	BEST ETA	OPTIMAL NUMBER OF NEURONS
0	94.08	92.16	0.08	=
1	96.72	95.32	0.2	30
2	92.88	92.12	0.7	50,60

Figure 28: COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

Using TensorFlow

ACTIVATION FUNCTIONN - SIGMOID

NUMBER OF HIDDEN LAYERS	TRAINING SCORE	TESTING SCORE	BEST ETA	OPTIMAL NUMBER OF NEURONS
0	91	90	0.08	=
1	92.11	91.15	0.2	30
2	88.52	88.16	0.7	50,60

Figure 29: COMPARISON FOR DIFFERENT NUMBER OF HIDDEN LAYERS

CONCLUSION

We can observe that we get higher training and testing accuracy by using our own implementation than the TensorFlow by using the tune hyperparameters we get by our own implementation.

PART 1 B)

Now we will consider the PCA-space representation of our data that we were provided for the previous assignment. This was a way of mapping the images to a lower-dimensional space, something that the neural net is also doing via its hidden units.

ACTIVATION FUNCTIONNN - SIGMOID

ONE HIDDEN LAYER

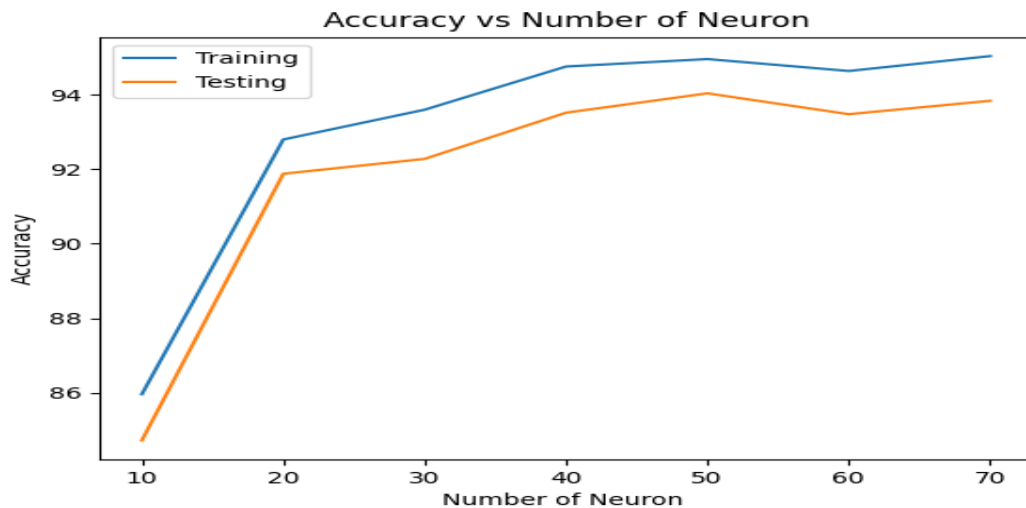


Figure 30: Accuracy vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs accuracy keeping number of iterations as 10 and eta = 0.8. In the plot we can observe that at Number of Neuron in the Hidden layer = 50 we get maximum testing accuracy = 94.04% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the Hidden layer = 50

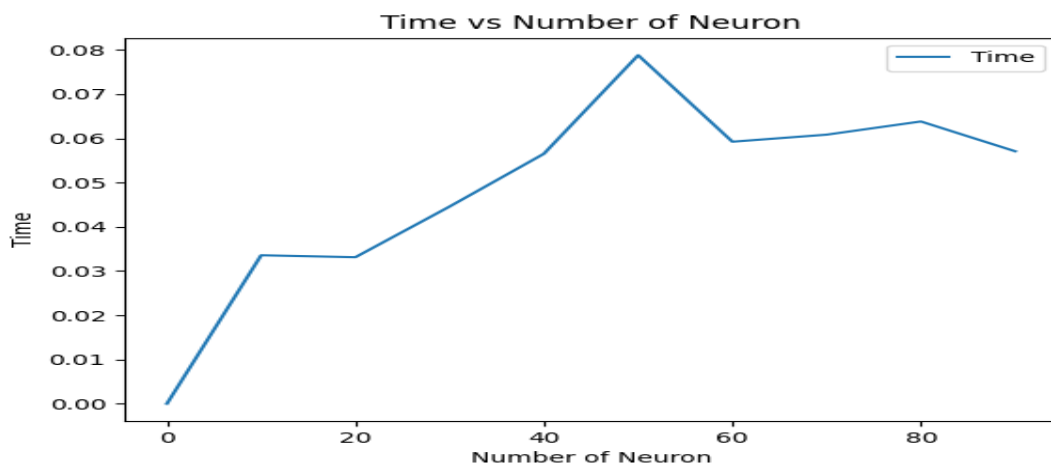


Figure 31: Time vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs Time keeping number of iterations as 10 and eta = 0.8.

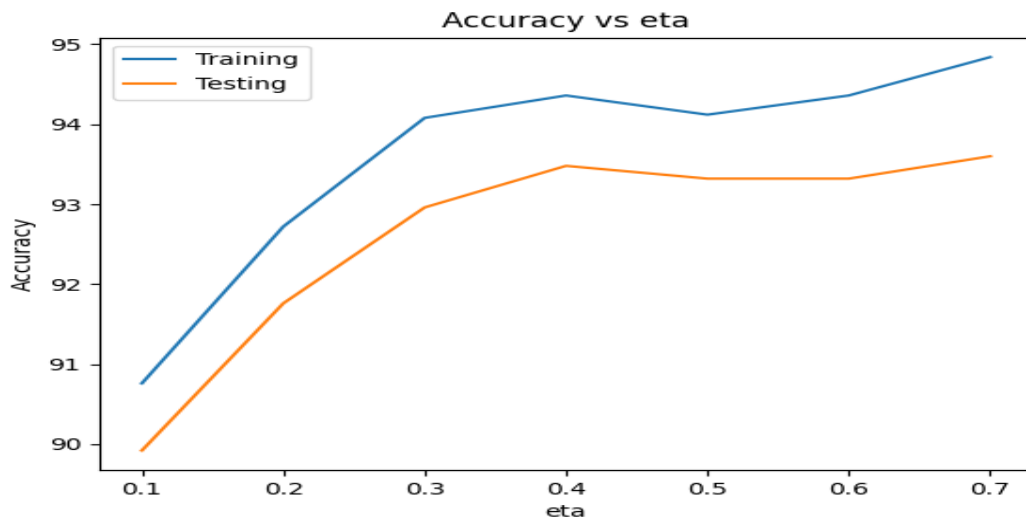


Figure 32: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 10 and Number of neurons in the hidden layer = 50. In the plot we can observe that at eta = 0.4 we get maximum testing accuracy = 93.12% after that model starts overfitting.

CONCLUSION

Optimal eta (learning rate) = 0.4

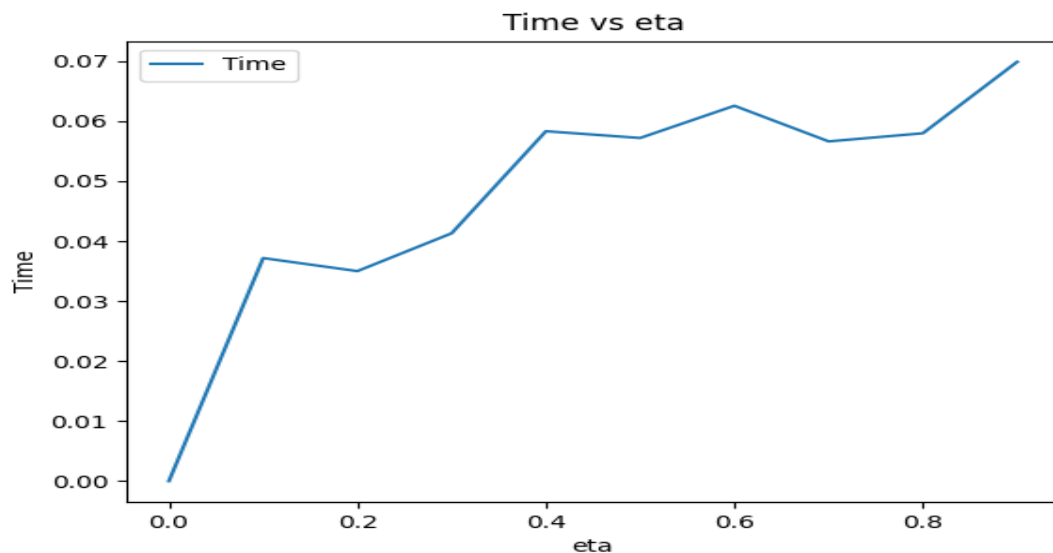


Figure 33: Time vs eta (learning rate) plot

The above graph shows variation of eta (learning rate) vs Time keeping number of iterations as 10 and Number of neurons in the hidden layer = 50.

ACTIVATION FUNCTIONN - TANH

ONE HIDDEN LAYER

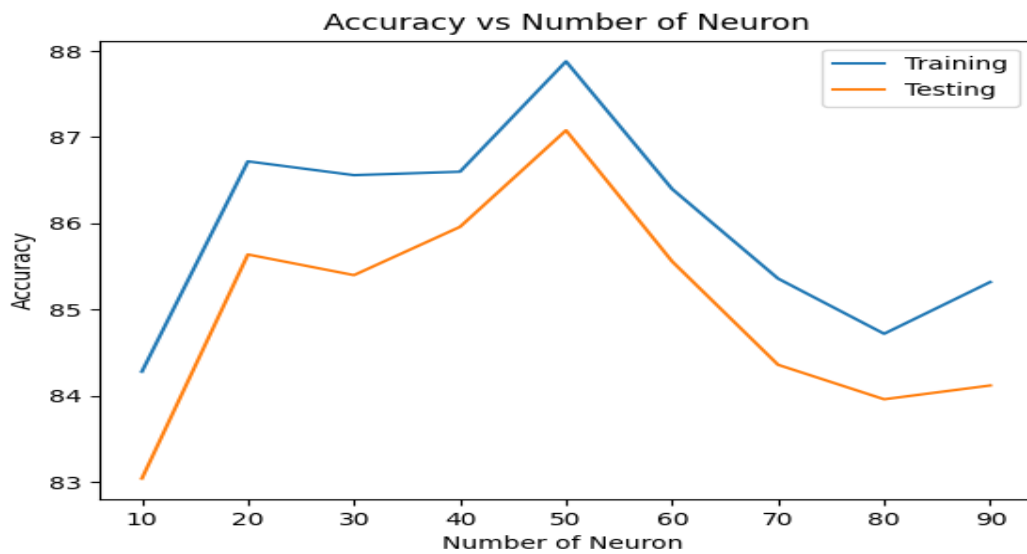


Figure 34: Accuracy vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs accuracy keeping number of iterations as 10 and $\eta = 0.008$. In the plot we can observe that at Number of Neuron in the Hidden layer = 50 we get maximum testing accuracy = 87.08% after that model starts overfitting.

CONCLUSION

Optimal Number of Neuron in the Hidden layer = 50

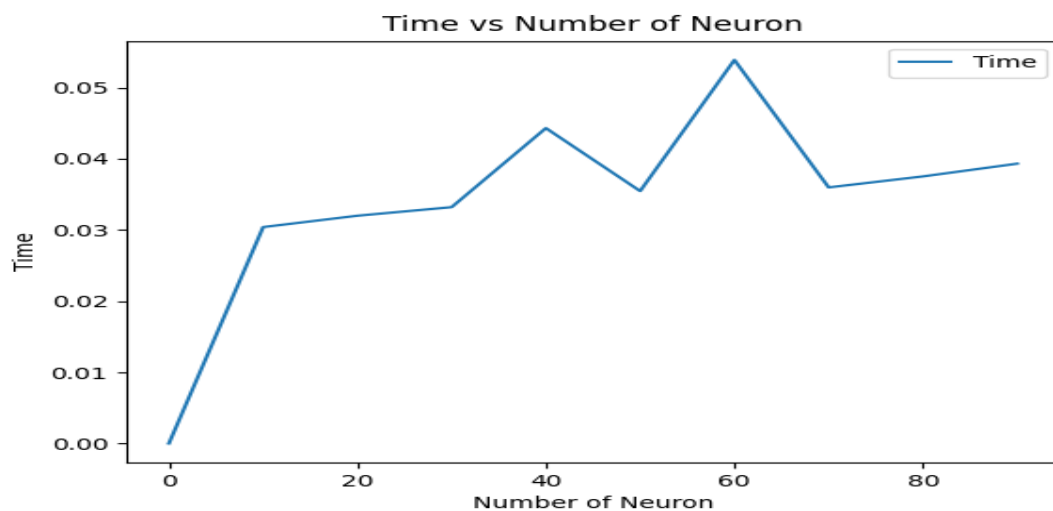


Figure 35: Time vs Number of Neuron in the Hidden layer plot

The above graph shows variation of Number of Neuron in the Hidden layer vs Time keeping number of iterations as 10 and $\eta = 0.008$.

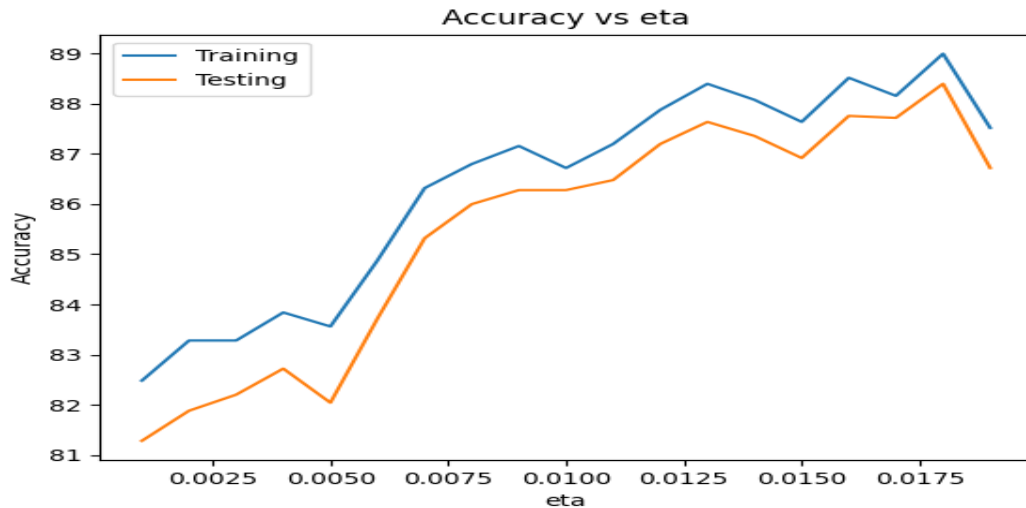


Figure 36: Accuracy vs eta (learning rate) plot

The above graph shows variation of eta vs accuracy keeping number of iterations as 10 and Number of neurons in the hidden layer = 50. In the plot we can observe that at eta = 0.013 we get maximum testing accuracy = 87.61% after that model starts overfitting.

CONCLUSION

Optimal eta (learning rate) = 0.013

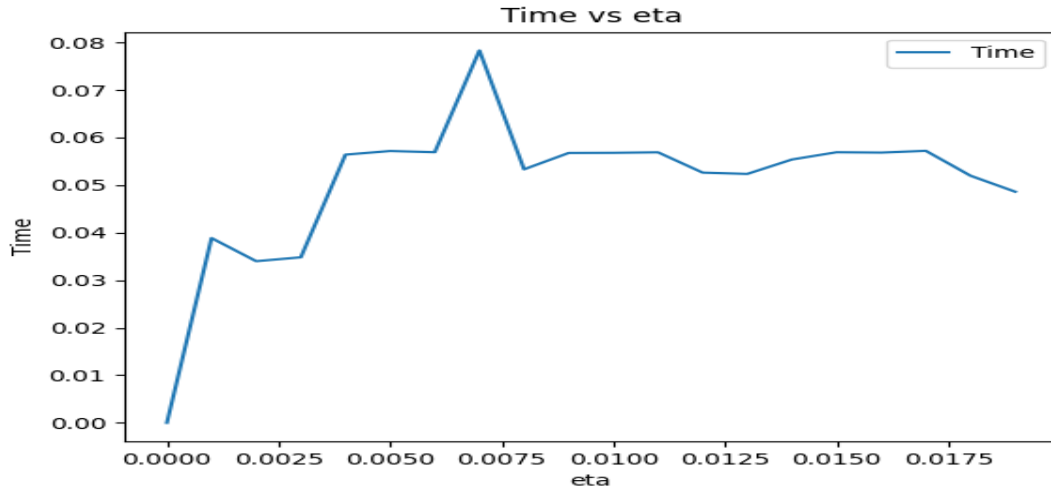


Figure 37: Time vs eta (learning rate) plot

The above graph shows variation of eta (learning rate) vs Time keeping number of iterations as 10 and Number of neurons in the hidden layer = 50.

CONCLUSION

- PCA features speed is much faster than raw pixels, it takes only half the time then the original pixels.
- Although the accuracy is not greater than raw pixels, as that provide more data.