

PROJECT REPORT
CSE574 MACHINE LEARNING

HAND-WRITTEN DIGITS CLASSIFICATION
ARTIFICIAL NEURAL NETWORKS AND KNN
CLASSIFIER

TEAM MEMBERS
Ankit Sarraf (50097190)
Karthick Krishna Venkatakrishnan (50098126)
Rahul Singh Chauhan (50097213)

Neural Networks:

Neural network is the system of programs and data structures that is analogous to the activity of a human brain. Neural network consists of large number of processors operating in parallel. Each processor in the neural network has its access to the data in the local memory. Similar to the function of a human brain, the neural network learns by example. The neural network trains its data and after training it can act upon the data to which it is tested.

K-Nearest Neighbor Classifier:

The KNN classifier acts on the notion that the elements which are separated by a farther distance are less likely to relate each other than the elements which are having lesser distance between them. In other words, the closer points are given more weightage than the farther points. All instances corresponds to a points in the Euclidean space. Classification is delayed till a new instance arrives. Classification is done by comparing the feature vector of different points. Target function may be distinct or real valued.

Choice of Hyper parameter of neural network:

The Hyper parameter (λ) value is changed from 0.1 to 100 and we can see that the training, validation and the testing accuracy are maximum for 0.8. We can also see that when the hyper parameter value is increased beyond 1, the training, validation and the testing accuracy starts decreasing.

Hyper parameter	Training Accuracy	Validation Accuracy	Testing Accuracy
0	94.34	93.87	93.94
0.1	94.48	93.96	94.19
0.2	94.35	93.87	94.17
0.3	94.92	94.32	94.5
0.4	93.94	93.64	93.77
0.5	94	93.52	93.96
0.6	94.53	94.43	94.32
0.7	94.64	94	94.5
0.8	95.25	94.6	94.74
0.9	94.79	93.79	94.63
1	94.75	94.14	94.53
1.2	94.11	93.23	93.72
5	94.208	93.34	93.92
10	93.542	92.84	93.53
50	93.192	92.83	93.27
100	92.16	91.91	92.37

The reason for such a variation is that the hyper parameter (λ) is a measure of how much approximation the neural network can make. Generalization of neural network increases with the increase in value of hyper parameter.

Hence, lower the values of lambda lead to over fit and higher values lead to under fit. At an optimal value of lambda (here 0.8), a best fit can be observed.

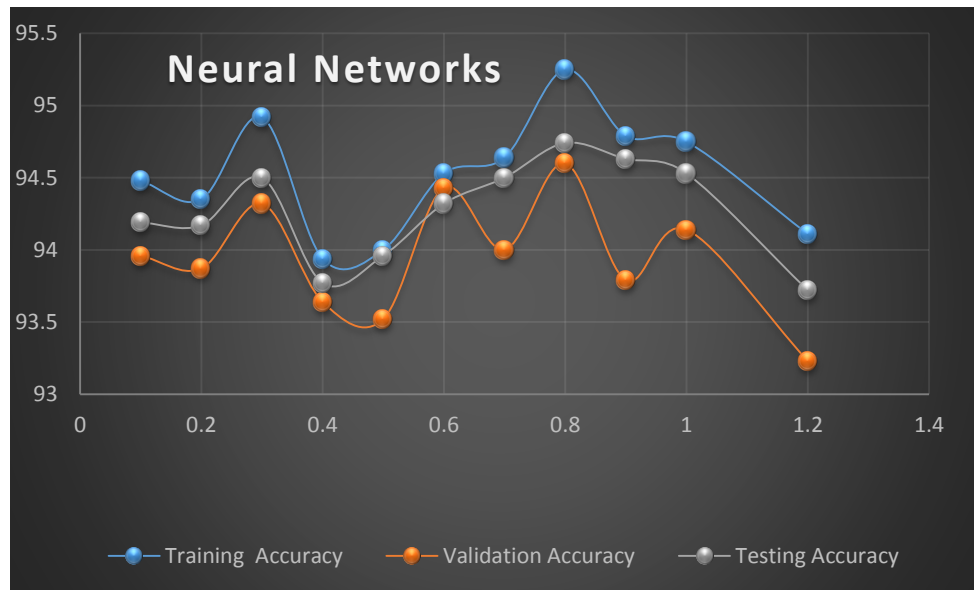


Figure 1: Variation of Accuracies (Y-Axis) with Lambda (X-Axis)

Next is the observations between the different values of hidden units and the corresponding training, validation, testing accuracies as well as the training time. Test depict that we have an optimum accuracy for training, validation and testing at 50 hidden units. Also, it can be noted that the training time first decreases, reaches a minimum and then it starts increasing. The observations depict the following. Initially, the decrease in accuracies is observed because there are lesser number of hidden nodes, and hence lesser parallelism. As the number of hidden nodes are increased the accuracies increase. After the optimal value, the accuracies start decreasing because complexity of network increases.

n_hidden	training accuracy	Validation accuracy	Testing accuracy	training time(seconds)
1	26.71	26.46	26.97	105
5	72.15	72.17	26.97	80.88
10	90.92	90.43	90.94	75.93
15	91.84	91.76	91.94	71.75
25	94.08	93.19	93.55	84.11
50	94.874	94.47	94.55	89.39268
75	93.688	93.06	93.69	101.7466
100	94.938	94.51	94.75	104.9056
125	94.938	94.51	94.75	104.9056
150	92.966	92.54	93.25	134.3064
200	92.154	92.05	92.38	176.5487

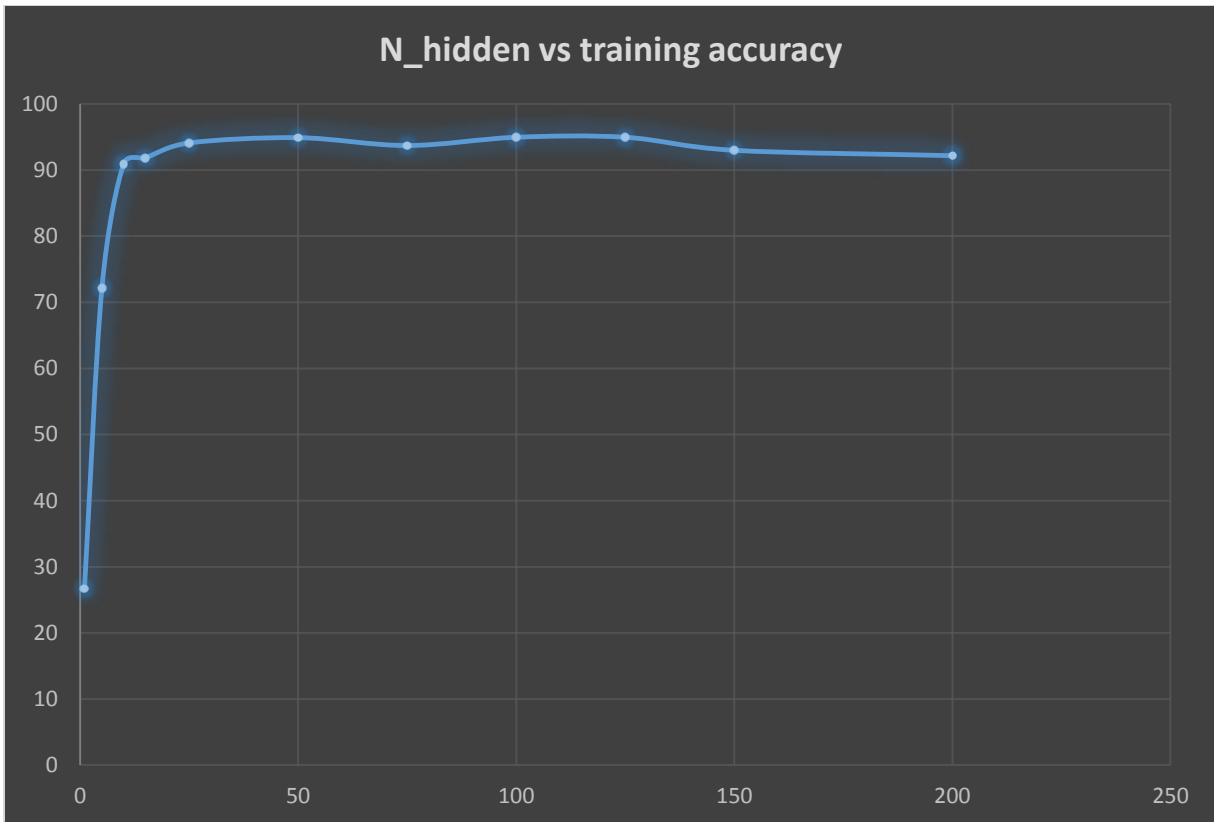


Figure 2: Variation of Training Accuracy (Y-Axis) with n_{hidden} (X-Axis)

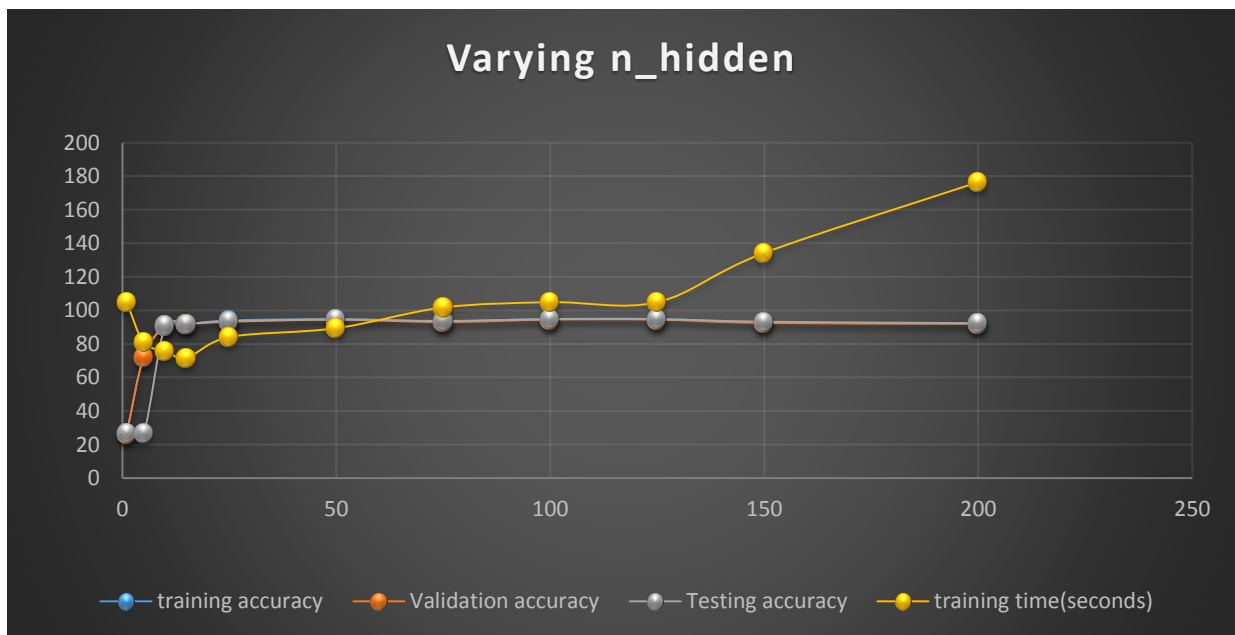


Figure 3: Variation of Training Accuracy & Training Time (Y-Axis) with n_{hidden} (X-Axis)

KNN-CLASSIFICATION:

For the KNN classifier, the validation accuracy and the testing accuracies are maximum for $k = 3$.

K-parameter	Validation accuracy	Testing accuracy
1	96.31	96.13
2	97.41	96.8
3	97.6	97.04
4	97.37	96.99
5	96.99	96.87
6	97.01	96.88

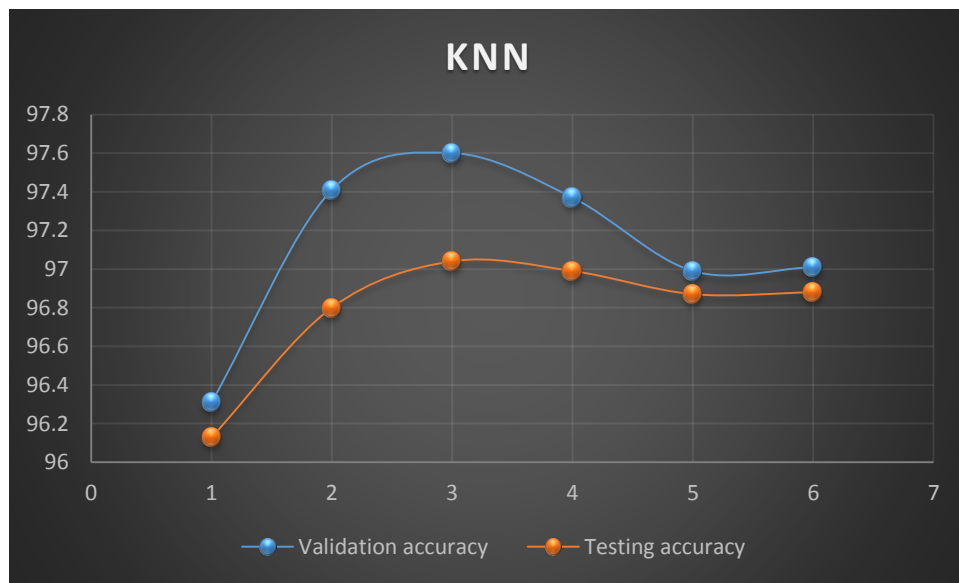


Figure 4: Variation of KNN-Classifiers Accuracies (Y-Axis) with k (X-Axis)

Comparison between the performance of Neural Network and KNN:

The neural network tends to take time to train its data. But after training it takes lesser time to test data. On the other side the KNN classifier does not need to train data but tends to take way more time to test data. The accuracy of KNN is more than that of the neural network and the learning time of the KNN is more than that of the neural network.

The neural network takes only 3 minutes for training even though it spends time to train data. But the KNN classifier takes an overall time of 13.5 minutes on our system even though it does not need to train data. The reason for this time consumption in case of KNN is because it needs to calculate the Euclidean distances between the given image and all the images in the dataset for classification.

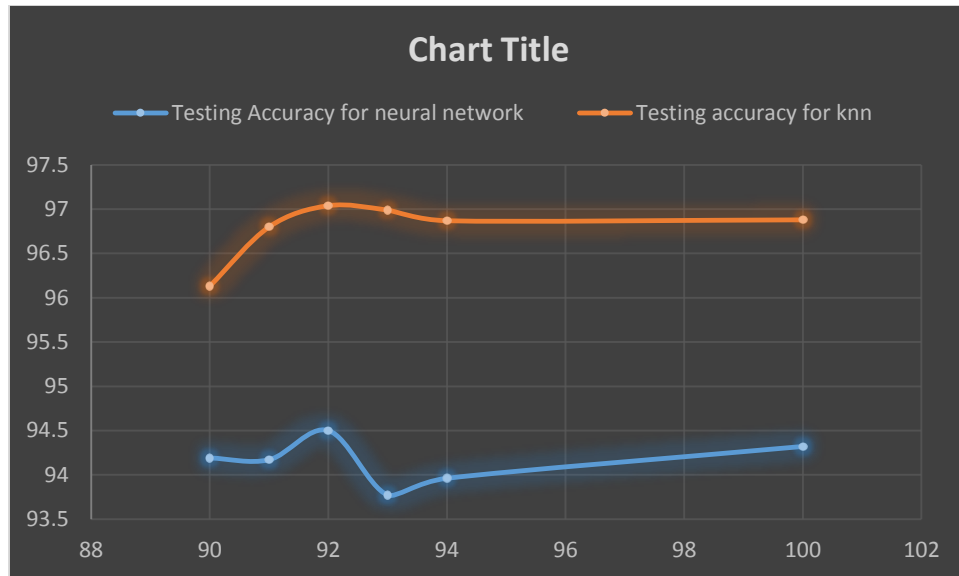


Figure 5: Variation of KNN - Accuracies (Y-Axis) with NN Accuracies (X-Axis)

Advantages and disadvantages of neural network:

Neural networks do not need an algorithm to solve the problem. Since it involves parallel methodology of solving problem, the neural network can still achieve its purpose even though one of its nodes fails out.

The time taken for computation is lesser than the KNN classification. Apart from it, neural networks offer a number of advantages such as requiring less formal statistical training, ability to implicitly detect complex nonlinear relationships between dependent and independent variables, ability to detect all possible interactions between predictor variables, and the availability of multiple training algorithms.

Disadvantages include its "black box" nature, greater computational burden, and the empirical nature of model development.

Advantages and disadvantages of KNN classification:

The main advantage of k-NN methods is their simplicity and lack of parametric assumptions. In the presence of a large enough training set, these methods perform surprisingly well, especially when each class is characterized by multiple combinations of predictor values.

There is one difficulty with the practical exploitation of the power of the k-NN approach. Although no time is required to estimate parameters from the training data, the time to find the nearest neighbors in a large training set can be prohibitive.

Resources:

- 1) <http://searchnetworking.techtarget.com/definition/neural-network>
- 2) <http://www.sciencedirect.com/science/article/pii/S0895435696000029?cc=y>