University of Burgundy

MsCV

VISUAL PERCEPTION

Report of project on Human Psychophysics

by

Gopikrishna Erabati

Supervisor: Dr. Elizabeth Thomas

The well documented code for the project can be found in the folder and there is a README.txt file which explains the execution of code for the project.

1. Introduction

The Self-Organizing Map is one of the most popular neural network models. One particularly interesting class of unsupervised system[1] is based on competitive learning, in which the output neurons compete amongst themselves to be activated, with the result that only one is activated at any one time. This activated neuron is called a *winning neuron*. Such competition can be implemented by having lateral inhibition connections between the neurons. The result is that the neurons are forced to organize themselves. For obvious reasons, such a network is called a Self Organizing Map (SOM).

For example, we can use the SOM for clustering data without knowing the class memberships of the input data. The Self-Organizing Map was developed by Prof. Kohonen. This is also used to produce a low-dimensional (typically two-dimensional), discretized representation of the input space of the training samples.

2. Project I

In the Kohonen network the cluster is formed with weights as a centroid of the cluster. So, after finding the weights, each test data is compared with weights to judge the cluster to which data belongs to[2]. The kohonen network is built for the given data in project 1. The learning rate is chosen by converging the synaptic weight values. So, for this the error in weights between two iterations is plotted to check the convergence of weights as shown in Fig. 1.

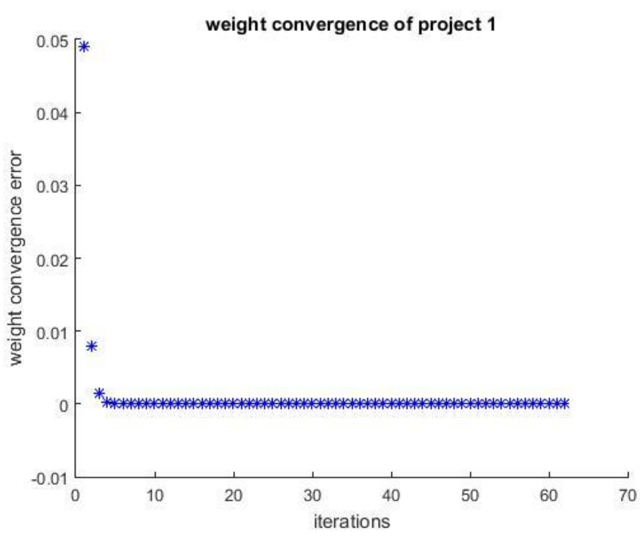


Fig. 1 Synaptic weight convergence vs iterations

The initial learning rate is chosen as 0.6 and it has been chosen to decay exponentially as shown in Fig. 2.

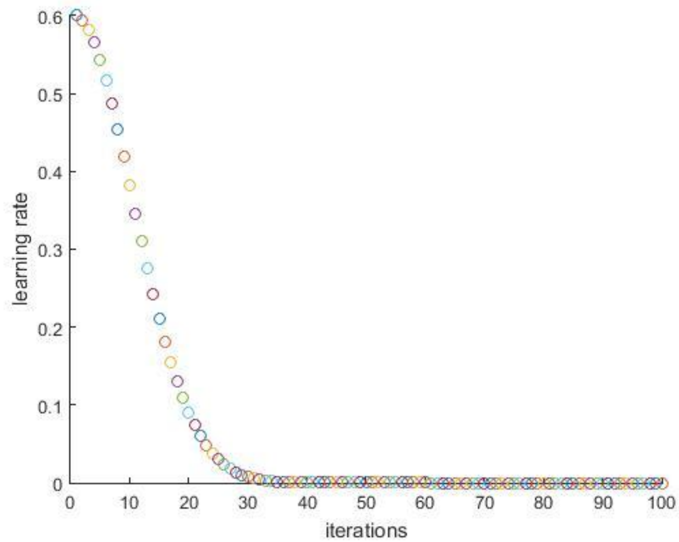


Fig. 2 Learning rate vs iterations

The final converged weights are as:

1.0000	0.0000
0.4972	0.0000
0.0000	0.5028
0.0000	1.0000

And the test data is tested on the trained network, and the results are as:
test data 1 [0 0 0 0.9] belongs to class II
test data 2 [0 0 0.8 0.9] belongs to class II
test data 3 [0.7 0 0 0] belongs to class I
test data 4 [0.7 0.9 0 0]belongs to class I

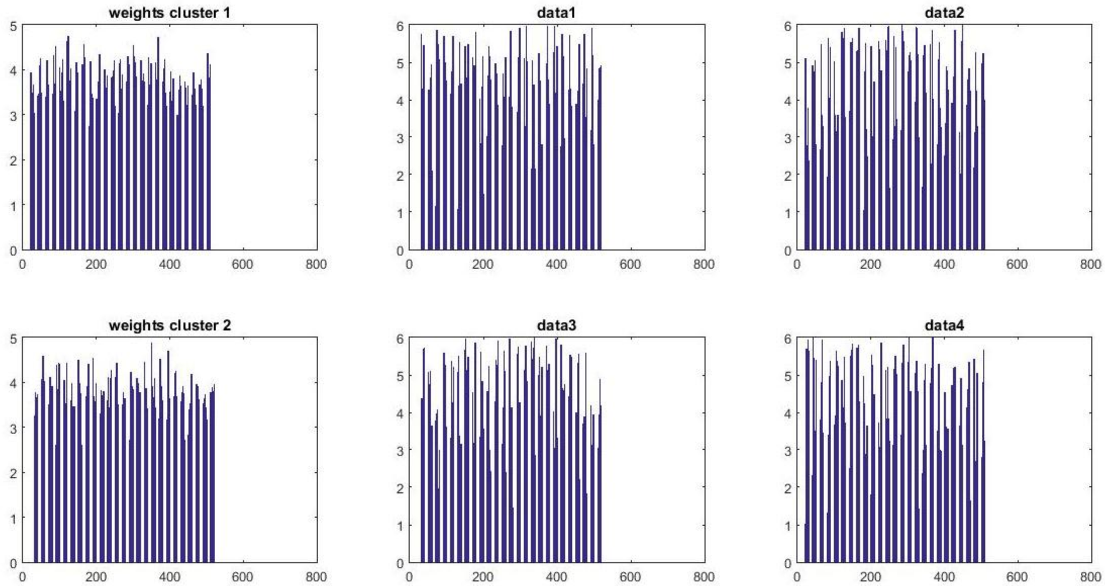
2. Project II

The kohonen network is trained for the control and patient data and converged synaptic weights are computed for appropriate learning rate of 0.6.

3. Project III

The test data 'gopi.txt' which has four subjects is tested and the results of identity are as:
test data 1 belongs to control
test data 2 belongs to patient
test data 3 belongs to control
test data 4 belongs to patient

We can also infer the information for which data belongs to which cluster by visualizing the bar graphs (histograms) of weights and data as shown below:



In the above figure, we can see that data1 and data3 belongs to weight cluster 2(control) and data2 and data4 belongs to weight cluster1(patient).

4. Differences between a bio-inspired algorithm like the Kohonen SOM and two other well-known similar algorithms – the k-means and k nearest neighbour (K-NN) algorithm:

Kohonen SOM	**K-means**	**K-NN**
Clustering algorithm	Clustering algorithm	Classification algorithm
Unsupervised learning Labeling, supervised	Unsupervised learning	Supervised learning
	It cares about global	It cares about Local (when k is small)
More robust learning	Sensitive to initialization	The target function may be either discrete-valued or real valued
Clusters by geometrically	Clusters are formed through centroid and cluster size	Use of Vornoi diagrams
Not better result than SOM, if more clusters	If k is large, better result than SOM	It is easily misled when instance space is high-dimensional
Less sensitive to the noise present in the dataset	More sensitive to the noise present in the dataset	Dominated by large number of irrelevant features
Useful when clustering is some kind of 2D visualization	Can be considered a special case of SOM were no neighbors are taken into account when modifying centroid vectors	

5. Conclusion

Advantages:

- It is very easy to understand.
- It’s very simple, if they are close together then they are similar.
- We can classify data well and then are easily evaluate for their own quality so we can actually calculate how good a map is and how strong the similarities between objects are.

Disadvantages:

- SOMs are very computationally expensive which is a major drawback since as the dimensions of the data increases, dimension reduction visualization techniques become more important, but unfortunately then time to compute them also increases.
- Problem with SOMs is getting the right data at starting.
- Every SOM is different and finds different similarities among the sample vectors.

References

[1] Teuvo Kohonen, Self-Organizing Maps, Springer Series in Information Sciences, Vol. 30, Springer, Berlin, Heidelberg, New York, 1995, 1997, 2001, 3rd edition

[2] lecture notes of Dr. Elizabeth Thomas.