

BLANCHON MARC
MsCV University of Burgundy
Class project for module Human Psychophysics

Matlab Code

The goal of this project was to develop a Kohonen Network in order to first make clusters between two different type of data (making this code generic to be useful in any situation) and then extend the process to a differentiation between a control and a patient.

Lot of approach was possible while sticking to the Kohonen method seen in the class but independently of this we could either use a straight code and run it only on our application or use a generic approach using functions to be able to expand the project to another dimension such as being able to use our homemade functions to divide the work, understand it better and reuse it after. In this case the approach chosen is the second.

Focusing on the second part of the work, we had to build the code able to output a weight matrix corresponding to the dimension of our input and use it to differentiate a control and a patient.

To avoid any further explanation on the subject, we will discuss more about the effectiveness and my interpretation of the work instead of sticking to an usual report showing results. In fact this kind of project allowed me to be able to separate and differentiate the tasks needed and by this expand my knowledge about supervised or un-supervised clustering. Being able to input any dimension matrix containing data to sort, and output a weight (or importance) matrix to describe our model and then use it to cluster a testing set of data was intuitive and allowed me also to understand more about other process of unsupervised clustering as tensor flow user for image classification.

This kind of functions could be used to differentiate lot of things and even images could be the input parameter. In fact after testing this kind of process my result was not that good but I think the main problem is the dimensions of the image (using 3D data RGB instead of 2D vectors is the main problem).

To end with this Matlab code part, we can say that the functions are able to differentiate the two clusters (for the first part) and for the second part is able to differentiate control from patient. I want to highlight that in the code there is the possibility of changing the learning rate and the number of iterations. At some parameters I could observe some incoherences such as interpreting patient as control. It happen mainly when we greatly decrease the learning rate and increase the iteration, it slows the process but it was in the objective of testing the limits of this program.

At the end we obtain with coherent values that the four data correspond to patient (cluster 1).

Difference between SOM , k-mean and k-nearest

We will consider these three algorithm as cluster algorithm. First the first differentiation is k-nearest is a supervised algorithm and this is the main difference, computing a number of neighbour to have the "agreement" but it is highly dependant of input parameters. We can say that this algorithm is because of the supervised side, out of the course.

For the SOM and the k-mean, we have two algorithm that are linked. K-mean is a subset of the SOM. For the SOM we wont provide any number of clusters and this is the strength and also the weakness of this algorithm, SOM can be really slow and this is a matter of optimization. k-mean is mainly different in this point, k is the number of clusters at the end of the clustering so if we have

this good k , this algorithm can be much more efficient. Another difference is the dimension of output, when SOM gives 2 dimension output, k -mean gives k dimension output.

To conclude, the first and only approach used is the difference between clustering method, and this is a special case of optimisation / classification. In fact I think that it was the best approach to have multiple proprieties in common to be discussed and compared.

We can never say that an optimisation method is better than another, to reach to goal we have to define our needs and this is exactly the purpose of this comparison.