

# Implementation of a Self-Organizing Map

Damien DELPY

ENSEIRB-MATMECA

October 31, 2023

Overview

Algorithm

Bibliography

# OVERVIEW

# What is a SOM ?

## Definition

It is a neural network of just one layer : the output layer.

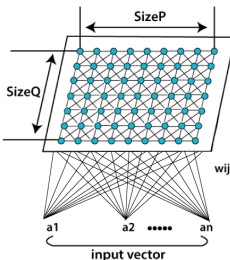


Figure: SOM Architecture (see 2 of Bibliography).

## Wikipedia

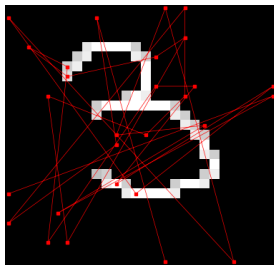
A self-organizing map (SOM) is used to produce a low-dimensional (typically two-dimensional) representation of a higher dimensional data set, while preserving the topological structure of the data.

# Motivation

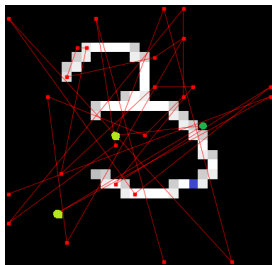
The Self-Organizing Maps permit to :

- ▶ Analyse and visualise the data. It represents complex data on a map of only two or three dimensions (see Convergence slide).
- ▶ Detect patterns from the data. Clustering (see K-means slide).
- ▶ Improve a deep neuronal network by sorting the data at the beginning.

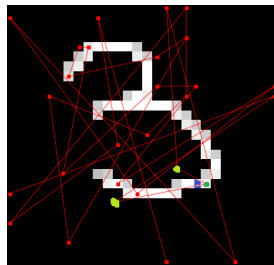
# Example



Step 1



Step 2



Step 3

1. Initialize the weight vectors (randomly or not) in red.
2. Select a data vector(blue), then chose the closest weight vector(green) to it.
3. Update the winner and its neighbors (all green ones).
4. repeat the process till reach max\_iteration

# Similarities with the Perceptron

## Perceptron

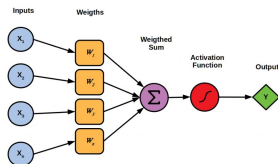


Figure: Diagram of a Perceptron.

It is also a one-layer neuronal network. However, this one is used to separate two different classes. The output is actually a binary one.

This is a supervised learning algorithm.

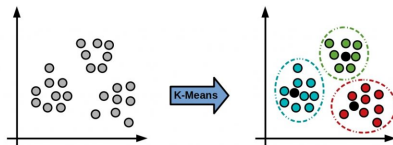
## SOM

The SOM can gather vectors due to their similarities.

The SOM is an **unsupervised learning algorithm**.

# Similarities with K-means algorithm

K-means algorithm is an unsupervised learning technique that can automatically gather data by creating **clusters**, which are subsets of data elements that share common characteristics.



**Figure:** Process of K-means algorithm

The user must define the number of clusters **K**. However, The SOM does not require this, it guesses the right amount of clusters.



# Convergence

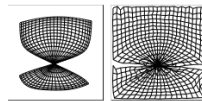
The convergence of the SOM algorithm is not guaranteed (1). There are actually 2 errors that can appear.



**Figure: Dimension Error.**

It happens when the number of neurons does not fit with the data.

The ideal number of neurons is  $5\sqrt{N}$  where  $N$  is the number of data vectors.



**Figure: Topological Error.**

It happens when a node is created. It looks like a butterfly.

# ALGORITHM

# Different Phases

Competition

Cooperation

Adaptation

# Algorithm

# BIBLIOGRAPHY

1. Self-Organizing Maps - Teuvo Kohonen (2001)
2. <https://www.baeldung.com/cs/som-algorithm>
3. [http://www.pspc.unige.it/drivsc/Papers/VanHulle\\_Springer.pdf](http://www.pspc.unige.it/drivsc/Papers/VanHulle_Springer.pdf)