

Machine Learning

Practical work 05 - Unsupervised Learning

Teacher: Andres Perez-Uribe (Email: andres.perez-uribe@heig-vd.ch)

Assistants: Hector Satizabal (Email: hector-fabio.satizabal-mejia@heig-vd.ch), Yasaman Izadmehr (Email: yasaman.izadmehr@heig-vd.ch)

Introduction

Use Self-Organized Maps (SOM) and K-Means as a means for clustering and visualization purposes using a collection of images.

Practical work

1. Explore the use of Self-Organizing Maps and K-Means

To start, let's use the animals database to try the K-means and SOM algorithms with different arguments and observe the results.

2. Clustering of wine data

Apply the K-means algorithm to the wine database. Use the 13 features of wine to find clusters in the data. a) Set $K=3$ and run K-means, 10 times each time, and b) given that you already know the type of wine for each observation, compute the average number (based on your 10 runs) of the number of observations that are correctly grouped together for each type of wine. Comment your results

3. Clustering of images application

We will provide you with a database of color images, a set of three feature extraction methods and a SOM library. You may setup diverse experiments with the database, which contains a lot of classes. You should apply the three feature extraction methods and observe the results, and you have to modify the configuration and learning parameters of the SOM algorithm

Report

Present your results and comments for point 2 and regarding points 3, simply follow the instructions on the notebook SOM_part2.

Summary for the organization:

- Submit the solutions of the practical work before Wednesday 27.5.2019, 23h55 via Cyberlearn.
- Modality: PDF report
- The file name must contain the number of the practical work, followed by the names of the team members by alphabetical order, for example 05_dupont_muller.pdf.
- Put also the name of the team members in the body of the report.
- Only one submission per team.