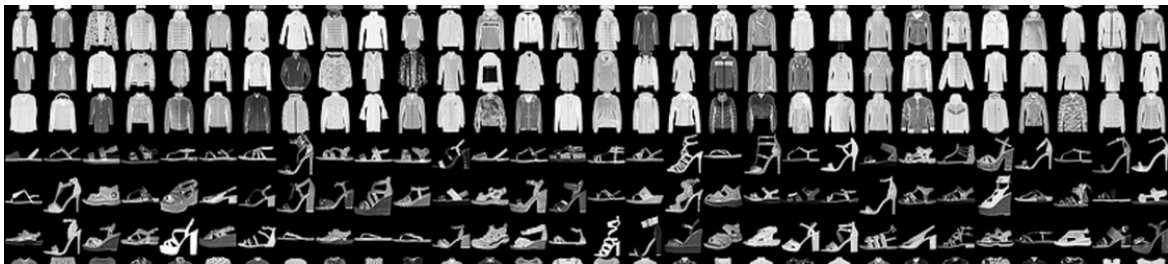


Machine Learning

Practical Sheet 6: Dimensionality Reduction (cont.)

Revisit the Fashion-MNIST dataset of Zalando's article images, consisting of a training set of 60,000 examples and a test set of 10,000 examples.



Data are provided in “csv” format, with one image per row. Each image has 28 pixels in height and 28 pixels in width (total 784 pixels). Each pixel has a single value, between 0 and 255. The data sets have 785 columns. The first column is the class labels, representing the article of clothing. The remaining columns are the pixel-values (features) of the image.

Your task is to compare the effectiveness of different variants of MLP classifiers:

- 1- A “Raw MLP” that receives the 784 features and learns the corresponding class;
- 2- A “Handcrafted Features MLP”. Use the [“fashionmnist_feature_extraction.py”](#) functions, available in the course web page, to obtain a set of handcrafted features that will feed the MLP.
- 3- A “PCA MLP” that receives the images data projected into the PCA space, while keeping 50%, 75%, 90%, 95%, 99% and 99.9% of the information (variability).
- 4- A “t-SNE MLP” that receives the images data with dimensionality reduced by t-SNE, using different number of projected components.
- 5- A “UMAP MLP” that receives the images data with dimensionality reduced by UMAP, using different number of projected components.