Machine Learning: Homework on NN for Classification and Clustering

Fabio Vandin

January 9th, 2020

Instructions

You have to complete 1 Jupyter notebook

The notebook has missing code: need to fill in what is missing

You also need to write some text (to explain choices or describe results)

Feel free to add cells with text if you need to explain or describe some ``non-standard" decision!

FIRST THING TO DO: you need to put your ID number in the notebook (as seed for random number generators).

IMPORTANT: code already there is a guideline, if you want to change or rewrite some/all of it go ahead, but make sure to answer all TO DOs!

Get the data

Download the following file (26MB):

http://www.dei.unipd.it/~vandinfa/courses/ML2019/data.zip

Unzip the file so that the “data” folder is within the “ML1920_HW3” folder you downloaded from elearning.

Deadline

Submit your completed notebook:

- **deadline: Thursday January 23rd, 11:55 PM**
- **use link in elearning website**

Submit 1 file - **Only submit your completed notebook!**

IMPORTANT: Use the following file name for the files that you have to submit:

- NN_FirstnameLastName_IDnumber.ipynb

Example: student Fabio Vandin, ID number 000001 will submit files:

- NN_FabioVandin_000001.ipynb

WRONG FILE NAMES = 0 POINTS

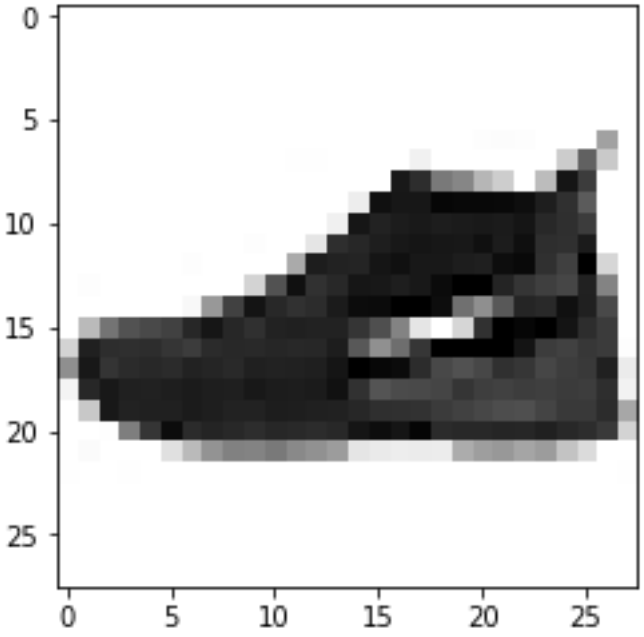
In lab assistance: **Thursday January 16th, room Ue, 2:30-4:15 pm**

Dataset

Fashion MNIST dataset:

- see this link for more info: https://pravarmahajan.github.io/fashion/
- 60,000 training instances, input =vector of 784 integer values in [0,…,255]= 28x28 matrix
- Instance = image from one of 10 clothes categories

INPUT:



TARGET/LABEL: 7
(= sneaker)