Machine Learning: NN for Classification, and Clustering

Fabio Vandin

December 20th, 2022

Instructions

You have to complete 1 Jupyter notebook:

- one for NNs on classification and clustering (clothes images data)

Both notebooks have missing code: need to fill in what is missing

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

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

Deadline

Submit your completed notebooks:

- **deadline: Tuesday, January 17th, 2023, 11:55 PM**
- **use link in elearning website**

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

IMPORTANT: Use the following file name for the 1 file that you have to submit:

- NN_FirstnameLastName_IDnumber.ipynb

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

- NN_FabioVandin_000001.ipynb

WRONG FILE NAMES = 0 POINTS

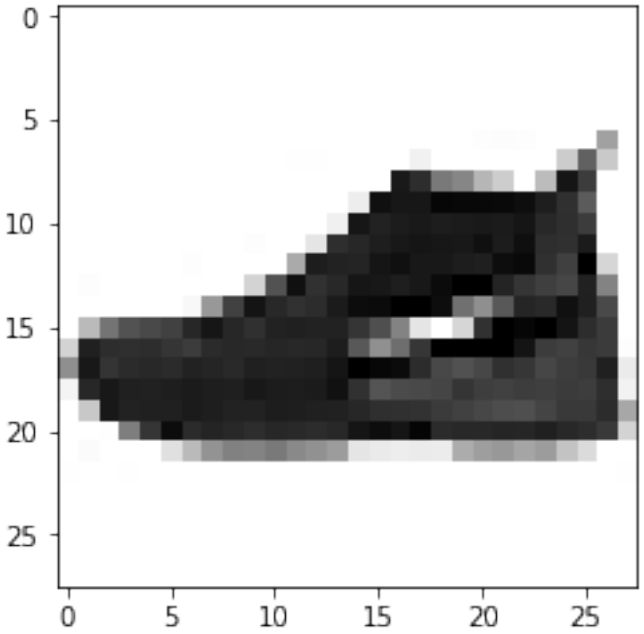
LATE SUBMISSION (e.g., email) = 0 POINTS

Dataset for NN classification and Clustering

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)

Get the data for NN classification and Clustering

Download the following file (26MB):

http://www.dei.unipd.it/~vandinfa/courses/ML2022/data.zip

Unzip the file so that the “data” folder is within the “ML2022_HW3” folder you downloaded from https://stem.elearning.unipd.it/