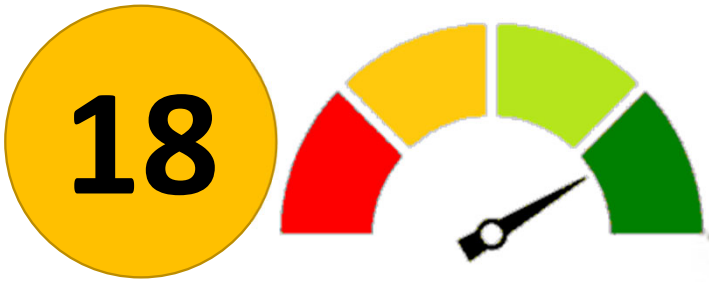
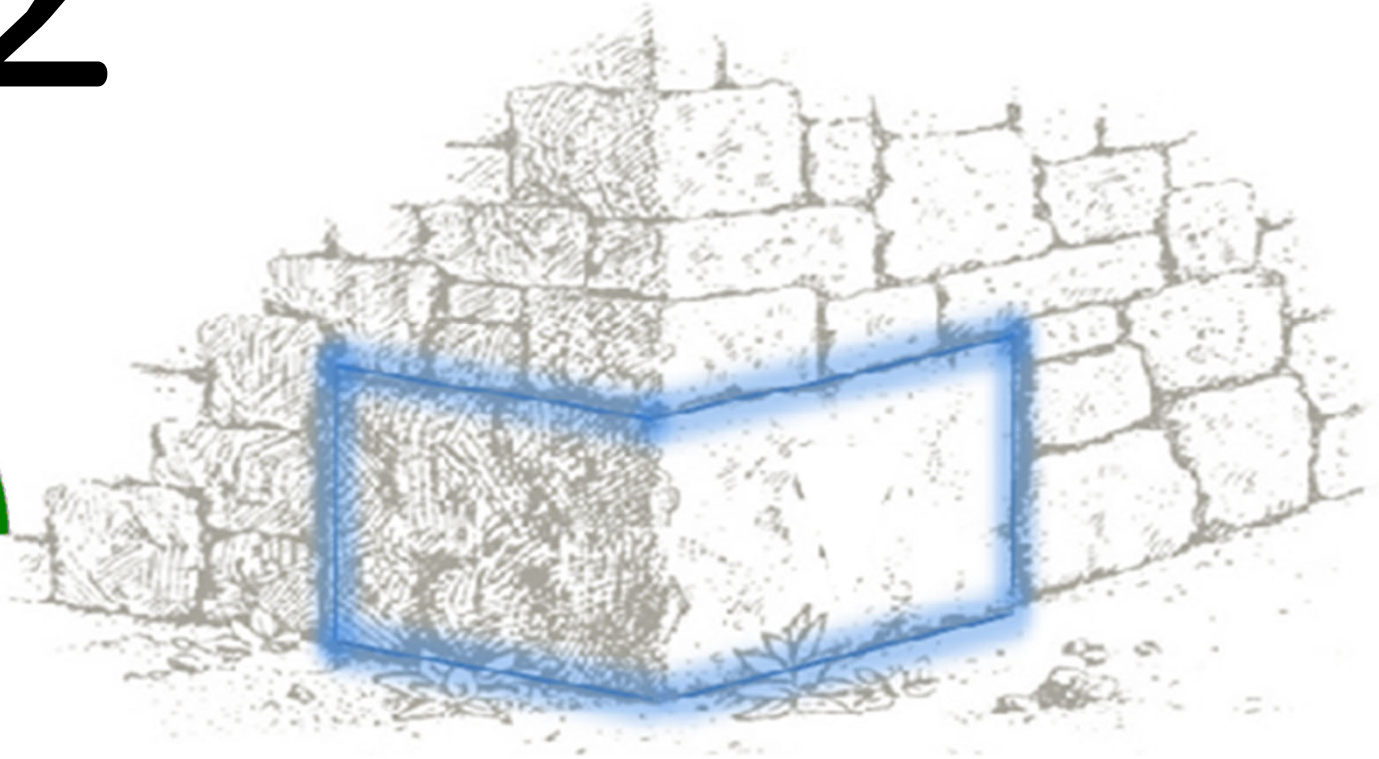
ML^2

Machine Learning [to Learn] Foundations





José Manuel Rey, IE

Machine Learning Foundations
Assignment 18.00
Practice Recap

Date: Mar/5 2024

Due Date: Apr/15 2024 [SESSION 26]

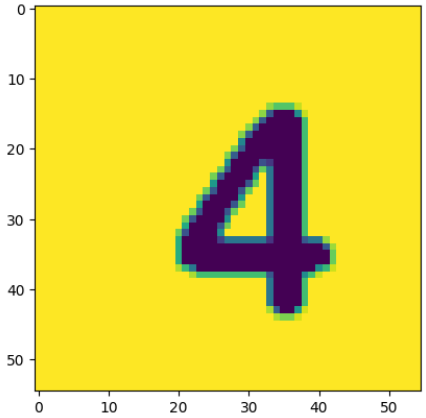
Ongoing grading (maximum)

60 Individual Assignment [mini-groups]

Submit your work in ZIP format
via Campus Blackboard
(one submission per mini-group)

MLF_18 - Assignment Description

- A ***simple custom digit recognizer*** implementation
- At this stage of the course the students will have a sufficiently wide knowledge of the main technical areas and components involved in the development of ML models that should allow them to make well directed designs regarding solutions or applications using these components for solving basic ML tasks and evaluating their expected performance



Ongoing grading (maximum): 60 Individual Assignment realized in mini-groups

Assignment Description – Overview (1)

Students will just be provided with the dataset of “digits” for training

- MLF_18 - digit_recognizer - Assignment.ipynb
- dataset_train.csv

Tasks:

- Create 4 models for multi-class classification of the input images
  - A,B) Two Basic Machine Learning Models
  - C) Multi-layer Perceptron
  - D) Simple Convolutional Network
- Development tools:
  - Stand-alone Jupyter Notebook
  - Solutions Implemented with Sklearn, Tensorflow and/or Pytorch (free choice by each student or mini-group)

Assignment Description – Overview (2)

Tasks:

- Understanding and EDA of the dataset provided
- Build the Models
- Train the Models
- Cross-Validation + Basic GridSearch hyper-parameter tuning
- Performance report / Comparison of results
- Submit:
  - Code (documented and executed)
  - Models
  - Selected model (ready to be scored independently with a “test” set)

Assignment Description - Overview (3)

Grading:

- As defined in the attached rubrics file, the grading will have 3 components:
  - 35% CONCEPTUAL
  - 35% CODING
  - 30% BAKE-OFF Ranking of results (performance)
- For the scoring and ranking of results an IID(*) "test" set will be used
  - Note: the ranking grade is based on the "discrepancy" of performance with respect to the average standard solution (only really under-performing models would have a significant penalty in this rubric)
- A "fun" application or demonstration of the use of the [winner] model created will be showcased

(*) Independent and Identically Distributed random variables

Assignment Description - Overview (4)

Deliverables:

SUBMIT YOUR WORK AS A STANDALONE .ZIP FILE BY ZIPPING TOGETHER IN A SINGLE FILE THE ADDITIONAL FILES (JUPYTER NOTEBOOK CODE, DATA, MODELS, REFERENCES, ETC) THAT YOU MAY DEEM NECESSARY

NAME YOUR SUBMISSION AS: "MLF_18_00 – mini-group number.zip" (see the attached .txt for the mini-groups)