

FML exam 20-21

Details and projects data

The project part of the FML exam

- The projects are related to several problems connected to some data I will give you once you decide the project, starting from the next week.
- Unfortunately, we did not have time to do lab in Python, at least I hope you have got a honest insight of the different techniques we discussed. In any case, the project can be done in Matlab or in python, as you prefer
- The project has to be carried out individually, and provided to the teacher one week before the exam, by email, with a link to a zip file containing the code and a readme file, and a report of what you have done
- During the exam, there will be a first part of discussion about the project, in which I will check your work, asking you the reasons of your implementations
- After having considered the project, we will have the

The oral part of the FML exam

- After having considered the project, we will have the oral part (which process has been already discussed in the first lesson)
 - you can start with a topic of your interest. In the case the explanation is not satisfactory, or to cover parts you have missed, or where I would like to ask more, I may ask more details
 - There will be a second main question, where there will be another topic of analysis
- For each question (the chosen topic and the one of mine) I expect to have written proofs

Details on the schedule

- As I told you, the days for the exam are indicative. In fact, I will give you the Tuesday's slot of office hours (2-5pm) of the period 1 June – 15 July, so that you can defrag your schedule appropriately
- I ask you to organize the schedule (with the representative?), providing me a calendar with day/hour/people. Each exam will take approx 40 minute.

Projects

1. Given a set of sales signals (time, number of sold items) perform clustering in order to highlight three classes: best-sellers, normal-sellers, low-sellers. The items have a time span in the market of 12 values.

Data available here: : [link](#)

2. Given a wrapper for google trends (you indicate a text and a time period, and you will get the signal), a set of topics and periods, (which we will give you all!) the goal is to understand if and how a topic is «hot»

- is not hot!
- is hot for a single time
- is hot *recurrently*

For example, given the topic «red», and the periods Jan-June, July-Dec, when the topic was highly searched. The data and google trends downloader [are available here](#).

The time interval for the download is: 2015-10-02 – 2019-12-22 (already defined in the code) and the the list of topics (two macro-categories to choose from) is contained in the file topics.txt

3. Given a set of scooter trajectories of an urban map, perform unsupervised classification
 - I. on the area covered, and the shape of the trajectories, answering to the question: which trajectories were on the same geographical area, and of the same shape?
 - II. on the time span and the area covered, answering to the question: which time that path was particularly crowded of scooters?
 - Data here: [link](#)
4. Given a set of clothes images, cluster them into meaningful and interpretable clusters
 - Data here: [link](#)

5. Face mask analysis: given a set of images of people faces, cluster them into clusters accordingly to some criteria (e.g. mask color)
 - Data here: <https://www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset>
6. Face mask analysis: given a set of images of people faces, classify them as "worn" and "not worn"
 - Data here: <https://www.kaggle.com/ashishjangra27/face-mask-12k-images-dataset>
7. Trajectory analysis: Given a set of coordinates representing a trajectory in time and space (2D), cluster them by some meaningful criteria (e.g. similarity)
 - Data here: https://1drv.ms/u/s!ApbwhsUbf_OliPjU7ecmMT7I-53i2g?e=yFULUr

8. Head direction: given some full-body people images, split the dataset into two sets, one for "training" the other for "testing" (e.g. 80:20. Evaluate carefully how to divide the dataset).

For each frame in training dataset annotate each person with:

- **bounding box** around the face (rect)
- **9-way direction** (straight forward, up, down, left, right, up-left, up-right, down-left, down-right) of the face
- **head keypoints location** (2d pixel): mouth, nose, left eye, right eye, left ear, right ear

Then classify the testing set in order to get the class (9-way direction of the head).

- Data here:

<https://drive.google.com/file/d/19SFbWzvhlv9OJzpdL8myELMrts8rmHxE/view?usp=sharing>

9. Human behavior analysis: given a set of images of full-body people and the position of each body skeleton (i.e. 17 2D coordinates) in the image, classify the person as "still" and "in movement".

- Data here: https://1drv.ms/u/s!ApbwhsUbf_OliPJP80VFFwu1hjh8g?e=VSnjZS