The exam

Steps towards results

- Understanding of the business problem and business need
- Existent data analysis
- Imagine a service
- Search for new possible data sources
- Feature identification and extraction
- Modelling
- Best results identification
- Presentation of results (not necessarily a prototype even just a mock-up) - storytelling

Schema

A. **Introduction**
Quick explanation of
- dataset (with some descriptive statistics)
- idea and motivations
- the aim

B. **Methodologies**
List all the algorithm and methods used explaining
- The principles at a high level (in few words)
- Theory and Formulas
- Pros and cons respect other possibilities

C. Experiments
List all the experiments you did indicating
- General and clear explanation
- How you used methods (indicating the parameters choice with reasons)
- Details to reproduce them

D. Results
List results of all experiments with easy understandable graphs and tables

C. **Conclusion(s)**
It has to be synthetic, precise and quantitative.

General points

It should be possible to read it in this way (try to be sure!):

1) read the introduction explaining the context,
2) go directly to the conclusions that say what has been found,
3) randomly look at various parts in the body, which must be easily found and understood (including the graphics with exhaustive captions so that they are understandable without necessarily reading the whole text)

Things to take care of:

- Graphics must have clear captions and dimensions with units
- Proofread the text carefully to minimize typos
- Give an explanation for the choices you make
- Use correct scientific words

Avoid these mistakes

- Do not use non quantitative and generic sentences (e.g. "forecast is not good”, "cookies are few", "we got satisfying results")
- Do not create graphs without information on axis or with the “code variables” (e.g "gr_inc")
- Do not linger on off topic difficulties (e.g. "CSV files are huge and we tried many things before finding a way to import them")
- Do not give partial or vague informations (e.g. "we used many variables, such as ... [list all!]”, "we cleaned up time series with several expedients” [how? which? what have you done?])
- Do not use improper words (mainly if technical, e.g. “higher trend”, “average” without indicating if daily, weekly, monthly ecc.)
- Do not forget to provide all necessary informations (e.g. "forecast error is 4%" without specifying how distant in the future)
- Do not use useless bombastic style
- Do not repeat the same concept many times in different part of the text
- Do not send the document without page numbers

How to send your work

The delivery has to be done in this way, otherwise it won’t be considered:

- sent by email from university accounts of one of the group’s members (@campus.unimib.it)
- sent to (field “To:”) both marco.fattore@unimib.it and silvio.gerli@unimib.it
- with all the members of the group in the “Cc:” field
- with subject: "DELIVERY DSLAB 2024" and then all the surnames of the students in the group
- in attachment: the text in pdf and the code for reproducibility (link to the code is acceptable)
- in the body the list of the names and surnames of all the members of the group with their “matricola” numbers
- sent it at least 10 days before the last date of the exam session