Our Aim:

The aim of the project is to examine a real-world process of deploying machine learning models. More specifically, the project component of this course allows you to explore a technology that assists in model deployment and evaluation, either directly or indirectly, and asks you to report your experience working with that technology (or multiple technologies) to achieve some overall deployment goal.

Group:

Groups of 2 students for this project component . Each member's contributions need to be reported in the final submission in a separate readme file.

Project notes

There is one due-date for the project deliverables. (le 17 dec 2022)

The deliverables are as follows.

- Project video: you should explain each group member's contributions in the project
- What was the problem and what type of treatment you did and how did you solve it and why this choice.
- Code and data: Code associated with the project (e.g. Jupyter notebooks +scripts in python), a small sample of the data/inputs/outputs, and all steps necessary to replicate your project should be provided along with/in the report/demo file.
- A link to your github/gitlab/ or other repository (write it on the front page of the report and don't forget to add the link directly and give me access to clone and run the project).
- The assessment is not done on the performance of the algorithm but rather on the explanation and coding skills. You are encouraged to provide as much details as you can in markdown cells while using jupyter, or file documentation documenting the scripts in python.

Each team should upload the report/ video (and code) on the Ametice page before the deadline. In case of difficulty uploading, send me a link by email.

Example Report /video components

-Group name,
-github link
- What was the goal business point, questions, Timeline

Part I: Data Engineering

- Data Cleaning : Look at the missing values and how you can impute them. (Dates and times, Filling factor)
- Compute a few descriptive statistics.
- Compute features.
- Challenges descriptions.
- and some insight and visualization of your findings.

Jupyter, python

Part II:

According to your problem and case study:(classification..etc)

- Choose to train more than one model.
- Check model results .

Use the right tools to compare the different models.
Don't forget to detail :
• The data used.
• The model and the hypothesis made to build it.
• The validation methodology and the validation results.
• The monitoring methodology.

Kedro , ML flow..etc

Grading

It is importantly to make a project plan that allocates sufficient tasks for each team member. It will be great if you can write the project plan.

- Projects will be graded based on the creativity shown in handling the technology and the insights drawn. (so show me your talent !)
- The documentation /video should be clearly written and presented.
- All external material/sources (code/idea/theory/insights) used should be cited.

  The use of pre-trained models, databases, API , visualization tools .. or any tool you find useful.. etc for your project is allowed and encouraged.

Note: Don't hesitate to contact me in case of confusion/difficulties/questions.