Advanced Data Analytics
Requirements for the Capstone Project 2024

Spring 2024

1 General Comments and formal requirements

To show the successful acquisition and mastery of the course content thought in the class "Advanced Data Analytics", all the students individually need to propose and carry out a final project that applies this knowledge to a problem of their choice.

All students individually will present their project via a 15 minutes long video that has to be submitted/uploaded/sent to the TAs of the course in the last week after classes of the semester—that is, on Sunday, **26th of May 2024, 23.59h**.

A pre-recorded presentation of the project is supposed to be 15' long, recorded, for instance, via Zoom.

Question regarding the details should be directed to Anna Smirnova (anna.smirnova@unil.ch) or Maria-Pia lombardo (mariapia.lombardo@unil.ch).

- 100% of the final grade will be based on submitting a project that provides evidence of the student's ability to apply what she/he learned during class.

- The grading will be based on several formal factors that are listed below as well as on the originality and complexity of the project and the presentation.

- By the **12th of April 2024** the latest, all students need to inform the TAs about the proposed project by submitting a 200-word proposal. Based on the TA's sign-off regarding the scope and complexity of the proposed project, it then can be undertaken.

- Due date of the semester project is Sunday, **26th of May 2024, 23.59h**. No late hand-in is accepted. A late hand-in will result in zero points for the project.

- The semester project will be submitted to Anna Smirnova (anna.smirnova@unil.ch) or Maria-Pia lombardo (mariapia.lombardo@unil.ch).

2   Grading

The grading will be based on three factors—that is, i) satisfying the formal factors (listed below), as well as on ii) complexity of the project and iii) the originality of the project.

2.1   Formal Factors

- At the due date, four parts have to be submitted. i) a research paper of about 10 pages length (min. 8 pages without references), ii) the source data used to carry out the project, iii) the source code to carry out the data analytics project, and iii) a recording of about 15 minutes length that presents and summarizes the project.

- The research paper will be submitted as a pdf format and in the SIAM conference style (style files and examples are provided)

- The submitted research paper needs to contain the following mandatory 8 sections

    - 1. Abstract
    - 2. Introduction
    - 3. Description of the research question and the relevant literature
    - 4. The methodology applied to address the research question
    - 5. A description of the data set (where it comes from, how it was cleaned, etc...)
    - 6. A discussion of the implementation (code)
    - 7. Results
    - 8. Conclusion
    - Appendix: A list of helper-tools (if used, such as Chat-GPT).

- Additional sections are permitted if they make sense in the context of the project.

You may use Chat-GPT, Co-pilot, and other resources aiding the code development and writing. However, all resources you use need to be clearly listed in the appendix.

2.2   Our policy in case there are no significant results

Working with data can be hard: Having a hypothesis does not always mean that one can confirm it with data. We want to encourage scholarly behavior: a project can still get full marks with a well-told story about why the hypothesis failed rather than someone submits a project with rigged data.