DS-GA 1007 Programming for Data Science

Final Project Instructions and Dates

The agenda for final project presentations comprises four main steps, as detailed below:

November 16th: Cover sheet submission

By November 16th students must upload at NYU Classes a cover sheet pdf file containing:

Title of the Project
Summary of the Project (one paragraph)
Name and NetID of each member of the group (at most 4 students per group)
Member responsible for uploads

There is an assignment opened at NYU Classes entitled _Final Project_, where each group must upload its pdf file.

NOTICE: Only the _member responsible for uploads_ needs to upload the pdf file. In other words, each group should have only one pdf file uploaded at NYU Classes.

November 23rd: Announcement of the order and duration of the presentations

The duration of each presentation depends on the number of groups. We are estimating 25 groups, what gives about 4 minutes per group. However, the exact duration of each presentation will be defined on November 23rd, after the groups have submitted their cover sheet files.

The order of the presentations will be defined by the date and time of cover sheet upload. The sooner a group uploads its cover sheet into NYU Classes the earlier it will be in the presentation section.

December 14th: Presentation submission

The slides of the presentation must be uploaded into NYU Classes by December 14th. Only pdf file format is accepted. You can prepare the presentation in Powerpoint or in any other slide editor, but you must export the slides to pdf format before uploading it into NYU Classes. December 14th is a hard deadline, so please do not miss it. Only the _member responsible for uploads_ needs to upload the pdf file.

December 19th: Presentation section and Project Report

The slides from each group will be merged in a single presentation, whose order will respect the order of cover sheet uploads. A computer with all presentations will be available for the groups during the presentation section. Groups will not be allowed to connect their own computer.

December 19th is also the deadline for uploading the project report at NYU Classes. The report should describe the problem under analysis, the methodology used to solve it, results with discussion, and conclusion. More specifically, the report should be structured as follows:

1. Title
2. Group Members
3. Abstract
4. Introduction and Motivation
5. Methodology
6. Results
7. Discussion
8. Conclusion
9. References

The report cannot exceed four pages. A template for the report can be downloaded from NYU Classes. Only the *member responsible for uploads* needs to upload the report pdf file.

Final Project Evaluation:

The final project will be graded based on three main aspects:

1. complexity and relevance of the problem,
2. quality of the presentation,
3. quality of the final report.

Complexity of the problem accounts for the size and degree of "cleanness" of the data sets involved in the problem, the complexity of the algorithms employed to model and/or handle the problem, as well as the importance of the problem in terms of social, scientific, and/or economic impact.

A good quality presentation and final report should:

1. clearly state the problem, pointing which are the hurdles and issues to solve it;
2. clearly present the methodology employed to solve the problem, pointing out:

a. the data sets used
b. the methods employed to (if necessary) handle missing data, transform data, combine data, etc.
c. the algorithms involved in the solution, as for example, SVM for classification, DBScan for clustering, etc.
d. present and discuss the results, highlighting the strengths and weaknesses of the proposed methodology
e. make some conclusion, emphasizing whether the chosen approach was success and, if not, why.

SUMMARY OF IMPORTANT DATES:

November 16th: Cover sheet submission
November 23rd: Announcement of the order and duration of the presentations
December 14th: Presentation submission
December 19th: Presentation Section and Final Project Report Submission