

Statistical Learning (MT7049) – Group Project

Task: Pick up a data-set from the UCI Machine learning repository:

<https://archive.ics.uci.edu/ml/index.php>

then analyze it using any method from the ESL book. You can use any program package in the analyses. This project should be solved by teamwork with 3 students in each team. Since we may not have number of students which is divisible by 3, it is also permitted (but not recommended) if 1-2 students carry out the project, but you should make an honest attempt at forming groups of 3.

Amount of work: The Group Project contributes the same weight to the final grade as other projects, so the amount of work for each student should be about the same as in one of the other projects. For a team of 3 students, the collective amount of work should therefore be about three times as in one of the other projects.

Guidelines for the written report:

- 1) Give the student names of the group.
- 2) State which data-set is picked and why it is picked
- 3) Describe what the data is about and the data properties with suitable exploratory analyses.
- 4) Clearly state the goal of your analysis.
- 5) Explain which methods in the ESL book are used, and justify why the methods are suitable for the chosen data.
- 6) Concisely state how the analysis is performed (e.g. which functions and packages are used, parameter setting, etc.) and the analysis results in terms of graphs, tables, etc. Do your graphs effectively present the message you want to convey?
- 7) **Most importantly: Explain and discuss the implications from the analysis results, what you learn from the analysis, what are the possible improvements, etc.**
- 8) Give appropriate references.
- 9) Your performance of the project will be judged on how clear, logical and organized your report and presentation are given, and on whether the suitable methods are employed and correctly implemented for the analyses. Please include enough explanations and discussion and avoid writing a technical manual only on what things are done. Moreover, a report size (not including the appendices) of around 5 pages is preferred.
- 10) The written report as a pdf file together with a source code file should be submitted to the course page.
- 11) At the same time, you should also send by email your report to the group that will review your work. You will be given their names and email addresses in due time. This means that you will also receive the report from another group for you to review.

Instructions for peer-reviews and oral presentations will be given in due time.