Project

Big Data Analytics 2020-2021
UMONS
Souhaib Ben Taieb

1 Task

The goal of this project is to study the topic of *Fairness in machine learning*[1][2]. We ask you to find and read some of the relevant references. Then, you should run some experiments to compare algorithms/methods on at least one dataset that you have chosen. Finally, you will write a report which describes the algorithms/methods you have studied and the experiments you have performed. You will provide the commented code (in Python) used for your experiments, which should run without errors. All the sources used in this work should be properly cited in the report. We would like to remind the students that plagiarism will be taken very seriously.

2 Project report

The report can be **a maximum of 10 pages**[3]. You are free to choose the presentation. You could for example divide your report in three sections as follows.

1. The first section presents the main topics/methods/algorithms you have considered, including the key (theoretical) properties/theorems.
2. The second section describes the data set you have chosen and the experiment you have performed, including a justification of your design choices (evaluation metric, optimization procedure, etc).
3. The third section includes a summary of the findings.

Overall, you will be graded based on clarity of writing, quality of presentation, level of machine learning content, and technical communication of main ideas. You should clearly explain what you have done, using figures to supplement your explanation. Your figures must be of proper size with labeled, readable axes. In general, you should take pride in making your report readable and clear.

3 Deadline

May 16, 11:59pm. Upload to Moodle your project report and commented source code to easily reproduce the experiments.

[1] `https://fairmlbook.org/`
[2] `https://fairmlclass.github.io/`)
[3] Few additional pages can be used for Figures only.