COMP0050 Assignment

The coursework involves a group project of your choice, where you will apply techniques and concepts discussed in the module to address a practical problem. You need to identify a problem and an associated dataset, frame the problem in terms of appropriate tasks (e.g., prediction, clustering, etc.), and propose a solution.

There are various ways to approach the project, which may depend on the specific problem and dataset you choose. For instance, you could compare the performance of different models for the problem at hand or focus on a single model, exploring it in greater depth. This might include comparing different regularization techniques or varying the model's specifications. Alternatively, if your dataset is imbalanced, you could analyze the impact of different rebalancing techniques.

In summary, there is no single "correct" solution to the project. However, one requirement is that you incorporate some of the techniques discussed during lectures. If you wish to explore methods beyond those covered in class, this is acceptable, provided you include results (e.g., as a benchmark) using techniques presented during the module.

You are encouraged to discuss your project ideas with me during lectures or office hours, or to send a brief project proposal to ensure your chosen topic and dataset are appropriate.

Written report

A brief written report (maximum of 8 pages) should be submitted to Moodle before the deadline. The report must specify the problem, provide a motivation for your study, describe the data and methodology used, and present your results. This will be a group submission.

Marking

This assignment is worth **40% of the overall mark.** The marking will be based on the following criteria (with uniform weights):

1. Clarity of the report (e.g. is the report clear and well structured, are figures informative, is the methodology explained well? Are modeling assumptions discussed? Are language and style appropriate for a scientific work?)

2. Results (e.g. Are the results sound? Are the methods appropriate and correctly used?)

3. Critical discussion (e.g. are results correctly interpreted? Is there a discussion of limitations and further challenges?)