COMP0087: Group project instructions (2023/24)

8 January 2024

Due *8 April 2024 (Monday) at 16:00 (Europe/London)*[1] and constitutes 100% of your mark for COMP0087. The mark is broken down into the following components: Project report (60%), 15-minute demonstration (10%), 5-minute recorded presentation (10%), and individual peer feedback (20%).

1 Group formation

Register your group of *exactly* five students[2] in the dedicated Moodle forum "Group project formation" by 19 January 2024 (Friday) 23:59 (Europe/London). In the post you make to the forum, you should provide the following details:

1. UCL e-mails for *all* five members,
2. a desired name for the group, and
3. a brief description of up to 100-words of the project goal.

Following the group formation each group will be assigned a responsible teaching assistant.
It should be noted that you are allowed to change the project goal after the deadline in light of for example discussions with your assigned teaching assistant.

2 Project report

For the assignment, work as a group on a machine learning project in natural language processing. You will prepare a report (maximum eight pages; excluding references) that delivers:

1. an introduction including a well-defined research hypothesis;
2. a literature review, critically discussing work related to your project;
3. a methods section describing your model/approach;
4. a clear description of your experiments;

[1] Unless extenuating circumstances applies.
[2] We reserve the right to assign students at random to groups that fail to reach this size.

5. a critical discussion of your findings including a quantitative and qualitative analysis of your empirical results; and

6. a conclusion and description of future work.

You can choose to either work on one of the default projects in Section 5 or propose your own project. We highly encourage you to choose the latter and propose your own project. Your project should have a strong machine learning component and make extensive use of natural language data.
Submit a report as a single PDF with a maximum of eight pages (excluding references and an optional appendix). The report should at least include the following:

1. An abstract briefly summarising the research problem and the outcome of your work.

2. An introduction describing the problem/research question on a high-level as well as the main outcomes of your work.

3. A related work section discussing how your work relates to prior work.

4. A methods section, detailing what you did, how you did it, and with what motivation. Make sure to formalise your models.

5. An experiments section, describing the empirical setup of your research and metrics.

6. A results and discussion section, presenting your findings as well as critically analysing your findings.

7. A conclusions section, summarising your research outcome, as well as discussion of future work if you or others were to build upon your work.

3 15-minute demonstration

You are to prepare a demonstration of your project to be carried out to at least one member of the teaching team. Prepare a few slides to describe the project and be prepared to answer probing questions regarding your approach and to be asked to run input provided by the teaching team through your model.

4 5-minute recorded presentation

In addition, we ask you to prepare and submit a recording of a 5-minute talk about your project.[3] For the recording, prepare a few slides that outline the problem setting, your approach, and results. You are encouraged to share materials between the 15-minute demonstration and 5-minute recorded presentation.

[3] You can use open source software such as OBS Studio (`https://obsproject.com`) for this.

5 Default projects

As potential default projects, you can develop a *novel* model for one of the following datasets:

- `https://nlp.stanford.edu/sentiment/code.html`
- `https://github.com/facebookresearch/QA-Overlap`
- `https://github.com/AkariAsai/XORQA`

6 Computational resources

We are sadly unable to offer substantial computational resources and we ask that you keep this in mind when deciding on the scope of your project.[4] However, there are a number of resources you can use.

6.1 UCL Computer Science Cluster

You can apply for access to the UCL Computer Science Cluster using the following URL: `https://hpc:comic@hpc.cs.ucl.ac.uk`. The cluster houses a substantial number of GPUs, but the application process can take some time and we *highly* advice that you apply as soon as possible.

6.2 Google Colab

Through Google Colab[5] you can gain access to free GPUs for a number of hours. You can of course create an account for each group member to gain additional access to GPUs. Given that the GPU allocation is time limited, you most likely want to consider regularly saving your models to be able to resume training if your allocation runs out (this is good practice in general, regardless of where you train your model).

7 Project report style

You are required to use LaTeX and the Transactions of the Association for Computational Linguistics (the leading scientific journal in natural language processing) style to typeset your report.[6] Submit a single PDF with at most eight pages (excluding references and an optional appendix) to Moodle. Use *only* your group name for the file (for example: `largest_language_models.pdf`). Do not make modifications to the template, by for example reducing or increasing the font size, margins, etc.

[4] At the time of writing, we have sent out requests for credits to use large language models. However, we have not yet received a positive from the commercial actors we have reached out to, but will keep you updated.

[5] `https://colab.research.google.com`

[6] See `https://transacl.org/ojs/index.php/tacl/about/submissions` under "Submission Format".

8 Marking and peer feedback

Your report, presentation, and recording will be assessed based on the criteria in Section 9. In addition, we ask you to anonymously comment on your collaborators regarding the quality of their contribution to your joint project using the form in Section 10.

9 Marking descriptors

Marking descriptors lifted from the UCL Computer Science MSc thesis marking guidelines for 2021/22.

9.1 90–100% Exceptional (Distinction)

- A clear contribution to the field, of publishable quality, excellent report
- Evidence of considerable extra-curricular reading, critical thought and original interpretation
- Challenging goals have been fully met, substantial deliverables, research level insight needed
- Close to faultless in execution and write-up, a high level of independence

This represents a really outstanding achievement. The project needs to clearly stand out above others. A mark in this range is hard to achieve and rare ($<$ 1%).

9.2 80–89% Outstanding (Distinction)

- Potential contribution to the field, could lead on to publishable work, very good report
- Evidence of significant extra-curricular academic, critical thought and original interpretation
- Goals met, only very minor faults in execution, depth of understanding or write-up
- Challenging project and substantial deliverables, largely self-directed

Excellent in most respects but doesn't fully meet the criteria for the top range. A small number of projects are in this range each year (2–3%).

9.3 70–79% Excellent (Distinction)

- Very well written report with a clear logical structure
- Good demonstration of critical thought, understanding and extra-curricular reading
- Some minor faults in execution or understanding, otherwise carried out effectively, most or all goals fully achieved

- A good level of challenge, substantial deliverables, and a good level of self-direction

This represents a straightforward distinction project. Most things have been done well, but there will be some faults or criticisms. The goals have been met. A reasonable number of projects can be expected to achieve this level (∼20%).

9.4 60–69% Good (Merit)

- Clear project write-up with logical structure
- Evidence of understanding, some extra-curricular reading, and critical thought
- May contain some ambiguities or faults, not all goals fully achieved
- Reasonable level of challenge, good quality deliverables, satisfactory self-management, with some supervision help needed

A good result, that is well on the way to delivering most features, but is not fully complete or finished, or has a lower level of challenge. The majority of projects are likely to be at this level.

9.5 50–59% Satisfactory (Pass)

- Adequate project write-up, lacking clarity or detail in places, or containing irrelevant material
- Limited evidence of extra-curricular reading or original thought, mostly demonstrates understanding of core issues
- Some significant deficiencies or incomplete goals, deliverables adequate but of limited quality
- Project not particularly ambitious or challenging, more significant supervision help required

A satisfactory but limited result. The core features are in place but may be buggy or not that well defined. Enough has been done to present a viable solution, of which at least some parts can be demonstrated. A minority of projects in this range.

9.6 45–49% Borderline fail (Could pass with extra work)

- Write-up is sub-standard, with noticeable errors or omissions, but could be made passable within a reasonable time
- Some clear flaws in understanding, limited or no extra-curricular reading
- Actual achievements not very substantial or challenging, deliverables of lower quality or incomplete, but could be improved fairly easily
- Not quite enough challenge or depth demonstrated, required significant extra supervision or there was a failure to attend tutorials

The project has enough substance to demonstrate it could be made into a pass in a relatively short length of time but is still missing significant features, or the write-up fails to describe what was actually achieved.

9.7 0–44% Unsatisfactory (Clear fail)

- Write-up is insubstantial, incoherent, rushed, has important omissions, or irrelevant material
- Serious flaws in understanding, little or no extra-curricular reading
- A lack of concrete achievements, substantial parts missing, few deliverables
- Serious lack of challenge or depth demonstrated, lack of engagement, required excessive supervision or there was a failure to attend tutorials

The basis of a viable project may be present but is a long way from being completed. A significant amount of additional work would be needed to reach a passable standard.

10 Peer evaluation form

This form was adapted from one developed at Johns Hopkins University in October 2006; has seen extensive use across a number of universities.

Student number:					
Attends group meetings regularly and arrives on time					
Contributes meaningfully to group discussions					
Completes group assignments on time					
Prepares work in quality manner					
Demonstrates cooperative and supportive attitude					
Contributes significantly to the success of the project					
Demonstrates energy and enthusiasm					
Total:					

Write *your own* student number in the first, left-most empty column and use the other columns for your fellow group members. Do note that you are expected to provide a self assessment by adding numbers for your own performance as well as for other members. In each column, write a number 0–3 in each column where 3 is "always", 2 is "most of the time", 1 is "sometimes", and 0 is "never" with respect to the statement in the left-most column. Lastly, please sum each column and fill in the total for the last row.