Project instructions

Marcel Bollmann

This document contains instructions for the project module in *Language Technology*.

Contents

1 Project structure

The project's main purpose is for you to investigate the capabilities and limitations of large language models (LLMs) by investigating a task of your choice, testing some models on that task, and reflecting on your results. The project work should be done in groups of ca. 6 students. A minimum project structure looks as follows:

1. **Choose a task** or an application of language technology.
2. **Choose one or more model(s)** or system(s) that you want to test on the chosen task; at least one of them needs to be an LLM.
3. **Evaluate the models** with respect to relevant properties.
4. **Draw conclusions** about the difficulty of the chosen task and the capabilities & limitations of the chosen models.

1.1 Choosing a task

"Default" project: Investigating an LLM benchmark task If you choose the "default project", you will pick an LLM benchmark task of your choice. You can use the pointers on the course website for places to find these kind of benchmark tasks, but you can also search for such tasks yourself. Here are some **guiding questions** that you can use to decide what to focus on in your project work & in your presentation:

- What benchmark task did you pick and why? What capability is this task supposed to test?
- How does the task work exactly? What are the inputs/outputs for the system? How does the evaluation work, e.g. what metrics are used?
- How do you judge the quality of the dataset? Did you find anything surprising, unusual, or problematic in the data?
- What model(s) did you evaluate and why did you pick them?
- How good were the models on the chosen task? How difficult would the task be for a human?
- What conclusions do you draw from your experiments? Do you see any limitations of the benchmark for actually testing the capability it is supposed to test?

"Individual" project: A task of your choice If you have your own ideas for a language technology task you'd like to explore, you can do that as well. In that case, you may have to do a bit more planning to decide what data to use for your project and how to evaluate your approach. You might have to modify the guiding questions a little bit, for example:

- What task did you pick and why? What capability would an LLM need to be good at this task?
- Did you use an existing dataset, or did you have to annotate your own data?

The rest of the guiding questions from above should still apply to your project.

1.2 Choosing models

You should do your own **implementation** of the chosen task and evaluate its results. For that, you will need to decide on one or more model(s) or system(s) to test. You will likely have to pick more than one model to draw meaningful conclusions about the results, as evaluation metrics rarely function as *absolute* measures of performance. However, how exactly you do that is up to you, as long as **you include at least one LLM** in the comparison. For example, you could:

- compare two (or more) LLMs against each other
- compare an LLM against a fine-tuned BERT model
- compare an LLM against a simpler statistical model like Naive Bayes
- conduct a small "human study" and compare an LLM against human performance

…provided that this makes sense for the chosen task.

1.3 Evaluating and presenting the results

If you picked an existing LLM benchmark task, you should evaluate your results in the same way that the benchmark has been evaluated before in existing work. If you defined your own task, you will need to plan yourself which evaluation approach is the best fit for the given task.

You present your conclusions in a short **group presentation** (see Section 2.3) as well as an **individual reflection paper** (see Section 2.4).

> **Important**
>
> There is **no "traditional" project report** in this course. The post-project reflection paper is not the same as a project report; please read the instructions below carefully before you start working on it!

2 Deliverables

Throughout the project, you will have to produce certain **deliverables**, which are designed to keep you on track and to give you feedback on your progress. For the exact due date on each deliverable, please see the course website and/or the Lisam submissions.

2.1 D1: Group contract

The project should be carried out in **groups of ca. 6 students**. As a first step, you will form your project group and **write a group contract**. The contract should spell out the behaviours that you expect of all group members, and procedures for resolving conflicts in the group. Specific questions to think about include:

- How will we communicate with each other?

- How often and where will we meet?
- How will we make sure that our meetings are productive?
- What will we do if somebody does not contribute to the project?
- What will we do if somebody breaks any rule set out in this contract?

> **Instructions**
>
> **Create a group contract** and have it **signed** by all members of the group. Include both the name and the LiU ID of each group member. Scan or take a good-quality picture of the signed contract and submit it as a PDF document.

After receiving your group contract, we will assign you a *group ID* that you should use in future submissions.

2.2 D2: Project plan

During the first weeks of the course, your group should communicate and/or meet regularly to plan the project. At the end of this phase, your group is required to **hand in a project plan** that should answer the following questions:

- Which language technology task do we want to investigate?
- What data are we going to use?
  - *Default project:* State what benchmark you picked and where you can access the data.
  - *Individual project:* State if you are going to use an existing dataset (and if so, which one), or whether you will have to do your own data annotation, and how.
- Which specific model(s) do we want to evaluate?
- How much implementation work will we need to do?
  - *For example:* Can you use an existing library or tool for this task, or will you have to write your own code?
- Who will be responsible for what in the project?

> **Instructions**
>
> **Write a project plan** answering the above questions and submit it as a single PDF document.

2.3 D3: Presentation & peer feedback

In the last course week before the exam, you will give a **group presentation** to present your project work. In addition, your group will be responsible for **giving peer feedback** to another group's presentation. Each group is given a time slot of 20 minutes, of which *at most* 15 minutes should be used for the presentation, and the remaining time will be used for questions and the peer feedback. Due to the number of course participants, we will have multiple presentation sessions on different days; you **only** need to attend the session at which your group is presenting (and giving peer feedback).

Presentation You are free to choose your presentation's content and structure, but we suggest that you **use the guiding questions** *(see above)* to help you think about what to include in it. The presentation needs to be understandable to everybody in the course, and the language of the presentation must be English.

Peer feedback Your peer feedback should clarify which aspects of the presentation you found most interesting and what you believe you learned about the presented task. Address at least the following points:

- Summarize the presented project in at most three sentences.
- Which parts of the presentation were clear, which parts could be improved?
- What did you find particularly interesting or surprising about the dataset and/or the evaluation method?
- Did you find the results surprising? In your opinion, what do the results tell us about the capabilities and limitations of the tested model(s)?

Instructions

Distribute the tasks for the presentation & peer feedback among the group members. It is entirely up to you how you want to divide the tasks, as long as you reach an agreement within your group.

At the presentation day, make sure that you have designated group members who will:

- **Present your project**, following the instructions above.
- **Give feedback to another group** in the same presentation session, covering at least the points mentioned above.

2.4 D4: Post-project paper

The final project-related assignment is an **individual reflection paper**. The purpose of this assignment is to make you critically reflect on what you have learned from the project and show that you achieved the intended learning outcomes. Please see the guidelines in the next section for detailed instructions on how to write this paper.

Instructions

Write a paper according to the instructions in Section 3. The paper's **length** should be around 1,500 words (approximately 3 pages). Submit your paper as a single PDF document via Lisam.

3 Guidelines for the post-project paper

For the post-project paper, you are asked to write a **reflection paper** on your project. A reflection paper is an academic essay where you reflect on the project experience. To "reflect" here means that you critically examine your experience and what you have learned.

Structure of the paper There are many ways to write reflection papers, but for this assignment, we ask you to structure your paper into three parts:

1. **Describe your work.** What task did you investigate in the project? What was your role in the project?
2. **Examine your experience.** What concepts from the course were relevant for the project?
3. **Articulate your learning.** What did you learn? How exactly did you learn it? Why does this learning matter?

> **Important**
>
> Submissions that do not clearly follow this structure may be returned without assessment.

3.1 Part 1: Describe your work

In this part, you should describe your project experience in an *objective* way. Start by giving a brief summary of what the project was about, before describing your own role and experiences in it. You can think about the following questions when writing this part:

- What was the project about? What task did you pick and why? What results did you get?
- How did you prepare yourself for the project? Did you consult any sources (websites, blog posts, scientific papers, …)? If yes, which ones, and what did you get out of them?
- What was your role in the project? What did the others do?
- Did you have to write code to run the experiments or the evaluation? If so, what exactly did you implement?
- What feedback did you get during the project and at the project presentation?

> **Tip**
>
> Focus on aspects relevant for the learning objectives, i.e., the difficulty of language technology applications and the capabilities & limitations of large language models.

3.2 Part 2: Examine your experience

In this part, you connect your work to the course content and critically examine your project experience. You should consider at least the following questions:

- What specific concepts and skills (from the course or from any additional sources you consulted) were relevant to your project? Explain those concepts and skills so that a fellow student who is not in the course can understand them.
- How exactly did you use these concepts and skills in the project?

For a **higher grade**, you should *additionally* address one of these more advanced questions:

- How did the project experience affect, change, or enhance your understanding of the course content?
- How did the project experience affect, change, or enhance your understanding of the capabilities & limitations of large language models?
- Was your understanding of the course content and any additional reading adequate, or did the project uncover some understanding that was still lacking?

> **Tip**
>
> Think about what you found especially interesting, surprising, or hard.

3.3 Part 3: Articulate your learning

In the last part, you summarize your reflections by explicitly articulating what you have learned. You should consider at least the following questions:

- What did you learn? Don't just state facts; focus on concepts and skills that are relevant to the course content.
- How exactly did you learn it? Refer back to specific details from the previous parts.
- What value does that learning have for you, and how will you use it?

4 Assessment

This section describes the exact assessment criteria for the deliverables that contribute to your grade, as well as how your final grade is computed.

4.1 Group part: Project complexity and presentation

Passing requirements These are the minimum requirements to receive a passing grade:

1. Your project represents an appropriate amount of work for this course.
2. Your presentation is clear about the method, results, and analysis, so that any student in the course would be able to follow.
3. The methodology and analysis is correct and appropriate for the chosen task.

Higher grade requirements For a higher grade, you need to fulfill all the criteria given above for passing, and additionally fulfill the following:

1. Your project thoroughly investigates the chosen task and contains more results and/or analyses than expected.
2. Your presentation is very clear and pedagogical, so that even students who did not take the course would be able to follow.
3. Your conclusions are convincingly supported by proper experiments, comparison to related work, and/or a good discussion of limitations of your approach.

4.2 Individual part: Post-project paper

Passing requirements These are the minimum requirements to receive a passing grade:

1. The paper needs to be clearly written, with appropriate language for an academic essay.
2. You give a clear description of the chosen task, the main results of your project, and your role in it. *(Part 1)*
3. You connect the work you did in the project to relevant concepts and skills from the course, and explain them so that a student not in the course could understand them. *(Part 2)*
4. You clearly describe what you learned, how exactly you learned it, and how this is relevant to the learning objectives. *(Part 3)*

Higher grade requirements For a higher grade, you need to fulfill all the criteria given above for passing, and additionally fulfill the following:

1. You refer to reading material or other sources outside the course, and describe how they were relevant for the project and significant for your learning.
2. You critically reflect on the concepts and skills you used for the project, and your understanding of them.
3. You critically reflect on the value of what you learned and how it extends beyond this specific project.

4.3 Final grade

Your overall grade for the project module takes into account both the group part and the individual part, with the individual part being given slightly more weight. Based on the criteria above, we will use a holistic assessment to assign your overall grade, following these principles:

- **Grade F/Fx**: You do not meet all of the passing requirements.
- **Grade E**: You meet all of the passing requirements, but none of the higher grade requirements.
- **Grades D, C, B**: You meet all of the passing requirements, and partially fulfill the higher grade requirements.
- **Grade A**: You meet all of the passing requirements, and all of the higher grade requirements.

> **Reminder**
>
> If your course code is TDP030, the ECTS grades are converted to your grading scale as follows:

729G86	Fx	E	D	C	B	A
TDP030	U	3	3	4	4	5