NLP project

Fegyó Benedek

January 2026

1 Introduction

The following document explains the task for the Natural Language processing assignment and describes the methods and design choices that were used to create a prompt-chosen-reject dataset based on Hungarian Matura exams and Wikipedia articles.

2 Input data

The LLM was used with a few-shot prompting technique for which a custom prompt was curated for the task specified. It contains three different examples that are added to ensure that the final generated output aligns with the task. The sources are from three different time periods to widen the context. All three examples are from past Hungarian higher level Matura exams, in a way that the correct answer is used as the chosen example, while an incorrect answer is used for the rejected one.

3 Structure of the data

The examples and the final generated data both follow the structure of

```
... = {
    "prompt": [
        {
            "role": "user",
                "content": f"Forrás szöveg:  Kérdés: ..."
        }
    ],
    "chosen": [{"role": "assistant", "content": "..."}],
    "rejected": [{"role": "assistant", "content": "..."}],
}
```

4 Source files for generation

The sources for the generation were crawled from the Hungarian Wikipedia in a recursive manner. A depth of 2 was chosen so that the dataset covers the entire time period. This is necessary as the Matura exam also covers everything up to the modern age. **This came with one drawback**: which is that due to the structure of the sources a depth of three would be overly specific and a depth of two would only yield approximately 500 pages. I choose the latter because of time and resource constraints as the generation of 1500 triplets took approximately 25 hours without accounting for the crawling and pre-processing steps.

5 Validating the generaton

The data was validated using the following formula: $score = s(chosen) + [s(chosen) - s(rejected)]$, where $s(.)$ is given by the model based on relevance and accuracy. This yields a score between $[-5; 10]$ where -5 means that the rejected answer is extremely accurate and relevant while the chosen one is inaccurate. Sampling 50 examples from the final dataset consistently yielded a score greater than 9.