Natural Language Processing and Large Language Models

Corso di Laurea Magistrale in Ingegneria Informatica

Lesson 16

Final Project

Nicola Capuano and Antonio Greco

DIEM – University of Salerno

Outline

- Goal of the project

- Tools to use for the project

- Chatbot evaluation procedure

Goal of the project

Goal of the project

- The goal of the project is to create a chatbot that answers questions about the NLP and LLM 2024/2025 course.

- The questions can be not only related to the topics covered in the course, but can also concern more general information, such as the course teachers, recommended books, etc.

- Furthermore, the chatbot must answer ONLY questions related to the course, recognizing out-of-context questions and replying that it is not enabled to provide answers on topics outside the context of interest.

- The groups will deliver the code of the chatbot and a report in which they describe in detail their solution.

Tools to use for the project

Tools to use for the project

- Students are free to use any tool or technology analyzed in the course, both the more recent ones associated with LLMs and the more classic ones of NLP.

- To achieve the required result, the students can develop parts of the chatbot with LLMs and other parts with more classic tools, with total freedom of choice as long as the various solutions identified are justified in the final report.

- Any LLM or other already available model can be used with or without modifications, as long as the groups are fully familiar with the tools they used and are able to answer questions about every aspect of the code and models.

Chatbot evaluation procedure

Chatbot evaluation procedure

- Before the project discussion, the course instructors will question the chatbot of the group in real time with a pre-defined set of questions on the course and evaluate the answers from the following points of view:

  - **Relevance**: Judges if generated text answers the query.
  - **Fluency**: Evaluates the readability and grammar of the output.
  - **Coherence**: Checks logical flow and consistency.

- Then, the course instructors will do another set of pre-defined questions to verify:
  - **Robustness**: Evaluates resistance to adversarial misleading prompts (e.g. are you sure?).
  - **Precision**: Evaluates the capability to recognize out-of-context questions (e.g. who is the king of Spain?)

- The course instructors will give a grade to the performance according to these aspects.

Natural Language Processing and Large Language Models

Corso di Laurea Magistrale in Ingegneria Informatica

Lesson 16
Final Project

Nicola Capuano and Antonio Greco

DIEM – University of Salerno