LucidQA: knowing when and how to answer

Courses: MEI; MCD
Supervisor: Sérgio Matos (aleixomatos@ua.pt)
Collaborator: Tiago Almeida (tiagomeloalmeida@ua.pt)

Context and objectives

Recent advancements in large language models (LLMs) have revolutionised how we interact with technology, enabling models to "understand" and generate human-like text, and directly respond to natural language queries. Apart from the popular and sometimes amusing "hallucinations", LLMs display other limitations such as the lack of depth in specific subjects. In question answering, an effective approach is to condition a generative language model on a text containing information that helps answer the question. However, a key challenge is ensuring entanglement between the answer and the context provided.

The objective of this work is to investigate and implement methods for: a) determining which parts of an answer are based on information from the context and which are based on the model's internal information; b) validating if the context provides relevant information for answering the question with high confidence; c) deciding if and how to answer, minimizing hallucination.

This work will also give continuity to the group's successful participation in the BioASQ challenge (http://www.bioasq.org/participate/eleventh-challenge-winners).

Technologies and frameworks: The work will be based on deep learning frameworks, namely PyTorch and HuggingFace.

Work plan

- Literature review on LLMs and neural information retrieval (IR)
- Development of context and answer validation methods
- Validation on available benchmark datasets
- Application and validation of the methods on the BioASQ challenge (subject to results and timeline of work)
- Dissertation writing