AI Application -Proposal

Team Insiköörit - 10.2024

1. The Problem

HAMK has a lot of documentation and guidelines for authoring the theses, but it can be overwhelming and difficult to navigate, since the information is all over the place in different documents and websites. Our idea for the project is to expand a pretrained LLM model to include HAMK's thesis documentation. This way the user of the application can easily search for information and the technical details of how to author the thesis. This way searching for help will become easier and faster so that the user can focus on what is the most important. The contents of the thesis.

2. The Model

The base model for the project is either Llama 3.2, Gemma 2 or a similar LLM. Pretrained Large Language Models that will provide better general knowledge and can provide help with formal writing techniques and general writing tips. These models will be fed HAMK's thesis documents and guidelines, so that it will be specialized in helping with HAMK's specific assumptions and technical details of writing.

3. The Development Approach

How do we do this? This project will be developed together by the team by feeding the base model with additional information on top of the base. The model will be uploaded to Hugging Face -spaces where it can be interacted with using a Gradle based web-UI. The user can use this web-UI to interact with the model and tweak some of its' features. The additional information provided to the LLM will be given in either text form or as PDFs, that get tokenized and then trained on.

This approach will change along the way as the limitations and possibilities of the technical details become clearer, but this is the initial plan for the project. The specifics are not set in stone since we want everyone to learn as much as possible along the way. For additional information and questions, contact our group.

-The Insiköörit group
Jaakko Saarikko, Saku Karttunen, Samppa Sirnö, Wais Atifi