



IDP Project: natural language assistant of an open-source multi-physics library and simulator

Project Overview and Goal

Generally, the learning curve for open-source library is quite steep because the lack of documentation, especially when the library is large and complicate, for instant the multi-physics libraries. Nowadays, the advanced AI tools open a new way for decrease the steepness of learning curve by fine-tuning a baseline LLM model to an expert on a specific open-source library.

This project aims to **build up a continuous fine-tuning workflow for the open-source multi-physics library SPHinXsys and its python simulator SPHinXsim**. The focus will be on exploring variable approaches by which the model can be trained directly based on the Github repositories.

Key Highlights

- **Parameter-Efficient Fine-Tuning (PEFT)**, a technique like **LoRA (Low-Rank Adaptation)** will be used.
- No code-generation is required, and the simulator already has been using LLM for structured JSON generation.
- The integrated platform will be built on python3 for easy access.
- Continuous fine-tuning (retrain on new development of the libraries).

References: To explore SPHinXsys and SPHinXsim, check out from the repository:
<https://github.com/Xiangyu-Hu>