

HelpDeskAssistant: A Specialized Language Model for Operational Maintenance in Institute of High Energy Physics

Large language models, including ChatGPT, ERNIE, and Llama, have exhibited remarkable capabilities and diverse applications in the realm of natural language understanding. These models are frequently utilized for commonplace conversations and standard question resolution. Nevertheless, they fall short in providing high-quality responses to domain-specific inquiries, largely due to the absence of pre-training on relevant field text.

The Computing Center of the Institute of High Energy Physics established the HelpDesk platform to manage routine operational maintenance issues and requests. In this study, we assembled a Q&A dataset through the collection and purification of data from the HelpDesk platform. Utilizing the Xiwu large language model for high energy physics, released by the institute, we fine-tuned the model, yielding a specialized operation and maintenance model. The precision of the model's responses was augmented by integrating internal manuals and system logs into the knowledge base via Langchain technology, culminating in the release of HelpDeskAssistant.

Owing to its expertise in high energy physics and operational knowledge derived from the HelpDesk Q&A dataset and associated documents, HelpDeskAssistant significantly surpasses other existing open-source models in responding to operational queries from the High Energy Physics Research Institute, thereby showcasing superior performance.

Primary author: Mr 张, 张 (张张张张张)

Presenter: Mr 张, 张 (张张张张张)

Track Classification: Track 10: Artificial Intelligence (AI)