

Artificial Intelligence Intern (IDX: 2024-00047)

あなたの仕事内容

Description:

Retrieval-Augmented Generation (RAG) is one of the most demanded applications of LLM by any industry, which involves refining the output of a large language model by incorporating insights from an authoritative knowledge base beyond its original training datasets prior to generating a response. We are seeking a talented and motivated individual to lead a capstone project focused on Generative AI, specifically in the domain of requirements engineering using Retrieval Augmented Generation (RAG). This project offers an exciting opportunity to explore cutting-edge AI techniques, exploring the possibilities and resolving the difficulties with usual RAG techniques, domain specific finetuning for open source LLMs and contribute to advancing the field of natural language processing.

Primary Goals:

- Develop a Retrieval Augmented Generation (RAG) framework to standardize requirements expressed in natural language.
- Explore and implement methodologies to enhance the effectiveness and efficiency of requirement normalization using RAG.
- Explore and implement methods to make open-source LLMs to assist in requirement domain specific generation, including data pre-processing, model-finetuning etc.
- Evaluate the performance of the developed model through rigorous testing and comparison with existing methods.
- Document the project findings and present them in a clear and comprehensible manner.

This internship provides an excellent opportunity for talented individuals passionate about Artificial Intelligence, Generative AI, Retrieval Augmented Generation and Prompt Tuning. As a part of our team, you will collaborate with experienced practitioners and researchers, gain practical experience in state-of-the-art technologies, and contribute to pioneering research in the field of Generative AI assisted by RAG.

Activities:

- Research and understand the principles and applications of Generative AI, with a focus on Retrieval Augmented Generation.
- Design and implement a RAG framework and Generative AI model capable of normalizing requirements expressed in various forms of natural language.
- Develop methods to assist and augment the knowledge retrieval from multiple sources.
- Collect and preprocess a diverse dataset of requirements for training and evaluation purposes.
- Fine-tune the LLM parameters and optimize its performance through



ジョブID

REF572310

業務分野

エンジニアリング

勤務地

シンガポール

法的事項

**Continental Automotive
Singapore Pte. Ltd.**

iterative experimentation.

- Develop evaluation metrics and conduct thorough testing to assess the effectiveness and efficiency of the model.
- Document the project progress, methodologies, and results in a comprehensive report.
- Prepare and deliver presentations to communicate the project findings to stakeholders.

あなたのプロフィール

Requirements:

- Proficiency in machine learning and natural language processing techniques.
- Experience with deep learning frameworks such as TensorFlow or PyTorch.
- Familiarity with Generative AI models, particularly Retrieval Augmented Generation (RAG), is highly desirable. Prior experience with libraries like langchain, streamlit is a plus.
- Prior experience with storage and retrieval from vector DBs is a plus
- Strong programming skills in languages such as Python. Experience with implementing machine learning algorithms or NLP projects in academic or otherwise is advantageous.
- Ability to work independently and collaboratively in a dynamic team environment.
- Excellent communication skills, both written and verbal.
- Prior experience with AI-related projects or research is a plus.

オファー

Ready to drive with Continental? Take the first step and fill in the online application.