

LLM Engineer (Researcher)

Ho Chi Minh

Remote - Full-time

About our company

Jumpstart Disruptive Innovations, former name Tech JDI, is established by experienced operators in the IT sector, to help foreign companies establish presence in Vietnam to access its fast-growing market and large talent pool for IT development.

Talent JDI, a subsidiary of **Jumpstart Disruptive Innovations**, is a leading recruitment firm that helps link top IT talents in Vietnam to Singapore-based startups and growing businesses. With a dedicated & experienced recruitment team, who specialized in both the Singapore and Vietnam market, we enable IT talents to find the right profession in no time.

Our clients are rapidly scaling IT companies in variety industries: Covering e-commerce, IT security Consumer applications, Business software, Online-to-offline businesses.

About our Client

- Our client's company is an AI-focused start-up company. They are a small team dedicated to develop experimental AI products.

Responsibilities

- Our client is currently seeking for a talented LLM Engineer.
- As a LLM engineer, you will be responsible for developing and deploying LLM applications and infrastructure to meet business needs.
- You will also train and fine-tune foundational LLM models using state-of-the-art techniques and tools.
- You will also design LLM inference infrastructure to scalably deploy LLMs within infrastructural constraints. You will also keep up with the latest research and use-cases of LLMs and collaborate closely with LLM research teams to participate in foundation model research, specifically for training productivity-related LLMs.

Job requirements

- Experience with LLMs, including popular foundation models like Llama2, MPT
- Experience with training and fine-tuning foundational LLM models using techniques such as PEFT, Lora, QLora
- Experience with quantization techniques, such as llama.cpp, GPTQ, etc
- Experience with LLM related development, such as Llamaindex, Langchain, Vector DBs, Prompt Engineering etc
- Good to have experience running LLMs in production, such as Triton Inference Server, etc

Language:

- English: Good



About our recruitment process

In this job title, we would like to take you into the follow recruitment process:

- **Round 1: Technical interview (Online)**
- **Round 2: Paid Work Test**
- **Round 3: Technical interview (Online)**

Benefits

- Standard Benefit

For further support, do not hesitate to contact our consultant, or contact us via email recruitment@talentjdi.com or hotline 028 3526 0292

