Machine Learning Internship (Spring 2025)

Location: Fully Remote
Duration: TBD
Department: Machine Learning Team
Company: Nebesta

Nebesta is a leading innovator in improving customer service experiences using state-of-the-art machine learning technology. We seek a **Machine Learning Intern** to join our fully remote team for the Spring 2025 internship program. This is an exciting opportunity to work with the latest advancements in Large Language Models (LLMs) and contribute to the evolution of our proprietary stack.

About Nebesta:

At Nebesta, we use machine learning to enhance the customer service experience. Our team focuses on exploring and implementing cutting-edge technologies to optimize interactions between companies and their customers. Unlike many organizations that rely on third-party APIs, Nebesta has built our LLM stack from the ground up (based on open-source solutions and models), providing unique opportunities to innovate directly with our internal models and infrastructure.

Internship Overview:

As a **Machine Learning Intern** at Nebesta, you will be involved in exploring and experimenting with the latest open-source breakthroughs in LLMs. Your primary goal will be to identify new research papers, codes, and models from the field, testing and adapting them within our proprietary machine learning stack. This internship will provide hands-on experience with GPUs and state-of-the-art resources, allowing you to test the latest technologies directly.

Key Responsibilities:

- **Research and Exploration:** Stay up to date with the latest developments in the field of large language models (LLMs), including papers, open-source code, and model architectures.
- **Experimentation:** Test recent LLM innovations into Nebesta's proprietary stack.
- **Optimization and Fine-Tuning:** Collaborate with the technical team to optimize and fine-tune models for improved accuracy, performance, and efficiency.
- **Documentation and Reporting:** Document experimental results, model performance, and key insights and share findings with the team to drive continuous improvements.

Skills and Qualifications:

- **Educational Background:** Currently pursuing or recently completed a degree in Computer Science, Data Science, Artificial Intelligence, or a related field.
- **Technical Skills:** Familiarity (or willingness to learn) with machine learning frameworks such as PyTorch, TensorFlow, and Vllm/TGI. Experience with programming languages such as Python. A solid understanding of machine learning algorithms and model evaluation. Experience with code containerisation (docker/kube).
- **Research Ability:** Strong interest in staying current with the latest advancements in machine learning and natural language processing (NLP), particularly about LLMs.
- **Collaboration:** Ability to work effectively in a fully remote, collaborative environment.

What You'll Gain:

- Hands-on experience working with cutting-edge **LLM technologies** and **GPU-based machine learning models**.
- Exposure to real-world applications of machine learning in customer service automation.
- Opportunities to collaborate with a team of experienced engineers, gaining mentorship and professional development.
- Insight into the full process of integrating research into a proprietary machine learning stack.

Compensation:

- **Hourly Rate:** TBD
- **Additional Benefits:** Access to advanced machine learning tools and resources, as well as the opportunity to work with a dynamic and innovative remote team.