Software Engineer, PyTorch – Deep Learning

Description

NVIDIA is hiring Deep Learning Compiler Engineers to design and build tools used by AI engineers across the world to design, develop, and deploy AI applications scalable across thousands of GPUs. Our team is responsible for the continual delivery of outstanding experience on NVIDIA's hardware with [PyTorch](). Apply for an opportunity to collaborate with many multi-disciplinary engineering teams within NVIDIA and internationally in the PyTorch open-source community to deliver our customers the best of NVIDIA software.

What you will be doing:
As a Deep Learning Compiler Engineer, you will be an integral part of our team, contributing to the advancement of distributed deep learning training workloads through compiler technologies. Your responsibilities will include:

- Conducting in-depth performance analysis of deep learning workloads, identifying bottlenecks, functional errors, and system inefficiencies.

- Correlating these performance issues with compiler bugs or missed optimization opportunities and developing strategies to address them.

- Collaborating with our team to extend existing program transforms or craft new ones based on the recommendations from performance analyses.

- Staying up-to-date with the latest advancements in deep learning compilers, and proposing innovative solutions to improve the efficiency of deep learning frameworks.

- Rigorously testing and validating compiler optimizations to ensure the highest quality and performance of model training.

What we need to see:

- Bachelors, Masters, or Ph.D. in Computer Science or a related technical field (or equivalent experience).

- Proficiency in Python.

- Experience using machine learning frameworks such as PyTorch or JAX.

- Some knowledge of compiler concepts such as abstract interpretation, code representations (e.g. SSA form, AST), code generation, and program transformations.

- Demonstrated experience developing large software projects.

- Strong verbal and written communication skills.

Hiring organization
Candidate-1st

Employment Type
Full-time

Beginning of employment
asap

Job Location
Finland, Helsinki

Working Hours
40

Base Salary
euro EUR 57K - 106K *

Date posted
May 22, 2024

Ways to stand out from the crowd:

- Previous contributions to open-source deep learning compiler projects, such as TVM, or deep learning frameworks.

- Understanding of the internals of PyTorch and/or JAX.

- Knowledge of distributed systems, parallel computing, and CUDA programming.

- Participation in the open source community.

- Demonstrated experience working with multi-disciplinary teams.

Job Benefits

EUR 57K – 106K *