Machine Learning Systems Engineer

Description

Summary

Posted: May 15, 2024
Weekly Hours: **35**
Role Number:**200551577**

As part of Apple Cloud Platform, our team is responsible for building libraries and services which form the foundations of critically important systems at Apple. Specifically, as a ML Systems Engineer, you will focus on building software stack to run large machine learning models (generative AI, LLMs) over multiple high performance Apple Silicon SoCs.

Key Qualifications

- Practical experience running and evaluating machine learning models for quality and performance metrics
- Experience in system-level code optimisation and power/performance evaluation for ML acceleration hardware
- Large-scale server side development experience
- Experience programming in Swift, C, C++, iOS/macOS, XCode
- Bonus qualification: familiar with Apple ML stack (ANE, CoreML, MPS/Metal), high-level general distributed ML stack ([PyTorch](#)-distributed, NCCL) and high throughput inter-chip communication systems.

Description

We write software in Swift and C++ to build services and infrastructure around innovative generative AI and machine learning. We write performant and scalable frameworks to distribute and coordinate ML inference tasks to different hardware acceleration IP blocks on different SoCs; we build jobs to deploy and load models, and support high level machine learning platforms. Our software has a growing user base and the team is looking to expand to build amazing products making it a truly exciting place to work.

We're a collection of highly skilled and friendly engineers who value each other's opinions and experience. We strive for excellence and believe strongly in the quality of our output. We have formed a team of domain experts who specialise in specific core subject areas, and also have broad experience of cloud software services and platforms. You will integrate inference code into a full service stack to ensure that user traffic is served reliably and performantly, and will have a strong focus on developing code that is easy and safe to develop, update and monitor.

Education & Experience

B.S. in Computer Science or other numerate subject, with 6+ years experience; or M.S. in Computer Science or other numerate subject, with 4+ years experience

Hiring organization
Candidate-1st

Employment Type
Full-time

Beginning of employment
asap

Job Location
London, Greater London, United Kingdom

Working Hours
40

Base Salary
euro USD 50K - 93K *

Date posted
May 17, 2024

Additional Requirements

- Quality focus – produce reliable, maintainable, deliverable software
- Comfortable diving deep – working across multiple levels of abstraction
- Good at handling relationships & communication – collaborate well with colleagues across a
- wide range of functions.

How the process will look like

Your teammates will gather all requirements within our organization. Then, once priority has been discussed, you will decide as a team on the best solutions and architecture to meet these needs. In continuous increments and continuous communication between the team and stakeholders, you're part of making data play an even more important (and understood) part withing Brand New Day.

Job Benefits

USD 50K – 93K *