Machine Learning Data Engineer

ABOUT INTUITION MACHINES

Intuition Machines™ specializes in deep learning and visual domain machine learning at scale. Intuition Machines has decades of software and ML expertise. We build and operate massively scalable systems to tackle some of today's hardest problems.

In the near future it is planned to open a branch of the company in Minsk.

Want to work on cutting edge, large scale image/video data problems with Machine Learning, Computer Vision, and Distributed systems?

We are looking for experienced ML Data Engineers / Scientists, who can work alongside both our ML research and production engineering teams on projects bringing research level Deep Learning methods into production at large scale; applying efficient image data collection, cleaning, annotation, transformation, and automation techniques to scale up to Millions and Billion size image sets and apply the latest ML-ops strategies to distributed Training and Inference systems.

IMI builds ML (Machine Learning) products and services used by some of the largest companies in the world, with millions of users.
Our hCaptcha.com data labeling service already has substantial traction and is growing rapidly. Join us as we work to apply insights in active and semi-supervised learning, image retrieval and representation learning to our production ML systems.

You will have the opportunity to work alongside Computer Vision and ML experts who publish our work at top conferences like ICLR, NeurIPS, CVPR; security and distributed systems experts hailing from Apple, Amazon, Google, Stanford, MIT, etc.

You should be an expert in Python, NumPy, Pandas, with experience of deploying ML models, training via PyTorch, and familiarity of Computer Vision, Image Processing, and applications shuch as Image Classificaiton, Detection, Segmentation, localization. Knowledge of data analysis, statistics, probability theory is desirable.

Also general software development, other languages such as JavaScript, C/C++, CUDA, systems infrastructure and architecture experience, test-driven development, deployment, operations, security, and distributed systems (Kubernetes, Docker etc), are all valuable areas of competency.

Your day-to-day responsibilities will consist of:

- Creating systems and pipelines to process large volumes of image data; deploying, monitoring, running scripts. Analyzing results, creating logging and reporting systems, delivery of results to clients.
- Working with our ML research teams to bring the latest models and ideas into production. Create efficinet procesisng pipelines for these systems.
- Working with our distributed systems engineers to improve data annotation and ML training / inference systems; create, deploy, monitor and analyze jobs, add automation and improve these systems.
- Designing, architecting, communicating technical designs and improvements to high capacity, highly available systems
- Writing clearly structured, maintainable, well documented and tested code that meets our requirements and goals

Please send your CV, github, and a brief description of what most interests you, along with your dates of availability and preferred location.

You will need to issue a "sole proprietor" and attach a card in order to receive monthly wages for it.

Payment for work done is from the office in San Francisco.