JONATHAN JIN

me@jonathanj.in
linkedin.com/in/jinnovation
github.com/jinnovation
huggingface.co/jinnovation

SELECT WORK EXPERIENCE

Hinge

Staff Machine Learning Engineer — New York, NY, Apr 2024 – Current

Leading the democratization and platformization of AI at Hinge.

- Spearheading the 0-to-1 development of **generative AI platform and infrastructure** at Hinge, with a starting focus on **evals and online monitoring**. Leading vendor evaluations for solutions to monitoring generative AI systems, prompt management, A/B testing, etc. Implemented a platform for generative AI systems, prompt management, and judge-based evaluations based on OpenTelemetry, OpenInference, and Phoenix in collaboration with stakeholders across Backend, AI, and Data disciplines.
- Developed and drove year-long company-wide **model serving strategy and roadmap**. Delivered generally-available platform based on MLflow, Ray, and Spark, reducing time-to-production for net-new models by ~50%. Aligned AI department on core focuses; drove close collaboration across cross-functional platform partner teams in Data and Engineering.

Humane Intelligence

Engineering Advisor — New York, NY, Feb 2024 – Aug 2024

"Fractional" tech lead and architect. Partner with the CEO to define and execute on engineering strategies. Contribute in hands-on capacity for technical proofs-of-concepts.

- Led 0-to-1 development of platform for crowd-sourced auditing of foundation models using Vercel, Vercel AI SDK, Next.js, and the Hugging Face API; enabled over 100 "bias bounty" participants to contribute evaluations of open-source LLMs.
- Hired and managed a team of engineering contractors and an external development agency; defined, communicated, and coordinated development roadmap, led platform evaluations, and defined technical ways-of-working for a brand-new team;

Spotify

Senior Machine Learning Engineer — New York, NY, Mar 2021 – Mar 2024

Senior member of the ML Platform org. Primarily focused on managed multi-tenant compute infrastructure for Kubeflow and Ray in GCP. Later shifted focus to AI/ML governance and artifact/experiment management solutions.

- Drove early-stage evaluation and implementation of centralized, multi-tenant Ray infrastructure based on KubeRay;
- Led development of Backstage-based model governance solution in preparation for the EU Digital Services Act;
- Defined general-availability roadmap for Kubeflow-based orchestration product, centered on SLOs and SRE best practices.

NVIDIA

Senior Systems Software Engineer, AI Infrastructure — New York, NY, Dec 2019 – Jan 2021

Worked on AI infrastructure for autonomous vehicle research and development.

Twitter

Machine Learning Software Engineer — New York, NY, Aug 2018 – Dec 2019

Member of Cortex, Twitter's central ML platform organization. Worked on: workflow orchestration; experiment management.

EDUCATION

University of Chicago *B.S. Computer Science, B.A Economics* — 2011 – 2015

SPEAKING

Data Council, "Trimming the Long Tail of Production Model Ownership at Hinge" — Apr 2025
Ray Summit, "Accelerating Time-to-Production for ML at Hinge" — Sep 2024
MLconf, "Empowering Traceable and Auditable ML in Production at Spotify with Hendrix" — Mar 2022
KubeCon + CloudNativeCon, "Scaling Kubeflow for Multi-tenancy at Spotify" — Oct 2021

SKILLS

- **AI, ML, and Data**: Phoenix, OpenInference, Ray, PyTorch, Spark, Flink, MLflow, Airflow, Kubeflow, TensorFlow
- **Infrastructure and Distributed Systems**: Kubernetes, Terraform, Helm, Prometheus, OpenTelemetry, gRPC, Docker
- **Enterprise Platforms**: Amazon Web Services (AWS), Google Cloud (GCP), Databricks, Arize
- **Programming Languages**: Python, Go, C++, Java