

Jonathan Jin

Machine Learning Infrastructure Engineer

jonathanjin.in [@me@jonathanjin.in](mailto:me@jonathanjin.in) linkedin.com/in/jinnovation github.com/jinnovation

SELECT EXPERIENCE

Present 03/2021	Spotify SENIOR MACHINE LEARNING SOFTWARE ENGINEER <i>Member of ML Platform. Working on: multi-cluster, ML-centric cloud infrastructure with Kubernetes; and ML engineering SDKs for feature engineering, data processing, and model training, analysis, and deployment.</i> <ul style="list-style-type: none">➤ Driving platform unification initiative centered around ontologically-sound declarative configuration systems, based around Kubernetes Operators, for compute configuration, deployment of scheduled training pipelines, model serving, etc.;➤ Drove team's primary product (managed Kubernetes cluster infrastructure with companion SDK) to general availability; collaborated with engineering and product leadership to identify and establish baseline product excellence, e.g. API design philosophy, documentation tooling, explicit reliability guarantees (SLOs), etc.;➤ Implemented formalized SLO tracking for multi-cluster, pipeline execution infrastructure; used Terraform to canonicalize SLO definitions for monitoring and violation reporting in Google Cloud Monitoring; implemented custom Kubernetes listener to implement nuanced and domain-specific SLIs.➤ Onboarded four new junior engineers in less than a year; held biweekly one-on-ones with each to provide technical and career mentorship and address ongoing concerns. <div>Kubernetes Ray GCP Terraform TensorFlow TFX KubeFlow Prometheus gRPC</div>	New York
01/2021 12/2019	NVIDIA SENIOR SYSTEMS SOFTWARE ENGINEER, AI INFRASTRUCTURE <i>Member of AI Infrastructure. Contributor to MagLev, NVIDIA's AI infrastructure for autonomous vehicle development. Also contributed to Modulus, the deep learning SDK for autonomous vehicle R&D.</i> <ul style="list-style-type: none">➤ Initiated development of solution for "hybrid data/model parallelism" using a Ray-based parameter server design and Horovod to enable horizontally-scalable multi-task training;➤ Co-delivered a Kubernetes-based scheduling mechanism to enable priority access to cluster resources for select use cases, e.g. prep for upcoming external demos, via virtual "resource shares"; <div>Kubernetes TensorFlow Horovod Ray gRPC Bazel SwiftStack</div>	New York
12/2019 08/2018	Twitter MACHINE LEARNING SOFTWARE ENGINEER <i>Member of Cortex, Twitter's central ML platform organization. Worked on: workflow orchestration; experiment management/iteration; and overall ML engineering productivity.</i> <ul style="list-style-type: none">➤ Spearheaded initial integration of TensorFlow Extended (TFX) with legacy Airflow-based orchestration platform to increase agility of workflow development, iterative execution/experimentation, etc.➤ Enabled distributed training of TensorFlow models in Apache Mesos from an Airflow pipeline via Deepbird, Twitter's TensorFlow-based model training/evaluating/serving framework <div>Apache Airflow Apache Aurora TensorFlow</div>	New York

SKILLS

Programming Languages	Python, Go, Bash, C++, Java
Machine Learning	Kubeflow, TensorFlow Extended (TFX), TensorFlow, Ray
Distributed Systems	Kubernetes, gRPC, Docker
Infrastructure Tooling	Bazel, Prometheus, Grafana, M3, Cassandra, Apache Airflow
Cloud Infrastructure	Google Cloud Platform (GCP), Terraform

EDUCATION

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