Terence Zhi Liu

Machine Learning / Platform Engineering

linkedin.com/in/terence-z-liu terence.z.liu@gmail.com

(419) 936-4932

For detailed projects and write-ups please visit terencezl.github.io.

EXPERIENCE

- **Head of Machine Learning / Principal Engineer | Clearview AI, NY** 2021 - Present
  - Created a facial recognition algorithm that ranked **US #1** and **World #2**.
  - Researched & implemented a **deca-B** nearest-neighbor search engine **100+ times** the SOTA throughput. Fused high-profile open-source C++ libraries to push the industry boundary. Near-term 1/10 server cost & much higher capacity. Long-term $multi-M annual savings.
  - Perfected efficient model training using tens of nodes with machine & job orchestration tools.
  - Drove model efficiency using distillation training, quantization, pruning & trimming, and hardware acceleration. Streamlined deployment with model encryption and serving engines.
  - Led cross-functional collaboration to design & review multiple compact ML data pipelines & inference infra for images/videos, resulting in $100k cost reduction per recurring batch job.
  - Established foundational ML practices, tooling & platform and built a team of ML engineers.
  - Planned & executed product development & GTM of PAD feature with team.
  - Co-wrote marketing materials, and public letters. Engaged with policy officials.
  - ➢ *Liu, Ton-That, Scalable Training Data Pipeline… (2022) U.S. Patent No. 11,443,553*
  - ➢ ***New:*** *PyTorch, OpenCV, ONNX Runtime, OpenVINO, TensorRT, AWS SageMaker, GCP Vertex AI, Ray, Airflow, MLFlow, DVC, Faiss, hnswlib, FastAPI, Socket.IO, Protobuf & gRPC, RocksDB, Datadog, GitLab CI/CD*

- **Senior Software Engineer, Platform | Bloomberg LP, NY** 2017 – 2021
  - Maintained & monitored widely-used bare-metal SFTP infrastructure with 7M daily logins.
  - Designed & implemented reliable account management, auth, routing, caching, messaging for cloud-based next-gen SFTP. Wrote OS-level modules and web servers for auth interfacing with OpenSSH. Managed deployment on internal Kubernetes-based PaaS.
  - Leveraged S3 Storage as SFTP subsystem with file system emulation and cross-data-center replication & failover.
  - Spearheaded successful multi-year high-stake account migration as technical lead.
  - ➢ ***New:*** *C++, Golang, OpenSSH, Flask, SQLAlchemy, Postgres, Redis, MongoDB, RabbitMQ, Kafka, Flink, Spark, Grafana, Splunk, Humio, Jenkins, Chef, Docker, Kubernetes, OpenStack*

- **Researcher & Developer | University of Toledo, OH** 2013 – 2017
  - Specialized in materials simulations with parallel computing clusters. 14 papers published.
  - Calculated electronic ground states with gradient descent & residual minimization schemes, (non-)linear regression - similar routines as in ML/DL frameworks. Had various ML projects.
  - Authored open-source projects: *ScriptsForVASP*, *pydass_vasp*, *pyvasp-workflow*.
  - Built a materials database website with a modern stack, supporting tabulation/graphing, user auth/contribution. Project helped get a $100k research grant.
  - ➢ *Linux, Python, R, Java, NumPy, SciPy, scikit-learn, XGBoost, LightGBM, pandas, matplotlib*

EDUCATION

- **Ph.D. in Computational Physics | University of Toledo, OH | *GPA 4.0/4.0*** 2012 – 2017
  Dissertation: *A First-Principles Simulation Approach with Cluster Expansion Predictive Models*

- **B.S. in Materials Physics | Nanjing University, China** 2008 – 2012
  Trained in computer science, statistics, operations research, math, physics, etc.