Tung-San (Haley) Lai

+1 (858) 220-1205 | haleylai1617@gmail.com | haleylai.com | Haley-1617 | tslai

Education

University of California, San Diego — *La Jolla, CA*

Master of Science in Computer Science — *Sep. 2025 - Jun. 2027*

- Relevant Coursework: Agentic AI, Recommender Systems, Web Mining, Unsupervised Learning, Machine Learning, Probabilistic Reasoning in AI, Computer Vision, ML Systems

University of California, San Diego — *La Jolla, CA*

Bachelor of Science in Mathematics-Computer Science — *Sep. 2019 - Jun. 2021*

- GPA: **3.76/4.0**
- Relevant Coursework: Data Structures, Algorithms, Object-Oriented Programming, Computer Networking, Computer Architecture, Database, Statistics, Software Design, Computer Graphics, Graph Theory, AI Search Algorithms

Work Experience

TSMC, Senior Software Engineer — *Hsinchu, Taiwan*

Python | FastAPI | Spring Boot | MariaDB | REST APIs | Shapely | Three.js — *Apr. 2023 — Aug. 2025*

- Developed multithreaded, real-time parsers (10000 logs/sec) to power exposure-aware models, increasing throughput by 100+ masks/year
- Built an end-to-end ML pipeline automating data ingestion, training, deployment, and monitoring, reducing deployment time by 70%
- Designed a scalable file system using state machines and message queues to orchestrate 50000+ TB of production data across 30+ services
- Built 3DIC geometric verification engines to validate 100+ layout configurations, detecting die-gap, collision, and enclosure violations
- Collaborated with a 20+ member team to deploy a full-line 3DIC data pipeline, automating 100% of new 3DFabric business workflows

TSMC, Software Engineer — *Hsinchu, Taiwan*

Python | Java | TensorFlow | TypeScript | PL/SQL | React.js | Vue.js | Docker — *Jul. 2021 — Apr. 2023*

- Reduced mask-scrap rates by 60% (saving $3.5M/year) by developing a time-series ML anomaly detection system for mask writers
- Achieved 90% precision by developing a hybrid statistical/deep-learning model for distance and pattern-based detection for mask writers
- Re-architected seismic detection system from sequential file polling to a multithreaded socket-stream design, reducing latency by 88%
- Increased production efficiency 30% by building a full-line management platform that supports 200+ users across manufacturing sites
- Refactored codebase to Vue/React and migrated 50+ systems/services to Azure DevOps, reducing development cycles by 60%

Deloitte, Software Engineer Intern — *Taipei, Taiwan*

Python | PyTorch | Selenium | MySQL | React Native | Azure Cloud — *Sep. 2020 — Mar. 2021*

- Improved BERT-based NLP model accuracy by 10% for risk advisory data via t-SNE analysis, correcting 1000+ labels within 5000+ articles
- Developed an automated Python/Selenium pipeline to scrape and preprocess datasets for NLP models, cutting cycle time by 90%
- Engineered a cross-platform (Web, iOS, Android) workflow system to automate review and approval requests, improving efficiency by 30%

Projects

Multi-Agent Student Support System (AI Agents Intensive: Google X Kaggle) — *Nov. 2025*

Python | Google Agent Development Kit | MCP | Vertex AI | GCP

- Implementing Model Context Protocol to query ArXiv API to validate AI-generated content and personalize knowledge for students
- Deploying an adversarial agent loop where a 'Critic Agent' autonomously challenges the 'Idea Agent' to refine project proposals

Google Restaurant Rating Predictor (Recommender Systems + Open Source) — *Nov. 2025*

Python | Pandas | Scikit-learn | NLP | Sentiment Analysis | pgmpy

- Integrated NLP (TF-IDF) with statistical and probabilistic models (Latent Factor, Bayesian Network) to predict ratings, achieving 0.64 MSE
- Contributed to pgmpy by supporting node-specific priors, enhancing flexibility and accuracy in Bayesian network parameter learning

Publications and Patents

Tung San Lai, et al, "New Applications on Multi-Beam Mask Writers to Enable Mask-Making in 3nm and Beyond," in *Proc. SPIE Photomask Technology 2024*, doi: 10.1117/12.3034678

Tung San Lai, et al, 2022. Real-Time Machine Learning-Based Multivariate Time Series Anomaly Detection Control System for Reticle Manufacturing. Patent pending.

Skills

Languages	Python, Java, JavaScript, TypeScript, HTML, CSS, Shell Script, SQL, C++
Web / Backend	Spring Boot, React.js, Vue.js, Node.js, FastAPI, Selenium, REST APIs, PL/SQL, MySQL
ML / Data	TensorFlow, PyTorch, Jupyter Notebook, NumPy, Pandas, Matplotlib, Seaborn, RegEx
DevOps	Git, Linux, Docker, Azure DevOps, Azure CI/CD Pipelines, Kubernetes, ArgoCD, Grafana, Azure Cloud, Google Cloud Platform