Max Ma, PhD

hellomaxma@gmail.com, 425-985-0189, Redmond, WA

Career Summary

AI Architect with depth and breadth for real world Machine Learning solution in healthcare and life science, strong science, and engineering discipline with creative and curious mind. **7** years ML, **3** years big data and **15** years enterprise development & consulting, Built and lead mid-size ML team for ML model and solution development, MLOps and Production (GxP) in Azure/AWS/GCP. Leadership in ML **strategy, vision & roadmap** for model development, ML solution/platform, MLOps, cross-team collaboration/ML talent development and **ML best practice & culture**.

Professional Experience

ML Manager/AI Architect, Johnson & Johnson 2019.11-2023.10

Provided leadership in ML team building and management, ML model and solution development and MLOps. Developed ML best practice: model strategy with ML depth, science discipline for model development, engineering discipline (design pattern and best practice) for MLOps and ML solutions development and led several innovations for high precision ML solutions for healthcare and life science.

- **GenAI**: application in drug discovery: -omics, RWE/clinical trial, and large model training MLOps
- **Delivery**: delivered **15** projects (model and ML solution development, MLOps) and **34** Model PoC. **2** GxP Productions, designed/delivered no-code ML self-service for JnJ business of MedTech, Version, Supply Chain, Manufacture, Compliance, Quality Control, Biochemistry R&D, and Drug Discovery.
- **Research**: CV model edge optimization, image anomaly detection, GNN for casualty and GAN for missing data. QnA for NER, information retrieval for nonconformity, SSL for Microscope image.
- **Team**: started with 3 FTE and built **20** people in **7** different countries at low cost (<**$40**/hour), cultivated a culture of open/trust, and being yourself. I managed 11 projects simultaneously in 2022.

Senior Machine Learning Engineer, Microsoft Research 2016.7-2019.11

Founding member of Health AI Team

- Project and PoC: Sleep quality causal inference using wearable data, pathology image, surgical wound image, medical pill identification, and sleep study/patient bedside safety video, medical translation & outpatient care notes key event, oncology patient readmission, mental health, and Cytometry
- Research: mining weak signals from vast wearable data, deep learning model nonlinearity/fragility and Health AI Service for number of healthcare models in millions.

Senior Engineer (Big Data and ML), Microsoft 2012.2-2016.6

Founding member of Microsoft Band (wearable device), Lead big data management/pipeline/mining (3TB/day) for BI/ML and ML production integration. I joined the Band Data Science team in early 2016 as Big Data and ML engineer. Microsoft Band was discontinued after 3 years.

Founder (ZeuPa, a startup) **and Engineering Consultant** 2005.3-2012.2

Support my own startup via consultant work, then fulltime consultant of SharePoint/data/cloud.

- ZeuPa: online services for real estate agents and post-sale relationship management. (’05-’07)
- Microsoft: SharePoint Hub & Fast Search for Microsoft IT, SharePoint development for Microsoft License, Video management for Microsoft Training and wearable device backend for Microsoft Xbox
- Re-architected real estate lead generation system/SEO generates nearly $100M revenue. Data pipeline, mining/reporting, and GIS/map for crime control reporting. Data engineering for global stock trading ($70B trading/year) system, Media copyright management for SONY/ UNIVERSAL, EMI

Max Ma, PhD

hellomaxma@gmail.com, 425-985-0189, Redmond, WA

Director of Engineering/Chief Architect, Tegris/RealTech, Bellevue, WA 1999.9-2005.2

Tegris was a $100M funded startup providing online service for real estate agents and brokers. I started as the 3rd developer before becoming dir. of Eng./Chief Architect 2 years later. I managed 15-person engineering team, and architected/managed all projects ($2M/year revenue)

Project Engineer, Golder Associates Inc, Redmond, WA 1996.12-1999.9

I was a junior member of an elite group of world class experts. We did projects for US/EU governments and the United Nation. TMS (Technology Management System) was a US DOE sponsored, United Nation project. The US congress asked for a project status report after 10 years and only gave us 6-weeks in late 1996. DOE has 15+ field offices, it is impossible to collect information via phone call and fax. My manager (PhD in physics/rock blasting), co-worker (PhD in groundwater) and I decided to try out the Internet. We developed a web application to collect information, and impressed US congress and DOE, we were awarded a 2-year, $4M to continue the work. I decided to change my career to work in High Tech because I realized that the internet would change the world at the end of this project.

High Profile Civil Eng. Projects

Expert Consultant: confidential strategic projects in the Middle East ('92)**,** China Three Gorges Dam ('93), Yucca Mountain nuclear waste repository ('94), Historical slender masonry tower in Tuckey ('94), SF Bay Bridge ('95)**,** Hawaii H-3 highway ('95-'96), Hoover Dam ('97)**,** Earth dam design in Philippine ('98-'99)

EDUCATION

PhD of Science in Civil Engineering, 1995, Clarkson University, Potsdam, NY

My PhD work, ML depth with Science Discipline: My PhD was in numerical computation for highly nonlinear systems. Deep Learning feature extraction is numerical computation for highly nonlinear systems. The difference is one partial differential equation vs a set of partial differential equations: each deep learning layer is a partial differential equation problem. My PhD work has a huge influence on my ML research and practice.

Data Centric AI: Deep Learning has no boundary conditions in numerical computation, and its training data defines boundary conditions. Model is very much determined by data than algorithms, which provides mathematical "theory" for Data Centric AI, and the "know your data like your life depends on it" culture and practice I established at JnJ

Millions small task specific model: Model robustness is in context of data distribution which decides deep learning boundary conditions. Often it is in small scale (task specific) vs in the wild and would reduce degree of model nonlinearity: "learning space" is the concept I developed while I was at Microsoft Research and was able to validate this in 2 CV real world projects in JnJ: I was able to reduce 80% model size with same performance. For most real-world ML problems in healthcare, model high precision requirements will push model nonlinearity to fragility state, it would result in many (millions) small tasks specific to the model.

ML depth and science discipline: data science is science, apply open-source code to your own dataset, it is not science work, "it is lab technician work or try your luck". My PhD rigorous training of science discipline really helps me to develop ML depth, which enables me to answer questions like OpenAI GPT-4 and Google PaLM2 have different architecture, which one is better? like object detection localization and classification, which is more important in what situation? why "no intelligence at bottom of gradient descent" is true. It is hard to develop model strategy without ML depth. Without the right model strategy, ML projects could be on the wrong track for a long time and all budgets wasted.