

Matthew Wang

972-975-7628 | mat.wang25@gmail.com | [linkedin.com/in/maw2025](https://www.linkedin.com/in/maw2025) | github.com/TheMattWang

EDUCATION

Texas A&M University

Honors Bachelor of Science in Computer Science and Applied Mathematics

College Station, TX

August 2021 – December 2025

EXPERIENCE

Incoming Platform Engineer Intern

Equilibrium Energy

May 2025 – Present

Remote

Undergraduate Research Assistant

Texas A&M University

August 2024 – Present

College Station, TX

- Published Undergraduate Research Thesis on Contrastive Multimodal Electronic Health Record Learning
- Developed a SOTA model that beats benchmarks by 15% through incorporating more modalities
- Created two zero-shot evaluations by extracting data from the MIMIC-IV dataset in order to evaluate against benchmarks

Software Engineer (AI/ML)

Mercor

March 2025 – Present

Remote

- Worked with and evaluated LLMs in contract engagements with leading AI research organizations and major tech companies, enhancing model reliability, safety, and performance
- Developed standardized rating rubrics for code-related info-seeking prompts, establishing clear criteria to classify major vs. minor issues and critical items for correctness
- Implemented JSON-based rating outputs to integrate with the evaluation pipeline, ensuring consistent and efficient auto-rating of AI responses

Machine Learning Engineer Intern

Capital One

May 2024 – August 2024

Philadelphia, PA

- Implemented a model migration and optimization plan for an Ensemble Classification Model for Sensitive Data and co-authoring a paper in the Capital One Tech Blog about distributed training
- Improved Training times by over 23% saving thousands of dollars through distributed training algorithms
- Collaborated with Machine Learning Team to quantize models reducing sizes by 4x and inference times by 3x

Software Engineer Intern

Walmart Global Tech

May 2023 – August 2023

Sunnyvale, CA

- Designed a Data Pipeline to enrich data analytics on Seller Setup Metrics by creating 15 new metrics and increasing data-processing throughput by 14%
- Integrated the Canadian Catalog into the enterprise Product Classification Model with confidence filtering capabilities for a web based application leveraged by 10 Walmart Catalog Teams

PROJECTS

NoteDown | *React.JS, Docker, Flask, Vite, Gemini, OpenCV, Git*

January 2025 – May 2025

- Built an end-to-end ML pipeline that transforms handwritten notes into LaTeX using OCR, LLMs, and summarization models averaging an 80% accuracy.
- Used by 5 Professors in the Department of Computer Science to comply with ADA requirements to increase accessibility for students with disabilities
- Implemented CI/CD pipelines that deploy to containerized Docker Containers that compiled LaTeX

GPU Accelerated Matrix Multiplication | *CUDA, OpenMP, High Performance Computing*

November 2023

- Implemented high-performance matrix multiplication in CUDA, achieving a 15 % improvement in weak scaling.
- Designed thread block and grid configuration to maximize occupancy and coalesced memory access.
- Benchmarked on matrices up to 4 K×4 K; achieved a 15 % weak-scaling throughput gain

TECHNICAL SKILLS

Languages: Java, Python, C/C++, SQL (Postgres), JavaScript, HTML/CSS

Frameworks: React, Node.js, Flask, PyTorch, TensorFlow, scikit-learn, Hugging Face Transformers

Developer Tools: Git, Docker, Google Cloud Platform, AWS

Certifications: AWS Cloud Practitioner