

Xinyi Li

Ann Arbor, MI 48104 • xinyiade@umich.edu • + 1 (612) 453-8775

EDUCATION

University of Michigan-Ann Arbor

Ann Arbor, MI

Major in Computer Science; Minor in Economics

B.S expected Dec 2025

Courses: Data Structure & Algorithm, Object-Oriented Programming, Computer Organization, Advanced Regression, Web Systems, Natural Language Processing, Machine Learning, Database Management Systems

GPA: 3.91/4

PROFESSIONAL EXPERIENCE

Glimpse Diagnostic LLC.

Arden Hills, MN

Data Engineer Intern

May 2023-August 2023

- Built generative adversarial networks model using **Tensorflow** and **PyTorch** in AWS Sagemaker to augment synthetic children tympanic membrane images from 726 to 2000 for training abnormal tympanic membrane recognition
- Optimized quality metrics including Fréchet Inception Distance(FID) derived during model training process to classify images under cross validation to enable an effective diagnostic source for medical practitioners
- Pre-processed image samples using **NumPy** and **TensorFlow** (keras) for image transformation (resizing, color uniforming) and utilized CUDA for GPU acceleration to optimize model training efficiency

RESEARCH EXPERIENCE

Statistics Online Computational Resource (SOCR) Lab - University of Michigan

Ann Arbor, MI

Undergrad Research Assistant

May 2024-Jan 2025

- Run supervised and unsupervised classification to see if any experiment group cluster emerge on recasted clinical datasets (conditional mouse treatment) in R and Python
- Work on the deployment of applications on a Nursery School compute server and AWS S3, streamlining automated healthcare privacy document management in collaboration with Prof. Ivo Dinov's team

PROJECT

Instagram Ads Generator

- Fine-tuned multiple instruction-tuned LLMs (LLaMA-3.1-8B, Qwen2.5-7B, Mistral-7B) using LoRA with 4-bit quantization and Flash Attention 2 to efficiently learn influencer-specific writing styles on a single A100 GPU
- Engineered robust influencer recommendation pipeline by designing SBERT-based embeddings over KeyBERT-filtered influencer metadata and implementing FAISS similarity search with cosine and L2 metrics for fast, scalable vector retrieval
- Led multi-metric evaluation framework combining lexical diversity, novelty, cosine similarity, and LLM-as-a-Judge for content generation assessment; created GPT-based rubric to score influencer-post alignment by niche and writing style

Facebook Clone Database System Integration (Oracle SQL, Java, MongoDB): Engineered a relational database system with 40% improved query efficiency, Java application development for processing 10,000+ records with 95% reduced response time, and a SQL-to-MongoDB data pipeline achieving 60% faster retrieval for complex hierarchical data.

Webpage Development (JavaScript, SQLite3, HTML): Developed a full-stack Instagram-like web app with user authentication, file uploads, likes, comments, and follow/unfollow features using Flask, Python, HTML, CSS, and JavaScript(React.js). Built a RESTful API, integrated AJAX for dynamic updates, and optimized the frontend for a responsive single-page application experience. Deployed on AWS EC2 for public access.

MapReduce Search Engine (Python): Developed a distributed MapReduce system on Hadoop for efficient master-worker coordination and parallel task execution. Optimized job scheduling and resource allocation across nodes to improve throughput, reduce latency, and enhance fault tolerance.

TECHNICAL SKILLS

Programming Language: C++, C, Python, CSS, Java, R

Web Technologies: Javascript, HTML, MySQL, Restful APIs, AWS, Git