

Zichen Zhang (Charlie Zhang)

📞 +1-248-710-6414 | 🏠 zichenz.me | ✉️ zhangzzc@umich.edu

Last Updated: March 25, 2025





RESEARCH & TECHNICAL INTERESTS

I build at the intersection of machine learning, systems, and entrepreneurship. My research explores scalable **multimodal** foundation models—integrating touch, vision, language, and audio—by efficiently fine-tuning unimodal models and transferring cross-modal knowledge. Beyond research, I love engineering full-stack and AI-powered **software systems** and leading **consumer**-focused teams.

EDUCATION

- **University of Michigan** April 2022 - Present
B.S., Computer Science Ann Arbor, MI, USA
 - **Relevant Course Topics:** Computer Vision, LLMs, Robot Learning for Planning and Control, ML Research, ML, Distributed Systems, Operating Systems, Web Systems, Data Structures and Algorithms, Computer Organization, Logic Design, Linear Algebra, Calculus, Probability
 - **Organizations & Activities:** LSA Honors, Michigan Hackers, MHacks hackathon, V1 (Startup Hub), Undergraduate Research Opportunity Program (UROP)

EXPERIENCE

- **Supercell**  March 2025 - Present
Incoming AI Engineering Intern, AI Innovation Lab Helsinki, Finland
 - Integrating Multimodal LLMs into games to create richer, more immersive end user experiences, pushing beyond traditional input methods like mouse, screen, and keyboard.
- **Collage**  March 2024 - Present
Co-Founder & CTO Remote
 - Engineered scalable backend services in Flask and MySQL to support user data, course metadata, and AI workflows; containerized with Docker and deployed on DigitalOcean with CI/CD pipelines.
 - Built Retrieval-Augmented Generation (RAG) pipelines and web search & text-to-SQL retry agent workflows using LlamaIndex and Google Custom Search API, boosting course recommendation accuracy by 40%.
 - Designed responsive UIs in Next.js and React Native for web and mobile platforms, enabling students to explore personalized career paths; increased user satisfaction scores by 35% from user surveys.
 - Integrated Firebase for authentication and real-time updates; implemented analytics tracking and feature flags for A/B testing, improving feature rollout velocity by 50%.
 - Led a cross-functional startup engineering team using Agile Scrum to develop an AI-driven platform designed to help students find academic advising resources, clubs, internships, job opportunities, and connect with peers through shared schedules.
 - Achieved **324 users and 100+ daily active users (DAU)** as of March 24, 2025, demonstrating strong user engagement and platform adoption.
- **U-M Minji Lab**  May 2024 - Present
Research Intern Ann Arbor, MI, USA
 - Developed MIA-Sort, a Python library for multiplex chromatin interaction analysis that processes 4B+ genome-scale fragments across 6 biologically meaningful sorting schemes.
 - Achieved 5.2× speedup over baseline methods (280s → 54s per dataset) and reduced peak memory usage by 60% by reducing search space
 - Designed scalable visualizations (plots + histograms) and auto-generated summary statistics; supported parameterized batch processing across hundreds of chromatin complexes.
 - Published miasort to PyPI and actively maintained the codebase with open-source contributions.
- **U-M Direct Brain Interface Laboratory**  September 2022 - April 2023
Research Intern Ann Arbor, MI, USA
 - Deployed complex branching logic for a brain-computer interface survey in Qualtrics, supporting 150+ unique question paths across 5 participant groups and increasing survey completion rates by 28%.
 - Implemented interactive frontend functionalities using JavaScript and Qualtrics API, reducing user drop-off by 35% and improving average session time by 1.7×.
 - Streamlined survey updates and testing workflows, cutting down researcher setup time by 50% through modular and reusable code components.

PROJECTS

- **Babysitting a Small Language Model through One-Step Tree-of-Thoughts Knowledge Distillation** December 2024
Keywords: SmolLM, semi-supervised learning, causal language modeling, Chain-of-Thought (CoT), Tree-of-Thoughts (ToT) [🌐]
 - Proposed and implemented the One-Step ToT prompting framework, which simplifies multi-step reasoning into a single structured prompt, achieving a 19% success rate on the Game of 24—outperforming CoT (7%) and Multi-Step ToT (4%).
 - Fine-tuned SmolLM-360M (360M parameters) via knowledge distillation using 144 high-quality ToT-style examples generated by GPT-4o, improving its accuracy from 1% to 9%—surpassing GPT-4o with CoT prompting (7%).
 - Replicated and improved Multi-Step ToT performance with GPT-4o on Game of 24, reaching an 82% success rate, outperforming the 74% result reported for GPT-4 in Yao et al. 2023.
- **VTMo: Unified Visuo-Tactile Transformer Encoder with Mixture-of-Modality-Experts** October 2024
Keywords: multi-modality, transfer learning, InfoNCE loss, image-to-tactile retrieval [🌐]
 - Proposed VTMo, a unified visuo-tactile transformer encoder with modality-specific and cross-modal experts, combining the speed of dual-encoders and the reasoning strength of fusion-encoders.
 - Achieved 57.27% Recall@1 accuracy on the Image-to-Touch Retrieval benchmark, significantly outperforming the frozen-attention baseline (15.11%) while requiring fewer FLOPs than CLIP-style dual encoders.
 - Demonstrated faster convergence and improved training efficiency using contrastive InfoNCE loss over 3.6K vision-touch image pairs from the Touch and Go dataset.
- **GenHint (🏆 MHacks 2024 Best Developer Tool Winner)** September 2024
Tools: Groq’s AI Inference API, Llama-3-70B, Node.js, VS Code API, TypeScript, Warp Terminal [🌐]
 - Spearheaded the development of GenHint, an AI-powered VS Code extension designed to provide step-by-step coding guidance, helping students learn by guiding them through the coding process rather than just giving answers.
 - Integrated Groq’s fast inference engine to generate tailored todo steps from user comments, enabling real-time, interactive learning through a structured task breakdown.
 - Led the development of a fully functional VS Code extension, deployed on the Visual Studio Marketplace, offering features such as task structuring, elaborating with detailed explanations, and code review with best practice suggestions.
 - Collaborated with a team of 4 to develop and deploy the project within 24 hours at MHacks 2024, resulting in winning the **Best Developer Tool award**.
 - Delivered a comprehensive demo of GenHint to showcase its impact on reducing dependency on traditional AI tools by fostering deeper understanding and critical thinking among beginner developers.

PUBLICATIONS C=CONFERENCE, J=JOURNAL, P=PATENT, S=IN SUBMISSION, T=THESIS

[S.1] **Zichen Zhang, Minji Kim. (2024). MIA-Sort: Multiplex Chromatin Interaction Analysis by Efficiently Sorting Chromatin Complexes.**

CHALLENGE PRIZES

- **Honorable Mention Best Developer Tool** September 2024
MHacks, Major League Hacking [🌐]
 - One of the largest hackathons in the U.S., attracting over 550 students from leading universities in North America.

HONORS

- **University Honors** December 2022, April 2023, December 2023, May 2024, March 2025
University of Michigan [🌐]
 - Awarded to students who earned a 3.5 GPA or higher during a term.
- **James B. Angell Scholar** March 2024, March 2025
University of Michigan [🌐]
 - Awarded to students who achieve an “A” record for two or more consecutive terms.
- **William J. Branstrom Freshman Prize** March 2023
University of Michigan [🌐]
 - Award to first-term freshmen who rank in the upper 5%.
- **Yale Young Global Scholars (YYGS)** July 2021
Yale University [🌐]
 - A competitive program for outstanding high school students from around the world focusing on discussing and tackling global challenges.