

YIHENG ZHANG

608-209-7701 | zhang2968@wisc.edu | [in](#) | [G](#) | [HomePage](#)

EDUCATION

University of Wisconsin–Madison (UWM)

Madison, WI

Bachelor of Science in Computer Science (with Honors), GPA: 3.95 (Dean's list)

Sep 2024 – Jun 2026

University of Shanghai for Science and Technology (USST)

Shanghai, China

Bachelor of Science, GPA: 3.87 (WES) / 3.84 (Official)

Sep 2022 – Jun 2024

Main Courses: *Artificial Intelligence, Foundation Models, Matrix Methods in Machine Learning, Big Data Systems, Optimization, Java, Data Structures, Linear Algebra*

RESEARCH EXPERIENCE

Large Language Models for Arithmetic Reasoning | *Machine Learning*

Feb. 2025 – Jun. 2025

Advisor: Prof. Grigorios G. Chrysos

University of Wisconsin-Madison

- Designed and executed experiments covering arithmetic tasks such as addition, multiplication, parity and sort.
- Implemented modular plug-in components to enhance arithmetic reasoning ability in transformer-based large language models (LLMs);
- Compared Muse and Transformer architectures on arithmetic reasoning tasks; Muse achieved faster convergence and superior generalization on tasks with more digits across addition, multiplication, and parity, outperforming Transformers with or without learnable embeddings.

Diffusion Language Models and Remasking | *Machine Learning*

Jun. 2025 – Present

Advisor: Prof. Grigorios G. Chrysos

University of Wisconsin-Madison

- Integrated the core **LLaDA** (Large Language Diffusion Models) module for iterative denoising-based text generation; Built efficient model training codebase using **PyTorch Lightning** and **Hydra**.
- Investigated the output distribution behavior of diffusion language models; observed a tendency toward overconfident predictions on selected digits rather than uniform uncertainty.
- Developed entropy-based and edit-distance regularization methods to mitigate this issue and improve diffusion stability and calibration.

INTERNAL READING GROUP ON DIFFUSION AND LARGE LANGUAGE MODELS

Group Reading and Discussion on Diffusion-based Models | *Research Group*

Jan. 2025 – Present

- Presented weekly readings and discussions on diffusion model research within the group, covering theoretical foundations, architectural and imaging applications, and hallucination detection methods, along with other state-of-the-art studies.
- Discussed with others based on the talks; Applied insights from the reading group to the ongoing senior thesis on LLaDA-based Diffusion LLMs and interpretability of denoising trajectories.

EXTRACURRICULAR PROJECTS

Molecular Architect: 2D Bridge Construction Game Project (Unity) | *Programmer*

Jun. 2024 – Jul. 2024

- Developed three increasingly difficult game mode levels, with the third level introduces complex mechanics requiring the use of reflective properties along the borders for problem-solving to elevate player's sense of interaction.
- Implemented additional functionalities such as scene transitions and model imports, enhancing the game's versatility, and optimized the original mono instance game framework by integrating new modules, including custom music and UI panels, to improve player engagement and overall immersion.

- Game Link: 

Thousands of Phases: AVG Game Project (Unity) | *Programmer*

Sep. 2023 – Dec. 2023

- Developed the framework and implemented the game play mechanics based on the traditional Chinese background, focusing on the diversity development of individuals.
- Created a combat system with a turn-based card fighting mechanic like Slay the Spire to enrich players' experience.
- Finalized the the existing framework using **Fungus** to enhance the user experience and more accurately reflect the game's emotional tone.

LEADERSHIP EXPERIENCES

Technical Department, USST

Sep. 2023 – Jun. 2024

Department Head

- Delivered technical workshops and mentorship for the school's tech club, providing hands-on training in tools such as **After Effects (AE)** and **Photoshop (PS)**.
- Organized team members in planning, filming, and editing promotional videos for school clubs, ensuring high-quality and engaging final outputs.
- Developed and cooperated on a digital art museum in **Unreal Engine**, allowing students to explore virtual exhibitions.

TECHNICAL SKILLS

Programming Languages: Java, C/C++, SQL, Pytorch, Python, C #, Linux

Languages: Chinese(Madarine), English (Toefl: 106, GRE: 325 (Quant: 170))

Developer Tools: VS Code, Docker, gRPC, Aider

Extra Skills: Overleaf, Beamer, PS, AE, Word, Excel