Mengze Tang

+1 (608) · 722 · 0463 ◇ mengze.tang@wisc.edu
github.com/madmtang ◇ mengze.org

EDUCATION

University of Wisconsin-Madison **Dec 2024 (expected)**
B.S. in Computer Sciences & Mathematics - ***GPA: 4.00/4.00*** *Madison, WI*

- Distinctive Scholastic Achievement
- Hilldale Fellow, Advisor: Shivaram Venkataraman
- Selected Courses: (Computer Sciences) Big Data Systems, Operating Systems, Computer Networks, Database Management Systems, Algorithms; (Mathematics) Linear Algebra, Linear Optimization, Probability, Combinatorics, Calculus, Mathematical Methods in Data Science, Matrix Methods in Machine Learning

RESEARCH INTERESTS

Big Data Systems, Machine Learning Systems, Distributed Systems

RESEARCH EXPERIENCE

Research Assistant - Harvard University **July 2024 - Present**
Advisor: Juncheng Yang *Remote*

- Leverage and optimize machine learning techniques, such as Bayesian optimization, to adaptively auto-tune the configuration of slab rebalancing strategy in caching systems, in order to achieve global optimal miss ratio.
- Analyze caching traces with various slab rebalancing strategies and eviction algorithms (collected from industrial data), and identify patterns to better motivate and guide the tuning process.

Research Assistant - University of Wisconsin-Madison **Dec 2023 - Present**
Advisor: Shivaram Venkataraman *Madison, WI*

- Develop *Compass*, an easy-to-use framework that 1) automatically and efficiently optimizes the configuration of billion-scale, hierarchical, partitioned online approximate nearest neighbor search systems; and 2) provides simulated and real-world dynamic workloads for benchmarking online indexes throughout their lifetime, which will be the artifact coming along with our paper.
- Contribute to a geometric model of indexing that provides recall estimation for optimal search parameter selection by profiling indexes across different datasets, experimenting with search parameters, and measuring their performance.
- Implement, optimize, and evaluate the effectiveness of product quantization (PQ) for large-scale distributed vector search, aiming to alleviate data movement bottlenecks and improve search efficiency.

Research Assistant - Wisconsin Institute for Discovery **Jan 2023 - Present**
Supervisor: Claudia Solís-Lemus *Madison, WI*

- Train and deploy machine learning models on biological sequences. Research and develop modern transformer-based language models for tasks involving protein sequences.
- Leverage variational autoencoders (VAEs) to encode biological sequences into latent spaces for downstream ancestral sequence reconstruction tasks.

PUBLICATIONS

[2] Jason Mohoney, Devesh Sarda, *Mengze Tang*, Anil Pacaci, Shihabur Rahman Chowdhury, Ihab F. Ilyas, Theodoros Rekatsinas, Shivaram Venkataraman ***Quake: Adaptive Indexing for Online Vector Search*** - Under Submission

[1] Evan Gorstein, *Mengze Tang*, Hailey Bruzzone, Claudia Solis-Lemus ***Ancestral Sequence Reconstruction Assisted by Variational Autoencoders*** - In Preparation

AWARDS & HONORS

Hilldale Undergraduate Research Fellowship *University of Wisconsin-Madison* **2024**
Dean's List every semester at *University of Wisconsin-Madison* **2022, 2023, 2024**
University Academic Excellence Award (Top 5%) *Xi'an Jiaotong-Liverpool University* **2021, 2022**

PROFESSIONAL ACTIVITIES

Intro to Operating Systems (Comp Sci 537) **Fall 2024**
Peer Mentor for undergrad OS course at UW-Madison *Madison, WI*

· Hold office hours to mentor students in understanding operating systems concepts and assist them with projects coded in C within Linux and xv6 toy-OS environments.

Intro to the Theory of Probability (Math 431/331) **Fall 2024**
Grader for undergrad probability course at UW-Madison *Madison, WI*

· Set rubrics for and grade exams and weekly assignments for an undergraduate probability course.

SKILLS

Programming Languages	Python, Java, C/C++, SQL, Bash
Development Tools	PyTorch, Keras, Git, Docker, Spark, Faiss, LaTeX, Markdown
Languages	English (fluent), Chinese (native)