

Hongtao Hao, CS (ML) PhD Candidate

Madison, WI | hongtaoh@cs.wisc.edu | hongtaoh.com | [Google Scholar](#)

PROFESSIONAL SUMMARY

CS PhD candidate specializing in machine learning for healthcare, probabilistic modeling, GenAI systems, and data science & visualization. **Expertise** in large-scale HPC pipelines (9k+ distributed jobs), time-series modeling, LLMOps, and Retrieval-Augmented Generation (RAG). **Strong record** of transforming research into production-ready AI tools, including three open-source ML packages for multimodal clinical datasets, and full-stack RAG application. **Passionate about** designing scalable, high-performance AI systems that connect research innovation with real-world impact.

RELEVANT WORK EXPERIENCE

University of Wisconsin–Madison

PhD Researcher, Healthcare ML

Madison, WI

Aug. 2021 — May 2026

- Designed and implemented novel Bayesian probabilistic models for healthcare disease progression and subtype discovery that are validated against real-world clinical datasets (ADNI, NACC), achieving 21-46% ordering accuracy improvements and 89% disease staging accuracy improvements over state-of-the-art methods across diverse synthetic datasets.
- Designed and executed 9,000 + distributed computing jobs on CHTC HPC clusters for large-scale experiments, demonstrating strong experience with scalable ML/AI systems.
- Built and maintained three production-ready open-source Python packages (`pysaebm`, `pyjpm`, `bebms`) for probabilistic disease progression modeling, demonstrating ability to transform complex research algorithms into well-documented, pip-installable tools adopted by the broader research community.

Robert Bosch LLC

Research Intern. LLM Applications

Sunnyvale, CA

Jun 2023 — Aug 2023

- Built full-stack LLM powered data analysis application using Python, MongoDB, and Streamlit, enabling domain-specific querying and automated insight generation from corporate datasets.
- Implemented scalable evaluation pipeline using Microsoft Azure for metrics such as response quality and relevance, contributing to 2 peer-reviewed publications on LLM-human interaction patterns and LLM capabilities of domain-specific data analysis (CHI EA '24, IUI '24 Companion).

YY Lab, Indiana University Bloomington

Research Assistant (Full-time), Data Visualization

Bloomington, IN

Aug 2020 — May 2021

- Maintained COVID-19 trend visualization dashboard (<https://covid19-dashboard.pages.dev>) in D3.js under Prof. Yong-Yeol Ahn, achieving Top 10 Most Liked notebooks on Observable platform (2020).
- Automated data workflow with CI/CD pipeline, ensuring daily synchronization with Our World in Data (OWID) COVID-19 Pandemic dataset covering 200+ countries and regions.
- Implemented interactive visualization features including temporal controls (delay parameter, play speed, time-reversal), data range filtering, and user interface.

TECHNICAL SKILLS

- Languages: Python, JavaScript, R, Julia, SQL, Bash, Java
- ML & AI: pandas, NumPy, Numba, scikit-learn, PyTorch, TensorFlow, XGBoost, Hugging Face
- GenAI & LLM Systems: LangChain (retrievers, prompts, RAG chains), custom retrievers, Pinecone/Qdrant vector DB, embedding-based semantic search, LLM evaluation pipelines, prompt engineering
- Infrastructure: HPC, Docker, CI/CD, Git, Linux, Spark, MongoDB, FastAPI, Next.js, React.js, Vercel deployment, Google Cloud Platform, Microsoft Azure
- Healthcare domain: time-series modeling, Bayesian modeling, MCMC, ranking models, event-based models, disease progression, multimodal data modeling, regulatory compliance (HIPAA), cross-sectional and longitudinal EHR.

EDUCATION

University of Wisconsin-Madison

PhD in Computer Sciences

Madison, WI

Sept 2021 — May 2026

SELECT PROJECTS

Maintainer • **LLM-Powered Movie Recommendation System** • <https://htmovies.vercel.app>

- Engineered a retrieval-augmented generation (RAG) system using LangChain retrievers, vector stores (e.g., embeddings for similarity search), and LLMs for personalized movie recommendations.
- Deployed as a production web app on Vercel, enabling scalable real-time queries from users efficiently.

Maintainer • **Deep Learning for American Time Use Survey (ATUS)** • <https://atus.hongtaoh.com>

- Trained TensorFlow-based deep learning models to predict demographic time-use patterns from survey data.
- Deployed as an interactive web interface, demonstrating skills in low-latency interactive predictions.

Maintainer • **pysaebm: Open-Source Python Package** • <https://pypi.org/project/pysaebm>

- Developed and maintained an open-source Python package for event-based modeling, demonstrating strong documentation & written communication skills, and showcasing experiences with transforming research into production ready tools.

PUBLICATIONS

† These authors contributed equally to this work.

1. **Hao, H.**, & Austerweil, J. L. (2025). Bayesian Event-Based Model for Disease Subtype and Stage Inference. *Machine Learning for Health (ML4H) Symposium, & NeurIPS 2025 Workshop (Learning from Time Series for Health)*. https://raw.githubusercontent.com/hongtaoh/bebms/main/ml4h_bebms.pdf
2. **Hao, H.**, & Austerweil, J. L. (2025). Joint Progression Modeling (JPM): A Probabilistic Framework for Mixed-Pathology Progression. *Machine Learning for Health (ML4H) Symposium, & NeurIPS 2025 Workshop (Learning from Time Series for Health)*. https://raw.githubusercontent.com/hongtaoh/jpm/main/ml4h_jpm.pdf
3. **Hao, H.**, Prabhakaran, V., Nair, V. A., Adluru, N., & Austerweil, J. L. (2025). Stage-Aware Event-Based Modeling (SA-EBM) for Disease Progression. *Machine Learning for Healthcare (MLHC) Conference*. <https://proceedings.mlr.press/v298/hao25a.html>
4. Guo, J., Mohanty, V., Piazzentin Ono, J. H., **Hao, H.**, Gou, L., & Ren, L. (2024, May). Investigating interaction modes and user agency in human-llm collaboration for domain-specific data analysis. *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems* (pp. 1-9). <https://doi.org/10.1145/3613905.3651042>
5. Guo, J., Mohanty, V., **Hao, H.**, Gou, L., & Ren, L. (2024, March). Can LLMS infer domain knowledge from code exemplars? A preliminary study. *Companion Proceedings of the 29th International Conference on Intelligent User Interfaces* (pp. 95-100). <https://doi.org/10.1145/3640544.3645228>
6. Hatfield, H. R.†, **Hao, H.**†, Klein, M., Zhang, J., Fu, Y., Kim, J., Lee, J., & Ahn, S. J. (2024). Addressing Whiteness in communication scholar composition and collaboration across seven decades of ICA journals (1951–2022). *Journal of Communication*, 74(6), 451-465. <https://doi.org/10.1093/joc/jqae019>
7. **Hao, H.** 2023. Selfie-editing among young Chinese women may have little to do with self-objectification. *Current Psychology*, 1-18. <https://doi.org/10.1007/s12144-023-04616-y>
8. **Hao, H.**, Cui, Y., Wang, Z., & Kim, Y.S. (2022). Thirty-two Years of IEEE VIS: Authors, Fields of Study and Citations. *IEEE Transactions on Visualization and Computer Graphics*. <https://doi.org/10.1109/tvcg.2022.3209422>