



XINXING WU, PH.D.

ML Researcher | Biomedical AI Researcher | AI Data Scientist (U.S. Permanent Resident)

GitHub — <https://github.com/xinxingwu-uk> Homepage — <https://xinxingwu.github.io>

Phone: 617-259-6772 Lexington, KY, USA (Open to relocation) E-mail: xinxingwu@gmail.com

RESEARCH SUMMARY

Machine Learning (ML) researcher with first-author publications at **NeurIPS**, **AAAI**, **IJCAI**, and **IEEE TNNLS**, working at the intersection of ML theory, deep learning, and biomedical data science. I develop statistically grounded and interpretable ML algorithms, including feature selection, Deep Learning (DL), Graph Neural Networks (GNNs), survival analysis models, and algorithmic stability frameworks, and apply them to clinical, genomic, transcriptomic, and longitudinal health datasets, particularly Alzheimer's disease and neurodegeneration.

My research integrates rigorous algorithm design and transparent methodology with real-world biomedical impact, producing open-source pipelines (e.g., [FAE](#), [UFS](#), [DGAE](#)) that support clinical prediction, biomarker discovery, circadian analysis, and disease progression modeling. I emphasize reproducibility, transparent validation, and cross-disciplinary collaboration with clinicians, neuroscientists, and computational biologists. More recently, I am expanding my research to include practical and methodological advances in Large Language Models (LLMs).

My background aligns strongly with labs focused on ML for health, computational biology, neuroimaging, translational AI, and projects requiring both mathematical rigor and real-world clinical insight.

RESEARCH GOALS

My goal is to advance robust, interpretable, and generalizable ML methods for biomedical and multimodal data, particularly in genomics, neuroimaging, longitudinal health records, and mechanistic disease modeling. I aim to bridge ML theory with translational impact, developing algorithms that are mathematically principled yet deployable in real clinical or scientific contexts.

I seek to contribute to labs focused on ML for health, computational biology, precision medicine, and neurodegenerative disease, where methodological innovation supports diagnosis, risk prediction, biomarker discovery, and mechanistic understanding. I am particularly motivated by research that integrates representation learning, stability theory, and multimodal biomedical data to uncover interpretable and actionable biological insights.

RESEARCH INTERESTS

- **ML Theory & Algorithms:** feature selection, algorithmic stability, generalization guarantees.
- **Graph & Deep Learning:** graph autoencoders, GNNs for link prediction and structured biological/clinical data, representation learning.
- **Biomedical & Clinical AI:** survival analysis, disease progression modeling (AD/LATE), biomarker discovery, circadian patterns.
- **Multimodal Data:** omics (e.g., bulk RNA-seq and microarray), EHR, longitudinal ML (time-to-event modeling), interpretable AI.
- **LLMs & Education/Science:** model evaluation/fine-tuning, AI-generated content for teaching and scientific communication.

EDUCATION

- Postdoctoral Scholar, ML & Data Analysis - University of Kentucky, United States, 2019 - 2022
- Visiting Researcher, ML - Boston University, United States, 2018 - 2019
- Ph.D., Computer Applications Technology - East China Normal University, China, 2007 - 2011

RESEARCH & PROFESSIONAL EXPERIENCE

External Experts Program at Amazon MTurk (Part-time), 2025 - Present
Designed challenging math/analysis tasks to probe LLM reasoning; created rubrics and analyzed failure modes; suggested model improvements.

Assistant Professor at Midway University, Kentucky, United States, 2023 - Present
Conducted applied ML/AI research with open-source outputs, including LLM-driven educational tools (SSVA). Developed and delivered a broad portfolio of ML, analytics, and computing courses. Mentored 50 students in research, internships, and technical career development.

Principal Investigator at Shanghai Threebio Technology Co., Ltd. (Part-time), Shanghai, China, 2016 - 2018
Developed statistical computing and demand-forecasting pipelines for practice-oriented AI/ML applications.

Associate Professor at Shanghai Technical Institute of Electronics and Information, Shanghai, China, 2013 - 2018
Taught programming languages, mobile development, databases, and web development; led hands-on projects integrating Python, Java, and embedded systems (e.g., Arduino and Raspberry Pi).

Algorithm Engineer at Shanghai Advanced Research Institute of CAS, Shanghai, China, 2012 - 2013

Optimized data-parallel performance in Python/C-adjacent environments on many-core Tiler.

Algorithm Developer at Shanghai Alcatel Network Support Systems Co., Ltd., Shanghai, China, 2011 - 2012
Collaborated with platform engineers to integrate feature analysis algorithms into production services.

SELECTED FIRST-AUTHOR PUBLICATIONS

[1] **Xinxing Wu**, Junping Zhang, Wang Fei-Yue. Stability-based Generalization Analysis of Distributed Learning Algorithms for Big Data. *IEEE Transactions on Neural Networks and Learning Systems*, 2020, 31 (3), 801-812. [↪Paper link](#)

[2] **Xinxing Wu**, Qiang Cheng. Fractal Autoencoders for Feature Selection. *The 35th AAAI Conference on Artificial Intelligence (AAAI 2021)*. 2021. [↪Paper link](#) [↪Paper codes](#)

[3] **Xinxing Wu**, Qiang Cheng. Algorithmic Stability and Generalization of An Unsupervised Feature Selection Algorithm. *The 35th Conference on Neural Information Processing Systems (NeurIPS 2021)*. 2021. [↪Paper link](#) [↪Paper codes](#)

[4] **Xinxing Wu**, Chong Peng, Donna M. Wilcock, et al. PRIME Uncovers Circadian Oscillation Patterns and Links with Alzheimer's Disease in Untimed Genome-Wide Gene Expression Data across Multiple Regions of Human Brain. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 2021, 17 (S5), e053266. [↪Paper link](#)

[5] **Xinxing Wu**, Chong Peng, Peter T. Nelson, et al. Random Forest-Integrated Analysis in AD and LATE Brain Transcriptome-Wide Data to Identify Disease-Specific Gene Expression. *PLOS One*, 2021, 16 (9), e0256648. [↪Paper link](#) [↪Paper codes](#)

[6] **Xinxing Wu**, Chong Peng, Peter T. Nelson, et al. Deep Learning Algorithm Reveals Probabilities of Stage-Specific Time to Conversion in Individuals with Neurodegenerative Disease LATE. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 2022, 8(1):1-9. [↪Paper link](#) [↪Paper codes](#)

[7] **Xinxing Wu**, Qiang Cheng. Deepened Graph Auto-Encoders Help Stabilize and Enhance Link Prediction. *The 31st International Joint Conference on Artificial Intelligence and the 25th European Conference on Artificial Intelligence (IJCAI 2022)*. 2022. [↪Paper link](#) [↪Paper codes](#)

[8] **Xinxing Wu**, Chong Peng, Peter T. Nelson, et al. Machine Learning Approach Predicts Probability of Time to Stage-Specific Conversion of Alzheimer's Disease. *Journal of Alzheimer's Disease*, 2022, 90 (2), 1-13. [↪Paper link](#) [↪Paper codes](#)

- [9] **Xinxing Wu**, Peter T. Nelson, Qiang Cheng. Positive effect of moderate alcohol intake for AD. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 2023, 19, e063758. ↗[Paper link](#)
- [10] **Xinxing Wu**. Singing Syllabi with Virtual Avatars: Enhancing Student Engagement Through AI-Generated Music and Digital Embodiment. <https://arxiv.org/abs/2508.11872>, 2025. ↗[Paper link](#)

SELECTED PROJECTS - OPEN SOURCE

- **FAE** - Fractal Autoencoders for Feature Selection (AAAI'21). TensorFlow+Keras implementation for robust feature selection on high-dimensional data: [Repo](#)
- **UFS** - Stability & Generalization for Unsupervised Feature Selection (NeurIPS'21). Theory-backed algorithmic stability and generalization analysis for unsupervised feature selection: [Repo](#)
- **DGAE** - Deep Graph Neural Network for Link Prediction (IJCAI'22). Residual multi-scale encoder for robust graph link prediction: [Repo](#)
- **AD_Survival** - DeepSurv pipeline for clinical survival analysis. DeepSurv-based pipelines to estimate stage-specific conversion risk and provide feature importance: [Repo](#)
- **SSVA** - Singing Syllabi with virtual avatars (2025). AI-generated music + virtual avatars to present course syllabi; integrates LLMs, TTS, and multimedia pipelines: [Repo](#)
- **JS Projects** - Educational apps for teaching and ML concepts: [Repo](#)

ACADEMIC SERVICE & GRANTS

- **Program Committee Member (Selected):** NeurIPS, ICML, AAAI, IJCAI, CVPR, ICLR, ICCV, ECCV, WACV, WWW, AMIA, FSDM, MLIS, CAIT, ICBDS (Technical Chair)...
- **Journal Reviewer:** TMLR, CAAI Transactions on Intelligence Technology, Artificial Intelligence in Medicine, Heliyon, ...
- **Grant:** PI, Shanghai Talent Development Fund.

TECHNICAL SKILLS

- **ML & Stats:** feature selection, DL, GNNs, survival analysis, anomaly detection, multimodal learning, LLM finetuning, etc.
- **Data & Visualization:** Pandas, NumPy, SciPy, Matplotlib; dashboards (Plotly/Streamlit); SQL.
- **Libraries:** TensorFlow/Keras/PyTorch, Scikit-learn, Transformers, Gradio, etc.
- **Programming:** Python, R, Java, C#, JavaScript, Ruby, React/React native, PHP, etc.
- **Workflow:** GitHub, JupyterLab/PyCharm, Power BI/Tableau, Linux, GPU training (TACC; Colab)
- **Certifications:** Red Hat Certified Engineer (RHCE); Red Hat Certified System Administrator (RHCSA)

TEACHING & MENTORING EXPERIENCE

Midway University, United States

2023 – Present

Teaching: Designed and taught 19+ undergraduate and graduate courses, including ML, AI in Analytics, HCI, Database Systems, Networking, Cybersecurity, Mobile Computing, Business Analytics, and MIS Security; consistently earned 4.75–5.00/5.00 student evaluations.

Mentoring: Academic advisor for about 50 CS/MIS students; guided research projects, ML/AI portfolios, graduate school and industry placements.

Service: Math Search Committee (2023–2024); CS Search Committee (2024–2025); Faculty Leadership Council (2024–2025); Institutional Review Board (2025–Present).