Jingxuan Wu

jingxwu@unc.edu jingxuan.com Google Scholar

Last updated: October, 2025

Research Interest

I have published several papers at top-tier conferences including *NeurIPS*, *EMNLP*, *NAACL*, *ICCV*, and received the Best Paper Award at *ICLR workshop*. My research focuses on building reliable and trustworthy AI systems that bridge the gap between machines and the real world. I am particularly interested in **Human-AI Interaction** [**yang-etal-2025-ucfe**, **liu2025presenting**], **Trustworthy NLP** [**li-etal-2025-feature**] and **Socially Aware NLP** [**yang2025twinmarketscalablebehavioralso**] My vision is to enhance the social intelligence of (vision) language models, empowering them to not only understand the principles of the physical world but also to gain insight into complex social environments, enabling reliable and meaningful interactions with humans.

Education

The University of North Carolina at Chapel Hill Aug. 2025 – Dec. 2026 (Expected)
M.S. in Statistics *Chapel Hill, NC, USA*

The Chinese University of Hong Kong, Shenzhen Sep. 2021 – May 2025
B.S. (Hons) in Data Science and Big Data Technology *Shenzhen, China*

Publications

(∗ indicates equal contribution)

Conference & Workshop Papers

Journal Papers

[J1] Chi Li, Jianfeng Mao, Lingyi Li, **Wu, Jingxuan**, Lianmin Zhang, Jianyu Zhu, and Zibin Pan. **"Flight delay propagation modeling: Data, Methods, and Future opportunities"**. In: *Transportation Research Part E: Logistics and Transportation Review* 185 (2024), p. 103525.

Preprints & Technical Reports

[P1] **Jingxuan Wu**, Zhenglin Wan, Xingrui Yu, Yuzhe Yang, Bo An, and Ivor Tsang. **"OSCAR: Orthogonal Stochastic Control for Alignment-Respecting Diversity in Flow Matching"**. 2025.
Under Review at ICLR 2026.

[P2] Zhenglin Wan, **Jingxuan Wu**, Xingrui Yu, Chubin Zhang, Mingcong Lei, Bo An, and Ivor Tsang. **"FM-IRL: Flow-Matching for Reward Modeling and Policy Regularization in Reinforcement Learning"**. 2025.
Under Review at ICLR 2026.

[P3] Wang, Hanzhao*, **Wu, Jingxuan***, Pan, Yu, Li, Yumeng, Wang, Yansong, Liu, Helang, Wang, Fuqiang, and Chen, Guanting. **"LLM-Powered Predictive Decision-Making for Sustainable Data Center Operations"**. OpenReview. 2024.

[P4] Chi Li, Mingcong Lei, **Jingxuan Wu**, Yuzhe Yang, Zibin Pan, Xiongwen Qian, and Jianfeng Mao. **"Integrative Mean-Field Epidemic Model and Adaptive Graph Learning for Network-wide Delay Propagation Dynamics Prediction"**. 2024.
In Preparation for *Transportation Research Part B*.

Research Experience

Online Learning Algorithms in Multi-Agent Systems Jul. 2025 – Present
Advisors: Guanting Chen (Dept. of Statistics & Operations Research, UNC) *Chapel Hill, NC, USA*

Guidance Mechanisms to Enhance Diversity in Flow Matching Jun. 2025 – Sep. 2025
Advisors: Xingrui Yu (CFAR, A*STAR)
Ivor Tsang (Director, A*STAR CFAR; Adjunct Prof., SCSE, NTU) Singapore
• Proposed a training-free, inference-time guidance scheme for flow/rectified-flow sampling: maximize a semantic feature volume (log-det) together with time-dependent noise, while enforcing orthogonality of both controls to the base flow velocity to preserve alignment and image quality.
• Used Heun's second-order extrapolation to predict endpoints and optimized volume in semantic feature spaces (e.g., CLIP/image-tower); propagated gradients to the latent space via two VJPs, reducing complexity to $O(K^2)$.
• Added two quality safeguards: (i) project deterministic control and stochastic noise onto the subspace orthogonal to the base flow; (ii) redundancy-aware reweighting using leverage scores to upweight under-covered samples and downweight redundant ones.
• On COCO concept classification and text-to-image tasks, achieved higher diversity (Vendi score, coverage, entropy) than CADS/DPP/PG baselines under the same sampling budget, while maintaining strong quality/alignment (FID/BRISQUE, CLIP-Score) and robustness to CFG noise schedules.

Flow-Matching Rewards & Policy Regularization Apr. 2025 – Sep. 2025
Advisors: Xingrui Yu (CFAR, A*STAR)
Ivor Tsang (Director, A*STAR CFAR; Adjunct Prof., SCSE, NTU) *Singapore*
Bo An (President's Chair Professor; Head of Division of AI, College of Computing and Data Science, NTU)
• Proposed FM-IRL, a teacher–student framework: a Flow-Matching (FM) teacher provides a reward model and a policy regularizer, while a lightweight MLP student interacts with the environment and is updated online.
• Designed an FM-enhanced discriminator that turns distribution-level FM distances (expert vs. agent, with a binary condition) into rewards via a softmax transform, enabling informative and stable shaping for online RL.
• Regularized the student policy using state–action pairs generated by the FM teacher, balancing exploration and exploitation and stabilizing training under limited or suboptimal demonstrations.
• Demonstrated improved learning efficiency, generalization to perturbed initial/goal states, and robustness to suboptimal expert data across six environments, outperforming DP/FP and IRL baselines (GAIL/AIRL/VAIL/WAIL/DRAIL).

Resource Allocation Optimization for GPUs using LLMs May 2024 – Dec 2024
Advisors: Guanting Chen (Dept. of Statistics & Operations Research, UNC)
Xiaocheng Li (School of Business & Analytics and Operations, IC) *Remote*
• Predicted job runtime and energy across GPU types by encoding program code into embedding vectors and training a lightweight linear probe.
• Modeled time-varying job arrivals with a non-homogeneous Poisson process and estimated parameters via maximum likelihood.
• Designed an online scheduling algorithm that uses predictions to allocate GPUs and route jobs under uncertainty.
• Built a simulation environment using data-center traces; outperformed a greedy baseline in throughput and energy efficiency.

Research on Network-wide Delay Propagation Dynamics Prediction Dec. 2023 – Present
Supervisor: Prof. Jianfeng Mao, School of Data Science, CUHK-Shenzhen *Shenzhen, China*
• Adapted a heterogeneous Susceptible–Infected–Susceptible (SIS) contagion process to model airport congestion spreading across a network.
• Upgraded SIS by turning transmission parameters into time-dependent functions; learned them via an *adaptive graph learning* (AdapGL) structure with GAT.

• Introduced heterogeneous, dynamic, and negative recovery states; used AdapGL to predict infection and recovery trajectories.
• Benchmarked against ODE-based methods, LSTM, STGCN, and ASTGCN, where the adaptive graph approach (GAT+AdapGL) achieved superior delay-status prediction.
• Provided theoretical conditions under which delay propagation converges even with node-varying infection/recovery rates.
• Integrated a Kalman filter to enhance multi-step forecasts, significantly improving multi-horizon performance.
• Ran simulations with realistic infection rates and transmission processes to validate model correctness and the theoretical derivations.

Study on Flight Delay Propagation Modeling Oct. 2023 – Mar. 2024
Advisor: Jianfeng Mao, School of Data Science, CUHK-Shenzhen *Shenzhen, China*

• Reviewed 40 + studies on airport delay propagation across statistical, econometric, and queueing models; organized findings by model family and data scale.
• Synthesized each approach's assumptions, strengths/limitations, and suitable problem settings; drafted a comprehensive survey manuscript.

Projects

Handwriting Recognition | PyTorch, OpenCV, OCR, Transformers, LSTM, GAN Feb. 2024 – Apr. 2024

End-to-end pipeline for handwriting dataset curation, auto-labeling, and robust recognition.

• Collected and cleaned ∼10k handwriting images by suppressing page background and guideline artifacts for higher signal-to-noise.
• Applied OCR-based word boundary detection to auto-label images and organized them into a structured corpus for downstream training.
• Trained transformer/LSTM recognizers and validated labels post-recognition, yielding a +4% accuracy improvement.
• Modified a GAN-based recognition/generation setup to better capture font/style characteristics, improving robustness to handwriting variability.

Mathematical Contest in Modeling: Optimization of Great Lakes Water Level | Jan. 2024 – Feb. 2024
Python, CCA, DEMATEL, AISM, SARIMA, ANN, GA, Sobol
Meritorious Winner at MCM. Data-driven control and forecasting for Great Lakes water-level management.

• Collected historical hydro-meteorological data (evaporation, precipitation, inflow/outflow, etc.) across the Great Lakes system.
• Applied Canonical Correlation Analysis to estimate the impact matrix linking environmental factors to water-level changes; combined DEMATEL and AISM to identify key drivers and their influence levels.
• Modeled nonlinear dynamics with SARIMA and Artificial Neural Networks to obtain a dynamic network-flow formulation of the lakes.
• Established water level–flow differential equations and injected perturbations to improve robustness under environmental variability.
• Optimized control strategies via a multi-iteration Genetic Algorithm and assessed sensitivity with Sobol indices (static-impact and dynamic-flow models).

Skills

Programming Languages: Python, C/C++, HTML/CSS, LaTeX
Developer Tools: Pytorch, Transformers, LangChain, Faiss, Git, Slurm

References available upon request.