Sizhe (Alex) Xu

Website: sizhexu.com
Location: 370 Jay St, Brooklyn, New York, 11201
Github: MazelTovy
Email: sx2490@nyu.edu
Mobile: (347)-712-0812

Education

New York University Sept 2024 - June 2026
Master of Science - Urban Data Science

- **Courses**: Urban Computing & AI, Data Science, Deep Learning, Computer Vision, Large Language and Vision Models, Transportation and Logistics, Innovative City Governance, Probability and Stochastic Processes

Dalian Jiaotong University Sept 2020 - June 2024
Bachelor of Engineering - Electronic Engineering

- **Courses**: Analog Electronics, Object-Oriented Programming, Algorithm Design, Machine Learning

Publications

[1]. **Sizhe Xu***, Renzhao Liang*, Chenggang Xie, Jingru Chen, Feiyang Ren, Shu Yang, Takahiro Yabe, "Abstain Mask Retain Core: Time Series Prediction by Adaptive Masking Loss", *Advances in Neural Information Processing Systems* (2025) **Spotlight (top 5%)**. * Denotes equal contribution.

[2]. **Sizhe Xu**, Renzhao Liang, Jun Han, Qitong Sun, "A Hybrid Framework for Evaluating and Enhancing Syntactic and Semantic Diversity in Low-Resource Text Generation", *Under Review* (2025).

[3]. **Sizhe Xu**, Boyang Li, Donghak Lee, Takahiro Yabe, "Thinking on the Move (ToM): A Framework for LLM-Agent-based Reinforcement Learning in Urban Mobility Simulation", *In Progress* (2025).

[4]. Boyang Li, **Sizhe Xu**, Yulin Wu, Takahiro Yabe, "A Generalized RoPE for n-Dimensional Position Embedding", *In Progress* (2025).

Research

Abstain Mask Retain Core: Time Series Prediction by Adaptive Masking Loss
Co-first Author, Neural Information Processing Systems (NeurIPS) 2025 Spotlight Mar 2025 - Aug 2025

- **Theoretical Innovation**: Challenged conventional "long-sequence information gain hypothesis" through systematic experimentation; discovered that appropriately truncating historical data paradoxically enhances prediction accuracy by eliminating redundant features and noise.
- **Methodological Framework**: Developed AMRC (Adaptive Masking Loss with Representation Consistency) framework based on information bottleneck theory; integrated dynamic masking loss for discriminative temporal segment identification and representation consistency constraints for stable mapping relationships.
- **Performance Achievement**: Achieved significant improvements across multiple datasets; over 50% of samples exhibited improved predictive performance while effectively suppressing redundant feature learning.

A Hybrid Framework for Evaluating and Enhancing Syntactic and Semantic Diversity
First Author, Under Review Jan 2025 - July 2025

- **Framework**: Designed and implemented a hybrid framework integrating probabilistic generative models with LLMs for synthetic data generation in low-resource scenarios.
- **Innovation**: Pioneered reference-free metrics for syntactic and semantic diversity evaluation, eliminating dependency on gold standards; leveraged LLMs for realistic text refinement while maintaining semantic integrity.
- **Performance**: Enhanced downstream classification with 13% F1-score improvement; generated text achieved higher human evaluation scores for naturalness and fluency while preserving key linguistic patterns.

Urban Mobility Prediction for Commercial Site Selection Guided Study
NYU Center for Urban Science and Progress (Mentor: *Prof. Takahiro Yabe*) June 2025 - Present

- **LLM-enhanced Agent Modeling**: Integrated Dewey and Cuebiq mobility datasets for Brooklyn Downtown analysis; pioneered LLM-generated user profiles in agent-based modeling to simulate individual decision-making and aggregate mobility behaviors for realistic forecasting.
- **Predictive Analytics Framework**: Established evaluation framework benchmarking LLM-enhanced models against conventional discrete choice and gravity models; developed commercial site selection optimization through predictive mobility analytics.

Intelligent School District Advisory Service Guided Study
NYU Center for Urban Science and Progress (Mentor: *Prof. Zhaoxi Zhang*) Dec 2024 - Apr 2025

- **Data Engineering**: Designed scalable ETL pipeline for processing and integrating 15+ diverse educational datasets; implemented data quality monitoring with Great Expectations and optimized PostGIS queries for spatial joins.
- **Geospatial Analysis**: Developed interactive platform visualizing geographic and school performance data; implemented Gemini-powered recommendation system for personalized school selection advice.

Experience

Global Data Dive Competition - Skyline & Sustainability New York, NY
Best Technical Contribution Award Winner Feb 2025

- **Advanced LLM Integration**: Implemented LLaMA-3-8B model with custom RAG architecture for real-time building energy efficiency analysis; optimized vector retrieval for 3D building data querying and developed prompt engineering techniques for domain-specific responses.
- **Urban Sustainability Platform**: Created interactive 3D visualization of NYC buildings with energy compliance metrics; built end-to-end solution for policymakers and investors to assess sustainability impact through natural language queries and geospatial analytics.

National College Student Smart Car Competition - Intelligent Vision Group Dalian, China
Team Leader & Primary Contributor - First Prize Winner Dec 2022 - July 2023

- **System Integration & Leadership**: Led 5-member multidisciplinary team through complete autonomous vehicle development cycle; designed and implemented integrated control algorithms for mechanical structure, embedded systems, and machine learning components.
- **Hardware & Control Systems**: Engineered Mecanum wheel-based chassis with 3D-printed modular storage compartments; developed custom PCB with NXP RT1064 microcontroller, DRV8701 motor drivers, and sensor fusion from ICM-20602 gyroscope and 4-directional encoders for precise motion control using Kalman filtering and PID optimization.
- **Computer Vision & Automation**: Implemented MobileNetV2-based object classification system with custom dataset captured using OpenART; achieved end-to-end automation for card detection, pickup, sorting, and storage tasks through integrated mechanical arm with SPT5410 servos and electromagnetic actuators.

SnowFox Technology Co., Ltd. Remote
Embedded System Developer (Entrepreneurship) Nov 2022 - Apr 2024

- **Motion Capture & Real-time Processing**: Developed sensor fusion firmware for STM32F4 microcontrollers with MPU9250 9-DoF IMU arrays; implemented multi-rate EKF and quaternion-based orientation estimation with 200Hz sampling and efficient fixed-point computation for power-constrained wearable devices.
- **3D Visual Analytics**: Created real-time biomechanical analysis pipeline with Unity HDRP renderer, IK solver optimization, and TensorFlow Lite for edge-deployed BiRNN movement classification (96.4% accuracy); integrated performance metrics dashboard and RESTful API for cloud synchronization through AWS IoT Core.
- **Impact**: Deployed at 5 ski resorts, increasing user satisfaction by 50% and expanding customer base over 1,000 users.

Projects

Multi-modal Context-aware RAG System Dec 2024 - June 2025
Developed RAG system handling text, images, and structured data with hybrid search and hallucination detection.
Tech: LangChain, Pinecone, FAISS, HuggingFace Transformers, CLIP

Small Object Image Segmentation Mar 2024 - June 2024
Active contour framework integrating improved YOLOv8, enhancing boundary precision and segmentation accuracy.
Tech: PyTorch, OpenCV, YOLOv8, FPN, EMA, TensorRT, ONNX

Ecological Model for Fungal Biocontrol Aug 2022 - Mar 2023
Fungal growth management system for optimizing ecological balance and maximizing agricultural profitability.
Tech: Tensorflow, RLib, MADDPG, PSO, PostgreSQL, Optuna, ArcGIS, Folium

Skills Summary

Languages:	Python, Rust, C++, R, SQL, Go, JAVA, LaTeX, Swift
Frameworks:	Scikit, PyTorch, LangChain, CUDA, Django, Spring Boot, Unity, NodeJS
Tools:	SolidWorks, Docker, Kubernetes, ArcGIS, vLLM, Git, ONNX
Platforms:	Ubuntu, Kali, Raspberry Pi, ROS, NVIDIA Jetson, GCP, AWS

Honors and Awards

- Spotlight Poster Presentation at NeurIPS 2025 - Sept, 2025
- Best Technical Contribution Award - Global Data Dive Competition - Feb, 2025
- Finalist of NYU CUSP Public Data Challenge - Oct, 2024
- NYU CUSP Experiential Scholars - Sept, 2024
- First Prize of National Intelligent Car Competition - Aug, 2023
- Bronze Medal in the China Collegiate Programming Contest (CCPC) - Oct, 2023

References

Takahiro Yabe: Assistant Professor at the Department of Technology Management and Innovation and the Center for Urban Science + Progress, New York University.

Zhaoxi Zhang: Assistant Professor at the College of Design, Construction and Planning, University of Florida (formerly Postdoctoral Researcher at NYU CUSP during the mentorship period).

Tamir Mendel: Postdoctoral Researcher at the Department of Technology Management and Innovation, New York University.