

Yuhan (Victoria) Nian

📍 Houston, TX ✉ vn23@rice.edu ☎ +1-832-955-9956

Education

Rice University

BS in Statistics and BA in Asian Studies; Minor in Data Science

Aug 2022 – May 2026

- **GPA:** 3.86/4.0 | **Honors:** Rice President's Roll (2023, 2024, 2025)
- **Relevant Coursework:** (GR) Statistical Inference, Causal Analysis, Graphical Models; (UG) Probability, Time Series, Statistical Machine Learning, Linear Regression, Bayesian Statistics, Computer Algorithm, Database Tool

IES Abroad

Study Abroad in Tokyo, Japan

May 2023 – Jul 2023

Research Experience

Neuron Redistribution for Multi-Layer Transformer Interpretability

Research Assistant | Advisor: Dr. Hanjie Chen

Rice University

Mar 2025 – Present

- Fine-tuned Qwen3-1.7B, LLaMA3.1-8B, and GPT-2 baselines on classification (SST-2, IMDB), text generation (WritingPrompts), and math reasoning (GSM8K) tasks in PyTorch
- Applied cross-layer regularization on high-attention tokens to encourage monosemantic representations and on correlated token pairs to manage polysemantic interactions
- Achieved average improvements of 65% in key-token detection and 9% in model accuracy over fine-tuned baselines

Cultural Mapping Visualization Toolkit

Research Assistant | Advisor: Dr. Corey M. Abramson

Rice University

Aug 2024 – Present

- Developed an open-source toolkit for qualitative text analysis and visualization in computational social science
- Provided RoBERTa, Jaccard, and TF-IDF embeddings to user-uploaded interview texts for co-occurrence and word similarity analytics based on user-defined seed words
- Built interactive semantic-network, heatmap, and tSNE plots by Plotly and Dash to support qualitative research

Online Conformal Prediction by Super Learner Ensembling

Research Assistant | Advisor: Dr. Meng Li | Manuscript Available upon Request

Rice University

May 2024 – Present

- Applied online conformal prediction with ARCH and LSTM base learners in Scikit-learn, statsmodels, and TensorFlow, ensuring 90% empirical coverage in sequential forecasting
- Developed a Super Learner ensemble in NumPy/Pandas using convex weighting of calibrated base predictors, reducing prediction interval width by ~30% compared to the best single learner
- Benchmarked the Super Learner against AdaHedge majority aggregation on jump-diffusion simulations, achieving a ~14% reduction in interval width while preserving 90% coverage.

Publication

Conference Presentations

- Abramson, C. M., Turner, K., Arteaga, I., Hernández de Jesús, A., Ginn, B., **Nian, Y.**, & Dohan, D. (2025). *Pragmatic sensemaking: Mapping the cultural work of living with dementia*. American Sociological Association Annual Meeting — Computational Approaches to Culture and Cognition (sponsored by the Section on Social Psychology). Chicago, IL.

Under Review

- Abramson, C. M., Turner, K., **Nian, Y.**+, Arteaga, I., Hernández de Jesús, A., Ginn, B., Dohan, D. (2025). *“Pragmatic Sensemaking: Mapping the Cultural Work of Living with Dementia”*. Submitted to Social Science and Medicine.
- Abramson, C. M. and **Nian, Y.**+. (2025). *“The Cultural Mapping and Pattern Analysis (CMAP) Visualization Toolkit: Open Source Text Analysis for Qualitative and Computational Social Science”*. Submitted to Journal of Open Source Software.

Professional Experience

Xiamen C&D Paper&Pulp Group Co., Ltd.

Data Engineer Intern

Chongqing, China

Apr 2025 – Jun 2025

- Engineered data preprocessing pipeline by pandas, NumPy, SQLAlchemy to standardize paper product codes and grades, reducing data mismatch by 12%
- Developed routing optimization by using Google OR-Tools to reduce shipping costs by 32%; created a delay classifier using XGBoost for shipping department
- Deployed real-time monitories on **Streamlit** for cost estimation, route planning, and sales tracking, reducing decision turnaround from 3 hours to 30 minutes

Xiamen C&D Inc. (Fortune 500)*Data Engineer Intern***Xiamen, China***May 2024 – Jul 2024*

- Indexed MySQL tables reducing key mismatches between the warehouse and client sales systems by 35%
- Developed a Python–SQL ETL pipeline processing 5M+ nightly sales rows, with automated validation for duplicates and nulls, ensuring reliable and consistent data loads into the warehouse
- Containerized the pipeline with Docker and orchestrated runs in **Airflow**, automating deployment across dev, test, and on-prem environments and reducing new-warehouse rollout time **from days to hours**

China Merchants Bank*Financial Analyst Intern***Chongqing, China***Jul 2023 – Aug 2023*

- Analyzed 20k+ records in SQL for customer demographics, transactions, and product adoption
- Implemented KMeans clustering by pandas, and scikit-learn on normalized transaction frequency, average balance, and product mix, identifying three distinct customer groups that guided targeted cross-selling opportunities
- Built Tableau dashboards to visualize customer acquisition, retention, and cross-sell adoption, reducing manual reporting time by 40% and enabling weekly campaign reviews

Other Related Experience

Rice Machine Learning Club*Machine Learning Engineer***Rice University***Dec 2024 – Present*

- Built a real-time image synthesis tool using React for game engine mask-based input and visualization
- Developed a low-latency pipeline linking segmentation masks to PyTorch diffusion models via vector embeddings, enabling interactive, precision-guided generation with sub-2s response time

STAT410 Linear Regression*Teaching Assistant (S24, F24, S25)***Rice University***Jan 2024 – May 2025*

- Graded assignments and exams for 300+ students; wrote weekly reports summarizing common mistakes and provided feedback to the professor
- Led tutorials and prepared supplementary materials based on these recurring mistakes to strengthen students' conceptual understanding and application of linear regression in R

Rice Women in STEM*Head of Outreach***Rice University***Aug 2023 – Dec 2024*

- Led a 5-member team in managing the WiS Mentorship Program, reaching out to more than 100 high schools across Texas to promote STEM careers through educational seminars
- Developed detailed STEM-focused college preparatory courses designed for high school students

Rice Center for Language and Intercultural Communication*Language Consultant***Rice University***Aug 2023 – May 2024*

- Organized and led discussions on current Chinese news and policy to deepen cultural awareness
- Created over 80 pages of learning materials; analyzed the best practices outlined in language acquisition and Mandarin textbooks; graded daily homework and offered individualized feedback for 20+ students

Skills

Programming Languages: Python, R, SQL, C/C++, Java, JavaScript, HTML**Tools & Frameworks:** React, Figma, Excel, Tableau, Power BI, Linux, CUDA**Data Science & Machine Learning:** scikit-learn, PyTorch, TensorFlow, NumPy, Pandas, Apache Spark, Docker, AWS, Hadoop, RoBERTa, Gensim, NLTK, A/B Testing**Languages:** Native in Chinese, Proficient in English, Intermediate in Japanese