Yilun Wang

Email: ywang323@ncsu.edu
Nelson Hall 4125
Phone: +1 (857)4457232(US);+86 18588974976(China)
931 Vintage Franklin Way, Apt 320, Raleigh, NC 27606, US

Education

Ph.D. in Economics, North Carolina State University, Raleigh, North Carolina 2020 – 2025

- Dissertation: *"Essays on Asset Pricing with Deep Learning Methods and Large Language Models"*
- Committee: Profs. Mehmet Caner (Advisor), & Ilze Kalnina, & Denis Pelletier, & Zheng Li

M.S. in Economics, Chinese University of Hong Kong, Shatin, Hong Kong 2017 – 2018

B.A. in Economics, Southwestern University of Finance and Economics, Chengdu, China 2013 – 2017

Research Fields

Applied Econometrics, Language Models(LMs), MAchine Learning, Deep Learning, High Dimension Causal Inference, Policy Analysis

Job Market Paper

"Which Data Tells the Truth? A Multimodal Deep Learning Framework for Stock Movement Prediction"

- **Abstract**: This paper introduces a novel multimodal data framework for stock price movement prediction, combining stock graphical, market, and text modalities with state-of-the-art language models such as BERT. Our model addresses two key challenges in return prediction about how to extract signals from different types of data and which type of data dominates others. Using a 20-year US stock market dataset, we show that deep learning and language models efficiently capture critical features, with time series data proving more influential than graphical and text modalities. The framework's attention mechanisms and weight allocation effectively reduce conflicts between modalities. Our best-performing model, Fusion(AW), achieves a higher balanced accuracy and a Sharpe ratio of 3.05 annually, outperforming single- and dual-modal approaches. Moreover, our model shows its ability and robustness in three recent periods—the COVID-19 crash (2020), the Russia–Ukraine shock (2022–2023), and the AI-driven expansion (2023–present). This research has broad implications for investment decision-making and paves the way for further exploration of multimodal data in financial modeling.

Working Papers

"Theory-Inspired Task-Relevant Representation Learning for Incomplete Multi-View Multi-Label Learning", with T.Xu, Q.Li and T.Luo In Processing of 2026 ICLR

- **Abstract**: Multi-view multi-label learning is commonly hindered by dual data incompleteness, arising from constraints in feature collection and prohibitive annotation costs. To address the intricate yet highly practical challenges and enhance the reliability of representation extraction, heterogeneous feature fusion, and label semantic learning, we propose a Theory-Inspired Task-Relevant Representation Learning method named TITRL. From an information-theoretic standpoint, we identify the sources of view-specific information that interfere with shared representations. By introducing dual-layer constraints on feature exclusivity and label integration, TITRL constructs a general framework for task-relevant information extraction. Besides, through variational derivation, we demonstrate the existence of tractable bounds for the mutual information model that guides the optimization direction. Regarding label semantic learning, we establish flexible relationships between label prototypes by promoting the expression of sample-level label correlations. During the multi-view integration process , TITRL simultaneously incorporates early fusion through distribution information aggregation and late fusion weighted by prediction confidence, which improves the semantic stability while enabling dynamic view quality assessment. Finally, extensive experimental results validate the effectiveness of TITRL against state-of-the-art methods.

"Improving multimodal models in automated artworks evaluation with optimizing shot selection", with F.Li In Processing of 2026 CVPR

- **Abstract**: Automated artwork evaluation via Multimodal Large Language Models (MLLMs) is crucial for applications in art education and analysis, yet their performance is often unreliable in few-shot learning scenarios due to non-optimal example selection. This limitation severely hinders their practical use. To address this challenge, we investigate an optimized shot selection method designed to enhance MLLM accuracy. The core idea is to leverage clustering algorithms, specifically comparing Random, K-Means, and K-Medoids selection, to identify the most representative examples for the model's learning context. Extensive experiments demonstrate that K-Medoids-based selection significantly improves model performance and outperforms other methods, particularly in assessing dimensions like Color Richness and Picture Organization and showing robustness to data outliers. Our study concludes that strategic shot selection is a potent and vital component for developing accurate and reliable automated artwork evaluation systems.

"Learning from memory: Asset Pricing via the Sequence model and Attention Mechanism" Submitted

- **Abstract**: We propose a novel asset pricing model that combines the latent factor model with deep learning, specifically tailored for time series data. Recognizing the temporal dependencies, our model employs a recurrent neural network to define latent factors and a deep neural network to learn nonlinear factor loadings, guided by observed covariates like asset characteristics. Moreover, we propose an efficient solution with Attention Mechanism to long dependency challenge. Our model is tested on a 60-year dataset of US equity returns. It significantly outperforms benchmarks, achieving higher out-of-sample R^2 values and better Sharpe ratios for both long-short and long-only portfolios. The model's robustness is further confirmed through simulations in low signal-to-noise scenarios, typical in financial markets.

"CVRA: Asset Pricing via the Conditional Variational Recurrent Autoencoder in Asian Market" Submitted

- **Abstract**: The paper proposed a novel probabilistic dynamic factor model for investment strategies in Asian market, which improves handling of complex and noisy market conditions. Our model integrates deep learning's variational recurrent autoencoder with advanced temporal dependency modeling. A key innovation of our model is its prior-posterior learning method, which refines the model using future data to optimize posterior factors. Designed for volatile stock markets, our model effectively estimates variances from latent space distributions and predicts returns. Statistical and empirical experiments on China and Japan stock market data demonstrate the model's superior performance compared to established traditional and machine learning methods.

Work in Progress

"Views from the Fed and Market News: A multi-agent system for financial crisis prediction" with X.Han

"Can Large Language Models Predict War-Related Black Swans via Leadership Personality? " with N.Li, X.Zhang

"Relaxation-and-Refinement Policy in Improving Soft Actor-Critic (SAC) Agent for Portfolio Optimization"

Working Experience

Research Fellow, Big Data and Computer Science Department, Jinan University, China 03/2025-Now

- Conducted research on applying Large Language Models and NLP techniques to forecast Chinese and global macroeconomics from market news;
- Developed LLM-based methods and Agent System to interpret central bank and government communications (e.g., FOMC meetings, Announcements of the Ministry of Finance) and quantify their influence on financial markets through sentiment analysis, policy-volatility modeling, and macro shock extraction;
- Corporated with firms to explore how LLM-driven macro and geopolitical risk scores enhance multi-factor machine learning strategies, and cross-market asset allocation under high uncertainty environment

Quantitative Algorithm Engineer, Hithink Flush Information Network Co., Ltd, China 03/2025-Now

- Developed LLM-based multi-agent systems to detect early-warning signals of financial crises, integrating market and the Fed's news, macro data, historical events and scenario-based Chain-of-Thought reasoning to enhance crisis prediction accuracy;
- Developed an automated factor mining framework based on RD-Agent where LLM agents generate, evaluate, and refine quantitative factors using historical A-share market datasets, improving factor discovery efficiency and producing stronger IC-ranked signals;
- Participated and applied Deep Reinforcement Learning for Portfolio Construction in A-share Markets

Teaching Experience

Independent Instructor, NC State University — 2021 – Present

- Principles of Macroeconomics: Fall 2022, Spring 2023, Summer 2023, Fall 2023, Spring 2024, Fall 2024

Graduate Teaching Assistant, NC State University — 2020 – Present

- Fundamentals of Microeconomics (Master-Level): Fall 2024
- Intermediate Microeconomics: Fall 2020, Spring 2021
- Intermediate Econometrcis (Master-Level): Spring 2025

Lab Instructor, NC State University — 2020 – Present

- Principles of Macroeconomics: Fall 2021, Spring 2022

Fellowships, Awards and Grants

Jon & Kathryn Bartley Scholarship, NC State University	2020
Owens Graduate Fellowship, NC State University	2021
Toussaint Scholarship, NC State University	2022
Goins Economic Graduate Education Scholarship, NC State University	2022
Andrew & Thelma Scholarship	2023
Econ Graduate Fellowship, NC State University	2023, 2024

Conference Presentations

Southwestern Finance Association, San Antonio	2025
European Winter Meeting of the Econometric Society, Spain	2024
Asia Meeting of the Econometric Society,East&Southeast Asia conference, Vietnam	2024
CES China Conference, Hangzhou	2024
Agricurical & Applied Economics Association(AAEA) Annual Meeting (Poster Session),New Orleans	2024
NC State University BrownBag Student Seminar, Raleigh	2023,2024

Other Experience

Quant Researcher Intern for Positive Venture DAO (PVD) Inc	2024 Summer
Data Science Intern (DouBao AI Direction) for ByteDance China	2023 Summer
Market Investment Analyst for Baoneng Investment Group	2018-2019

Other Information

Programming: Python, R, C++, STATA, MATLAB, Rust, Java, Julia, MySQL

Language: English(fluent), Chinese(native)

References

Prof. Mehmet Caner (Advisor)
Thurman-Raytheon Distinguished Professor
Dept. of Economics
NC State University, Raleigh, NC
mcaner@ncsu.edu

Prof. Ilze Kalnina
Associate Professor
Dept. of Economics
NC State University, Raleigh, NC
ikalnin@ncsu.edu

Prof. Zheng Li
Associate Professor
Dept. of Agricultural & Resource Economics
NC State University, Raleigh, NC
zli42@ncsu.edu

Prof. Denis Pelletier
Professor
Dept. of Economics
NC State University, Raleigh, NC
dpellet@ncsu.edu