YIFEI DING

Department of Economics,
University of California at Riverside,
900 University Avenue,
Riverside CA 92521, USA

Tel: (951) 231-5175
Email: yding067@ucr.edu

EDUCATION

Department of Economics, University of California, Riverside *09/2019 –Present*
Ph.D. in Economics

Wang Yanan Institute for Studies in Economics, Xiamen University *09/2015 –06/2018*
Master of Arts in Applied Statistics

Wang Yanan Institute for Studies in Economics, Xiamen University *09/2013 –06/2015*
Bachelor of Arts in Mathematic Finance

College of Chemistry and Chemical Engineering, Xiamen University *09/2011 –06/2015*
Bachelor of Science in Chemistry

RESEARCH INTERESTS

Machine Learning (Prediction and Causal Inference), Econometric Theory, Applied Econometrics

WORKING PAPERS

• **Deep Learning for Individual Heterogeneity with Generated Regressors by Adversarial Training**

We develop a semiparametric framework that uses machine learning to capture individual heterogeneity and simultaneously estimate the control function using generated regressors. It allows us to deal with endogeneity or sample selection bias in various semiparametric models with individual heterogeneity.

- **Key Contributions**: Our approach involves using a well-suited deep learning framework to uncover the parameter functions and integrate the control function with generated regressors that can be easily adjusted to fit the structural configuration of the economic model. We establish Sup-norm convergence rates using adversarial training, achieving optimal min-max rates. This enables valid inference for second-stage parameters and machine learning (ML) model functions. By deriving influence functions or orthogonal scores, we enable valid inference for any second-stage parameter and various ML model parameter functions.

- **Methodology**: By employing PyTorch' s automatic differentiation, influence functions can be applied directly to data, allowing broad applicability in studying individual heterogeneity.
- **Impact**: This framework bridges advanced machine learning techniques with econometric models, which can be broadly applied to various structure parameters studying their individual heterogeneities.

• A Comparative Study of Machine Learning Models for Prediction: Insights from Tree-Based Models and Deep Neural Networks

This paper conducts a systematic comparison of prediction models across different machine learning techniques, including tree-based and deep neural network models. The research explores the effectiveness of various prediction combination methods.
- **Key Models**: Tree-based models include Boosting smooth transition regression trees, Boosting symmetric smooth additive regression trees, traditional boosting regression trees, and random forests. Deep neural network models include multilayer perceptrons and deep residual learning.
- **Combination Methods**: We categorize prediction combination methods into three main types: 1. Simple Methods: Simple average, median, Bates and Granger(1964), Newbold and Granger(1974) and inverse rank combination. 2. Regression-based Methods: OLS and constrained least squares regression combinations. 3. Eigenvector Methods: Standard eigenvector combination methods are also evaluated.
- **Findings**: Boosting-based models, especially the newly developed BoostSmooth and SMARTBoost models, perform superiorly in prediction accuracy compared to other tree-based and deep learning models. In terms of prediction combinations, the regression-based prediction combination methods and median combination robustly predominate other prediction combination methods in our simulations and empirical applications.

• Estimating Partial Effects Using Machine Learning

This work investigates causal inference by comparing partial derivative estimations using machine learning models as nonparametric alternatives for recovering regression functions.
- **Key Contributions**: Through simulations and empirical applications, we showcase the BoostSmooth model's advantages in estimating partial derivatives compared to other machine learning methods across various signal-to-noise ratios and redundant variable settings.
- **Findings**: BoostSmooth effectively recovers nonlinear function structures and performs well in identifying linear function structures, making it versatile in various modeling scenarios.

RESEARCH & PROFESSIONAL EXPERIENCE

• Applied Scientist Intern, Snap Inc. *06/2023 – 09/2023*

- **Project**: Conducted research on causal inference using double machine learning (DML) techniques for continuous treatment effects, focusing on dose-response curve estimation and marginal effects.

- **Key Methodologies**: Applied entropy balancing and DML for continuous treatments on both synthetic and semi-synthetic data derived from Snapchat user data to evaluate performance.
- **Findings**: Tree-based models like XGBoost and BooST outperformed DNNs, with BooST showing exceptional accuracy, surpassing XGBoost.
- **Outcome**: Presented findings at CODE@MIT 2023, demonstrating the application of ML in continuous treatment estimation.

• **Master Thesis, Advisor: Prof. Haiqiang Chen, Xiamen University** *09/2017 – 06/2018*
Title: *Network of Shareholders and Stock Price Synchronicity: Empirical Evidence from the A-share Market*
- Developed a social network model for the top 10 shareholders to measure network topology variables (centrality, validity, and length).
- Demonstrated a significant positive causal relationship between shareholder networks and stock price synchronicity using spatial panel data and network models.

• **Research Assistant, Prof. Haiqiang Chen, Xiamen University** *09/2016 – 08/2017*
- **Project**: *Estimating and Testing Threshold Models with Time-Varying Thresholds* funded by National Science Foundation of China (71571152).
- Developed research ideas on shareholder networks and stock market crash risk, built project databases using web crawlers (Rcurl and Rvest), and assisted in empirical analysis and paper writing.

• **Research Assistant, Prof. Haiqiang Chen, Xiamen University** *06/2016 – 08/2016*
- Employed R and NetLogo to simulate the formation and evolution of shareholder networks and their impact on financial markets.

PUBLICATIONS & CONFERENCE PAPERS

Yifei Ding and Meng Xu. *Comparing Methods for Continuous Treatment.* Presented at CODE@MIT, 2023.

Haiqiang Chen, Yang Chen, Yifei Ding. *Does "Too-Connected" Network of Shareholders Exacerbate Crash Risk?* China Economic Quarterly (经济学季刊), 2023(03):1070-1087. Available at: https://www.nsd.pku.edu.cn/pub/chnsd/docs/20230719150300278598.pdf

JOURNAL REFEREE

Journal of Quantitative Economics

TEACHING EXPERIENCE

• **Instructor, University of California, Riverside** *09/2024*
Course: Statistics in Mini-Math Camp

• **Teaching Assistant, University of California, Riverside** *09/2020 – Present*

- Graduate Courses: Econometric Method I, Micro theory (Grader)
- Undergraduate Courses: Data Analysis in Economics, Introductory Econometrics I, Statistics for Economics, Intermediate Microeconomic Theory I, Introduction to Microeconomics

• **Teaching Assistant, Xiamen University** *09/2016 – 06/2017*

- Graduate Courses: Advanced Econometrics II, Time Series Analysis
- **Award**: Best TA

PROGRAMMING SKILLS

Languages: R (Extensive), Python, GoogleCloud, SQL, LaTeX, Eviews, Stata

CONFERENCE & SEMINAR PRESENTATIONS

- The Causal AI Conference, San Francisco *2024*
- The Conference on Digital Experimentation (CODE@MIT), MIT *2023*
- Econometrics Colloquium, UC Riverside *09/2019 – Present*
- 2018 Workshop of Resource Security and Economic Sciences, Xiamen, China *05/2018*
- 1st Forum of China Econometrics, Xiamen, China *12/2017*

AWARDS & HONORS

Dissertation Year Fellowship, University of California, Riverside *2024*
Dean's Distinguished Fellowship, University of California, Riverside *09/2019 – Present*
Graduate Scholarship, University of Xiamen *09/2015, 09/2017*
Academic Excellence Scholarship, University of Xiamen *05/2015, 09/2013, 09/2012*
"Yangliangli" Scholarship, University of Xiamen *04/2013*
"Chongqing Ziguang" Scholarship, University of Xiamen *04/2012*