YEHU CHEN

☎ (734) 846-0174 ✉ chenyehu@wustl.edu ⌂ yahoochen97.github.io

SUMMARY

Ph.D. researcher at WashU STL specializing in machine learning and quantitative methods. With vast experience and deep understanding in data analytics, causal inference, experimental design, predictive and forecasting models, natural language processing, software engineering and cross-functional team collaboration. Fluent in modern programming languages and visualization tools. Research work has led to publications in top-tier conferences/journals. **Actively looking for full-time opportunities in data science starting Spring 2025.**

EDUCATION

- **Ph.D. in Data Science**, Washington University in St Louis. GPA: 3.9/4.0.   2019 - Present
  Research topics: Bayesian Machine Learning, Gaussian Process, Quantitative methods, Causal Inference

- **Bachelor of Science in Computer Science**, University of Michigan. Summa Cum Laude.   2017 - 2019
  Selective courses: Computer Vision, Reinforcement Learning, Data Mining, Deep Learning, Autonomous Robotics

RESEARCH PROJECTS

- **Dynamic Forecasting for 2020 U.S. Senate Elections**   *Publish at* **Political Analysis, 2023**
  Collaborate with *CNN* and design forecasting models for senatorial elections by collecting polling data from major polling companies using web-scraping techniques. Successfully forecast outcomes of 33/35 races in 2020 with lower MSE of predicted vote shares than other forecasters, including *The Economist* and *FiveThirtyEight*.

- **Treatment Effect Estimation in Panel Text Data**   *Publish at* **AISTATS, 2023**
  Propose and implement a novel difference-in-difference method based on Gaussian Process with more precise causal effect estimation and calibrated event predictions. Apply the model and natural language processing to analyze broadcast transcripts that deepens the understanding of supply-side roles by mainstream news media.

- **Personalised Assessment of Big-Five Personality and Behaviors**   *Publish at* **Neurips, 2024**
  Design and execute pilot studies on personalized assessment by conducting experience sampling surveys and building novel measurement models under deep learning framework. Substantive findings manage to reconcile a long-lasting psychological debate and contribute to grant award from National Science Foundation of $500,000.

- **Quantitative Latent Measurement using Dynamic Item Response Theory**   *Publish at* **APSA, 2023**
  Develop new quantitative models and publish R packages of Bayesian item response theory that provides better justification and prediction of economic and legal behaviors. Analyze large data archives such as American Panel Survey and Supreme Court dispositions that yield meaningful insights to researchers and clients.

- **Public Opinion Survey using Small-Area Estimation with Post-Stratification** *Work in progress*, **2024**
  Cooperate with National Geospatial-Intelligence Agency and build machine learning models for efficiently estimating small-area public opinion from large nationally representative surveys. Significantly reduce required sample sizes by over 25% and hence alleviating the operational and time-related demands associated with data collection.

WORK EXPERIENCE

- **Software Engineer Intern, Foxit Software Inc, Fremont, CA**   Summer 2018
  Engineered key features for enterprise solutions including automation and UI for advanced PDF integration.

- **Research Intern, Shanghai Fudan Microelectronics Group, Shanghai, China**   Winter 2017
  Conducted research to support advancements in microelectronics engineering solutions.

TECHNICAL SKILLS

- **Programming**   C/C++/C#, Python, R, Matlab, SQL, Java, JavaScript, HTML, Latex, Linux, Tableau
- **Machine Learning**   Tensorflow, Pytorch, GPyTorch, AWS, Anaconda, Jupyter, Google Colab, Pyro, Stan
- **Data Science**   Statistics and analysis, Database, Data visualization, Critical thinking, Communication