YEHU CHEN

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

SUMMARY

Software Engineer in AIML at Google with experience building, deploying, and operating large-scale machine learning systems in production. Currently working on audience modeling for YouTube Ads, owning end-to-end ML pipelines from modeling and experimentation to serving and monitoring. Strong background in ML systems, user modeling, and A/B experimentation, with PhD training in machine learning and causal inference.

EXPERIENCE

- **Software Engineer, AIML** YouTube Ads, Google Nov 2025 – Present

– Own the design and implementation of production lookalike audience modeling pipelines powering advertiser targeting on YouTube Ads. Also Build and maintain data quality dashboards to detect data drift and integrity issues in contextual taxonomy models at large-scale traffic.
– Drive successful model launches and iterations through offline evaluation and online ramp-up experimentation.
– Author design docs, lead code reviews, and influence engineering best practices across the team.

- **Machine Learning Engineer** Supply Chain Outbound Science, Chewy Aug 2025 – Nov 2025

– Designed and productionized ML models for routing optimization and demand charge forecasting, incorporating causal attribution models to isolate the impact of operational decisions in a large-scale fulfillment network.
– Implemented agent-driven orchestration on AWS, leveraging Jenkins CI/CD to automate model deployment, scenario simulation, and decision logic across downstream operational systems.

- **AI Research Assistant** Division of Comp. Data Science, WashU Sept 2019 – Aug 2025

– Led survey design and developed PyTorch-based models for experience-sampled personality traits, producing findings that resolved long-standing debates in behavioral research and contributed to a $500k NSF grant.
– Collaborated with *CNN* to build an early, high-accuracy polling-based election forecasting pipeline, achieving ˜95% accuracy in the 2020 cycle and outperforming FiveThirtyEight and The Economist.
– Design predictive and computational models in Python/R/SQL and conducted causal inference analyses with diff-in-diff method. Built end-to-end data workflows for preprocessing, feature engineering, and large-scale survey analytics. Communicated insights through dashboards, peer-reviewed publications, and stakeholder reports.
– Built research prototypes and production-quality tooling at the intersection of emerging AI technologies, legal frameworks, ethical accountability, and digital privacy.

EDUCATION

- **Ph.D. in Data Science**, Washington University in St. Louis, GPA: 3.9/4.0 2019 – 2025
  Focus: Machine Learning Systems, Experimentation, Causal Inference
- **B.S. in Computer Science**, University of Michigan, Summa Cum Laude 2017 – 2019

TECHNICAL SKILLS

- **Programming** C/C++/C#, Python, R, Matlab, SQL, Java, JavaScript, HTML, Latex
- **Machine Learning** MLOps, Deep Learning, LLM, AWS, Google Colab, Jupyter, TensorFlow, PyTorch, Pyro
- **Software Engineer** Linux, Git, CI/CO, Automated Testing, Data Pipelines, Docker/Kubernetes
- **Soft Skills** Communication, Collaboration, Experimental Design

SELECTED PUBLICATION

- Idiographic Personality Gaussian Process for Psychological Assessment **NeurIPS 2024**
- A Multi-Task Gaussian Process Model for Time-Varying Treatment Effects **AISTATS 2023**
- Polls, context and time: a dynamic Bayesian forecasting model for US senate elections **Political Analysis**
- Compressive Big Data Analytics: An Ensemble Algorithm for High-dimensional Datasets **Plos One**