XIAOTONG (ERIN) TAN

Phone: +86 159-6662-1692 | Email: xt254@cornell.edu | Job Interest: Data Science

EDUCATION

Cornell University Ithaca, USA
Master of Engineering in Operations Research and Information Engineering Aug. 2024 – Dec. 2025
- GPA: 4.22/4.0
- Related Courses: Machine Learning, Deep Learning, Modeling and Computation, Big Data Technologies, Data-Driven Marketing, Consumer Data Analytics & Strategy

New York University Shanghai Shanghai, CHN
Bachelor of Arts in Economics, double major in Data Science Sept. 2020 – May 2024
- GPA: 3.95/4.0; Honors: Dean's Award in Arts and Sciences, Summa Cum Laude Honor, Recognition Award

WORK EXPERIENCE

The Home Depot, Advanced Analytics Team Remote, USA
Data Science Student Consultant Jan. 2025 – Apr. 2025
- Developed and deployed **Prophet** and **Silverkite** time series forecasting models to conduct **counterfactual analyses** of key ecommerce metrics—including add-to-cart, checkout, and full-funnel conversion rates. Leveraged these predictive insights to quantify performance gains from the Apple Pay launch, informing the payment module rollout.
- Applied **K-means** algorithms to segment promotions into 10 distinct impact tiers and developed weighted promotional indicators to adjust for peak underestimations (e.g., Black Friday). These enhancements reduced forecasting error from 12% to 4.8% and boosted conversion rate predictions by 2 percentage points.

Gucci, Product Care Department Shanghai, CHN
Data Analyst Intern Feb. 2024 – Jun. 2024
- Engineered a **SQL/Python ETL pipeline** to analyze 10,000+ complaint emails using NLP techniques and summarized shopping–complaint–resolution data network. Leveraged **XGBoost with SHAP** to pinpoint recurring complaint drivers (e.g., holiday shutdowns), driving a 5 pp improvement in first-contact resolution year-over-year.
- Developed an **integer programming** model to optimize the repair sequence for 100 products, boosting profits by 5%.
- Established five key service performance metrics by benchmarking top-performing staff. Implemented an **automated Python monitoring system** with a real-time Power BI dashboard that boosted management efficiency by 30%.

Kuaishou Technology, Search Analysis Center Beijing, CHN
Data Analyst Intern May 2023 – Aug. 2023
- Mapped the conversion funnel and segmented users for search incentive tasks, uncovering low conversion rates caused by exploitative behavior. Developed a **logistic regression** model to predict conversion likelihood and implemented targeted exit strategies, achieving a projected 1.8× improvement in conversion efficiency.
- Quantified surges in search volumes under trending events by analyzing billions of records in SQL. Identified three behavioral patterns through **path and cohort analysis**, informing the launch of trending event reminder feature.
- Analyzed **A/B tests** and behavioral heterogeneity using the **Causal Tree algorithm**. Refined user profiles with the TGI index and recommended targeted strategies. Contributed to the optimization and launch of 3 new features.

PROJECT EXPERIENCE

Impact of AI Adoption on Industry Concentration in AI-using Industries Sept. 2023 – May 2024
- Integrated data from Compustat and LinkedIn to build an enterprise-level panel dataset of over 200,000 companies.
- Implemented long difference, **difference-in-difference**, and **event study** models in R to analyze the causal impact of AI adoption on industry concentration, and presented the significant and robust results in an academic report.

ML-Based Causal Inference for Effect of Education on Emergency Room Visits Sept. 2023 – Dec. 2023
- Employed Python EconML **double machine learning** model to analyze Medical Expenditure Panel Survey data, estimating the average treatment effect of education on emergency room visits and uncovering its significant impact.

ADDITIONAL INFORMATION

- **Technical Skills**: Statistical inference, A/B testing design and analysis, causal inference (including PSM, uplift modeling, double machine learning, etc.), data visualization, machine learning, and deep learning.
- **Software Skills**: Python, R, SQL, Hive, Spark, SPSS, Stata, Tableau, Power BI, AMPL, Excel
- **Language Proficiency**: TOEFL: 111, GRE Verbal: 162; proficient in English for professional use