Xintong Shi

linkedin.com/in/xintong-shi-277763114 | github.com/ericshi123 | xintongshi.dev

EDUCATION

University of Wisconsin-Madison | Madison, WI | 2020 – 2021 **M.S. in Data Science**
Carnegie Mellon University | Pittsburgh, PA | 2016 – 2020 **B.S. in Statistics and Machine Learning; Minor in Computer Science**

TECHNICAL SKILLS

- **LANGUAGES:** Python (Pandas, NumPy), SQL, Spark (PySpark), R, GraphQL, C.
- **MACHINE LEARNING:** Recommendations, Fraud Detection, Deep Neural Networks (PyTorch, TensorFlow), Graph Neural Networks (GraphFrame), XGBoost, LightGBM, LLMs, Unsupervised Learning.
- **AI & VIBE CODING:** Claude Code, Metamate (Internal), DevMate, Agent-Assisted Development, Prompt Tuning, LangChain.
- **TOOLS & PLATFORMS:** Google Cloud Platform (BigQuery, Vertex AI), Azure Databricks, MLflow, Airflow, Looker, Neo4j, Spanner DB, Apache Hive.
- **SOFT SKILLS:** Cross-functional Leadership, Data-backed Strategy, Team Player, Results-Driven.

EXPERIENCE

Meta | **Data Scientist** | 2026 – Present **Threads Trust & Safety**

- **Metric Strategy:** Spearheaded the ownership and strategic roadmap of topline performance metrics for Account Graph and Discovery surfaces within the Threads ecosystem.
- **Optimization Framework:** Evaluated and audited quality metrics for Suggested Users, People You May Know (PYMK), New User Experience (NUX), and Chaining surfaces across precision, recall, goalability, interpretability, sensitivity, and coverage.
- **Product Impact:** Engineered and shipped a refined metric architecture that significantly improved precision and goalability by reducing data noise, directly enhancing long-term surface health tracking.
- **AI-Native Development:** Leveraged **Metamate** and **Claude Code** to achieve high-velocity iteration, utilizing 'vibe coding' workflows to automate routine data engineering tasks and codebase navigation.
- **Cross-Functional Leadership:** Partnered with Product, Engineering, and Design leads to provide data-backed directions for feature prioritization and graph surface improvements.

Walmart (Sam's Club) | **Data Scientist / Senior Data Scientist** | 2022 – 2026 **GenAI & Fraud Detection Systems**

- **GenAI Innovation:** Developed an end-to-end **GenAI Agent** using Google Agent Development Kit and **Vertex AI** to automate blocklist explanations, integrating **Natural-Language-to-SQL** for real-time data insights.
- **Network Fraud Detection:** Engineered a scalable, unsupervised graph-based algorithm using **GraphFrame** to identify fraudulent return networks, analyzing data from **69M+ members**.
- **Financial Impact:** Reduced Year-over-Year (YoY) return fraud costs by over **$20 million**, receiving the Sam's Club Make A Difference Award.

- **Model Optimization (EDEL):** Improved Email-Delivery Fraud Detection (V4) by implementing a **LightGBM** model with custom fraud window and location features; boosted precision by **7%** and recall by **10%** via rigorous A/B testing.
- **Real-time Systems:** Co-designed a real-time **XGBoost** model for Sam's Cash POS fraud detection, reducing chargeback rates by **25%** (to 0.69 bps) while maintaining sub-50ms latency.
- **Pipeline Orchestration:** Built automated model training and deployment pipelines using **MLflow** and **Airflow** to ensure seamless production integration.

WALMART (Sam's Club) | **Data Scientist Intern** | Summer 2021 **Search & Personalization**

- **Recommendation Systems:** Developed a substitute item recommendation module for checkout pages using Latent Dirichlet Allocation (LDA) and the Sceptre Model.
- **Feature Engineering:** Extracted user-item interaction features and propensity scores from raw session data to drive personalized recommendations.

RESEARCH & PUBLICATION

Shi, X., Cao, W. & Raschka, S. "Deep neural networks for rank-consistent ordinal regression based on conditional probabilities." Pattern Anal Applic 26, 941–955 (2023).

- **Ordinal Regression Research:** Designed a novel rank-consistent method for ordinal regression in neural networks.

Code available on GitHub: https://github.com/ericshi123/ordinal-conditional-network