

Sara Ahmed

Data Scientist

Contact

- ✉ sara.ahmed@email.com
- ☎ +1 (415) 555-0172
- 📍 San Francisco, CA
- 🌐 linkedin.com/in/sara-ahmed-ds
- 🐙 github.com/sara-ahmed

Technical Skills

Languages

Python, R, SQL

ML & AI

scikit-learn, XGBoost, PyTorch

Statistics

A/B testing, Bayesian inference, time-series

Visualization

Tableau, Power BI, matplotlib

Big Data

Spark, Hadoop, Airflow

Education

MSc Statistics & Data Science

Carnegie Mellon University
2019

BSc Mathematics

University of Michigan
2017

Profile

Senior Data Scientist with 6+ years delivering end-to-end ML solutions across e-commerce, fintech, and SaaS. Expertise in predictive modelling, causal inference, and large-scale data pipelines. Passionate about translating complex analytical findings into clear business value.

Experience

Senior Data Scientist

Feb 2022 – Present

Acme Commerce Inc. — San Francisco, CA

- Built a real-time recommendation engine with XGBoost; lifted conversion rate by **18%** and incremental revenue by \$4.2M annually.
- Designed causal A/B testing framework (bootstrapped CI, CUPED variance reduction) cutting experiment duration by 30%, enabling 2× faster shipping cadence.
- Deployed churn prediction model (AUC 0.91) integrated with Salesforce; reduced monthly churn by **22%** saving \$1.8M in ARR.

Data Scientist

Jul 2019 – Jan 2022

Finova Analytics — Chicago, IL

- Developed credit-risk scoring pipeline (gradient boosting + SHAP explainability) achieving AUC of 0.88, cutting manual review costs by **35%**.
- Automated ETL workflows with Apache Airflow & Spark on AWS EMR, reducing data-refresh latency from 6 h to **45 min**.
- Applied Bayesian hierarchical models to forecast regional loan demand; improved forecast accuracy by **12%** vs. prior ARIMA baseline.

Projects

Customer Lifetime Value Engine

PyTorch, Lifetimes, SQL, Tableau

- BG/NBD + Gamma-Gamma CLV model retrained weekly; drove **24%** uplift in targeted-campaign ROI across 2M+ customers.

Real-Time Anomaly Detection Dashboard

Spark Streaming, sklearn, Power BI

- Isolation-Forest pipeline processing 50K events/sec; reduced fraud-detection latency from 4 h to **under 2 min** in production.

Achievements

- Kaggle Expert** — Top 3% finish in Tabular Playground Series (May 2023); gold medal solution open-sourced with 400+ GitHub stars.
- Co-authored “*Scalable Causal Inference for Online Platforms*” published in *NeurIPS 2023 Workshop on CausalML*; 180+ citations.