

TARU TAK

New York, NY · taktaru@gmail.com · linkedin.com/in/taru-tak · github.com/tarutak

PROFESSIONAL SUMMARY

Senior Data Scientist with **6+ years** building and deploying applied ML systems for targeting, personalization, and decisioning. Strong background in **predictive modeling and applied uplift modeling**, with hands-on work in **agentic AI, deep learning methods**, and end-to-end production ML. Delivered **\$19M–\$48M in annual incremental revenue** at Nike and built Virtuous's ML foundation serving **1,500+ nonprofits**.

EXPERIENCE

Virtuous | Senior Data Scientist (Remote)

Sep 2023 – Present

- As Virtuous's **founding data science hire**, built **Virtuous Insights / Donor360**, unifying first-party CRM data with third-party and public data into a donor-intelligence platform powering targeting, prioritization, analytics, and fundraising workflows across **1,500+ nonprofits**.
- Designed and productionized **4 predictive ML models** for donor reactivation, acquisition, retention, and gift-ask optimization, owning feature engineering, automated retraining, validation, batch inference, and CRM delivery; achieved **5.8x, 2x, and 1.7x lift vs. random targeting** and **1.9x over the incumbent gift-ask heuristic** in out-of-time validation.
- Built Donor360's **identity-resolution system**, matching fragmented CRM donor records against **260M US consumer records** and stitching in public-record, real-estate, demographic, and wealth data to create enriched profiles for **100M+ donor households**.
- Shipped the customer-facing **Insights Research Agent** from hackathon to production, building an **LLM-powered, multi-step agentic workflow** that combines unified donor profiles with live web research to generate source-cited fundraiser briefings in **1–2 minutes**, with LLM evaluations for response quality and grounding.
- Cut **data and ML debugging from ~2 days to 2–4 hours** using an agentic engineering workflow with **Claude Code** and tool access across Snowflake, Azure SQL, dbt/GitHub, and Jira to trace issues across source data, CRM sync, transformations, and modeling logic.

Nike (via Launchpad Inc. & TEKsystems) | Data Scientist, Consumer Data Science

Mar 2020 – Sep 2023

- Generated **\$19M–\$48M in annual incremental revenue** across **200+ campaigns** by replacing propensity-based targeting with **causal uplift models (EconML/CausalML)**, using holdouts to isolate customers whose response was incremental to treatment.
- Expanded the **SNKRS sell-through model** to **3 additional retail geographies**, incorporating product-image and user-product taste embeddings and improving sell-through by **5–8%**.
- Built **campaign experimentation, audience, and measurement infrastructure** in Databricks/PySpark, expanding reusable training data from **300K to 8M+ records** and reducing audience sizing, A/B testing, and post-campaign reporting from **~3 days to ~6 hours**.
- Developed a **Nike By You trend-discovery system** using CLIP embeddings, UMAP, and HDBSCAN to surface emerging custom-design patterns and assortment gaps, reducing merchandiser analysis time by **~80%**.
- Built a **90-day launch-demand forecasting pipeline** with **10% lower forecast error** for regional inventory and investment decisions.
- Presented causal-inference methodology and results to **60–80 senior leaders and non-technical stakeholders** at Nike Data Science Labs.

Fellowship.AI (Launchpad Inc.) | Machine Learning Fellow

Sep 2019 – Dec 2019

- Led **30 ML fellows** developing a mobile computer-vision system with PyTorch and FastAI, improving model accuracy from **82% to 90%**.

TrueFort Inc. | Machine Learning Intern

Feb 2019 – May 2019

- Built behavioral threat detection with **FP-growth pattern mining on Apache Spark**; automated client reporting saved **~17 hours/week**.

EDUCATION

M.S. Information Systems, Stevens Institute of Technology, Hoboken, NJ · **GPA 3.85/4.0**

2019

B.E. Computer Engineering, Savitribai Phule Pune University, India · **First Class with Distinction**

2017

CORE SKILLS

Applied Machine Learning: Predictive Modeling · Feature Engineering · XGBoost / Gradient-Boosted Trees · Causal Inference & Uplift Modeling · Experimentation / A/B Testing · Recommender Systems · Forecasting · Embeddings & Clustering · Identity Resolution · Model Explainability (SHAP)

ML, Data & AI Systems: Python · SQL · PyTorch · Databricks · PySpark · Snowflake ML · Snowpark Python · dbt · MLflow · Snowflake Model Registry · Model Evaluation & Monitoring · Batch Inference · CI/CD · LLM Agents · Tool Calling · LLM Evaluation · Docker