

IAN FINN

Senior Data Scientist | Agentic AI & LLM Systems | PhD Economist

Los Angeles, CA (open to remote) • (518) 526-9929 • ian.patrick.finn@gmail.com • linkedin.com/in/ianpatrickfinn • github.com/dr-ian-finn

SUMMARY

Senior data scientist with a Ph.D. in economics and five years shipping production ML and LLM systems at ServiceTitan (NASDAQ: TTAN): churn and customer-health models scoring ~14K accounts weekly, a 10+-pipeline GenAI platform with measured cost and hallucination control, agentic evaluation and data-migration systems, MILP portfolio optimizers, and a 14-study causal-inference portfolio. I build the model, the pipeline it runs in, the evaluation that proves it, and the documentation that survives me.

EXPERIENCE

ServiceTitan • Los Angeles, CA (Remote)

May 2021 – Jun 2026

Field-service management SaaS; IPO Dec 2024 (NASDAQ: TTAN)

Senior Data Scientist

Aug 2022 – Jun 2026

- Architected a config-driven GenAI platform on Azure ML powering **10+ daily production pipelines** across the customer lifecycle (onboarding, go-live, success transition, escalation, critical issue, churn), turning call transcripts, support cases, and product-usage data into per-account intelligence briefs in Snowflake, Google Docs, and Salesforce — new pipelines ship as single YAML configs, cutting delivery from weeks to days with zero code changes
- Cut LLM cost per account up to **96% (\$2.79 → \$0.11)** via prompt caching and two-stage prompt consolidation; raised structured-output reliability from **~95% to 99.9%** with an LLM-based repair loop; built production hallucination detection, remediation, and prevention gates
- Built the churn-driver analysis behind ServiceTitan's flagship onboarding-overhaul program: statistically significant usage thresholds (e.g., 11% vs. 33% first-year churn by early activity) became the program's success metrics, and its AI account briefs ran on the platform above — the 29-account pilot cut onboarding churn from a **16–20% baseline to 2.5%** before GA rollout
- Built an agentic LLM-as-judge evaluation of the customer-facing support chatbot: mined **19,082 canonical questions from 51,000 customer calls** and graded 5,000 answers with tool-grounded verification against the full help-center corpus and product source code — establishing an **82.0% correct-answer baseline and 10.0% hallucination rate** at ~\$128 per fully automated run
- Built the **fully agentic** tier of the team's data-migration app: autonomous Claude agents explore any competitor system's export, plan a transformation strategy, then **write and execute their own** validated Python transformations in a sandbox — with per-job cost caps and automatic model escalation on failure; owned Okta SSO auth, security hardening, and Azure deployment
- Conceived (2023 personal prototype), productionized, and twice re-architected CoachGPT, an LLM pipeline QA-scoring every support call and chat daily — replacing minimal manual sampling with **100% coverage at 75% supervisor adoption** and **~\$82K/month** in avoided manual-QA cost
- Owned the production churn/health-score system (weekly LightGBM scoring of **~13,800 accounts** from 81 hand-written SQL feature modules, delivered to Tableau and Gainsight): found and fixed a silent data-leakage bug in retraining, added a 7-check validation suite, and rebuilt monitoring from a fragile manual process into <2-minute automated reporting
- Shipped the Ops Data Science team's **first real-time recommendation endpoint** (Azure ML Kubernetes) plus its pricebook sibling — zero-regression deploys, per-call audit logging, startup smoke tests — with a three-tier evaluation reaching 100% precision and recall on half of audited accounts, failures root-caused to a single industry-coverage gap
- Balanced Customer Success and specialist books of business with MILP optimization (PuLP) and custom redistribution heuristics: **\$26.6M MRR** across 4,000+ accounts distributed at **0.00% deviation** on multiple rebalances; same-day absorption of 3 simultaneous CSM departures (222 accounts, \$774K MRR)
- Ran a **14-study causal-inference portfolio** (difference-in-differences, panel fixed effects, propensity-score matching, mixed effects, power analysis) evaluating training certification, chatbot rollout, pricing changes, and onboarding programs for Success, Support, and Onboarding leadership

Data Scientist

May 2021 – Aug 2022

- Deployed weekly Prophet forecasts of support contact volume (phone and chat; email added 2024) used directly in workforce-management staffing decisions — email model achieved **8% MAPE** at a 3-month horizon, 5 points better than the incumbent process (~350 fewer forecast errors/month)
- Built fairness-adjusted performance benchmarks for support agents — predicting expected escalation rate and handle time from case mix (logistic regression, LASSO) with per-agent SHAP coaching drivers — delivered to Support leadership
- Productionized LDA topic-detection pipelines (self-built Airflow DAGs + Docker) across 8+ text domains, joining topics to resolution-time KPIs in Tableau
- Trained early NLP/transformer models on proprietary support corpora — BERT article recommender, custom GPT-2 tokenizer, and a fine-tuned GPT Q&A prototype built before the ChatGPT API existed

Economist

- Valued intangible property in inter-company transfer-pricing transactions among Fortune 500 taxpayers using local-regression and econometric models in Python and Stata; proposed a precedent-setting tax adjustment
- Researched industries, functions, and supply chains to evaluate potential tax avoidance, forecast profit potential, and direct the course of taxpayer examinations
- Explained and defended quantitative findings to taxpayers' representatives, team members, and managers of varying statistical backgrounds

Welch Consulting • Los Angeles, CA

Aug 2016 – Aug 2018

Economist

- Designed and trained a surname-based ethnicity-inference algorithm that reduced test error 7%; deployed it to help defeat class certification in an employment suit seeking hundreds of millions in damages
- Processed, cleaned, and analyzed employment and demographic data spanning millions of observations across disparate client databases (SQL, Stata)
- Managed a team of researchers and generated **\$1M+ in annual billed revenue** through client development and presentations at legal conferences

EDUCATION

Ph.D. & M.A., Economics — University of California, Irvine

2016

*Herbert F. York Global Security Dissertation Fellow, UC Institute on Global Conflict and Cooperation***B.A., Economics (minor: International Studies) — Boston College**

2008

SKILLS

Languages & core: Python (pandas, scikit-learn, LightGBM, SHAP, statsmodels, Prophet, PuLP) • SQL (Snowflake, Redshift) • Stata • R**LLM & agentic AI:** Anthropic Claude API (prompt caching, extended thinking, tool use, batch) • Azure OpenAI • agentic pipelines & Claude Agent SDK • LLM-as-judge evaluation & hallucination measurement • prompt engineering • cost governance • MCP**ML & statistics:** Churn/propensity modeling • time-series forecasting • causal inference (DiD, fixed effects, PSM, mixed effects) • experiment design & power analysis • MILP optimization • NLP (LDA, BERT, embeddings)**Platform & MLOps:** Azure ML (Kubernetes endpoints, scheduled pipelines) • Airflow • Docker • Snowflake • Flask • TeamCity CI/CD • Okta OIDC • Tableau • Salesforce/Gainsight integration