

Mishika Ahuja

+1 (949) 283-4554 | Irvine, CA, USA | ahujamp@uci.edu

[GitHub](#) | [LinkedIn](#) | [Portfolio](#)

EDUCATION

University of California, Irvine

M.S. Data Science

Irvine, CA

Sep 2025 – Dec 2026

Vellore Institute of Technology

B.Tech Computer Science Engineering, CGPA: 8.72/10

Vellore, India

Sep 2021 – May 2025

EXPERIENCE

Quantiphi | Data Science / Machine Learning Engineer Intern

Jan 2025 – Jun 2025

- Built the RAG pipeline for an AI support-assist tool on GCP within a 6-person team, retrieving the top-5 similar tickets and prompting Gemini 1.5 Pro to recommend next actions at sub-2-second latency.
- Engineered data pipelines across 4 sources, cleaning duplicates, HTML, and noise to raise retrieval context precision to 85% and context recall to 84%.
- Deployed a Vertex AI vector search index over historical tickets, FAQs, and docs, generating embeddings with text-embedding-005 and tuning 1500-token chunking.
- Established an evaluation harness (DeepEval, BLEU, F1), confirming 95% answer relevance and 92% faithfulness before every release.

SKILLS

Languages & Analytics: Python, SQL (window functions, CTEs, cohort analysis), R, A/B Testing, Experimental Design, Hypothesis Testing, Bootstrap, Regression, RFM Segmentation, Forecasting

ML & AI: Machine Learning, Deep Learning, LLMs, RAG Systems, Recommender Systems, NLP, Transformers, Anomaly Detection, Explainable AI (SHAP), Feature Engineering, Model Deployment

Data & Cloud: PySpark, Apache Spark, Kafka, Snowflake, Databricks, PostgreSQL, GCP (Vertex AI, BigQuery), AWS, Docker, FastAPI, Git

Visualization & BI: Tableau, Power BI, Chart.js, Streamlit, Dashboard Design, Matplotlib, Pandas, NumPy, scikit-learn

Certifications: AWS Solutions Architect · GCP Professional ML Engineer · ML Engineer (Quantiphi)

PROJECTS

RouteLLM: Cost-Aware LLM Router — [\[Code\]](#)

XGBoost, FastAPI, OpenAI

- Spearheaded a two-tier router predicting cheap-model success from query embeddings, cutting inference cost 69% for a 2.5-point accuracy tradeoff.
- Charted a cost and quality frontier across 16 configurations, validating each against a random-routing baseline.
- Showed why cascade verification fails: recall rose 0.14 to 0.72 moving from multiple-choice to free-form math.

FinSight: Multi-Agent Financial RAG with Observability — [\[Code\]](#)

LangChain, FAISS, Groq

- Designed a 4-agent pipeline (Planner, Retriever, Critic, Synthesizer) over 10-K filings for Apple, Amazon, and JPMorgan.
- Initiated a deterministic verification layer that re-parses each cited table and auto-corrects wrong-year figures.
- Authored AgentLens, a self-scoped observability layer that surfaces per-agent cost and latency, diagnosing 90% of latency as silent SDK stalls.

Cookie Cats A/B Test: Retention Experiment — [\[Code\]](#)

z-test, Bootstrap, SQL

- Analyzed a 90,189-player mobile-game experiment, delivering a NO-GO decision after a gate change dropped 7-day retention 4.3% (two-proportion z-test, $p=0.0016$).
- Caught a sample-ratio mismatch ($p=0.009$) and proved via balanced resampling that the result held; confirmed significance with a 10,000-sample bootstrap.
- Segmented retention by engagement tier in SQL, showing the drop hit the highest-value players hardest.

Olist E-Commerce Analytics: 100k Orders — [\[Code\]](#)

SQL, RFM, Chart.js

- Turned 100k orders across 8 tables into a prioritized recommendation, identifying a 3.1% repeat-purchase rate as the growth ceiling via cohort-retention SQL.
- Quantified that late delivery drops review scores 1.72 stars (4.29 to 2.57), isolating the clearest driver of dissatisfaction.
- Communicated R\$7.9M of at-risk revenue to stakeholders via RFM segmentation (NTILE window functions) and an interactive dashboard.

FleetPulse: Streaming Anomaly Detection — [\[Code\]](#)

scikit-learn, Kafka, Spark

- Constructed a streaming pipeline with per-vehicle windowed anomaly detection, measured at 0.81 recall against labeled faults, mapping the logic onto Kafka and Spark patterns.
- Swept the alert threshold to chart the precision-recall tradeoff, framing the operating point as a missed-fault versus wasted-inspection business decision.
- Presented anomalies as a ranked, cost-weighted inspection dashboard built for direct action.