Raunak Sood

Los Angeles, CA • raunak.sood@gmail.com • (650) 996-2932 • linkedin.com/in/raunak-sood • github.com/quantraunak

Experience

INNOVIUS CAPITAL — San Francisco, California
Machine Learning Engineer, Investment Platform (part-time during school) — *Aug 2026 - Present*
Machine Learning Research Intern — *May 2026 - Aug 2026*
- Rebuilt the firm's startup-ranking model (CatBoost + GPT-4o rubric): 90% of its top-20 picks advanced, vs 20% before
- Built an LLM-as-judge grader that matched real investor decisions at human-level agreement and gated every model release
- Took the model to production: a weekly GitHub Actions pipeline scoring 14,000 companies, with drift checks and alerts
- Built an AI search agent over 268,000 companies (embeddings + LLM + web research): recall 49% to 93% in backtests
- Built a Claude tool-use agent that answers deal questions from CRM data and call transcripts with code-verified citations

QIAGEN — Redwood City, California
Machine Learning Intern — *Jun 2025 - Aug 2025*
- Built a GraphRAG tool on QIAGEN's knowledge graph that ranks drug-target paths and has an LLM write cited explanations
- Fine-tuned and deployed an LLM tool for gene-variant lookups, cutting research time by 70%

QIAGEN — Dusseldorf, Germany
Data Science & Global Treasury Intern — *Jul 2024 - Sep 2024*
- Built a LangChain RAG chatbot for financial dashboards, cutting internal query time by 50%
- Analyzed currency risk on $10M+ of transactions and executed FX trades on Bloomberg

YORK UNIVERSITY — Toronto, Canada
Data Science Intern — *Sep 2023 - Dec 2023*
- Used topic modeling and regressions on earnings calls to study how investment firms discuss and adopt fintech

PRECANTO — Los Gatos, California
Data Science and Product Intern — *Jul 2023 - Dec 2023*
- Forecast hiring and attrition with time-series models; automated payroll-tax validation, raising its efficiency 90%

GREENFIELD RESEARCH PARTNERS — Palo Alto, California
Private Equity Research Associate — *Feb 2023 - May 2023*
- Sourced 100+ real estate acquisition targets (5 potential deals worth $50M+) and supported LBO modeling

Education

UNIVERSITY OF SOUTHERN CALIFORNIA — Los Angeles, California
Master of Science in Computer Science — *Expected May 2027*
- **GPA:** 3.85 | **Coursework:** Analysis of Algorithms, Applied NLP, Database Systems, Deep Learning, Operating Systems

SANTA CLARA UNIVERSITY — Santa Clara, California
Bachelor of Science in Economics — *Jun 2024*
- **GPA:** 3.9 (Upper-Division), Dean's List | **Coursework:** Machine Learning, Time Series, Algorithms, Data Structures

Projects and Research

LEAKPROBE — Open source on PyPI: pip install leakprobe
Sole author, Python library — *Sep 2026 - Present*
- Flags ML features that use future data: caught 8 of 9 planted leaks in 5 datasets, zero false alarms on 5 clean pipelines
- Runs as a test in CI, failing the build when a feature reads data it was never supposed to see

BIAS FINGERPRINTS AND CARDINALITY-EVAL — Sole-author papers
Financial data quality; LLM evaluation — *2025 - 2026*
- Found that period-end dating of financial data inflates the average signal of 11 filing-based stock factors by 59%
- Showed LLM extraction output stops being valid JSON as lists grow (21/25 at 16 items) unless a schema grammar is used

DEAL DESK — github.com/quantraunak/deal-desk-voice
Voice agent with an evaluation harness — *Oct 2026*
- Voice agent answering deal questions from notes with citations: 87/90 on a 30-case eval with a rule-based grader
- Audited the platform's LLM judge against hand labels: it was wrong in 14 of its 17 disagreements with the grader

MELANGE — Live on the iOS App Store and web
Sole developer: Next.js, React Native (Expo), Supabase Postgres — *Feb 2026 - Present*
- Built and shipped a full-stack app where creatives post projects, match and chat in real time, on iOS and the web

Skills

Languages: Python, TypeScript, SQL, C++, Java, R
AI/ML: LLM agents and tool use, RAG, structured outputs, LLM-as-judge evals, fine-tuning, PyTorch, CatBoost, embeddings
Systems: PostgreSQL, pgvector, AWS (S3, Lambda, App Runner, ECR), Docker, GitHub Actions, Next.js, React Native
Athletics: Ranked top 6 in Northern California tennis; NCAA Division I Men's Tennis at Santa Clara University