

Chinmay Sonar

(805) 895-4475 | chinmaysonar96@gmail.com | Google Scholar | LinkedIn

Staff Machine Learning Scientist specializing in transformer-based representation learning, agentic systems, and graph ML with experience developing and evaluating large-scale models for production financial applications.

EXPERIENCE

VISA | Staff ML Scientist

(Jan 2024 - present, Austin, TX - Hybrid)

Transaction Sequence Modeling for User Behavior Understanding

- Led a team of three to develop an encoder-based transformer for transaction sequence modeling, delivering a 25% uplift in fraud dollars captured and generating account-holder embeddings used across multiple downstream fraud applications.
- Outperformed the production XGBoost baseline by \$2M+ in monthly fraud capture; pretrained account-holder embeddings also achieved 60% detection on first-party fraud, a 15-point improvement over the prior rule-based system.
- Trained a 10M-parameter model on 200M+ transaction sequences using distributed GPU training and one patent filing covering the sequential behavior architecture and self-supervised pretraining approach.

Graph Representation Learning for Identity and Fraud-Network Detection

- Led the design of unified visa identity based on graph linking 1B+ cardholder accounts across 5.4B+ transactions, improving first-party fraud models by 8–15% across return abuse, serial disputing, and transaction-by-mistake use cases.
- Developed proximity embeddings that quantified each account's structural relationship to known fraud networks, surfacing risk patterns that were not visible from account-level features alone.
- Independently built a petabyte-scale evaluation framework using PySpark, quantifying graph coverage, representation quality, and downstream model impact to drive adoption across multiple fraud teams.

Risk Modeling and Evaluation for Agentic Commerce

- Formulated intent alignment as an emerging agentic-commerce risk problem after observing elevated return rates for agent-initiated transactions. Designed a two-stage transformer that transfers representations learned from non-agentic returns and disputes before fine-tuning on limited agentic-commerce data and agent-specific behavioral features.
- Built a synthetic-data pipeline to bootstrap development before sufficient labeled data was available, generating LLM-based intent-action pairs with controlled misalignment patterns and an evaluation framework for measuring intent-alignment risk.

Agentic System for Automating Data Science Workflows (ML-Agent)

- Designed and built an agentic ML system that converts Jira requests into end-to-end fraud modeling workflows, automating a nine-stage process from problem framing and feature engineering through model evaluation and delivery.
- Added human-in-the-loop review gates that convert expert corrections into DPO and SFT training data, enabling the system to improve from repeated use.
- Implemented a RAG-based compounding knowledge base using SQLite and ChromaDB that carries validated experiments, features, and expert corrections across teams, allowing institutional knowledge to accumulate across projects rather than reset; currently piloted at Visa and deployed using Docker and Kubernetes.

RESEARCH IMPACT

- Research in representation learning, transformer models, agentic AI, fraud detection, and graph algorithms
- 11 peer-reviewed publications | 200+ citations
- Lead author on one patent filings covering transformer architectures and loss-function optimization, plus three defensive publications; resulting methods supported production fraud models and generated incremental fraud capture.
- Reviewer/Program Committee: NeurIPS, AAAI, IJCAI, AAMAS, EMNLP, SODA, SoCG.

EDUCATION

UC SANTA BARBARA, CA | PHD IN COMPUTER SCIENCE | GPA: 4/4 | SEPT 2019 - DEC 2023

IIT GANDHINAGAR, INDIA | MS IN COMPUTER SCIENCE AND BS IN MECHANICAL | GPA: 9.83/10 | JULY 2019

ADDITIONAL PROJECTS

Mule and Scam Detection for Visa Protect A2A

- As the sole data scientist on a team of five, built mule and scam detection models for Visa Protect A2A across East Asian financial institutions processing 2M+ transactions per day; achieved 70% mule detection with a two-month lead over police reporting.
- Extended the initial market-specific work into a single global scam model that generalized across financial institutions in Europe, Asia, and Latin America, outperforming country-specific models across 50M+ monthly transactions.

PREVIOUS EXPERIENCE

VISA | PhD Research Intern

Fine-Tuning Llama for Transactional Data Exploration

(June 2023 - September 2023)

- Fine-tuned Llama with LoRA to answer natural-language queries over transactional datasets, enabling data analysts and product leads to explore data via plain English rather than SQL while keeping all sensitive transaction data on-premise.
- Designed a data anonymization layer ensuring no raw PII left Visa's infrastructure during LLM inference — addressing the key compliance constraint for deploying open-source LLMs on financial data.

Shared-Model Onboarding and Distribution-Shift Detection

(June 2022 - September 2022)

- Proposed and validated a scalable onboarding strategy for Visa Protect's shared fraud model, improving performance for new financial institutions while bounding degradation for existing clients with minimal additional compute.
- Developed distribution-shift detection methods to determine retraining triggers under both organic data drift and shifts introduced by newly onboarded institutions.

CLEMSON UNIVERSITY | Undergraduate Research Intern

(May 2017 - August 2017)

Computational Modeling and Visualization of Hydrogel Resonance

- Developed C/C++ simulation software to model resonance behavior in hydrogel samples and generate outputs for scientific analysis.
- Built interactive visualization interfaces in MATLAB and Mathematica, enabling researchers to explore simulation results and analyze resonance patterns.

SKILLS

Machine Learning: Transformers • RAG • Representation Learning • Agentic AI • Gradient Boosting • Tabular ML • Graph ML

Languages & Frameworks: Python • SQL • PyTorch • XGBoost • scikit-learn • LangGraph

ML Systems: Spark/PySpark • Distributed Training • Large-Scale ML • Model Evaluation & Deployment • Docker • Kubernetes