Kaushik Agastyaraju

Los Angeles, California | kaushik8agastyaraju@gmail.com | 805-638-8721

Summary

Data Science student at UC San Diego (Minor in Business Analytics) with internships at Adobe and Kaiser Permanente spanning production LLM evaluation, machine learning, data engineering, and business analytics. Experienced in turning data into metrics, models, and decisions for engineers and executive stakeholders.

Education

University of California, San Diego — La Jolla, CA

Bachelor of Science in Data Science, Minor in Business Analytics — *Expected March 2027*

- **Relevant Coursework:** Probabilistic Modeling & Machine Learning; Representation Learning; Data Science w/ AI; Data Structures & Algorithms; Scalable Analytics Systems; Data Management; Data Visualization; Advanced Business Analytics

Experience

Adobe — San Jose, CA

Data Science Intern — *June 2026 – September 2026*

- Shipped an end-to-end LLM evaluation pipeline to production for a live agentic AI assistant, shadow-testing candidate models against the production model on real user traffic with zero added latency and cross-provider support.
- Designed the evaluation methodology for agent outputs, pairing objective metrics with an LLM-as-judge layer that accounts for known judging biases and audits its own reliability.
- Created a self-supervised routing-evaluation method that flags AI assistant mis-routes from disagreement between two independent classifiers, with no labeled ground truth required.
- Built a taxonomy of customer-built AI agents and skills using unsupervised clustering, surfacing redundancy and gaps against the platform's built-in agent capabilities.
- Defined and measured power-user behavior across two product surfaces from product usage data, presenting findings in a leadership-facing capstone report.
- Contributed to user frustration and abandonment detection for the AI agent product to improve the assistant's user experience.

Kaiser Permanente — Pasadena, CA

Data Analyst Intern — *June 2025 – September 2025*

- Built scalable Python ETL pipelines integrating 25+ disparate systems into unified, analysis-ready datasets powering enterprise reporting and high-traffic dashboards.
- Built data pipelines and defined new business metrics in Databricks for executive-level decision-making, working cross-functionally with engineers and stakeholders.
- Designed SQL validation frameworks reconciling Oracle and Databricks, catching schema drift, data gaps, and inconsistencies.
- Automated data validation and quality checks in Python, reducing manual effort and improving reliability of downstream workflows.
- Managed multiple concurrent projects in a fast-paced environment, consistently meeting tight deadlines.

Halicioglu Data Science Institute — La Jolla, CA

Outreach & Research Fellow — *December 2024 – June 2025*

- Built AI-enabled robotics systems with Python, OpenCV, and Raspberry Pi, iteratively improving computer vision pipeline accuracy and responsiveness through experimentation and model evaluation.
- Designed and taught LLM curriculum modules on tokenization, data representation, and generative modeling using real-world text data.
- Explained technical concepts to K–12 students and non-technical audiences through hands-on workshops and visual demos.

Leadership

Project RISHI – UC San Diego — La Jolla, CA

Computer Literacy Committee Lead — *October 2025 – Present*

- Lead a team of 10 designing and implementing computer literacy initiatives for a rural community in India.

Eagle Scout – Boy Scout Troop 127 — Agoura Hills, CA

Eagle Scout — *September 2017 – July 2023*

- Led 50+ volunteers on a community service project to earn Eagle Scout rank, and mentored peers.

Technical Skills

Languages & Libraries: Python, SQL, R, PySpark, Pandas, NumPy, scikit-learn, OpenCV, Matplotlib, Plotly

Machine Learning: Classification, Clustering & Unsupervised Learning, Predictive Modeling, Probabilistic Modeling, Representation Learning, NLP, Computer Vision, Model Evaluation

LLMs & AI Systems: LLM Evaluation, LLM-as-a-Judge, Shadow Testing, Agentic AI, Agent Routing & Orchestration, Prompt Engineering, Model Context Protocol (MCP), Langfuse (Tracing & Observability), Multi-Provider LLM APIs

Data Engineering & Cloud: ETL/ELT Pipelines, Data Validation & Quality, Schema Design, Databricks, AWS, DynamoDB, Oracle

Business & Product Analytics: Business Analytics, Product Analytics, Metric Definition, User Behavior Analysis, EDA, Tableau, Power BI, Data Storytelling, Stakeholder Communication