Kavya Sridhar

858-372-3468 | k1sridhar@ucsd.edu | linkedin.com/in/kavyasridhar | github.com/kavyasridhar1501

Summary

Master's student in Data Science with a strong foundation in Machine Learning, NLP, Large Language Models, and Algorithm design. Passionate about leveraging data-driven solutions to drive meaningful impact.

Education

University of California San Diego — La Jolla, CA
Master of Science, Data Science (GPA 3.916/4.0) — *Sep 2024 – Mar 2026*

College of Engineering Guindy, Anna University — Chennai, India
B.E. Computer Science and Engineering (GPA 3.98/4.0) — *Aug 2019 – Apr 2023*

Experience

Software Engineer — Jul 2023 – Aug 2024
Société Générale Global Solution Centre — *Chennai, India*

- Optimized SQL-based CI/CD pipelines in Palantir Foundry by integrating ML models and automating query execution logic, reducing AML transaction query processing time by **20%** and accelerating compliance decision-making by **3 business days** for stakeholders.
- Developed dashboards tracking AML alert classifications (reportable status, resolution time, escalations) with real-time filtering and automated alert triggers, eliminating **100%** of manual analysis time and improving compliance team efficiency by **50%**.
- Streamlined data storage and workflows for **2** AML business units using Scala Spark transformations, Parquet partitioning/compression, S3, and Oozie orchestration, reducing data processing time by **40%**.
- Migrated 7 critical AML datasets from Palantir OSv1 to OSv2 by resolving schema incompatibilities and decimal precision constraints, successfully validating and transferring **50M+** transaction records with zero data loss.
- Monitored and maintained **20+** production ETL pipelines processing daily AML transactions, resolving data quality issues and failures, achieving **98%** on-time delivery for compliance reporting.

Projects

HemsHappen App RAG-Based Medical Chatbot — Sep 2025
Python, RAG, FAISS, LLM Evaluation — *Link*

- Developed a RAG-based medical chatbot using FAISS vector search and Claude Sonnet 4 over curated GI clinical guidelines, evaluated across medical accuracy, safety, and patient communication by LLM-as-judge and human experts.
- Achieved **100%** clinical safety pass rate, outperforming medical-tuned baseline by **42%** percentage points.
- Submitted to DDW 2025.

LegalInsight, Self-RAG Legal Contract Analyzer — Jan 2025
Python, JavaScript, Flask, Self-RAG — *Link*

- Developed a legal contract analysis tool using Self-RAG with EigenScore validation for hallucination detection, implementing Flask backend and FAISS vector store for semantic search across 6,858 LegalBench-RAG queries.
- Achieved **74.65%** semantic consistency in AI-generated contract reviews and **4.26x** faster review time compared to manual analysis across multiple LLM providers.

Publications

Understanding DeepFool Adversarial Attack and Defense with Skater Interpretations — May 2023
2023 International Conference on WiSPNET, IEEE — *Link*

- Improved ResNet50 robustness against adversarial attacks by **43%** using adversarial training techniques and quantified attack impact through Skater visual interpretation analysis.

Technical Skills

Programming Languages: Python, JavaScript

Data Engineering & Databases: SQL, Apache Spark, FAISS

Machine Learning & AI: PyTorch, LLMs, RAG, Adversarial Training, Skater, MLOps

Cloud & DevOps: Git, GitHub, AWS(S3), CI/CD, Palantir Foundry

Web Frameworks: Flask, JavaScript, React, FastAPI,Docker

Relevant Coursework: Machine Learning, Deep Learning, Natural Language Processing, Computer Vision, Data Mining, Statistical Models, Probability and Statistics, Linear Algebra