AJAY SAI MIRYALA

Atlanta, GA | +1 (240) 360-7905

sai.ajaysai@gmail.com - linkedin.com/in/ajay-sai - ajaymiryala.com - github.com/ajay-sai

PROFESSIONAL SUMMARY

Data Scientist & Generative AI/ML Engineer with **9+ years** in Data analytics and Data Science, including **3+ years** building production GenAI systems at Fortune 50 scale | Expertise in **GraphRAG**, **Agentic Design Patterns**, LLM fine-tuning, and MLOps | **5x Google Cloud Certified** | Delivered **$20M+** business impact | Shipped chatbots serving **200+ daily users** | Reduced inference latency by **40%** | Optimized **$10M+** marketing spend (18% ROI) | Mentored **8 engineers** (3 promoted).

WORK EXPERIENCE

Data Scientist | Generative AI/ML Engineer   Jan 2025 - Present

The Home Depot   *Atlanta, GA*

- Spearheaded the **GenAI Center of Excellence**, and served as AI champion, defining architectural standards and agentic workflows adopted by **300+ data scientists** across the USA and Canada.
- Engineered the Home Depot **CoreAI** platform for agentic workflows, reducing time-to-market for new agents by **40%** via standardized evaluation (**DeepEval**), **Grafana** observability, and **MLflow** tracking.
- Deployed a production **GraphRAG** (RAG 2.0) system utilizing **Neo4j** and vector hybrid-search. Serving **200+** internal IPR analysts, achieving **92% accuracy** across **1M+** internal documents.
- Implemented advanced RAG strategies (Adaptive, Corrective, Fusion) with fine-tuned embeddings on **100GB+ multi-modal data** using **Gemini 2.5 Pro** and **Flash**.
- Implemented Reinforcement Learning from Human Feedback (**RLHF**) pipeline for chat-based support agents, achieving **92% user satisfaction** (+38 pts) across **10k+ daily interactions** with sub-second **p95** latency.
- Led fine-tuning of LLMs using **LoRA, QLoRA, and PEFT** methods with Hugging Face Transformers, achieving **28% improvement** in supply chain IPR domain-specific F1 scores while reducing training time by **65% using Unsloth**.
- Engineered a high-throughput inference layer using **vLLM** and custom semantic caching. Reduced **P99 latency by 40% (2s → 1.2s)** and cut API costs by **35%** via intelligent prompt compression and routed model cascading (Gemini Flash Lite → Flash → Pro).
- Developed **image-to-text pipelines** integrating Stable Diffusion and Vision Transformers for product tagging, improving search relevance by **42%** and enabling automated metadata generation for **1M+ SKUs**.
- Built generative AI systems using **transformer architectures** (GPT-4, BERT, Gemini) for text summarization, Q&A bots, and **Text-to-SQL data agents** on multi-TB data sets for **automated data reporting** needs.

Senior Data Scientist - Product Analytics | Decision Analytics   Jun 2023 - Jan 2025

The Home Depot   *Atlanta, GA*

- Developed deep learning models (**ResNet, Vision Transformers**) for image classification and object detection on 50K+ MLS listing images to **predict House renovation score**, achieving **87% accuracy** vs 64% baseline, piloted with marketing team to segment 5M+ properties for campaign targeting, projecting **$20M** in marketing savings.
- Led an offshore **engineering team of 4** to optimize and migrate **30+ ETL** workflows (BigQuery to Google Dataform and dbt), integrating **50TB+ data** (2.5 TB daily ingest) from **15 sources** and delivering **40% faster ETL runtime**, **35% lower compute cost**, 99.9% SLO-backed data quality, and **180 FTE hours/year** freed from manual refresh and troubleshooting.
- Designed and executed **40+ A/B experiments** using Bayesian sequential testing, hypothesis testing frameworks, and multi-armed bandit algorithms for dynamic content optimization, increasing conversion rates by **12%** while reducing time-to-decision by **50%**.
- Delivered **10+ real-time analytics dashboards**, **OLAP** cubes using **BigQuery and Tableau** for critical Tier 1 events (Black Friday, Cyber Monday), enabling A/B testing and data-driven decisions impacting **$100M+** in GMV. Adoption metrics: Avg 12 dashboard views/day per user, 60% of optimization decisions now self-executed.
- Engineered **Marketing Mix Models** using Bayesian hierarchical methods and time series forecasting via Google Meridian to optimize **$10M+ annual media spend** across **12 channels**, improving marketing ROI by **18%** through data-driven attribution and budget reallocation strategies.
- Applied **causal inference techniques** (Difference-in-Differences, propensity score matching, synthetic control) to measure incremental lift of marketing campaigns, informing **$25M+** in strategic budget decisions with **95% confidence** intervals.
- To improve guided product search on Home Depot online website, built a **multi-modal image generation pipeline** combining Google Image Gen-3, Stable Diffusion (generate images), LLMs (Gemini-1.5 Pro, Text-Bison-32k) (generate dynamic product filters) and text embeddings (**Text-embedding-004, Gecko@002**) (image-keyword alignment), yielding 45% higher visual relevance scores than legacy static catalog images.

Senior Data Analyst   Mar 2022 - Jun 2023

The Home Depot   *Atlanta, GA*

- Delivered **customer behavior analytics and executive reporting** for Category Experience and Brand Advocate teams supporting **300+ associates** with ad-hoc analysis and automated dashboards.
- Closely partnered with **Adobe Analytics** data architect (AAPEs) team to design eVar/prop/event mappings, leading tag management implementation and QA to ensure lossless capture of key behavioral signals.
- Implemented analytics tracking for **15+ Tier 1 events** - (Black Friday, Christmas sales, etc) to optimize Click-Through Rate (CTR) and Conversion Rate (CVR) across product pages.

- Acted as **Adobe Analytics and CJA subject matter expert**, training **80+** associates on self-serve analytics and building reusable workspaces and reports that reduced ad-hoc reporting requests by **20%.**
- Directed a cross-functional task force of **7** in the Voice of Associates (VOA) initiative, spearheading the Liftoff platform to streamline **data science onboarding**, reducing onboarding time by **20%** and increasing satisfaction by **10%** through department-specific insights shared by senior leadership.

Data Engineer | Data Analyst
Feb 2020 - March 2022

Harley Davidson Motor Company
Milwaukee, WI

- Analyzed General Merchandise data and EDI transactions to identify trends, **providing recommendations** for product placement and new product launches with **weekly metrics to leadership.**
- Built **optimized ETL pipelines** and data models using **SQL** and **Alteryx** to support BI dashboards in **Tableau, Power BI, and SAP Analytics Cloud**, reducing data processing time by **80%.**
- Drove **Root Cause Analysis** with SMEs for supply chain inefficiencies, achieving **55%** decrease in open purchase orders and **30%** reduction in inventory count mismatches.
- Maintained the master dataset and performed batch inserting/updating of accounts, product information, lead times, BOM, dealer information, and other objects in SAP using FLEX PLM.

Marketing Analyst
May 2019 - Dec 2019

Anahata Art and Design Pvt - Stoned Santa
College Park, MD

- Created and managed **3-week** Google Ads Campaign with **$300** budget, with **110 ad copies** with **6,000 keywords**, achieving **200%** increase in traffic and **113 sales** (**$3,100** revenue) while minimizing CPC to **$0.11**.
- Analyzed multi-source data (web analytics, external) to propose market expansion, product updates, and campaign optimization recommendations for gifting startup client.

Data Scientist
Aug 2019 - Dec 2019

Principal Financial Group
College Park, MD

- Predicted market regime of Russell 1000 companies to evaluate prospective investments, performing data aggregation on multi-TB **high-frequency financial time-series data** using SQL.
- Implemented machine learning algorithms (Logistic Regression, Random Forest, XGBoost) with hyperparameter tuning achieving **78%** accuracy and **5%** improvement in prediction accuracy for investment decisions.

Graduate Assistant
May 2019 - Dec 2019

University of Maryland
College Park, MD

- Led a team of **10** undergraduate assistants to triage student inquiries and maintain records for **4,000 students**.
- Informed academic advisors on undergraduate program improvements using SQL data extraction visualizations to analyze career opportunities and industry trends.

Data Analyst
May 2017 - May 2018

Bridge Solutions
Bangalore, India

- Introduced **Docker** application to deploy **IBM OMS 9.5 and WMS 9.5**, adopted by **75%** of workforce for streamlined deployments.
- Created interactive dashboards in Tableau and Excel, reporting key KPIs for multiple clients; built relational databases using Microsoft SQL Server for root-causing board failure data in an agile environment.
- Developed inventory target models using data mining and analytical modeling, resulting in **$1M** cost reduction.

TECHNICAL SKILLS

Languages & Frameworks: Python (Pandas, NumPy, Scikit-learn, matplotlib, etc), SQL, R, PySpark, PyTorch, TensorFlow, Hugging Face, XGBoost, FastAPI, PyTest, Bash, DBT.

GenAI & Agents: LangChain, LangGraph, AutoGen, CrewAI, Google ADK, GraphRAG, OpenAI/Gemini/Claude, Neo4j, Pinecone, DeepEval, Stable Diffusion, LlamaIndex, DSPy, W&B, claude-mem, Ollama.

LLM Infra & MLOps: vLLM, Ray Serve, LoRA/QLoRA/PEFT,DPO, Quantization, MLflow, Docker, Kubernetes, GitHub Actions, Grafana, Vertex AI, Terraform, Prometheus, CI/CD.

Cloud & System Design: GCP (BigQuery, Dataflow), Airflow, Dataform, Distributed Systems, Microservices, Event-Driven Architecture, Kafka, Redis, AWS(Lambda, Glue, Bedrock, EC2).

Analytics & Science: Tableau, Power BI, A/B Testing, Causal Inference, MMM, Bayesian Methods, NLP, CV, RLHF, Vision Transformers, Time Series, Experiment Design.

Leadership: Agile/Scrum, Technical Leadership, Stakeholder Management, Mentorship, OKRs, Code Review.

CERTIFICATIONS

Google Cloud Platform (GCP):

Professional Data Engineer | Professional Cloud Database Engineer | Generative AI Leader | Associate Data Practitioner

AI/ML Specializations:

AI Agents in LangGraph | AI Agentic Design Patterns with AutoGen | Multi AI Agent Systems with crewAI
Practical Multi AI Agents and Advanced Use Cases with crewAI | Grokking the Generative AI System Design

Additional: Google Ads Professional, Data Scientist (DataCamp), and Data Analyst (Python) - Edureka.

EDUCATION

University of Maryland, College Park
Aug 2018 - Dec 2019

Master of Science in Business Analytics and Statistics (MSBA)
Robert H Smith School of Business, Maryland, USA

SRM University
May 2013 - May 2017

Bachelor of Technology in Computer Science
Kattankulathur, India