AJAY SAI MIRYALA

(240)-360-7905 | sai.ajaysai@gmail.com | https://www.linkedin.com/in/ajay-sai/

SUMMARY

Lead AI Data Scientist with 8+ years of experience designing and deploying robust, production-grade data ecosystems. I specialize in building the full-stack infrastructure that powers modern AI, from architecting scalable ETL and analytics platforms (BigQuery, GCP) to operationalizing the complete lifecycle of Generative AI models, including RAG and multi-modal systems. My focus is on creating resilient, efficient, and scalable solutions that serve as the backbone for enterprise-level data science.

Career Highlights:

- Architected LLM evaluation pipelines integrating offline metrics, synthetic/adversarial data, online A/B tests, and CI/CD gates tied to OKRs.
- Engineered Neo4j Knowledge-Graph RAG assistant deployed in Slack/Teams for 200+ analysts, improving discovery speed and answer groundedness.
- Deployed multi-LLM orchestration (Gemini 2.5 Pro/Flash/Flash-Lite, Llama) with structured JSON outputs, function-calling, and safety filters.
- Established PromptOps & RAI guardrails, including prompt registries, canary testing, rollback playbooks, PII masking, and bias checks.
- Delivered multimodal AI prototypes (text + vision) and LLM-generated executive insights, reducing manual reporting by 60%.

EDUCATION

University of Maryland, College Park, Robert H Smith School of Business - *Master of Science in Business Analytics (MSBA)*
Big Data and Artificial Intelligence, Data Analysis with Python, Data Mining and Predictive Analytics, Data Models and Decision Making, Database Management System, Operation Analytics, Decision Analytics, Price Optimization and Revenue Management, Google Analytics.

SRM University, Kattankulathur - *Bachelor of Technology in Computer Science*
Web Technology, Software Engineering, Pervasive computing, Operating Systems, Network Security, Microprocessor and Interfacing, Linux Internals, Database Management Systems, Data Structures and Algorithm Design, Data Mining, Artificial Intelligence and Expert Systems.

TECHNICAL SKILLS AND CERTIFICATIONS

Languages: Python (Pandas, NLTK, Spacy, NumPy, Scikit, Seaborn, Matplotlib, Plotly), R, Spark, SQL, DAX, MATLAB
Technologies: GCP (Big Query, Vertex AI, Dataflow, Cloud Storage), BI Tools (Tableau, Power BI, SAP Analytics Cloud), Web Analytics (Google Analytics, Adobe Analytics), ETL Tools (Analytics Workbench, Alteryx), AWS (Redshift, S3), Other tools (GitHub, Advanced Excel (VBA), Microsoft SQL Server, Shell scripting, Hadoop, SAP)
Certifications: Google Professional Data Engineer (**Google**), Google Ads Professional (**Google**), Certified Data Scientist (**Data Camp**), Certified Data Analyst in Python (**Edureka**)

PROFESSIONAL EXPERIENCE

Gen AI / ML Engineer, *The Home Depot Management Company*   Jan 2025 – Present

- Designed and developed scalable generative AI systems using transformer-based architectures (GPT-4, BERT, Gemini, Longformer) for text summarization, Q&A bots, and contract parsing.
- Led fine-tuning of LLMs with **LoRA, QLoRA, and PEFT** methods using HuggingFace Transformers to improve model alignment with Home Depot-specific customer and vendor datasets.
- Built and deployed image-to-text pipelines integrating Stable Diffusion and Vision Transformers for intelligent product tagging and visual search enhancements.
- Created custom prompt optimization frameworks and integrated GenAI tooling for internal analytics automation, **reducing turnaround time by 55%.**
- Implemented RLHF to train chat-based support agents and automated documentation tools, increasing task **resolution accuracy by 38%.**
- Established MLOps pipelines with **TensorFlow Extended (TFX), Vertex AI Pipelines, and MLflow** for model versioning, validation, and deployment in staging and production environments.
- Partnered with engineering, product, and marketing teams to align model outcomes with business KPIs through dashboards and experiment tracking.
- Developed a modular GenAI microservice architecture on **GCP using Docker, Cloud Run**, and Firestore to power multiple real-time analytics and automations.
- Published internal technical documentation and model cards to guide ethics review, reuse, and reliability tracking for all deployed GenAI solutions.

Senior Data Scientist - Decision Analytics, *The Home Depot Management Company*   Jun 2023 – Jan 2025

- Architected a robust, scalable **dynamic image generation pipeline** utilizing state-of-the-art **vision models** (Google Image Gen-3, Stable Diffusion) and advanced **text generation models** (Gemini-1.5 Pro, Text-Bison-32k), transforming Home Depot's guided search with enhanced visual relevance and accuracy.
- Integrated multiple AI models, including **multi-modal** and **text embeddings** (Text-embedding-004, Gecko@002), to automate image-keyword alignment, optimizing search coherence and improving product discovery across Home Depot's platform.
- Deployed **LLM-powered automation** for end-to-end process reporting, delivering scalable, on-demand PDF documentation with code explanations for technical and non-technical stakeholders, cutting manual reporting effort by **60%.**
- Developed predictive models including **image classification**, **object detection**, and **house renovation score prediction** using Res-Net and Vision Transformers on MLS listing images provided by CoreLogic, achieving an accuracy of **87%** and aiming to save **$20 million** in marketing budget.
- Led a team of **4 offshore resources** (TCS) to optimize ETL workflows and develop ETL scripts in Analytical Workbench and Big Query, integrating data **(~50TB)** from **15 disparate sources** (clickstream, orders, marketing data), resulting in a **40% increase in data processing**.
- Built OLAP data cubes and architected databases, data warehouses to support Tableau Dashboards (SAIM Deck) and advanced data analysis, leveraging **SQL optimization, clustering, partitioning, stored procedures, and functions** on Google Cloud Big Query.
- Partnered with the Data Engineering team to migrate **25+ workflows** from AWB Workbench to **Google Data form**, targeting significant process efficiencies by automating Big Query SQL workflows, dependency management, and job scheduling—expected to cut project timelines by **30-50%** upon completion.
- Advancing data pipeline reliability and scalability through Data form's integrated testing, GIT version control, and optimized SQL transformations, with anticipated benefits of a **50% reduction in data error rates** and enhanced handling of large datasets for

- improved operational efficiency.
- Empowered **BACE (Brand Advocates and CEX) partners** with advanced analytics tools, enabling real-time monitoring and post-event analysis during critical events such as Black Friday and Cyber Monday to track key metrics, improving strategic decision-making for various projects to enhance customer experience.
- Created data standards and implemented new methods of capturing tagging information in **Adobe Analytics Tag Manager** by working with the Adobe Analytics team (AAPES team) to gain new analytical insights on customer interaction across all Home Depot online platforms.
- Enabled **5 internal organizations** to devise strategies for performing full category refreshes across all Home Depot online platforms to maintain and improve foundational stability of online categories (display taxonomy) and perform full-funnel analysis and optimization.

Senior Data Analyst, *The Home Depot Management Company* — March 2022 – Jun 2023
- Analyzed customer behavior across Home Depot platforms to provide key insights to the **Category Experience team (CEX)** and **Brand Advocate Team (BA)** with over **300 associates** by providing ad-hoc data, standardized real-time reporting, and offering business recommendations for senior executives.
- Enhanced the full-funnel customer experience by providing insights into online **Category Pages**, **Product Listing Pages (PLP)**, and **Product Information Pages (PIP)**.
- Constructed analytical dashboards using visualization tools like **Tableau** and **Google Data Studio**. Leveraged job orchestration tools such as Analytical Workbench and performed data manipulation using **Big Query** and **Python** (~70 hours/month).
- Delivered website performance analytics using **Adobe Analytics**, Tableau, and Big Query to derive analysis of **15+ events** (Black Friday, Winter Sale, etc.) over the year to improve the **Click-Through Rate (CTR)** and **Conversion Rate (CVR)** by optimizing content placement.
- **Led a team of 7** in the Voice of Associates (VOA) initiative, leveraging the Liftoff platform to streamline data science onboarding, reducing onboarding time by 20% and increasing satisfaction by 10% through department-specific insights shared with senior leadership.
- Acted as an **Adobe Analytics Workspace SME**, leading weekly Adobe Analytics Office Hours to provide live training and creating training resources and best practices.
- Fostered cross-functional partnerships, **mentored junior analysts**, delivered technical training and spearheaded various project initiatives.

Data Analyst and Engineer, *Harley Davidson Motor Company* — Feb 2020 – March 2022
- Performed in-depth analysis of general merchandise data to identify opportunities and develop proposals and recommendations for use by management.
- Designed, prepared, and manipulated data using Business Intelligence tools—**Tableau**, **Power BI**, and **SAP Analytics Cloud**—to identify user behavior and analyze trends and patterns, both independently and in collaboration with product managers and data modeling resources.
- Extracted, cleaned, and analyzed multiple data sources, and built optimized data models and **ETL pipelines** to support dashboard requirements using **SQL** and **Alteryx**, which improved the performance of existing reporting dashboards in SAP Analytics Cloud and helped reduce data processing time by **80%**.
- Maintained, enhanced, and drove **Root Cause Analysis** in conjunction with SMEs to identify and resolve business process problems, leading to a decrease in open purchase orders by **55%** and inventory count mismatches by **30%**.
- Maintained the master dataset of the General Merchandise department and performed batch inserting/updating of accounts, product information, lead times, BOM, dealer information, and other objects in SAP using **FLEX PLM**.
- Worked with the Supply Chain Analyst and warehouse coordinators to perform error analysis on **EDI transactions (IDoc Resolution)**, providing recommendations and analyzing all error data established for new product builds and launches, compiling and communicating weekly metrics to leadership.
- Led multiple rounds of **User Acceptance Testing (UAT)** by identifying appropriate stakeholders and building test scripts for each to execute.

Marketing Analyst, *Anahata Art and Design Pvt* — May 2019 – Dec 2019
- Created, maintained, and managed a **3-week Google Ads Campaign** with a total budget of $300 for an online gifting startup in India to understand their business, market competitors, popular selling products, and target audience.
- Managed to make **110 ad copies** with **6,000 keywords**, minimized cost-per-click to **$0.11**, achieved a **200% increase in website traffic** (92% new Users/week), achieved **113 sales** of products with **$ 3100 in revenue**, and improved the landing page experience.
- **Proposed multiple recommendations** for the potential new markets, product updates, and marketing campaign changes to the client after analyzing visits, page views, purchases, revenue, and conversion metrics from multiple data sources (web analytics data as well as external data).

Data Scientist, *Principal Financial group* — Aug 2019 – Dec 2019
- Devised the **KPI's** and predicted market regime of companies from Russell 1000 to evaluate prospective investments for the client.
- Performed data aggregation on 5.5 million rows of data using **SQL aggregation** techniques.
- Implemented data cleaning and hyperparameter tuning using **machine learning algorithms** in **Python** that resulted in an accuracy of **78%.**
- **Increased the prediction accuracy** of investing in one company by **~5%** by identifying the most important variables that increased the client's confidence in carrying out deals/investments.
- Tools and Libraries- Python, R, Microsoft SQL Server, Logistic Regression, Random Forest, XGBoost, dplyr, scipy.io, keras, numpy, matplotlib, sklearn.

Graduate Assistant, *University of Maryland* — May 2019 – Dec 2019
- Assessed, created and maintained student records (~4000 students) to identify low performing students to improve academic standing and disciplinary records.
- Performed **data extraction using SQL** to assist the academic advisors to implement various programs to improve a student's performance.
- Led a team of 10 undergraduate students to answer student inquiries that improved the **satisfaction rate by 10%**.

- **Designed visualizations** to understand and analyzed the career opportunities chosen by the students and helped advisors to improve the academic program based on the visualizations.

Data Analyst, *Bridge Solutions*   May 2017 – May 2018
- Created visually impactful and **interactive dashboards in Tableau** and Excel to report various key KPIs of various clients of Bridge Solutions.
- Handled and built relational databases, designed queries using **Microsoft SQL Server**, and created reports for analyzing and root-causing board failure data. Well- versed in finding patterns and trends in complex, multivariable data sets using **Python** in an **agile** environment.
- Created inventory targets by employing analytical abilities, data mining skills, and experience that resulted in a **cost reduction of $1M**.
- Introduced and developed **Docker** application to deploy **IBM OMS 9.5** and **WMS 9.5 which** is used by **75% of the workforce**.

ACADEMIC PROJECTS

Relational Database modeling *(Tableau Visualizations, Dashboards, MS SQL Server, Excel, Wix)*
- Built a relational database by **scrapping online data** of apartments near College Park, MD.
- Performed data cleaning, **ETL operations**, populated and normalized the database to **3NF** using **SQL queries**.
- Developed **interactive dashboards using Tableau** to help users compare the price, location, amenities for various houses and published the same to a website developed using Wix.
- Gained 2000 users in a span of 1 month and converted 250 users to active leads for the apartments.

Recommender System *(Machine Learning, Text Mining, Predictive Analysis, NLP, Python)*
- Implemented a model using Amazon reviews data (10 GB) to predict two most similar products based on item-item collaborative filtering using Python.
- Cleaned, trained, cross-validated, performed **NLP operations** like tokenization, lemmatization, stop word removal on the dataset and used KNN model to achieve **highest accuracy of 84.4%.**
- Built visualizations to analyze the reviews and built word clusters to discover the most used words for each review level using matplotlib and seaborn.

Age and Gender Prediction *(Machine Learning, Python, Predictive Analysis, Neural Networks, Transfer Learning, Python)*
- Built a **machine learning (CNN)** model to predict the age and gender of a person via their photos using python.
- Acquired 7 GB of data from IMDB-WIKI to perform **data cleaning, data aggregation and data manipulation**.
- Implemented **transfer learning** from VGG-face model, tuned the model to prevent overfitting and implemented the model on real time video stream.
- Achieved a **mean absolute error rate ±4** years in predicting the age and an **accuracy of 97%** in predicting the gender.

Healthcare Predictive Data Analysis *(RStudio, Machine Learning, SQL, Predictive Analysis, Visualization, R)*
- Analyzed a 60K+ dataset to predict patient revisit rate for emergency care hospital patients and discovered the correlation with other features.
- Performed **exploratory data analysis, multiple imputation and feature engineering** using R.
- **Implemented and evaluated** models like Logistic, KNN, Ensemble models, Decision Trees, XGboost.
- Secured 2nd position by achieving ~ **83 % accuracy**, thereby reducing the patient **readmission rate by 4%.**

LEADERSHIP EXPERIENCE

Executive Vice President, *Robert H. Smith School of Business / SMSA*   Dec 2018 – Dec 2019
- Planned and evaluated leadership growth programs, utilizing a fund of **$80,000 per semester**.
- Coordinated logistics for numerous events with club presidents under the **Smith Master Student Association (SMSA).**
- Launched an **alumni networking platform** to foster relationships and mentorship for incoming and existing students.

Track Representative, *Robert H. Smith School of Business / MSBA*   Aug 2018 – Dec 2019
- Collaborated with the **program manager** and **director** to improve the **MSBA** program.
- Resolved cohort grievances, achieving a **30% reduction** in grievances from previous batches.