Aishanee Sinha

San Jose, CA | aishanee.sinha@gmail.com | (669) 265-2905 | linkedin/aishanee-sinha | github/aishanee-sinha

Professional Summary

Data Scientist with 4+ years of experience in marketing and portfolio analytics within the financial services sector. Proven track record of architecting machine learning solutions that delivered $800k+ in fraud mitigation and a 23% revenue uplift. Expert in bridging the gap between advanced predictive modeling (Python, Scikit-learn, LLMs) and business strategy (A/B testing, KPI optimization) to drive high-impact insights for executive leadership.

Skills

Languages & Core: Python (Pandas, NumPy), SQL, SAS
Machine Learning: Scikit-learn, XGBoost, K-Means, A/B Testing, Decision Tree, Random Forest, Logistic Regression
Gen AI: PyTorch, Keras, vLLMs, VectorDB (Pinecone), LLM Fine-Tuning (PEFT), NLP, RAG, Agentic AI, MCP, A2A Protocol
Cloud & Big Data: PySpark, AWS, GCP, Docker, Snowflake, dbt, Airflow, ETL/ELT Data Pipelines, CI/CD
Other Tools: Tableau, PowerBI, Github, Jira

Work Experience

Business Systems and Data Manager (Intern)   May 2025 – Aug 2025
PACCAR Financial   Bellevue, WA, USA

- Built a python tool using NLP techniques to analyze auction pdf reports and visualize insights to identify pricing and market trends in used truck sales, reducing manual market trend analysis time by **80%.**
- Optimized marketing budget allocation by analyzing lead conversion data, recommending targeted ad placements that identified **$200K in potential savings** per quarter and projected a **15% increase in Campaign ROI.**
- Developed a robust Tableau inventory dashboard linked to live Snowflake data, to automatically flag data anomalies during daily updates. It improved **data integrity scores by 30%** and streamlined the reporting process for the sales team.

Lead Data Scientist   Aug 2020 – Aug 2024
Citibank *(Global Decision Management)*   Bengaluru, India

- Led **automation of 14+ campaign KPI dashboards**, leveraging PySpark and SQL for data orchestration and achieving a **70% reduction** in manual workload.
- Championed Citi Ignite 2023 global ideation challenge as a semifinalist, proposing an ML-driven win-back model to **re-engage 1.5M inactive customers** at a lower acquisition cost.
- Architected scalable PySpark data pipelines on Hadoop clusters, integrating portfolio, transaction, and CRM data for unified customer analytics.
- Designed and deployed customer segmentation models using K-Means, Decision Trees, and Random Forests, boosting campaign targeting efficiency and revenue by **23%.**
- Improved existing **anomaly detection models** to identify fraudulent campaign activity, **saving $800K.**
- Designed **A/B tests** to optimize campaign offer strategy and contact frequency, driving **20% higher conversion rates.**

Academic Project Highlights

Agentic AI-Powered Workforce Assistants

- Autonomous AI agents integrate with Slack, Jira, email to manage routine tasks such as scheduling meetings, summarizing lengthy documents, organizing tasks, and proactively anticipating and resolving productivity bottlenecks.
- **Fine-tuning LLMs** using **LangChain framework** for natural language understanding and using **Chain of Debate (COD)** Method for resolving bottlenecks.

Multi-Modal AI Chatbot for Health Insights

- Built a context-aware **multimodal chatbot using LLMs** and Pinecone retrieval to generate conversational answers from text and images. Designed an end-to-end chatbot with query reformulation, citation handling, and conversation memory.

Education

MS Applied Data Intelligence | San Jose State University **(GPA: 3.9/4)**   Aug 2024 - May 2026 (Grad)

- **Graduate Student Assistant** - *Generative AI, Big Data Technologies*
- *Relevant Coursework : Machine Learning, Deep Learning, Gen AI, Big Data, Data Engineering, Distributed Systems*

MS Economics | Indira Gandhi Institute of Development Research   Jul 2018 – Jul 2020
BE Electronics & Telecommunication | Jadavpur University *(Gold Medalist)*   Aug 2013 – Jul 2017