Sai Charan Somineni

Fairfax, VA — Open to Relocate

📞 +1-832-931-1949 | ✉️ ssominen@gmu.edu | 💼 linkedin.com/in/sominenisaicharan | 🐙 github.com/saicharansom

EDUCATION

George Mason University — Fairfax, VA
M.S. in Data Analytics Engineering, GPA: 3.8/4.0 — *Aug 2023 – May 2025*

SRM Institute of Science and Technology — Chennai, India
B.Tech in Computer Science and Engineering (First Class with Distinction) — *Jun 2018 – May 2022*

TECHNICAL SKILLS

Languages: Python, SQL, R, C++, Java, JavaScript
ML/AI: Supervised/Unsupervised Learning, GPT, BERT, RAG, LLM Finetuning, Transformers
Frameworks: Scikit-learn, TensorFlow, Keras, PyTorch, FAISS, spaCy, Pandas, GeoPandas, Plotly
Cloud/Data: AWS, GCP, Azure, PySpark, Databricks, Apache Spark
Databases: MySQL, MongoDB, AWS Neptune, Neo4j, GraphQL
DevOps & Tools: Docker, Kubernetes, Prometheus, Git, Jira, Tableau, Power BI, Looker Studio, Jupyter, Streamlit, VSCode, OpenRefine, Figma

PROFESSIONAL EXPERIENCE

Data Analysis and Visualization Intern — June 2025 – Present
Pearl Matibe — International Security and Counterterrorism — *Washington, DC*
- Analyzing and visualizing data related to U.S.-Africa relations, leadership travel, and open-source intelligence (OSINT).
- Developing high-impact dashboards, maps, and reports that inform journalism, public discourse, and international policy analysis.
- Cleaning, standardizing, and enriching data from 100+ diplomatic and security engagement events across 50+ countries.
- Collaborating directly with Pearl Matibe to align analytical outputs with public communications and investigative goals.

Research Assistant, CAPS Lab — Jan 2025 – May 2025
George Mason University — *Fairfax, VA*
- Conducted NLP-driven literature synthesis using NLTK/spaCy over 250+ research papers on adaptive authentication.
- Modeled usability-privacy-security trade-offs using Scikit-learn, informing secure design prototypes.
- Researched trends in infrastructure security and adversarial AI, shaping strategic lab publications.

Business Analyst — Aug 2022 – Jun 2023
The Detailing Mafia Velachery — *Chennai, India*
- Automated performance dashboards for 25+ franchises using Power BI, driving a 22% ROI boost.
- Analyzed customer data via Python & SQL, increasing upsell rates by 12% and cutting operational overhead by 15%.
- Delivered weekly data insights that shaped cross-channel marketing strategies.

Implementation Consultant — Jan 2022 – Jul 2022
Digite Infotech Pvt Ltd — *Remote*
- Improved cross-system data syncs by 30% through streamlined ETL designs.
- Led client-facing sessions to align SaaS tools with business workflows, boosting adoption.
- Applied AI-based diagnostic metrics, increasing operational throughput by 40%.
- Integrated observability and monitoring solutions using Prometheus and custom instrumentation for Java-based systems.
- Built tooling and dashboards to improve visibility into SaaS platform performance and alerting workflows.
- Collaborated with engineering teams to implement instrumentation strategies for key microservices.

Research Intern, SPIHER Lab — Oct 2020 – Dec 2020
SRM University — *Remote*
- Built a personalized cross-modal recommendation system with DLRM, improving engagement by 25%.
- Integrated MongoDB with a Keras-trained backend for Bookworm-Innovations.
- Conducted A/B testing for user experience enhancements.

PROJECTS

Augment Linkage: Evidence & Exploration | *AWS Neptune, Neo4j, Gremlin, Python* — Jan 2025 – May 2025
- Designed and deployed a scalable graph database system for TraCCC, mapping 10K+ nodes and 15K+ relationships to uncover illicit fentanyl supply networks.
- Modeled and automated evidence-weighted linkage algorithms, improving network detection efficiency by 42% and reducing false-positive connections.
- Implemented Gremlin queries and AWS Neptune visualizations for 5+ law enforcement agencies to explore high-risk clusters.
- Streamlined entity normalization, reducing manual investigation time by 60%.

PrivAware Voice Assistant | *GPT-3.5 Turbo, Streamlit, Python* — Sep 2024 – Dec 2024
- Built and fine-tuned a privacy-preserving voice assistant with real-time semantic filtering and role-based admin control.
- Integrated RAG to improve response accuracy by 35% while ensuring ethical data handling.
- Reduced latency by 28% and enhanced user interaction in pilot trials.

Domain-Specific QA for Pavement NDE | *BERT, GPT-2, FAISS* — Aug 2024 – Nov 2024
- Built a technical QA engine for the Federal Highway Administration using fine-tuned LLMs.
- Curated and embedded pavement engineering datasets with FAISS for real-time QA retrieval.
- Achieved 86% answer accuracy on domain-specific queries.