



SAI CHARAN SOMINENI

Fairfax, VA

+1-832-931-1949 ✉ ssominen@gmu.edu  <https://www.linkedin.com/in/sominenisaicharan/>  <https://github.com/saicharansom>

Education

George Mason University

Master of Science, Data Analytics Engineering

- GPA: 3.78/4.0

Aug 2023 - May 2025

Fairfax, Virginia

SRM Institute of Science and Technology

Bachelor of Technology, Computer Science and Engineering

- **Achievements:** Graduated with First Class and Distinction

Jun 2018 - May 2022

Chennai, Tamil Nadu

Technical Skills

- **Programming Languages:** Python, R, SQL, C++, Java, JavaScript, HTML/CSS
- **Machine Learning / AI:** Supervised & Unsupervised Learning, Neural Networks, NLP, LLMs (GPT, BERT, LaMA), RAG, Finetuning Large Language Models (LLMs), Generative AI
- **Frameworks & Libraries:** Scikit-learn, TensorFlow, Keras, Pandas, NumPy, Matplotlib, Seaborn, Plotly, BeautifulSoup, ggplot2, Common Vulnerabilities and Exposures
- **Deep Learning & Tools:** PyTorch, Megatron-LM
- **Data Visualization & BI:** Power BI, Tableau
- **Cloud & Big Data:** AWS, Azure, GCP, PySpark, Databricks, Apache Spark
- **DevOps & Tools:** Docker, Kubernetes, Prometheus, Git, Jira
- **Databases:** MySQL, NoSQL, MongoDB, AWS Neptune, GraphQL, Graph database, Neo4j
- **Other Tools:** Jupyter, VSCode, RStudio, Anaconda, Cursor

Professional Experience

CAPS Lab - George Mason University

Research Assistant

- Conducted in-depth literature reviews and performed thematic analysis using Python libraries such as NLTK and spaCy to uncover emerging patterns in adaptive authentication technologies.
- Analyzed academic and industry research to identify critical usability, security, and privacy trade-offs, supporting evidence-based recommendations for adaptive authentication strategies.
- Designed clear, insightful dashboards and visual reports using Tableau, Matplotlib, and Seaborn to effectively present findings to faculty and stakeholders.
- Utilized Scikit-learn to explore basic predictive modeling techniques, translating theoretical research insights into initial data-driven validations.

Jan 2025 - Present

Fairfax, VA

Digite Infotech Pvt Ltd

Implementation Consultant

- Improved data flow efficiency by 30% through streamlining data integration processes with enterprise-class architecture.
- Implemented AI-driven analytics and conducted statistical analysis on datasets to improve efficiency by 40%.
- Collaborated cross-functionally and enhanced project lifecycle with strong data validation and reporting tools.

Jan 2022 - Jul 2022

Remote

SPIHER Lab - SRM University

Research Intern

- Developed a multimodal recommendation system using DLRM and cross-modal learning, enhancing user experience by integrating MongoDB for data management and Keras for model training
- Increased engagement for Bookworm-Innovations by 25% with personalized strategies.

Oct 2020 - Dec 2020

Remote

Projects

AUGMENT LINKAGE – Evidence & Exploration

TRACCC(Terrorism, Transnational Crime and Corruption Center)

- Led as Product Owner for a 7-member team partnering with TraCCC to design a graph-based system uncovering hidden connections in illicit fentanyl supply chains using AWS Neptune and Neo4j.
- Automated association discovery across 10,000+ entities (companies, emails, IPs) by resolving identifier inconsistencies (e.g., phone/email formats) and integrating raw TraCCC datasets (sanctions, PDFs), improving linkage accuracy by 35% and reducing pre-processing time by 30%.
- Enabled law enforcement to prioritize high-risk trafficking clusters through dynamic graph exploration and evidence-weighted filtering, scaling the framework for future anti-fraud investigations.

Jan 2025 - May 2025

Fairfax, VA

Developing a QA System for Pavement NDE Technologies

George mason university Academic Project

- Designed a domain-specific QA system for Federal Highway Administration to address challenges in accessing pavement evaluation technology insights.
- Implemented fine-tuned BERT, GPT-2, and Groq LLMA 3.1 models for precise and fluent responses.
- Utilized FAISS for efficient data retrieval and built a structured dataset from complex technical PDFs.

Aug 2024 - Nov 2024

Fairfax, VA

PrivAware: A Privacy-Centric Voice Assistant for Ethical Ad Experiences

George mason university Academic Project

- Developed a full-stack voice assistant using OpenAI's GPT-3.5 Turbo and Streamlit that enables privacy-preserving ad interactions.
- Incorporated voice-based semantic classification of sensitive categories like gambling, mental health, etc., ensuring user-controlled content filters.
- Integrated user authentication, secure data handling, and admin interfaces to support anonymized analytics.

Sep 2024 - Dec 2024

Fairfax, VA

Limitless Tales: A Fine-tuned GPT Model for Story Generation

George mason university Academic Project

- Adapted a pre-trained GPT model on custom text datasets to generate personalized narratives.
- Developed a user-friendly interface with speech synthesis and AI-driven image generation.

Jan 2024 - May 2024

Fairfax, VA