Saahil Sanganeriya

Atlanta, GA | saahilsanganeriya@gatech.edu | linkedin.com/in/saahil-sanganeriya | github.com/saahilsanganeriya | 404-353-0965

Education

Georgia Institute of Technology **Atlanta, GA**
Master of Science in Machine Learning *Graduating December 2025*
Bachelor of Science in Computer Science | Threads: *Theory, Intelligence* | *Major GPA: 3.85* *Graduating December 2024*

Professional Experience

The Coca-Cola Company, *Data Engineer Co-op* *May 2024 – August 2024*
- **Revamped Internal Ticketing System:** Redesigned the backend data model & created a scalable web interface using Django, HTML, & JavaScript, significantly improving functionality for 1,000+ users & facilitating seamless cross-team adoption.
- **API Development:** Developed a robust API for CRUD operations across two large databases, streamlining data management for internal teams & consultants like McKinsey. Created comprehensive documentation that reduced onboarding time & facilitated seamless integration into existing workflows, enhancing data accessibility & operational efficiency for over 300 datasets.

The Coca-Cola Company, *Data Engineer Co-op* *September 2023 – December 2023*
- **AI:** Envisioned & led the development of an AI chatbot for query resolution, leveraging open-source technologies like Haystack & RoBERTa based NLP models from Hugging Face to implement a question-answering system that retrieves information (BM25Retriever & FARMReader) in real time from text-based indexed datasets (Elasticsearch Document Store)
- **Metrics Dashboard:** Spearheaded design of a proprietary monitoring platform by reengineering the CPS Python Library to enhance inference accuracy & embed metrics collection, pre-empting data integrity issue identification to reduce external complaints by 63% & optimizing daily pipeline run times from 9 to 6.5 hours, enhancing global processing efficiency
- **Cost Saving Projects:** Identified production bottlenecks at the Buenos Aires plant during plant visit, subsequently using predictive analysis on historical data to curate a Power BI dashboard, generating substantial cost savings; integrated internal training REST API & Workday API (Postman) to create training metrics dataset, side-stepping expensive external solutions

The Coca-Cola Company, *Data Engineer Co-op* *May 2023 – September 2023*
- **ETL:** Orchestrated Azure Data Factory pipelines, retrieving data from SAP backends & automating 300+ datasets (60% time reduction) through PySpark, NumPy, Pandas, & SQL transformations on large-scale compute Clusters using databricks
- **CI/CD:** Championed iterative dataset development/enhancement, using mathematical preprocessing, employing rigorous issue-tracking mechanisms & transition strategies accompanied by daily pipeline run reviews to maintain compatibility
- **Storage:** Pioneered dual-layered storage using ADLS for raw data & Azure SQL Data Warehouse for structured repositories

Research

Precision Biosystems Lab, *RA - Machine Learning for Label-Free Biological Cell Segmentation* *August 2024 – Present*
- Improved SAMCell's segmentation accuracy using computational geometry to delineate densely packed cell boundaries.
- Developed a Voronoi-based post-processing pipeline to enhance cell boundary detection under Dr. Craig Forest.

Quantum Computing with Future Computing with Rogues Gallery, *Neural Net Team* *January 2022 – May 2022*
- Formulated neural net architecture (keras, tensorflow) for cross-checking LSTM-driven circuit optimization algorithm
- Developed one-hot encoding to match categorical input (categories are gate types) to pass into both neural networks

Projects

Real-time Accident Risk Assessor (ML) *January 2023 – May 2023*
- Formulated a neural net architecture & trained it using vehicle accident data from Maryland; divided 155,000 data points over 7 years into dynamic features like weather, road/traffic conditions, & static features like vehicle condition
- Applied gradient boosting & random forest to make a predictor that takes real time data & provides risk assessment

Autonomous Robot Code base development—ME Creative Design Competition *January 2022 – April 2022*
- Designed an autonomous robot for the Creative Decisions & Design Machine Competition; developed a time sensitive code base in Arduino C++ to perform 3 tasks: engaging a door, a throwing mechanism & a sliding mechanism

Tideman Voting Algorithm Emulator *August 2020 – October 2020*
- Studied the Tideman Voting Algorithm & implemented a graph-based approach in C by using recursive functions
- Ensured Acyclic graph formations by 'locking' edges with based on edge weight, used graph to find ranking-based winner

Leadership

Invention Studio at Georgia Tech*, Prototyping Instructor* *January 2021 – May 2023*
- Provided design consultation, supervision, & training for prototyping technologies including 3D printing, water jetting, laser cutting, electronics, soldering, welding, woodworking, metalworking, CNCs, Mills, Lathes, etc. [200+ hrs]
- Maintained & repaired equipment, ordering parts & consumables, and ensured a high standard of reliability.

Skills

C/C++, Java, Python, SQL, NumPy, Pandas, PySpark, Django, Javascript, Deep Learning, Convolutional Neural Networks, Algorithms, Motion Planning, Keras, TensorFlow, Scikit-learn, pytorch, Azure Data Factory, Databricks, DevOps, Data Lake, Postman, REST & SOAP APIs, Haystack NLP, Rasa, JSON, PowerBI, Linux, Docker, Android Studio, Assembly, Qiskit