Samir Varma

732-401-6181 | sv773@scarletmail.rutgers.edu | linkedin.com/in/samir-varma-1a1166281/ | github.com/SamirVarma01

Education

Rutgers University, New Brunswick - Honors College GPA: 3.967/4.000

B.S. in Computer Engineering, B.S. in Mathematics, Minor in Physics *Aug. 2023 – May 2027*

- Involvement: IEEE Honors Society, Honors Academy Events Committee, Data Science Club, Rutgers Blueprint

Experience

Data Science Intern June 2025 – Aug 2025

AT&T *Middletown, NJ*

- Built and deployed **Isolation Forest-based intrusion detection services** via **Scikit-Learn** to flag malicious IPs from high-volume network telemetry, processing **3M+ traffic records** with low-latency inference.
- Designed modular Python training and evaluation pipelines with **GridSearchCV** and parameter sweeps, improving anomaly detection precision by **8%** while maintaining stable runtime performance.
- Engineered a scheduled ETL workflow using **Apache Airflow** to orchestrate data ingestion, validation, and feature preprocessing from **Azure Databricks**, enabling reliable daily model refreshes.
- Refactored legacy notebooks into production-ready components with clear interfaces, logging, and failure handling, reducing manual intervention and cutting end-to-end pipeline runtime by **40%.**

Software Engineering Intern Jan. 2025 – Apr. 2025

Stealth Mode Startup *San Francisco, CA*

- Engineered vision-language photo editing pipelines by integrating **OpenCV, Grounding DINO and Segment Anything (SAM)**, translating natural-language prompts into segmentation masks and targeted image edits.
- Developed Python orchestration scripts for prompt parsing, object grounding, segmentation, and post-processing, reducing manual workflows by **60%** and enabling fully automated, prompt-driven transformations.
- Containerized end-to-end inference pipelines with **Docker** and instrumented **MLflow** for reproducible evaluation, tracking latency, segmentation quality, and model versions.
- Built a **Selenium-based ETL pipeline** to scrape **5,000+** images from Reddit and store datasets in **Google Cloud Storage** for testing and evaluation.

Machine Learning Research Assistant June 2024 – Apr. 2025

Rutgers University *Piscataway, NJ*

- Implemented **Actor–Critic reinforcement learning agents** in **PyTorch** within the CARLA autonomous driving simulator, designing modular policy, value, improving policy learning by **14%.**
- Built an experimental evaluation framework to compare **DQN** vs. **Actor–Critic architectures**, instrumenting reward curves, convergence behavior, and policy stability across parallel simulation runs.
- Optimized training throughput via **batched environment rollouts, GPU-accelerated inference (CUDA), and structured logging (TensorBoard)**, reducing experiment iteration time by **26%**

Projects

NexusML | *Go, Python, FastAPI, GitPython, Typer, Scikit-Learn, PyTorch, Amazon S3, Google Cloud SDK, Docker*

- Engineered a **two-plane ML inference platform** combining a **Go-based inference proxy** using dynamic batching to achieve **10×** GPU utilization and a **FastAPI model server** supporting **sklearn/PyTorch** models and **S3/GCS** loading; containerized and orchestrated via **Docker/Docker Compose.**
- Designed a **Git-integrated model versioning CLI** that ties ML artifacts to commits via lightweight JSON metadata, enabling reproducible deployments while keeping repositories lean with **S3/GCS storage**.

NextLevel | *JavaScript (Node.js, Next.js/React), TypeScript, MongoDB, Tailwind CSS, AWS (Lambda, S3)*

- Developed a **full-stack game review platform** using **Next.js, React, TypeScript, Node.js, and MongoDB**, integrating the IGDB API to serve metadata for **400K+** games, with secure authentication, **RESTful API endpoints**, and scalable data models supporting user reviews and ratings.
- Implemented review creation and engagement features with a responsive **Tailwind CSS UI** and **AWS (S3, Lambda)** backend services, supporting **30+** active users with reliable media storage and low-latency page loads.

Technical Skills

Languages: Java, Python, Golang, C/C++, SQL, JavaScript, TypeScript, HTML/CSS
Frameworks & Libraries: React, Node.js, Next.js, Tailwind CSS, Flask, Django, FastAPI, PyTorch, TensorFlow, Scikit-Learn, Pandas, NumPy, Matplotlib, BeautifulSoup, Selenium
DevOps, Cloud & Databases: Git, GitHub, GitLab, AWS, GCP, Databricks, MLflow, Docker, Kubernetes, Apache Airflow, MongoDB, Postgres, MySQL