

Tanay Shah

+1 (240) 413-7915 | tanayshah2024@gmail.com | linkedin.com/in/tanayshah11 | github.com/tanayshah11

EDUCATION

University of Maryland

Bachelor of Science in Computer Science Honors, Minor in Statistics

College Park, MD

Aug 2022 - Dec 2025

EXPERIENCE

Systems Software Engineering Intern

June 2025 – Aug 2025

Intuitive Labs

San Francisco, CA

- Built distributed market research platform with **AWS Fargate**, FastAPI, and **Redis**, processing 500K+ companies daily across 8 containerized microservices with auto-scaling and load balancing, reducing analysis time by 30x.
- Engineered ensemble ML pipeline integrating **Llama 4** & 5 scikit-learn text classification models achieving 97.3% accuracy across 125+ data points with anti-hallucination validation (backed by ML Models).
- Implemented **asyncio**-based browser automation with Playwright, Docker orchestration, and context pooling managing 1500 concurrent sessions with memory optimization and resource recycling across 25 industry verticals.
- Developed real-time analytics dashboard using **Streamlit**, Redis pub/sub, **WebSocket** protocols, and asyncio event loops handling 20K+ events/hour with cross-container telemetry aggregation and sub-150ms latency.

Software Developer - Scalable ML Infrastructure

May 2024 – Present

College of Information Studies, UMD

College Park, MD

- Built and deployed an end-to-end machine learning pipeline for bias classification on Wikipedia talk pages, achieving **90%+ accuracy** using NLP techniques on 20M+ AfD comments.
- Engineered **multithreaded** Python web scrapers in Selenium with **aihttp** and **asyncio**, boosting data collection efficiency by **8300%** as compared to standard MediaWiki API calls.
- Developed **distributed data processing pipelines** with Apache **Kafka** & **Spark**, improving data throughput by 40% across real-time ML workflows.
- Enhanced model performance by 25% through fine-tuning of **Llama3.2** using **LoRA** on scraped data.
- Implemented NLP-based bias detection using **TF-IDF** and topic modeling, improving precision by 40%.

Machine Learning Intern

May 2022 – Aug 2022

Softvan Pvt. Ltd.

College Park, MD

- Developed CNNs using Pytorch and TensorFlow, achieving an **88–90% accuracy** rate in spatial object categorization, enhancing predictive modeling capabilities for applications in vehicle navigation and surveillance systems.
- Optimized data preprocessing workflows using OpenCV and NumPy; implemented image resizing, grayscaling, and Gaussian blurring to **reduce data processing time by 30%** & enable real-time analysis for applications in traffic monitoring and video analytics.
- Boosted model robustness by adversarial training (used FGSM), model ensembling, & automated Bayesian hyperparameter optimization; reduced validation error by 15% & improved reliability under real-world conditions.

PROJECTS

Distributed E-Commerce Platform | Node.js, Docker, AWS

- Built a **microservices** architecture supporting 2K+ concurrent users; integrated **Redis caching** and **RabbitMQ** for asynchronous order processing, ensuring sub-second latency and **scalable** deployment via **AWS ECS**.

G.E.O.P.A.L | Python, Flask, React, Machine Learning, Object Recognition

- Developed GEOPAL's geospatial analytics engine using Flask, **Python**, and Google Earth Engine (GCP); **parallelized** metric extraction across 20+ features using **multithreading** to reduce API latency by **99.93%**; integrated facial-based authentication using OpenCV and NFT-based decentralized tokenization via the Verbiwire API.

TECHNICAL SKILLS

Languages: Java, Python, C/C++, C#, **SQL (MySQL)**, NoSQL (MongoDB), JavaScript, R, Rust, Golang

Frameworks/Libraries: React, Node.js, Flask, Django, PyTorch, Numpy, Pandas, **Kubernetes**, **Agile**

Tools: Git, AWS, Google Cloud Platform, Azure, Linux, **JIRA**, Docker

Other Skills: **REST**, GraphQL, Web Services, **Load Balancing**, Microservices, CI/CD