Vatsal Malkari

https://www.linkedin.com/in/vatsal-malkari | vatsalmalkari@gmail.com | https://github.com/vatsalmalkari | https://vatsalmalkari.github.io/portfolio/ | (925) 968-8233

EDUCATION

Rutgers University, New Brunswick Expected Dec 2026

- *Master of Science in Data Science*
- **Coursework**: Probability & inference, Regression & time series analysis, Data Structures & algorithms

Rutgers University, New Brunswick Sep 2022 – May 2025

- *Bachelor of Science in Computer Science & Data Science*
- **Coursework**: Data Structures & Algorithms, Systems Programming, Computer Architecture, Database Management, Deep Learning, Regression & Statistical Methods

EXPERIENCE

New Jersey Cybersecurity and Communications Integration Cell Jan 2025 – May 2025

Data Engineering Extern

- Built a scalable data pipeline processing 2M+ weekly login events to identify top 1% highest-risk accounts.
- Implemented anomaly detection in BigQuery ML, improving threat detection accuracy by 22.4%.
- Automated threat classification using PCA and vectorized workflows, reducing analyst investigation time.

PROJECTS

Bristol Myers Squibb Challenge Winner: **Protein Modification Engine** Oct 2025 – Dec 2025

- Designed a high-performance Python engine to model protein modification probabilities across 100+ sites.
- Optimized algorithmic complexity using memoization, LRU caching, convolution, achieving 41× runtime speedup.
- Deployed user-friendly executable for 500+ Bristol Myers Squibb biomedical researchers to use.

Spoiler-Controlled TV Show RAG Agent Nov 2024 – Jan 2025

- Built a spoiler-controlled RAG system with separate offline embedding and online inference pipelines.
- Embedded 400+ documents using a PySpark pipeline, achieving ~20s indexing time.
- Improved retrieval robustness with ChromaDB vector search and RapidFuzz fuzzy matching.
- Deployed a FastAPI + Gradio application backed by Google Gemini, delivering sub-second latency.

Custom C++ Deep Learning Engine Sep 2025 – Jan 2026

- Constructed a C++ deep learning engine to optimize neural network execution and memory management.
- Built CNN layers with automatic differentiation and backprop, reaching 45% FER-2013 accuracy.
- Integrated real-time webcam inference (<100 ms latency) using Python and OpenCV.

Python Interpreter in C Aug 2025 – Nov 2025

- Developed a Python interpreter in C to explore runtime design, parsing, and memory management.
- Implemented tokenization, recursive-descent parsing, and execution; validated with 40+ programs.
- Optimized performance using manual memory management and O(1) dictionary lookups.

Low-Level Task Scheduler Mar 2025 – Jun 2025

- Built a thread-safe job scheduler in C with priority scheduling (min-heap) and O(1) task lookup.
- Cut CPU overhead by replacing busy-waiting with condition variables and synchronization primitives.
- Stress-tested with 2,000+ concurrent tasks and implemented logging to debug race conditions.

SKILLS

- **Languages**: Python, C++, C, Java, SQL, JavaScript
- **Machine Learning**: PyTorch, TensorFlow, Transformers (Hugging Face), OpenCV, Scikit-Learn, Pandas, NumPy
- **Backend**: FastAPI, Flask, Spring Boot, Docker, PostgreSQL, ChromaDB, Git, CI/CD, Django
- **Systems Engineering**: Multithreading (pthreads), Concurrency, Memory Management, OS Scheduling, Linux/Unix
- **Concepts**: Object-Oriented Programming, Data Structures & Algorithms, REST APIs, System Design, JWT Auth

LEADERSHIP

Alpha Phi Omega Co-Ed Service Fraternity

Webmaster, Brotherhood Engagement Committee, Alumni Chair

- Managed chapter website (Squarespace) and coordinated outreach events to boost alumni and brotherhood engagement.