Aayush Anand

Brooklyn, NY, 11223 +1(347)481-8897
aa13015@nyu.edu aayushanand2710 AayushA10 Portfolio

Technical Skills

Programming Languages: Python, C, C++, SQL, HTML/CSS, R
Frameworks & DevOps: Django, Flask, FastAPI, React, Node.js, Next.js, Spring Boot, Docker, Kubernetes, Jenkins, Git, JUnit, Mockito
ML / Big Data / AI: TensorFlow, PyTorch, Scikit-learn, OpenCV, H-F Transformers, Hadoop, Spark, Kafka, Hive, MapReduce, HDFS, FAISS, MongoDB, NoSQL
Cloud / Platforms: AWS, Microsoft Azure, GCP, Linux, Windows, Web, iOS, Arduino

Work Experience

Zof AI — San Francisco, USA **May 2025 – Present**
Software Engineering Intern

- **Engineered a scalable model-validation pipeline** in **Python + PyTest**, achieving over **90%** unit and integration test coverage across 120+ models, significantly reducing post-deployment failures.
- **Optimized PyTorch inference workflows** using CUDA kernel tuning, mixed precision, and TorchScript; reduced average GPU inference latency by **25%** and increased throughput to 18 QPS on A100 GPUs.
- **Benchmarked 3 open-source LLMs** (7B–13B parameters) for latency, accuracy, and memory usage; enforced performance regressions via **GitHub Actions and Docker**, preventing over 1,000 unstable PRs annually.

Indiana University — Indianapolis, USA **Jun 2025 – Aug 2025**
Data Analytics Research Intern

- **Processed 2.1 million GC/MS spectra** (12 GB mzML files) using **R/xcms, eRah, and Pandas**, reducing feature extraction time from 3 hours to **45 minutes**.
- **Increased XGBoost model accuracy from 68% to 82%** through Bayesian hyperparameter optimization and SHAP-based feature selection; deployed using **FastAPI**.
- **Designed and deployed an R Shiny dashboard** with **Plotly and dplyr** for real-time QC monitoring, adopted by **5+ academic departments**; powered by a **SQLite data mart with Airflow ETL**, reducing query latency by **70%**.

Projects

AI Career Coach (**GitHub**)
Flask, OpenAI API, SQLAlchemy, HTML/CSS, Python, Jinja2, REST API

- Built an AI assistant used by **50+ users**, leveraging the **OpenAI GPT API** to provide tailored resume feedback, job recommendations, and interview tips.
- Achieved **85%+ matching accuracy** between resumes and job descriptions using **NLP-based semantic similarity** with embedding models.
- Designed a responsive frontend using **HTML/CSS + Jinja2** and integrated REST APIs with **Flask** and **SQLAlchemy** for modular full-stack architecture.

Chip Design Optimization using GNNs (**GitHub**)
PyTorch Geometric, Python, Graph Neural Networks, NumPy, NetworkX

- Modeled VLSI chip floorplans as graphs; trained **GNNs** to reduce delay and congestion, improving placement efficiency by **22%**.
- Achieved **15% faster convergence** over baseline ML models by customizing message-passing layers in **PyTorch Geometric**.
- Validated design on **10+ synthetic benchmarks**, bridging hardware and ML workflows for EDA (Electronic Design Automation).

Education

New York University **Expected May 2026**
Master of Science in Computer Science *New York, USA*

- **Relevant Coursework:** DAA, ML, AI, HCI, Big Data, Principles of Database

SRM Institute of Science and Technology **Sep 2020 – May 2024**
Bachelor of Technology in Computer Science and Engineering *Chennai, India*

- **Relevant Coursework:** OOPS, OS, CN, Data Mining, Software Engineering, Advance Programming