Ayush Yadav

(513) 461-4375 | aesh.03.23@gmail.com | linkedin.com/in/ayush-yadav-developer | github.com/yadava5 | ayush-yadav.com

SUMMARY

Computer Science graduate (May 2026) who builds machine learning and full stack systems end to end, grounded in a year of data engineering work. Five live, public projects, including a Gmail classifier that runs inference fully in the browser via int8 ONNX (90 to 23 MB), and a JDK 25 compression engine 6.4x faster than single-threaded java.util.zip.

EDUCATION & AWARDS

Miami University, Oxford, OH | B.S. in Computer Science, Major GPA 3.65 May 2026

Relevant coursework: High-Performance Computing, Data Structures & Algorithms, Machine Learning, Deep Learning, Databases

Certificates (2026): Azure AI Essentials, Snowflake Data Engineering, Data Analysis & GitHub (Microsoft)

- Dean's List: Fall 2023, Spring 2025, Fall 2025.
- Finalist, MUCAT Design Innovation (Feb 2025): LiDAR visual assistance proposal; $2,500 prototyping grant.
- Social Innovation Weekend (Mar 2025): co-built **LifeQuest**, a React and NestJS goal tracking platform (7-person team).

EXPERIENCE

ITSM Data Integration Intern, Miami University | Oxford, OH Jun 2025 – May 2026

- Engineered a **Python** pipeline that ingested 1.6M+ Oracle Analytics (OAS) query logs (5 years; 1,153 users, 66 dashboards) into a 57.8M-row field usage table in **SQLite**, flagging fields in use for an OAS-to-Tableau migration.
- Unified asset data siloed across **Tableau** and **Workday** into one 10,453-row, 35-field master inventory of dashboard ownership, using hash-based deduplication and the **Tableau REST API**, backed by **ruff and pytest CI**.
- Refactored a legacy **Laravel** compliance reporter (TeamDynamix ITSM, GitLab, End-of-Life API) into an ETL feed powering a 37-month **Tableau** dashboard, taking tracked code compliance from 0% to 97% across 61 projects.

PROJECTS

Applied | Gmail Job Application Classifier 2026

Solo Developer | getapplied.vercel.app

- Developed a Gmail job application classifier combining 201 regex rules, **e5 embeddings**, and a fine-tuned **SetFit** (PyTorch) model, routing low-confidence cases to human review; the rules stage reaches **0.979 macro-F1** (8 classes), and CI blocks builds below 0.95.
- Deployed inference in the browser with an int8 **ONNX** build served by **transformers.js**, shrinking the model from 90 to 23 MB, and isolated each user's data with database-enforced **PostgreSQL** row-level security (a non-BYPASSRLS role scoped to the request identity).

jetpack | Parallel gzip-Compatible Compression Engine 2026

Solo Developer | jetpack-compress.vercel.app

- Designed a gzip-compatible parallel compression engine on **JDK 25** that runs one **virtual thread** per block with a bounded in-flight window, reaching **6.4x the throughput** of single-threaded java.util.zip (422 vs 66 MB/s, 3-fork JMH).
- Vectorized Adler-32 by hand with the **Vector API** at 2.8x scalar (4.26 GB/s), matching java.util.zip.Adler32 bit-for-bit, with memory-mapped I/O via the **Foreign Function & Memory API**. Documented the limits: the hand-vectorized Adler-32 still trails the JDK's own intrinsic (14.06 GB/s) and DEFLATE entropy coding is delegated to zlib by design; 72 tests, 68.1% coverage with the SIMD vector package at 98.9%.

Glyph | SIMD-Optimized MNIST Neural Network 2025 – 2026

2-person team | getglyph.vercel.app

- Hand-wrote **SIMD** kernels across 4 instruction sets (**AVX-512, AVX2, NEON, wasm-simd128**) and parallelized the matmul with **OpenMP** for a 3.536x speedup, on an existing 2-layer **C++** MLP that hits 97.01% on the 10,000-image MNIST test set.
- Shipped the Glyph landing page, a live in-browser **WebAssembly** benchmark timing SIMD against scalar on the visitor's device.

Agentic AutoML Platform | Senior Capstone 2025 – 2026

Backend Developer | 2-person team | agentic-automl-platform.vercel.app

- Co-built a **LangGraph** AutoML platform and its backend with 44 agent tools (12 over **MCP**) spanning preprocessing, feature engineering, and training, running LLM-generated Python in a locked-down **Docker** sandbox (non-root, read-only, capped CPU and memory).
- Automated the pipeline from dataset and docs through preprocessing, feature engineering, and training behind human-review gates.

Cadence | Natural Language Calendar & Tasks 2024 – 2026

Solo Developer | usecadenceapp.vercel.app

- Translated plain English into calendar events, tasks, and **Google Meet** invites with a 4-stage NLP parser (chrono, hashtag, priority, compromise), packed into one serverless function of 37 routes to fit Vercel's 12-function cap.
- Enforced multi-tenant isolation in production across 7 tables with **PostgreSQL** row-level security (non-bypassing role, per-transaction identity) and owner-scoped checks on 6 services, proven by an isolation and **IDOR** suite among 1,186 tests (635 frontend, 551 backend), 0 skipped.

TECHNICAL SKILLS

Languages & Web: Python, Java, C++, TypeScript, SQL, HTML, CSS, React, Next.js, Tailwind CSS, Node.js, Express, FastAPI

Machine Learning & AI: PyTorch, scikit-learn, SetFit, sentence-transformers (e5), ONNX, transformers.js, XGBoost, LightGBM, SHAP, LangGraph, MCP, RAG, OpenAI API, Optuna, feature engineering, model evaluation

Data & Infrastructure: PostgreSQL (row-level security), Supabase, pgvector, DuckDB, Polars, Vector API, Foreign Function & Memory API, WebAssembly, Docker, Vercel, Hugging Face Spaces, Google APIs, JMH, Git, GitHub Actions (CI/CD), Linux, pytest, Vitest