Nhoojah Maharjan

nhoojahm@gmail.com | +1 206-910-1893 | Hattiesburg, MS
Portfolio | GitHub | Linkedin

Education

University of Southern Mississippi May 2028
B.S. Computer Science, Minor in Economics & Data Analysis | ***GPA: 4.0*** *Hattiesburg, MS*
- Honors: President's List (All Semesters), Academic Excellence Scholarship
- Coursework: Data Structures & Algorithms, Discrete Mathematics, Probability & Statistics, Secure Software Development

Technical Skills

Languages & Frameworks: Python, TypeScript, Go, C/C++, SQL, Java, React, Next.js, Node.js, FastAPI, LangChain
AI/ML: PyTorch, LLM pipelines, RAG, agentic workflows, Transformers, Scikit-Learn, OpenRouter
DevOps & Cloud: GCP, AWS (ECS, Fargate, S3), Azure, Docker, Terraform, GitHub Actions, OpenTelemetry
Databases: PostgreSQL, MongoDB, Elasticsearch, Redis, Firebase, Supabase

Experience

Research Software Engineer Nov 2025 – Present
Institute for Advanced Analytics and Society (IAAS) | NOAA-funded *Hattiesburg, MS*
- Shipped **AQUAVIEW**, a production Python library with **7+ REST endpoints** for NOAA oceanographic data access, surpassing **1,000+ downloads** as the primary data layer for downstream research pipelines.
- Processed **17M+ records** from 125 years of the World Ocean Database, generating MongoDB metadata indexed via **Elasticsearch** on cloud infrastructure.
- Deployed an **LLM-based tagging and classification engine** standardizing Platform IDs across **20M+ data points** from **10+ heterogeneous NOAA sources**, eliminating manual intervention at scale.
- Developed **LLM-powered agentic workflows** on GCP, enabling NOAA stakeholders to assess environmental data impact through multi-source tool integrations and cloud-hosted inference infrastructure.

Research Assistant Apr 2025 – Present
University of Southern Mississippi *Hattiesburg, MS*
- Published IEEE IoT Journal paper establishing the first comprehensive benchmark of **100+ deep learning architectures** across **15+ datasets** spanning biomedical, infrastructure, and remote sensing domains.
- Stress-tested **vision-language foundation models** under zero-shot, few-shot, and prompt-engineered conditions — quantifying cross-dlevel segmentation precision by **up to 12%** across biomedical and infrastructure benchmarks.
- Formalized a **5-family architectural taxonomy** mapping CNNs, Transformers, GANs, diffusion models, and promptable architectures to segmentation task requirements — adopted as the paper's primary model-selection framework across **3 imaging domains**.

Data Engineering Intern Jan 2026 – Present
Optimal Answers LLC *Gulfport, MS*
- Architected prescriptive analytics models using **Dataform (Optimal Modeling Language)** on **Azure** to solve **Linear and Integer Programming** problems, modeling qualification criteria, solution constraints, and **reduced-cost marginal value** outputs across procurement, portfolio, and supply chain pipelines.
- Modeled large economic datasets from the **Gulf Coast Business Council** and Restoration Fund to optimize grant selection and resource allocation, defining multi-period constraint hierarchies and attribute logic to maximize goal achievement under strict bounds.

Software Developer May 2025 – Aug 2025
Dabuli (Nonprofit, 501(c)(3)) *Seattle, WA*
- Architected a microservices backend in Node.js/Express sustaining **15,000+ req/min** at **sub-100ms** p99, cutting query latency **40%** via Redis caching and compound indexing.
- Instrumented **OpenTelemetry** distributed tracing and automated recovery pipelines, reducing MTTR from hours to **under 10 minutes**.

Projects

Enterprise Data Toolkit (EDT) | *Next.js, Go, ConnectRPC, Protobuf, GCP, Vertex AI, Docker, Terraform*
- Orchestrated long-running AI coding agents in **Daytona sandboxes** via Vertex AI with real-time incremental output over **WebSocket/SSE fallback** and reconnect-safe session handling for durable workspace state.
- Owned Protobuf API contracts and shipped a **ConnectRPC Go microservice** abstracting full sandbox lifecycle behind a typed RPC surface, with **OpenTelemetry** tracing and Terraform-managed GCP infrastructure.

Research.it | *Next.js, FastAPI, Python, WebSockets, Gemini 2.5 Flash, Sentence Transformers, Sigma.js*
- Shipped a multi-stage **LLM pipeline** (extract → normalize → gap-detect → verbalize) with JSON-schema prompting and a regex-first/LLM-fallback citation extractor, cutting parse failures by **∼85%**.
- Rendered real-time citation graphs via **Sigma.js + ForceAtlas2** with sub-second layout on **500+ node graphs**; live progress streamed via **FastAPI WebSockets + uvloop**, cutting perceived latency by **60%**.