

Mithun Venkatesh Gowda

NY, US | (607) 381-0177 | gowdamithunv@gmail.com | linkedin.com/in/mgowdav | github.com/mgowda23 | [Portfolio](#)

PROFESSIONAL SUMMARY

Python & AI Software Engineer with 2.5+ years building scalable, high-throughput backend services (500K+ users, 2,000+ RPS). M.S. in CS (Binghamton) & AWS AI Practitioner specializing in Python backend systems, Generative AI, RAG, and LLMs. Proven delivery in ML and cloud microservices across CodePath and Pfizer, backed by Docker, Kubernetes, and CI/CD

PROFESSIONAL EXPERIENCE

Extern | Pfizer Advanced: AI-Powered Document Insights & Data Extraction **Sep 2026 – Present**

- Building Python pipelines with PyMuPDF and pdfplumber to extract and structure data from pharmaceutical PDFs
- Implementing OCR with Tesseract, PaddleOCR, and EasyOCR to read labels and scanned documents
- Creating RAG pipelines with LlamaIndex, FAISS, and Chroma; comparing Gemini, Mistral, and Phi-2 on document Q&A
- Developing a Gradio chatbot for uploading documents and querying their content in real time

Code Path | AI/ML Engineering Training **Aug 2026 – Present**

- Building hands-on skills in Python, machine learning, and Generative AI/LLM application development
- Developing technical foundations relevant to AI-powered workflow automation and intelligent application development

ITC Infotech Pvt Ltd | Software Engineer | Bangalore, India **Jan 2022 – Jun 2024**

- Developed production Java Spring Boot microservices and REST APIs for loyalty and payment systems for 500,000+ members
- Engineered Node.js and Express REST APIs as a routing layer handling 2,000+ concurrent requests per second at peak
- Built ETL pipelines streaming PostgreSQL and MongoDB data into Snowflake, using Python to validate JSON payment webhooks
- Integrated Razorpay and Pine Labs gateways (UPI webhooks); optimized PostgreSQL B-Tree indexes, cutting search delays 20% for 15,000+ daily users
- Developed React and Angular loyalty dashboards, reducing frontend code duplication by 10% through reusable components
- Maintained Jenkins CI/CD pipelines, integrated the AWS SDK for S3 document retrieval, and followed Git PR workflows

PROJECTS

AI-Powered Analytics & RAG Engine | Independent Project | [GitHub](#) **Aug 2026**

- Built a Python S3-to-Snowflake pipeline with dbt, processing 10M+ orders and 23M+ order items
- Enriched 300,000+ reviews with OpenAI GPT-4o-mini via prompt engineering into structured sentiment and topic fields
- Implemented a RAG chat interface over review embeddings (vector search) and a guarded text-to-SQL assistant via Streamlit
- Orchestrated the load-transform-enrich-analyze pipeline as a daily Airflow DAG running in Docker
- Used Claude Code and Copilot to debug Airflow and dbt failures, validating all AI-suggested changes with unit tests

Cloud Native Portfolio on AWS EKS | Independent Project | [GitHub](#) **Nov 2025**

- Built a React and Vite single-page application, structuring reusable components to optimize client-side rendering
- Containerized the application with multi-stage Docker builds and deployed it onto an AWS EKS Kubernetes cluster
- Provisioned infrastructure with Terraform and built GitHub Actions CI/CD pipelines for linting, testing, and zero-downtime deploys

Hybrid Movie/TV Recommender | Independent Project | [GitHub](#) **Apr 2026**

- Built a Python recommender over 8,800+ titles using TF-IDF vectors and cosine similarity on genres, cast, and descriptions
- Blended similarity with a popularity prior via a tunable weight in an interactive Streamlit app (scikit-learn, pandas)

EDUCATION AND CERTIFICATIONS

Binghamton University, SUNY | M.S in computer science (GPA: 3.67 / 4.00) **May 2026**

Certifications: AWS Certified AI Practitioner | Redis for Python Developers

TECHNICAL SKILLS

Languages: Python, SQL, Java, JavaScript, C++, C; Data Structures and Algorithms

Generative AI and ML: Generative AI/LLMs, RAG, prompt engineering, embeddings, vector DBs (Pinecone, FAISS, Chroma), LlamaIndex, OpenAI API, OCR (Tesseract, PaddleOCR, EasyOCR), PyMuPDF, pdfplumber, Gradio, text-to-SQL, NLP, TensorFlow/Keras, Scikit-learn; Claude Code, Copilot

Cloud and Data: AWS (S3, EKS), Docker, Kubernetes, Terraform, PostgreSQL, MongoDB, Snowflake, Airflow, ETL/ELT

Backend and APIs: Spring Boot, Node.js, Express, Django, FastAPI, REST APIs, JSON, microservices, Kafka, Redis

Testing, CI/CD and Practices: Unit/integration testing (JUnit, Mockito), Git, code review, CI/CD (Jenkins, GitHub Actions), Agile