

Ashutosh Agrawal

(929)-643-3424 | New York, NY | ashh.agrawal11@gmail.com | [LinkedIn](#) | [Portfolio](#) | [GitHub](#)

EDUCATION

New York University Master's in Computer Science (3.9 GPA)	Sep 2024 - May 2026
University of Mumbai Bachelor's in Engineering – Information Technology (3.6 GPA)	Aug 2017 - Jul 2021

SKILLS

Languages: Python, Java, JavaScript, TypeScript, React, Angular

Frameworks & Libraries: Next.js, Redux, Django, FastAPI, LangChain, LangGraph, TensorFlow, PyTorch, Spring Boot

GenAI / ML: Hugging Face, LLMs, LangChain, LangGraph, NLTK, AWS Sagemaker, AWS Strands

Data & Big Data: Pandas, PySpark, Hadoop MapReduce, Spark MLlib, ETL, MySQL, MongoDB, Hive

Cloud & DevOps: AWS, GCP, Kubernetes, Docker, Kafka

PROFESSIONAL EXPERIENCE

New York University - Research Assistant New York, NY	Jan 2025 - Present
---	--------------------

- Designed an **AI trading agent** using **LangChain** and **Llama 3-Instruct** to automate screening, ranking, and backtesting undervalued stocks, integrating **reinforcement learning** for adaptive investment decisions.
- Built **ML/DL models** using Python (TensorFlow, PyTorch, Decision Tree, Random Forest, XGBoost) on large YFinance and AlphaVantage datasets (**60+years data**), achieving **92% prediction accuracy** for investment strategy.
- Tech Stack:** LangChain, Llama 3, Python, TensorFlow, PyTorch, XGBoost, CUDA, LLM, TensorRT

STL Digital - AI Intern Danville, CA	Jun 2025 - Aug 2025
--	---------------------

- Built an **Agentic AI Resume Intelligence System (Llama 3-70B / GPT-4 + LangChain)** to parse and match resumes to JDs, automating screening and **boosting hiring accuracy by 90%**.
- Engineered **LangGraph + FastAPI** multi-agent workflows integrating **FAISS** retrieval and **Sentence-Transformer** embeddings, improving match **accuracy by 35%** and reducing **latency by 50%**.
- Architected a cloud-native **GCP** pipeline (**VMs, CI/CD, Pub/Sub, Docker, TensorRT**) handling **500K+** resumes/day with **continuous pipeline** optimization improving **sub-second latency** by 80%.
- Achieved **87% LLM precision** through automated **A/B validation**, iterative **prompt optimization**, and **vector feedback fine-tuning**, enhancing **candidate matching accuracy** and improving **system reliability**.
- Tech Stack:** Python, FastAPI, LangGraph, Llama 3, FAISS, Sentence Transformers, Angular, GCP, Docker, TensorRT

LTIMindtree Disney - Senior Software Engineer Mumbai, India	Jul 2022 - Jul 2024
---	---------------------

- Redesigned **AWS Cloud** modules by **containerizing microservices with Docker** and optimizing **ETL** pipelines, achieving a **70% reduction** in workflow complexity and **99.99% platform uptime**.
- Resolved critical service bottlenecks by engineering modular **REST API** connectors and event-driven synchronization between **Python** orchestration modules and **AWS S3**, boosting data **scalability by 40%**.
- Automated** region-specific audit and authorization workflows using **Python schedulers** and **MySQL triggers**, scaling to **2M+ users** across **50+ regions** and managing a **50B+ record database**.
- Led a **5 developers** team to enhance the **AWS S3 document retrieval system**, reducing retrieval time by **70%** through collaborative **problem-solving, code reviews and sprint planning**.
- Tech Stack:** Python, AWS (S3, Lambda), MySQL, Docker, REST APIs, Redis, ETL Pipelines

LTIMindtree CitiBank - Senior Software Engineer Mumbai, India	Jul 2021 - Jul 2022
---	---------------------

- Resolved **500+ L3 issues** in Citibank's GCP systems using **FAST APIs, Python automation, and SQL**, improving **performance** and **reliability** across **50M+ daily transactions**.
- Architected **React, Python, and SQL** proof-of-concepts with **BAs and product owners** to resolve critical production bottlenecks and **streamline** enterprise workflows.
- Tech Stack:** Python, Java, FAST APIs, GCP (Pub/Sub, Cloud SQL, Load Balancer), Kubernetes, Microservices

PROJECTS

CodeSensei – Autonomous Code Reviewer | LangGraph, Claude/GPT-4, GitHub API, OpenAI Embeddings

- Developed CodeSensei, an autonomous code-review agent using GPT-4 and AST parsing to detect logic flaws and propose optimized patches, reducing manual review effort by 55%.

Yuka Food & Cosmetic Scanner | Python, LangChain, OpenAI GPT-4, FAISS, REST APIs

- Built an Agentic AI layer over Yuka using RAG, vector search, and GPT-4 to autonomously analyze ingredients and suggest safer alternatives, reducing manual lookups by 65%.

CERTIFICATIONS

- Oracle Cloud Infrastructure 2025 Certified Data Science Professional
- Oracle Cloud Infrastructure 2025 Certified Generative AI Professional