

MALLIKARJUN AITHA

New York Metropolitan Area | mallikarjunaiitha@gmail.com | Phone: +1-203-737-0149
mallikarjunaiitha.com · linkedin/mallikarjun-aitha · github.com/mallikarjun

PROFESSIONAL SUMMARY

Data professional with 3+ years of experience in **data analysis, data engineering geospatial analysis, LLMs and cloud-based ETL pipelines** using **Spark, AWS, and Azure**. Proven expertise in optimizing **data workflows, deploying scalable analytics solutions**, and driving business insights through cross-functional collaboration. Adept at **building automated pipelines, developing dashboards**, and applying **machine learning** to real-world datasets in both industry and research setting.

SKILLS & COMPETENCIES

- LLMs & Generative AI: LangChain, HuggingFace, FAISS
- ETL Pipelines: Python, SQL, AWS Lambda, Spark, Kafka
- Geospatial & GIS Analysis: ArcGIS, DBSCAN, Hot Spot Analysis, KDE
- Cloud Platforms: AWS, Azure, GCP
- Data Engineering & Automation: Docker, Git, Cron Jobs, Shell
- Machine Learning: ARIMA, Random Forest, scikit-learn, PyTorch, R programming
- Business Intelligence & Performance Metrics: Power BI, Tableau, JavaScript
- CRM & Campaign Automation: AEM, React.js, JSON, CI/CD
- Data Visualization & Reporting: Power BI, Tableau, ArcGIS API, Matplotlib
- Statistical Analysis & Forecasting: Regression, Correlation, Bayesian Methods
- Advanced Excel: Pivot Tables, VLOOKUP, Macros
- Open-Source Models: Llama, Mistral, Phi, Falcon, Gemma
- Cross-Functional Collaboration

PROFESSIONAL EXPERIENCE

Research Assistant- *University of New Haven, West Haven, CT | Oct 2024 – Present*

- Built automated **ETL pipelines** for the **New Haven Police Department** using **Python, AWS Lambda**, and **PostgreSQL** to ingest and process daily crime reports from multiple sources, reducing manual overhead by 15 hours/month
- Conducted **crime analysis** for **NHPD** using **ArcGIS, Python**, and **SQL**, leveraging clustering models like **DBSCAN** and **Hot Spot Analysis**. Identified high-crime areas and improved resource allocation by 20%.
- Created 10+ **interactive dashboards** using **Power BI** and **JavaScript (ArcGIS API)**, visualizing NHPD **crime hotspots** and time series forecasts with **ARIMA** and **Random Forest models** for proactive policing.

Student Associate- *Sodexo, West Haven, CT | Jan 2024 – May 2025*

- Managed inventory and financial performance for a team serving 500+ students daily.
- Utilized **data visualization tools** to track **financial KPIs** and improve **operational efficiency**.
- Analyzed **budget allocations** to reduce waste and increase **cost efficiency**.

Data Intern- *Downtown Evening Soup Kitchen, New Haven, CT | July 2024 – Dec 2024*

- Designed **ETL workflows** to consolidate donor data, improving accuracy by 30% and operational efficiency.
- Built a **predictive analytics** pipeline using **Python, SQL, and Power BI** to forecast donation trends, manage inventory better which reduced food waste by 25%
- Optimized 5+ **Power BI** dashboards to track **KPIs** like meal distribution efficiency and volunteer retention, integrating **PantrySoft, SQL, and Python** for **real-time insights**, improving **decision-making** by 30%.

Associate Data Engineer- *Cognizant Technology Solutions, Hyderabad, India | Sept 2021 – July 2023*

- Engineered **data pipelines** using **Python** and **SQL** to process **10M+ user logs** across 15+ countries, driving **A/B testing** insights that boosted campaign CTR by 18%.
- Developed and deployed **125+ AMEX campaigns** with **AEM, JSON, React.js, and CI/CD**, automating **online enrollment** and reducing reliance on offline promotions.
- **Automated global web operations**—including **data validation, UI updates, and content deployment**—using **Python, Shell, Git, Docker, and cloud CRM tools**; streamlined **200+ service requests** and multilingual site updates across 15+ countries

EDUCATION

Master's in Data Science (M.S), University of New Haven-West Haven, CT | GPA:3.86/4.0
Relevant Courses: **Data Analytics, Power BI, NLP, Deep Learning, AI, Data Engineering**

Bachelors in Electronics and Communication Engineering (B.Tech), Jawaharlal Nehru Technological University-Hyderabad, India | GPA: 7.43/10.00
Relevant Courses: **Advanced Mathematics, Data Analytics, Digital Communications**

CERTIFICATIONS

- LLM Applications with Prompt Engineering– NVIDIA Deep Learning Institute (2025)
- Power BI & Dashboarding – University of New Haven (2024)
- ArcGIS Professional – ESRI Academy (2025)

TECHNICAL PROJECTS

- **RAG-Based Mental Health Chatbot (LangChain, HuggingFace, FAISS)**
Developed Retrieval-Augmented Generation (RAG) based chatbot using LangChain and open-source LLMs such as **Mistral, Llama, Phi** and **Falcon**. Evaluated chatbot accuracy using coherence and **perplexity** metrics.
- **Real-Time Smart City Monitoring (AWS, Spark, Kafka)**
Designed a streaming pipeline to process 10K+ daily IoT data points; reduced latency by 40% and visualized traffic insights in **Power BI**.
- **Incept-Face (Facial Recognition System)**
Developed a facial recognition pipeline using **InceptionResNetV1** and **MTCNN**; deployed with **Gradio** for real-time inference.