Yahya Masri

Junior CS Student at George Mason University

yahya.masri@yahoo.com · GitHub · LinkedIn · Google Scholar · ORCID

Education

George Mason University — Fairfax, VA
B.S. in Computer Science — 2025 - Present

Northern Virginia Community College — Annandale, VA
A.S. in Computer Science — 2023 - 2025

Publications

Automating Map-Making through Enhanced Geographic Information Extraction
Currently Drafting, 2025
Using Retrieval-Augmented Generation with Open-source Large Language Models

Comparative Analysis of BERT and GPT for Classifying Conflict News Preparing to Submit, 2025
With Sudan Conflict as an Example

Optimizing context-based location extraction by tuning open-source LLMs with RAG Under Review, 2025
IJDE International Journal of Digital Earth

Work Experience

Undergraduate Research Assistant — 2024 - Present
NSF Spatiotemporal Innovation Center, George Mason University

- Developed and evaluated ML pipelines using BERT fine-tuning and GPT with RAG, ICL, and hyperparameter tuning to classify conflict incidents, achieving 91% accuracy in single-label classification
- Extracted conflict locations from news articles using RAG, CoT, and few-shot LLM prompting, achieving $>$ 90% F1 in context-aware information retrieval
- Developed end-to-end pipelines integrating large language models, QGIS, Neo4j, and Streamlit for automated conflict mapping and analysis
- Installed and maintained Linux-based operating systems (Red Hat 8, Rocky Linux, Ubuntu) across 7 rack-mounted server units to support high-availability computational workflows
- Designed and implemented a 3D data center digital twin using NVIDIA Omniverse, visualizing real-time server metrics for intuitive infrastructure monitoring

Research Projects

Data Center Digital Twin Development — Summer 2025

- Developing real-time monitoring and visualization systems using Prometheus and Grafana
- Implementing comprehensive data center performance analysis tools

- Creating intuitive infrastructure monitoring solutions

Multi-Agent Conflict Mapping System Summer 2025

- Automating conflict mapping via multi-agent system using LLMs, Neo4j, and PyQGIS
- Developing automated extraction and visualization of conflict-related data
- Implementing real-time data processing and analysis pipelines

ML Pipeline Development Spring 2025

- Contributed to machine learning pipelines using BERT, GPT, FAISS, and OpenAI APIs
- Implemented advanced NLP techniques for data analysis
- Developed classification systems with high accuracy rates

Location Extraction Research Fall 2024

- Researched context-specific location extraction for conflict articles
- Implemented Retrieval-Augmented Generation techniques
- Developed and optimized information extraction algorithms

Technical Skills

Programming Languages: Python, JavaScript/TypeScript, SQL
Technologies & Tools: Neo4j, QGIS, Streamlit, Large Language Models, Prometheus, Grafana, NVIDIA Omniverse
Areas of Expertise: Machine Learning, Geographic Information Systems, Data Visualization, System Administration

Research Interests

- Large Language Models and Natural Language Processing
- Inference Optimization and Prompt Engineering
- Geographic Information Systems and Spatial Analysis
- Data Center Management and Monitoring

Personal Projects

BlueTemp

- AI platform supporting marine environments by predicting sea water temperatures
- Applications in hurricane forecasting and coral reef preservation

Crushor

- Retro-style platform-based game with advanced mechanics
- Implemented wall-jumping, dashing, and stomping features