Abdelrahman Elewah

Oshawa, Ontario, Canada :abdelrahman.elewah@gmail.com 289 939 6665 elewah.github.io
abdelrahman-elewah elewah Google Scholar

Summary

Artificial Intelligence (AI) and Machine Learning Engineer (ML) with over 5 years of experience developing scalable AI and IoT systems. Expertise in LLMs, RAG, and real-time IoT data frameworks. Proven success in designing and implementing innovative solutions that connect AI with practical applications. Skilled in software development, cloud-based platforms, e-commerce systems, and data analytics. Experienced in conducting research and adept at collaborating with cross-functional teams to achieve impactful results.

Technologies

LLM Frameworks & Tools: LangChain - LangGraph - LangSmith - RAG

Programming Languages: Python - C/C++ - JavaScript

Web Technologies: REST APIs - React - HTML - CSS - Bootstrap

Automation Tools: Jenkins - GitLab CI/CD - GitHub Actions Workflow

Cloud & Deployment: AWS EC2 - AWS App Runner - Elastic Container Service - Elastic Beanstalk - CI/CD Pipelines

Database Systems: PostgreSQL - MongoDB - MySQL - NoSQL - SQL - Spark - Hadoop

DevOps Tools: Docker - Docker Compose - Kubernetes - Dev Containers

Development Environments: GitHub - GitLab - VS Code - Anaconda

Experience

Instructional Specialist (part-time), 2U / University of Toronto — Remote, Jan 2023 – Apr 2025

- Contributed to the success of the **University of Toronto** 's online **Data Analytics Boot Camp**, supporting 100+ learners in mastering practical skills for data-driven careers.
- Facilitated hands-on workshops in **Python, Database, Machine Learning, and Data Visualization**, resulting in a **15% increase** in student satisfaction scores.
- Supported the deployment of **real-world capstone projects**, helping learners apply techniques in domains such as **healthcare, HR, and finance**.

Graduate Research Assistant, Ontario Tech University — Oshawa, ON, Jan 2020 – Apr 2025

- Designed the **SensorsConnect** framework, enabling real-time **IoT device interoperability** and scalable data exchange, inspired by the principles of the **World Wide Web**. Publication
- Developed an **Agentic IoT Search Engine (ASE-IoT)** that integrates **LLMs**, **RAG**, and autonomous agents to enable **natural language querying** of live IoT data, enhancing search precision and user interaction. publication
- Co-led an **OVIN-funded project** to develop **autonomous vehicle curriculums** and conduct applied research on **software-defined vehicles (SDVs)** and **digital twins**, contributing to workforce development in emerging mobility technologies.
- Collaborated with **Eagle Aerospace** to prototype an **Aircraft Deceleration Early Warning System**, enhancing runway safety through predictive analytics and early alert mechanisms.

Co-Founder(part-time) Tamra-IoT, — Toronto, ON, May 2019 – Jan 2024

- **Architected** secure and scalable **IoT platforms** by integrating **MQTT over TLS**, **cloud infrastructure**, and **mobile control interfaces**, enhancing real-time communication and remote device management.
- **Collaborated** on **business management** and strategic planning, contributing to key decisions that optimized resource allocation, improved product direction, supported **mentorship**, and accelerated go-to-market execution.
- **Developed a curriculum** to teach **IoT concepts** to high school students, promoting early STEM engagement and empowering the next generation with practical, hands-on IoT experience.
- **Designed** and deployed **Over-The-Air (OTA) firmware update mechanisms** and implemented **robust IoT device management systems**, significantly reducing maintenance costs and improving system resilience. Publication

Education

PhD **Ontario Tech University**, Electrical and Computer Engineering — Jan 2020 – Mar 2025

- GPA: 4.22/4.3 Link to Transcript issued by Ontario Tech University
- **Coursework:** Real-Time Data For IoT, Communication Networks, Knowledge Discovery & Data Mining, Data Visualizations
- **Thesis**: SensorsConnect: World Wide Web for Internet of Things.

MSc **Benha University**, Electrical Engineering — Feb 2013 – Jan 2018

BSc **Benha University**, Electrical Engineering — Sep 2008 – Jun 2012

Projects

Localelive: Agentic Search Engine for Real-Time IoT Data Live Demo — github.com/repo

- **Developed** a real-time IoT search engine powered by **LLMs and RAG**, enabling users to query complex sensor data using natural language, improving query efficiency and decision making.
- **Implemented** a semantic search pipeline using **Sentence-BERT and HNSW indexing**, reducing query latency by 73% and enhancing relevance in top-k retrieval across diverse IoT datasets.
- **Managed** over 37,000 real-time IoT documents from 500+ service types in **MongoDB** with geo-indexing, ensuring scalable and location-aware data access for time-sensitive decision-making.
- **Achieved** 92% top-1 accuracy in complex intent detection and information retrieval, surpassing systems like Gemini, and significantly improving user satisfaction and task completion rates.
- **Applied** in real-time urban scenarios—such as locating least-crowded clinics, nearest available parking, and lowest gas prices—demonstrating direct utility for smart city applications.
- **Technologies:** Leveraged LangGraph, Tavily API, OpenRouteService, VectorDB, and Sentence-BERT to build a modular and extensible architecture for dynamic IoT data exploration and retrieval.
- **Deployed** the system in **AWS** using **Docker Compose and Traefik** Traefik live dashboard, enabling seamless container orchestration, automated HTTPS provisioning, and scalable reverse proxy management for reliable production-grade deployment.

Story-to-Movie Recommender Chatbot (RAG-based) Live demo — github.com/repo

- Developed a **retrieval-augmented generation (RAG)** system combining vector search with LLMs to deliver context-aware answers, reducing hallucination rates by **30%**.
- Built a **semantic search pipeline**, enabling retrievals from 1K+ documents.
- Fine-tuned user prompts and applied **prompt chaining** techniques to improve answer relevance, validated through user feedback and precision metrics.
- Leveraged **Pandas**, **OpenAI API**, and **Tenacity** to ensure resilient API usage and robust data handling under real-time loads.

Apply Lightweight Fine-Tuning to a Foundation Model — github.com/repo

- **Built** an end-to-end NLP pipeline using **PyTorch** and **Hugging Face Transformers**: loaded a pre-trained **GPT-2** model and prepared the **AG News** dataset for news-topic classification.
- **Applied** parameter-efficient fine-tuning (PEFT) using **LoRA adapters** to fine-tune GPT-2 while keeping the base model's weights frozen, **reducing training time and memory usage by over 60% compared to full fine-tuning**.
- **Achieved** a significant improvement: **boosted accuracy from 83.16% to 88.95%** on the AG News dataset using LoRA-fine-tuning, **demonstrating the effectiveness of PEFT in enhancing model performance with minimal compute**.

RadViz-Plotly — github.com/repo

- **Developed RadViz-Plotly**, an **open-source Python package** that generates **2D and 3D Radial Visualization (RadViz)** plots for **high-dimensional datasets**, enabling broader accessibility to dimensionality reduction techniques in research and industry. publication
- **Enabled data scientists** and analysts to explore and interpret **complex data distributions** interactively using **Plotly**, significantly improving **model explainability** and **decision-making** in analytics workflows.
- **Facilitated** deeper insights into **high-dimensional data** by revealing **hidden clusters, outliers, and trends**, increasing user engagement through intuitive visual interfaces.