

Abdulrahman Elattar

AI & Data Engineer

abdodev@tuta.io | linkedin.com/in/abdulrahman-elattar-6b83102a5

github.com/Abdulrahman532 | kaggle.com/abdulrahmanex | Egypt

TECHNICAL SKILLS

Certifications	AWS Certified Cloud Practitioner; AWS Academy Cloud Foundations; DEPI AWS Cloud Solution Admin & Architect; NTI NLP Engineer; EPIC Applied GenAI Program; Samsung Innovation Campus (Big Data Track)
Languages	Python; SQL; NoSQL; C++; C
ML / NLP	Scikit-Learn; TensorFlow; PyTorch; LangChain; FAISS; ChromaDB; RAG; HyDE; MCP (FastMCP); Prompt Engineering; TF-IDF; NumPy; Pandas; Matplotlib
Backend & Apps	FastAPI; Streamlit; Gradio; BeautifulSoup4
Data Engineering	Apache Spark (PySpark, Structured Streaming); Airflow; dbt; Snowflake; Kafka; Databricks; NiFi; Hive; Flink; Hadoop (HDFS); HBase; Amazon Redshift
Databases	MongoDB; Redis; SQL
DevOps / Cloud	AWS (S3, Lambda, VPC, CloudWatch, AWS CLI); Docker; Git / GitHub; Linux

EXPERIENCE

Big Data Track Trainee | Samsung Innovation Campus August 2025 — November 2025

- Completed the Big Data track of Samsung Innovation Campus, engineering scalable ETL pipelines using PySpark to process, transform, and join large-scale e-commerce datasets (orders, customers, products, logistics).
- Modeled the combined data into a consolidated, analysis-ready fact table, preparing it for business intelligence reporting and visualization in Power BI.

NLP Engineer Trainee | NTI July 2025 — August 2025

- Developed end-to-end NLP applications, including a RAG chatbot (LangChain, FAISS, LLMs) and a movie recommender (TF-IDF, Cosine Similarity).
- Deployed services via FastAPI backends and built interactive user interfaces with Streamlit and Gradio.

Cloud Solution Architect Trainee | DEPI April 2024 — October 2024

- Automated AWS infrastructure deployment using AWS CLI scripting, focusing on secure networking solutions based on project requirements.
- Architected and configured a VPC Peering connection, including Route Tables and Security Groups, to enable private cross-VPC communication, monitored via CloudWatch.

PROJECTS

ChefBot — Domain-Specific Cooking Assistant (RAG + MCP)

- Built a RAG pipeline over 21 recipe documents (~73 chunks) using ChromaDB and HyDE (Hypothetical Document Embeddings) retrieval, with Llama3.2:3b for generation and prompt engineering for grounded answers.
- Developed an MCP server (FastMCP) exposing 3 tools (ingredient substitution, nutrition estimation, semantic recipe search) and 1 resource, usable from Claude Desktop.
- Evaluated on a 10-query test set: 100% top-5 retrieval precision and 4.7/5 average answer relevance; HyDE improved relevance on abstract queries by about 1 point.

NYC Taxi Data Pipeline (Batch + Real-Time) — Team Project

- Co-built an ELT pipeline for NYC Yellow Taxi data: Airflow-orchestrated ingestion from AWS S3 (Lambda) into Snowflake, with dbt models in a medallion architecture (staging, intermediate, marts) and custom data-quality tests.
- Built a real-time layer with FastAPI, Kafka, Spark Streaming, and Redis that flags fraudulent trips using 15+ rules, with a live Streamlit dashboard; containerized with Docker.

Medical Assistant Chatbot with RAG — Team Project

- Co-built a FastAPI backend with a RAG pipeline (LangChain + FAISS) that retrieves context from medical PDFs and CSVs, then generates answers using Llama 3 via Together AI.
- Co-developed a Streamlit frontend with persistent chat history and a pytesseract OCR module to extract text from uploaded medical prescriptions.

E-commerce ETL Pipeline (PySpark) — Team Project

- Collaborated in a team to build a PySpark ETL pipeline that joins multiple raw CSV sources (orders, customers, products, logistics) into a single clean fact table (fact_orders_logistics.csv) ready for Power BI reporting.

EDUCATION

BSc in Artificial Intelligence | Kafr El-Sheikh University | GPA: 3.66 / 4.0 October 2022 — July 2026

EPIC Diploma – Applied GenAI Program | EPIC Institute of Technology September 2025 — September 2026

- Two-semester program covering advanced programming languages, algorithms and data structures, data science, math and statistics, real-time backend, distributed systems architecture, and applied generative AI; completed a final project building an AI-powered solution.

EXTRACURRICULAR

Dev Days | GitHub Community Event (Participant) October 2026

- Attended a hands-on workshop on prompt engineering and AI-assisted development with GitHub Copilot.