Mohamed Atwan

DATA SCIENTIST | NLP SPECIALIST

Egypt, Cairo | +20 155 120 9560 | Mo7amed3twan@gmail.com

LinkedIn : linkedin.com/in/mo7amed3twan
Kaggle : kaggle.com/mo7amed3twan
GITHUP : github.com/MO7AMED3TWAN

Summary "We Can Change the World Through the Data"
_____

Results-driven Data Scientist with **2 years** of experience in **Machine Learning(ML), Deep Learning(DL), Natural Language Processing(NLP), and Data Science**, skilled in **extracting insights** and **building data-driven solutions** with <u>Focusing Now</u> In **The AGENTIC AI & LLMS.**

Proven track record of <u>Leading a Team to a top 5</u> **in MTC-AIC2 AI Competition** by leveraging **Deep Learning** and **advanced modeling techniques in Speech Recognition Systems.** And <u>Leading Other Team to Top 8</u> **in Huawei Spark Infinity North Africa 24 Competition** by leveraging **AI Techniques In Our Supply Chain Solution** That We Introduced it.

NOW Actively engaged in **community knowledge-sharing** as **Huawei Cloud Developers Group (HCDG) Organizer & Huawei Cloud Student Developer (HCSD) Ambassador**, with a focus on fostering **collaboration** and **innovation** in **data science, AI and Cloud Computing Areas.**

Education
_____

Bachelor's Degree of Science in Artificial Intelligence, **Kafr Elsheikh University**   *2021 - 2025*

Specialized coursework in Machine Learning, Deep Learning, Data Science, Cloud Computing

Experience
_____

Shabab Mobtakeron (Under Supervised Of Ministry of Youth & Sports)

AI Instructor (Part - Time)   **October** *2024 – April 2025*

- **Deliver interactive lectures on AI and deep learning concepts.**
- **Monitor students' progress through assignments and projects.**
- **Supervise and guide students in their graduation projects.**
- **Track and assess students' performance to provide feedback.**
- **Assist students in troubleshooting technical issues in their projects.**

Shabab Mobtakeron (Under Supervised Of Ministry of Youth & Sports)

Technical Support & Organizer (Part - Time)   **October** *2024 – April 2025*

- **Provide technical support during events, activities, and workflow to ensure smooth operations.**
- **Organize events and training sessions while coordinating trainees' schedules efficiently.**
- **Assist participants in troubleshooting and resolving technical issues for a seamless experience.**
- **Improve the training environment by managing logistical and organizational tasks.**

Information Technology Institute (ITI)

Business Intelligence Development Trainee (Full - Time)   *July 2024 – September 2024*

- **Gain experience in data warehousing, advanced SQL, Power BI, and Tableau.**
- **Design and optimize data warehouses for efficient management.**
- **Create an interactive dashboards to provide actionable insights.**
- **Applied business intelligence tools {PowerBI-SSMS-SSIS-SSRS} to solve business challenges.**

Digital Egypt Pioneers Initiative (DEPI)

Data Engineering Trainee (Full - Time)   *July 2024 – October 2024*

- **Master SQL for data management and optimization.**
- **Design scalable data warehouses for efficient storage and retrieval.**
- **Gain hands-on experience in Python for data processing and automation.**
- **Applied machine learning concepts and MLOps tools like MLflow and Hugging Face.**

Volunteering Locally
_____

IEEE Kafr Elsheikh Branch

Head of Natural Language Processing   (Part - Time)   *October 2024 – February 2025*

- **Defined project roadmaps, ensuring team expertise in RNNs, LSTMs, GRUs.**
- **Mentored team members, designed training materials, and assessed skills in Machine learning, Deep learning and NLP.**
- **Collaborated with Students to build scalable pipelines for large-scale text processing (Sentiment Analysis Solution Both Arabic/English).**

Microsoft Learn Student Ambassadors, KFS Branch
Database Administrator (Part - Time)   *April 2023 - September 2024*

- **Manage the club's database infrastructure, ensuring optimal performance and data integrity.**
- **Design and optimize schemas for scalability and efficiency.**
- **Implement monitoring solutions to ensure continuous system availability.**
- **Collaborate with other teams to deliver tailored database solutions and troubleshooting support.**

Microsoft Learn Student Ambassadors, KFS Branch
Head of Data Engineering (Part - Time)   *September 2024 – March 2025*

- **Developed and executed the strategic roadmap for the Data Engineering track.**
- **Led and mentored a team of student mentors, ensuring effective execution of learning objectives.**
- **Fostered a collaborative, educational environment and provided ongoing guidance and support.**

Volunteering Globally
_____

Huawei Cloud Developer Group (HCDG) Organizer &
Huawei Cloud Student Developer (HCSD) Ambassador   *December 2024 – Present*

I Actively work to **Promote Huawei Developer Certifications and other educational programs across universities**. in addition to **Organize and participate in tech events** such as hackathons, workshops, and seminars. and **Build and manage a community of students** interested in **cloud computing**, **AI**, **and other emerging technologies.**

Microsoft Learn Student Ambassadors
"Alpha" then "Beta" Ambassador   *September 2024 – May 2025*

I Actively work to **empower students with advanced technology skills**. I <u>focus on areas such as</u> **Data Science**, **Artificial Intelligence**, and **Cloud Computing**, with a commitment to **sharing knowledge** through **workshops**, **events**, and **collaborative projects**. My role allows me to collaborate with fellow students to **leverage technology to solve real-world problems**, while continuously learning and growing alongside a global community of innovators.

Projects
_____

Comprehensive Hadith Chatbot System (My Graduation Project)   *September 2024 - june 2025*
Role: Team Leader

- **Led the development** of an **advanced chatbot system** that delivers authenticated hadiths, explanations, and actionable lessons, empowering users to apply prophetic teachings in their daily lives.
- **Curated and preprocessed a custom dataset** of **authenticated hadiths** for **model training and knowledge retrieval**.
- Fine-tuned large language models (LLMs) for **text generation and optimized embedding models** for semantic understanding using **Python, Hugging Face, and PEFT (LoRA)**.
- **Developed a Retrieval-Augmented Generation (RAG)** system with **LangChain and FAISS** to ensure accurate and contextually relevant responses.
- **Deployed the chatbot on web and Android platforms** using **Docker** and **Azure**, ensuring accessibility for a global audience.

Automatic Arabic Speech Recognition (MTC-AIC2 – 5th Place)   *May 2024 - September 2024*
Role: Team Leader

Developed a Deep Learning-based Speech-to-Text system using DeepSpeech 2 (NVIDIA) for Modern Standard Arabic & Egyptian dialects in real-world noisy environments, And we Ranked 5th in the MTC-AIC2 competition, competing against top AI teams.

- Designed a robust **Automatic Speech Recognition (ASR) system** capable of handling **noisy speech data** from **movies, street conversations, and non-standard Arabic dialects**.
- Developed a **DeepSpeech 2-based architecture** using **CNNs & Combination (GRU & LSTM)** for **high-accuracy transcription**.
- Applied **advanced signal processing**: **noise reduction, silence removal, and data augmentation** to enhance recognition accuracy.
- Built **custom preprocessing pipelines** for transcript cleaning, feature extraction (**MFCCs, Spectrograms**), and model training.
- Overcame computational limitations by **leveraging cloud computing (Colab Pro)** for large-scale model training.

Additionally, there is more work available on my GitHub, where shared several projects related to
Data Science | Machine Learning & Deep Learning | Natural Language Processing (NLP)

Certificates & Licenses
_____

Huawei Certified ICT Associate – Artificial Intelligence (HCIA-AI)
Issued Mar 2025 · Expires Mar 2028   Credential ID: 01010200185581002629I409

HCCDA-AI – Huawei Cloud Certification
Issued Mar 2025 · Expires Mar 2028   Credential ID : HWENDCAIDA658987

HCCDA-Tech Essentials – Huawei Cloud Certification

Issued Jan 2025 · Expires Jan 2028   Credential ID: HWDCTEDA364117

Certified AI Fundamentals – DataCamp

Issued Aug 2024   Credential ID : N/A

Machine Learning Specilization – DeepLearning.AI

Issued Apr 2024   Credential ID : XPB5KNX3QT2C

Courses

_____

Python Programming Basics – Information Technology Institute (ITI)	*May - 2021*
Applied Deep learning – Information Technology Institute (ITI)	*May - 2021*
Database Fundamentals – Information Technology Institute (ITI)	*September - 2021*
AWS Machine Learning Foundation – Udacity	*October - 2022*
Data Analysis With Excel And PowerBI – Integrated Training Academy (ITA)	*October - 2023*
Assessment for Data Analysis and Visualization Foundations – IBM	*October - 2023*
Introduction to Data Analytics – IBM	*October - 2023*
Excel Basics for Data Analysis – IBM	*October – 2023*
Data Visualization and Dashboards with Excel and Cognos – IBM	*October – 2023*
Python for Data Science, AI & Development – IBM	*December - 2023*
Machine Learning Specialization – Stanford University & DeepLeaning.AI	*April – 2024*

Skills & abilities

_____

Machine Learning & Deep Learning Modeling

- **Traditional ML:** Supervised (Regression, Classification) & Unsupervised (Clustering, Dimensionality Reduction) techniques.
- **Deep Learning:** CNN (Computer Vision), RNN/LSTM (Time Series), Transformers (NLP), GANs (Generative Models).
- **Transfer Learning:** Experience with fine-tuning and adapting pre-trained models (e.g., ResNet, EfficientNet) for domain-specific tasks.

Natural Language Processing (NLP)

- **LLM Fine-Tuning:** Falcon, JAIS, BERT, T5, GPT-2, QWEN-2.5/3, Phi, DeepSeek, Mistral, Llama 2/3.
- **Training Techniques:**
  - **Efficient Training:** PEFT, LoRA, QLoRA, RLHF, SFT.
  - **Frameworks:** Hugging Face Transformers, LLaMA-Factory, Unsloth, Lightning-AI, WandB, VLLM.
- **RAG & Agents:** LangChain, LangGraph, CrewAI, MCP and A2A.
- **NLP Tasks:** Text summarization, translation, sentiment analysis, NER, chatbots, and AI Agencies.

Programming & Data Management

- **Python:** Advanced EDA (Pandas, NumPy, Matplotlib/Seaborn), PyTorch, TensorFlow, Scikit-learn, Hugging Face. analysis, manipulation, machine learning, and deep learning tasks.
- **Databases:** SQL Server, PostgreSQL, MySQL, FAISS, Qdrant, Milvus (vector DBs for RAG applications).

Cloud Computing & MLOps

- **Azure:** Data Factory, Synapse Analytics, ML for orchestration, transformation, and model deployment.
- **Huawei Cloud:** Managed cloud-based ML workflows and data processing.
- **Automation Tools: N8N**, **Together AI** for workflow automation and scalable AI solutions.

Version Control & Deployment

Proficient in **Git**, **GitHub**, and **Docker** for **version control**, **code management,** and **containerization**

Data Pipeline Development & Data Analysis

- **ETL/ELT:** SSIS, SSRS, SSAS for data integration and reporting.
- **BI & Visualization:** Power BI, Excel for dashboards and actionable insights.

Other Tech Skills

Terminal/CLI | Code Quality | Agile/CI-CD | Security & Best Practices

Soft Skills

Leadership | Emotional Intelligence | Communication | Team Working | Problem Solving | Adaptability & Flexibility Mentorship & Coaching | Persuasion & Influence | Critical Thinking | Storytelling

Languages

English: B2 (EF-TEST - 4 Skills – 90MIN Test) | German: A1 (Beginner)