

Mostafa Saeed

Position: Ph.D. Student in NLP at MBZUAI

Location: Abu Dhabi, UAE * *Telephone:* (+971) 50 740 8568

E-mail: mostafa.saeed@mbzuai.ac.ae

[LinkedIn](#) * [Google Scholar](#)

Education

Ph.D. in NLP

Mohamed bin Zayed University of Artificial Intelligence (MBZUAI)

Abu Dhabi, UAE

2026 – Present

B.Sc. Computer Science

MSA University and University of Greenwich

Cairo, Egypt and London, UK

2018 – 2022

Work Experience

Research Assistant – Full Time

New York University Abu Dhabi – CAMEL Lab

Abu Dhabi, UAE

July 2024 – August 2026

- Conducted research on Arabic morphological analysis and morphosyntactic tagging, with a focus on enhancing the morphosyntactic tagging components of **CAMEL Tools**.
- Reformulated Arabic lemmatization as a classification task across multiple genres, and developed sequence-to-sequence lemmatization models for dialectal Arabic.
- Led the development and release of **MAQAM**, the most richly annotated resource for Quranic Arabic, covering annotations for 22 morphosyntactic features.

NLP & ML Developer – Contract Based

Ministry of Communications and Information Technology

Cairo, Egypt

May 2023 – July 2024

- Developed NLP and LLM-based solutions for government digital services, including a Retrieval-Augmented Generation (RAG) system and automated analysis of user complaints.

Teaching and Research Assistant – Full Time

MSA University

Cairo, Egypt

September 2022 – July 2024

- Served as a teaching assistant for Artificial Intelligence, Machine Learning, Natural Language Processing, and Probability & Statistics courses.
- Co-supervised undergraduate graduation projects alongside the main faculty supervisor.

Academic Researcher – Freelancer

Islamic University of Madinah

Remote – Madinah, KSA

August 2022 – July 2024

- Contributed to two research papers on automatic short answer grading for Arabic.

Data Scientist – Freelancer

Upwork

Remote – North Carolina, United States

December 2019 – June 2024

- Performed data analysis and built predictive models for live stock market prediction.
- Designed and deployed RAG systems for multiple client companies across diverse use cases, integrating different database backends.

1. **Mostafa Saeed**, Mohammad Amjad, Mahmoud Kassem, Rawan Bondok, Mayar Nassar, Adel Mahmoud Wizani, Salman Elgamal, Mahmoud Reda, Nizar Habash. 2026. MAQAM: A Richly Annotated Corpus for Quranic Arabic Morphological Analysis and Disambiguation. *In Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing, Budapest, Hungary. Association for Computational Linguistics.*
2. Saeed Alzaabi, Salam Khalifa, **Mostafa Saeed**, Shahd Dibas, Hanan Aldarmaki, Nizar Habash. 2026. Gahwa: A Computational Lexicon for Emirati Arabic. *In Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing, Budapest, Hungary. Association for Computational Linguistics.*
3. Khalid N. Elmadani, **Mostafa Saeed**, Hanada Taha, Nizar Habash. 2026. A Benchmark for Arabic Lemma-Level Readability Assessment. *In Proceedings of the 2026 Conference on Empirical Methods in Natural Language Processing, Budapest, Hungary. Association for Computational Linguistics.*
4. **Mostafa Saeed** and Nizar Habash. 2025. [Lemmatization as a Classification Task: Results from Arabic across Multiple Genres](#). *In Proceedings of the 2025 Conference on Empirical Methods in Natural Language Processing, Suzhou, China. Association for Computational Linguistics.*
5. **Mostafa Saeed**, Nizar Habash. 2025. [Lemmatizing Dialectal Arabic with Sequence-to-Sequence Models](#). *In Proceedings of The Third Arabic Natural Language Processing Conference, Suzhou, China. Association for Computational Linguistics.*
6. **Mostafa Saeed**, Rana Waly, Abdelaziz Ashraf Hussein. 2025. [AMAR at BAREC Shared Task 2025: Arabic Meta-learner for Assessing Readability](#). *In Proceedings of The Third Arabic Natural Language Processing Conference: Shared Tasks, Suzhou, China. Association for Computational Linguistics.*
7. Emad Nabil, **Mostafa Saeed**, Rana Waly, Safiullah Faizullah, Wael Hassan Gomaa. 2025. [Leveraging Large Language Models for Arabic Short Answer Grading and Feedback Generation](#). *International Journal of Innovations in Engineering and Technology (IJIET).*
8. Wael H. Gomaa, Abdelrahman Ezzeldin, **Mostafa Saeed**, Abdulmohsen Algarni and Emad Nabil. 2023. [Empowering Short Answer Grading: Integrating Transformer-Based Embeddings and Bi-LSTM Network](#). *Big Data and Cognitive Computing (BDCC).*
9. Abdelrahman Ezzeldin, **Mostafa Saeed**, Shereen El-Feky and Ali Khater. 2022. [Hyper-parameters Optimization of Deep Convolutional Neural Network for Detecting COVID-19 Using Differential Evolution](#). *In Decision Sciences for COVID-19: Learning Through Case Studies. Cham: Springer International Publishing.*