Omar Farag Ahmed Ismail Hamed

New Cairo, Egypt   (+20)1091161886

in omar-farag-ismail   𝕏 Omar Farag   ⚙ omarfarag95   ✉ omar16farag16@gmail.com

PUBLICATIONS

- **Hamed, O., Bakkali, S., Blaschko, M., Moens, S., Van Landeghem, J. (2024). Multimodal Adaptive Inference for Document Image Classification with Anytime Early Exiting. ICDAR 2024. [link](#)**

- **M. S. Darweesh, M. Adel, A. Anwar, O. Farag, A. Kotb, M. Adel, A. Tawfik, and H. Mostafa, "Early Breast Cancer Diagnostics Based on Hierarchical Machine Learning Classification for Mammography Images," *Cogent Engineering*, vol. 8, no. 1, 2021. [link](#)**

- **M. Adel, A. Kotb, O. Farag, S. M. Darwish, and H. Mostafa, "Breast Cancer Diagnosis Using Image Processing and Machine Learning for Elastography Images" in IEEE MOCAST 2019, Greece. [link](#)**

- **Farag, Omar, et al. "Integrated Sensors for Early Breast Cancer Diagnostics." 2018 IEEE 21st International Symposium on Design and Diagnostics of Electronic Circuits & Systems (DDECS). [link](#)**

EXPERIENCE

Specright   Tustin, California, United States
Applied Scientist - Full Time / Remote   Sep 2025 - present

- Document Intelligence and Understanding

Microsoft   Cairo, Egypt
Data and Applied Scientist II - Full Time   Nov 2022 – Aug 2025

- **Team**: Microsoft AI - Bing Search - Index Quality and Spam team
- **Core skills**: NLP, LLMs, Information Retrieval, Explainability, Finetuning, Distributed Training, Quantization
- **Core contributions:** Owned the 1.2 trillion Microsoft spam pipelines powering the training of Microsoft next-generation foundation models.
  * Implemented post-training quantization using frameworks like VLLM, BnB, and GPTQ to optimize deep learning models for production pipelines, enabling hardware-accelerated inference.
  * Designed and implemented distributed training and inference pipelines for SLMs using multi-GPU architectures.
- Developed an LLM-based judge system to evaluate for spam detection across the full web-index.
  * Enabled Microsoft Index Quailty team to replace crowdsourced spam data labeling.
- Fine-tuned unimodal/multimodal SLMs for spam detection and web content filtering.
- Built and deployed deep learning models at web scale for spam detection.
  * Worked on model explainability using Shapley values for deep and tree-based models.
- Worked on evaluating VLMs for visual/GUI grounding for agentic solutions
- Collaborated with the ranking and relevance team to integrate spam signals into ranking

Zai   Cairo, Egypt
Cofounder & Founding Engineer / Remote   Sep 2024 - Present

- Zai is reimagining knowledge discovery in the Arab world by combining real experiences with advanced NLP, LLMs, and dynamic knowledge graphs to deliver instant, context-aware answers

Netcall (Contract.fit)   Brussels, Belgium
Grad. NLP Researcher - Internship   Feb 2022 - Jul 2022

- Researching and developing techniques for optimizing Transformers (size/latency) for production settings.

* Experience with knowledge distillation, pruning, quantization techniques and embeddings compression.
  - Investigating new *embeddings* initialization approaches for custom domain vocabulary finetuning.

Botit   Cairo, Egypt
NLP Engineer - Full Time / Remote   *Apr 2019 – Oct 2021*

- Developing Natural Language Understanding (NLU) pipeline for a conversational e-commerce chatbot
- Developing **Synthetic Datasets** (English/Arabic) to simulate real user behaviour
- Developing NLP algorithms including **NER, Intent Classification, Spelling Correction**
- Researching and Developing approaches for e-commerce Recommendation System
- Applying **Text Mining** techniques for attributes extraction
- Building **information retrieval and ranking** algorithms for items retrieval.

Opto-Nano-Electronics (ONE) research lab   Cairo University
Machine learning Researcher   *Aug 2018 – Oct 2018*

- Developing machine learning algorithms for breast cancer classification (including all data pre-processing steps, data visualization and classification)

EDUCATION

KU Leuven   Leuven, Belgium
Masters of Science in Artificial Intelligence   *Sep. 2021 – Sep. 2022*
  * **Specialization**: Speech and Language Technology
  * **Thesis**: Optimizing Transformers for Model Serving
  * **Grade**: Distinction (Honours)

German university in Cairo (GUC)   Cairo, Egypt
Bachelor of Science in Electronics Engineering   *Oct. 2013 – Jul. 2018*
  * **Thesis Grade**: Excellent
  * **Thesis**: Integrated Sensors for Early Breast Cancer Diagnostics
  * **Overall Grade**: Very Good
  * **Activities and Societies**: The Insider GUC, Nawwar GUC

LANGUAGES AND SOFT SKILLS

Languages

  - **Arabic:** Native   - **English:** Fluent - IELTS : 7.5   - **German:** Beginner - B1   - **Dutch:** Beginner.

Soft Skills

  - **Self Learning Skills**   - **Presentational Skills**   - **Time Management Skills**.

TECHNOLOGIES AND PROGRAMMING SKILLS

- **Programming Languages**: Python, JAVA, C#
- **Machine/Deep Learning frameworks**: Pytorch, TensorFlow, Keras, Sklearn, VLLM
- **Cloud Services:**: Azure ML, Azure Singularity, Azure ADF, AWS Sagemaker

INDEPENDENT COURSE WORK

- **MITx Datascience Micromasters** *(WIP)*   - **MITx**