Ahmed Khaled

Data Scientist | Machine Learning Engineer

📞 +358-465376682/ ✉ mail 💼 ahmedkhaledsaleh ⭘ Ahmed-Khaled-Saleh 🖥 BLOG

Education

University of Ottawa May 2022 – Feb. 2023

Master of Engineering, Electrical and Computer Engineering *Ontario, Canada*

Suez Canal University Sept. 2016 – Aug. 2020

Bachelor of Computer Science *Ismailia, Egypt*

Experience

University of Oulu - Center for Wireless Communication Sept 2023 – Present

Research Assistant

- Research assistant in the Intelligent Connectivity and Networks/Systems (ICON) Group working under the supervision of Dr. Mehdi Bennis.
- Topics include Large Language models, Federated optimization, Federated Edge LLMs, and personalized Federated learning models.

RDP - Revolution Data Platforms - Canada-based Aug 2023 – Present

Data Scientist *Remote*

- NLP engineer working on the frontiers of LLMs and interacting with customers to optimize their strategy on how to leverage the latest revolution in generative AI.
- Leaveragees Azure OpenAI services, Azure ML stack, and Azure Cognitive services to build an end-to-end solution.

AUC - The American University in Cairo April 2023 – Sept 2023

Research Assistant - School of Science and Engineering

- A joint research project sponsored by the AUC and University of Oulu, Finland.
- Research focuses on the intersection of 6G-related technologies, Machine learning, and Federated Learning.

DEBI - Digital Egypt's Builder Initiative Feb 2023 – Present

Teaching Assistant

- Teaching assistant in the financial technology track of the University of Sains Malaysia.
- Courses taught include machine learning, data analysis, statistics, and Python scientific computing stack.

DEBI - Digital Egypt's Builder Initiative Feb 2022 – Feb 2023

Machine Learning Engineer

- Developed an NLP-based system for quotation extraction and attribution that will be utilized internally in the Ministry of Communication and Information Technologies in Egypt .
- The system is utilized and deployed as an internal tool in the MCIT of Egypt.

Suez Canal University August 2019 – Oct. 2019

Python Developer Intern

- Assisted in creating and automating some researchers' work using Python and scientific tools like pandas and NumPy.
- Worked with various user interface libraries like TKinter and PyQT to deliver desktop applications.

Projects

MIRA-Q | *Python, PyTorch, Transformers, Spacy, MongoDB, Pinecone, Streamlit*

- Developed a system to extract and attribute quotes from major news websites like NewYorkTimes, Guardian, etc...
- Used MongoDB to store the scraped text as documents and Pinecone as a vector database to store the embedding of the quotes.
- created the quotation extraction pipeline using CofeNet neural architecture. Adjusted the original code of the authors to create an inference module that can be directly used inside the project.
- Built the system's user interface using Streamlit and integrated it with the two databases utilized.

Consensus ratio in recorded meetings | *Python, PyTorch, Transformers, Librosa*

- A system to analyze meetings in an end-to-end automated way and provide a detailed comparison of the consensus ratio of each speaker in each topic discussed in the meeting.
- Natural language inference (NLI) along with semantic textual similarities (STS) techniques are employed. A few-shot topic modeling is used to define the topic in inference time.

- Built a speaker diarization module to segment the recording by a speaker at each time frame and pass the segmented audio to a Speech-to-text module. Nvidia Nemo Toolkit is used and OpenAI's Whisper is utilized as a speech-to-text module.

Buildings counter from satellite images | *Python, PyTorch, OpenCV, Pillow*
- The project aims to provide a solution for counting the buildings in crowded areas in Egypt using satellite images.
- A custom data set is built and transfer learning is utilized to apply a pretrained model to the data.
- Context-aware convolution neural network (CACN) is used as the generator in a GAN Min-Max game.

TLS-encrypted Flow Packet-length Sequences Analysis | *Python, PyTorch, scikit-learn*
- The project tries to build a classification system that assigns a label to a malicious attack such that every attack belongs to a specific type or family of attacks (a total of 8 types of attacks).
- A Bi-GRU with Attention is applied for obtaining a representation of the sequence. This architecture wasn't used to address this specific problem (TLS traffic classification) before.
- A newly-released dataset is used as part of the cybersecurity mining competition.

Skills and Tools

Skills: Machine learning, Deep learning, NLP, Data Science, Computer Vision,
AI for Cybersecurity, Cloud Analytics.
Languages: Python, R, C, C++, JavaScript, SQL.
ML Tools: PyTorch, TensorFlow, Scikit-learn.
BigData: Hadoop, Spark, MongoDB.
DevOps tools: Docker, Linux, Git, Github, Gitlab.

Certificates

AWS Certified Machine Learning Specialist   March 2022

- Machine learning specialist in AWS stack related to Data Analytics, Machine Learning, NLP, and big data.
- Topics covered include services like SageMaker, Transcribe, Textract, Personalize, Forecast, Rekognition, Polly, Lex.

AWS Certified Cloud practitioner   May 2022

- Cloud practitioner certification aims to build essential building blocks for anyone utilizing AWS as a cloud infrastructure.
- Topics covered include Cloud front, Fargate, EC2, AWS lambda, CloudWatch, AutoScaling, Route53, Direct Connect.

Leadership / Extracurricular

Google Developer Group - Ismailia branch   Aug 2018 − Present
Educator, Speaker, Volunteer
- * Conducting mini-talks, activities, and sessions ranging from machine learning to tools like Keras and Tensorflow.
- * Led the TOT (Training of trainers) program that aims to build machine learning knowledge for other members to conduct educational sessions.

Langauges

Arabic: Native speaker.
English: Fluent.