KHALED EL-SAKA

WORK EXPERIENCE

Machine Learning Engineer
Ogoul | Jan 2023 – Present

- **Innovations:**
  - Fine-tuned NLLB200 across 98 language directions to improve translation quality.
  - Extending translation capabilities from 9 to 200+ languages by updating the Helsinki models to NLLB200.
  - Improved translation quality significantly , including:
    - Increased En-Ar BLEU score from 11 to 63 and other pairs.
    - Enhanced Egyptian dialect to En from 23 to 36 BLEU.
  - Introduced emoji translation support with an Aligner model.
  - Applied manual pruning (model surgery) for Marian models.
  - Enhanced video violence classification models (safe recall from 0.48 to 0.94, violent recall from 0.82 to 0.98) and improved image violence classification accuracy (94% to 98%) using Vision Transformers (VIT).
- **AI System Development:**
  - Developed multiple chatbots, including a web-search chatbot, a Retrieval-Augmented Generation (RAG) bot, and other helper bots using Langchain, LangGraph, and Llama 3.1.
  - Created an evaluation pipeline using LLM-as-a-judge for assessing Chatbots outputs.
  - Developed a Generative AI system supporting 12 API endpoints for tasks involving text, image, video, audio, music, and multimodal requests.
  - Developed text translation project supporting 200+ languages.
  - Developed real-time transcription and translation systems for speech-to-text tasks using Whisper.
- **Optimization & Tools**:
  - Eliminated web search costs by deploying the open-source SearxNG search engine locally with DuckDuckGo as a backup.
  - Achieved 2x latency reduction and doubled requests/sec rate for LLM endpoints by replacing Ollama with VLLM.
  - Improved model inference speed (4x on CPU, 5x on GPU), loading time (11x faster), and memory efficiency (8x) through CTranslate2 and int8 quantization.
  - Built an efficient YouTube Shorts crawler using RapidAPI, multithreading, Redis Queue, multiprocessing, R2 Cloudflare, and MySQL, with optimized job handling to remove redundant processes.
  - Created a simple GUI for video annotation after splitting bloody movie content into unique scenes.

Internship
Microsoft | Aug – Dec 2023
- Automated detection of new words and meanings in Egyptian Arabic, tackling challenges like distinguishing new words from misspellings.
- Proposed and implemented a novel solution leveraging MARBERT models, clustering, and an autocorrect system.
- Experimented with training BERT model and tokenizer from scratch for Egyptian Arabic.
- Built a Flask web application and applied analytic visualization using Tableau.

SUMMARY

Self-motivated Machine Learning Engineer with extensive experience leading and independently executing diverse AI projects, including NLP, computer vision, and translation systems. Skilled in designing scalable solutions, optimizing performance, and making critical technical decisions. Passionate about tackling challenging projects that push boundaries and elevate expertise. Seeking a senior role to drive innovation and deliver impactful AI solutions.

EDUCATION

Master of Electrical and Computer Engineering – Ottawa University:

May 2022 – Jan 2023   **Major**: AI and Data Science.
Projects:

- **Chinese sign language Image Classification [link]:**
  Keywords: (Image processing, Feature Engineering, ResNet, Zernike Moments, Data Augmentation)
- **Static Vs Dynamic, Cybersecurity attack detection [link]:**
  Keywords: (Preprocessing, Analysis, Feature Selection, Modeling, Kafka, Docker, Static Vs Dynamic Model)
- **Healthcare Chat Bot [link]:**
  Keywords: (Preprocessing, DialogFlow, NLP, Classification, Clustering, Weighted Association Rules, Evaluation, Error analysis)
- **Bitcoin price prediction using Sentiment Analysis [link]:**
  Keywords: (Data collection and preprocessing, Sentiment Analysis, Ml, LSTM, Evaluation, Error analysis)
- **Fake task detection [link]:**
  Keywords: (CGAN, RF, AdaBoost)
- **Intrusion Detection [link]:**
  Keywords: (Dimensionality reduction vs Feature selection, Deep Neural Network, HP Tuning, Flow Chart, TSNE Visualization, Tensorflow)
- **Sentiment Analysis for Arabic tweets [link]:**
  Keywords: (Data preprocessing, BERT, araBERT, Fasttext, araXLENT)

Faculty of Computers and Artificial Intelligence – Cairo University:
2015 – July 2019
Major: Computer Science.   **Minor**: Information System.
Graduation project: **School bus-tracking system [link]:**
Keywords:(Android, ReactJS, Google Maps API, Firebase, NodeJs, Express, Socket I.O., TypeORM, MySql)

PROFESSIONAL TRAINING

(Feb 2022–Jan 2023) Digital Egypt Builders Initiative (DEBI):
- (Feb – May) 240 hours of Berlitz English conversation course [link]
- (Feb – May) 200 of Dale Carnegie training certificate [link].
- (Apr – May) IBM Big Data [link].
- (Mar – Apr) IBM Predictive Analytics Modeler [link].
- (Feb – Mar) Huawei HCIA-AI v3 [link].

CONTACTS

+20 114 882 7429

Giza, Egypt

khaled.elsaka25@gmail.com [link]

khaled-el-saka-962700161[link]

Sa2a [link]

kaggle khaledelsaka [link]

ONGOING DEVELOPMENT

Currently enhancing expertise in MLOps, advanced system design, and scalable AI deployments through Practical MLOps, Designing Machine Learning Systems, and self-driven projects.

TECHNICAL SKILLS

Programming languages (Python, Java…etc.), NLP, Computer Vision, Generative AI, Model Fine-tuning, Translation Systems, PyTorch, TensorFlow, Model Optimization, FastAPI, and Flask.