Ammar Khairi

+44 737 750 6528 | @ ammarnasraza@gmail.com | LinkedIn | GitHub | Portfolio | London, UK

Education

The University Of Edinburgh — Edinburgh, UK
M.Sc. in Artificial Intelligence; GPA: 7.5/10 — *Sep 2022 – Sep 2023*

- Master's Thesis: Efficient Fine-Tuning of Large Language Models for Code Generation.
- Investigated efficient fine-tuning of LLMs on limited compute for code generation across 4 programming languages
- Reduced training costs by 70% while maintaining performance through LoRA fine-tuning and smaller LLMs.
- Published and presented work at NeurIPS Black in AI Workshop; Code and models open-sourced on HuggingFace

University of Khartoum — Khartoum, Sudan
B.Sc. in Electrical and Electronics Engineering; GPA: 8.64/10 — *Sep 2014 – Jun 2020*

- BSc Thesis:Text to Image Synthesis from Text Semantics using Attentional Generative Adversarial Networks
- Proposed incorporating sentence-level semantics to enhance attention mechanism in text-to-image synthesis
- Improved Inception Score by +4.13% and Fréchet Inception Distance by +13.93% on CUB birds dataset
- Published and presented work at International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE)

Work Experience

Clingendael - the Netherlands Institute of International Relations — Remote
NLP Consultant — *Oct 2023*

- Working as an NLP consultant on a research project analyzing polarization in Twitter discussions related to the conflict in Sudan.
- Manually annotated part of the Arabic dataset for stance detection, sentiment analysis and bot identification.
- Collaborated with senior researchers to design a pipeline for stance classification and bot detection using BERT models fine-tuned on our annotated dataset.
- Applying state-of-the-art NLP techniques like transformers and transfer learning to enhance ML models for low-resource Arabic text.

School of Informatics, The University of Edinburgh — Edinburgh, UK
Teaching Assistant — *Jan 2023 – June 2023*

- Served as a TA for the Computer Graphics course, instructing over 100 students.
- Designed and implemented 3 programming assignments on core graphics algorithms including raytracing, meshes and optimization.
- Hosted lab sessions to mentor students in C++ and OpenGL.
- Achieved an overall student satisfaction rating of 4.8/5.0 in course evaluations.

AmunData, Data-Driven Research, and Advisory — Khartoum, Sudan
Data Scientist — *July 2021 – September 2022*

- Led the development of an open-source precision agriculture web application using Python, Streamlit, and Azure to provide crop monitoring and insights to clients. The application leveraged Sentinel-2 satellite imagery and vegetation indices to enable low-cost, customized field monitoring.
- Created automated workflows to process and store raw Sentinel-2 data by extracting vegetation insights like NDVI, LAI then storing raw processed imagery in Azure Blob storage.
- Designed and implemented a crop classification pipeline using Random Forest and Sentinel-2 multispectral bands. Achieved 89.22% accuracy in categorizing key crops in the state of El Gezira. Analysis provided area estimates for strategic planning on food security.
- Developed predictive yield models by combining satellite data with ground truth data from surveys. Models were used by clients like FGM International to monitor crop health and compare seed performance.
- Led migration of the company's crop monitoring platform from a licensed solution to open source, which reduced running costs by 85% and enabled greater customization.
- Managed interns, provided technical guidance, wrote project proposals, and maintained existing data solutions. Contributed to various agricultural projects using remote sensing and ML.

CODE Software Sudan, Khartoum
Full Stack Developer *Dec 2020 – Jun 2021*

- Led migration of monolithic Code Management System (CMS) from Java to Golang microservices architecture, enabling 100% reduction in downtime during upgrades and integration of new services.
- Implemented concurrent modules using Goroutines in Go which increased frequency of health checks for critical services by 2x.
- Enhanced CMS frontend built with NodeJS by adding mentor assignment and email alerts for faster issue resolution, significantly lowering client complaints.
- Migrated Sudatel's Direct Carrier Billing from Java to GoLang, allowing more flexible concurrent implementations and increasing revenue by 10% for their Anghami subscription service.
- Contributed as a full stack developer on specialized solutions encompassing systems integration, web development, process management, and mobile development.

Projects

Bridging The Data Gap: Estimating Cultivated Lands Using Satellite-based Machine Learning | *GitHub*
- Led end-to-end development of a methodology to estimate cultivated agricultural lands in Sudan using Sentinel-2 satellite imagery and XGBoost classification.
- Solely managed large-scale data collection, preprocessing, and labeling of over 100GB of satellite imagery data.
- Developed XGBoost classifier optimized through hyperparameter tuning, achieving 92-99% accuracy in distinguishing cultivated vs uncultivated croplands.
- Generated interactive visualizations and maps to present yearly cultivated land insights at state and scheme levels.
- Authored technical report summarizing methodology, results, and potential applications.
- Presented analysis estimating 40-60% decline in cultivated lands in 2023, providing data-driven insights into Sudan conflict's impact on food security.

Multi-Lingual Image Captioning Using Pretrained Models | *GitHub*
- Developed an end-to-end pipeline for multilingual image captioning in English, German and Arabic using CLIP, GPT-2 and a transformer adapter architecture.
- Fine-tuned language-specific GPT-2 models and trained adapter networks achieving BLEU scores of 15.11 for English, 11.78 for German and 14.17 for Arabic.
- Our approach matches or exceeds performance of prior work on small datasets while requiring less data and compute than large scale models.
- Performed comprehensive quantitative benchmarking using BLEU, ROUGE, CIDEr and qualitative evaluation.
- Open-sourced code, models and datasets to support further research.

Knowledge-Based Question Answering System | *GitHub*
- Developed a knowledge-based QA system enabling businesses to build chatbots matching user questions to similar pre-defined questions/answers.
- Led deployment on AWS as serverless API integrating Lambda and a simple frontend, allowing clients to easily retrain adapters by providing their own CSV data.
- Utilized state-of-the-art BERT NLU model and fine-tuned it for a semantic similarity task using efficient LoRA parameter tuning.
- Reduced operational costs by 90% compared to previous versions, making the system affordable for SMEs.
- Integrated solution into EnigmaAI's platform improving accessibility of AI-powered chatbots.
- Awarded $20,000 grant to onboard service onto SingularityNET marketplace, increasing accessibility.

Publications

- **Training with Single Class Images and Generalizing for Multi-Class Images using Kasami Orthogonal Classification Layer.** | *Paper*
  Mohamed Saadeldin, Ammar Nasr, Amel AbdElraheem, Arjun Pakrashi, Brian MacNamee.
  Northern Lights Deep Learning Conference (NLDL) 2022.

- **SemGAN: Text to Image Synthesis from Text Semantics using Attentional GANs** | *Paper*
  Ammar Nasr, Ruba Mutasim, and Hiba Imam.
  International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), 2021.

- **Modeling, Visualization, and Analysis of African Innovation Performance.** | *Paper*
  Muhammad Omer, Moayad El-Amin, Ammar Nasr, Rami Ahmed.
  International Conference on Learning Representation (ICLR), 2020.

Skills

Programming Languages: C, C++, Java, Python, GoLang, JavaScript, MySQL

Machine Learning Frameworks: PyTorch, SciKit-Learn, HuggingFace Transformers

Cloud Computing: AWS SageMaker, Azure Storage

GIS and Remote Sensing: ArcGIS, Google Earth Engine, Geopandas, Folium

Version Control and Tools: Git/GitHub, Docker, LaTeX, Blender

Interpersonal Skills: Technical writing, Project Management, Mentoring

Languages: English (Professional), Arabic (Native)