Besher Alkurdi

NLP Research Engineer

 [LinkedIn] [icon] mrbesher Google Scholar @besheralkurdi

 mrbesheralkurdi@outlook.com Jyväskylä, Finland (Relocatable)

Experience

Research Engineer *at* Huawei   Aug 2023 - Aug 2024
- Enhanced custom in-house distributed training library with parameter-efficient methods, context expansion, and training stabilization techniques, achieving 8x resource reduction without performance degradation.
- Conducted customer requirements analysis sessions, architected ML solutions for business needs, and defined PoC implementation roadmaps with cross-functional teams.
- Coded LLM architectures enabling compatibility with standard libraries, delivering significant inference speedup.
- Built scalable multi-backend inference platform with RESTful/streaming APIs, supporting scalable deployment and advanced sampling methods.
- Wrote multilingual synthetic data generation toolkit and fine-tuning pipelines for customer LLM adaptations.
- Ran ZeRO/FSDP/DeepSpeed multi-node training, tested models and supplied teams with improved data pipelines.

Undergraduate Research Fellow *at* Yildiz Technical University   June 2022 - March 2023
- Curriculum learning research in *publicly funded project*, implementing cyclical and difficulty-based training strategies.
- Developed data augmentation pipeline and experimented with auxiliary classifiers for early exit strategies.
- Deployed distributed training systems using TensorFlow, PyTorch, and Horovod on lab computers.
- Trained CLIP-style multimodal models with contrastive learning.

Research Intern *at* New Mind AI   July 2021 - October 2021
- Conducted research on state-of-the-art natural language processing techniques and their application in a system assisting lawyers in lawsuits focusing on Document AI applications in the Turkish language.

Education

MSc. in Aritifical Intelligence   Thesis Phase
University of Jyväskylä
BSc. in Computer Science   June 2023
Yildiz Technical University   *2nd in class, GPA: 3.84/4*

Publications

Scaling Up Paraphrase Generation Datasets with MT and Semantic Similarity Filtering   2024
- An extension of our prior work. Experimented with data augmentation using paraphrasing.

Semantic Similarity Based Filtering for Turkish Paraphrase Dataset Creation   December 2022
- A method for creating paraphrase datasets from parallel corpora along the largest paraphrase datasets in Turkish.

Linguistic-based Data Augmentation Approach for Offensive Language Detection   2022
- Augment classification datasets with a linguistic-based approach improving the detection of offensive language.

Skills

ML/DL: PyTorch, TensorFlow, HuggingFace, OpenCV, NLP, TRAX, ZeRO, FSDP, DeepSpeed

Programming: Python, C, Rust, SQL, Java, Julia

Tools: Linux, Git/GitHub, ROS, Docker

Languages

Arabic: Native • **English**: Advanced • **Turkish**: Advanced • **Finnish**: Beginner • **French**: Beginner

Certificates

Deep Reinforcement Learning Course | *HuggingFace*   May 2024
- Q-Learning, PPO.

Generative Adversarial Networks (GANs) Specialization | *DeepLearning.AI*   July 2023
- GAN components, data augmentation, privacy.

Natural Language Processing Specialization | *DeepLearning.AI*   June 2023
- NLP, Transformers, Attention models, Machine Translation.

Deep Learning Specialization | *DeepLearning.AI*   February 2023
- Deep learning, including GANs and VAEs.

Machine Learning Specialization | *DeepLearning.AI*   August 2022
- ML fundamentals and practical applications.

Introduction to Computational Thinking and Data Science | *edX*   May 2021
- Introductory course in data science and ML.

Verification: All certificates are verifiable at *https://www.linkedin.com/in/mrbesher*.

Volunteer Experience

Volunteer Researcher *at* COSMOS AI Research Group   December 2024 – Present

Conducting research on improving the performance of large language and vision-language models using unlabeled data and reinforcement learning without reliance on stronger teacher supervision.

Tutor *at* Schoolhouse.world   February 28, 2021 - May 20, 2021

Tutoring students worldwide in subjects such as Math, Computer Science, and more.

Tutor *at* Aqsa Youth Institution   November 30, 2020 - August 31, 2021

Taught computer, networking, and programming basics to middle school and high school refugee students.

Translator *at* Ask Me   June 30, 2019 - August 31, 2019

Assisted tourists visiting Istanbul as part of Fatih Municipality's "Voluntary Tourism Delegates" program.