

DR. HASSAN ISMAIL FAWAZ

Senior Machine Learning Engineer | PhD in Artificial Intelligence

✉ hassanismaillfawaz@gmail.com

📍 Saudi Arabia

🔗 hfawaz

🎓 publications



EXPERIENCES

Senior Machine Learning Engineer

Beyond Limits

📅 Feb 2024 – Present

📍 Al Khobar, Saudi Arabia

- Scaling up and **load balancing** Large Language Models (LLMs)
- **Project Lead** at the client's site for a large oil and gas operator
- **Data pipelines** for **completion** and **instruction tuning**
- Maintain and develop LLMs in production (**GenAI** / LLMOps)
- Benchmark and evaluate various **vector databases** (e.g. OpenSearch, Weaviate, Qdrant) for information retrieval
- Run various experiments for advanced **Retrieval Augmented Generation (RAG)** systems
- Fine tune **embedding** models on private data using **contrastive** learning approaches
- Evaluate various techniques for fine tuning LLMs such as **QLORA** with **Mistral**, **LLama**, **Gemma**

Senior Machine Learning Engineer

Ericsson

📅 May 2022 – Jan 2024

📍 Paris, France

- **Data Architect**: Scaling **MLOps** framework for deep learning
 - Build on top of managed Elastic **K8s** in **AWS**
 - Setup **CI/CD** pipelines for infrastructure as code (**IAAC**)
 - Implement **experiment tracking** and serving using **MLFlow**
 - Ensure **reproducibility** via **Docker** and **Kubeflow**
- **Tech Lead**: In-house research **project** on **transfer learning**
 - Setup **best practices** for **Deep Learning** experimentation
 - Code review & **mentoring** of PhD students & data scientists
 - Implement state of the art models for **domain adaptation**
 - **Scale up Deep Learning** experiments on **Kubernetes**
 - Define the research project **roadmap** & team responsibilities
 - Lead a monthly **reading group** on Artificial Intelligence
 - Animate training **workshops** for software best practices
- **Machine Learning Engineer**: developing a microservice
 - Follow **software engineering** best practices
 - Review code, unit & integration tests, **CI/CD**
 - Interact with **Kafka** based communication system
 - Develop **ML Ops** & model **Life Cycle Management (LCM)**
 - Implement and investigate various design patterns
- **Data Engineer**: develop big data processing pipelines
 - **Spark** for extracting & normalizing **parquet** data
 - **Clustering** & **Forecasting** of time series

Machine Learning Engineer

Besedo

📅 Oct 2020 – April 2022

📍 Paris, France

EDUCATION

Université Haute-Alsace

PhD in Machine Learning

📅 Sep 2017 – Sep 2020

📍 Mulhouse, France

- Advancing **Deep Learning** for **Time Series Classification (TSC)** ([link](#))
- Contributing to **Keras.io** with a tutorial on Convolutional Neural Networks for TSC ([link](#))
- Open sourcing the **dl-4-tsc** framework with **1k+** GitHub ⭐ stars ([link](#))
- Publishing in top conferences and journals while reaching **#1 of all time** in DMKD ([link](#))
- Student travel **award** for **IEEE Big Data** 2018 conference at Seattle ([link](#))
- **International** collaborations: Monash University (**AUS**), Alan Turing Institute (**UK**), The Open University (**NL**)
- Teaching a **Deep Learning** course for masters students at ENSISA engineering school

Université de Bourgogne

Masters in Databases

📅 Sep 2016 – Sep 2017

📍 Dijon, France

- SQL, NoSQL, Graph
- Data Mining, Semantic Web
- Linear & Constraints Programming

Université Antonine

Masters in Software Engineering

📅 Sep 2011 – Sep 2016

📍 Beirut, Lebanon

- JAVA, C#, Multi-threading, Web Services
- Networking, CCNA Routers & Switches

DATA SCIENCE SKILLS

Computer Vision, NLP & LLMs

Torch, Pandas & HuggingFace

Hydra, & MLFlow

Algebra & Probability

●●●●●●

●●●●●●

●●●●●●

●●●●●●

DEVELOPMENT SKILLS

Object Oriented Programming

Software Design Patterns

Algorithms & Data Structures

Python, SQL, NoSQL

Time & Memory Complexity

MapReduce, Java, C++

●●●●●●

●●●●●●

●●●●●●

●●●●●●

●●●●●●

●●●●●●

- Content moderation using Machine Learning
- Benchmark latest **Computer Vision Deep Learning** models
- **Guide & mentor** junior data scientists and linguistics
- Evaluate **NLP transformer** models for **text classification**
- Develop **microservices** for live computer vision inference
- **Reduce latency** using quantization, pruning, JIT & tensorRT
- **Monitor model performance** and decay (e.g. accuracy)
- Evaluate concept & data **drift**, covariate & label **shift**
- **Orchestrate** and automate model training using Airflow
- Contribute to **HuggingFace's** open source datasets ([link](#))
- Deliver **proof-of-concepts** for ML models using **Streamlit**

Visiting Machine Learning Researcher

Monash University

📅 Nov 2019 – Dec 2019 📍 Melbourne, Australia

Classifying satellite image time series.

Deep Learning Lecturer

Université Haute-Alsace

📅 Sep 2018 – 2021 📍 Mulhouse, France

Giving an advanced course on deep neural networks.

Software Developer Intern

Orange Labs

📅 Mar 2017 – Sep 2017 📍 Nice, France

Java recommendation system based on **RDF triplet data**

Freelance Web Development

MradMCC

📅 Mar 2016 – Aug 2016 📍 Beirut, Lebanon

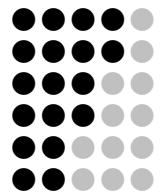
Creating a WordPress website that can be found [here](#).

JOURNAL PUBLICATIONS

- Pialla, G. et al. (2023). "Time series adversarial attacks: an investigation of smooth perturbations and defense approaches". In: *International Journal of Data Science and Analytics*.
- Ismail Fawaz, Hassan, Benjamin Lucas, et al. (2020). "InceptionTime: Finding AlexNet for Time Series Classification". In: *Data Mining and Knowledge Discovery*.
- Ismail Fawaz, Hassan, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, et al. (2019a). "Accurate and interpretable evaluation of surgical skills from kinematic data using fully convolutional neural networks". In: *International Journal of Computer Assisted Radiology and Surgery* 14, pp. 1611–1617.
- – (2019e). "Deep learning for time series classification: a review". In: *Data Mining and Knowledge Discovery* 33, pp. 917–963.
- Forestier, Germain et al. (2018). "Surgical motion analysis using discriminative interpretable patterns". In: *Artificial Intelligence in Medicine* 81, pp. 3–11.

OPERATION SKILLS

Git, Code Review, Unit Tests
 Docker, Kubernetes
 GitLab CI/CD Pipeline
 High Perf. Computing (HPC)
 Amazon Web Services (AWS)
 REST, gRPC, Message Brokers



CONFERENCE PAPERS

- Ismail-Fawaz, A. et al. (2023). "ShapeDBA: Generating Effective Time Series Prototypes using ShapeDTW Barycenter Averaging". In: *ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data*.
- Pialla, Gautier et al. (2022). "Smooth Perturbations for Time Series Adversarial Attacks". In: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*.
- Mathis, Florian, Hassan Ismail Fawaz, and Mohamed Khamis (2020). "Knowledge-driven Biometric Authentication in Virtual Reality". In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems Extended Abstracts*.
- Rakhshani, Hojjat et al. (2020). "Neural architecture search for time series classification". In: *IEEE International Joint Conference on Neural Networks*.
- Ismail Fawaz, Hassan, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, et al. (2019b). "Adversarial Attacks on Deep Neural Networks for Time Series Classification". In: *IEEE International Joint Conference on Neural Networks*.
- – (2019f). "Deep Neural Network Ensembles for Time Series Classification". In: *IEEE International Joint Conference on Neural Networks*.
- Ismail Fawaz, Hassan, Germain Forestier, Jonathan Weber, François Petitjean, et al. (2019). "Automatic alignment of surgical videos using kinematic data". In: *Artificial Intelligence in Medicine*.
- Ismail Fawaz, Hassan, Germain Forestier, Jonathan Weber, Lhassane Idoumghar, et al. (2018b). "Evaluating surgical skills from kinematic data using convolutional neural networks". In: *International Conference On Medical Image Computing and Computer Assisted Intervention*, pp. 214–221.
- – (2018d). "Transfer learning for time series classification". In: *IEEE International Conference on Big Data*, pp. 1367–1376.