

Hafedh Hichri

📍 Sousse, Tunisia ✉️ hhichri60@gmail.com ☎️ +21694993040 🌐 in/hafedh-hichri 🏠 not-lain.github.io

SUMMARY

Machine Learning Engineer and **Hugging Face Fellow** with expertise in AI, NLP, computer vision, and LLMs. Proficient in PyTorch, TensorFlow, and transformers, with a track record of **contributing to open-source AI libraries**.

SKILLS

Technical Skills: Transformers, PyTorch, TensorFlow, Keras, PEFT, Python, JavaScript, Gradio, Docker, Git, PyPI, NLP, Computer Vision.

EXPERIENCE

Software Engineer

Chonkie

August 2025 - Present, Remote, USA

- Maintaining Chonkie, our lightweight chunking library.

Machine Learning Researcher

PramaLLC

May 2025 - July 2025, Remote, USA

- Conducting research in image segmentation, focusing on background removal as well as image inpainting.

Machine Learning Engineer

Mermory

October 2024 - Present, Remote, USA

- Improved customer satisfaction by 20% through the integration of a real-time visual question-answering pipeline, enhancing user experience
- Developed a suite of AI-powered image manipulation APIs to support advanced visual editing tasks, including segmentation, infilling, and outfilling.
- Built robust file conversion pipelines to standardize and transform diverse formats (PDF, Word, PPT, image, markdown) into the desired format.

Machine Learning Engineer

Juridoc

March 2024 - May 2024, Tunis, Tunisia

- Enhanced data extraction by benchmarking various OCR models across multiple languages and document types, ensuring optimal performance.
- Optimized the retrieval pipeline by implementing reranking strategies based on relevance, release date, and hierarchical norms, improving data accuracy.
- Identified and fixed issues with existing libraries and shared best practices in maintaining them with coworkers.

AI Research Engineer Intern

InstaDeep

February 2023 - August 2023, Tunis, Tunisia

- Aggregated online handwriting datasets by introducing a new format, improving data quality and efficiency, and lowering errors by 5% CER.
- Spearheaded the development of an LLM for Arabic online handwriting recognition with a 7.8% CER.
- Conducted research by analyzing 8 scholarly papers, leading to the extraction of key elements in the project.

EDUCATION

Master of Science (M.S.) - Computer Science

Minor in Engineering • National School of Electronics and Telecoms of Sfax • Sfax • 2023

- Coursework: Data Analysis, Software Engineering, Algorithms, Artificial Intelligence, Machine Learning

OPEN SOURCE

- Contributed to improving more than 5 AI libraries, including **Transformers**, **Unsloth**, **Autotiktokenizer**, and **FAL**.
- Helped open-source state-of-the-art AI models, including BEN2, RMBG1.4, BiRefNet, MatAnyone.
- Advocated for the ModelHubMixins classes adopted by more than **3000** public AI models across HuggingFace.
- Maintaining **Chonkie** an AI library that offers a wide variety of chunking methods, including semantic chunking.
- Wrote a **blogpost** about the inner workings of the transformer architectures, and tensors dimensions and how they propagate through the architecture
- Authored an **article** about **KV Caching** and how it operates under the hood, explaining its role and what allows it to optimize the inference speed.

CONTRIBUTIONS AND AWARDS

- **PXIA** • Developed a transformer-like AI library with **AutoModel** and **trainer API** support.
- **LOADIMG** • Published a library for loading and converting images with API compatibility, reaching 1,000,000 downloads.
- **RAG** • Authored a blog post about RAG, captivated 30,000 readers, and earned 80 upvotes.
- **BACKGROUND REMOVER** • Collaborated with the creator of BiRefNet in hosting BiRefNet on HuggingFace amassing 2000 likes and **1 million** visits.
- **BEN2** • Set up the library for ben2, the current state-of-the-art AI model for background removal, and all the necessary integrations with HuggingFace.
- **GRADIO SUCCESS MODAL** • Added gradio success modal, now used with zero GPU across the hub
- **Teacher** • Taught more than 2000 students in more than 70 AI events.
- **Administering social media** • Administered a social media group with 410K members and a FB page with 10K members
- **TensorFlow developer** • Awarded a voucher as one of the top 3 most active Google clubs in the MENA region to sit for the exam, in which, after you pass, you become a TensorFlow developer certified.

