



# Haroun Hassan BRAHIM

📞 07 58 83 77 78 | ✉ haroun-hassan.brahim@etu.univ-amu.fr  
🌐 LinkedIn | 🐙 Github | 📧 Researchgate | 📍 Marseille, France

**PROFIL** | Etudiant en Intelligence Artificielle, spécialisé en Apprentissage Automatique et Apprentissage Profond.

## COMPÉTENCES

**IA:** Apprentissage Automatique, Apprentissage Profond, Vision par Ordinateur, Traitement Automatique du Langage Naturel, (LLMs, Transformers), IA Générative  
**Data:** Traitement, analyse, visualisation  
**Frameworks:** Scikit-learn, keras/TensorFlow, PyTorch  
**Langages:** Python, Java, C/C++, Dart/Flutter, Matlab, R, SQL  
**Outis/MLOps:** Git, MLflow, Docker, AWS, Vercel,  
**Autres:** Scrum (Agile), Pytest, JUnit, Slurm

## FORMATIONS

🎓 **Master en Intelligence Artificielle et Apprentissage Automatique** 2024 – en cours  
Faculté des Sciences AMU | [Aix-Marseille Université](#)

🎓 **Master de Recherche en Technologies Numériques pour la Santé Physique et Mentale** 2020 – 2022  
Institut Supérieur d'Informatique et de Multimédia de Sfax | [Université de Sfax](#)

🎓 **Licence Fondamentale en Sciences Informatiques** 2017 – 2020  
Faculté des Sciences de Gabès (FSG) | [Université de Gabès](#)

## EXPÉRIENCES PROFESSIONNELLES

**Assistant chercheur stagiaire** 📅 Mai 2025 – Juillet 2025  
[Centre de Recherche en Psychologie et Neurosciences \(CRPN\)](#) 📍 Marseille, France

- ✎ Bases Cérébrales du traitement de la parole chez le nourrisson: EEG et Apprentissage Automatique
- Visualisation, prétraitement de données : filtrage, suppression des artefacts et segmentation
  - Mise en place du Pipeline d'apprentissage pour la classification des propriétés sonores à partir des données
  - Décodage EEG : cas des phonèmes /da/ et /ba/, analyse et évaluation
- Outils: Scikit-learn, VsCode, Jupyter-Lab, Slurm (Cluster), Git, Gitlab, MNE-Python

**Assistant chercheur stagiaire** 📅 Février 2022 – Novembre 2022  
[Multimedia, InfoRmation systems and Advanced Computing Laboratory \(MIRACL\)](#) 📍 Sfax, Tunisie

- ✎ Détection du MCI à partir des données EEG et FNIRS, basée sur le Deep Learning (Transformers)
- Analyse, nettoyage et organisation de données
  - Visualisation détaillée et précise de données
  - Personnalisation de l'architecture Transformer
  - Construction du modèle final de classification, entraînement, test et évaluation du modèle
- Outils: Tensorflow, Keras, Colab, Jupyter-Lab, Matlab, MNE-Python, Git, Gitlab, CUDA et cuDNN Toolkit

**Stagiaire data-analyste** 📅 Mars 2021 – Avril 2021  
[Multimedia, InfoRmation systems and Advanced Computing Laboratory \(MIRACL\)](#) 📍 Sfax, Tunisie

- ✎ Visualisation et analyse des signaux électroencéphalographiques (EEG)
- Enregistrement et acquisition de données EEG avec le casque Emotiv EPOC+
  - Visualisation de données avec le logiciel EmotivPro
  - Traitement de données avec élimination d'artefacts
  - Développement d'une interface pour la visualisation et l'analyse de données EEG
- Outils: Colab, Emotiv EPOC+, EmotivPro, Django, Vuejs

## Stagiaire Développeur

Unité de Recherche RTIM (Research Team in Intelligent Machines)

📅 Février 2020 – Juin 2020

📍 Gabès, Tunisie

### ✎ Développement d'une application OCR pour scanner des algorithmes SCRATCH

- Traitement par filtrage approfondi des images
- Segmentation des blocs composants les algorithmes Scratch, basée sur les contours et le Deep Learning
- Détection et Extraction du texte dans différents blocs segmentés
- Génération de graphe par DFS et Conversion d'algorithme Scratch en python.

Outils: OpenCV, Tesseract-OCR, EAST (An Efficient and Accurate Scene Text Detector), Mask-RCNN, Graphviz

## Stagiaire Développeur

Intellect Academy | Département IT

📅 Janvier 2020 – Février 2020

📍 Gabès, Tunisie

### ✎ Développement d'une application OCR pour scanner des algorithmes SCRATCH

- Construction de la base de données à partir des images contenant d'algorithmes Scratch
- Catégorisation et annotation des images avec le logiciel d'annotation VIA

Outils: Scratch, VGG Image Annotator (VIA)

## PROJETS ACADÉMIQUES

### Aix-Marseille Université

📅 2024 – 2026

#### ✎ Conception et implémentation d'un modèle de langage neuronal basé sur le (MLP) en PyTorch

- Entraînement du modèle et optimisation par descente de gradient
- Implémentation de la génération de texte par échantillonnage aléatoire à partir des distributions de proba prédites
- Calcul de la perplexité sur un corpus de test pour évaluer la qualité du modèle.
- Optimisation des performances via le traitement par lots (batching) et recherche de la taille de batch optimale

#### ✎ Implémentation d'un modèle de langage markovien (unigram, bigram, trigram)

- Génération automatique de phrases à partir d'une amorce (<s> <s> pour trigram) en tirant aléatoirement selon les distributions conditionnelles estimées
- Utilisation de techniques d'estimation par maximum de vraisemblance et lissage de Laplace pour gérer les proba nulles.
- Évaluation des modèles via la perplexité et le taux de bonnes prédictions sur des textes masqués
- Analyse comparative des performances entre les modèles unigram, bigram et trigram sur des corpus segmentés

### Laboratoire MIRACL

📅 2021 – 2022

#### ✎ Détection et prédiction de l'état pathologique de la rétine dans des images OCT

- Prétraitement des données, usage du modèle de classification Inception-V3, entraînement, test et évaluation du modèle

#### ✎ Détection du cancer de la peau

- Prétraitement des données, construction du modèle de classification CNN, entraînement, test et évaluation du modèle

#### ✎ Analyse du Cancer du sein

- Analyse des données, visualisation des données, sélection de caractéristiques et extraction de caractéristiques.

## CERTIFICATIONS

### CCNA1: Certifié 2020

Cisco Certified Network Associate, Routing and Switching

📅 Juillet 2020 – Pas de date d'expiration

## LANGUES

🇫🇷 Français



🇬🇧 Anglais



🇸🇩 Arabe



## PUBLICATIONS

B. Haroun Hassan et al. New MCI Detection Method Based on Transformer and EEG Data European Signal Processing Conference (EUSIPCO) (2023) [🔗](#)