

Baye Lahad Mbacke

10 rue de Roumanie – Rennes, France

☎ 06 01 32 80 22 • ✉ [regular] lahadmbacke00@gmail.com

in baye-lahad-mbacke • 🌐 LahadMbacke • 🐦 LahadMbacke

Fort d'une formation en Traitement Automatique du Langage et Machine Learning (Master ATAL, Nantes Université) et d'expériences concrètes en R&D (géotagging, extraction et classification d'entités), je suis résolument motivé pour une thèse en Intelligence Artificielle visant à renforcer la compréhension automatique du langage naturel via des modèles avancés et des méthodes innovantes de fine-tuning.

Formation Académique

Nantes Université

Master ATAL (Apprentissage et Traitement Automatique de la Langue)

Spécialisation en techniques avancées de NLP et d'apprentissage automatique

Nantes, France

2023–2025

Université de Strasbourg

L3 en Informatique

Approfondissement en algorithmique et programmation

Strasbourg, France

2022–2023

Université Cheikh Anta Diop de Dakar

Licence en Mathématiques-Informatique

Formation aux fondamentaux en sciences mathématiques et informatiques

Dakar, Sénégal

2018–2022

Expérience Professionnelle

Ouest-France

Rennes, Bretagne, France

Assistant Data Scientist – Recherche&Développement, Contrat en alternance Sept. 2024–Aujourd'hui

- Réalisation de tâches de "geo tagging" : identification de références de localisation à partir de bribes d'adresse dans des articles journalistiques
- Rattachement à des tags éventuels de lieux
- Travail dans l'extraction de connaissances : extraction d'entités nommées et mise en relation d'entités nommées, désambiguïsation d'entités géographiques (adresse, place, monument, ville etc..)

Laboratoire des Sciences du Numérique de Nantes

Nantes

Stagiaire Recherche

Jan. 2024–Mai 2024

Comparaison de systèmes de reconnaissance d'entités nommées, sur des articles de presse

- Contribution à la comparaison de systèmes de reconnaissance d'entités nommées
- Utilisation de techniques d'apprentissage automatique pour extraire des informations clés
- Participation à l'annotation et à l'analyse de données structurées et non structurées
- Évaluation de différentes approches en traitement automatique des langues

Nantes Université

Nantes

Projet Académique

Fév. 2024–Mars 2024

Compétition DEFT2013 - Tâche 2 : Classification des recettes de cuisine

- Développement de modèles de classification de recettes (Plat principal, Entrée, Dessert)
- Utilisation du traitement automatique du langage et de l'apprentissage automatique
- Classification basée sur les titres et les instructions des recettes

Auto-encodeur et classification

- Évaluation de l'intérêt des auto-encodeurs pour la classification de données tabulées
- Implémentation de réseaux de neurones pour la compression et la représentation de données
- Construction de fonctions de compression dans un espace latent de dimension réduite
- Analyse comparative des performances de classification avec et sans auto-encodeur

Compétences Techniques

Machine Learning & Deep Learning: PyTorch, Scikit-learn, Large Language Models (LLM), Fine-tuning, Quantization, Distillation

Traitement Automatique du Langage: Spacy, NLTK, LangChain, RAG, Agents

Langages de Programmation: Python, R, Java, C/C++

Bases de données: SQL, MySQL, Oracle, PostgreSQL, MongoDB

Langues

Wolof: Langue maternelle

Français: Langue courante

Anglais: Lu, écrit, parlé

Qualités Personnelles

- Méthodique
- Adaptabilité
- Rigueur

Centres d'Intérêt

Lecture: Journaux, Actualités technologiques, Veille informatique

Cultures: spectacles, pièces de théâtre, musique mbalax

Sport: Football