

ELHADJI ATIKA

Élève ingénieur en **Data Science** à l'Institut National de Statistique et d'Économie Appliquée (**INSEA**), passionné par le **Data Engineering** et la création de pipelines de données.

A la recherche d'un stage d'application

0651170297

elhadjiatikapfa@gmail.com

Atila EL HADJI



Formation

Institut National de Statistique et d'Économie Appliquée (INSEA)

RABAT, MAROC

2022 -Présent

ÉLÈVE INGÉNIEURE EN DATA SCIENCE

Lycée IBN ABDOUN, Khouribga

Khouribga, MAROC

2020-2022

**CLASSES PRÉPARATOIRES
MATHÉMATIQUES-PHYSIQUE (MP)**

Lycée IBN TOFAIL, Oued Zem

Oued Zem, MAROC

2019-2020

BACCALAURÉAT SCIENCES MATHÉMATIQUES A



Compétences

- Statistique:** Probabilités, statistiques descriptives, inférence-statistique, Séries-chronologiques, apprentissage statistique
- Programmation :** Python, Java(Spring boot), TypeScript(Angular), R, SQL
- Technologies :** Pandas, Numpy, Scikit-learn, TensorFlow, Git
- data engineering:** Apache airflow, ETL, data warehousing
- SGBD:** Mysql, Postgresql
- DATA VIZ :** Power BI
- LATEX**



Activités parascolaires

Cheffe de cellule Formation-Club INSEA INNOVATION EDGE

- ORGANISATION ET ANIMATION DE SESSIONS DE FORMATION EN DATA ET DÉVELOPPEMENT ET SOFTSKILLS

Cheffe de cellule Technique-Insea charity

- GESTION DES ASPECTS TECHNIQUES ET NUMÉRIQUES DES INITIATIVES DU CLUB



Soft skills

- Travail en équipe
- la lecture
- gestion de stress



LES PROJETS

★ **Traitement du Signal – PROTECTION DE L'IDENTITÉ VOCALE PAR TRAITEMENT DU SIGNAL**

- Conception et implémentation d'un algorithme en Python visant à modifier les caractéristiques vocales pour préserver l'anonymat des locuteurs. Utilisation des techniques avancées de traitement du signal avec Librosa dans le cadre du compétition national de data science organisé par Inwi, Cleverlytics et l'UM6P.

★ **Projet Data Engineering – Analyse des avis clients sur les banques marocaines**

- Objectif : Construire un pipeline de données automatisé pour analyser les avis clients sur les banques au Maroc.
- Scraping Web de données à partir de Google Maps (récupération de noms, notes, commentaires, localisations) via Selenium et BeautifulSoup sur Python.
- Nettoyage, normalisation et structuration des données textuelles.
- Création de modèles DBT pour structurer les données dans un data warehouse PostgreSQL.
- Orchestration du pipeline complet sur un système Ubuntu, avec planification automatique des tâches via Apache Airflow.
- Déploiement d'un dashboard interactif avec Looker Studio pour visualiser les tendances clients, scores par banque, évolution des avis...
- Stack technique :** Python, DBT, PostgreSQL, Apache Airflow, Looker Studio, Git, Ubuntu

★ **Projet Web – Plateforme de gestion de club + chatbot IA intégré (INSEA Innovation Edge)**

- Développement d'un site web interactif pour le club INSEA Innovation Edge (gestion des membres, événements, et publications).
- Frontend avec Angular : authentification, affichage des activités, tableau de bord pour admin.
- Backend en Spring Boot : API REST sécurisées pour gérer les utilisateurs, les événements et les rôles.
- Intégration d'un chatbot intelligent basé sur une architecture RAG (Retrieval-Augmented Generation) pour répondre aux visiteurs sur le club et l'école.
- Utilisation de Together AI pour l'inférence avec le modèle LLaMA 3 (Meta) via API.
- Traitement et vectorisation des documents du club pour l'indexation des réponses.
- Stack technique :** Angular, Java Spring Boot, Python, RAG, FAISS, Together AI, LLaMA 3, Git

★ **Projet data visualisation et machine learning – Modèle de prédiction de prêt bancaire**

- Réalisé une analyse exploratoire des données (EDA) incluant le traitement des valeurs manquantes et l'encodage des variables catégorielles.
- Construit un modèle de machine learning pour prédire l'éligibilité à un prêt bancaire à partir de données socio-économiques.
- Testé plusieurs algorithmes (Logistic Regression, Decision Tree, Random Forest) et optimisé les performances avec GridSearchCV.
- Présenté les résultats via des visualisations claires (matrices de confusion, courbes ROC, importance des variables).

Stack technique : Python, Pandas, Scikit-learn, Matplotlib, Seaborn



Les expériences

OCP-GROUPE SIDI CHENANE

KHOURIBGA, MAROC

01.07.2023 - 30.07.2023

- Analyse approfondie de la base de données des produits du magasin principal du département électrique au sein du service SIDI-CHENANE.
- Conception et développement d'une interface de gestion des opérations du magasin (ajout de produits, gestion des stocks, sortie de matériel, etc.).
- Technologies utilisées :** Python (Jupyter) pour l'analyse et la visualisation des données, HTML/CSS et JEE pour le développement de l'interface.

**Ministère de l'Agriculture et de la Pêche Maritime –
Département de la Pêche Maritime .**

Rabat, MAROC

01.07.2023 - 29.08.2023

- Conception et développement d'un site web permettant d'optimiser la communication entre les délégations de pêche à travers tout le Maroc et le Ministère.
- Mise en place d'une architecture robuste en utilisant Spring Boot pour le backend et Angular pour le frontend.
- Utilisation de MySQL pour la gestion des données et de Git/Trello pour le suivi et la gestion du projet.