



ABDELGHAFOR ELGHARBAOUI

DATA ENGINEERING STUDENT

Email: abdelghaforelgharbaoui@gmail.com
Phone: 0607514544
Github: Abdelghafor-az
Adress: Essaouira, Morocco

SUMMARY

Data engineering student seeking a graduation project opportunity to leverage cutting-edge technologies in developing innovative data solutions. Passionate about addressing real-world challenges through effective data management and processing.

EXPERIENCE

AI Developer Intern - Remote	Shiftbricks	Jun 2024 - Sep 2024
<ul style="list-style-type: none">Developed an AI Legal Assistant app that allows users to browse, search legal documents, and receive instant responses via a chatbot.Utilized MongoDB Atlas as a vector store for semantic search, delivering precise results for search service and enhancing chatbot response generation.Implemented citation extraction for both AI-generated responses and search results.Developed a browser for browsing and exploring legal documents, where search citations redirect users to relevant sections.Integrated next-intl to provide seamless multilingual support across the platform.Set up a GitLab CI/CD pipeline to automate containerization using Docker Compose.Deployed the application in Kubernetes using Azure Kubernetes Service (AKS) and leveraged Azure Container Registry to securely store and manage Docker images, ensuring that all company assets remained protected and accessible only to authorized users.Technology Stack: FastAPI - Mongo Atlas - Langchain - NextJS - Tailwind - Postgres - Docker Compose - GitLab CI/CD - Azure Kubenetes Service - Azure Container Registry.		
Backend Developer Intern - Remote	3D Smart Factory	Jul 2024 - Oct 2024
<ul style="list-style-type: none">Collaborated on an e-commerce microservices project, Implemented Authentication and Authorization in Users Service using Spring Security and leaving OpenFeign for seamless communication with the Product Management and Payment Services.Developed a Notification Service using Kafka to handle asynchronous requests, enhancing user experience.Utilized Grafana and Prometheus for real-time monitoring, along with Loki for log aggregation and Tempo for distributed tracing, ensuring comprehensive observability and efficient troubleshooting across microservices.Integrated Cloudinary for efficient media file storage, ensuring optimized performance and scalability.Technology Stack: Spring - Kafka - Docker Compose - Cloudinary - Prometheus - Grafana - Garfana Loki - Grafana Tempo		

EDUCATION

Data Engineering National School of Applied Sciences AI Hoceima ENSAH	Sep 2020 - July 2025
---	-----------------------------

CERTIFICATIONS

Machine Learning Specialization [DeepLearning.AI]	September 4, 2023
Python for Data Science, AI & Development [IBM]	February 14, 2024

SKILLS

- **Programming Languages** : Python - Java - R - PL/SQL - Bash - C - TypeScript
- **Databases** : SQL (Postgres - Oracle - MySQL - SQL Server)
NoSQL (Mongo - Cassandra - Redis)
- **Cloud Data Platforms** : Azure Data Factory - Databricks - Snowflake
- **Big Data & ETL** : Hadoop - Spark - Kafka - Airflow - dbt - Selenium - Power BI
- **Machine Learning & AI** : Scikit Learn - Keras - LangChain - nltk - spaCy - Hugging Face
- **DevOps** : Linux - Git - Docker - Kubernetes - Github Actions - GitLab CI/CD - Prometheus - Grafana - Azure
- **Web Development** : Spring (Web - Security - JPA - Cloud) - FastAPI - Flask - NextJS - Tailwind

PROJECTS

Patent Analysis

- Empowering users to explore and analyze patents in the field of Education and AI.
- Leveraging web scraping, APIs, and data warehousing for comprehensive patent insights.
- Established comprehensive dashboards in PowerBI and integrated Chart.js for dynamic visual insights, powered by background queries to the data warehouse.
- **Technology Stack:** Flask - Spark - Postgres (Data Warehouse) - Mongo Atlas - PowerBI - Chart.js

Real Time People Monitoring

- Real-time foot traffic monitoring in different zones, providing dynamic insights.
- Utilized lambda architecture and computer vision frameworks to process streaming data.
- Employed Kafka for seamless event streaming and HDFS for scalable storage, ensuring reliable data flow and management.
- Developed Power BI dashboards to display real-time graphs for each zone, providing dynamic insights into foot traffic patterns.
- **Technology Stack:** Docker - HDFS - Kafka - Spark - Cassandra - ClickHouse - Power BI - Yolo

HelpDesk Dashboard

- Developed a dashboard to track created and resolved tickets, as well as incidents, by employee and team.
- Performed data preprocessing using Power Query and modeled the data in Power BI.
- Created various report pages with key indicators and interactive visualizations.
- Published the dashboards in Power BI Service, set up the workspace, managed access, and created an application in the workspace.
- **Technologies:** Power BI, DAX, Power BI Service.

Political Opinions Analysis on Palestinian-Israeli Conflict

- Gauged international opinions on the conflict through analysis of YouTube comments.
- Involved web scraping, manual labeling, and text preprocessing.
- Fine-tuned BERT for sentiment analysis and opinion classification.
- **Technology Stack:** Bert (Hugging Face) - Azure Translator - Scikit Learn - Apify

ADDITIONAL INFORMATION

- **Languages:** English, French.
- **Curiosity and Continuous Learning:** Commitment to staying updated on the latest data technologies and techniques.
- **Social Work:** Active participant in social and educational initiatives with the Quran Club at school.
- **Community Engagement:** Participation in national events such as MNPC23 (Moroccan National Programming Contest), contributing to the development of the community.