Mohamed Dyn

+212 657 722 751 | dynmedo28@gmail.com | Mohamed Dyn | dynmohamed | Rabat, Maroc

Profil

Final-year Engineering Student (Data Science) at INSEA, seeking an end-of-studies internship to apply my knowledge in Data Engineering, Machine Learning, and Data Analytics on concrete projects.

Education

Institut National de Statistique et d'Économie Appliquée (INSEA) — Rabat, Morocco
Engineering Degree (MEng), ***Data Science*** — *Sept. 2023 – Present*

Classes Préparatoires aux Grandes Écoles (CPGE) — Tangier, Morocco
Preparatory Diploma in Mathematics, Physics, and Engineering Sciences — *Oct. 2021 – July. 2023*

Professional Experience

TEMACONCEPT — Temara, Morocco
Data Scientist (Intern) — *July. 2025 – Sept. 2025*

Development of an optimization engine for CPGE teacher assignment : modeled the problem as a needs grid and implemented the Hungarian algorithm (Python) to maximize profile/position matching.Design of a RAG (Retrieval-Augmented Generation) system integrated into a chatbot to query assignment results using natural language.

⇒ ***Performance Optimization :*** Achieved **100%** automation of teacher assignment, reducing administrative time from **5** days to **2** hours via the implementation of the Hungarian algorithm in Python..

⇒ **Technologies** : Python, Pandas, LangChain, LLM, Vector DB.

MINISTRY OF RELATIONS WITH THE PARLIAMENT — Rabat, Morocco
Data Analyst (Intern) — *July. 2024 – Aug 2024*

Design of interactive dashboards (**Power BI**) for monitoring parliamentary activities, facilitating strategic decision-making by management.Data Cleaning and Modeling to transform raw data into reliable Key Performance Indicators (**KPIs**).

⇒ Facilitated decision-making through clear and interactive report representation.

Projects

Deployment of a Serverless ELT Architecture on AWS — March 2025

— Designed an **automated data pipeline** for commercial analysis (Ingestion S3 → Transformation Glue → Warehousing Redshift).
— Orchestrated workflows via **Apache Airflow** ensuring reliability and error recovery.
— *Stack :* AWS (S3, Glue, Lambda, Redshift), Airflow, Power BI.
⇒ **Gain :** Reduced data availability latency from **24h à 15 min**,enabling near real-time steering.

Automation of Big Data ETL Flows (Azure & Databricks) — Apr. 2025 – May 2025

— Implemented a **Medallion Architecture** (Bronze/Silver/Gold) on Azure Databricks to clean and structure raw data.
— Developed optimized **PySpark** jobs and continuous ingestion via **Auto Loader**.
— *Stack :* Azure Data Factory, ADLS Gen2, Databricks, PySpark.
⇒ **Bénéfice :** Accelerated massive data volume processing by **40%** because of Spark distributed computing.

Development of a Medical AI Assistant (RAG RLHF) — May 2025 – June 2025

— **RAG** Architecture with vector indexing in **ChromaDB** to query medical PDF documents.
— Implemented an **RLHF** loop stored in **PostgreSQL** to fine-tune the model.
— *Stack :* Python, LangChain, FastAPI, Docker, PostgreSQL.
⇒ **Impact :** Improved response accuracy and reduced hallucinations compared to a standard **LLM**.

Technical Skills

Cloud & DevOps : AWS (S3, Glue, Lambda, Redshift, Athena, SageMaker), Azure (Data Factory, Synapse, ADLS), Docker, Git, CI/CD

Data Engineering : Apache Spark (PySpark), Databricks, Apache Airflow, Kafka, Hadoop (HDFS/MapReduce), DBT, ETL/ELT Pipelines

Data Science & IA : Machine Learning (Scikit-learn, TensorFlow), GenAI (LLMs, RAG, LangChain, Ollama, Hugging Face), MLOps

Programmation & Bases de Données : Python (Advanced), SQL (Advanced), Scala, Java, NoSQL (MongoDB, Cassandra, HBase), PostgreSQL, DuckDB

Développement Web & API : FastAPI, Flask, React, REST APIs

Business Intelligence : Power BI (DAX, Power Query), Tableau, AWS QuickSight

Languages

English (Fluent) French (Fluent) Arabic (Native))