

Data Scientist Resume - English

AYYOUB EL-HYANI

elhyani.ayyoub@gmail.com

Professional Summary

- Data Science professional with expertise in predictive modeling, machine learning, and NLP
- **Currently seeking Data Scientist apprenticeship position** starting September 2025
- Experience in healthcare analytics, agricultural AI solutions, and predictive maintenance

Education

Université Paris Cité (Starting September 2025)

Master 2 – Distributed Systems and Data Science Technologies (apprenticeship) | Paris, France - Apprenticeship-based program focused on distributed systems, artificial intelligence, big data, and cloud architectures

École des Sciences de l'Information (ESI) (2022 – 2025)

Engineering Degree – Data Science and Knowledge Engineering | Rabat, Morocco - Advanced courses in machine learning, data engineering, and statistical modeling

Preparatory Classes for Engineering Schools (2020 – 2022)

Mathematics-Physics and Industrial Sciences Track (MPSI/MP) | Marrakesh, Morocco

Professional Experience

Somei Maroc (February 2025 – Present)

Data Science Intern – Predictive Maintenance | On-site, Rabat - Developed predictive models (Random Forest, XGBoost) on **Dataiku** to anticipate pipe failures using historical, sensor, and geospatial data - Implemented **survival analysis** algorithms (Random Survival Forest) to estimate remaining equipment lifetime - Collaborated with business teams for model interpretation and integration into management tools

Nucleon Security (June 2024 – August 2024)

Data Engineering Intern | On-site, Rabat - Migrated 10M+ malware records from MongoDB to PostgreSQL, improving performance and data integrity - Designed Python scripts (pandas, SQLAlchemy) for ETL with 35% reduction in processing time - Built real-time threat dashboards using Apache Superset

Broccoli Data (July 2023 – September 2023)

Data Engineering Intern | Remote - Created dbt models to automate transformation of complex statistical data - Improved query performance by converting datasets to Parquet format

Machine Learning Projects

Neuroblastoma Cancer Outcome Prediction

- Developed ML models to predict survival and recurrence using integrated clinical and genomic data
- Conducted high-dimensional data analysis using PCA, t-SNE, and clustering
- Evaluated models using cross-validation, ROC-AUC, and survival metrics

Agricultural Chatbot with RAG (Retrieval-Augmented Generation)

- Built a chatbot tailored for Moroccan farmers using LLMs and a RAG pipeline
- Performed vector indexing (FAISS + SentenceTransformers) on official documents and guides
- Generated reliable responses via OpenAI API and custom prompt engineering

Technical Skills

- **Languages:** Python, R, MATLAB, Scala, Java
- **ML/AI:** scikit-learn, XGBoost, TensorFlow, PyTorch, Random Forests, Survival Analysis
- **NLP/LLM:** Transformers, RAG, SentenceTransformers, FAISS, OpenAI API
- **Data Engineering:** SQL, PostgreSQL, MongoDB, dbt, Dataiku, Apache Superset
- **Visualization:** Power BI, Matplotlib, Seaborn, Tableau
- **Tools:** Jupyter, Git, Docker, REST APIs

Soft Skills and Languages

- **Languages:** French (C2), English (C1), Arabic (native)
- **Soft skills:** Clear communication, critical thinking, autonomy, teamwork, adaptability, time management