**HAJJAJI Imad**
**Data Scientist / Data Analyst**
Morocco • imad.hajjaji@gmail.com • +212638362815 • www.linkedin.com/in/imad-hajjaji-0003a0132/

## Education

**National Institute of Posts and Telecommunications (INPT)** Rabat, MA
PhD, Application of Data Science in Smart Cities. 01/2021 – Expected 2025
Fields: Data Science, Data Analysis, Smart Grid, Optimization, Telecommunication, IoT, Cyber Security.

**National Institute of Statistics and Applied Economics (INSEA)** Rabat, MA
Operational Research and Data Science Engineer. 09/2016 – 09/2020

**PREPARATORY CLASSES** Laayoune, MA
Mathematics and Physics (MP) Classes 09/2014 – 06/2016

## Experience

**Policy Center for the New South – PCNS** Rabat, MA
**Program Officer – AI Engineer** 07/2024 – Now
**Work Environment:**

- Coordinated a cross-functional team of 4 data scientists and engineers to deploy AI solutions (chatbot, translation platform) adopted by 60+ users, reducing operational costs by 80%.
- Implementation of AI solutions in a policy research organization setting
- Agile development methodology with biweekly sprints

**Technology Stack:**

- **AI Models & Frameworks:** LangChain for orchestration, tested and implemented various LLMs including LLAMA_3_1_70B, LLAMA_3_1_8B, MISTRAL_LATEST, LLAMA_3_1_70B_GROQ_TOOL_USE, GPT-4o, and GPT-4o mini
- **API Integration:** DeepL API for translation services, ChatGPT API for translation correction and content enhancement
- **Web Development:** Flask RESTful and FastAPI for backend development, deployment of web services with user-friendly interfaces
- **DevOps:** Docker for containerization, CI/CD pipeline implementation
- **Data Processing:** NLTK, Spacy for NLP tasks, Scikit-learn for machine learning components
- **Automation:** Microsoft Power Automate for workflow automation, Task Scheduler for periodic tasks
- **Database & Storage:** SQL databases, cloud storage solutions for document management
- **UI/UX:** HTML, CSS, JavaScript for frontend development

***Project**: Translation Plateform*

- Proposal and implementation of an internal translation platform using the DeepL API.
- Integration of table translations while maintaining their format, reducing translation costs by 80% with usage by 60 people.

***Project:** Archiving Solution*

- Implementation of an online archiving solution connected to a dedicated email via Power Automate.
- Automatic launch process triggered by incoming emails with attachments, used by all staff at the Policy Center.

***Project**: CHATBOT POLICY CENTER*

- Data collection and cleaning for the creation of a chatbot.
- Conversion of videos, podcasts, and roundtable discussions into text.
- Creation of a RAG-based chatbot to explore internal data, with integration of the OpenAI API and deployment as a web service.

***Project**: REPEC Structure Analysis*

- Analysis of new tags and research documents.
- Extraction of the most relevant tags and automatic structuring of files via a script executed at the beginning of each month.

***Project**: Dashboard for Social Media*

- Data collection and analysis, creation of a Power BI dashboard for real-time insights on social media

engagement.

***Project**: Salesforce CRM Tracking*

- Tracking the creation of contacts for each event.
- Updating BeeFree templates for newsletters and invitation emails.
- Creation and linking of Power BI dashboards for improved event tracking.

Tools: Python, Flask, HTML, CSS, JavaScript, Ubuntu, Git, Microsoft Outlook, Power Automate, Linux, Task Scheduler, Excel, Power BI, Salesforce, BeeFree, NLTK, Spacy, Scikit-learn, Docker.

**Sofrecom Services Morocco – Orange Luxembourg** Remote
**Data Scientist / Data Analyst** 06/2023 – 06/2024
**Work Environment:**

- Fully remote position working with international telecom client
- Collaborative work with data engineering and business teams in Luxembourg
- Strict data privacy compliance (GDPR) environment
- Weekly progress tracking and monthly deliverables

**Technology Stack:**

- **Cloud Platform:** Google Cloud Platform (GCP) ecosystem for all development and deployment
- **Code Repository & CI/CD:** GitLab for version control and CI/CD pipeline integration
- **Migration:** Led migration efforts from on-premises solutions to cloud-based implementations
- **Cloud Services:** Google Cloud Functions for serverless computing, VertexAI for machine learning operations, BigQuery for data warehousing and analysis, Cloud Scheduler for automated workflow management
- **Data Processing:** Python ecosystem (Pandas, NumPy, Scikit-learn, SQLAlchemy)
- **NLP Tools:** NLTK for natural language processing tasks, particularly for address matching
- **Visualization:** Power BI dashboards for business intelligence and reporting
- **Data Science:** Implementation of machine learning models for customer behavior prediction and address matching algorithms

**Projects:**

- Upsell Prediction: Enhancing Revenue Leveraging ML to Identify Mobile upgradable low-cost offers' customers Performed statistical analysis and prediction model on data sets.
- Pre2post Prediction: predict which prepaid mobile customers can switch to a postpaid offer.
- Address Matching: link the postal addresses present in the OLU database to the postal addresses present in the POST database, in order to assign them a postalID, useful for giving a customer fiber eligibility.

Tools: Python (Pandas, Numpy, Scikit-learn, SQLAlchemy, NLTK, Scipy, Seaborn …), BigQuery, GCP(Cloud Functions, VertexAI, Cloud Scheduler, Power BI)

**Digital for Research (D4R) - UM6P** Benguerir, MA
**Data Scientist** 04/2021 – 06/2023
**Work Environment:**

- Led end-to-end development of a **labor market platform**, directing 3 data scientists to integrate real-time CNSS/ANAPEC data and NLP keyword extraction

***Project**: Labor Market Platform (OCP-UM6P) (marssad.ma)*

- Analyzed real-time data (CNSS, ANAPEC, Ministry of Labor).
- Extracted data describing the labor market from more than 8 Moroccan work websites.
- Keyword Extraction Engine: extract keyword from labor description.
- Built a Monitoring dashboard to better measure key model performance metrics.
- Coordinated a team of 3 data scientists working on different fields.

Tools: Pandas, Beautiful Soup, Plotly, Geopy, NLTK, Spacy, Scikit-learn, Pyspark, Flask, Power BI, SQL.

***Project:** Driver's License Automatization (NARSA-UM6P)*

- Developed 2 direction change model detection of vehicles (turn left/right, straight) using CNN model and Temporal Recurrent Networks.
- Programmed a detection model of road Lanes using OpenCV and CNN with 87% accuracy.
- Implemented face recognition model of license participants.

Tools: NumPy, TensorFlow, Keras, Pytorch, Scikit-Image, OpenCV, Git.

***Project**: Course Recommendation System (prepadigitale.com)*

- Convert +700 videos to text.
- Extracted 3 to 4 key words of each course using NLP methods.
- Developed RS: RS based on content; RS based on collaborative filtering; RS based on popularity; hybrid RS.
- Deployed the systems as web application.

Tools: Pandas, NumPy, MoviePy, selenium, SpeechRecognition, NLTK, Scikit-learn, Flask, Git.

***Project**: Data Extraction*

- Trained extraction model on +30000 files using OCR, PyTesseract and neural network to extract table/image/graph/curve data.
- Improved the performance of the extraction model.

Tools: NumPy, TensorFlow, Keras, Scikit-Image, PyTesseract, OpenCV, Git.

**Python Developer - Freelance** 04/2020 – 04/2021

- Web Application with Django: Stock Chart Pattern Recognition using Market Internals Machine Learning and Neural Network with Python (TD Ameritrade API, Web Scraping, DataViz, Trading, Deep Learning).
- Web Application with Django: Scraping data and doing a Multi-aspect sentimental Analysis with Python (Twitter-API, Nltk,TextBlob, Geopy).
- Credit Card Fraud Detection with Machine Learning in Python (Scikit-learn, Decision Tree, Logistic Regression, KNN).

**Digital for Research (D4R) - UM6P** Benguerir, MA

**Data Scientist Intern -** *Rainfall Prediction in Morocco* 04/2020 – 10/2020

- Designed and implemented a machine learning and deep learning model that predicts rainfall with more than 94% accuracy.
- Deployed the systems as web application.

## Publications

**Hajjaji I.,** El Alami H., El-Fenni M. R., & Dahmouni H. (2021, June). Evaluation of Artificial Intelligence Algorithms for Predicting Power Consumption in University Campus Microgrid. In 2021 International Wireless Communications and Mobile Computing (IWCMC) (pp. 2121-2126). IEEE.

**Hajjaji I.,** El Alami H., El Alami R., & Dahmouni H. (2021, June). Energy Consumption Characterization in University Campus Microgrid Based on Power Data Analysis. In 2022 The 9th International Conference on Future Internet of Things and Cloud (FiCloud).

**Hajjaji I.,** Kassri N., Ennouaary A., & Dahmouni H. A Dense-free Hybrid Neural Network for automatic Modulation Classification in Cognitive Radio Networks. In Physical Communication Journal (under publication).

Rhesri A., Bennani R., **Hajjaji I.**, Ben Maissa Y., Tamtaoui A., & Dahmouni H. Devlopment of a Low-cost Internet of Things Architecture for Energy and Environment Monitoring in a University Campus. In 2023 The 9th International Conference on Future Internet of Things and Cloud (FiCloud).

**Hajjaji I.,** El Alami H., Ait Mansour A. & Dahmouni H. (2021, June). Smart Energy Management System for University Microgrids: A Classification and Prediction Approach. In Planning and Operation of Distributed Energy Resources in Smart Grids II, Energies MDPI journal (under revision).

## Skills & Interests

**Computer :** Python, Java, MS Excel (VBA excel), SQL, Power BI, Git, Latex.
**Technical Leadership**:,Team Mentoring, Agile Methodologies, AI Roadmap Planning, Cloud Architecture (GCP, VertexAI)
**Language :** Fluent in Arabic, French and English, Reading knowledge in Spanish.