



# Kpodjro KPATOUKPA

## Apprenti Data Scientist

A la recherche d'une alternance en Data Science à partir de Septembre 2025, je suis dynamique, d'un esprit d'analyse et de synthèse.

Rythme : 3 jours Ecole/2 jours puis temps plein en entreprise à partir d'Avril.

✉ kpatoukpakpodjro@gmail.com

☎ +33 6 14 97 75 51

🏠 Bezons 95870

### EXPÉRIENCES PROFESSIONNELLES

#### Machine Learning Engineer Intern

03/2024 - 08/2024 ATTIJARIWafa Bank Casablanca

- Extraire les données à partir de fichiers PDF et de fichiers texte et traitement des images contenues.
- Incorporer du texte dans la base de données Chromadb pour la recherche sémantique de texte.
- Développer un système de RAG augmenté pour la gestion des questions des clients (B2B).
- Affiner ( Fine tuning) le modèle Mistral et comparaison des résultats par rapport au RAG.
- Mettre en place deux interfaces : une pour l'utilisation du chatbot, une pour une supervision basé sur 4 KPIs.

**Technologies** : Mistral AI, Langchain, Python, Pytorch, Streamlit, NLP, Chromadb, MongoDB(NoSQL), Tesseract

#### Assistant Data scientist

08/2023 - 09/2023 MLView Tanger

- Analyse exploratoire des données marketing en collaboration avec l'équipe marketing.
- Entraîner un modèle Deep Learning pour le ciblage des clients pour des campagnes futures.
- Fine tuning du grand modèle de langage ( LLM) LLAMA-2 pour la génération de l'accroche personnalisée.

**Résultat** : Amélioration du score de prédiction de 23% grâce au fine-tuning.

**Technologies** : Python, Langchain, LLAMA-2, NLP, Tensorflow, Sk-learn, SQL.

**Portfolio** : <https://kpatoukpakpodjro.github.io>

#### Projets tutorés

2022 - 2024 ..

- Implémentation d'un système de recommandation basé sur l'IA générative et les graphes : LightGNN, Factorisation matricielle, Gemini.
- Développer d'un système d'approvisionnement par anticipation d'un supermarché précis à 91% : Extraction (SQL), analyse et visualisation (R), modélisation (Python).
- Modélisation du risque de crédit via une approche statistique de scoring avancé (Recall : 93%).
- Optimisation de la fidélisation client par un modèle de machine learning précis à 92% (AUC).

### COMPÉTENCES TECHNIQUES

#### Informatique

- Machine Learning, Deep Learning, RLHF.
- Développement web & Dataviz: Java, RShiny, Streamlit, FastAPI, PowerBI, Tableau.
- Analyse de données: Excel(TCD), Python, Statistiques, R, Spark(Big Data), SQL, NoSQL.
- WebScraping : selenium, scrapy, BeautifulSoup, Google-search.

### DIPLÔMES ET FORMATIONS

#### Master Machine Learning pour la Data Science

09/2024 Université Paris Cité Paris, France

Apprentissage par renforcement et Graph learning, Factorisation & recommandation, Gestion des données non structurées, Clustering et réduction de la dimension, NLP, text minig, Times series, Mixture Models, Vertex AI GCP.

#### Ingénierie Logiciel et Système Intelligents

10/2019 - 07/2024 Faculté des Sciences et Techniques Tanger, Maroc

Machine Learning, Deep learning, méthodologie de l'Intelligence Artificielle, vision artificielle, data mining, statistique inférentielle, systèmes multi-agents, analyse de données, système décisionnel (BI).

### SOFT SKILLS

#### Soft Skills

- Travail collaboratif efficace démontré dans mes précédents projets.
- Capacité d'adaptation face aux changements de priorités en projet de fin d'étude.
- Esprit de synthèse et de création prouvé au poste d'assistant Data scientist.

### LANGUES

**Français** : Langue maternelle **Anglais** : Professionnel

### RÉSEAUX SOCIAUX

**in** @kpodjro-kpatoukpa-9a5643261

**🐙** <https://github.com/kpatoukpakpodjro>