

# ELHADJI NGOM

Seeking a challenging permanent role as a data scientist where I can contribute on the frontiers of innovation around Artificial Intelligence enabled business trends. I really like creating optimized and cutting edge algorithms to help for decision making.



## CONTACT

✉ ngomel.ehn@gmail.com  
☎ +33 (0)7 81 78 31 33  
📍 75007 Paris (Bureau)  
🏠 personal-webpage-link  
📧 @engom  
🌐 @elhadji-ngom

PERMIS : CATÉGORIE B

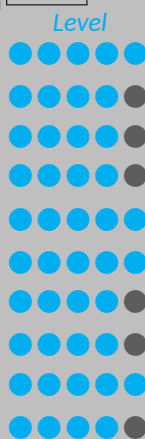
## COMPÉTENCES

### Langages de programmation

Python R SQL Spark LaTeX

### ML-DL frameworks

Numpy Pandas  
Sklearn Mlflow  
TensorFlow Keras  
Pytorch  
Matplotlib  
Streamlit  
spaCy  
Selenium  
BeautifulSoup  
Transformers



### Cloud computing & Versioning

AWS + Docker  
Git-DVC + GitHub



### Data science & AI

**Machine learning** : Regression (Cox/Linear), Classification (kNN, SVM, ...), Clustering (k-means, HC, GMM-EM), PCA, Anomaly detection, Artificial Neural Network (ANN), Deep learning (CNN), NLP & CV.

### Mathématiques appliquées :

Algèbre linéaire, Statistiques, Optimisation, Analyse de survie, Probabilités.

### Data engineering :

Data Wrangling, ETL tools, Hadoop, Web Scraping, Features Engineering.

### Langues

Français  
English

### Niveau



## CENTRES D'INTÉRÊT

♥ Football  
♣ Jeux de société  
❄ Cyclisme

## EXPÉRIENCES

📅 06/2022 - à présent

📍 Paris | OIF

Organisation inter. de la francophonie

Summarization Models (NLP with BERT), Build ML apps with Streamlit, Web Scraping with Selenium, Data warehousing (ETL - SQL) and Dashboard & Putting model into production via UI deployment.

Data scientist | NLP

📅 01/2018 - 08/2021

📍 Versailles |

Ministère de l'Éducation Nationale.

Enseigner les mathématiques en lycée général - Éduquer & Orienter des jeunes gens.

Développer une ingénierie didactique en algorithmique et en programmation avec Python.

Professeur de mathématiques

📅 09/2017 - 12/2017

📍 Grenoble | Ivalis

Mener des études statistiques & Analyser des données relatives aux performances des spécialistes d'inventaire. Créer des tableaux de bord, Gérer & Alimenter la data warehouse au quotidien.

Chargé d'études statistiques

## FORMATIONS

📅 10/2021 - 08/2022

📍 PARIS | Data Science Tech Institute

Big Data & Hadoop Machine Learning Deep Learning Advanced Statistics for ML Optimization  
Time Series Mlops & Devops Semantic Web SQL & NoSQL SAS Cloud AWS Data Law & AI

Applied MSc in Data Science & AI

📅 09/2019 - 06/2021

📍 PARIS | Université de Paris (Paris 7)

Mathématiques fondamentales - Histoire des sciences mathématiques - Didactique des sciences et des algorithmes - Approche didactique de l'algorithmique au secondaire en France.

Master II Didactique des sciences

📅 08/2015 - 06/2016

📍 MBOUR | AIMS-SENEGAL

Statistics - Big data - Machine learning - Problem solving - Operations research - Optimization.

MSc in Mathematical Sciences

## CERTIFICATIONS

📅 01/2022 - 02/2022

📍 Neo4j (GraphAcademy) - NoSQL

Neo4j Fundamentals, Cypher Fundamentals, Indexes and Query Best Practices in Neo4j 4.x., Graph Data Modeling for Neo4j, Importing CSV Data into Neo4j graph database (using Cypher or the Data Importer).

Neo4j Certified Professional

📅 06/2021 - 08/2021

📍 Coursera - By Stanford University

Image Classification (computer vision), Logistic Regression, Neural Networks, Spam filter (NLP), SVM, K-Means, PCA, Recommender Systems.

Machine learning

📅 08/2021 - 09/2021

📍 LinkedIn Learning

Formation basée sur la résolution de challenges: Savoir interagir avec une base de données relationnelle, effectuer des jointures de tables et résoudre des problèmes complexes.

SQL : Défis Code

## PROJETS

**FAREX WEB:** Digitalisation du journal hebdomadaire du service des représentations extérieures (SRE) de l'OIF à l'aide d'outils NLP avec BERT et spaCy (NER). Extraction de textes sur PDF, Classification automatisée du documents puis Déploiement de microservices (UI+DataBas+Admin-UI) en production (tableau de bord).

**Mots-clés:** Classification de documents, Clustering de thématiques, NER, NLP, BERT, spaCy.

**KONVOO ETP:** Mise en place d'un outil d'extraction (Web-scraping) et de traitement des messages des patients postés sur le forum santé *doctissimo.fr* (+12M de commentaires) au sujet du diabète, de covid-long ... Le but du projet est de créer d'abord une base de données de connaissance du type *neo4j* alimentant une AI d'analyse de sentiment et d'émotion avec BERT qui aide le pharmacien à optimiser davantage le parcours des patients.

**Mots-clés:** Labélisation de sentiments et d'émotions, CamemBERT, Attention Mechanism, NLP.