

DHIA HMILA

Senior Data Scientist at AXA France

@ dhiahmila@gmail.com
https://hmladhia.github.io/

+33610201606
hmladhia

92700, Colombes, France
dhia-hmila Dhia Hmila

dhia-hmila



PROFESSIONAL EXPERIENCE

Senior Data Scientist

Data Science Team, AXA France

June 2021 - Present Nanterre, France

- Development of document classification model using Computer Vision (Vision Transformer, VGG16, openCV, ...) to automatically index AXA's Document management system (GED).
- Building a reusable Azure ML Pipeline to automatically retrain an AI System (Yolov5, tesseract, VGG16, Vision Transformer, ...) used for information extraction (Driver's license, National ID card, Passports ...)
- Defining CI (Continuous Integration), CD (Continuous Deployment) & CT (Continuous Training) Templates for the Team's projects.
- Leading AXA's "Python COP" and organising a series of monthly programming talks/events (e.i. testing, Web Scraping, packaging, code quality, ...)
- Leading a series of training sessions/courses on Python, Data Manipulation (pandas & pyspark), Software Engineering best practices for AXA collaborators

Graduate Program - Data Scientist

Data, Fraud, Waste & Abuse Team, AXA France

May 2020 - June 2021 Nanterre, France

- Development of an AI-based Fraud, Waste, and Abuse Detection Tool that acts on various scopes (Claims, Health Providers, Beneficiaries, ...) & Deployment on Databricks
- Creation of various fraud detection models for fraudulent patterns as part of the FWA detection tool.

Data Scientist

Shift Technology

April 2019 - March 2020 Paris, France

- Use of NLP technologies for clustering Claim managers' notes and visualizing results for detecting new patterns of fraud
- Build AI Models for Shift's fraud detection solution & Deployment for a new customer

PROJECTS

nbmanips: Split, merge and convert IPython Notebooks

Jupyter Databricks Zeppelin

May 2021 - Present

An open-source package containing a collection of tools to manipulate Notebooks via a python or CLI.

Spot-Language: Programming Language Detection

ABOUT

I'm a passionate Data Scientist and a programming enthusiast. I've worked on fraud/anomaly detection and document classification problems

EDUCATION

Engineering double degree

ENSTA ParisTech

ENIT

2016 - 2019

France - Tunisia

Specialized in Artificial Intelligence

Preparatory School

IPEIT

Esprit Prépa

2014 - 2015

France - Tunisia

Specialized in Mathematics & Physics Ranked 10th in the entrance exam for engineering schools among 3,000 candidates.

SKILLS

Artificial Intelligence

Anomaly Detection Tensorflow TFX
scikit-learn NLP nltk spaCy Gensim
pyod

Data Wrangling

pandas pyspark SQL ETL

Programming Languages

Python C/C++ C# JavaScript

MLOps

Github Workflows Azure Pipelines Git
Linux Docker Azure ML Databricks

Web Scraping

Classification Flask NLP Model Compression

📅 January 2020

Spot-Language is a classification model to detect the language used in code snippets.

- Construction of a Training dataset from github public repositories.
- Deployment of Web Application to serve as demo of the ML Model.

Dmail

email markdown downloads>100k

📅 April 2020 - May 2020

A python package to write & send emails in markdown format

selenium requests BeautifulSoup4

LANGUAGES

Arabic
English
French



CERTIFICATES



Cloud-Native Development with Open-Shift and Kubernetes Specialization

Coursera

📅 2023 2023



Machine Learning Engineering for Production (MLOps)

Coursera

📅 2022 2022



Machine Learning & General Coding Logic

Triplebyte

📅 2023 2023