

DHIA HMILA

Data Scientist Senior chez AXA France

@ dhiahmila@gmail.com +33610201606 92700, Colombes, France
hmladhia dhia-hmila Dhia Hmila dhia-hmila

hmladhia.github.io



EXPÉRIENCE PROFESSIONNELLE

Data Scientist Senior

Équipe Data Science & IA, AXA France

Jun 2021 - Aujourd'hui Nanterre, France

- Développement de modèles de classification de documents basé sur des technologies de Computer Vision (keras, Vision Transformer, openCV, ...) pour indexer les documents de la GED (Gestion électronique des documents) d'AXA.
- Construction de Pipelines Azure ML réutilisables pour réentraîner automatiquement un système IA (Yolov5, tesseraact, VGG16, Vision Transformer, ...) utilisé pour l'extraction d'informations (Permis, CNI, ...)
- Construction de Templates CI (Continuous Integration), CD (Continuous Deployment) & CT (Continuous Training) pour les projets de l'équipe
- Responsable de la "COP Python" d'AXA et animateur d'événements mensuels de programmation (testing, Web Scraping, packaging, qualité de code)
- Animation d'une formation Python et Manipulation de données en utilisant pandas & pyspark pour des collaborateurs AXA

Graduate Program - Data Scientist

Équipe Data, Fraud, Waste & Abuse, AXA France

Mai 2020 - Juin 2021 Nanterre, France

- Développement d'un outil de détection automatique de fraudes qui agit sur différents périmètres (Prestataires de santé, Bénéficiaires, Procédures médicales, ...) et déploiement sur Databricks.
- Création de différents modèles de détection de fraude pour alimenter l'outil de détection FWA.

Data Scientist

Shift Technology

Avril 2019 - Mars 2020 Paris, France

- Utilisation de technologies NLP pour le clustering des notes de questionnaires de sinistres et création de nouveaux modèles de détection de fraude
- Amélioration et déploiement de la solution de détection de fraude Force de SHIFT pour un nouveau client

PROJETS

nbmanips: Découper, fusionner et convertir vos Notebooks

Jupyter Databricks Zeppelin

Mai 2021 - Aujourd'hui

Un package open-source d'outils de manipulation de Notebooks via des scripts python ou une CLI.

CURSUS

Double diplôme d'ingénieur

ENSTA ParisTech

ENIT

2016 - 2019

France - Tunisia

Spécialisé en Intelligence Artificielle

Cycle préparatoire

IPEIT

Esprit Prépa

2014 - 2015

France - Tunisia

Spécialisé en Math & Physique Rang 10 au concours d'entrée aux écoles d'ingénieurs en Tunisie parmi plus de 3000 candidats.

COMPÉTENCES

Intelligence Artificielle

Anomaly Detection

Tensorflow

TFX

scikit-learn

NLP

nlTK

spaCy

Gensim

pyod

Data Wrangling

pandas

pyspark

SQL

ETL

Langages de programmation

Python

C/C++

C#

JavaScript

MLOps

Github Workflows

Azure Pipelines

Git

Linux

Docker

Azure ML

Databricks

Web Scraping

selenium

requests

BeautifulSoup4

LANGUES

Arabe



Spot-Language: Prédiction du langage de programmation

[Classification](#) [Flask](#) [NLP](#) [Compression de Modèles](#)

 Janvier 2020

Spot-Language est un modèle de classification des langages utilisés dans des sources de code.

- Construction de dataset d'entraînement provenant de repositories github publiques.
- Déploiement d'une application Web pour servir de démo au modèle IA.

Dmail: Envoi d'e-mails au format Markdown

[email](#) [markdown](#) [downloads>100k](#)

 Avril 2020 - Mai 2020

Un package python qui permet d'envoyer des e-mails au format markdown

Anglais
Français



CERTIFICATIONS



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