

AMINE FERDJAOUI

Chef de projet | PhD

@ aminefrjdzr@outlook.com +33 6 44 75 74 20 Paris / Île-de-France aminefrj.github.io linkedin.com/in/aminefrj

EXPÉRIENCES PROFESSIONNELLES

Chef de projet - Lead data Scientist

SogetiLabs - Capgemini (PhD)

2022 - 2024 Issy-les-Moulineaux, Paris

- Responsable de projets internes à SogetiLabs France.
- Pilotage et suivi opérationnel des équipes de développement.
- Définition du besoin et échanges avec les clients pour réaliser les maquettes.
- Développement de modèles de clustering et de recommandation en utilisant des techniques de NLP.

Data Scientist

Orange - Enovacom (Stage de fin d'études)

Mar-Sep 2021 La Défense, Paris

- Développement d'une plateforme de recherche et de visualisation de données non structurées dans un entrepôt de données de santé.
- Collecte, préprocessing et classification des données de santé de Harvard Medical School.

Data Scientist - Full Stack developer

Numendo (Prestations de service pour Orange)

Sep-Dec 2020 Châtillon, Paris

- Maintenance et ré-entraînement de modèles de classification de texte.
- Réalisation d'une plateforme de streaming de flux vidéos rtsp (React et Flask).

Data Scientist - Full Stack developer

Orange SA (Alternance)

Sep 2019 - Août 2020 Châtillon, Paris

- Classification de tickets d'incidents cloud.
- Détection (en temps réel) du port d'équipements de sécurité dans les chantiers.
- Développement et déploiement d'applications web pour utiliser les modèles IA en React et Flask.
- Développement de l'api de la plateforme AI Marketplace en php-GraphQL.

Data Engineer

Saynova (Projet de fin d'études Master 1)

Février-Mai 2019 Nanterre, Paris

- Analyse et visualisation des données avec Elasticsearch et Kibana.
- Collecte et transformation de fichiers de Logs avec Logstash et Beats.

FORMATION

Doctorat CIFRE

Centre Borelli / Capgemini

2022 - 2025 Université Paris Cité

- Sujet de thèse : Inférence de relations causales par apprentissage non-supervisé appliqué au Traitement Automatique du Langage Naturel.

Master 2 Miage

Parcours Valorisation et protection de la donnée de l'entreprise

2020 - 2021 Université de Paris

- Mention Très Bien.

Master 1 & 2

Machine Learning for Data Science

2018 - 2020 Université de Paris

- Mention Bien.

Licence

Informatique Academique Générale

2015 - 2018 USTHB Alger, Algérie

- Majeur de promotion.

COMPÉTENCES

Pilotage et suivi de projet

Gestion des Attentes

Planification Stratégique

Coordination des Équipes

Gestion Agile

Leadership

Choix Technologiques

PROJETS

Application de recommandation de profils inspirants

- Recommandation de profils internes (350k profils) pour aider les collaborateurs à tracer leur trajectoire de carrière et réduire le turnover.
- Collecte et analyse de CV internes.
- Proposition de formations personnalisées.

Application de gestion de projets SogetiLabs

- Outil de suivi et de gestion des projets et des consultants de SogetiLabs.
- Recommandation de profils aux missions en utilisant des techniques de NLP.

Outil d'extraction de données médicales

- Extraction d'entités nomées à partir de dossier patient informatisé à l'aide de NER NLP et modèles LLM.
- Application web d'analyse et exploration de données de santé non structurées du dossier patient informatisé à l'aide de techniques de NLP.

AI Marketplace

- Le premier projet introduisant une place de marché de modèles IA en Europe basé sur Acumos (Linux Foundation et AT&T).
- Utilisation des modèles en mode API ou applications Web.

PUBLICATIONS ET COMMUNICATIONS

Articles scientifiques

- A. Ferdjaoui, S. Affeldt et M. Nadif (2024a). **Une variante pondérée de K-means adaptée aux données textuelles.** In: SFC 2024.
- – (2024b). **Modèles graphiques causaux interactifs pour les données textuelles.** In: Conférence EGC 2024 (Prix du meilleur article scientifique).
- – (2024c). **WordGraph: a python toolkit for reconstructing interactive causal graphical models from text data.** In: Conférence ACM WSDM 2024.
- – (2023a). **Modèles graphiques causaux interactifs pour les données textuelles.** In: Conférence APIA 2023.
- A. Ferdjaoui A. Tlati, S. Affeldt et M. Nadif (2023b). **CORPEX: Analyse exploratoire d'un corpus biomédical à l'aide de la classification croisée.** In: EGC, pp. 597–604.

Communications

- Présentation de l'article: "Une variante pondérée de K-means adaptée aux données textuelles." (2024). In: **Marseille, France: Conférence SFC.**
- Présentation et démonstration de l'article: "WordGraph: a python toolkit for reconstructing interactive causal graphical models from text data." (2024). In: **Mérida, Mexique: WSDM.**
- Démonstration de l'outil: Modèles graphiques causaux interactifs pour les données textuelles. In: **Dijon, France: Conférence EGC 2024 (Prix de la meilleure démonstration).**
- Présentation et démonstration de "CORPEX: Analyse exploratoire d'un corpus biomédical à l'aide de la classification croisée." (2023). In: **Lyon, France: Conférence EGC.**
- Présentation et démonstration de l'article "CAEclust: A Consensus of Autoencoders Representations for Clustering" (2022). In: **Paris, France: Image Processing On Line, pp. 590–603.**

TECHNOLOGIES

Développement web

React, Django, FastAPI, Flask, Figma

Outillage

Git, Docker, Linux, Cloud, Jira, Azure DevOps

Machine Learning

Pytorch, Tensorflow, Scikit-learn

LANGUES

Français

Anglais

Diplome Anglais Pre-advanced

Arabe



AWARDS



Conference EGC 2024

Prix de la meilleure démonstration.

Codingame summer 2023

Première place Sogeti.

Hackathon Sogetilabs fall 2023

Seconde place sujet FFR.

CENTRES D'INTERÊT

Tennis : Pratique depuis 2010

Formation initiateur tennis