



# Islem HABIBI



Data scientist - CDI

+33 (0)6 52 13 82 77

71100 Chalon-sur-Saône

[islem.habibi.upec@gmail.com](mailto:islem.habibi.upec@gmail.com)

Permis B

## A PROPOS

En tant que Data Scientist, je maîtrise l'analyse et l'exploration des données, j'assure aussi le développement, l'amélioration et l'évaluation des modèles de machine learning. Mes compétences en travail d'équipe, associées à mon sens de l'écoute et de la communication, me permettent de m'intégrer activement dans différents environnements.

## COMPETENCES

**Programmation:** Matlab, Python, C#, Git, GitHub, Markdown, Scilab, SQL

**Microsoft Office:** Word, Excel, PowerPoint...

**Natural Language Processing (NLP)**

**PowerBI**

**AWS**

**Project Management (Agile, waterfall)**

**Gestion des conflits**

**Gestion de projet et d'équipe:** Trello,

Ms project, confluence, Slack

**Systèmes d'exploitation:** Windows, Linux ubuntu

## Langues

**Français:** Bilingue

**Anglais:** Courant

## CENTRE D'INTERET

Cinéma

Sport (Taekwondo pendant 6 ans)

Montage vidéo

Jeux de société

Piano (débutant)

## EXPERIENCES PROFESSIONNELLES

**Clinicog: Data Scientist**

09/2022 - Présent

- Développement des outils de diagnostic basée sur l'IA.
- Développement d'outils d'aide à la révision basés sur l'IA pour les élèves.
- Accompagnement et coaching des entreprises pour la maturation de leurs projets, de l'idée au marché.
- Analyse et exploration des données psychométriques
- Natural Language Processing des corpus avec Python
- Rédaction des articles scientifiques en anglais
- Nettoyage des données (formatage, suppression des doublons...).
- Animation des réunions d'équipe
- Gestion et encadrement de l'équipe de gestion et collecte de données
- Construction des données d'entraînement des modèles.
- Amélioration continue des modèles
- Gestion de projet

**Laboratoire Architecture et Fonction des Macromolécules Biologiques (AFMB) -**

**Équipe glycomique:** Ingénieur bioinformatique

02/2021 - 07/2021

- Analyse et nettoyage des données
- Nettoyage et split des fichiers textes

## Projets

- **Analyse des données de systèmes éducatifs**
- **Conception une application au service de la santé publique**
  - Nettoyage et exploration des données
  - Effectuer une analyse statistique multivariée et univariée
  - Représentation graphique des données
- **Anticipation des besoins en consommation de bâtiments**
  - Mise en place du modèle d'apprentissage supervisé adapté au problème, adapter ses hyperparamètres et évaluer ses performances
- **Segmentation des clients d'un site e-commerce**
  - Adaptation et amélioration des hyperparamètres d'un algorithme non supervisé
  - Évaluation des performances d'un modèle d'apprentissage non supervisé
  - Clustering non supervisé
- **Classifiez automatiquement des biens de consommation**
  - Prétraitement des données texte et image
  - Mise en œuvre des techniques de réduction de dimension
  - Élaboration d'un modèle d'apprentissage profond
  - Clustering supervisé
  - Évaluer la performance des modèles
- **Implementez d'un modèle de scoring**
  - Creation et testing des modèle de scoring
  - Creation d'API
  - Déploiement d'API sur le web (Heroku)
  - Création de dashboard avec streamlit
  - Test du drifting des données
- **Déployer un modèle dans le cloud**
  - Prétraitement de données dans le cloud (AWS)
  - Lecture et enregistrement de données sur Amazon S3

## ETUDES / DIPLOMES

2022 - 2024

**OpenClassrooms**

Parcours Data Scientist - diplôme RNCP niveau 7

2020 - 2021

**Faculté de Médecine de Créteil - Université Paris-Est Créteil (UPEC)**

Master Biologie-Santé parcours Bio-ingénierie pour la santé (MBIOS)