



Data Scientist passionné par la Data et IA

Expérience professionnelle

08/2019 - **Data Scientist**, *Quinten Maroc*, (5 ans)

Aujourd'hui Collaborer avec l'équipe DEV pour exploiter la valeur des données et créer des solutions data-driven, nettoyer et structurer les données brutes, développer des modèles prédictifs à l'aide d'algorithmes de machine learning pour anticiper des comportements futurs ou automatiser certaines tâches et faciliter l'intégration de ces modèles dans les pipelines de production.

- Création de script **python** pour **préparer** et **transformer** des données sous format texte ou des données tabulaires
- **Optimisation** des temps de calculs et de consommation des ressources des pipelines de **dataprep** existantes
- **Scraping** de différents fichiers **PDF** et de ressources **Web** pour collecter des **dataset** et créer des bases d'apprentissage pour des modèles de **deep learning**
- Créer des **parsers** pour exécuter des **grammaires** à partir de données textuelles.
- Explorer et concevoir architectures de **modélisation** des données pour créer des modèles **prédictifs**
- Explorer différentes **métriques** pour **évaluer** la performances les modèles de deep learning
- Collaborer avec l'équipe DEV afin d'**intégrer** et **déployer** les modèles dans les projets existants
- Participer aux différentes réunions d'analyse et de conception

02/2024 - **Stagiaire Data Scientist**, *Quinten Maroc*, (6 mois)

07/2024 Développement d'un système de scoring basé sur des techniques de Text Mining pour créer un jeu de données à partir d'un dictionnaire de médicaments, élaborer une fonction pour annoter les données puis entraîner un modèle de machine learning SVM pour la régression sur la criticité des prescriptions médicales.

Éducation

2016 - 2019 **Formation d'ingénieur**, *Ecole nationale d'informatique et d'analyse des systèmes (ENSIAS)*, Rabat
Ingénierie Informatique et Systèmes Embarqués et mobiles

2014 - 2016 **Classe préparatoire aux grandes écoles (CPGE)**, *Lycée Omar Ibn Abdelaziz*, Oujda
MPSI/MP

2013 - 2014 **Baccalauréat**, *Oued Eddahab*, Oujda
Science mathématiques option A

Compétences générales

Data Preparation, Data Visualisation, Machine Learning, Statistical Analysis, Time-Series Forecasting, Classification, Regression, Clustering, Dimensionality Reduction, Natural Language Processing, Text Embeddings, Semantic Search, Text Parsing, Image Processing, Unit Testing, Web Scrapping, Agile Methodologies

Algorithmes

Decision Trees, Random Forest, Support Vector Machines (**SVM**), Linear Regression, Logistic Regression, Gradient Boosting, K-Means, Synthetic Minority Over-sampling Technique(**SMOTE**), Principal Component Analysis (**PCA**), Neural Networks (**CNN**, **RNN**, **LSTM**, **Autoencoders**), Bidirectional Encoder Representations from Transformers(**BERT**)

Compétences techniques

Systèmes d'exploitation: **Linux** (Debian, Ubuntu, Raspbian), **Windows**

Langages de programmation: **Python**, **Java**, **C**

Bases de données: **SQL** (PostgreSQL, MySQL, Oracle DB), **NoSQL** (Mongo DB)

Data Transformation & analysis: **Pandas**, **Numpy**, **PySpark**, **Dask**

Machine Learning: **PyTorch**, **Tensorflow**, **Scikit-learn**, **XGBoost**, **Transformers**

Model Serving: **Flask, TensorFlow Serving**

Data visualisation: **Grafana, Seaborn, Matplotlib**

Indexing: **FAISS, Solr**

Image Processing: **OpenCV, Scikit-image, Pillow/PIL**

Autre compétences: **Docker, Git, Jupyter Notebooks, Selenium, PyParsing, PdfQuery,**

Langues

Arabe **Maternelle**

Anglais **Couramment** TOEFL iBT Novembre 2018 score: 86

Français **Couramment** TCF Octobre 2018 niveau C2

Centre d'intérêt

Natation, Raod trip à Moto.