

PARCOURS ACADÉMIQUE

Master 2 Apprentissage automatique pour la science des données , Université Paris Cité, Paris	2021 - 2022
Modules pertinants: Machines/Deep Learning, statistics, probability.	
Diplôme d'ingénieur , École des Sciences de l'information, Rabat Maroc	2017 - 2020
Modules pertinants: Python, SQL, Information thoery, Science de données, Gestion de connaissances.	
Classes Prépa , École préparatoire AL-KHANSAA, Casablanca Maroc	2015 - 2017
Modules pertinants: Python, SQL, Algèbre Linéaire, Probabilité ...	
Baccalauréat sciences mathématiques Groupe Scholaire Ouhoud, Casablanca Maroc	2014 - 2015

EXPÉRIENCES PROFESSIONNELLES

Data Scientist - CDI Aqsone	09/2022 - Present <i>boulogne-billancourt, France</i>
---------------------------------------	--

- Mission Nexans**, correction d'incohérences dans les données historiques d'achats et de fournisseurs . Enrichissement de données manquantes impactant +440 M€ en matériaux grâce à des méthodes de deduplication et clustering ML.
- Collaboration à la création d'une solution de calcul des coûts numériques pour articles vestimentaires basée sur images et descriptions. Utilisation de réseaux neuronaux convolutifs et transformateurs.
- Collaboration au developpement d'une solution d'approvisionnement 360° avec Python, AWS MySQL et Google Data Studio. Tableaux de bord interactifs incluant prévisions, analyse des dépenses, audit des fournisseurs, émissions de CO2, etc.
- Participation au développement et à la maintenance d'une solution de planification des successions utilisant l'apprentissage automatique et d3js pour les visualisations.

Data Scientist - Stage fin d'étude Laboratoire de recherche Astek - 6 mois	02/2022 - 08/2022 <i>boulogne-billancourt, France</i>
--	--

- Implémentation d'un modèle qui utilise les données tactiles, d'orientation et d'accélération d'un utilisateur pour déterminer s'il est le propriétaire ou un imposteur.
- Développement d'une application avec React et MongoDB qui collecte les données d'entrainement grâce à des API de capteurs intégrés.
- Précision du modèle et détection de téléphones : Le modèle atteint 94% de précision et un score F1 de 88% sur 7 utilisateurs, et peut également identifier les téléphones utilisés par un même utilisateur.

Data Scientist - CDI Alexsys Solutions - 9 mois	12/2020 - 08/2020 <i>Casabalanca, Maroc</i>
---	--

- Mission ONCF (l'office National des Chemins de Fer du Maroc)**. Implémentation d'un modèle de classification et de reconnaissance optique de caractères (OCR) en utilisant YOLOv4 sur des scans de documents juridiques en arabe. Le système développé réduit le temps de traitement de chaque document à la durée de détection, permettant la détection en lot de 1000 documents en 5 minutes (3,33 documents par seconde).
- Mission TotalEnergies**, Développement d'une application de détection et lecture de plaque d'immatriculation pour les camions de distribution de gaz naturel. Ce système a réduit le temps d'accueil des camions entrant dans le centre de distribution de gaz par 2 minutes et 53 secondes en moyenne.

PROJETS PERSONNELS

Réduction de la dimensionnalité/regroupement de données provenant d'articles scientifiques/de résumés de wikipedia/de titres de journaux télévisés

- Utilisation des méthodes Tandem: Technique de réduction de la dimensionnalité (PCA, t-SNE, UMAP, Autoencodeurs) suivi par chacune des techniques de clustering suivantes: Kmeans, Spherical Kmeans, factorial Kmeans, clustering hiérarchique (WARD, Complete, Linkage et Single Metrics), et HDBSCAN

Mise en œuvre d'un réseau neuronal pour l'apprentissage semi-supervisé afin de prédire les données MNIST.

- Dans ce projet je participe à Implémentation d'un réseau de neurones pour l'apprentissage semi-supervisé pour prédire les données MNIST (en utilisant seulement 100 points le modèle atteint 95% accuracy)

Application Streamlit ML qui contient 3 modèles pour aider à prédire le diabète, la maladie de Parkinson et les maladies cardiaques sur la base de valeurs de formulaire.

- Cette application utilise des classificateurs binaires tels que la régression logistique et le SVM pour prédire avec précision ces trois types de maladies. Déploiement sur AWS.

COMPÉTENCES

Machine Learning: Ensemble methods (Bagging, Boosting) / kernel methods (SVM, SPCA), Deep Learning, Natural Language Processing, Optimisation

Data Science Data science pipeline (cleansing, wrangling,visualization,modeling, interpretation),Git.

Software Engineering: Développement en Python (Streamlit / Flask), SQL, AWS, and Docker.

Base de donnée: Postgres, MongoDB, MySQL (on Premise and Cloud)

Data Visualization: PowerBi, Qlicksense, D3js.

Transfer Learnings: LLMS, CNNs, Transformers ...

CERTIFICATIONS ET COURS

- Improving Deep Neural Networks : Hyperparameter Tuning, Regularization and Optimization**
- Machine Learning : Université Stanford**
- SQL for Data science: Université de Californie à Davis**
- Neural Networks and Deep Learning**
- NLP - Natural Language Processing with Python**