

COMPAORE Yolemba Harold

+212-658-809976
haroldcompaore07@gmail.com
<https://github.com/202422> 
<https://www.linkedin.com/in/harold18/> 

COMPÉTENCES

- **Programmation** - python, java, SQL, R, C, php
- **Analyse statistique** - Analyse exploratoire de données, Tests statistiques
- **Machine learning & Deep learning** - Sklearn, Tensorflow, keras, HDBSCAN
- **Visualisation de données** - Tableau, Matplotlib, seaborn, plotly express
- **Développement Web** - J2EE, html, css, Eclipse, MySQL, PHP, LARAVEL
- **Compétences comportementales** - Communication en français (langue maternelle) et anglais (Moyen); Problem-Solving; Méthodologie Agile (SCRUM); Travail en équipe; Adaptabilité; Esprit critique

EXPERIENCE

Stagiaire Data scientist: Prédiction du statut opérationnel des Startup India (A distance)

Technocolabs Software

2024.06-2024.08

- Prétraitement des données : suppression des informations redondantes et non pertinentes, gestion des données manquantes et des valeurs aberrantes, transformation des variables de type date
- Analyse exploratoire des données: analyse univariée, bivariée et multivariée
- Feature engineering: sélection de variables, transformation logarithmique et standardisation, création de nouvelles variables, encodage des variables
- Modélisation : suréchantillonnage, classification binaire et classification multiclasse
- Détails du projet disponibles ici: <https://github.com/202422/Startup-Operational-Status-prediction-using-Machine-learning>

Stagiaire de recherche en Machine et Deep Learning: Imputation de séries temporelles Rabat, Morocco

ENSIAS

2024.06-2024.09

- Implémentation d'un algorithme pour convertir les séries temporelles en séquences de taille fixe (10) afin de le préparer à l'apprentissage supervisé.
- Conception d'un algorithme pour analyser systématiquement les séries temporelles afin de détecter toutes les occurrences de données manquantes et entraîner un modèle pour imputer ces valeurs manquantes.
- Entraînement de modèles de machine learning/deep learning tels que SVR, RNN et LSTM sur l'ensemble d'entraînement et comparaison de leurs performances respectives.
- Outils: Numpy, Pandas, sklearn, Tensorflow, metrics
- Détails du projet disponibles ici: https://github.com/202422/Time_series_Imputation_using_ML-DL

FORMATION

École nationale supérieure d'informatique et d'analyse des systèmes (ENSIAS)

Diplôme d'Ingénieur

2023.09 - En cours

Faculté de sciences/Université Abdelmalek ESSAADI Tétouan

DEUG/Mathématiques et informatique

2021.10 - 2023.06

PROJETS

- **Analyse de l'attrition des employés et prédictions de leur départ:** Ce projet vise à fournir des informations sur les facteurs influençant l'attrition des employés et à prédire quels employés sont susceptibles de quitter l'entreprise à travers des analyses statistiques et du machine learning: <https://github.com/202422/Employee-Attrition-Analysis-and-Predictions>
- **Optimisation des hyperparamètres - Prédiction du prix des maisons et survie sur le Titanic:** Ce travail couvre l'ensemble du workflow du machine learning en mettant l'accent sur l'optimisation des hyperparamètres en utilisant des techniques telles que le Zooming In et l'optimisation bayésienne: <https://github.com/202422/Hyperparameter-tuning-House-Price-and-Titanic-Survival-Prediction>

CERTIFICATIONS

Machine Learning with Python — IBM

Mai 2024

Exploratory Data Analysis for Machine Learning — IBM

Août 2024