

Alexandre Chen

Data Scientist - Stage

+33 7 63 31 54 95 | chenalexx@hotmail.fr | [LinkedIn](#) | Paris

En tant qu'aspirant Data Scientist, je suis à la recherche d'un contrat alternant pour mettre en pratique mes aptitudes d'analyse et de modélisation. Avec une solide expérience en Data Science, je souhaite contribuer à l'efficacité de vos équipes grâce à mes compétences en programmation et mon engagement pour résoudre des problèmes complexes.

Expériences professionnelles

Metro France

92000, Nanterre

Data Scientist – Alternance

2022 – 2024

- **Analyse Ad Hoc :**
 - Collaborer avec les équipes pour collecter et analyser des données, concevant des stratégies axées sur les données qui ont accru l'efficacité opérationnelle.
 - Présenter des visualisations de données et des rapports détaillés aux parties prenantes, traduisant des résultats complexes en insights actionnables et en recommandations pertinentes.
 - Assurer la définition du problème, la collecte des données, l'analyse statistique, l'interprétation des résultats et la présentation des conclusions.
 - Outils : GCP (Google Cloud Platform) – SQL – Python – Power BI – Excel
- **Automatisation :**
 - Analyse Ad Hoc automatisée sous GitHub : Transformer des calculs complexes en un rapport non technique clair et accessible de tous, présenté dans un fichier Excel.
 - Outils : GitHub – Excel – GCP – SQL – Python
- **Machine Learning :**
 - Développer et mettre en œuvre des modèles avancés (classification, NLP, clustering) pour améliorer l'analyse prédictive, optimiser la prise de décision et classer les clients selon leurs consommations en magasin, raison sociale et base INSEE.
 - Outils : GCP – SQL – Python (packages : scikit-learn, Matplotlib, Seaborn, Plotly, Pandas, NumPy)

Projets

- **Text analysis :**
 - Développer un modèle NLP pour catégoriser les titres de livres, en utilisant des techniques de traitement de texte et des embeddings, classer les titres par genre, thème ou public ciblé.
Ce projet a automatisé la catégorisation, améliorant l'efficacité du catalogage et des recommandations.
- **Time series forecasting :**
 - Analyser des données historiques sur la consommation de charbon et utilisé des modèles ARIMA et SARIMA pour prédire les tendances futures.
En optimisant les données, j'ai amélioré la précision des prévisions, permettant par exemple de prédire une baisse de la demande de charbon et de soutenir des décisions stratégiques pour la transition énergétique.

Formations académiques

MASTER Informatique - Machine Learning pour la Science des Données

75006, Paris

Université Paris Cité

2024 – 2026

- Deep Learning, Text-mining, NLP, Traitement d'image, Détection d'anomalies, Systèmes de recommandation, Optimisation combinatoire, Apprentissage machine

BUT SD (Ex-STID) - Parcours : Exploration et modélisation statistique

75016, Paris

IUT de Paris – Rives de Seine (Université Paris Cité)

2021 – 2024

- Statistique, Informatique, Décisionnel, Mathématiques appliquées

Compétences

Langues : Anglais niveau B2, Chinois niveau B1

Outils Informatiques

- **Langage :** Python, R, SAS, SQL, NoSQL, Bash
- **Logiciels :** GCP, Bigquery, VSCode, Docker, Github, Kubernetes
- **BI :** Tableau, PowerBI, Excel/VBA, Shiny, Streamlit

Centres d'intérêts

Voyage, IA, E-sport, Natation