

Qi Wang

34 avenue de Lingenfeld
77200 Torcy, France
☎ +33 (0)6 47 36 96 20
✉ wq2021@outlook.com
📄 wq2021.github.io

*Titulaire du double master en Data Science et NLP. Curieuse, rigoureuse avec l'esprit d'organisation et d'analyse acquis à travers le parcours professionnel et académique.
Je recherche une première opportunité en tant que Data Scientist / Data Analyst.*

Études et formation

- 2021–2022 **Master Machine Learning for Data Science**, Université Paris Cité (ex. Université Paris-Descartes), Mention Bien, Supervised Learning, Non-Supervised Learning (Dimensionality Reduction, Clustering, Co-Clustering), Deep Learning, Reinforcement Learning, Text Mining, NLP, Big Data, Recommender Systems.
- 2019–2021 **Master Traitement Automatique des Langues (TAL/NLP) parcours Ingénierie Multilingue**, Inalco, Mention Bien, Algorithmique, Traitement de Corpus, Text Mining, Documents Structurés, Ingénierie des Connaissances, Linguistique Fondamentale.

Expériences professionnelles

- Février à **Stagiaire en IA/Data Science**, Astek, Boulogne-Billancourt, France.
- Août 2022
- Construire un système de recommandation d'articles techniques français personnalisé aux spécialistes du domaine en utilisant les techniques récentes en Deep Learning.
 - Proposer un protocole d'évaluation au niveau de la pertinence ainsi que le classement des résultats pour le système construit.
 - Synthétiser les travaux à travers la capitalisation documentaire.
 - Outils utilisés : GitHub, PyTorch, Tensorflow, Pandas, Selenium, etc.
- Avril à **Stagiaire en Informatique Linguistique**, BNP Paribas, Paris, France.
- Octobre 2021
- Détecter les faits d'actualité parmi un ensemble d'articles de presse du domaine économique et financier à partir d'une veille documentaire interne.
 - Implémenter et comparer trois différentes approches respectivement sur les propriétés lexicales, la statistique du corpus ainsi que les réseaux de neurones profonds.
 - Outils utilisés : Pandas, SpaCy, SBERT, lxml, etc.

Projets universitaires

- Construction d'un modèle de Scoring à l'aide des algorithmes supervisés appliqués aux données bancaires et relationnelles (Pandas, Scikit-Learn, Numpy, SciPy, Matplotlib, etc).
- Analyse des métadonnées des articles scientifiques avec la classification non supervisée (Pandas, Scikit-Learn, Numpy, SciPy, Seaborn, NLTK, etc).

Compétences informatiques et outils

Maîtrise	Python, SQL, HTML5, CSS3, XML, MS Office, iWork
Intermédiaire	R, Shell, Git, \LaTeX
Notions	Java, PHP, JavaScript

Langues

Natif	Mandarin
Avancé	Français (DALF C2), Anglais (B2)