



Récemment diplômé, mes expériences m'ont permis d'acquérir des compétences et une expérience en tant que Data Scientist ou ML Engineer. Je souhaite donc poursuivre cette trajectoire et évoluer dans ce rôle.

Disponible à partir du 16/09/2024.

- ✉ hrywatiotienne@gmail.com
- 🏠 Mons-en-Barœul (59)
- 📅 23 ans
- 📄 Permis B
- 🚗 Véhicule personnel
- 📞 06 52 87 79 30

## Compétences

**Anglais : Intermédiaire**

**Langages de programmation**

- Python (numpy, matplotlib, pandas, Scikit-learn, pytorch, vLLM)
- R
- C++
- Java
- SQL

**Outil informatique**

- Git
- Docker
- Office
- GCP

**Retrouvez tous les projets ici :**

**GitHub:**  
<https://github.com/WatiotienneH>

## Centres d'intérêt

**Jeux de société et stratégie**

**Sport** Musculation, Natation, Tennis de table

**Association** Membre actif d'une association de jeu ancien : La butte Carnière

## Expériences professionnelles

**Data Scientist & ML engineer Stagiaire ( 6 mois )**  
**Entreprise : Picomto, Lieu : La Madeleine**

De mars 2024 à septembre 2024

**Développement d'un LLM et de plugins spécialisés :**

- Implémentation d'outils spécialisés disponibles par le LLM via appel de fonction
- Interaction avec des endpoints de l'application.

**Optimisation et évaluation du LLM :**

- Techniques de Retrieval-Augmented Generation (RAG) , Fine-tuning, Utilisation de vLLM
- Création de benchmarks

**Développement de modèles de vision (Détection d'objets) :**

- Création de données annotées
- Fine-tuning des modèles

**Compétences :**

- Organisation de projets, animation de réunions
- Rigueur dans le développement et la documentation, analyse d'articles scientifiques
- Communication et rédaction de rapports techniques.

**Outils et technologies :**

- Python : HuggingFace, PyTorch, TensorFlow, vLLM, Scikit-learn, Matplotlib
- GitHub, Docker, Google Cloud Platform (GCP)

**Projet d'entreprise : En partenariat avec le groupe VitamineT** De septembre 2023 à février 2024

**Objectifs :** Mise en évidence des patterns sur le parcours des individus et prédiction du statut de sortie

**Actions réalisées :**

- Analyse descriptive, identification d'outliers
- Nettoyage des données (suppression des doublons, traitement spécifique des colonnes de texte libre, ex : analyse de polarité, extraction de mots-clés)
- Exploitation des données, dont certaines non structurées, pour identifier des patterns
- Création d'indicateurs clés
- Algorithme de classification(Random Forest) et de clustering (K-means)

**Outils utilisés :**

Python (Keras, scikit-learn, matplotlib, seaborn)

## Projets

- **Travail de recherche encadré :** "Régression Bayésienne : Des Fondements Théoriques à l'Application Climatique" . *Modèle de régression linéaire, bayésien semi-paramétrique, Algorithme de Gibbs en R*
- **Optimisation de la supply chain :** Modélisation du problème et linéarisation des contraintes. Résolution à l'aide du simplexe et étude des optimisations.
- **Projet d'application machine learning :** Classification de statut d'un logicielle (malware/goodware). Étude des outliers, implémentation d'algorithmes de classification (Random Forest, SVM, Régression logistique), rééchantillonnage des données, optimisation des hyperparamètres.

## Diplômes et Formations

**Master : Ingénierie statistiques et numérique - Data science**  
**Université de Lille, Villeneuve-d'Ascq**

De 2022 à 2024

Compétences :

- **Apprentissage autonome :** apprentissage par renforcement, Q-learning, SARSA, réseaux de neurones, Actor-Critic, pytorch .
- **Machine learning :** régression logistique, k plus proches voisins, CART, Random forest, SVM, bagging, boosting, k-means, classification ascendante hiérarchique.
- **Statistiques Spatiales, Séries temporelles**
- **Recherche opérationnelle**
- **Outils informatiques pour les big data**
- **Probabilités, statistiques et modèles statistiques**
- **analyse factorielle**
- **Anglais**

**Licence Mathématiques - option renforcé recherche**  
**Université de Lille, Villeneuve-d'Ascq**

De 2019 à 2022

**Classe Préparatoire BCPST**  
**Lycée Albert Chatelet , Douai**

De 2018 à 2019