

Adrien SKRZYPCZAK

Ingénieur en formation, stagiaire Analyste Data/IA chez AI Builders (Paris).
En maîtrise IA & Data, à la croisée du NLP, du Machine Learning et de l'IA générative.

📍 Paris + Hauts-de-France - ✉ adrien.skrzyp@gmail.com - 🌐 adrien-skr.dev
🌐 linkedin.com/in/adrienskrzypczak - 🐙 github.com/AdrienSkr

EXPÉRIENCE PROFESSIONNELLE

Stagiaire Analyste Data/IA, AI Builders – Paris Sep. 2026 – Fév. 2027

- Veille sur les LLMs et l'IA agentique, comparaison de solutions Data/IA et développement de POCs.

Stagiaire en IA Générative, Nicolaudie Group – Montpellier Juin 2025 – Août 2025

- Concevoir un système de transformation du langage naturel en configurations JSON en combinant extraction d'entités, RAG et reclassement avec FAISS, LangChain et OpenAI.
- Industrialiser une API REST et prototyper une interface Streamlit, conteneurisées avec Docker.
- Constituer et annoter un jeu de données d'entraînement et de test avec Label Studio, puis générer des prédictions initiales pour accélérer le processus d'annotation.

FORMATION

Université de Sherbrooke, Maîtrise en IA et en sciences des données – Sherbrooke, Canada Août 2025 – Août 2026

- Machine Learning et méthodologie en Python : apprentissage supervisé et non supervisé, prétraitement, validation croisée et optimisation de modèles.
- Deep Learning, NLP et IA avancée : réseaux de neurones (CNN, RNN, LSTM, GRU), Transformers, modèles de langage, apprentissage par renforcement, recommandation.

École d'Ingénieurs du Littoral Côte d'Opale, Informatique – Calais Sep. 2023 – Sep. 2026

- Base solide en traitement du signal, image, IA et statistiques appliquées.
- Gestion de projets d'équipe en contexte académique et associatif, en méthodologie agile.

PROJETS

SemEval 2025 Task 9 – Détection explicable de risques alimentaires Jan. 2026 – Avr. 2026

- Concevoir un modèle NLP multi-tâches basé sur RoBERTa-large pour classifier des risques alimentaires et extraire des segments textuels justificatifs via un étiquetage BIO.
- Mettre en place une architecture à deux branches Hazard/Product combinant têtes de classification, extraction de segments et pondération automatique des pertes.
- Orchestrer une stratégie d'annotation BIO par ensemble de LLM (Gemini, DeepSeek, Qwen) via OpenRouter, avec soft labeling pour gérer la subjectivité des justifications.
- Améliorer la baseline BERT de 0,71 à 0,79 en F1 macro Hazard, avec explications extractives évaluées par suffisance et complétude.

Détection d'astéroïdes dangereux (Python, scikit-learn) Déc. 2025

- Prétraiter 90 000 objets géocroiseurs NASA en réalisant nettoyage, sélection de variables, normalisation et gestion du déséquilibre des classes, puis optimiser les modèles en privilégiant le rappel afin d'atteindre environ 99% sur la classe « astéroïde dangereux ».
- Concevoir une architecture Python orientée objet pour comparer six classifieurs et produire des visualisations améliorant l'interprétabilité des modèles.

COMPÉTENCES ET INTÉRÊTS

- **Compétences techniques** : Python, PyTorch, scikit-learn, Hugging Face, NLP / LLMs, RAG (LangChain, FAISS), API REST, Docker, Streamlit, Git.
- **Compétences humaines** : Curiosité, esprit critique, rigueur, adaptabilité.
- **Langues** : Anglais (B2, TOEIC 795/990), espagnol (A2), italien (A1).
- **Intérêts** : Pratiquer régulièrement la course à pied.