



Lucas POTIN

Data Scientist | Graphes et Modélisation

Ingénieur doctorant en data science, spécialisé en fouille de motifs et classification de graphes. J'ai conçu des modèles interprétables pour détecter des comportements atypiques dans des données hétérogènes (marchés publics, sessions web, réseaux biologiques). Je cherche un poste de Data Scientist pour développer des modèles robustes et explicables sur des données tabulaires ou relationnelles complexes.

🏠 Rouen
📄 Permis B
✉️ lpotin888@gmail.com
☎️ +33 6 51 71 01 81

Portfolio

<https://lucaspotin98.github.io/>

Projets :

- **DramaPang** : Classification supervisée de pièces de théâtre.
- **API de sirétisation** : Service FastAPI pour retrouver des SIRET.
- **Dashboard de ventes** : Application Streamlit de prévision des ventes.

Compétences

Mathématiques & Modélisation

Statistiques, Probabilités, Optimisation (linéaire, combinatoire), Théorie des graphes.

Langages

Python, SQL, C++, CUDA.

Traitement et manipulation de données

Pandas, NumPy, SciPy.

Graphes et classification

NetworkX, PyTorch Geometric, Fouille de motifs.

Machine Learning et Deep Learning

Scikit-learn, XGBoost, PyTorch.

Déploiement et APIs

Streamlit (interfaces), FastAPI (APIs).

Environnement technique

Git, LaTeX, Bash, Linux.

Savoir-être

Adaptabilité, Autonomie, Créativité, Travail d'équipe, Vulgarisation.

Langues

Français

Langue maternelle.

Anglais

Courant (C1) : 970/990 au TOEIC.

Centres d'intérêt

Tennis

Classement 15/2.

Bridge

Classement : 3e série trèfle.

Expériences professionnelles

Thèse de Doctorat

Depuis 2021 Laboratoire d'Informatique d'Avignon (Avignon)

- Sujet : Détection de corruption dans les marchés publics à l'aide de réseaux complexes.
- Conception de deux bases open data sur les marchés publics français (FOPPA et BeauAMP), intégrant plus de **300 000** entreprises sirétisées avec **80 % de réussite**, à partir de données hétérogènes et bruitées.
- Développement d'un framework de classification de graphes basé sur la fouille de motifs fréquents, atteignant un **F1-score de 95 %** sur des benchmarks publics et des cas réels de détection de fraude.
- Obtention du **Prix Science Ouverte 2024** pour la base BeauAMP.
- Technologies : Python (Pandas, NetworkX, Scikit-learn, XGBoost), SQL, Git.

Stage Ingénieur

De mars 2020 à septembre 2020 Aubay (Boulogne-Billancourt)

- Amélioration d'un générateur de sessions web synthétiques, en modélisant **3 types d'agents** utilisateurs (aléatoire, rationnel, récurrent) et **2 profils** (client, candidat).
- Conception et entraînement des modèles de classification et de prédiction comportementale sur sessions de navigation simulées, via des réseaux LSTM ; obtention d'une **précision de 94,5 %** pour l'identification du profil utilisateur.
- Technologies : Python (Pandas, Keras), méthodologie agile (Kanban).

Assistant Chercheur

De juin 2018 à septembre 2018 James Hutton Institute (Dundee, Écosse)

- Implémentation d'un algorithme de modélisation cellulaire parallèle en C++/CUDA pour simuler des processus biologiques sur GPU.
- Réduction du temps d'exécution par un **facteur de 2 à 50**, selon les cas, via des optimisations bas niveau ciblées.
- Collaboration avec une **équipe de chercheurs internationaux**, dans un environnement interdisciplinaire (informatique, agronomie, biologie).
- Technologies : C++, CUDA, Linux.

Diplômes et Formations

Doctorat en Informatique

De 2021 à 2025 Laboratoire Informatique d'Avignon (Avignon)

Formation Ingénieur option Génie Mathématique

De 2015 à 2020 INSA Rouen (Saint-Étienne-du-Rouvray)

Master Sciences et Ingénierie des Données

De 2019 à 2020 Université de Rouen (Saint-Étienne-du-Rouvray)

Baccalauréat Scientifique, mention européenne allemande

2015 Lycée Georges Brassens (Neufchâtel-en-Bray)

Liens externes

- [GitHub : LucasPotin98](#)
- [LinkedIn : lucas-potin](#)
- [Google Scholar : Lucas Potin](#)