

FLORENT LIN

📍 Paris, France ☎ +33 7 81 63 66 63 📩 1florentlin@gmail.com

🔗 github.com/flin1 💬 linkedin.com/in/florentlin

FORMATION

- ENSAE Paris - IP Paris** Palaiseau
Diplôme d'Ingénieur (Grade Master) - Data Science & Mathématiques Appliquées Sept. 2021 - Juin 2024
- **Majeures** : Probabilités Avancées, Statistique Bayésienne, Transport Optimal, Séries Temporelles.
 - **Spécialisation IA** : Deep Learning, NLP, Apprentissage en ligne, Statistique en grande dimension.

- Lycée Saint-Louis** Paris
CPGE - PC (Préparation intensive aux concours Mathématiques & Physique) Sept. 2019 - Juin 2021

EXPÉRIENCE PROFESSIONNELLE

- Aubay Data & AI (Client : SNCF Réseau)** Boulogne-Billancourt
Data Scientist Consultant - Maintenance Prédictive Janv. 2025 - Présent
- **Production ML** : Déploiement de modèles prédictifs pour les défauts ferroviaires (XGBoost/Random Forest).
 - **Ingénierie** : Construction de pipelines ETL robustes pour données géospatiales et temporelles multi-sources.
 - **Optimisation** : Ajustement des modèles (Recall/F1-Score) pour la sécurité ; SHAP pour l'interprétabilité.
 - **Aide à la Décision** : Conception de tableaux de bord et interprétation d'indicateurs en KPIs opérationnels.

- Aubay Data & AI** Boulogne-Billancourt
Data Scientist R&D (Stage) - Machine Unlearning Août 2024 - Déc. 2024
- **Recherche** : Étude de l'état de l'art sur le Machine Unlearning (Concept Erasure, DPO).
 - **Implémentation** : Fine-tuning de **LLaMA** (LoRA) et entraînement de classifieurs RoBERTa pour le dés-apprentissage de connaissances.
 - **Métriques** : Benchmark efficacité vs dégradation via métriques spécifiques (Truth Ratio, Prob. Cond.).
 - **Innovation** : Développement de techniques cross-lingues pour prévenir les fuites d'information.

- Deloitte Conseil AIM (via ENSAE)** Paris
Data Scientist - Recherche Opérationnelle Nov. 2022 - Mai 2023
- **Modélisation** : Formalisation de contraintes de planning infirmier en Programmation Linéaire (MIP).
 - **Algorithmes** : Développement de solveurs de programmation par contraintes pour le routage multi-objectifs.

PROJETS SÉLECTIONNÉS

- Analyse Documentaire Multi-modale (RAG)** 🔗 lien
Client : MEWE Architects Avr. 2025 - Présent
- Construction d'un workflow RAG en production (*Mistral OCR + Gemini*) pour des documents complexes.
 - Stratégies d'injection contextuelles pour reconstruire la mise en page et améliorer la précision du retrieval.
 - Implémentation Pydantic pour générer des sorties structurées déterministes depuis la vision.

- Matching Lead-Franchise** 🔗 lien
Client : Franserve Consultants Juil. 2025 - Présent
- Architecture d'un système de recherche sémantique (Embeddings) pour le matching automatique.
 - Pipeline Scraper-vers-DB (Supabase) intégrant les logs CRM (GoHighLevel) pour enrichissement temps réel.

- **Bandits Manchots** : Benchmark d'algos UCB/Thompson Sampling en apprentissage en ligne.
- **Statistique Bayésienne** : Gibbs Sampling et Lasso Bayésien pour la sélection de variables.

COMPÉTENCES TECHNIQUES

Deep Learning	Python, PyTorch, JAX , HuggingFace, Scikit-learn
Ingénierie	SQL (PostgreSQL), AWS, Docker, Git, Firebase, Supabase
Outils Dev	TypeScript, Golang, C, R, Cursor (AI-assisted dev)
Langues	Français (Maternel), Anglais (C1+), Chinois (B2)

FORMATION CONTINUE

Backend Engineering (Python & TypeScript)

Boot.dev

 *lien*

Sept. 2025 - Présent

- **Informatique Fondamentale** : Complexité (Big-O), Gestion mémoire, POO et Programmation Dynamique.
- **Systèmes** : Serveurs HTTP scalables, Architectures Pub/Sub, Conteneurisation Docker.
- **RAG from Scratch** : Implémentation de pipelines RAG complets sans frameworks.
- **Algorithmes de Recherche** : Construction de systèmes TF-IDF, BM25 et Recherche Hybride.