

Du notebook à la production : déployer un modèle de Machine Learning comme en entreprise

Intervention – Master TRIED

 Mercredi 28 janvier 2026

1. Présentation de l'intervention

Cette intervention propose une immersion concrète dans le cycle de vie **industriel** d'un modèle de Machine Learning, depuis son développement expérimental dans un notebook jusqu'à son déploiement en production, tel qu'il est réalisé en entreprise.

À travers des exemples pratiques et des outils standards du marché, les étudiants découvriront les bonnes pratiques **MLOps** qui permettent de rendre un modèle **tracable, reproductible, déployable et maintenable**.

L'accent sera mis sur la transition entre la data science académique et les contraintes industrielles : mise en production, services d'inférence, déploiement cloud et exploitation.

 **6h ensemble dont 3h de TP : ordinateurs chargés à 100 %, neurones aussi (enfin on va essayer).**

2. Objectifs pédagogiques

À l'issue de l'intervention, l'étudiant sera capable de :

- Comprendre le cycle de vie industriel d'un modèle de Machine Learning
- Logger un modèle, ses métriques et ses paramètres avec MLflow
- Transformer un modèle ML en service d'inférence à l'aide de FastAPI
- Comprendre comment ce service peut être déployé en production
- Bonus: Identifier les principales options de déploiement cloud utilisées en entreprise

3. Stack technique

- **uv** (pour créer l'environnement du projet)
- **MLflow** (tracking, modèles, reproductibilité)
- **FastAPI** (API d'inférence)
- **Docker** (pour servir un modèle ML)
- Balayage de notions de MLOps (**CI/CD, airflow, ..**)
- Balayage de notions cloud (exemples sur **GCP**)

 **Intervenante: Carole Ibrahim** Data Scientist au pass Culture

- Linkedin: <https://www.linkedin.com/in/carole-ibrahim-049883160/>
- Repo équipe Data du pass Culture: <https://github.com/pass-culture/data-gcp>