

## Note technique présentant une étude sur les approches et outils MLOps

### 1. Pipeline de codage des steps d'élaboration du modèle

La mise en place d'un pipeline de codage standardisé permet d'assurer la cohérence, la reproductibilité, et la gestion de toutes les étapes du processus, de la préparation des données à l'élaboration des modèles.

Outils :

Kedro : intègre des pipelines de données, standardise le processus

MLFlow : assure suivi d'exécution, la reproduction des expériences, et la gestion des modèles

### 2. Suivi de performance

Model Drift: Changements graduels dans la distribution des données affectant la performance du modèle.

Data Drift: Modifications dans les caractéristiques des données impactant la qualité des prédictions.

Concept Drift: Évolution des relations entre les entrées et les sorties du modèle au fil du temps.

Outil :

EvidentlyAI: Tableaux de bord pour surveiller les dérives de modèle et de données.

Prometheus: Système de surveillance et d'alerte pour suivre les métriques de performance.

Popmon: Outil pour la détection et le suivi des dérives de données.

Pour le suivi de performance, utilisez des métriques telles que le Jaccard score, l'accuracy, ...

Etablir un seuil déclenchant une alerte en cas de variation importante des résultats en se basant sur le Jaccard score (plus adapté pour de la prédiction de tags).