RER – Monitoring & Testing

**Contexte :**

Identifier et comparer des solutions de monitoring pour la garantie de la qualité et d'effectuer des tests unitaires. Intégrer une solution de solution de monitoring et testing d’un modèle de machine learning. Le monitoring est aussi utilisé pour corriger et détecter des éventuels disfonctionnements.

# **Problématique(s) :**

Comment intégrer un outil de monitoring pour un modèle ? Comment et pourquoi effectuer des tests unitaires ?

# **Mots clés :**

* **Testing**

Le testing est l’ensemble des activités qui permettent de vérifier que les fonctionnalités d’un système ou d’une application répondent aux besoins exprimés dans les spécifications et les exigences. Il est utilisé pour détecter les bugs, les erreurs et les incohérences dans le système ou l’application respecte les normes et les standards en vigueur.

Il existe plusieurs types de testing :

1. Tests unitaires
2. Tests d’intégration
3. Tests fonctionnels
4. Tests de performance
5. Tests de sécurité

* **Test unitaire**

L’objectif est de tester une unité de code à un niveau de granularité élevé (fonctions, méthodes).

* **Fixture**

Un état fixe des objets utilisés comme base pour l'exécution de tests, garantissant un environnement de test cohérent.

Les fixtures sont utilisées pour garantir un contexte spécifique avant l'exécution des tests, assurant ainsi des conditions reproductibles et fiables pour évaluer le comportement d’une fonction.

* **Mock**

En programmation orientée objet, les mocks (*simulacres* ou *mock object*) sont des objets simulés qui reproduisent le comportement d'objets réels de manière contrôlée.

* **Assertion**

Instruction python permettant de vérifier une condition et de lever une exception de type “Assertion Error” lorsque l’évaluation de la fonction retourne la valeur False.

Les assertions peuvent être désactivées dans python et il ne faut pas les utiliser pour vérifier des valeurs. Elles sont utilisées pour le testing. Une seconde expression peut être passée à l’instruction, elle sera retournée dans l’erreur AssertionError si elle est levée.

* **Unittest**

Bibliothèque built-in de Python qui permet de réaliser des tests unitaires.

* **Pytest**

pytest est une bibliothèque de test pour Python qui offre une approche flexible et puissante pour écrire et exécuter des tests. Elle étend et améliore le module de test intégré (unittest) de la bibliothèque standard de Python en fournissant une syntaxe concise et expressive pour la création de tests.

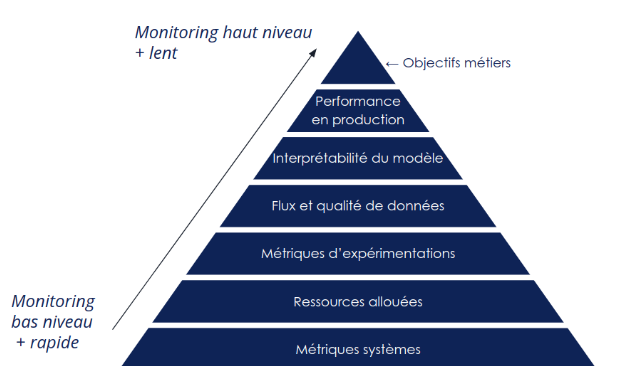
Elle fait des tests unitaires, des tests d’intégration et des tests fonctionnels.

* **Monitoring**

Action de conserver un historique des métriques afin de créer des insights (agrégation, classification, résumés) et d’en rendre compte sous forme de visualisation.

Il y a plusieurs niveaux de monitoring :

1. Bas niveau : CPU, RAM, etc.
2. Haut niveau : Métriques, performance du modèle...



* **Observabilité**

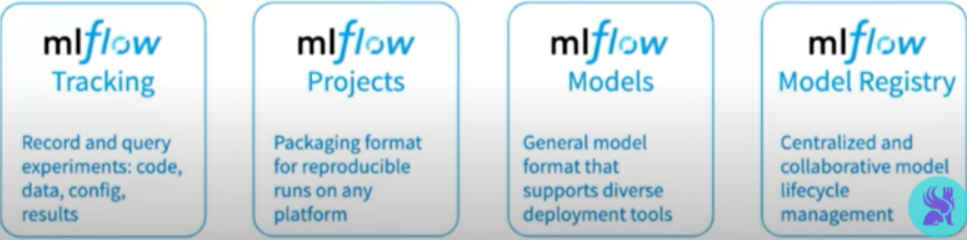
Système observable est un système qui donne des données sur lui-même.

* **MLFLow**

MLFLow est une plateforme open source développée par databricks destinée à la gestion du cycle de vie complet des projets de machine learning. Elle prend en charge la gestion des expériences, le suivi des modèles, l'emballage et le déploiement des modèles.

MLFlow est indépendant des modules installés et il existe des SDK Python, R et Java.

Il existe plusieurs fonctions : projects, tracking et models.



* **Arize**

Plateforme d’observabilité pour le machine learning. Permet de monitorer, évaluer et débuger les applications LLM.

- Gestion des performances de l’apprentissage automatique

- Détection des dérives

- Vérification de la qualité de données

- Validation des modèles.

* **superwise.ai**

C’est est une plateforme d'analyse et de surveillance des modèles d'intelligence artificielle. Elle permet aux entreprises de surveiller en continu leurs modèles AI afin d'identifier les déviations de performance, les biais et les erreurs potentielles. Grâce à des fonctionnalités avancées telles que la détection automatique des anomalies et la génération de rapports détaillés, [superwise.ai](http://superwise.ai/) offre une solution complète pour garantir la fiabilité et la transparence des modèles d'IA.

Elle peut se connecter à n’importe quelle solution de machine learning en production.

* **Collecteurs**

Composants ou outils qui rassemblent des méta-données ou des métriques à partir de systèmes pour le monitoring ou d'autres analyses.

* **Déclencheurs**

Conditions ou événements prédéfinis qui, lorsqu'ils se produisent, déclenchent une certaine action ou alerte dans un système de monitoring.

* **Intégration continue**

Approche CI qui consiste à automatiser le déploiement des changements de code réalisé par plusieurs contributeurs dans un seul projet de développement.

* **Déploiement continue**

Approche CD qui consiste à automatiser le déploiement d'une version d’une application.

* **Livraison/Distribution continue**

Étape au milieu du Pipeline CI/CD

* **Grafana**

Plateforme d'analyse et de monitoring pour visualiser et créer des tableaux de bord dynamiques pour les données de temps réel et les métriques.

* **Métriques système**

Métriques bas niveau : CPU, RAM, etc.

* **Métriques de performance**

Métrique de haut niveau : Métriques du modèle, performance du modèle.

# **Hypothèses :**

* Briand : Il n’y a pas d’ordre de l’exécution du testing et du monitoring.

Vrai. Tester en premier est recommandé.

* Briand II : Effectuer du testing reviens à observer la courbe d’évolution de nos métriques choisis sur différents cas de figure et sur la durée.

Faux, il s’agit de monitoring.

* Adrien : Il faut faire des tests sur toutes les fonctions de notre programme.

Vrai

* Axel : Pytest est plus simple que Unittest pour faire des tests unitaires en python.

Vrai, meilleure syntaxe et plus de fonctionnalités.

* Aude : Les tests unitaires peuvent être intégrés au code grâce à Docstring.

Vrai. Non recommandé.

* Tetyana : Il est obligatoire de faire du testing et du monitoring pour de meilleures performance du modèle.

Vrai

* Jean-Paul : Contrairement au testing, le monitoring se fait en continue.

Vrai

* Seydou : Les dashboards vus précédemment sont les mêmes que ceux du monitoring.

Faux, le monitoring utilise des dashboard mais va plus loin que ça.

* Étienne : On peut passer par plusieurs technologies pour réaliser du testing et du monitoring (réseau, console, fichiers de log).

Faux pour le testing, vrai pour le monitoring

# Plan d'action :

* Explorer les ressources
* Traiter les mots clés
* Répondre aux problématiques
* Réponde aux hypothèses
* Rendre les livrables : RER et workshops
* Comparaison pytest / unittest



* Comparaison MLflow / Arize / superwise.ai

|  |  |  |
| --- | --- | --- |
| **MLflow** | **Arize** | [**superwise.ai**](http://superwise.ai/) |
| Une plateforme open-source pour la gestion du cycle de vie des modèles d'apprentissage automatique | Une plateforme de surveillance des modèles ML pour les performances et la fiabilité | Une plateforme d'analyse et de surveillance des modèles d'intelligence artificielle |
| Fournit des fonctionnalités pour le suivi des expériences, le packaging des modèles et le déploiement en production | Permet de capturer et d'analyser les données de prédiction en temps réel pour détecter les dégradations et les erreurs | Permet de surveiller en continu les modèles d'IA pour identifier les déviations de performance, les biais et les erreurs potentielles |
| Prend en charge plusieurs frameworks d'apprentissage automatique tels que TensorFlow, PyTorch, et scikit-learn | Intégration avec des frameworks populaires tels que TensorFlow, PyTorch et XGBoost | Offre une solution complète pour garantir la fiabilité et la transparence des modèles d'IA |
| Possède une interface utilisateur conviviale pour suivre et gérer les expériences et les modèles | Offre des tableaux de bord et des visualisations pour analyser les performances des modèles | Fournit des fonctionnalités avancées telles que la détection automatique des anomalies et la génération de rapports détaillés |
| Documentation complète et communauté active | Documentation complète et support client disponible | Documentation complète et support client disponible |

