

Rien ne dure éternellement : améliorer les prévisions en automatisant la détection des dérives

Gado, Moubarak, mgado@mathworks.com, MathWorks, Paris (Orateur)

Harouimi, Pierre, pharouim@mathwork.com, MathWorks, Bordeaux

Thématique : Data Ops / ML Ops

Résumé : Une solution de Machine Learning n'est aussi fiable que les données ayant servies à son apprentissage. En conditions réelles, les données peuvent évoluer au fil du temps et différer de l'échantillon initial. Dès lors, ces dérives soulèvent trois problématiques concernant le modèle : l'évaluation de la dégradation de ses performances, sa mise à jour à partir d'informations actualisées, et son déploiement dans un cycle automatisé déjà en production. A partir d'un cas concret de maintenance prédictive, nous présentons une solution qui répond à chacune de ces problématiques.

Mots clés : Data Drift, Machine Learning, AutoML, Interpretable AI, State of Health, Dégradation

1. Introduction

Prenons l'exemple d'une flotte de véhicules électriques utilisés pour la livraison autonome de colis. Leurs batteries se dégradent avec le temps, ce qui augmente le temps de charge et diminue l'autonomie des véhicules. Les batteries sont volumineuses et coûteuses à remplacer, et le fait de s'appuyer sur une estimation statistique de leur durée de vie conduit inévitablement à remplacer certaines batteries trop tôt et d'autres trop tard. Une approche plus rentable consiste à recueillir des données sur l'état et les performances des batteries de chaque véhicule et à utiliser des modèles d'apprentissage automatique pour prédire la durée de vie restante de chaque batterie. Toutefois, les changements dans l'environnement d'exploitation peuvent introduire une dérive dans les données, qui par conséquent peut impacter la performance des batteries. La température externe, par exemple, affecte le taux de charge et de décharge maximal de la batterie, ce qui rend les prédictions du modèle moins précises.

2. Méthodologie

Nous présentons une solution qui répond à chacune des problématiques soulevées par la dérive des données dans un environnement de production :

- Pour détecter la dérive, des algorithmes évaluent périodiquement la variabilité des observations et la précision des prédictions du modèle
- Des modèles à haute-fidélité labelisent les nouvelles données
- L'intégration et le déploiement continus se font grâce à des pipelines de Machine Learning conformes aux standards industriels.

En automatisant à la fois la détection de la dérive et la labélisation des données, nous réduisons le niveau d'expertise requis pour faire fonctionner le système en production.

Dans l'exemple d'une flotte de véhicule électrique, les données de la batterie sont transmises via Kafka aux sous-systèmes de tests (apprentissage) et de production : un modèle MATLAB déployé sur le serveur de production qui prédit la durée de vie restante de chaque batterie et un modèle Simulink physique thermodynamiquement précis de la batterie qui labélise automatiquement les données en vue de leur utilisation pour l'apprentissage de nouveaux modèles d'IA. La labélisation basée sur le modèle physique est préférée dans ce cas, étant beaucoup plus rapide que celle basée sur la simulation, rendant cette dernière moins adaptée à une utilisation en production. Le sous-système de production supervise en continu à la fois le modèle déployé et les données pour détecter les dérives. Si le modèle se dégrade, et que cette dégradation est liée à une dérive des données les plus récentes, le modèle est réentraîné voire recréé. Les modèles nouvellement obtenus sont transférés dans un registre de modèles où le système de production peut les identifier et les intégrer dans le pipeline d'apprentissage automatique déployé.

Tout en tenant compte du cycle de vie DevOps de vos modèles, l'approche proposée permet de développer localement des analyses en temps réel, puis de déployer ces solutions sur le cloud et appliquer des stratégies de Machine Learning automatisé pour maintenir les modèles à jour et leurs prédictions aussi précises que possible.

3. Originalité / perspective

La classification automatique des données et la détection fiable des dérives créent ainsi une architecture adaptable qui réduit la complexité du développement et du déploiement de solutions de Machine Learning pour les problèmes de maintenance prédictive. Le choix d'une suite d'outils prenant en charge le cycle de vie DevOps du modèle facilite la collaboration des équipes et garantit que la solution s'adapte au problème.

Références

<https://www.mathworks.com/solutions/artificial-intelligence.html>

<https://www.sia.fr/evenements/274-ia-pour-nouvelles-mobilites-2022>