**Détails de l’article**

* **Titre** : Anomaly Detection in Microservice-Based Systems
* **Auteurs** : João Nobre, E.J. Solteiro Pires, Arsénio Reis
* **Année de publication** : 2023
* **Journal** : Applied Sciences

**Résumé**

Cet article étudie l'utilisation d'un modèle de machine learning supervisé, le **Perceptron multicouche** (MLP), pour détecter les anomalies dans les systèmes basés sur des microservices. En créant une infrastructure de microservices, un module d'injection de fautes pour simuler des anomalies, et un dataset de surveillance, l'étude montre l'efficacité du MLP dans la détection des anomalies avec une précision élevée, particulièrement au niveau du service. Ces résultats suggèrent que l'automatisation de la surveillance pourrait améliorer la gestion des systèmes distribués.

**1. Domaine traité**

Le domaine principal abordé est la **surveillance et détection d’anomalies** dans les systèmes logiciels basés sur des microservices.

**2. Problème résolu dans cet article**

* **Problématique scientifique** : Les systèmes de microservices deviennent complexes et difficiles à surveiller manuellement pour identifier et corriger les anomalies.
* **Problème de société** : Les interruptions de service ou les anomalies affectant la qualité des services ont un impact direct sur la satisfaction des utilisateurs finaux.
* **Question de recherche** : Comment automatiser efficacement la détection d'anomalies dans des environnements de microservices pour assurer leur disponibilité et leur résilience ?

**3. Intérêt d’avoir une solution pour le domaine d’étude**

Dans le cadre de la **surveillance distribuée des applications basées sur des microservices** en environnement DevOps, résoudre le problème de détection d’anomalies permettrait de renforcer la réactivité et la résilience des systèmes. En automatisant la détection des anomalies, il devient possible d’anticiper et de traiter rapidement les défaillances, ce qui est crucial pour garantir une disponibilité continue des services. Cela contribue également à améliorer la gestion des performances et la qualité des services rendus aux utilisateurs, des priorités fondamentales en DevOps pour des architectures microservices robustes.

Cette solution va également permettre de libérer les équipes d'opérations pour qu'elles puissent se concentrer sur l'amélioration continue, tout en ayant une vue globale de la santé du système, ce qui est essentiel pour une surveillance en temps réel dans un cadre DevOps.

**4. Approche adoptée**

L’approche consiste à utiliser un modèle de machine learning **supervisé** (MLP) entraîné avec des données de performance recueillies d’un système de microservices open-source (Sock-shop). L’algorithme est testé pour sa capacité à identifier des anomalies à l’échelle du service et de l’application.

**5. Solution proposée**

La solution détaillée implique la création de datasets spécifiques pour les anomalies d’application et de service, et le prétraitement de données pour le MLP. La performance du modèle est optimisée avec GridSearchCV (**Technique d'optimisation des hyperparamètres utilisée en machine learning pour trouver la meilleure combinaison de paramètres d'un modèle, maximisant ainsi sa performance**), en utilisant les métriques de précision, rappel, F1-score, et taux de faux positifs.

**6. Discussion**

* **Points positifs** :
  + Capacité élevée de détection d’anomalies au niveau des microservices.
  + Utilisation de métriques de surveillance en temps réel (Prometheus, Grafana).
  + Précision et rappel élevés pour les anomalies au niveau du service.
  + Faible taux de faux positifs dans la détection d’anomalies.
* **Manquements relevés** :
  + Dépendance aux données synthétiques pour l'entraînement, limitant la généralisation.
  + La détection des anomalies au niveau de l’application est moins performante.
  + Nécessité d’étendre le modèle pour gérer des benchmarks plus diversifiés.
  + Explication et interprétation des résultats du MLP pour l’utilisateur sont complexes.

**7. Intérêt pour votre problème**

L’approche de détection d’anomalies présentée dans cet article est directement applicable au thème de la **surveillance distribuée** dans les architectures de microservices, essentiel en DevOps. L’utilisation de métriques de performance pour automatiser la détection d’anomalies, couplée à des outils comme **Prometheus** et **Grafana**, offre des perspectives intéressantes pour améliorer la gestion proactive des incidents et optimiser la résilience des applications en microservices, conformément aux pratiques DevOps. Cette méthode pourrait ainsi faciliter la mise en œuvre de systèmes de surveillance automatisés capables de réagir en temps réel, un atout pour des architectures microservices performantes.

**8. Travaux à regarder**

Les travaux référencés dans la bibliographie de l'article qui peuvent être utiles incluent les recherches sur la détection d’anomalies dans les systèmes distribués (ex. Du et al. sur les SVM et k-NN pour la détection d’anomalies, Mariani et al. Utilisant des séries temporelles pour la détection d'anomalies).

Les références de l'article qui pourraient être utiles sont :

* **Du et al.** pour leur utilisation de SVM, k-NN, Naive Bayes, et Random Forest en détection d'anomalies via des métriques de surveillance dans les microservices, ce qui s'aligne avec la surveillance distribuée.
* **Zhou et al.** qui appliquent des techniques de ML pour classer et localiser les anomalies dans les microservices.
* **Mariani et al.** sur l'utilisation de séries temporelles pour détecter des anomalies dans des applications multi-services, une approche précieuse pour l’analyse de la performance continue dans DevOps.
* Les références sur l’utilisation de **Prometheus et Grafana** pour la collecte et la visualisation de métriques, essentielles pour un pipeline de surveillance dans un environnement DevOps.