

## ABSTRACT

---

Machine Learning Operations (MLOps), the task of coordinating machine learning projects with multiple models and team members, is growing in importance and interest. Cloud computing resources are a popular option in this scenario due to many reasons like easily accessible computing resources, billing by usage time and further available services like a fully managed environment. Two approaches to monitor models in an MLOps environment are compared by using two popular statistical time series forecasting models and two datasets from a widely known forecasting competition. One approach is the default Amazon Web Services (AWS) model drift monitoring and the other is a tracking signal monitoring. The goal is to reduce economical and ecological costs generated by retraining deployed models with more recent data in a cloud environment. The tracking signal monitoring is shown to serve as a more generic approach which can reduce costs when a decreased model performance is accepted for lower training costs. The AWS monitoring with in-sample error metrics as monitoring threshold used as retraining trigger shows a better performance at a comparable level of retraining counts.

## ZUSAMMENFASSUNG

---

Projekte des maschinellen Lernens mit mehreren Modellen und Teammitgliedern zu koordinieren (MLOps) ist eine Aufgabe wachsender Relevanz. Aus verschiedenen Gründen wie dem leichten Zugang zu Rechenressourcen und weiteren angebotenen Leistungen wie einer vollständig verwalteten Rechenumgebung sind Cloud Computing Ressourcen eine beliebte Option in diesem Szenario. Die Abrechnung ist dabei meist an die Nutzungszeit gekoppelt. Um zu evaluieren, wie sich die Kosten in einer Cloud Computing MLOps Umgebung in Verbindung mit Modellen aus dem Bereich der Zeitreihenvorhersage abhängig von der Überwachungsvariante verhalten, werden verschiedene Herangehensweisen verglichen. Dabei sollen ökonomische und ökologische Kosten, die durch das Nachtrainieren von Modellen mit neueren Daten entstehen, reduziert werden. Dazu werden zwei beliebte statistische Modelle und zwei Datensätze aus einem bekannten zeitreihenprognose Wettbewerb evaluiert. Als Überwachungsmethode wird die standardmäßige Überwachungsvariante für MLOps Umgebungen von Amazon Web Services (AWS) mit einer Tracking Signal gestützten Überwachung verglichen. Die Tracking Signal Überwachung bietet einen generischeren Ansatz, welcher Kosten reduzieren kann, wenn eine reduzierte Vorhersagekraft von Modellen für verringerte Kosten durch das Nachtrainieren akzeptiert werden kann. Die AWS Überwachung, welche die Fehlermetriken aus den Trainingsdaten als Grenzwert nutzt um ein Nachtrainieren auszulösen, zeigt eine bessere Leistung bei einem vergleichbaren Level von Nachtrainierungsiterationen.