Analyse

Die Erkennung von Anomalien erfolgt dabei durch Erlernen des normalen Verhaltens durch maschinelles Lernen. In den letzten Jahren hat maschinelles Lernen gezeigt, dass es möglich ist, Muster in Bildern zu erkennen, beispielsweise Objekte oder Gesichter zu erkennen oder Anomalien in elektrischen Signalen mithilfe von Erfahrung aus maschinellem Lernen zu erkennen.

Ich habe mehrere Methoden ausgewählt, die auf unbeaufsichtigtem Lernen basieren, da es den Vorteil hat, dass es einfacher zu implementieren ist als überwachtes Lernen. Methoden, die unbeaufsichtigtes Lernen verwenden, haben auch gezeigt, dass sie flexibel lernen können und daher für jede Art von Prozessangriff verwendet werden können.

Diese Methoden lassen sich in zwei Arten der Anomalieerkennung einteilen: Erkennung und Vorhersage basierend auf Zeitreihen. In diesem Bericht werde ich die Erkennungs- und Vorhersagedefinitionen verwenden, die auf der Taxonomie von [3] basieren und die Erkennung von Anomalien als Erkennung basierend auf den uns vorliegenden Daten definieren und vergleichen, um einen Fehler zu erkennen. Prognosen basierend auf Zeitreihen werden als eine Methode definiert, die die Werte der nächsten Zeitreihen vorhersagt und mit den beobachteten Zeitreihen vergleicht.

1. CNN

Das Convolutional Neural Network (CNN) ist eine Methode, die auf neuronalen Netzen und Faltung basiert, um Muster und Muster in Signalen oder Bildern zu lernen. Während eines CNN wird der Wert jedes Punktes der nächsten Schicht aus den benachbarten Werten dieses Punktes der vorherigen Schicht berechnet. Diese Operation wird durch die Zahl "" dargestellt.

Figure

Wenn wir zum Beispiel für den Kernel [-1,1, -1] haben, ist der Wert des i-ten Punktes (i-1) \* (- 1) + i + (i + 1) \* (- 1) . Die Werte des Kernels ändern sich natürlich während des Trainings, um die Kostenfunktion zu reduzieren.

CNN kann sowohl zur Erkennung als auch zur Vorhersage verwendet werden. In meinem Fall werde ich dasselbe neuronale Netzwerk wie in [4] verwenden, das ihre Methode zur Vorhersage verwendet. In [4] verwenden sie ein CNN, das aus 5 Faltungsschichten mit einem Kern der Größe 3 und einem Kern der Größe 1 besteht. Das neuronale Netzwerk ist in Abbildung „“ dargestellt.

Figure

Die Schichten 3, 4 und 5 sind verkettet, um eine Überanpassung des neuronalen Netzwerks zu verhindern.

Bei der Eingabe des neuronalen Netzwerks werden die Zeitreihen vom Zeitpunkt t bis t + T-1 jedes Sensors und Aktors injiziert, und bei der Ausgabe haben wir die Vorhersage von Zeitreihen vom Zeitpunkt t + T bis t + 2-1 für jeden Sensoren und Aktoren. Die Ausgabe wird mit den wahren Werten der Zeitreihen verglichen. Die Kostenfunktion wird anhand des mittleren quadratischen Fehlers (MSE) berechnet.

1. Autoencoder

Der Autoencoder ist ein neuronales Netzwerk, das die Dimensionsreduktion verwendet, um die Eigenschaften seiner Eingabe zu kennen. Der Autoencoder ist ein neuronales Netzwerk, das zur Erkennung von Anomalien verwendet wird. Jede Schicht wird mit der Funktion "Vollständig verbundene Schicht" mit der nächsten verbunden, dh jedes Neuron einer Schicht wird an alle anderen Neuronen der nächsten Schicht weitergeleitet, wie in Abbildung "" gezeigt.

Figure

Die verwendete Autoencoder-Struktur ist dieselbe wie in [5] mit 9 vollständig verbundenen Schichten mit den Abmessungen 256, 196, 136, 76 und 14. Das Funktionsprinzip ist in Abbildung "" dargestellt.

Figure

Als Eingabe für den Autoencoder senden wir für jeden Sensor ein Fenster mit 40 Werten. Der Autoencoder versucht, diese Fenster wiederherzustellen und dabei die Überanpassung zu vermeiden, da sonst der Autoencoder der Gerätefunktion entspricht. Die Kostenfunktion wird anhand des mittleren quadratischen Fehlers berechnet.

1. LSTM
2. SVM

SVM kategorisiert die Signale in zwei Klassen. Eine Klasse besteht aus allen richtigen Werten der Signale und die andere besteht aus aller falschen Werter etwa den Anomalien.

Für das SVM habe ich den Parameter "nu" geändert, der eine Obergrenze für den Anteil der Trainingsfehler und eine Untergrenze für den Anteil der Unterstützungsvektoren darstellt. Durch Verringern dieses Parameters erhalten wir bessere Ergebnisse für die Erkennung vom Normalverhalten, aber schlechte Ergebnisse für die Erkennung von Anomalien. Durch Erhöhen dieses Parameters können wir bessere skalierungsabhängige Ergebnisse erzielen.

1. LOF

Für LOF muss mit der Anzahl der Nachbarn experimentiert werden, um die Dichteabweichung einer Probe zu berechnen. Durch Ändern dieses Parameters werden je nach Scaler unterschiedliche Verhaltensweisen beobachtet, was bei einigen Scalern zu guten Ergebnissen bei einem niedrigen Wert der Anzahl der Nachbarn führen kann, während bei anderen Scaler eine große Anzahl von Nachbarn erforderlich ist, um eine korrekte Anomalieerkennung zu erzielen.

1. DBSCAN

Für DBSCAN experimentierte ich mit dem Parameter, der den maximalen Abstand zwischen zwei Samples darstellen, damit sie in demselben Sample berücksichtigt werden, und mit dem Parameter, welches die maximale Anzahl von Samples in einer Nachbarschaft darstellt, damit dies als Kernpunkt betrachtet wird. Diese beiden Parameter wurden so gewählt, dass DBSCAN ein einzelnes Cluster erkennt, wobei der eine das normale Verhalten des Prozesses neu gruppiert und alle Ausreißer die Punkte sind, die eine Anomalie erkennen.

[3] Grohmann, J.; Herbst, N.; Chalbani, A.; Arian, Y.; Peretz, N.; Kounev, S. A Taxonomy of Techniques for SLO Failure Prediction in Software Systems. Computers **2020**, 9, 10. https://doi.org/10.3390/computers9010010