

Generation of sequential data using machine learning models

In Zusammenarbeit mit der European Spallation Source (ESS) entwickelt die Zürcher Hochschule für Angewandte Wissenschaft (ZHAW) ein System für den Schutz des sich im Bau befindlichen Protonenbeschleunigers. Kleinste Abweichungen eines solchen hochenergetischen Strahls können zu grossem Schaden an der Anlage führen. Um solche Abweichungen schnellstmöglich zu detektieren, um darauf zu reagieren, werden verschiedene Sensoren und Algorithmen eingesetzt, welche in Zukunft durch Methoden aus dem Bereich des maschinellen Lernens unterstützt werden sollen. Solche Methoden benötigen in der Regel grosse Datenmengen, um Eigenschaften aus den Daten zu extrahieren. Die Datengewinnung stellt sich in der Entwicklungsphase eines Systems als schwierig dar, weshalb auch hier auf datengenerierende Modelle aus dem Bereich des maschinellen Lernens gesetzt werden könnte. In den letzten paar Jahren wurden neue datengenerierende Modelle vorgeschlagen, um Zeitreihendaten zu erzeugen.

Im Rahmen dieser Bachelorarbeit wurden geeignete Modelle, die mit Zeitreihendaten umgehen können, umgesetzt und für verschiedene Datensätze evaluiert. Zu den untersuchten Modellen gehören der Variationale Autoencoder (VAE), das Generative Adversarial Networks (GAN) Framework sowie das TimeGAN Framework, welches speziell für sequenzielle Daten entwickelt wurde. Ausserdem wurde ein eigener Ansatz verfolgt, welcher das GAN Framework mit einem Vorhersagemodell verknüpft. Alle Modelle weisen im Allgemeinen ein grosses Potenzial auf. Der VAE zeigt mithilfe von dimensionsreduzierenden Visualisierungen eine besonders gute Abdeckung der Originaldaten mit den synthetischen Daten. Modelle mit der Fähigkeit Zeitabhängigkeiten zu lernen, weisen eine bessere Klassifizierungs- und Vorhersagewertung auf. Zu diesen Modellen gehört das TimeGAN Framework und unser eigenes Konzept, welches wir PredictiveGAN nennen. Das TimeGAN Framework erweist sich als das rechenintensivste Modell. Unser PredictiveGAN erreichte neben dem VAE auch eine sehr gute visuelle Abdeckung der Originaldaten.

Schlüsselwörter: Deep Learning, Generative Modelle, Zeitseriedaten

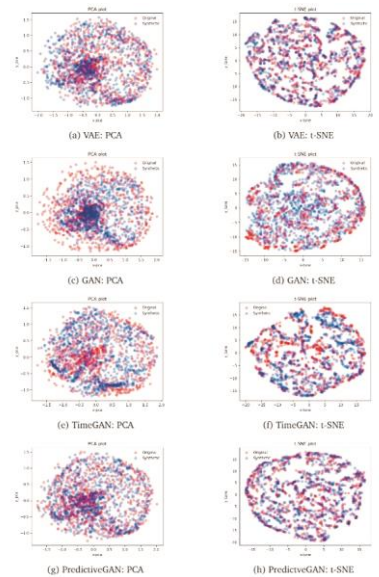


Diplomierende

Jenestin Anton Anthonipillai
Daniel Bauer

Dozentin

Monika Ulrike Reif



Verschiedene generative Modelle für Zeitreihendaten wurden untersucht. Mithilfe von Dimensionsreduktionstechniken (links PCA, rechts t-SNE) wurden die Originaldaten (rot) mit den synthetischen Daten (blau) verglichen.