Anomaly Detection Using LSTM Autoencoder Network In Ensuring Reliable Operation At Offshore Utilities System

Authors: Afidalina Tumian, M Hafrizal Azmi and M Zakwan M Sahak

Abstract

Continuous monitoring of offshore utilities system performance is essential in ensuring smooth operation and avoiding unplanned deferments (UPD). A proper prediction tool to identify anomalous patterns could help flag potential problems that may lead to a UPD event. With very few recorded and validated UPD events, formulating a suitable methodology that gives high accuracy score and low false positive signals is a challenge.

This paper presents our unsupervised machine learning approach using long short-term memory (LSTM) autoencoder network, in identifying anomalous pattern in offshore utilities system. This network learns the normal behavior of multiple sensors of interest. It first performs a compression process to represent data in lower dimension and then decompress them back to its original dimension.

This compression-decompression method produces a distribution of expected reconstruction error. When the trained model is presented with data exceeding its normal behavior, anomaly is identified when the residual between actual and reconstructed data is beyond the maximum value of reconstruction error; defined as T_{hard}. T_{soft} is an additional threshold that we provide for the system to give earlier warning before it reaches T_{hard}.

This network has been trained on two years of a fuel gas system data with frequency of two minutes. When tested with more than two years of data, our results have shown that this model successfully captures all five UPD events during the tested period, with some showing much earlier deviation pattern than the real UPD event.

This robust method has shown its potential in identifying anomalous patterns across multiple sensors, thus able to alert operators timely for further action. This could prevent events that can cause disruption at our facilities leading to unwanted production loss.

Keywords: anomaly, unsupervised machine learning, long-short term memory network

Introduction

Anomalies are data points that exhibit significantly different behavior compared to the rest. In brief, they can be categorized into the following categories: 1) point anomalies: anomalous individual instances compared to most other instances, 2) conditional/contextual anomalies: individual instances which are only anomalous at certain context, otherwise acknowledged as normal; and 3) group anomalies: consisting of more than one point of anomalous instance (Chandola et al., 2009).

Due to their rarity, heterogeneity and complexity in nature, anomaly detection has becoming active research widely applied across industries, owing to the availability of massive data and the advances of computing power which have eased the method training, testing and deployment. Several review papers have provided a comprehensive overview on the landscape of anomaly detection methodologies (Chandola et al., 2009; Feroze et al., 2021; Pang et al., 2021).

Pang et al., (2021) has highlighted six challenges that anomaly detection methodology needs to address: 1) low anomaly detection recall rate, 2) anomaly detection in high-dimensional and/or not-independent data, 3) data-efficient learning of normality/abnormality, 4) noise-resilient anomaly detection, 5) detection of complex anomalies, and 6) anomaly explanation (see the paper for further explanation). They then observed that there has been an increasing trend in experimenting deep learning methods to overcome some of these challenges. The ability of deep learning methods to understand complex patterns in highly dimensional data giving them the advantage in terms of accuracy and performance.

This paper focuses on exploring the application of one of the deep learning techniques in identifying anomalous patterns exhibited in industrial sensor data. If these events are captured much earlier, it will help flag issues and problems well ahead, allowing prompt measures to be taken prior to a more catastrophic situation, such as unplanned deferments (UPD).

The challenge faced in the chosen use case, which has been based on operation data from fuel gas system is due to low number of recorded UPD events. Having information on these events is essential during model validation phase, in ensuring that our methodology is providing high accuracy score and low false positives signals (as highlighted in Pang et al. (2021)'s paper). Thus, the approach we adopt here is by applying an unsupervised machine learning approach (or unsupervised learning, in short), using long-short term memory (LSTM) autoencoder network, on multivariate data. Potentials of LSTM autoencoder network in anomaly detection have been shown in some other similar applications (Devshali et al., 2022; Patel, 2019).

Methodology

Unlike supervised machine learning, unsupervised learning (UL) approach can discover the hidden patterns or structures in data without requiring data labels. This is achieved by performing representation learning, i.e., attempting to present the trained data with a set of parameters smaller than the number of trained datasets (Bengio et al., 2013). In brief, let dataset $\mathrm{X} = \{x_1, x_2, \dots, x_n\}$ with $x_i \in \mathbb{R}^D$, and $\mathrm{Z} \in \mathbb{R}^K (K \ll N)$. A chosen UL approach is trained to learn representation learning mapping function $f(\cdot): \mathrm{X} \mapsto \mathrm{Z}$.

Autoencoder is a type of unsupervised deep learning that is able to learn the behaviour of normal instances, and thus able to signal potential anomalies when they occur (Hinton & Salakhutdinov, 2006; Patel, 2019). This is achieved by first performing a compression process to represent data in lower dimension and then decompress them back to its original dimension (Kramer, 1991) (see Figure 1). The earlier representation learning mapping function that we saw is denoted as "encoder" network and the similar "decoder" network can be constructed to recover the data from the latent representation, Z. Amongst the common application for autoencoder is for noise removal in images (Patel, 2019).

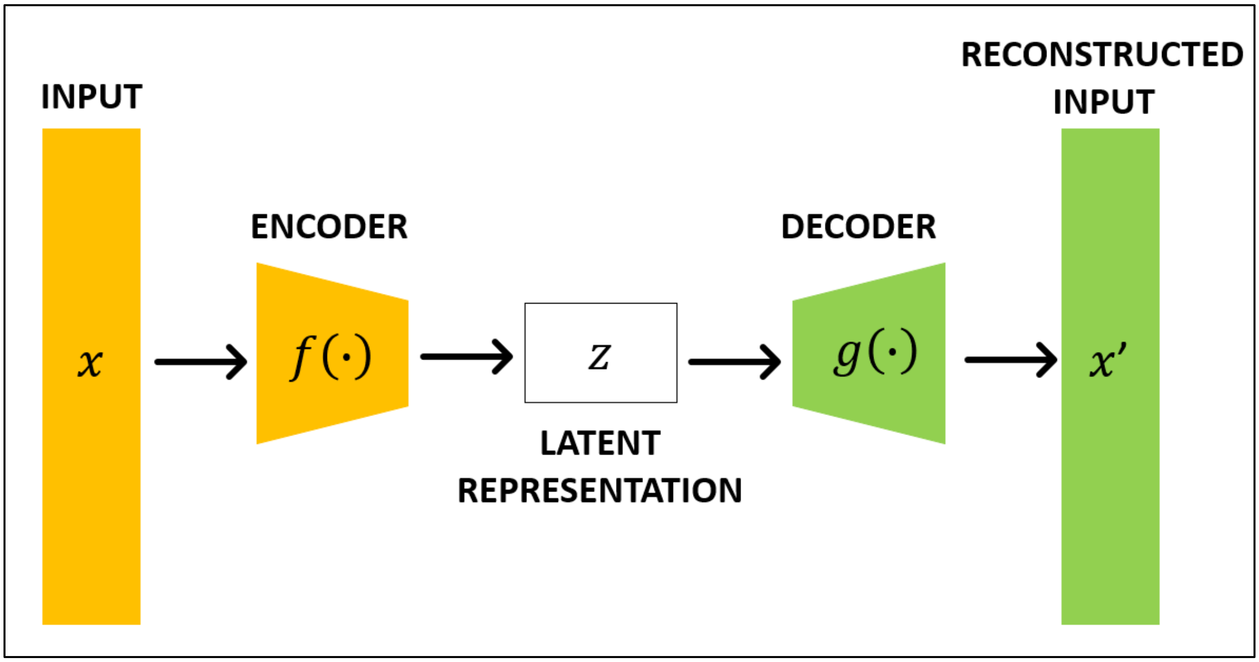

Figure 1: Diagram illustrating a typical autoencoder structure, comprising of encoder and decoder.

Once the network has been trained, residual between actual and reconstructed data, $R_{expected}$ is calculated. Anomaly is then identified when residual between actual and reconstructed data, is beyond the maximum of $R_{expected}$ (defined as T_{hard} – see Figure 2). We also incorporated LSTM layers in encoder and decoder networks since LSTM is able to capture the patterns of sequence data including time series sequence data and retain information better than recurrent neural network (Hochreiter & Schmidhuber, 1997).

Results

As explained in the earlier section, the key idea to detect anomaly in this work is to train our deep network for normality feature learning. We have designed our anomaly detection method, which is based LSTM autoencoder network, and tested on data obtained from an offshore utilities system (Sahak et al., 2022). Prior to model training, normal behaviour of the data used in our use case was defined together with subject matter experts. This task requires additional steps in performing data treatment.

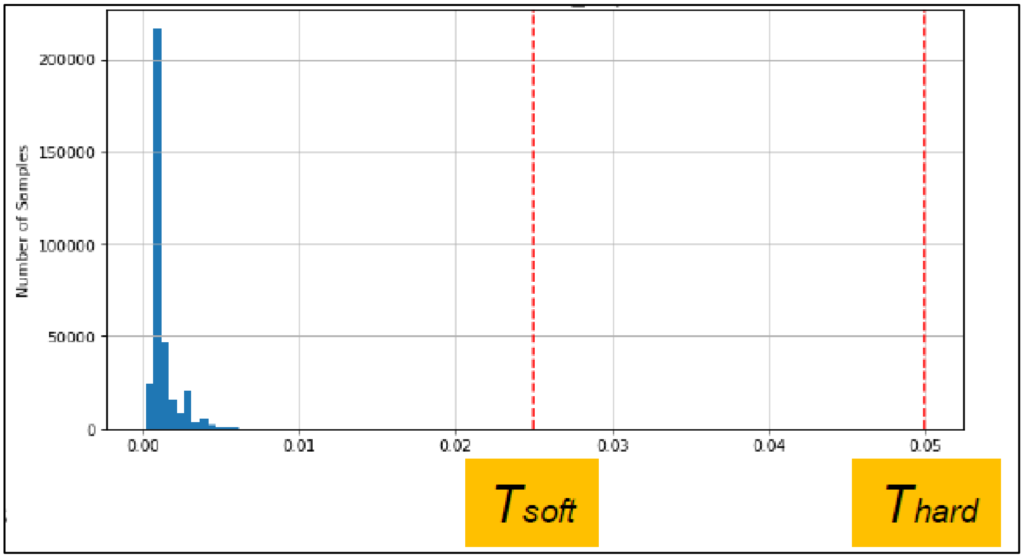

Figure 2: This figure shows the distribution of expected reconstruction error, R-expected. When the trained model is presented with data exceeding its normal behavior, anomaly is identified when the residual between actual and reconstructed data is beyond the maximum value of R(expected); defined as T_{hard}. T_{soft} is an additional threshold that we provide for the system to give earlier warning before it reaches T_{hard}. T_{soft} is an additional threshold that we provide for the system to give earlier warning before it reaches T-hard.

Distribution of $R_{expected}$ when trained with the use case data is shown in Figure 2. When presented with rare events, the trained LSTM autoencoder network is able to capture more than 90% of recorded UPD events from our use case. In fact, in some cases, the method is able to capture before some events took place. One example when validated with a recorded UPD event is shown as in Figure 3. The first instance when the reconstruction error (shown in brown line) crosses T_{hard} happens more than three hours earlier than the recorded UPD event (shown in fuchsia vertical line). The brown line continues to stay between T_{soft} and T_{hard} until slightly before the UPD event where it continues to escalate even further.

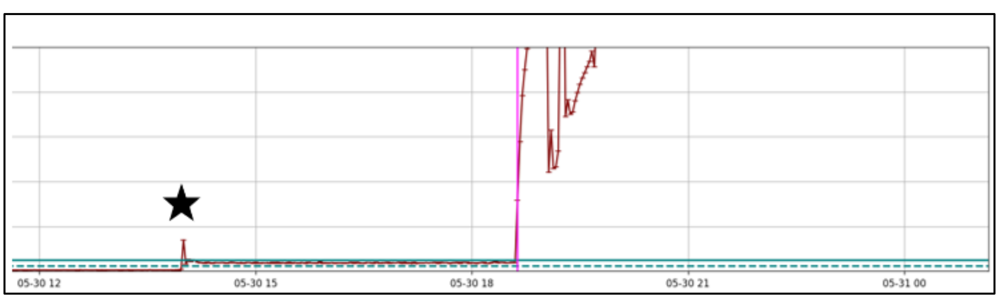

Figure 3: Brown line shows the patterns of reconstructed error plotted over time (y-axis not shown). Dotted and solid horizontal lines in green shows T_{soft} and T_{hard} respectively. Solid vertical line in fuchsia shows the time when a real UPD event took place. The star icon shows when the reconstruction residual surpasses T_{hard} prior to the UPD event.

Discussion and future work

We have shown the potential of LSTM autoencoder network in detecting anomalies by providing earlier warnings on potential events. From operational point of view, this could result in cost saving and opportunity for process improvement and optimization. Future work may involve further experimentation on similar use cases and comparison with the more recent network architecture such as transformer network.

Reference

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828. https://doi.org/10.1109/TPAMI.2013.50

Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, *41*(3), 1–58. https://doi.org/10.1145/1541880.1541882

Devshali, S., Tripathi, S. K., Dodda, D., Kumar, M., Uniyal, R., Yadav, M., & Malhotra, S. (2022). Predicting ESP failures Using Artificial Intelligence for Improved Production Performance in One of the Offshore Fields in India. *Society of Petroleum Engineers - ADIPEC 2022*. https://doi.org/10.2118/211031-MS

Feroze, A., Daud, A., Amjad, T., & Hayat, M. K. (2021). Group Anomaly Detection: Past Notions, Present Insights, and Future Prospects. *SN Computer Science*, *2*(3), 219. https://doi.org/10.1007/s42979-021-00603-x

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the Dimensionality of Data with Neural Networks. *Science*, *313*(5786), 504–507. https://doi.org/10.1126/science.1127647

Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. https://doi.org/10.1162/NECO.1997.9.8.1735

Kramer, M. A. (1991). Nonlinear principal component analysis using autoassociative neural networks. *AIChE Journal*, *37*(2), 233–243. https://doi.org/10.1002/AIC.690370209

Pang, G., Shen, C., Cao, L., & Hengel, A. Van Den. (2021). Deep Learning for Anomaly Detection. *ACM Computing Surveys (CSUR)*, *54*(2). https://doi.org/10.1145/3439950

Patel, A. A. (2019). *Hands-On Unsupervised Learning Using Python*. O'Reilly Media, Inc.

Sahak, M. Z. M., Thant, M. M. M., Tumian, A., Harun, Z., Castillano, E., Chee, S. B., & Tan, K. (2022). A Holistic Approach to the Development of Digital Twin Advisor for Upstream Production Process. *Society of Petroleum Engineers - ADIPEC 2022*. https://doi.org/10.2118/210979-MS