Automatic Anomaly Prediction

Koçfinans

Sabancı Üniversitesi FACULTY OF ENGINEERING AND NATURAL SCIENCES

Student(s)

Ali Erinç AYAZ
Göktuğ ÇALIŞKAN
Alpaslan HAMZAOĞLU

Faculty Member(s)

Anıl KOYUNCU
Cemal YILMAZ

Company Advisor(s)

Alp TUNÇAY
Oktay GÜNGÖR

ABSTRACT

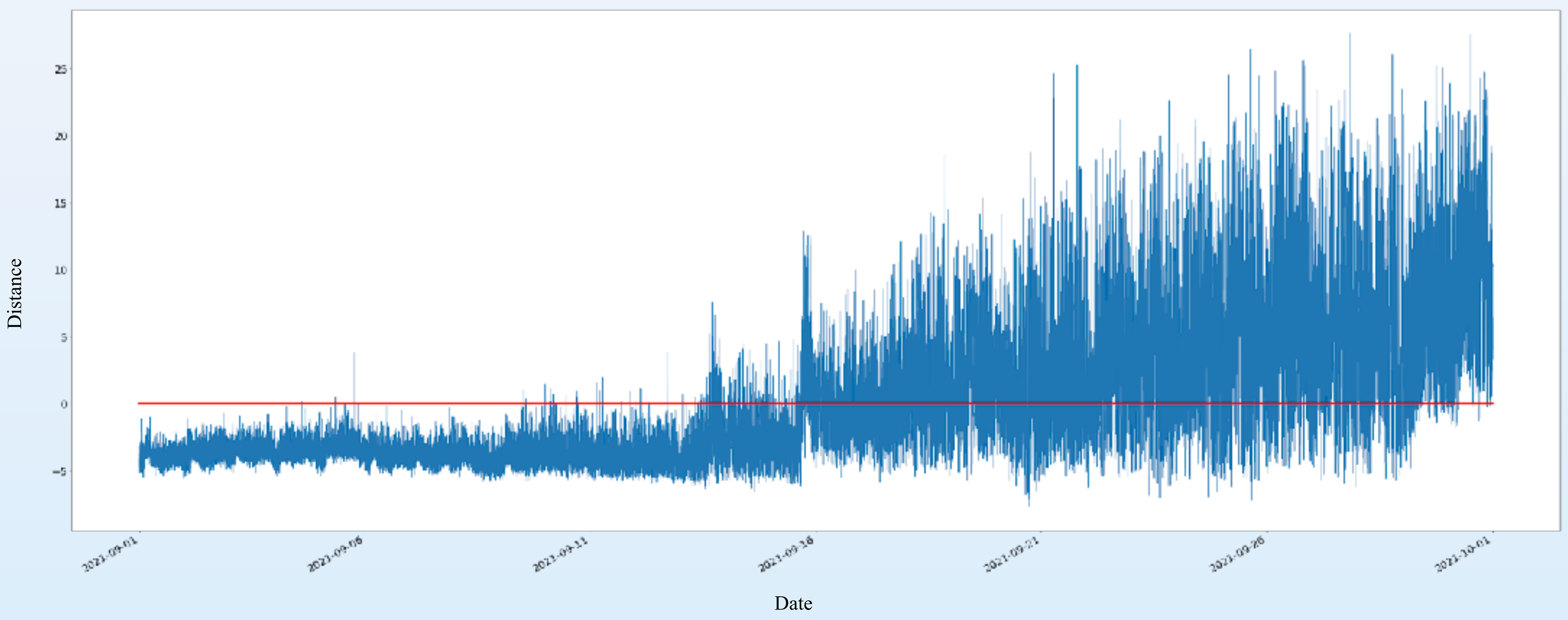

Figure 1. Notes: The figure above shows the decision function graph of the September data. The figure also shows the anomalous data which is determined by the threshold shown as a red straight line on the graph.

Organizations and companies that collect data originating from sales, transactions, client/server communications, IoT nodes, devices, engines, or any other data generator/exchanging need to analyze this data to reveal insights about the running activity.

Koç Finansman also uses a microservices for running their financial services. As most applications do, sometimes these containers can run into failures due to anomalies occurring at hardware or software layers. The main objective of this project is to detect and predict the anomalies that occur in the microservices of Koç Finansman. Koç Finansman provided their telemetry data from their microservices, so that the model to be trained could use real-world data. The collected telemetry data involves various attributes of the microservices, such as memory, disc, and CPU usage. As the dataset is a time series data set, LSTM Autoencoder (Figure 2) was the main choice to be used in this project due to the temporal and inter-dependency among the data. Note that LSTM Autoencoders are specialized for taking the history into account when making predictions.

In this work, by using an unsupervised learning technique, we train and test an LSTM Autoencoder model to find the anomalies by treating the telemetry data as time series data. The results of our empirical studies demonstrated that our trained LSTM Autoencoder succeeded in identifying the anomalous behavior of the data with decent accuracies.

OBJECTIVES

- Develop an approach for detecting the anomalies on a multivariate times series to predict the future manifestation of microservice failures.
- Evaluate the proposed approach by using real data collected in the field.

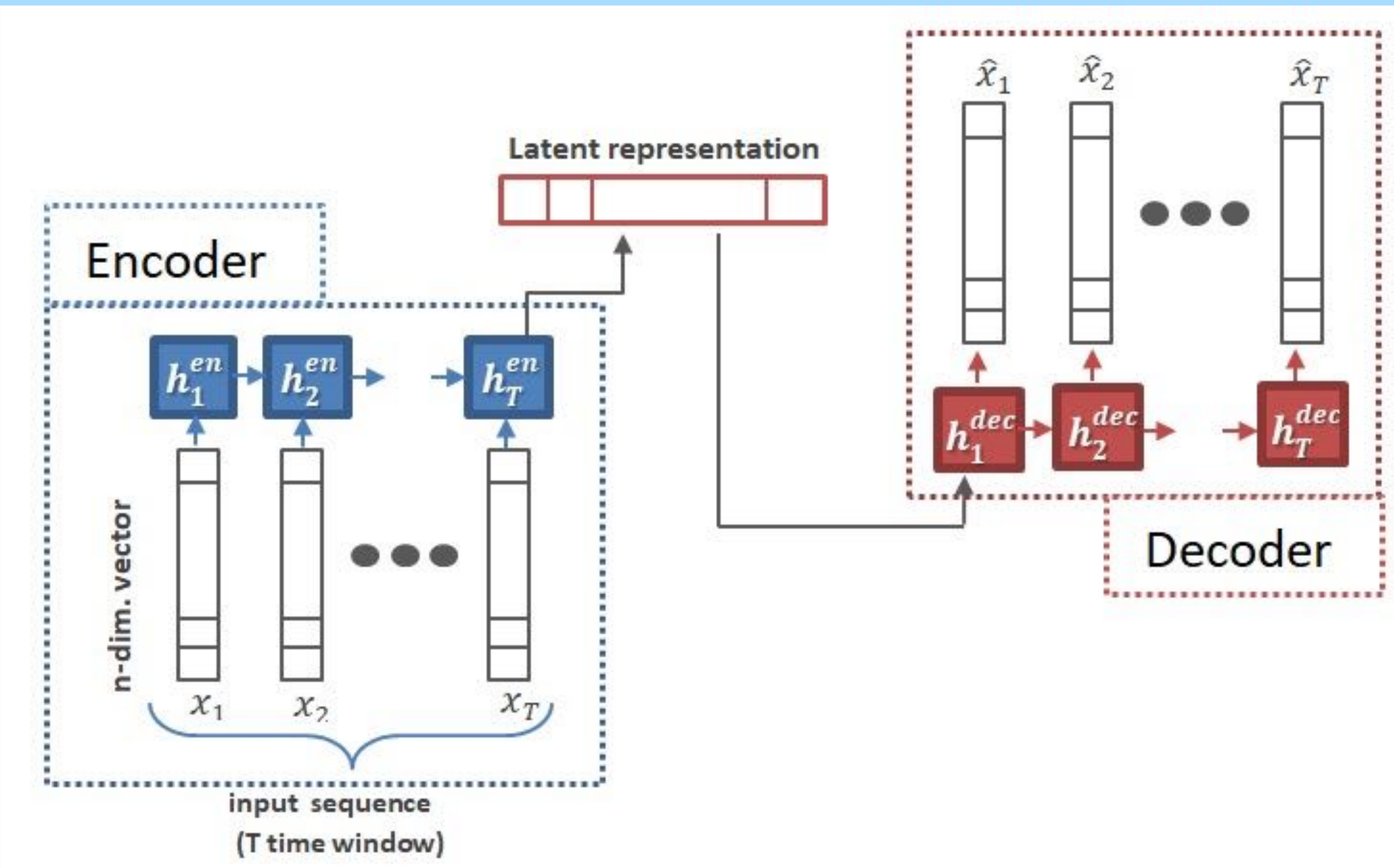

Figure 2. Notes: Process of LSTM Autoencoder. Autoencoder consists of Encoder and Decoder. From Ghrib et al., 2020.

PROJECT DETAILS

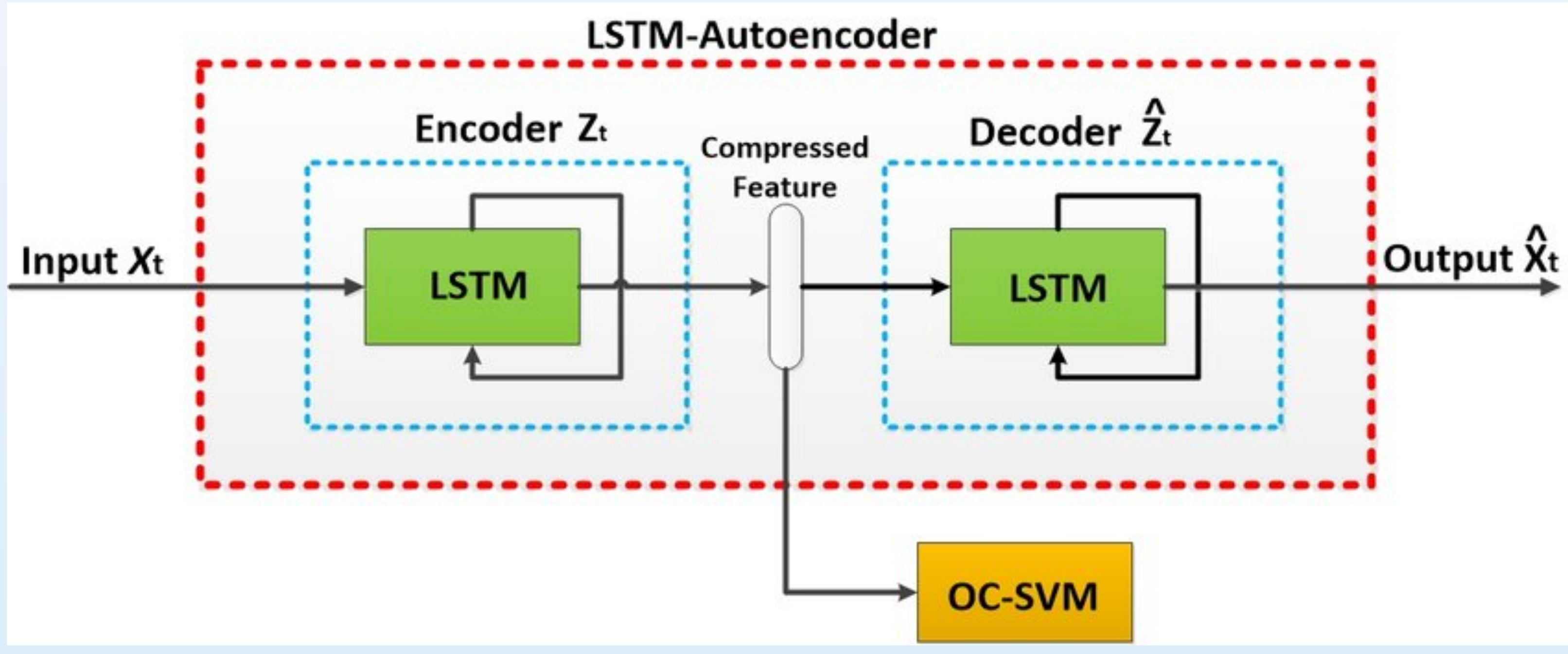

Figure 3. Notes: Latent representation obtained from LSTM Encoder is fed into One-Class SVM for anomaly detection. From Said Elsayed et al., 2020.

Ghrib et al. (2020) indicate a hybrid approach for anomaly detection, in which a combination of LSTM Autoencoder and SVM Classifier is used. In this work, we inspired from the aforementioned approach. In particular, we train an LSTM Autoencoder, the output of which is then fed to a One-Class SVM for anomaly detection (Figure 3). SVM Models create a plane and separate the data points into two classes differentiated by a maximum margin hyperplane with the help of a decision function. In this work, we have empirically evaluated a number of decision functions. We have then carried out a series of experiments to evaluate the proposed model by using the telemetry data collected in the field. And, we discussed the predications made by the model with the DevOps teams at Koç Finansman. In one case, the model detected an anomaly, which is then conformed by the DevOps teams.

CONCLUSIONS

- Our preliminary results are promising
- One potential avenue for future research is to develop more sophisticated models with optimized hyperparameter settings and deploy the system in the field, such that further and more realistic evaluations of the proposed approach can be performed.

REFERENCES

- Ghrib, Z., Jaziri, R., & Romdhane, R. (2020). "Hybrid approach for anomaly detection in time series data". 2020 International Joint Conference on Neural Networks (IJCNN). https://doi.org/10.1109/ijcnn48605.2020.9207013
- Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011. Retrieved January 8, 2022, from https://scikit-learn.org/stable/modules/svm.html
- Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011. Retrieved January 8, 2022, from https://scikit-learn.org/stable/auto_examples/svm/plot_oneclass.html
- Said Elsayed, M., Le-Khac, N.-A., Dev, S., & Jurcut, A. D. (2020). Network Anomaly Detection Using LSTM based Autoencoder. Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks. https://doi.org/10.1145/3416013.3426457