Development of an Anomaly Prediction System for Multivariate Time-Series from Sensor Data

Thasorn Chalongvorachai
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
58070047@kmitl.ac.th

Kuntpong Woraratpanya
Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand
kuntpong@it.kmitl.ac.th

***Abstract* – A long short-term memory recurrent neural network (LSTM-RNN) is applied for an anomaly detection in multivariate time-series from sensor data. This paper extends its ability to an anomaly prediction by means of a fuzzy logic technique for data modification. The experimental results give us satisfaction in a low root-mean-squared error (RMSE) and the anomaly events are predicted.**

Index Terms – LSTM-RNN, Anomaly prediction, Time-series data, Sensor data, Multivariate time-series data

I. Introduction

In the past, industrial sensors were installed in machinery to detect anomaly events or malfunctions and then alarm to engineers, technicians, or workers who were responsible to those problems. The problem-solving was based on human intelligence. Nowadays, artificial intelligence (AI) becomes a new wave technology of predictive diagnostics in current industries. It can detect anomaly incidents before the actual event happens in the future. One of the existing works on anomaly detection in time-series data is the use of a long short-term memory recurrent neural network (LSTM-RNN) [1]. This paper applied the recurrent neural network (RNN) and LSTM-RNN to various time-series datasets for an anomaly detection performance comparison. Nevertheless, the solution of this paper is in the form of detection, not a prediction. In other words, this approach can only detect abnormal incidents but cannot forecast anomaly events which may happen in the future. Detection is insufficient for anomaly prevention monitoring in industries that require online learning for real-time prediction and updating.

As mentioned above, it becomes a challenge to extend the ability of LSTM-RNN for multivariate time-series sensor data-sets. Responding to this challenge, we apply the fuzzy logic technique to the dataset by providing a suitable membership function for event labelling. This makes the labeled event possible to forecast and notified anomaly event in advance.

II. Methodology

This section describes the proposed method as schematically shown in Fig. 1. In the scheme, a real sensor multivariate time-series dataset is modified with a fuzzy logic method to make the model more suitable with the problem-solving. Then the LSTM-RNN is used to train the modified dataset. An overall framework can be visualized as shown in Fig. 1.

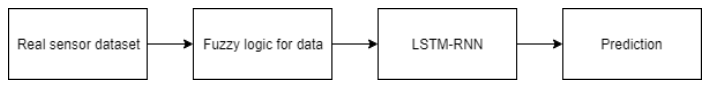

Fig. 1 An overall framework of an anomaly prediction system for multivariate time-series from sensor data.

A. Modification of a real sensor dataset with fuzzy logic

A real dataset is collected from the sensors installed in machinery, especially the motor used for a cooling system in the power plants. This dataset is a multivariate time-series platform that contains 15 features from 15 sensors classified into 9 temperature sensors, 2 vibration sensors, 2 electricity power sensors, 1 electricity current sensor, and 1 humidity sensor. This dataset gives us detected values from each sensor and time ranges of anomaly events when the cooling system starts working abnormally. It contains one anomaly event occurred from a complete breakdown of machinery and two abnormal events occurred from system shutdown by engineers before the actual abnormal event happens.

According to the previous approach [1], the algorithm can only detect anomaly events, but cannot forecast abnormal incidents in advance. In order to make the algorithm predicttive, the fuzzy logic technique is applied to label the dataset. With the fuzzy logic labelling, data can indicate the increase of anomaly incidents sign before actual event occurs in regression form, so prediction and notification of the abnormal event in advance are easily detected. The event label is defined by membership function in (1).

$$f(x) = \begin{cases} 0 & ,x < a \quad or \quad x > c \\ \frac{x-a}{b-a} & ,a \le x \le b \\ 1 & ,b \le x \le c \end{cases} \tag{1}$$

where x is the current time in time-series data, a is the beginning of anomaly event, b is the starting point of the absolute abnormal event, and c is the ending point of the anomalous incident as shown in Fig. 2.

As a result, the machinery status is labeled with 0 if the machine works normally. On the other hand, the machinery status is labeled with 1 if the machine is breakdown.

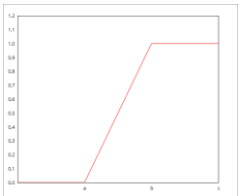

Fig. 2 A membership function for event labelling.

In order to validate our approach, we modified data when the sign of breakdown appeared with an interval between 0 and 1 by giving additional constant values per minutes.

B. *Long Short-Term Memory – Recurrent Neural Network*

The concept of a feed-forward neural network is the state-of-art solution for artificial intelligence prediction in various works. However, the feed-forward neural network has limitation. It can work only with non-sequential data, but cannot work well with sequential data such as time-series. In 1990, Elman proposed a recurrent neural network (RNN) solution [2]. To summarize the concept of this solution, he used the output as input again in the hidden state. This would make a neural network able to learn with sequential data. However, this approach still has a vanishing gradient problem. This means that the longer data will not be updated by the model. In 1997, Hochreiter proposed long-short term memory recurrent neural network [3]. They claimed that his model is more immune to the vanishing gradient problem. The concept behind architecture is using a cell state and four gates for calculation, including input gate, forget gate, update gate, and output gate. These gates are for updating the next hidden state to eliminate the vanishing gradient and return effective output.

III. Experiment

In the experiment, inputs for training and testing are variables from 15 real sensor data with event labeling. The data are divided into 62.5% for training and 37.5% for testing. Our models are developed with Python language and Keras library and are tested on Notebook with processor Intel(R) Core(TM) i7-4720HQ CPU 2.60 GHz, 2601 MHz, 4 Core(s), 8 Logical Processor(s). After training and testing with different units, batch size and layers; the results of root mean squared error are compared to each model as reported in Table 1.

TABLE I: Performance comparisons of test models.

Model	LSTM-RNN Units*	Epoch	Batch size	RMSE	Computational Time (Second)
a	(100)	100	100	0.02758	716
b	(100-100)	100	100	0.00886	1571
c	(100)	100	50	0.06096	1560
d	(100-100)	100	50	0.00536	2722
e	(100-100)	200	50	0.01849	3112
f	(200)	100	50	0.03055	3451
g	(200-100)	100	50	0.01520	5150
h	(200-200)	100	50	0.01636	5886

*(200-100) can be interpreted as 200 units in the first hidden layer and 100 units in the second hidden layer.

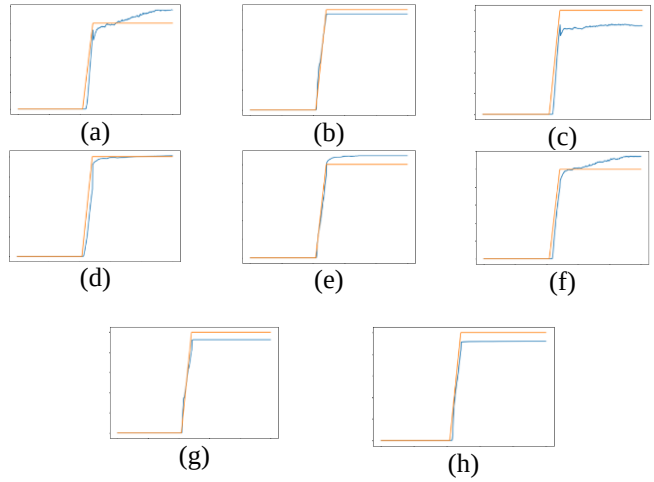

Fig. 3 Comparisons of actual abnormal data (orange) and predictions results from models (blue).

In prediction task, we show the visualization of a normal event compared with an abnormal event as shown in Fig. 3.

According to models (b) and (d), with a low RMSE and similar result of a prediction line compared with a test line, we could say that this model returns satisfaction results.

IV. Conclusion

In this study, we introduce an anomaly prediction method for multivariate time-series sensor data. A LSTM-RNN architecture in conjunction with a fuzzy-logic technique is implemented for anomaly event prediction. The results are satisfactory. However, this work still requires a diversity of abnormal event datasets for training and testing in order to improve the model. In future work, we are going to improve our model by modifying another architecture neural network such as GRU-RNN. We also eager to develop the online learning method by LSTM-RNN for prediction and updating in real-time.

References

[1] P. Malhotra, L. Vig, G. Shroff, P. , Long Short Term Memory Networks for Anomaly Detection in Time Series, Agarwal, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN, 2015, pp. 89-94

[2] J. L. Elman, Finding Structure in Time, Cognitive Science, University of California, 1990, pp.179-211.

[3] S. Hochreiter and J. Schmidhuber, Long Short-Term Memory, Neural Computation, 1997, pp. 1735-1780.