九州大学学術情報リポジトリ
Kyushu University Institutional Repository

Anomaly Detection Using LSTM-Autoencoder to Predict Coal Pulverizer Condition on Coal-Fired Power Plant

Henry Pariaman
PT Pembangkitan Jawa Bali

G M Luciana
PT Pembangkitan Jawa Bali

M K Wisyaldin
PT Pembangkitan Jawa Bali

Muhammad Hisjam
Industrial Engineering Department, Faculty of Engineering, Sebelas Maret University

https://doi.org/10.5109/4372264

出版情報：Evergreen. 8 (1), pp.89-97, 2021-03. Transdisciplinary Research and Education Center for Green Technologies, Kyushu University
バージョン：
権利関係：Creative Commons Attribution-NonCommercial 4.0 International

KYUSHU UNIVERSITY

Anomaly Detection Using LSTM-Autoencoder to Predict Coal Pulverizer Condition on Coal-Fired Power Plant

Henry Pariaman[1,2], G M Luciana[1,*], M K Wisyaldin[1], Muhammad Hisjam[3]
[1]PT Pembangkitan Jawa Bali, Indonesia
[2]Institut Teknologi PLN, Indonesia
[3]Industrial Engineering Department, Faculty of Engineering, Sebelas Maret University, Indonesia

*Author to whom correspondence should be addressed:
E-mail: gita.maya@ptpjb.com

(Received October 12, 2020; Revised March 19, 2021; accepted March 25, 2021).

Abstract: Coal pulverizing systems reliability can be ensured effectively by using prognostics and health management approach. A mathematical model of coal pulverizing system used for anomaly detection is hard to be constructed due to its dynamic and nonlinear high-dimensional system typically. This paper proposed the use of the Long-Short Term Memory Autoencoder model for anomaly detection of the coal pulverizing system on a coal-fired power plant. The LSTM will solve the gradient reduction problem, and Autoencoder will improve the generalizability of the model. As a result, the proposed model can detect the anomaly successfully before the Sequent of Events occurs.

Keywords: Long-Short Term Memory; Autoencoder; Pulverizer; Anomaly Detection

1. Introduction

Nowadays, industrial technology development and engineering systems complexity are increasing significantly[1–4]. System reliability and security are of great concern because they are driven by the increase of system failure costs[5]. Therefore, prognostic and health management (PHM) is a method that is widely studied and applied in industry to resolve this concern[3,4,6].

PHM is a method that can assess system reliability under actual application conditions. The system health status (for fault diagnosis), life prediction, and condition-based maintenance, which guides decision-making and reduces use and maintenance costs are monitored using equipment-mounted by the sensors[7,8].

Paiton 1-2 Coal-Fired Power Plant is a sub-critical thermal power plant constructed in 1993 and started to produce electricity in 1994. It has been designed to consume high coal quality rank (sub-bituminous A) with the calorific value of 6000 kcal/kg and the thermal efficiency about 37.84%[9]. To compete with other new power plants that have high efficiency to get the maximum dispatched load, Paiton performs a strategy to switch the coal calorific value with LRC (low-rank coal) so the production cost in Rp/kWh (BPP) becomes lower. Coal switching will give impact to the power plant equipment reliability. One of the major equipment affected by this improvement is the coal Pulverizer.

The coal Pulverizer system is important in coal-fired power plant since it is used to set up the qualified pulverized coal that meets plant combustion requirements[10]. To keep maintaining the coal Pulverizer work in normal condition is very important to ensure the safety and economical of the power plant. Prognostics and health management (PHM) is a potent method to analyze the reliability of the coal pulverizing system[11]. The PHM will observe the coal pulverizing system to work in conditions based on condition monitoring (CM) data and detect the anomaly of the equipment[12].

Generally, there are two classifications of anomaly detection approaches: model-based approaches and data-driven approaches[13]. In this case, the data-driven approach will have more concern. This approach uses historic CM data to bring up information finding and accurate decision-making[14]. The signal processing and analysis approach has been widely used among the data-driven approaches. The key principle of the signal processing and analysis strategy is to gain diagnostic information and generate a health index (HI) to detect anomalies. Machine Learning (ML) techniques have the potential to effectively detect this kind of anomalies as they can learn from past behavior, making them more efficient at detecting anomalies, as they are trained to observe data which do not fit with the previously observed behaviors[15,16]. Lin uses the ARIMA algorithm to predict the fault of a throttle valve[17]. Lin predicts the degradation trend of rotating machinery using a particle filter

algorithm[18]. Hence, machine learning techniques have the capability to handle high-dimensional data, to have appeal to more consideration. Choi used a long short-term memory (LSTM) network to find sensor faults[19]. Que proposed an XGBoost-based framework to detect anomalies in steam turbines[20]. Tarek proposed of using a neural-networks to do system identification for Quad-rotor parameter[21]. Artificial Neural Network (ANN) is one of the decision-making systems[22]. ANN is made up of artificial neurons configured with a complex interconnection that maps the inputs and outputs[23]. The majority of this machine learning approach belongs to supervised learning, which is sensitive to the quantity and quality of training data. The more the model is trained, the better the performance is gained. However, the machine learning process requires that the model is trained on the exact input and output datasets[24]. Supervised learning needs the training data in all working conditions and the amount of normal and abnormal data labels are identical to each other to improve performance.

The discrepancy analysis is established to overcome the problems mentioned above as a type of semi-supervised learning method. The anomaly detection method focused on discrepancy analysis generally has two modules: predicting working conditions and online detecting anomalies. In the field of predictive working conditions, several machine learning methods are suggested. Stief suggested an approach for sensor fusion based on a two-stage Bayesian method and principal component analysis (PCA) to condition monitoring induction motors[25]. These approaches, however, deal only with CM data as discrete times and neglect temporal information. A typical dynamic system with a large number of time details is the coal pulverization system. The recurrent neural network (RNN) and its variety were commonly used to consider temporal knowledge. To monitor machine health, Zhao combined CNN and bi-directional LSTM[26]. Qian suggested using the LSTM algorithms as a novel condition-monitoring approach for wind turbines[27].

In this paper, we purposed to combine the LTSM and Autoencoder methods to predict the condition of coal crusher in a coal-fired power plant by looking at the anomalies that occur in the equipment. LSTM will solve the problem of gradient reduction if the period is large, and then Autoencoder will improve the generalizability of the model by reducing the data[28].

2. Literature Review

2.1 Control Chart

Customer expectations can be fulfilled if a product is produced with a stable and repeatable process because the process will produce little variability around the target or nominal dimensions of product quality characteristics. Statistical process control (SPC) is an effective set of problem-solving tools, beneficial in making the process' stability and enhancing capabilities through reduced variability[29].

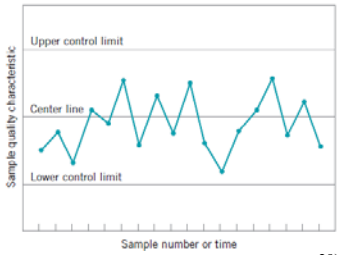

Fig. 1: Illustration of control chart[29].

Graphical display of a quality characteristic is represented by the control chart (see figure 1 above) to show that it that has been measured or computed from a sample quality characteristic versus the sample number or time. The control chart contains three horizontal lines which are a centerline, the upper control limit (UCL) line, and the lower control limit (LCL) line. The centerline is the average value of the quality characteristic corresponding to the in-control state. The process is assumed to be controlled if all sample points are between two control limits (UCL & LCL). When the sample points fall outside two control limits, it indicates the incontrollable process, so inspection and corrective action is needed to discover and abolish the determinable cause or causes responsible for this behavior[29,30]. Figure 1 shows the illustration of a typical control chart.

In general, the Upper and Lower Limit Control Chart formulations are as follows,

$$\text{UCL} = \text{center value} + \text{constanta} * \text{variance} \quad (1)$$

$$\text{LCL} = \text{center value} - \text{constanta} * \text{variance} \quad (2)$$

The process of determining the central value and variance varies is relied on the type of data used. Remote sensor data tends to be time-dependent/time series data. Creating Upper Limit and Lower Limit requires time series chart control methods. One of the methods that can be used in the case of time series data is LSTM-Autoencoder.

2.2 Long-Short Term Memory

LSTM is a type of deep learning that is especially used for time series. Deep learning can be approximate any complex function form and find a linear relation between non-linear data. Deep learning allows us to explore hidden relationships between data, then the maximum potential of data can be used. Deep learning has more benefits that cannot be provided by conventional machine learning. LSTM can avoid long-term dependencies by deliberate design[31].

As shown in Figure 2, the LSTM cell is mainly composed of 3 gates which are forget, input, and output gates. Sigmoid (g) is used to activate the gate then the status and gate input unit are converted to tan h. The weight of neurons and the bias are denoted by w and b, respectively. The LSTM gate can be expressed by the following equation[32].

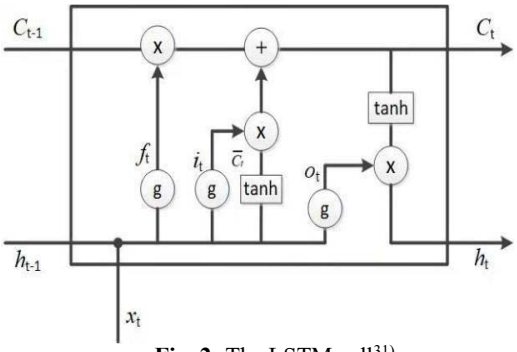

Fig. 2: The LSTM cell[31].

Forget gate : $f_t = g(w_f \cdot [h_{t-1}, x_t] + b_f)$ (3)

Input gate : $i_t = g(w_i \cdot [h_{t-1}, x_t] + b_i)$ (4)

State update : $\overline{C_t} = \tan h\,(W_c \cdot [h_{t-1}, x_t] + b_c)$ (5)

$C_t = f_t \cdot C_{t-1} + i_t \cdot \overline{C_t}$ (6)

Output gate : $o_t = g(W_o \cdot [h_{t-1}, x_t] + b_o)$ (7)

$h_t = o_t \cdot \tanh(C_t)$ (8)

2.3 Autoencoder

The impact of features on the final model is without being questioned as a raw material for machine learning systems. The performance of machine learning algorithms largely relies on the representation of the data or representation of the feature. Once the data can be characterized as features well, even with a simple model, good accuracy can be achieved. Therefore, an important aspect is how to prepare data before using machine learning algorithms to achieve good expression. Autoencoder (AE) is used to decrease the dimensions of the data to solve these problems. Autoencoders can learn to represent input data effectively through unsupervised learning. This process is referred to as encoding and has a smaller size than the input data, so AE can be used to decrease the dimension[31].

Figure 3 shows the input layer on the leftmost side, the hidden layer on the middle, and the output layer on the far. Suppose the input is X (x1, x2, x3, ⋯, xn), and xi ∈ [0,1], AE maps X to an implicit layer, denoted as H = (h1, h2, h3, ⋯, hs), and hi ∈ [0,1], this process is defined as Encode. The hidden layer output[31] is shown in the following equation,

$$H = g(w_1 \cdot X + b_1) \qquad (9)$$

The output H of the hidden layer is called an implicit variable, and the implicit variable reconstructs the Z of the output layer with the same structure as the X of the input layer. It is known as Decode. The output of the output layer[31] in the following equation,

$$Z = g(w_2 \cdot X + b_2) \qquad (10)$$

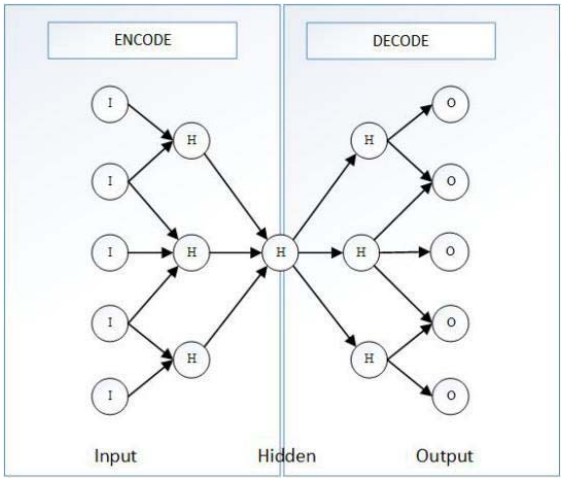

Fig. 3: Autoencoder[31].

Z is a prediction of data X using feature H. AE is using to decrease the data dimension and H as the result. The loss function shows the distinction between X and Y. By decreasing the value of the loss function, the weight between layers is adapted. It eliminates the loss of useful information during the reduction of dimensions[31].

3. Material and Method

3.1 Reconstruction time series data patterns using LTSM-Autoencoder

Autoencoder is one of the major branches of Deep Neural Network that is used to reconstruct the entered data. Autoencoder is widely used as a noise removal in pictures[32,33]. The anomaly detection algorithm compares the data with the prediction model. If the difference between the data and the prediction model is more than the defined limits, the data can be said as the anomaly data[34]. Autoencoder can also be used to detect anomalies by training models from normal data. When abnormal data is found, this model cannot be reconstructed into input data. Reconstruction that occurs will be far from the input data which makes the gap between input data and reconstruction data. After calculating the gap value, it will be determined whether the data are included in the anomaly.

One of the methods for reconstructing time series data patterns is Long-Short Term Memory. Long-Short Term Memory (LSTM) is a specialized type of neural network, which falls within the class of Recurrent Neural Network (RNN). Unlike the conventional feed-forward neural network architectures, RNNs employ feedback connections from their output layers back to the input layers, where each of these feedback connections can be used to serve as a time-delay gate[35,36]. Therefore, RNN architecture can represent explicitly the influence of past output values on the computation of the current output, making it ideal to model the autocorrelation structure of time series or sequence data.

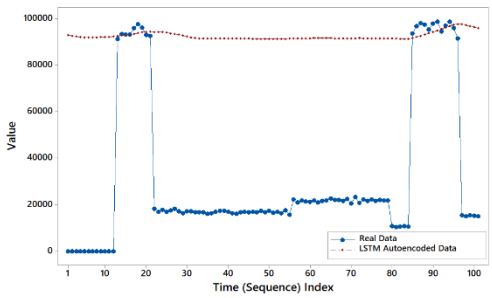

Fig. 4: Illustration of LSTM Autoencoder Application.

Pattern recognition model generated by LSTM can predict future value from previous behaviour data. The model generation by the LSTM method starts by accumulating the data history of the sensors. In this present study, we used 2 years of historical data. Then we did cleanse the data to evaluate abnormal data by defining the normal data using engineer justification. The engineer also mapped the connection between sensors that affected the failure of the equipment. The tool used in this study was Jupyter notebook (basis phyton) and library Tensor Flow & R. In this study we used 75% of historical data as the input for generating the model, then 25% of data was used as training data prediction. The iteration process was done until the accuracy of the model reached 95%, or the margin error maximum was 5%.

The objectivity of the model could be validated from the SOE (Sequent of Event) that has occurred. LSTM learnt how data behaved before SOE occurred. Then the data was iterated to produce a small margin of error. A small margin of error can be tested by dividing historical data into training data for model building and data testing (input data on the model that has been created). The LSTM output is a predictive value with a certain period (for example, every 5 minutes to 12 times) so that the LSTM value can predict the value up to 60 minutes ahead[37].

Table 1 shows that the LSTM model has the smallest MAPE (mean absolute percentage error) value compared to the conventional ARIMA and RNN models which is 3.8%. This shows that the LSTM model produces more accurate predictions than the conventional ARIMA and RNN methods. However, the LSTM method has a longer training time compared to other methods[37]. This issue can be resolved by using a more capable server computer[37].

Table 1. MAPE & Training Time for Each Model[37].

Algorithm	MAPE (%)	Training Time (second)
SARIMA (1,1,0) [i.e., ARIMA with seasonality tweak]	13,279.97	1.27
RNN with 2 layers: 120 & 60 neurons	5.59	3,403
LSTM with 2 layers: 120 & 60 neurons	3.80	13,579

3.2 Create Upper Limit and Lower Limit

After reconstructing the pattern using the LSTM Autoencoder, the next step is creating the Upper Limit and Lower Limit of the LSTM Autoencoder data. The Upper Limit and Lower Limit are determined based on the output of the autoencoder signal, then the average values and standard deviations are calculated so that the Upper Limit (UL) and Lower Limit (LL) are obtained.

Formulations for obtaining Upper Limit and Lower Limit are stated as follows,

$$\bar{x} = \sum_{i=1}^{w} \frac{1}{w} \; x_i \qquad (11)$$

$$std = \sqrt{\frac{\sum_{i=1}^{w} \frac{1}{w} \; (x_i - \bar{x})^2}{w-1}} \qquad (12)$$

$$SE = \frac{std}{\sqrt{w}} \qquad (13)$$

$$UL = \bar{x} + 3SE \qquad (14)$$

$$LL = \bar{x} - 3SE \qquad (15)$$

With:

w : specified time window (18 times window)
$\bar{x}$: average of the last (w) values
std : standard deviation
UL : Upper limit
LL : Lower limit

The formulations are based on the average and standard deviation moves each time with a time window of 18. The constant value is equal to 3 because the value is commonly used in Statistical Process Control which represents 3σ. Illustration of UL and LL results is shown in Figure 5.

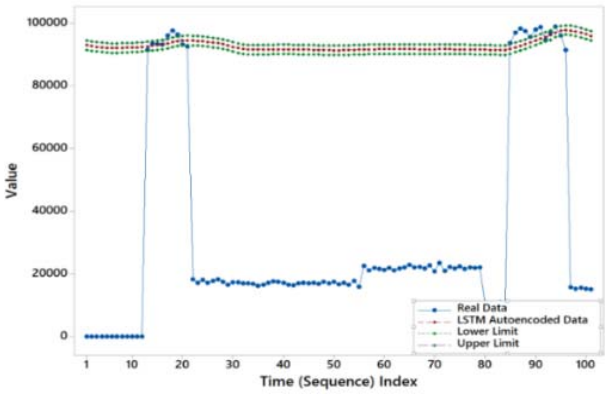

Fig. 5: Illustration of LSTM Autoencoder Application and Upper Limit and Lower Limit Result.

4. Result and Discussion

This study employed Pulverizer-A equipment as a trial to implement anomaly detection. The selection of this equipment was based on data validation in the form of a Sequence of Event (SOE). SOE will simplify the anomaly detection validation process that has been done. The sensor data history used was in the period of December 2016 to June 2017. The number of Pulverizer-A sensors available that used was 24 sensor tags with data sensor history per 30 minutes. The steps to anomaly detection on Pulverizer-A equipment are as follows,

4.1 Perform autoencoder modeling of each sensor data to get the Upper Limit and Lower Limit control chart

The Upper Limit and Lower Limit control were the maximum and minimum limit values of equipment operating parameters. Determination of the operating limits was needed to determine data that exceeded the normal limits of its operations. Figure 6 shows the result of the Upper Control Limit and Lower Control Limit from the autoencoder modeling of the Pulverizer motor current sensor (IT-BY001). The Upper Limit and Lower Limit are represented by the blue dash line while the solid line represents the real data of the Pulverizer motor current sensor (IT-BY001).

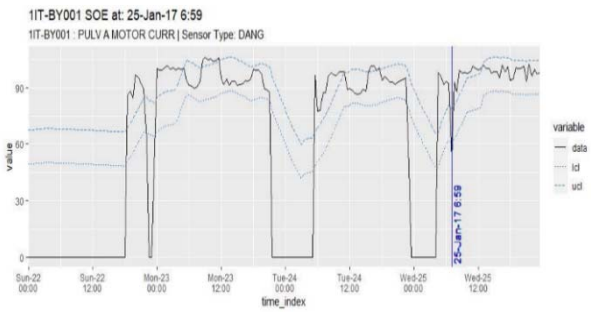

Fig. 6: Upper Limit and Lower Limit Autoencoder Sensor IT-BY001.

4.2 Accumulated amount of data that exceeds the Upper Limit and Lower Limit

The aims of the accumulated amount of data that exceeds the upper limit and lower limit control were to reduce misinformation in making odd decisions because not all data that exceeded the limit was always an anomaly, it could be just a spike. The accumulation applied in this study was per day.

Figure 7 shows the accumulation of the amount of data exceeded the upper limit for the Pulverizer motor current sensor (IT-BY001). Maximum accumulation of the amount of data that exceeded the upper limit per day of Pulverizer motor current sensor (IT-BY001) period December 2015 until July 2016 was 20. The sample accumulation of the amount of data that exceeded the upper limit for Pulverizer's sensors was tabulated in Table 2.

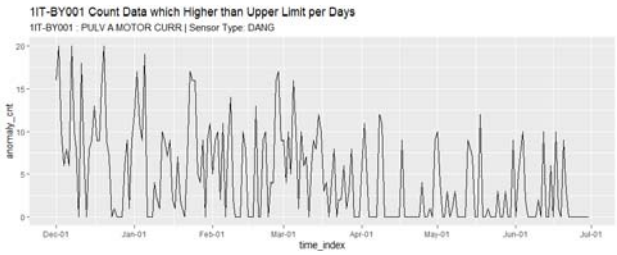

Fig. 7: Accumulation of Amount of Data Exceeds the Upper Limit for the IT-BY001 Sensor.

Table 2. Samples of Accumulation of Amount of Data above the Upper Limit on Some Sensors.

Date	1FI BA084	1FT BA560A	1FT BA560B	1FT BA560C	1FT BY517A	1FT BY518A	1IT BY001	1PDT BA087
12/1/2016	7	10	10	10	16	34	16	7
12/2/2016	13	17	15	15	19	26	20	16
12/3/2016	8	8	8	9	9	16	10	10
12/4/2016	5	7	8	8	7	11	6	7
12/5/2016	6	9	8	8	8	9	8	8
12/6/2016	7	7	7	7	7	14	6	6

4.3 Creating a threshold for anomaly identification based on the daily historical accumulation

The threshold was used to make criteria for an event to be classified into an anomaly. The threshold used was based on a 95% quintile of daily accumulation data that have cases that exceeded the limit. If an observation is equal to or exceeds the threshold, the observation is an anomaly. The threshold is made for both upper limit and lower limit control. Figure 8 and Figure 9 show observations that are anomalous and not anomalous.

The threshold accumulated amount of data above the Upper Limit Pulverizer motor current sensor (IT-BY001) is shown in Figure 8. Figure 8 shows that the threshold quantile for the Upper Limit is around 17. Then the accumulated amount of data that equals or exceeds 17 is identified as an anomaly. As shown in Figure 8, the anomaly occurred in December 2016, January 2017, and February 2017.

The threshold accumulated amount of data under the Lower Limit Pulverizer motor current sensor (IT-BY001) is shown in Figure 9. Figure 9 shows that the threshold quantile for the lower limit is around 48. The accumulated amount of data that exceeds 48 is identified as an anomaly. As shown in Figure 9, no data that exceed the Lower Limit threshold accumulated of the amount of data. There was no anomalous captured event notified as lower than the Lower Limit threshold.

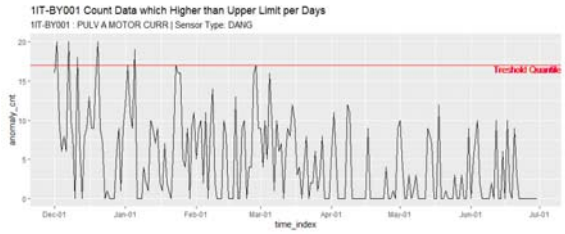

Fig. 8: Threshold Accumulated Amount of Data above the Upper Limit IT-BY001 Sensor.

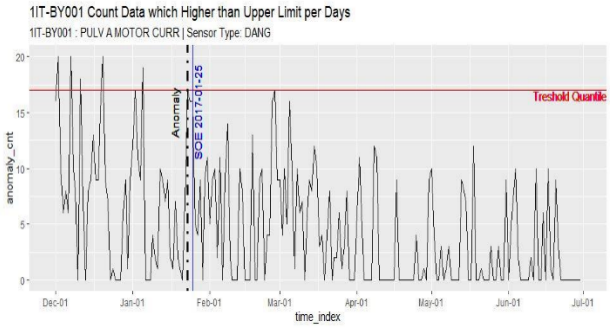

Fig. 9: Threshold Accumulation of Amount of Data under Lower Limit for IT-BY001 Sensors.

4.4 Validation of anomaly detection results

Validation of anomaly detection results was done to see whether the anomaly detection model can be captured before the actual event occured. We expected some sensors to capture anomalous events occurred before 1-3 days of SOE. Equipment Pulverizer-A case had 4 SOE. Details of the anomalous event date and the sensors that captured the pattern can be seen in Table 3.

Table 3 and Figure 10 show that anomaly detection focused on the data above the Upper Limit of the Pulverizer motor current sensor (IT-BY001) because accumulations below the Lower Limit tend to be unable to capture anomalies 1-3 days before SOE (see Figure 9). This study used real data from December 2016 until July 2017. As shown in Figure 10, the anomaly data captured on 23 January 2017, 2 days before SOE occurred (25 January 2017). The anomaly captured in Figure 10 is valid because it occurred 1 to 3 days before SOE. Conditions of accumulation before and after the occurrence of SOE were noticeable that many cases above the threshold just before the SOE. However, after SOE occurred it tended to be fewer cases above the threshold. This indicated that the Pulverizer's problem has been solved after SOE occurred. The pattern also occurred in several other sensors. They can be seen in Figure 11, Figure 12, and Figure 13.

Table 3. List of Sensor Tags Detecting Anomalies About 1-3 Days before SOE (Based on Upper Limit Threshold).

Date Captured	Sensor Tag	SOE Date
11 January 2017	1FI-BA084	12 January 2017
23 January 2017	1IT-BY001	25 January 2017
23 January 2017	1PDT-BA087	25 January 2017
23 January 2017	1TE-BY586	25 January 2017
24 January 2017	1PDT-BA087	25 January 2017
18 February 2017	1FI-BA084	19 February 2017
18 February 2017	1PDT-BA087	19 February 2017
18 February 2017	1TE-BY584	19 February 2017
18 February 2017	1TE-BY586	19 February 2017
18 February 2017	1FI-BA084	20 February 2017
18 February 2017	1PDT-BA087	20 February 2017
18 February 2017	1TE-BY584	20 February 2017
18 February 2017	1TE-BY586	20 February 2017

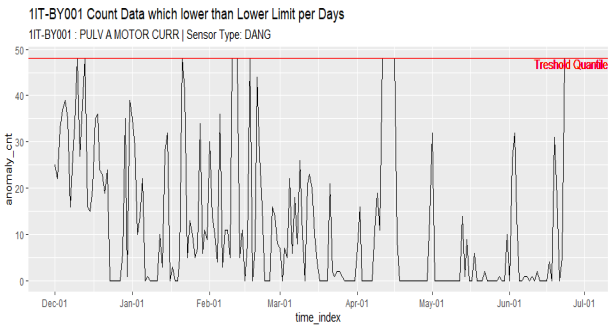

Fig. 10: Anomaly Captured on IT-BY001 Sensor 1-3 Days Before SOE Occurs (Based on the Accumulation of Data Amount above the Upper Limit).

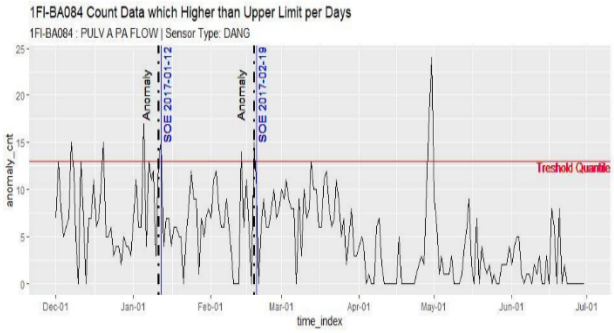

Fig. 11: Anomaly Captured on IF-BA084 Sensor 1-3 Days Before SOE Occurs (Based on the Accumulation of Data Amount above the Upper Limit).

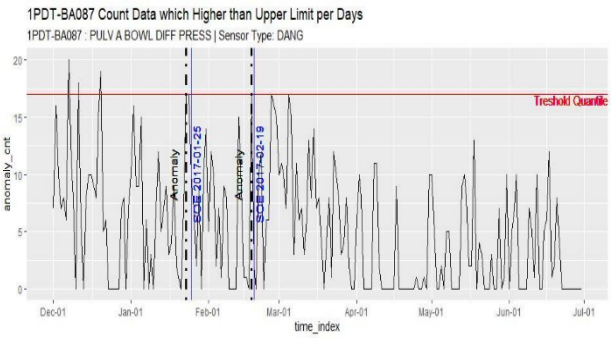

Fig. 12: Anomaly Captured on PDT-BA087 Sensor 1-3 Days Before SOE Occurs (Based on the Accumulation of Data Amount above the Upper Limit).

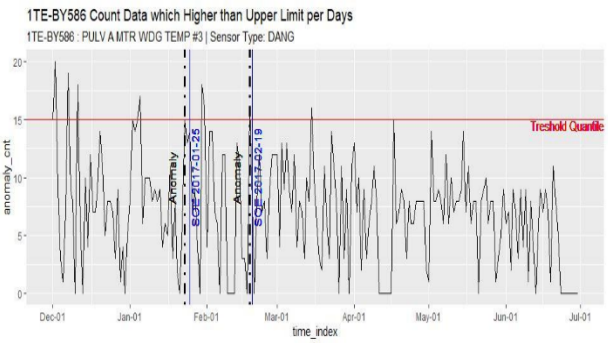

Fig. 13: Anomaly Captured on TE-BY586 Sensor 1-3 Days Before SOE Occurs (Based on the Accumulation of Data Amount above the Upper Limit).

Anomaly captured on Pulverizer PA Flow sensor (IF-BA084) is shown in Figure 11. Figure 11 shows the anomaly data captured on 11 January 2017 and 18 February 2017, one day before SOE occurred (11 January 2017 and 19 February 2017) respectively.

Anomaly captured on Pulverizer Bowl Different Pressure (PDT-BA087) is shown in Figure 12. Figure 12 shows, the anomaly data captured on 24 January 2017 and 18 February 2017, one day before SOE occurred (25 January 2017 and 19 February 2017) respectively.

Anomaly captured on Pulverizer Motor Winding Temperature (TE-BY586) is shown in Figure 13. Figure 13 shows the anomaly data captured on 24 January 2017 and 18 February 2017, one day before SOE occurred (25 January 2017 and 19 February 2017) respectively. The anomaly captured in Figure 11, Figure 12, and Figure 13 are valid because it occurred 1 to 3 days before SOE. The limitation of SOE data that captured from December 2016 to July 2017 resulted some anomaly data that cannot be validated.

The results that have been illustrated (see Figure 10 to Figure 13) approves that LSTM-Autoencoder can capture anomalous events before SOE happened. An anomaly that captured each sensor shows the current condition or the phenomena that occurred inside of the Pulverizer. After obtaining the anomaly data from each sensor, a symptom-cause logic is built by the engineer. A symptom-cause logic purpose to determine the condition of the equipment. This condition will be taken care by the plant operator to avoid or reduce the impact of the failure that came after. An anomaly that was detected before SOE can be used as an early warning for the operator to prepare and anticipate action before the failure occurs.

5. Conclusion

Reliability is one of the most critical points for engineering operations. Unexpected failures of some parts of the system could affect all of the operations. Failures can be predicted by knowing the anomalies that occurred in the equipment. This study made use of Pulverizer-A equipment as a trial to implement anomaly detection using the LTSM-autoencoder approach that learns to reconstruct 'normal' time-series behaviour and thereafter uses reconstruction error to detect anomalies. As a result, this method can capture anomaly events before SOE occurs (based on the Accumulation of Data Amount above the Upper Limit).

The accuracy of the proposed model points out the reliability of the architecture which might lead the power plant to get an alert before a breakdown happen. Therefore, management might have a chance to reduce maintenance costs while increasing revenue and service quality. Our future research is to calculate the Cost-Benefit Analysis of data unplanned maintenance cost versus data planned maintenance cost (the result of LSTM-autoencoder implementation) to convincing the economic feasibility of this research. The LSTM method has a longer training time compared to other methods. This must be prepared regarding big resources before running the model using LSTM method.

Acknowledgements

The authors gratefully acknowledge the support provided by colleagues of PT PJB UP Paiton.

Nomenclature

b	the bias
$\bar{C}_t$	state update
ft	forget gate
g	sigmoid
H / h	implicit layer
it	input gate
LCL	upper control limit
ot	output gate
$tan\ h$	status and gate input unit
UCL	lower control limit
w	weight of neuron
w	specified time window (18 times window)
X / x	the input
x切	average of the last (w) values
Z	the implicit variable reconstructs

Greek symbols

3σ	statistical process control (-)

Subscripts

std	standard deviation

References

1) D.H. Zhou, M.H. Wei, and X.S. Si, "A Survey on Anomaly Detection, Life Prediction and Maintenance Decision for Industrial Processes," *Acta Automatica Sinica*, **39** (*6*) 711–722 (2014). doi: 10.3724/SP.J.1004.2013.00711.
2) Z. Zheng, M. Wen, and R. Kang, "Belief Reliability: A New Metrics for Products' Reliability," *Fuzzy Optimization Decision Making*, **12** (*1*) 15–27 (2013). doi: 10.1007/s10700-012-9138-5.
3) M. Malinowski, S. Adams, and P. A Beling, "Risk Analysis and Prognostics and Health Management for Smart Manufacturing," Springer, 2010. https://link.springer.com/chapter/10.1007/978-3-030-00114-8_35.
4) X. Han, Z. Wang, Y. He, Y. Zhou, Z. Chen, and D. Zhou, "A Mission Reliability-Driven Manufacturing System Health State Evaluation Method Based on Fusion of Operational Data," *Sensors,* **19** (*3*) 442 (2019). doi: 10.3390/s19030442.
5) Y. He, Z. Chen, Y. Zhao, X. Han, and D. Zhou,

"Mission Reliability Evaluation for Fuzzy Multistate Manufacturing System Based on an Extended Stochastic Flow Network," *IEEE Transaction on Reliability,* **7** 1–15 (2019). doi: 10.1109/TR.2019.2957502.

6) T. Xia, and L. Xi, "Manufacturing Paradigm-Oriented PHM Methodologies for Cyber-Physical Systems," *Journal of Intelligent Manufacturing,* **30** 1659–1672 (2019). doi: 10.1007/s10845-017-1342-2.
7) J. Xu, Y. Wang, and L. Xu, "PHM-Oriented Integrated Fusion Prognostics for Aircraft Engines Based on Sensor Data," *IEEE Sensors Journal,* **14** (*4*) 1124–1132 (2014). doi: 10.1109/JSEN.2013.2293517.
8) Y. Peng, D. Liu, and X.A. Peng, "A Review: Prognostics and Health Management," *Journal of Electronic Measurement Instrumment,* **24** (*1*) 1–9 (2010). doi: 10.3724/SP.J.1187.2010.00001.
9) PT PJB UP Paiton, "Laporan Efisiensi Unit 1-2 April 2019," PT. PJB UP Paiton, 2019.
10) H. K Patel, and J. Purva, "Role of Optimization in Pulverization Process of Thermal Power Plant," *Proceedings of the 2016 International Conference on Control, Instrumentations, Communication and Computational Technology (ICCICCT),* 8–12 (2016). doi: 10.1109/ICCICCT.2016.7987831.
11) F, Guan, W. Cui, L. Li, and J.A. Wu, "Comprehensive Evaluation Method of Sensor Selection for PHM Based on Grey Clustering," *Sensors,* **20** (*6*) 1710 (2020). doi: 10.1007/s10700-012-9138-5.
12) D. Liu, J. Pang, G. Song, W. Xie, Y. Peng, and X. Peng, "Fragment Anomaly Detection with Prediction and Statistical Analysis for Satellite Telemetry," *IEEE Access,* **5** 19269–19281 (2017). doi: 10.1109/ACCESS.2017.2754447.
13) M. Swiercz, and H. Mroczkowska, "Multiway PCA for Early Leak Detection in A Pipeline System of a Steam Boiler–Selected Case Studies," *Sensors,* **20** (*6*) 1561 (2020). doi: 10.3390/s20061561.
14) A. Heureux, K. Grolinger, H. Elyamany, and M. Capretz, "Machine Learning with Big Data: Challenges and Approaches," *IEEE Access,* **5** 7776–7797 (2017). doi: 10.1109/ACCESS.2017.2696365.
15) V. G. Aquize, E. Emery, and F. B. de Lima Neto, "Self-organizing maps for anomaly detection in fuel consumption," *IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, (2017). doi: 10.1109/LA-CCI.2017.8285697.
16) S. Brady, D. Magoniz, J.Murphy, H.Assemy, A. O. Portillo-Dominguez, "Analysis of Machine Learning Techniques for Anomaly Detection in the Internet of Things," *IEEE Latin American Conference on Computational Intelligence (LA-CCI)*, (2018). doi: 10.1109/LA-CCI.2018.8625228.
17) C. Lin, Y. Hsien, and F. Cheng, "Time Series Prediction Algorithm for Intelligent Predictive Maintenance," *IEEE Robotics Automation Letters,* **4** (*3*) 2807–2814 (2019). doi: 10.1109/LRA.2019.2918684.
18) Q. Li, and S. Liang, "Degradation Trend Prediction for Rotating Machinery Using Long-Range Dependence and Particle Filter Approach," *Algorithms,* **11** (*7*) 89 (2018). doi: 10.3390/a11070089.
19) J. Choi, and S. Lee, "Consistency Index-Based Sensor Fault Detection System for Nuclear Power Plant Emergency Situations Using an LSTM Network," *Sensors,* **20** (*6*) 1651 (2020). doi: 10.3390/s20061651.
20) Z. Que, and Z. Xu, "A Data-Driven Health Prognostics Approach for Steam Turbines Based on Xgboost and DTW," *IEEE Access,* **7** 93131–93138 (2019). doi: 10.1109/ACCESS.2019.2927488.
21) T.N. Dief, and S. Yoshida, "System Identification for Quad-rotor Parameters Using Neural Network," *Evergreen,* **3** (*1*) 6-11 (2016). doi: 10.5109/1657380.
22) M. A. Berawi, S. A. O. Siahaan, Gunawan, P. Miraj, and P. Leviakangas, " Determining the Prioritized Victim of Earthquake Disaster Using Fuzzy Logic and Decision Tree Approach," *Evergreen,* **7** (*2)* 246-252 (2020). doi: 10.5109/4055227.
23) N. Weake, M. Pant, A. Sheoran, A. Haleem, and H. Kumar, " Optimising Parameters of Fused Filament Fabrication Process to Achieve Optimum Tensile Strength Using Artificial Neural Network," *Evergreen,* **7** (*3)* 373-381 (2020). doi: 10.5109/4068614.
24) H. Han, M. Hatta, and H. Rahman, "Smart Ventilation for Energy Conservation in Buildings," *Evergreen,* **6** (*1)* 44-51 (2019). doi: 10.5109/2321005.
25) A. Stief, J. Ottewill, J. Baranowski, and M. Orkisz, "A PCA and Two-Stage Bayesian Sensor Fusion Approach for Diagnosing Electrical and Mechanical Faults in Induction Motors," *IEEE Transactions on Industrial Electronics,* **66** (*12*) 9510–9520 (2019). doi: 10.1109/TIE.2019.2891453.
26) R. Zhao, R. Yan, J. Wang, and K. Mao, "Learning to Monitor Machine Health with Convolutional Bi-Directional LSTM Networks," *Sensors,* **17** (*2*) 273 (2017). doi: 10.3390/s17020273.
27) P. Qian, X. Tian, and J. Kanfoud, "A Novel Condition Monitoring Method of Wind Turbines Based on Long Short-Term Memory Neural Network," *Energies,* **12** (*18*) 3411 (2019). doi: 10.3390/en12183411.
28) A. Gensler, J. Henze, B. Sick and N. Raabe, "Deep Learning for Solar Power Forecasting – An Approach Using AutoEncoder and LSTM Neural Networks," *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC),* 2858-65 (2016). doi: 10.1109/SMC.2016.7844673.
29) D.C. Montgomery, "Introduction to Statistical Quality Control," John Wiley & Sons, 2009. https://www.wiley.com/en-us/Introduction+to+Statistical+Quality+Control%2C+8th+Edition-p-9781119399308.
30) D.N. Chorafas, "Quality Control Applications,"

Springer, 2013. https://link.springer.com/book/10.1007/978-1-4471-2966-0.

31) J. Zhang, X. Jiang, X. Chen, X. Li, D. Guo, and L. Cui, "Wind Power Generation Prediction Based on LSTM," *Proceedings of the 2019 4th International Conference on Mathematics and Artificial Intelligence,* 85-89 (2019). doi: 10.1145/3325730.3325735.
32) X. Ye, L. Wang, H. Xing, and L. Huang, "Denoising Hybrid Noises in Image with Stacked Autoencoder," *2015 IEEE International Conference on Information and Automation,* 2720-2724 (2015). doi: 10.1109/ICInfA.2015.7279746.
33) H. Chiang, Y. Hsieh, S. Fu, K. Hung, Y. Tsao, and S. Chien, "Noise Reduction in ECG Signals Using Fully Convolutional Denoising Autoencoders," *IEEE Access,* **7** 60806-60813 (2019). doi: 10.1109/ACCESS.2019.2912036.
34) G.D. Nugraha, B. Sudiarto, and K. Ramli, "Machine Learning-based Energy Management System for Prosumer," *Evergreen,* **7** (*2*) 309-313 (2020). doi: 10.5109/4055238.
35) Y. Bengio, P. Simard, and P. Frasconi, "Learning Long-Term Dependencies with Gradient Descent Is Difficult," *IEEE Transactions on Neural Networks*, **5** (*2*) 157-166 (1994). doi: 10.1109/72.279181.
36) S. Hochreiter, and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation,* **9** (*8*) 1735-1780 (1997). doi: 10.1162/neco.1997.9.8.1735.
37) M. K. Wisyaldin, G. M. Luciana, and H. Pariaman, "Using LSTM Network to Predict Circulating Water Pump Bearing Condition on Coal Fired Power Plant," *2020 International Conference on Technology and Policy in Energy and Electric Power (ICT-PEP),* (2020). doi: 10.1109/ICT-PEP50916.2020.9249905.