IAES International Journal of Artificial Intelligence (IJ-AI)
Vol. 11, No. 3, September 2022, pp. 923~929
ISSN: 2252-8938, DOI: 10.11591/ijai.v11.i3.pp923-929

The prediction of the oxygen content of the flue gas in a gas-fired boiler system using neural networks and random forest

Nazrul Effendy[1], Eko David Kurniawan[1], Kenny Dwiantoro[1], Agus Arif[1], Nidlom Muddin[2]
[1]Intelligent and Embedded System Research Group, Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia
[2]Project Development Division, PT. Pertamina (Persero), Jakarta, Indonesia

Article Info

Article history:

Received Sep 25, 2021
Revised Apr 7, 2022
Accepted Apr 21, 2022

Keywords:

Flue gas
Neural networks
Oxygen content
Prediction
Random forest
Soft sensor

ABSTRACT

The oxygen content of the gas-fired boiler flue gas is used to monitor boiler combustion efficiency. Conventionally, this oxygen content is measured using an oxygen content sensor. However, because it operates in extreme conditions, this oxygen sensor tends to have the disadvantage of high maintenance costs. In addition, the absence of other sensors as an element of redundancy and when there is damage to the sensor causes manual handling by workers. It is dangerous for these workers, considering environmental conditions with high-risk hazards. We propose an artificial neural network (ANN) and random forest-based soft sensor to predict the oxygen content to overcome the problems. The prediction is made by utilizing measured data on the power plant's boiler, consisting of 19 process variables from a distributed control system. The research has proved that the proposed soft sensor successfully predicts the oxygen content. Research using random forest shows better performance results than ANN. The random forest prediction errors are mean absolute error (MAE) of 0.0486, mean squared error (MSE) of 0.0052, root-mean-square error (RMSE) of 0.0718, and Std Error of 0.0719. While the errors using ANN are MAE of 0.0715, MSE of 0.0087, RMSE of 0.0935, and Std Error of 0.0935.

Corresponding Author:

Nazrul Effendy
Intelligent and Embedded System Research Group, Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada
Jl. Grafika 2, Yogyakarta, Indonesia
Email: nazrul@ugm.ac.id

1. INTRODUCTION

A steam boiler system is a closed vessel that uses fuel or electricity to generate steam to supply the power plant system [1]–[3]. The energy produced by the boiler in the form of heat comes from the combustion process in the combustion chamber or furnace. In general, the combustion reaction in the boiler furnace can be made using three main components, namely fuel, air, and fire from the lighter.

The combustion efficiency in the boiler furnace describes the ability of a burner to burn the entire fuel entering the furnace. The ideal furnace combustion reaction occurs when the oxygen in the air is sufficient to burn the whole fuel [4]–[6]. So, there is no remaining oxygen or energy in the flue gas. However, the oxygen volume will be insufficient to burn all the fuel when there is an imperfect mixture of fuel and oxygen. Therefore, it is necessary to have an appropriate combination of the required amount of energy and oxygen. In addition, more fresh air is needed to burn the entire fuel. This air is known as excess air. There is still oxygen content in the flue gas in combustion with excess air. Therefore, the amount of unburned fuel and the remaining oxygen excess air on the flue gas can be utilized to estimate the combustion efficiency [7]. In addition to saving fuel use, high combustion efficiency will also reduce air pollution generated from this combustion process [8], [9].

The optimal oxygen content depends on the type of furnace used. If the oxygen content is too low, unburned fuel will decrease air quality. On the other hand, if the oxygen content is too high, the furnace will be inefficient due to a large amount of energy lost through the flue gas.

The oxygen content of the flue gas of a steam boiler system can be conventionally measured by oxygen sensors, such as using zirconium oxide. Zirconium oxide is a material capable of measuring oxygen levels in flue gas. However, oxygen sensors with zirconium oxide tend to have the disadvantage of high maintenance costs. In addition, the absence of other sensors as an element of redundancy and when there is damage to the sensor causes manual handling by workers using portable measuring devices. It is dangerous for these workers, considering the environmental conditions with high-risk hazards.

To overcome the problems, we propose to use a soft sensor using two types of machine learning as the soft sensor: an artificial neural network (ANN) [10]–[12] and random forest [13], [14]. The soft sensor is a software-based method that utilizes an intelligent system to solve a problem based on the input-output relationship [15]–[19]. Several researchers have applied machine learning for prediction and pattern recognition as the primary key of the soft sensor [20]–[25]. Other researchers proposed the soft sensor to indicate the oxygen content of flue gas using the support vector model and mixed model [26].

The rest of this paper is described as follows. First, the research method is described in detail in section 2. Then, section 3 describes the experimental results, including the training and testing of the soft sensor using ANN and random forest. Finally, in section 4, the conclusion of this study is offered.

2. RESEARCH METHOD

In this paper, the proposed soft sensor to predict the oxygen content in the steam boiler flue gas is shown in Figure 1. This research was conducted in several stages: data collection, data preprocessing, soft sensor design using ANN and random forest, training, and performance evaluation of the soft sensor. The schematic diagram of a power plant's boiler is shown in Figure 2 [27].

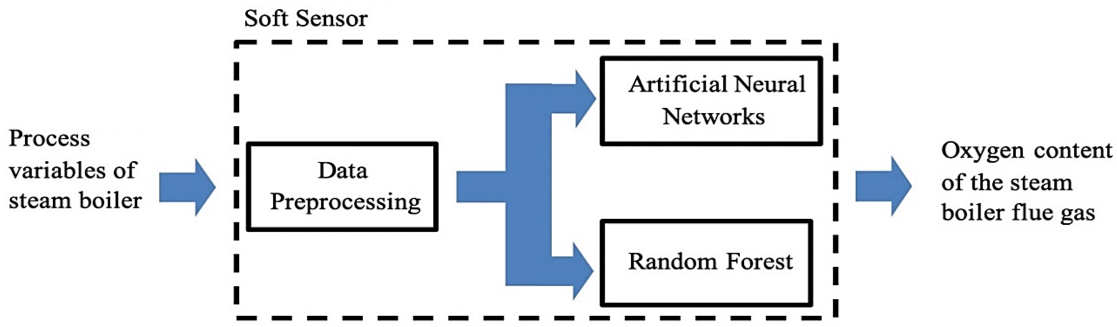

Figure 1. Diagram of the proposed soft sensor system to predict the oxygen content of the steam boiler flue gas using neural networks and random forest

2.1. Data collection

Data from the steam boiler system of a 32 MW power plant in an oil refinery unit, Indonesia, are collected from 1 January until 28 August. The boiler used in this study is a type of water tube boiler used to heat water to become superheated steam. The steam is fed to the steam turbine generator. The generator supplies all power requirements for processing operations at the oil refinery unit. The collected data consisted of 19 parameters, including oxygen content, and was acquired from a distributed control system historical data system. Table 1 lists the process variables of the steam boiler used in the research.

2.2. Data preprocessing

After data collection, the data preprocessing is carried out. The preprocessing consists of several steps: handling missing values, separating training data and test data, and data normalization. The results of the data preprocessing are then used to feed ANN and the random forest system.

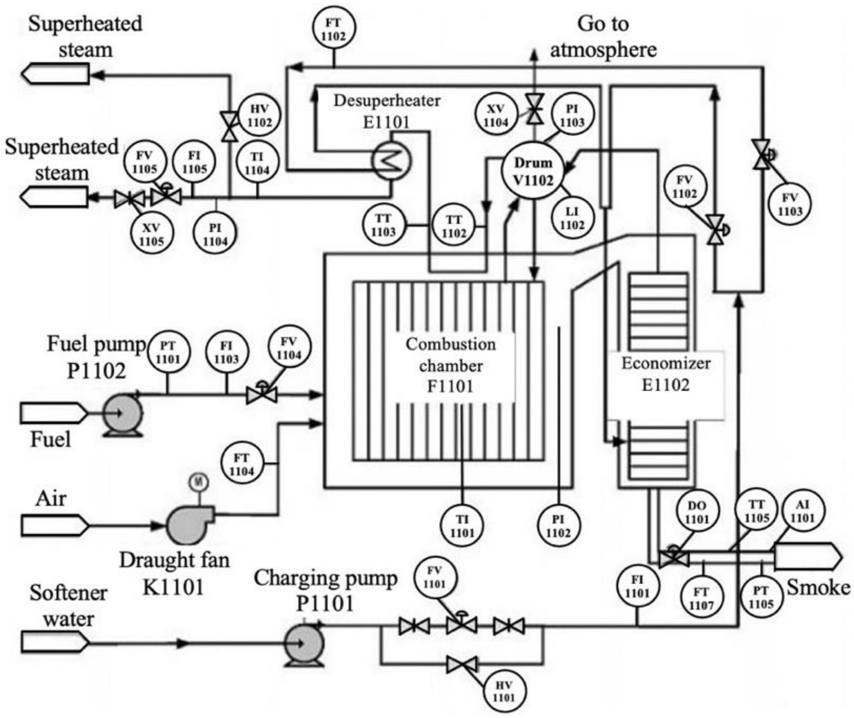

Figure 2. A schematic diagram of a steam boiler system [26]

Table 1. Process variables of the steam boiler system

No	Process variables	Unit
1	Deaerator level	mm
2	The feed flow rate of boiler water to the superheater	ton/hour
3	The temperature of advanced steam in the superheater	°C
4	Feedwater flow rate	Kg/hour
5	Main gas inlet flow rate to the furnace	N.m^3/hour
6	Fuel gas pressure behind the control valve	Kg/cm^2
7	Combustion air flow rate	Kg/ hour
8	Air pressure of burner box	mmH_2O
9	Main steam temperature	°C
10	Furnace exhaust gas pressure	mmH_2O
11	The temperature of boiler Flue gas	°C
12	Boiler steam pressure	kg/cm2
13	Wind box pressure	mmH_2O
14	Combustion air temperature	°C
15	Steam drum boiler levels	%
16	Primary steam header flow rate	kg/hour
17	The water inlet temperature of the economizer	°C
18	The water outlet temperature of the economizer	°C
19	Oxygen content	%

2.3. ANN soft sensor

After preprocessing the data, we design an ANN soft sensor. The ANN soft sensor is created using Python3 with libraries of NumPy, pandas, Matplotlib, Keras, and TensorFlow framework [28]–[30]. The hyperparameters used in the research are the number of neurons, the number of epochs, feature selection, and the early stopping strategy [31], [32]. The hidden layer consists of various neurons from 4 to 64. We used ReLU as the activation function in the hidden layer and mean squared error (MSE) as the loss function. The optimizer used as the determinant of the learning process is SGD [33]–[36]. After ANN soft sensor was designed, it was trained with a certain number of epochs. After the ANN training, the next step is to test the ANN. The test was conducted by comparing the prediction results of ANN with the target.

2.4. Random forest soft sensor

Besides using ANN, we also design a soft sensor using a random forest for the oxygen content prediction of the flue gas. The performance of both models is then compared. We compare the mean absolute error (MAE), MSE, root-mean-square error (RMSE), and Std Error of both models.

3. RESULTS AND DISCUSSION

After training with 1,000 epochs, ANN soft sensor performance was evaluated using test data. The experiments show that the best ANN architecture in this research is with 60 neurons in the hidden layer. Figure 3 shows the MAE of the ANN soft sensor with 60 neurons in the hidden layer. MAE in this experiment tends to decrease with increasing epoch. There is a slight increase in validation errors in certain epochs, even though this validation error generally tends to decrease with increasing epochs.

The experimental results indicate that this ANN successfully predicts the oxygen content of the steam boiler flue gas. Figure 4 shows the histogram of the prediction error of ANN. The distribution of error is typically distributed at the slightest error. It means that most of the prediction results have a relatively small error.

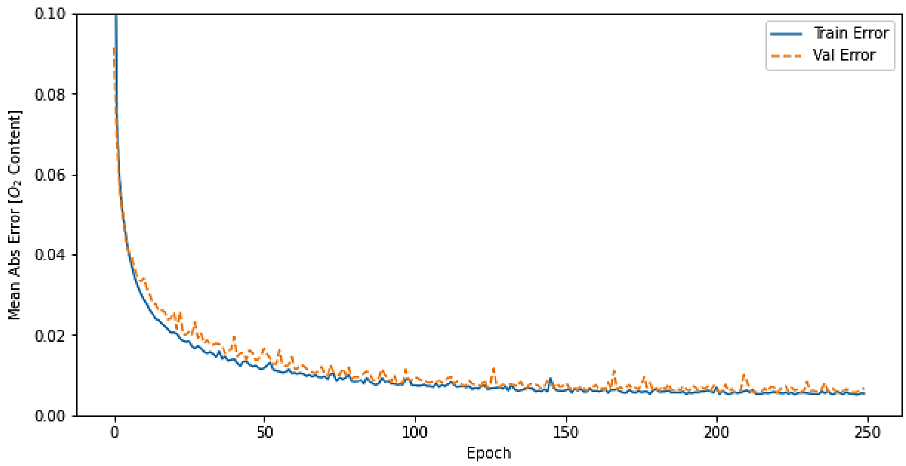

Figure 3. MAE of the ANN soft sensor with 60 neurons in the hidden layer

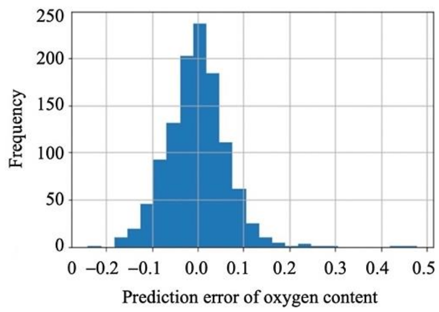

Figure 4. Histogram of the prediction error of the soft sensor with 60 neurons in the hidden layer

In this study, experiments were also conducted on the design of the oxygen content prediction system using a random forest. Figure 5 shows the system prediction error results compared with the oxygen content prediction system using ANN. The results of this study show that the random forest outperforms the ANN. The random forest prediction errors are MAE of 0.0486, MSE of 0.0052, RMSE of 0.0718, and Std Error of 0.0719. While the errors using ANN are MAE of 0.0715, MSE of 0.0087, RMSE of 0.0935, and Std Error of 0.0935. The model performance can also be investigated through the relationship between the flue gas's predicted and measured oxygen content. The relation of the random forest soft sensor is shown in Figure 6. The predicted and the measured values are almost close to the linear line.

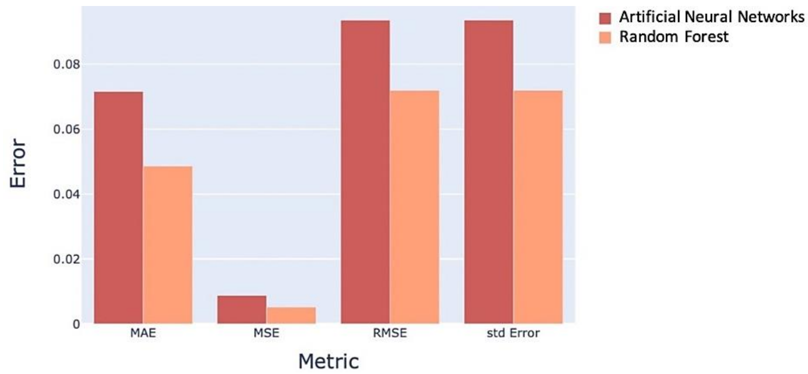

Figure 5. Comparison of the prediction errors of the soft sensors using artificial neural networks and random forest

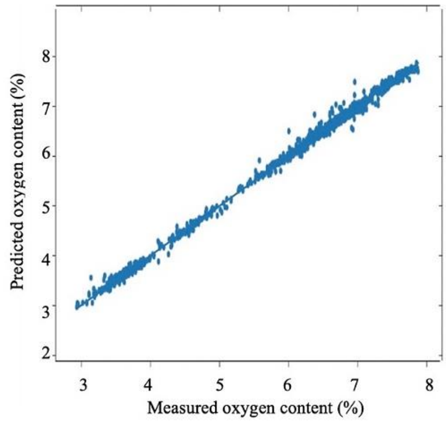

Figure 6. The predicted and measured oxygen content of flue gas using a random forest soft sensor system

4. CONCLUSIONS

In this paper, we propose a soft sensor system to predict the oxygen content of the steam boiler flue gas using ANN and random forest. From the experimental results, this soft sensor system has been proven to be successful in predicting the oxygen content in the flue gas of the steam boiler. The random forest oxygen content prediction system showed better performance than the ANN system. The random forest prediction errors are MAE of 0.0486, MSE of 0.0052, RMSE of 0.0718, and Std Error of 0.0719. While the errors using ANN are MAE of 0.0715, MSE of 0.0087, RMSE of 0.0935, and Std Error of 0.0935.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Universitas Gadjah Mada and PT Pertamina for the facilities while conducting this research. Part of this paper was developed using data from a thesis by the third author.

REFERENCES

[1] Y. Camaraza-Medina, Y. Retirado-Mediaceja, A. Hernandez-Guerrero, and J. Luis Luviano-Ortiz, "Energy efficiency indicators of the steam boiler in a power plant of Cuba," *Therm. Sci. Eng. Prog.*, vol. 23, p. 100880, Jun. 2021, doi: 10.1016/j.tsep.2021.100880.

[2] P. Madejski and P. Żymełka, "Calculation methods of steam boiler operation factors under varying operating conditions with the use of computational thermodynamic modeling," *Energy*, vol. 197, p. 117221, Apr. 2020, doi: 10.1016/j.energy.2020.117221.

[3] M. Trojan, "Modeling of a steam boiler operation using the boiler nonlinear mathematical model," *Energy*, vol. 175, pp. 1194–1208, May 2019, doi: 10.1016/j.energy.2019.03.160.

[4] Y. F. Wang *et al.*, "Fuzzy modeling of boiler efficiency in power plants," *Inf. Sci. (Ny).*, vol. 542, pp. 391–405, Jan. 2020, doi: 10.1016/j.ins.2020.06.064.

[5] J. Luo, L. Wu, and W. Wan, "Optimization of the exhaust gas oxygen content for coal-fired power plant boiler," *Energy Procedia*, vol. 105, pp. 3262–3268, May 2017, doi: 10.1016/j.egypro.2017.03.730.

[6] Y. Cheng, Y. Huang, B. Pang, and W. Zhang, "ThermalNet: a deep reinforcement learning-based combustion optimization system

for coal-fired boiler," *Eng. Appl. Artif. Intell.*, vol. 74, pp. 303–311, Sep. 2018, doi: 10.1016/j.engappai.2018.07.003.
[7] Z. Tang, Y. Li, and A. Kusiak, "A deep learning model for measuring oxygen content of boiler flue gas," *IEEE Access*, vol. 8, pp. 12268–12278, 2020, doi: 10.1109/ACCESS.2020.2965199.
[8] J. Li and W. Li, "Combustion analysis and operation adjustment of thermal power unit," in *Proceedings of the 2015 Asia-Pacific Energy Equipment Engineering Research Conference*, 2015, vol. 9., doi: 10.2991/ap3er-15.2015.1.
[9] Y. Ding, J. Liu, J. Xiong, M. Jiang, and Y. Shi, "Optimizing boiler control in real-time with machine learning for sustainability," in *Proceedings of the 27th ACM International Conference on Information and Knowledge Management*, Oct. 2018, pp. 2147–2154., doi: 10.1145/3269206.3272024.
[10] P. Tóth, A. Garami, and B. Csordás, "Image-based deep neural network prediction of the heat output of a step-grate biomass boiler," *Appl. Energy*, vol. 200, pp. 155–169, Aug. 2017, doi: 10.1016/j.apenergy.2017.05.080.
[11] C. C. Aggarwal, "Training deep neural networks," in *Neural Networks and Deep Learning*, Cham: Springer International Publishing, 2018, pp. 105–167., doi: 10.1007/978-3-319-94463-0_3.
[12] M. S. Khrisat and Z. A. Alqadi, "Solving multiple linear regression problem using artificial neural network," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 1, pp. 770–775, Feb. 2022, doi: 10.11591/ijece.v12i1.pp770-775.
[13] M. B. Nafouanti, J. Li, N. A. Mustapha, P. Uwamungu, and D. AL-Alimi, "Prediction on the fluoride contamination in groundwater at the Datong Basin, Northern China: comparison of random forest, logistic regression and artificial neural network," *Appl. Geochemistry*, vol. 132, p. 105054, Sep. 2021, doi: 10.1016/j.apgeochem.2021.105054.
[14] M. M. Islam, M. A. Kashem, and J. Uddin, "Fish survival prediction in an aquatic environment using random forest model," *IAES Int. J. Artif. Intell.*, vol. 10, no. 3, pp. 614–622, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp614-622.
[15] A. Ramakalyan, A. Sivakumar, C. Aravindan, K. Kannan, V. Swaminathan, and D. Sarala, "Development of KSVGRNN: a hybrid soft computing technique for estimation of boiler flue gas components," *J. Ind. Inf. Integr.*, vol. 4, pp. 42–51, Dec. 2016, doi: 10.1016/j.jii.2016.09.001.
[16] Y. Ye, H. Wang, and X. An, "An indirect online method for measuring the boiling rate of BFG boiler economizer," *Energy Reports*, vol. 6, pp. 703–709, Feb. 2020, doi: 10.1016/j.egyr.2019.11.141.
[17] R. S. Lakshmi, A. Sivakumar, G. Rajaram, V. Swaminathan, and K. Kannan, "A novel hypergraph-based feature extraction technique for boiler flue gas components classification using PNN – A computational model for boiler flue gas analysis," *J. Ind. Inf. Integr.*, vol. 9, pp. 35–44, Mar. 2018, doi: 10.1016/j.jii.2017.11.002.
[18] W. Yan, D. Tang, and Y. Lin, "A data-driven soft sensor modeling method based on deep learning and its application," *IEEE Trans. Ind. Electron.*, vol. 64, no. 5, pp. 4237–4245, May 2017, doi: 10.1109/TIE.2016.2622668.
[19] S. N. Sembodo, N. Effendy, K. Dwiantoro, and N. Muddin, "Radial basis network estimator of oxygen content in the flue gas of debutanizer reboiler," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 3, pp. 3044–3050, doi: 10.11591/ijece.v12i3.pp3044-3050.
[20] S. Nafisah and N. Effendy, "Voice biometric system: The identification of the severity of cerebral palsy using mel-frequencies stochastics approach," *Int. J. Integr. Eng.*, vol. 11, no. 3, pp. 194–206, Sep. 2019, doi: 10.30880/ijie.2019.11.03.020.
[21] N. Effendy, K. Shinoda, S. Furui, and S. Jitapunkul, "Automatic recognition of Indonesian declarative questions and statements using polynomial coefficients of the pitch contours," *Acoust. Sci. Technol.*, vol. 30, no. 4, pp. 249–256, 2009, doi: 10.1250/ast.30.249.
[22] N. Effendy, N. C. Wachidah, B. Achmad, P. Jiwandono, and M. Subekti, "Power estimation of G.A. siwabessy multi-purpose reactor at start-up condition using artificial neural network with input variation," in *Proceedings - 2016 2nd International Conference on Science and Technology-Computer, ICST 2016*, Oct. 2017, pp. 133–138., doi: 10.1109/ICSTC.2016.7877362.
[23] N. Effendy, D. Ruhyadi, R. Pratama, D. F. Rabba, A. F. Aulia, and A. Y. Atmadja, "Forest quality assessment based on bird sound recognition using convolutional neural networks," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 4, pp. 4235–4242, doi: 10.11591/ijece.v12i4.pp4235-4242.
[24] R. Y. Galvani, N. Effendy, and A. Kusumawanto, "Evaluating weight priority on green building using fuzzy AHP," in *Proceedings - 12th SEATUC Symposium, SEATUC 2018*, Mar. 2018, pp. 1–6., doi: 10.1109/SEATUC.2018.8788887.
[25] N. Effendy, E. Maneenoi, P. Charnvivit, and S. Jitapunkul, "Intonation recognition for Indonesian speech based on fujisaki model," in *8th International Conference on Spoken Language Processing, ICSLP 2004*, Oct. 2004, pp. 2973–2976., doi: 10.21437/interspeech.2004-746.
[26] S. Lingfang and W. Yechi, "Soft-sensing of oxygen content of flue gas based on mixed model," *Energy Procedia*, vol. 17, pp. 221–226, 2012, doi: 10.1016/j.egypro.2012.02.087.
[27] Z. Dong, L. Xie, and Q. Zhang, "Design of boiler control system based on PCS7 and SMPT-1000," in *Proceedings - 2015 7th International Conference on Intelligent Human-Machine Systems and Cybernetics, IHMSC 2015*, Aug. 2015, vol. 2, pp. 546–550., doi: 10.1109/IHMSC.2015.212.
[28] A. Gulli and S. Pal, *Deep learning with Keras : implementing deep learning models and neural networks with the power of Python*. Packt Publishing, 2017.
[29] S. Osah, A. A. Acheampong, C. Fosu, and I. Dadzie, "Deep learning model for predicting daily IGS zenith tropospheric delays in West Africa using TensorFlow and Keras," *Adv. Sp. Res.*, vol. 68, no. 3, pp. 1243–1262, May 2021, doi: 10.1016/j.asr.2021.04.039.
[30] F. Chollet, *Deep Learning with Python*, 2nd ed. USA: Manning Publications Co., 2021.
[31] W. Li, W. W. Y. Ng, T. Wang, M. Pelillo, and S. Kwong, "HELP: an LSTM-based approach to hyperparameter exploration in neural network learning," *Neurocomputing*, vol. 442, pp. 161–172, Jun. 2021, doi: 10.1016/j.neucom.2020.12.133.
[32] Y. J. Yoo, "Hyperparameter optimization of deep neural network using univariate dynamic encoding algorithm for searches," *Knowledge-Based Syst.*, vol. 178, pp. 74–83, Aug. 2019, doi: 10.1016/j.knosys.2019.04.019.
[33] G. Habib and S. Qureshi, "Optimization and acceleration of convolutional neural networks: A survey," *J. King Saud Univ. - Comput. Inf. Sci.*, Oct. 2020, doi: 10.1016/j.jksuci.2020.10.004.
[34] X. Feng, Q. M. J. Wu, Y. Yang, and L. Cao, "A compensation-based optimization strategy for top dense layer training," *Neurocomputing*, vol. 453, pp. 563–578, Sep. 2021, doi: 10.1016/j.neucom.2020.07.127.
[35] V. H. Nhu *et al.*, "Effectiveness assessment of Keras based deep learning with different robust optimization algorithms for shallow landslide susceptibility mapping at tropical area," *Catena*, vol. 188, p. 104458, May 2020, doi: 10.1016/j.catena.2020.104458.
[36] Z. Chang, Y. Zhang, and W. Chen, "Electricity price prediction based on hybrid model of adam optimized LSTM neural network and wavelet transform," *Energy*, vol. 187, p. 115804, Nov. 2019, doi: 10.1016/j.energy.2019.07.134.

BIOGRAPHIES OF AUTHORS

Nazrul Effendy received the B.Eng. Degree in Instrumentation Technology of Nuclear Engineering and the M.Eng. degree in Electrical Engineering from Universitas Gadjah Mada in 1998 and 2001. He received a Ph.D. degree in Electrical Engineering from Chulalongkorn University, Thailand, in 2009. He was a research fellow at the Department of Control and Computer Engineering, Polytechnic University of Turin, Italy, in 2010 and 2011 and a visiting researcher in Shinoda Lab (Pattern Recognition & Its Applications to Real World), Tokyo Institute of Technology, Japan, in 2009. Currently, he is an Associate Professor and the coordinator of the Intelligent and Embedded System Research Group, the Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada. He is a member of the Indonesian Association of Pattern recognition, the Indonesian Society for Soft Computing, and the Indonesian Artificial Intelligence Society. He can be contacted at email: nazrul@ugm.ac.id.

Eko David Kurniawan is a final year bachelor student in Engineering Physics, Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering Universitas Gadjah Mada, Yogyakarta, Indonesia. He is a research assistant at the Intelligent and Embedded System Research Group, Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada. His research interests are machine learning and its application in engineering. He can be contacted at email: ekodavid98@mail.ugm.ac.id.

Kenny Dwiantoro received the B.Eng. Degree in Engineering Physics from Universitas Gadjah Mada in 2020. He is a research assistant at the Intelligent and Embedded System Research Group, Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada. He is interested in conducting research in machine learning and its application in engineering. He can be contacted at email: kennydwiantoro14@gmail.com.

Agus Arif received the B.S. degree in nuclear engineering from Universitas Gadjah Mada, Indonesia, in 1991 and the M.Eng. degree in Engineering Physics from Institut Teknologi Bandung, Indonesia, in 2000. He was a research assistant in the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS, Malaysia, from 2009 to 2011. He is currently an assistant professor and a member of the Intelligent and Embedded System Research Group, the Department of Nuclear Engineering and Engineering Physics, Faculty of Engineering, Universitas Gadjah Mada, Yogyakarta, Indonesia. His research interests are instrumentation, control, machine learning, and applications. He can be contacted at email: agusarif@ugm.ac.id.

Nidlom Muddin received the B.Sc. Degree in Organic Chemistry from Universitas Sebelas Maret in 2008. Currently, he is a Sr. Analyst of the New Refinery Project for PT. Kilang Pertamina International. He is interested in conducting research in an area related to energy management systems, especially boiler and furnace operation. He proposed a standard boiler and furnace assessment method and periodically conducted boiler and furnace optimization training. He can be contacted at email: nidlom.muddin@pertamina.com.