Taylor & Francis
Taylor & Francis Group

Journal of Industrial and Production Engineering

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tjci21

Extreme learning machine – radial basis function (ELM-RBF) networks for diagnosing faults in a steam turbine

Arian Dhini, Isti Surjandari, Benyamin Kusumoputro & Andrew Kusiak

To cite this article: Arian Dhini, Isti Surjandari, Benyamin Kusumoputro & Andrew Kusiak (2022) Extreme learning machine – radial basis function (ELM-RBF) networks for diagnosing faults in a steam turbine, Journal of Industrial and Production Engineering, 39:7, 572-580, DOI: 10.1080/21681015.2021.1887948

To link to this article: https://doi.org/10.1080/21681015.2021.1887948

Published online: 01 Mar 2021.

JOURNAL OF INDUSTRIAL AND PRODUCTION ENGINEERING
2022, VOL. 39, NO. 7, 572–580
https://doi.org/10.1080/21681015.2021.1887948

ARTICLE

Extreme learning machine – radial basis function (ELM-RBF) networks for diagnosing faults in a steam turbine

Arian Dhini [a], Isti Surjandari [a], Benyamin Kusumoputro [b] and Andrew Kusiak [c]

[a]Department of Industrial Engineering, Faculty of Engineering, Universitas Indonesia, Depok, Indonesia; [b]Department of Electrical Engineering, Faculty of Engineering, Universitas Indonesia, Depok, Indonesia; [c]Department of Industrial and Systems Engineering, The University of Iowa, Iowa City, IA, USA

ABSTRACT

A fast and reliable fault diagnosis system for a steam turbine in thermal power plant is crucial. The system will detect and classify a potential or occurring fault, hence suitable precautions steps will be correctly determined, and unplanned breakdown will be prevented. This study proposes a new application of extreme learning machine-radial basis function networks (ELM-RBF) for steam turbine fault diagnosis system. ELM-RBF recently has been known for its extremely fast computation. The proposed system was tested with real fault historical data from a steam power plant in Jakarta. To evaluate the system performance, a comparison with backpropagation neural networks (BPNN) was conducted. Four scenarios using ELM-RBF and BPNN, with and without ReliefF for feature selection were designed. The results show high accuracy in almost all the scenarios tested. The BPNN shows better accuracy than ELM-RBF, however, ELM-RBF performs considerably faster computation than BPNN without significant decrease in accuracy.

ARTICLE HISTORY
Received 25 June 2019
Accepted 2 February 2021

KEYWORDS
Data-driven method; fault diagnosis; neural networks; extreme learning machine; steam turbine

1. Introduction

The supply of electricity is of vital importance in modern life. To ensure the availability of electricity, the reliability of power plants must be maximized, which also implies higher safety and quality requirements. Maintenance is the key to the efficient operation of electrical power plants because it prevents unplanned breakdowns.

The urgency of effective and efficient maintenance is even higher for aging power plants because failures can cause serious safety concerns and can even be life-threatening. The International Energy Agency [1] reported that the majority of electricity suppliers worldwide still relies on thermal power plants (about 66%). Most of these plants are aging, even the plants in Japan [2].

In response to these issues, predictive maintenance has emerged as a more efficient maintenance strategy. This maintenance strategy is also known as condition-based maintenance, which is based on equipment monitoring. It recommends maintenance only when an urgency appears [3].

Figure 1 depicts the process of condition-based maintenance. The initial stage in condition-based maintenance is monitoring the condition of the equipment. Once these data are acquired, they are pre-processed to prepare for analysis in the following stage. If a fault is detected and its type is diagnosed, the remaining useful life can be predicted by using a prognostic method. Hence, the maintenance time and activity can be determined.

The first critical analysis of condition monitoring is fault diagnosis, which is initiated by fault detection. A fault is defined as an unacceptable deviation of at least one parameter of a normally operating system [4]. Fault diagnosis should be fast and accurate, while ignoring faults degrades the safety and security of the process, such as catastrophic failures and loss of material and even life [5].

As one of the main systems of a steam power plant, a steam turbine hardly needs a fast and accurate fault diagnosis system, to avoid unplanned breakdown during operation. Any disturbance during steam turbine operation may cause trip and interrupt the electricity generation.

Previous studies developed fault diagnosis system by using model-based approaches, such as the Kalman filter or the parity relation [6]. However, in a complex industrial system, it is more challenging to develop a comprehensive understanding of the process, which is prerequisite to develop a model-based method.

On the other hand, distributed control systems (DCSs) have been widely adopted in industrial processes, including in aging power plants. DCSs involve process monitoring and automated control systems that generate a data stream and provide huge amounts of data for process monitoring. This situation

CONTACT Arian Dhini arian@ie.ui.ac.id Department of Industrial Engineering, Faculty of Engineering, Universitas Indonesia, Depok 16424, Indonesia

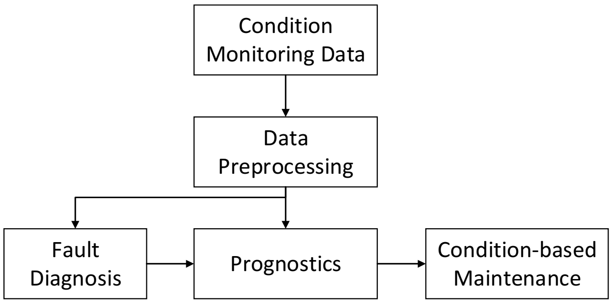

Figure 1. Condition-based maintenance scheme.

creates popularity for data-driven approaches, which use machine learning method. They are distribution free assumption methods, require no prior knowledge of the system, and offer satisfactory performance.

Neural networks (NN) have been known as data-driven self-adaptive model which can handle complex and nonlinear problem without prior knowledge. BPNN is the most popular NN-based technique due to its accuracy. Karlsson et al. [7], Changfeng et al. [8], and Chen et al. [9] used NN based methods to detect and diagnose faults, and compared with other methods, such as Bayesian networks (BN), support vector machine (SVM) and traditional multivariate method: linear discriminant analysis (LDA). Some studies used simulated data in their studies [7,10], while some recent studies used real data from power plants.

There is an urgency to lengthen the lifetime and preserve the reliability and availability of aging thermal power plant. Alongside, development in machine learning methods has created more advanced approaches. It creates an opportunity to investigate how the impact of advanced data-driven methods to fault diagnosis systems in the aging thermal power plant using recent machine learning based methods. Hence, this study proposes a new application of extreme learning machine-radial basis function networks (ELM-RBF) for steam turbine fault diagnosis system that uses real data for a steam turbine of a thermal power plant. The preliminary study on this topic has been published in [11].

A recent learning algorithm, Extreme learning machine (ELM), is proposed for RBF. ELM-RBF was first introduced by Huang and Siew [12], which was known to have extremely fast convergence and good generalization. It is due to random input weights selection and applied the generalized inversion method to get output weights [13]. Wang et al. [14] stated that ELM shown potential learning algorithm for single hidden-layer feedforward neural networks (SFLNNs).

ELM-RBF has attracted many researchers due to its capability, and there has been many studies to develop the derivation method [13,15]. However, the applications of ELM-RBF on power plant industries are still limited. Wong et al. [16], compared ELM and SVM for gas turbine fault diagnosis system using simulated data. Despite the fact that ELM only slightly outperformed SVM in accuracy, ELM proves significant faster time computation. Thus, the analysis done herein compares the performance of two NN-based approaches: ELM-RBF and BPNN.

The present study constitutes an empirical study of NN-based fault-diagnosis systems based on real data from a steam turbine of an aging thermal power plant in Jakarta, Indonesia. A snapshot of the data acquired from the DCS gives the historical faults and the normal operating conditions from 2014 to 2017. This study contributes to applying data-driven approach to create a real-time system, which detects and diagnosis faults accurately in the steam turbine.

This paper is organized as follows. Section 2 describes the steam turbine role in a power plant, steam turbine faults, and the significant role of fault-diagnosis systems in steam turbine maintenance. Section 3 explains in detail how the study was conducted, while Section 4 presents results and discussion. Section 5 concludes the paper with some limitations and future work.

2. Steam turbines

2.1. Steam turbine mechanism

As one of the main system in the power plant, steam turbine generator comprises of many subsystems: mechanical, electrical, hydraulic, heating, and related accessorial units. The mechanical unit alone contains hundreds of components, such as blades, rotor shaft and bearings, casings and seals, turbine pressure sections, and steam-flow control parts [8]. The 200 MW steam turbine used herein is a double-flow steam turbine that consists of two casings: one for both high and intermediate pressure, and the other for low pressure (LP). It operates as an impulse-reaction mechanism.

The work mechanism of a steam turbine can be explained as follows. First, superheated high-pressure steam passes through a small opening, i.e. nozzle, where the steam will attain a remarkably high steam velocity. After rotating turbine, the steam pressure is lowered and reheated in the boiler. The next cycle starts when the steam enters the intermediate- and low-pressure turbines.

2.2. Fault types in a steam turbine

Arjona-López et al. [17] state that one of the most critical periods during steam turbine operation is the start-up process. Turbine damages are often related with the transients such as turbine elements thermal fatigue, rotor brittle fracture, stress corrosion cracking, erosion of LP blading [18]. Figure 2 shows steam turbine fault types, which are classified based on turbine

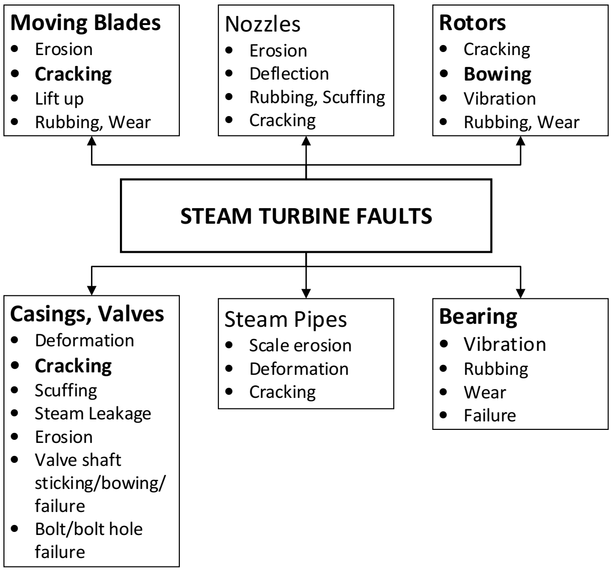

Figure 2. Degradation, damages, and failure modes of steam turbine components, adapted from [19].

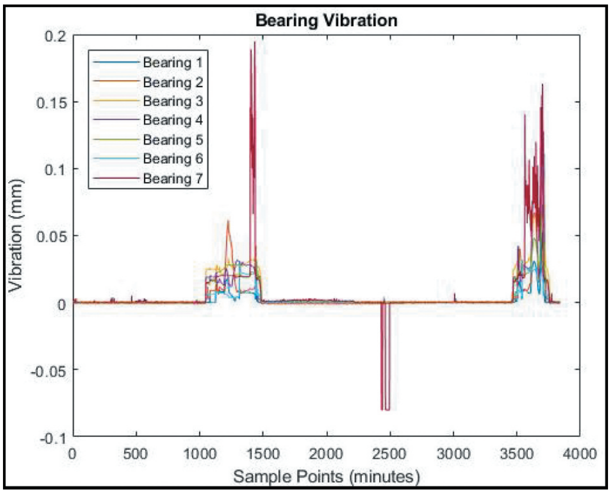

Figure 3. Vibration versus time indicating rotor bowing from July 14 to 15, 2014.

six main components, i.e. moving blades, nozzles, rotors, casing and valves, steam pipes, and bearings [19] .

The four faults (misalignment, rotor bowing, blade cracking, and casing cracking), were selected as outputs for this study. They were based on historical data and through in-depth discussions with experts.

Misalignment is a condition whereby the centerlines of coupled shafts do not coincide which was due to some reasons, such as burrs or dirt on shaft or housing shoulders. The high amplitude bearing vibration should not exceed 0.125 mm. Nonetheless, a vibration amplitude of around 0.04–0.05 mm for two successive bearings indicates misalignment, as well as the abnormal temperature increase of bearings. The inputs of misalignment data were obtained from event on 30 January 2017.

Rotor bowing is caused by the lack of space for rotor expansion. Rotor bowing creates high-amplitude vibration and rubbing on casings and bearings [19]. The vibration amplitude of all bearings increased from 0.04 to 0.3 mm, while the limit is 0.125 mm. Differential expansion and casing expansion can also be used as indicators of rotor bowing if they show abnormality. Data for this fault type were gathered from July 14 to 15, 2014, and the fluctuation is shown in Figure 3.

Blade cracking is the subsequent fault. Segura et al. [20] explain that out-of-range operating parameters are evidence of a crack in the final blades of the last stage; for example, excessive flow steam, low vacuum, or the impact of implosion particles in the nozzle area. The data for this event were collected on 19 December 2016, when an abnormality was appeared in the absolute condenser pressure, which was in the low-vacuum range (below atmospheric pressure). The tolerance of wet steam is only around 10% above the limit [18]. The main steam flow also reveals the presence of excessive wet steam.

The last type of fault is casing cracking. Chowdhury et al. [21] summarized that most of cracking is caused by thermal fatigue (65%), followed by brittle fracture (30%), and creep (5%). Long-term exposure to a large temperature difference between the upper and lower casings of a steam turbine leads to cracks in the casing. The data for this event were collected on 23 November 2015. The acceptable temperature range for the HP-IP upper and lower casings is −42 to 42 °C.

3. Proposed data-driven research methods and application

This study proposes and compares the performance of two neural-network-based approaches to detect and diagnose steam turbine faults. As described in Zhang et al. [22], NN has many advantages: (1) NN is data-driven self-adaptive model that require no prior assumptions or knowledge about the problem. (2) NN generally lends themselves to generalization. (3) NN has more general and flexible functional forms than any traditional statistical model. Finally, (4) NN can handle complex and nonlinear problems.

Previous studies in numerous applications of NN for fault detection in industries, including power plants, showed satisfactory results [7,16,23–25].

3.1. Research frameworks

Figure 4 depicts the proposed research framework to develop a data-driven fault-diagnosis system for the steam turbine. The figure shows the stages for developing the fault-diagnosis system.

There are 140 process parameters of the steam turbine, which covers the following categories:

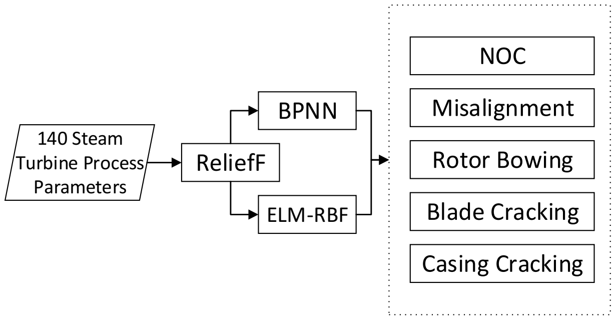

Figure 4. Framework of data-driven fault-diagnosis system.

temperature, pressure, flow, and vibration. They were used as inputs for the data-driven fault diagnosis model.

This research uses "snapshots" of both normal operating conditions (NOC) and fault events data. A comparison with and without feature selection using the ReliefF algorithm is also evaluated in this study. Redundancies may also exist among the features, and not all the features give high contribution to the certain condition: NOC or fault.

The next step is the main part of this study. Two NN-based approaches are applied to the data, and the results of each approach are compared with each other. Huang and Siew [12] found that the use of an ELM algorithm in RBF networks significantly improves the computation efficiency. The BPNN method, as the most popular approach of NN, is used as a benchmark. Details of both methods are elaborated in the following sections.

3.2. Backpropagation neural networks

The NN contains interconnected neurons, which denote knowledge by their assigned weights, as described in Figure 5. One typical NN, namely, multilayer perceptron, consists of an input layer, a hidden layer, and an output layer. The number of inputs (x_1, x_2, …, x_n) denotes the number of features in the process, which each is connected to each hidden neuron in the hidden layer. Finally, each hidden neuron is connected to each output neuron (y_1, y_2, …, y_m). The strength of each connection is called its "weight" (i.e. v_{ij} for the connection between each input neuron and each hidden neuron, and w_{jt} for the connection between each hidden neuron and each output neuron). This study applies sigmoid activation function.

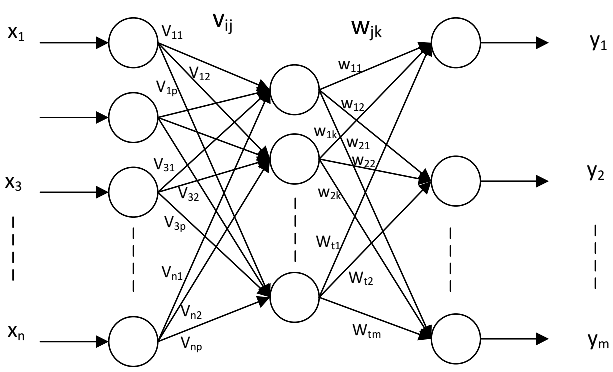

Figure 5. Architecture of neural network.

BPNN employs iterative gradient algorithm for minimizing the mean square error between the developed model and the real outputs by updating weights. When the output of the error function is satisfactorily small, then, the iteration can be terminated.

3.3. Extreme learning machine – radial basis function networks (ELM-RBF)

ELM-RBF is a modified RBF network. RBF networks are artificial NN which employs radial basis functions as activation functions. The structure of an RBF network consists of three-layered feedforward NN: the first layer is linear and distributes the input signal only, the second layer is nonlinear and uses Gaussian functions, and the third layer linearly combines the Gaussian outputs.

Traditionally, in the training process, the weights between the hidden layer and the output layer are adjusted. In this case, the RBF requires the following five parameters to be optimized: (1) weights between the hidden layer and the output layer, (2) the activation function, (3) the activation function centers, (4) the distribution of the activation function centers, and (5) the number of hidden neurons.

Huang, et al. [26] reported some drawbacks of the gradient descent algorithm which is used in traditional RBF: (1) If the learning rate is too small, the learning converges very slowly. In contrast, the learning diverges and is unstable. (2) The probability of a local minima is nonzero. (3) The NN may be over-trained, which hinders generalization. (4) The learning is time-consuming.

Thus, an ELM is introduced to overcome these weaknesses. Unlike the gradient descent-based algorithm, in a single-hidden-layer feedforward NN, the ELM arbitrarily picks the input weight and, therefore, the hidden biases of neurons and analytically define the output weights [12]. Furthermore, Huang, et al. [26] claimed that the use of an ELM results in better generalization and in an extremely high learning speed.

This study uses the ELM-RBF, which randomly generates kernel centers and the impact widths of RBF kernels and analytically calculates the output weights. As in Huang and Siew [12] (x_i,t_i), where $x_i = [x_{i1}, x_{i2}, \ldots, x_{in}]^T \in R^n$ and $t_i = [t_{i1}, t_{i2}, \ldots, t_{im}]^T \in R^m$, RBFs with $\tilde{N}$ kernels can be mathematically modeled as

$$\sum_{i=1}^{\tilde{N}} \beta_i \emptyset_i(x_j) = o_j, \quad j = 1, \ldots, N \tag{1}$$

where $\beta_i = [\beta_{i1}, \beta_{i2}, \ldots, \beta_{im}]^T$ is the weight vector linking kernel *i* and the output neurons and $\emptyset_i(x)$ is the output of kernel i. $\mu_i = [\mu_{i1}, \mu_{i2}, \ldots, \mu_{in}]^T$ is the center of kernel i and σ_i is the impact width. $\emptyset$ is the radially symmetric kernel function and it is assumed that $\emptyset_i(x)$ is a nonlinear bounded integral and is everywhere almost continuous:

$$\sum_{i=1}^{\tilde{N}} \beta_i \emptyset_i(\mu_i, \sigma_i, x_j) = t_j, j = 1, \ldots, N. \tag{2}$$

In short,

$$H\beta = Y \tag{3}$$

where H is the hidden-layer output matrix of the RBF network. Column i of H is the output of kernel i with respect to inputs $x_1, x_2, \ldots, x_N$. Because the number of kernels is unequal to the number of training samples, in this case $\tilde{N} << N$, H is a non-square matrix and $\beta_i = (i = 1, 2, \ldots, \tilde{N})$ may not exist, such that $H\beta = T$. Therefore, the unique smallest-norm least-squares solution $\hat{\beta}$ of the previous linear system is

$$\hat{\beta} = H^+ T, \tag{4}$$

where H^+ symbolizes the Moore–Penrose generalized inverse.

3.4. Application domain

The research framework as on Figure 4 were applied to data from an aging steam power plant in Jakarta. The power plant has been operating since 1980s and supplies around 26% of the highest electricity demand for the Greater Jakarta.

The data provide snapshots of four fault-type indicators: misalignment, rotor bowing, blade cracking, and casing cracking, which were observed through DCS. In addition to these fault types, the proposed fault diagnosis system also learns the normal operation condition (NOC) data of the steam turbine. It is assumed that only single fault may arise at any time.

3.5. Data preprocessing

The data preprocessing was conducted in the following stages: Firstly, data are transformed, from 1 s, into 1 min intervals, by averaging over groups of 60 sequential data. The next step in data preprocessing was data cleaning, whereby only relevant data are selected. Data cleaning resulted in 4320 one-minute-interval NOC data selected for classification learning (September 1 to 4, 2015). The process identified only 1296 one-minute-interval data (January 29 to 30, 2017) for the misalignment fault, 529 one-minute-interval data (July 13 to 15, 2014) for rotor bowing, 133 blade cracks from 19 December 2016, and 61 casing cracking from 23 November 2015.

To increase the accuracy and nullify any bias, the 10-fold cross-validation was used for model validation. According to Kohavi [27], the best k-fold cross-validation for model selection is 10-fold cross-validation because of the small bias reported from experiments. This process was applied before the data underwent further preprocessing to prevent leakage to the data testing.

Subsequently, because of the unbalanced data for each class, the NOC data were under-sampled, reducing the data count from 4320 to 4000. In addition, bootstrap, an over-sampling method, was applied to the rest of the fault-type data. Thus, there were 4000 data in each class.

The next step is data scaling, with the following reasons: (1) to avoid the domination of the larger-scale data over the smaller-scale data, and (2) to obtain a more rapid convergence of the learning process. This study applies the min-max formula to obtain scaled data in the range [−1,1] before classification by the BPNN approach. The formula is as on (5).

$$x_{inew} = \left[\left(\frac{x_i - \min(x)}{\max(x) - \min(x)} \right) * 2 \right] - 1 \tag{5}$$

ELM-RBF applied normalization for data scaling. It subtracts the mean from the value of each feature and divides it by its standard deviation.

3.6. Feature selection

Chandrashekar and Sahin [28] stated that feature selection helps to understand the data, reduces computation time, minimizes the curse of dimensionality, and improves the performance of data processing.

$$\begin{aligned} W[A] = P(&\textit{different value of } A | \textit{nearest} \\ &\textit{instance from different class}) \\ - P(&\textit{different value of } A | \textit{nearest} \\ &\textit{in stancefrom different class}). \end{aligned} \tag{6}$$

$$\begin{aligned} W[A] &: \\ &= w[A] - \frac{diff(A, R, H)}{m} \\ &\quad + \sum_{C \neq class(R)} [P(C) x diff(A, R, M(C))]/m \end{aligned} \tag{7}$$

Kira and Rendell [29] proposed Relief-based algorithm, one of the more powerful feature-selection methods. Relief is the only individual evaluation filter algorithm that can detect all feature dependencies [30]. An instance examines for its two nearest neighbors; namely, one from the same class (nearest hit) and another from a different class (nearest miss). Estimating the attribute weight of A, W [A], is done based on the difference between the following probabilities, as Equation (6).

Furthermore, Kononenko [31] explored this approach to handle multi-class problems with noisy and incomplete data. This study uses the extended version of Relief, called ReliefF. The ReliefF algorithm searches one near miss M(C) for each other class and calculates

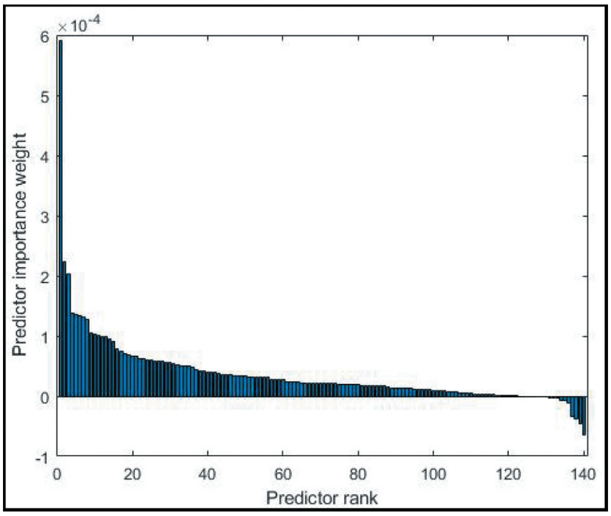

Figure 6. Distribution of predictor-weight importance (k = 15).

Table 1. Weights for ReliefF results.

Feature	k = 3	k = 10	k = 15
Condensate Pump 5B motor current	0.000755	0.000767	0.000590
SH Spray Flow	0.000341	0.000302	0.000223
No8 Circulating water pump 8 motor current	0.000329	0.000305	0.000222
Turbine HP-IP Horiz joint flange/bolt temp	0.000349	0.000221	0.000134
Condensate Flow	0.000183	0.000194	0.000103
CP-A/HC	0.000178	0.000183	0.000128
Deaerator 5D level (small)	0.000135	0.000182	0.000138
No 6 bearing vibration	0.000201	0.000175	0.000132
Condensate pump 5A motor current	0.000172	0.000168	0.000104
Aux cool water heat exc 5(7) cw inlet temp	0.000230	0.000158	0.000098
CP 5A coupling lube oil temp	0.000204	0.000149	0.000101
Feed water flow	0.000178	0.000141	0.000137
MT gov demand	0.000176	0.000131	0.000098

the average of the contribution for updating estimates W[A]. The next step is calculating weighted average with the prior probability of each class, as in Equation (7).

The study uses the ReliefF algorithm with k = 3, 10, and 15, where k is the number of nearest neighbors. Figure 6 shows the distribution of predictor weights for k = 15. A predictor with greater weight is more important than a predictor with less weight. The 13 features are selected based on consistency of weights over three values of k. These features are detailed in Table 1, with the associated weights of each k.

4. Results and discussions

This section presents the results of the BPNN and ELM-RBF algorithm for fault classification, with and without feature selection with ReliefF. All programs in this study were run in MATLAB version R2017b.

4.1. BPNN for fault diagnosis system

Tables 2 and 3 show the classification of each fold when using the BPNN approach. There are 140 input neurons and 75 hidden neurons. There are five output neurons, which is the same as the number of classes. Based on previous trial-and-error results, the learning rate and the momentum are set into 0.1. The result of each fold was averaged over 10 iterations. While the training and test recognition rate (RR) is mostly the same, the computation time for each iteration varies.

Table 2. RR (%) of BPNN.

Fold	Training RR (%)	Test RR (%)	Time (s)
1	100	99.03	155.24
2	100	100	157.68
3	100	100	159.38
4	100	98.7	156.15
5	100	100	155.40
6	100	100	162.23
7	100	100	158.15
8	100	100	155.76
9	100	100	158.69
10	100	100	192.55
Average	100	99.77	161.12

Table 3. RR (%) of BPNN with ReliefF.

Fold	Training RR (%)	Test RR (%)	Time (s)
1	100	100	810.19
2	100	100	800.76
3	100	100	801.09
4	100	100	750.73
5	100	100	730.46
6	100	100	771.23
7	100	100	800.50
8	100	100	784.53
9	100	100	774.46
10	100	100	785.76
Average	100	100	780.97

The results in Table 2 are highly accurate for both the training and testing data. The 100% training recognition reveals that no errors occurred during the training, and the 99.77% test recognition shows that the model performs very well with the testing data. The average computation time is 161.12 s, which is attributed to several sources: First, the data for training and testing may not significantly differ for each class. Second, BPNN is popular in NN applications because of the backpropagation steps in which the minimum error is found by revising the weights. However, this procedure seems to require significant computation time to achieve high accuracy.

Table 3 shows the results of fault diagnosis system using BPNN when applying the ReliefF algorithm prior to the BPNN algorithm for selecting features. The accuracy increases to 100% for all folds in testing data. Unfortunately, computation time increased almost five-fold with respect to the previous procedure. This increase may be due to the incomplete information that uses only 13 selected attributes to achieve high accuracy. More attributes should be included in the feature selection.

4.2. Results of ELM-RBF for fault diagnosis system

The RBF networks used in this study were modified by using the ELM for learning rather than the gradient

Table 4. RR (%) of ELM-RBF.

Fold	Training RR (%)	Test RR (%)	Time (s)
1	97.62	68.80	54.46
2	97.84	99.40	52.78
3	96.73	100	51.27
4	96.55	98.70	53.83
5	96.64	98.85	55.05
6	96.53	100	53.96
7	96.69	100	58.73
8	89.73	100	57.47
9	96.30	100	57.83
10	95.13	100	52.70
Average	95.98	96.58	54.81

Table 5. RR (%) of ELM-RBF with ReliefF.

Fold	Training RR (%)	Test RR (%)	Time (s)
1	98.34	74.30	9.05
2	96.94	94.50	9.01
3	98.68	96.50	9.95
4	96.38	100	10.03
5	95.75	100	9.15
6	94.32	100	9.80
7	96.39	100	9.82
8	96.17	97.7	8.84
9	89.58	81.2	9.04
10	94.22	97.4	10.14
Average	95.68	94.16	9.48

descent algorithm. It was expected that the results would better lend themselves to generalization and that the computation time would decrease, all without sacrificing accuracy.

Table 4 shows the results of the first scenario using ELM-RBF (without feature selection). Overall, only a slight decrease appears in the test-recognition rate (3.2%) compared with BPNN. Moreover, ELM-RBF offers much faster computation time, i.e. 2.74 times faster than the BPNN. A small decrease in testing accuracy occurs for two folds only, except for fold 1, which has the lowest test-recognition rate of 68.8%. Figure 7 plots the errors in the testing data, which were mostly fault types 2 and 3. The testing data of fold 1 for fault types 2 and 3 seem to differ significantly from the training data.

Table 5 summarizes the results of ELM-RBF when using 13 attributes selected by using the ReliefF algorithm. The results show no significant difference exists in terms of testing accuracy compared with the results given in Table 4. The average testing accuracy decreases by about 2.51%, and the computation time decreases. The computation time is 5.78 times faster than without feature selection.

This study shows that BPNN consistently has higher test-recognition rates both with and without feature selection. However, the high recognition rates are counter-balanced by longer computation times.

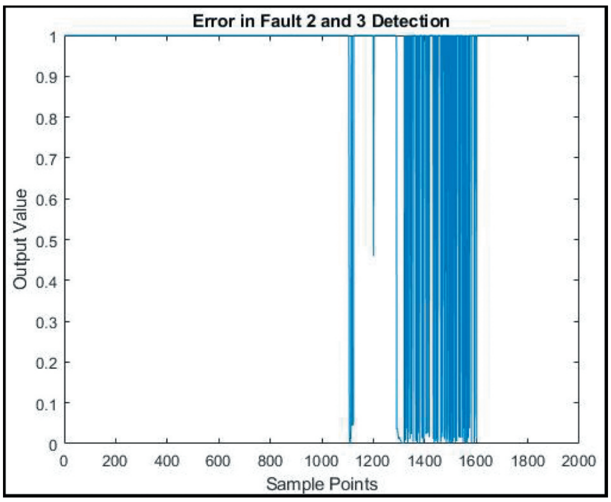

Figure 7. Error in diagnosis of faults 2 and 3.

Therefore, ELM-RBF seems more promising for obtaining high recognition rates and high generalization as stated in [14].

Application of these data-driven approaches for real-time fault diagnosis system significantly improves the fault diagnosis process in terms of accuracy and time. It replaces the manual interpretation of operators. Hence, it supports the reliability of CBM implementation.

5. Conclusion

The study uses two NN-based methods, BPNN and ELM-RBF. BPNN has been popular due to its accuracy, while ELM-RBF is known for its facile generalization and fast computation. The results prove that, although BPNN outperforms ELM-RBF in terms of accuracy, the former comes with long computation time for decision-making, which usually translates into high cost. The computation time for ELM-RBF is extremely better than that for BPNN.

This study also identifies attributes that contribute significantly to the relevant class. More analysis should be conducted to identify which attributes indicate certain fault.

The results of the study are consistent with the theory and results of previous studies. The modified learning method that uses ELM significantly improves the results when using NN learning algorithms for fault diagnosis. ELM-RBF is promising data-driven method to be applied for the next process of CBM, i.e. prognosis.

Time lag should be considered for the next research. The trend of certain attributes may create different fault types. It may create more complex model, but advanced machine learning methods or ensemble technique should be able to overcome the problems. More data sources should be added from various steam turbines. They will enrich the analysis and create more intelligent system.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research work was supported by the Doctoral Dissertation Research Grant from the Ministry of Education and Culture of the Republic of Indonesia [120/SP2H/PTNBH/DRPM/2018].

Notes on contributors

Arian Dhini received her Bachelor of Engineering degree and Master of Engineering degree in Industrial Engineering from Institut Teknologi Bandung, in 1997, and from Universitas Indonesia, in 2007, respectively. She finished her doctoral program from Department of Industrial Engineering, Universitas Indonesia in 2020. Her research fields include statistics, data mining, quality and reliability, and condition monitoring. Arian Dhini has been a member of ASQ on reliability section since 2015 and IEEE since 2019.

Isti Surjandari is a Professor and Head of Statistics and Quality Engineering Laboratory in the Department of Industrial Engineering, Faculty of Engineering, Universitas Indonesia. She holds a bachelor's degree in Industrial Engineering from Universitas Indonesia and a Ph.D. degree from the Ohio State University. Her areas of interest are industrial management, quality engineering and applied statistics. She is a senior member of American Society for Quality (ASQ) and has a vast experience in industrial, manufacturing and service systems.

Benyamin Kusumoputro is a Professor in Computational Intelligence and Intelligent Systems since 2004. He holds a bachelor's degree in Physics from Bandung Institute of Technology, Indonesia, and Magister of Engineering from Universitas Indonesia in 1981 and 1984, respectively. Professor Kusumoputro received his Dr. Eng. degree from Department of Electrical and Electronic Engineering, Tokyo Institute of Technology in 1993. His research interests include the development of 3D face recognition using Hemispherical Structure of Hidden Layer Neural Networks and Deep Learning, odor recognition system, and the development of autonomous control system of unmanned vehicle systems. He published more than 100 articles in academic and professional journals and serves as a frequent Invited Speaker at various academic conferences and professional meetings.

Andrew Kusiak is a Professor in the Department of Industrial and Systems Engineering at The University of Iowa, Iowa City and Director of the Intelligent Systems Laboratory. His current research interests include applications of computational intelligence and big data in automation, manufacturing, product development, renewable energy, sustainability, and healthcare. He has published numerous books and hundreds of technical papers in journals sponsored by professional societies, such as the Association for the Advancement of Artificial Intelligence, the American Society of Mechanical Engineers, Institute of Industrial Engineers, Institute of Electrical and Electronics Engineers, Nature, and other societies. He speaks frequently at international meetings, conducts professional seminars, and consults for industrial corporations. Professor Kusiak has served in elected professional society positions as well as various editorial boards of over 50 journals, including five different IEEE Transactions. He is a Fellow of the Institute of Industrial Engineers and the Editor-in-Chief of Journal of Intelligent Manufacturing.

ORCID

Arian Dhini http://orcid.org/0000-0003-2409-6644
Isti Surjandari http://orcid.org/0000-0002-2084-3254
Benyamin Kusumoputro http://orcid.org/0000-0003-3714-0106
Andrew Kusiak http://orcid.org/0000-0003-4393-1385

References

[1] International Energy Agency. Key world energy statistics. Paris: IEA; 2017.
[2] Yoshioka Y. Current status of Japanese thermal power plants and life assessments of high temperature steam and gas turbine components. Materials at High Temperatures. 2017;34:386–396.
[3] Jardine AK, Lin D, Banjevic D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing. 2006;20:1483–1510.
[4] Isermann R, Balle P. Trends in the application of model-based fault detection and diagnosis of technical processes. Control Engineering Practice. 1997;5:709–719.
[5] Gao Z, Cecati C, Ding SX. A survey of fault diagnosis and fault-tolerant techniques—part I: fault diagnosis with model-based and signal-based approaches. IEEE Transactions on Industrial Electronics. 2015;62:3757–3767.
[6] Venkatasubramanian V, Rengaswamy R, Yin K, et al. A review of process fault detection and diagnosis: part I: quantitative model-based methods. Computers & Chemical Engineering. 2003;27:293–311.
[7] Karlsson C, Arriagada J, Genrup M. Detection and interactive isolation of faults in steam turbines to support maintenance decisions. Simulation Modelling Practice and Theory. 2008;16:1689–1703.
[8] Changfeng Y, Zhang H, Lixiao W. A novel real-time fault diagnostic system for steam turbine generator set by using strata hierarchical artificial neural network. Energy and Power Engineering, 2009;1:7–16.
[9] Chen K-Y, Chen L-S, Chen M-C, et al. Using SVM based method for equipment fault detection in a thermal power plant. Computers in Industry, 2011;62:42–50.
[10] Salahshoor K, Kordestani M, Khoshro MS. Fault detection and diagnosis of an industrial steam turbine using fusion of SVM (support vector machine) and ANFIS (adaptive neuro-fuzzy inference system) classifiers. Energy. 2010;35:5472–5482.
[11] Dhini A, Kusumoputro B, Surjandari I. Neural network based system for detecting and diagnosing faults in steam turbine of thermal power plant. Proceeding of the IEEE 8th International Conference on Awareness Science and Technology (iCAST), 2017 Nov 8–10; Taichung, Taiwan; p. 149–154.
[12] Huang G-B, Siew C-K. Extreme learning machine: RBF network case. Proceedings of The Eighth International Conference on Control, Automation, Robotics and Vision (ICARCV); 2004 Dec 6–9; Kunming, China; p. 1029–1036.
[13] Zhang N, Ding S, Zhang J. Multi layer ELM-RBF for multi-label learning. Applied Soft Computing. 2016;43:535–545.

[14] Wang Y, Cao F, Yuan Y. A study on effectiveness of extreme learning machine. Neurocomputing. 2011;74:2483–2490.
[15] Li Q, Xiong Q, Ji S, et al. A method for mixed data classification base on RBF-ELM network. Neurocomputing. 2021;431:7–22.
[16] Wong PK, Yang Z, Vong CM, et al. Real-time fault diagnosis for gas turbine generator systems using extreme learning machine. Neurocomputing. 2014;128:249–257.
[17] Arjona-López MA, Hernández Flores C, Gleason García E. An intelligent tutoring system for turbine startup training of electrical power plant operators. Expert Systems with Applications, 2003;24:95–101.
[18] Leĭzerovich AS. Steam turbines for modern fossil-fuel power plants. The Fairmont Press, Inc; 2008. Lilburn, GA
[19] Fujiyama K, Nagai S, Akikuni Y, et al. Risk-based inspection and maintenance systems for steam turbines. International Journal of Pressure Vessels and Piping, 2004;81:825–835.
[20] Segura JA, Castro L, Rosales I, et al. Diagnostic and failure analysis in blades of a 300MW steam turbine. Engineering Failure Analysis. 2017;82:631–641.
[21] Chowdhury SG, Mukhopadhyay NK, Das G, et al. Failure analysis of a weld repaired steam turbine casing. Engineering Failure Analysis, 1998;5:205–218.
[22] Zhang G, Eddy Patuwo B, Hu MY. Forecasting with artificial neural networks: the state of the art. International Journal of Forecasting. 1998;14:35–62.
[23] Guglielmi G, Parisini T, Rossi G. Keynote paper: fault diagnosis and neural networks: A power plant application. Control Engineering Practice. 1995;3:601–620.
[24] Kusiak A, Song Z. Sensor fault detection in power plants. J Energy Eng. 2009;135:127–137.
[25] Rasaienia A, Moshiri B, Moezzi M. Feature-based fault detection of industrial gas turbines using neural networks. Turkish Journal of Electrical Engineering & Computer Sciences. 2013;21:1340–1350.
[26] Huang G-B, Zhu Q-Y, Siew C-K. Extreme learning machine: theory and applications. Neurocomputing. 2006;70:489–501.
[27] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. Proceedings of the 14th international joint conference on Artificial intelligence II; 1995 Aug 20–25; Montreal, Quebec, Canada.
[28] Chandrashekar G, Sahin F. A survey on feature selection methods. Computers & Electrical Engineering, 2014;40:16–28.
[29] Kira K, Rendell LA, Practical A. Approach to feature selection. Proceedings of the Ninth International Machine Learning Conference; 1992 Jul 1–3; Aberdeen, Scotland; p. 249–256.
[30] Urbanowicz RJ, Meeker M, La Cava W, et al. Relief-based feature selection: introduction and review. Journal of Biomedical Informatics. 2018;85:189–203.
[31] Kononenko I. Estimating attributes: analysis and extensions of RELIEF. Proceeding of European Conference on Machine Learning; 1994 Apr 6–8; Catania, Italy; p. 171–182.