

Predicting anomaly conditions of energy equipment using neural networks

Anastasia Sverdlova^{1,*}, and Artur Zaporozhets^{1,2}

¹Institute of Engineering Thermophysics of NAS of Ukraine, 2a Mariyi Kapnist Str., Kyiv, 03057, Ukraine

²State Institution “The Institute of Environmental Geochemistry of National Academy of Sciences of Ukraine”, 34a Palladin Ave., Kyiv, 03142, Ukraine

Abstract. In modern conditions for complex thermal power facilities, the issue of developing methods for predicting equipment failures is especially relevant. Methods based on the intellectualization of diagnostic systems and allowing to obtain predictive models based on the use of both current data received in real time from measuring equipment and retrospective information are considered promising. Intellectualization of the system in terms of the ability to learn allows to quickly adjust the parameters of forecasting models under changing conditions of equipment operation, to determine new deadlines for scheduled repairs and minimize equipment downtime. A limitation of the use of methods is the incompleteness of failure statistics, ie when equipment failures are rare or non-existent. Such diagnostics of energy equipment, especially thermal power facilities, contributes to a more environmentally friendly production.

1 Introduction

Modern technical facilities, including thermal power facilities, are complex systems built of different components – mechanical, electrical, electronic, etc., combined with the tasks of control of a particular process. Distinctive features of functioning of objects of heat power engineering are work in the conditions of irregular dynamic loadings applied practically to all elements of its design [1-3]. These loads can occur both due to internal factors and as a result of the influence of other equipment operating nearby [4].

At present, a large number of complex thermal power plants (nuclear and chemical reactors, turbines, powerful boilers, various furnaces for smelting metal, etc.) are equipped with automated control systems for technological parameters [5-7]. The controlled parameters, in addition to information about the flow of technological processes, also contains information about the current state of the equipment, the appearance and development of various faults [8-10]. Analysis of changes in these parameters, performed after accidents, unscheduled equipment stops, usually shows that there were signs of malfunctions that caused the accident or stop, long before the incident [11-13]. Thus, there is a fundamental possibility to identify abnormal conditions (faults) before they caused the accident and, thus, to prevent the accident or significantly reduce its consequences. Such cases are known to any operator who controls the technological processes of complex thermal power facilities (CTPF), but they are rare, exceptional due to limited human capabilities in the rapid processing of large amounts of information and lack of prior analysis of

processes in case of failure. Automated systems for collecting and processing information significantly expand both the operator's ability to identify faults and allow to automate the process of their detection, ie diagnosis [14,15]. For this purpose it is necessary to have a technique and corresponding diagnostic algorithms which use the information from sensors of technological control [16]. Such techniques and algorithms can significantly increase the effectiveness of CTPF diagnosis.

One of the important tasks of diagnostics of power equipment is to minimize the impact of power equipment on the ecological state of the environment. In this case, it is necessary to solve the following tasks:

- research and minimization of the negative impact on the environment at all stages of production, transmission and consumption of different types of energy [17-22];
- implementation of activities for the conduct of environmental studies, including the assessment of the impact on the environment [23-25];
- utilization / neutralization of energy equipment waste after completion of operation [26-30];
- participation in programs for the development of renewable energy sources, taking into account the application of the world practice of using market mechanisms for their support on the basis of “green” tariffs [31-34];
- formation of environmental requirements for the development of modern power equipment [35];
- analysis of developments and implementation of techniques, standards, criteria, requirements in the

* Corresponding author: science.sverdlova@gmail.com

field of environmental protection, environmental safety and rational use of natural resources [36-41];

- development of educational programs in the field of environmental safety in the production, transmission and consumption of electricity etc [42-46].

Thus, the diagnostics of power equipment should become one of the mandatory steps in improving the environmental friendliness of enterprises in the fuel and energy complex.

2 Anomaly detection based on neural networks

The method of detecting anomalies based on neural networks includes two stages: 1) the neural network learns to recognize classes of normal behavior in a training sample; 2) each instance of the instance comes as an input signal to the neural network. A neural network-based system can recognize one or more classes of normal behavior.

Replicative neural networks are used to find anomalies by recognizing only one class [47]. Deep Learning neural network technology has also been used successfully to solve this problem [48].

Various methods can be used to search for anomalies [49-51], the advantages and disadvantages are presented in the Table 1.

Machine Learning is a class of methods of artificial intelligence, which is characterized by indirect problem solving, and learning in the process of applying solutions to many similar problems [52]. In the process of constructing such methods can be used mathematical statistics, numerical methods, optimization methods, data mining, probability theory, graph theory.

Using machine learning possible the study of patterns in the data, which are then used to detect abnormal behavior. Machine learning tasks are usually divided into the following categories depending on the availability of a training “signal” or “feedback” available to the training system.:

- learning with a teacher using “reaction-stimulus” examples;
 - partial training with a teacher;
 - active learning;
 - reinforced training;
- learning without a teacher, suitable for tasks in which objects are described in detail and it is necessary to establish internal relationships between objects.

Next, deep neural networks will be considered as methods used to detect anomalies.

Table 1. Advantages and disadvantages of anomaly detection methods.

Method	Advantages	Disadvantages
Statistical analysis [52-55]	Lack of requirement for a priori information on signs of anomalies, which makes it possible to detect zero-day vulnerabilities against which protective mechanisms have not yet been developed	The difficulty of determining the threshold for optimal detection of anomalies, the impossibility of identifying anomalies as a result of malicious actions similar to normal actions or work and the requirement of statistical distributions in the absence of all elements of the process
Machine learning [56]	Improved system based on prior knowledge	High computational costs, as well as the complexity of adaptation to the subject area
Artificial neural networks [57]	Resistance to inaccurate input, as well as independence from the availability of information about the dependencies of the examples of input data	Complex and long training of neural networks, as well as demanding on the size of the training sample
Genetic algorithms [58]	Works well in solving large-scale optimization problems. Uses two decision mechanisms: deterministic and probabilistic. Uses multiple search space points	The complexity of the selection rules for selecting the best solutions
Hybrid methods [59-62]	The most flexible, as it allows to reduce the impact from the disadvantages of one method taking advantage of another	

The advantage of deep neural networks is the automatic selection of these important features. An inverse error propagation algorithm based on the gradient descent method is used to train neural networks [63]. In a deep neural network with several hidden layers, an error is calculated that is transmitted from one layer to another. In the first stage, the value of the error at the output of the neural network is calculated, for which the correct values are known. Then the error at the input to the output layer of the network is calculated, which will be used as an error at the output of the hidden layer. Therefore, the calculation continues until the error on the input layer is known. However, this algorithm is often not effective when the training sample is large, because it takes a long time to process all its elements. In practice, the method of

stochastic gradient descent or its modification is most often used to train neural networks [64,65].

However, the error back propagation algorithm is not sufficient for effective deep learning due to the problem of the vanishing gradient [66]. This problem is solved by the architecture of the neural network with long short-term memory (LSTM) [67,68]. Such networks contain special types of nodes that allow to remember values for a long time. The LSTM network unit contains a special neuron that is used as a memory element (Fig.1).

The output of a neuron is connected to its own input with a single weight. As a result, the value in the neuron is rewritten at each stage and thus stored. Neuron control is performed by three valves: input, output and forgetting valve. When the inlet valve is open, the value at the inlet

is written to the memory cell. If the inlet valve is closed, the inlet signals do not affect the contents of the cell. An open outlet valve allows to read values from a cell. When the value is no longer needed, it can be erased with the forget-me-not valve. Valves are connected to other nodes of a neural network which in the course of training define when it is necessary to open or close this or that valve.

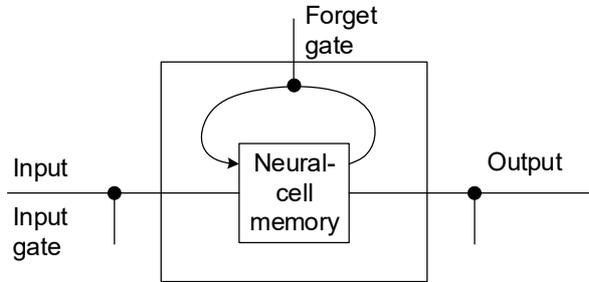


Fig. 1. LSTM network block diagram [67].

Through these networks cells LSTM can determine the importance of the events that occurred thousands of discrete time steps back and remember these events. Recurring networks previously used could remember the event for no longer than ten time steps [69].

The choice of neural network architecture is based on several factors. First, the sensors generate highly correlated multidimensional time series. In addition, these time series are aperiodic and synchronous (aligned in time), with fast (long-term) and slow (short-term) subprocesses. Under these conditions, conventional neural networks of direct propagation usually show poor results. An accurate data-driven predictive model can be developed using a neural network with LSTM cells [70].

Since not all time series can be predicted [71], it is possible to use an additional method of finding anomalies – autoencoder, which allows to apply training without a teacher during using the method of reverse error propagation. A synchronous architecture was chosen for the autoencoder, among the advantages of which is the possibility of using the streaming mode of data processing and a relatively smaller number of neural network parameters compared to other architectures.

3 Neural network architecture for anomaly prediction

3.1 The process of detecting anomalies

The process of detecting anomalies consists of the following stages:

1. Calculate the estimate of the anomaly of some observations or subsequences of a given time series using the detection method.
2. The obtained estimates are used to calculate estimates of anomalies of test time series. This application is performed in different ways, for example: (1) the average value of all anomaly indicators, (2) the average value of the upper anomaly estimate, (3) the average value of the logarithm of the anomaly estimates, (4) the anomaly estimates exceed the threshold, etc.

Test time series with the assessment anomaly that exceeds a threshold, denoted as abnormal.

The block diagram of failure forecasting is developed (Fig. 2) and the algorithm of detection of anomalies which consists of two parts is offered: forecasting and detection.

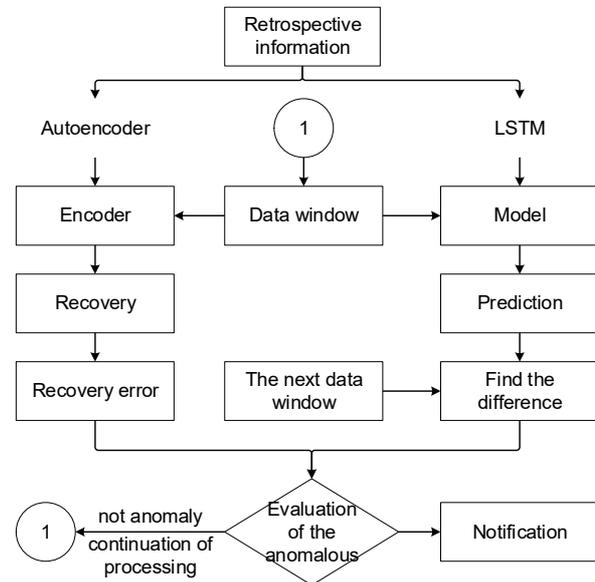


Fig. 2. Block diagram of the neural network for predicting anomalies.

First, all time series (collected at different points in time statistical material on the values of any parameters of the studied process) are divided into sequence levels of length w , namely $X(i) = x(j), x(j + 1), \dots, x(j + w - 1)$, where i – the sequence number, $j = w(i - 1) + 1$ – the number of the first countdown in the sequence. At the prediction stage, the values for the next sequence $X(i + 1)$ are selected using the data of already performed measurements $X(1), X(2), \dots, X(i)$.

All data points in the presented data set have the same time grid and have significantly different absolute values. To reduce these variations and combine different dimensions, the normalization transformation is applied to each dimension separately:

$$x_i^{*(j)} = \frac{x_i^{(j)} - \bar{x}_i}{\sigma_i}$$

where $i = \overline{1, n}$; \bar{x}_i and σ_i – mean and standard deviation for each measurement.

The detection stage is based on the search for time points, where the root mean square error between the measured values of $X(i + 1)$ and the predicted $X'(i + 1)$ becomes greater than the pre-calculated threshold.

To detect anomalies, a recurrent neural network with a long-term memory architecture and an autoencoder were chosen, which allows to apply training without a teacher during using the method of error back propagation. A synchronous architecture was chosen for the autoencoder, among the advantages of which is the possibility of using the streaming mode of data processing and a relatively smaller number of neural network parameters compared to other architectures.

Finding neural network parameters is an iterative task. Most often, developers select the optimal values of the

structural parameters of the network and the size of the training sample, based on personal experience and repeated trials and errors. Therefore, the structural parameters of the network and the size of the training sample may not be optimal in terms of some function of the approximation error.

Several main rules for neuronal isolation have been presented, for example, the number of hidden neurons should be less than twice the size of the input layer [72,73]. Several statistical methods have also been developed, some of which are presented in Table 2. Although such methods are proposed, they are difficult to implement using commercial software packages. Typically, these recommendations are applicable to the specific cases of a particular network topology.

Table 2. Review of methods for selecting the number of hidden neurons [74].

Method	Number of hidden neurons
Li et al. method [75]	$N_h = (\sqrt{1 + 8n} - 1)/2$
Tamura and Tateishi method [76]	$N_h = N - 1$
Fujita method [77]	$N_h = K \log \ P_c Z\ / \log S$
Zhang et al. Method [78]	$N_h = 2^n / n + 1$
Jinchuan and Xinzhe method [79]	$N_h = (N_{in} + \sqrt{N_p}) / L$
Xu and Chen method [80]	$N_h = C_f (N/d \log N)^{0.5}$
Shibata and Ikeda method [81]	$N_h = \sqrt{N_i N_o}$
Hunter et al. Method [82]	$N_h = 2^n - 1$
Sheela and Deepa method [83]	$N_h = (4n^2 + 3) / (n^2 - 8)$

The amount of training set is also a parameter that optimizes the quality of neural network model.

The greater the amount of training set, the greater the amount of memory required to store it, time spent on training the neural network and gather information about the object. A training sample with a small amount of data is not informative enough to characterize the behavior of an object with acceptable quality. This leads to the fact that the network is often unable to predict the behavior of the object outside the examples of the training sample.

Insufficient number of neurons in the hidden layer will not allow to fully approximate the behavior of the object, and the prediction error will be large. But the more

complex the neural network, the more time it takes to learn and work on the task. The predictive power of the network may also decrease due to the effect of retraining. The neural network will show insignificant or insignificant details in the studied dependence, such as noise, the output vector will change significantly with small deviations of the input vector and the neural network will not be able to generalize – prediction of the output vector with input data not included in the training sample.

Based on the above information, the decision to make a variation of the neural network training parameters:

- training periods are 1, 4 or 12 months;
- number of hidden layers – 2 or 4;
- the number of hidden neurons in the first layer – 8, 15, 20, 24, 30, 45, 60, 90;
- the number of hidden neurons is halved in each subsequent layer.

3.3. Neural network training

A total dataset (retrospective data) of 15,120,000 measurements over 35 months from more than 107 sensors [83] has the following values:

- date in the format – day/month/year;
- time in the format – hour: minute: second;
- sensor ID;
- sensor_measure temperature value.

Sensors generate highly correlated multidimensional time series, which are aperiodic and synchronous (aligned in time), have multiscale processes, fast (long-term) and slow (short-term) subprocesses.

Each created model is evaluated on the basis of a graphical overview of the model trend. Models are evaluated using the following hierarchical criteria:

- suitability of the model on educational data;
- suitability of the model on the data for monitoring the normal state;
- detection of anomalies.

To be considered reliable, the model must meet all of the above criteria. The evaluation criteria are hierarchical, as each model must meet the previous criteria before evaluating the next criterion.

Table 3. Evaluation criteria forecasting models.

Model criteria	Suitability of the model during the training period or monitoring of normal condition	Detection of anomalies
Very bad ($0 < R^2 \leq 0.2$)	Models with a constant very significant deviation from the measured values or with a clearly inconsistent forecast	Unable to detect anomalies model is not reliable
Bad ($0.2 < R^2 \leq 0.4$)	The models follow the measured trend to some extent, but are characterized by constant significant deviations and periodic very significant deviations from the measured values	Unable to detect anomalies, the model follows the trend measured in the same or a higher level
Satisfactory ($0.4 < R^2 \leq 0.6$)	Models can record the dynamics of the process, but are characterized by constant slight deviations and periodic significant or very significant deviations	Insignificant indicators of anomalies, the model follows the measured trend
Good ($0.6 < R^2 \leq 0.85$)	Most of the time, the models have an accurate forecast with frequent minor deviations	Indicators leading to a more detailed analysis
Very good ($0.85 < R^2 \leq 1.0$)	The models follow the measured trend and have only a few slight deviations from the measured values	Clear indication of anomalies

Validation and verification of the used neural network models was performed. The criteria for evaluating forecasting models during network training based on the coefficient of determination are given (Table 3).

In Table 3, the coefficient of determination R^2 of the model of the dependence of the predicted values on the measured values is determined by the formula:

$$R^2 = 1 - \frac{\sum_i (y_i - f_i)^2}{\sum_i (y_i - \bar{y})^2},$$

where y – predicted values, f_i – current values, \bar{y} – average value.

Determination coefficients R^2 together with scattering graphs provide a fairly simple tool for estimating model prediction. However, it should be emphasized that R^2 is a relative indicator of the adequacy of the model (test and training samples are representative, the network structure is not redundant) and should not be used only as an indicator of efficiency.

Since in most cases the result of errors depends on the shape of forecasting errors are more commonly used performance based on the likelihood of true or false prediction. Assume that for each set of test data, the predicted probability of success $f(x_i)$ is compared with a fixed threshold C . If the probability is greater than C , it was assumed that the model successfully predicts; otherwise it was believed that the prediction is unsuccessful. The result of this procedure of comparing the true and predicted values of the dependent variable for n observations in the test data set can be presented in the form of a contingency table (table 4).

A simple indicator of the effectiveness of the model is accuracy, the proportion of correct predictions in the entire data set of testing:

$$ACC_C = \frac{TP_C + TN_C}{n}$$

Metrics such as precision and completeness are also used to evaluate the model algorithm. Precision within a class is the proportion of objects that actually belong to this class relative to all objects that the system has assigned to this class:

$$Precision_C = \frac{TP_C}{TP_C + FP_C} = \frac{TP_C}{P_C}$$

Recall of the system – a lot classifier found objects belonging to the class with respect to all objects of this class in the test sample:

$$Recall_C = \frac{TP_C}{TP_C + FN_C} = \frac{TP_C}{S_C}$$

The resulting accuracy of the classifier is calculated as the arithmetic mean of its accuracy for all classes. The same with completeness. Technically, this approach is called macro-averaging.

The average harmonic value of precision and completeness determines the F-score:

$$F_c = 2 \cdot \frac{Precision_C \cdot Recall_C}{Precision_C + Recall_C}$$

This formula gives the same weight of accuracy and completeness, so the F-score will fall equally with decreasing precision and recall.

Table 4. Contingency table.

	True class: Positive	True class: negative	Measures
$f(x) > C$, predicted class: positive	TP_C	FP_C	P_C
$f(x) \leq C$, predicted class: negative	FN_C	TN_C	N_C
Measures	S	F	n

As a result, it was found that the performance of models with a training period of 12 months differs the most (Fig. 3). In the figure, the dotted lines are divided into three periods: training, normal operation and detection of anomalies.

Table 5 and Figure 4 present the results of the calculation of precision, recall, F-score and efficiency of the 2C8 model.

Table 5. Contingency table for model 2C8.

	0	1	Accuracy	Macro avg	Weighted avg
Precision	0.98	0.84	0.97	0.91	0.97
Recall	0.98	0.86	0.97	0.92	0.97
F-score	0.98	0.85	0.97	0.91	0.97
Support	1604	196	0.97	1800	1800

As a result, it is established that in order to diagnose the condition of the service station it is necessary to substantiate and build a model that provides detection of an anomaly with a minimum number of false alarms during operation in the normal state.

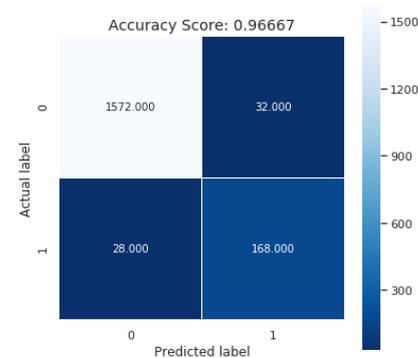


Fig. 4. Graphical representation of the contingency table for the 2C8 model.

4 Model testing results

It is accepted that during using models, they must be rated “good” or “very good” to be considered accurate enough for use in the diagnostic process (Table 3).

It is shown that 13 of the 96 created models meet the established criteria. Model 2C8 with a training period of 12 months, 30 automatically selected independent variables, 4 hidden layers and 90 hidden neurons showed very good results and was chosen as a reference model (Fig. 5). This made it possible to predict the anomalous state of the service station with a coefficient of determination of 0.965 based on the measurement of the surface temperature of thermal power equipment, which can be used as an indicator of the technical condition.

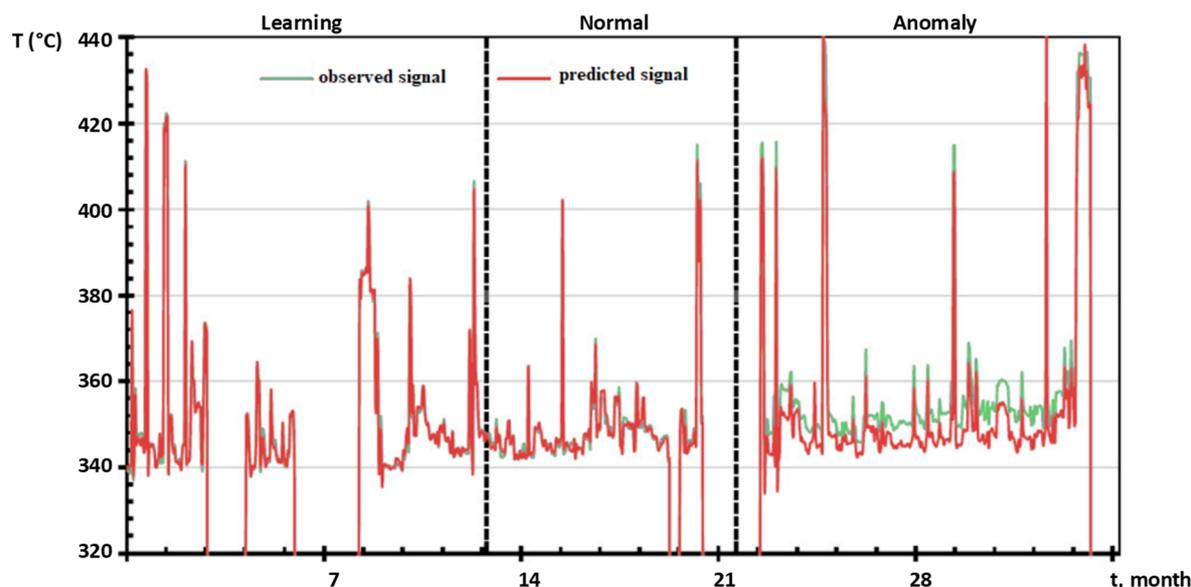


Fig. 5. The results of direct measurements (according to kaggle.com) and the predicted surface temperature of the boiler based on the reference model 2C8 for 35 months. Dotted lines indicate three periods: training, normal operation, detection of anomalies.

According to the results of testing the diagnostic system, it was found that models with a training period of one month show similar behavior to each other, regardless of other learning parameters. These models are characterized by satisfactory adaptation to these exercises. It is shown that the training period of one month is too short for the network to learn about the full features of the process.

It is shown that models with a training period of 4 months have better forecasting characteristics. Models 1B5–1B7 and 2B5–2B8 with 30 hidden neurons in the first hidden layer have relatively better overall R^2 values at all stages of training, monitoring and detection of abnormalities. These models are the group with the best performance, as they account for 7 out of 13 reliable models.

The accuracy of the models increases with the addition of two hidden layers. The set of variations among these models and the influence of the number of hidden neurons on the efficiency of the model are analyzed.

5 Conclusions

The structure of the neural network was developed on the basis of stacking algorithms of the recurrent neural network with the architecture of long short-term memory and autoencoder as part of the technical diagnostic system of the service station, which allowed to predict equipment failures in a small number of anomaly precedents. The use of the neural network in diagnostic systems increases the reliability of predicting anomaly conditions by 9%.

Models of neural networks are implemented and tested. Model 2C8 with a training period of 12 months, 30 automatically selected independent variables, 4 hidden layers and 90 hidden neurons showed very good results and was selected as a reference model. This made it possible to predict the anomalous condition of the service station with a coefficient of determination of 0.965 based on the measurement of the surface temperature of

complex thermal power facilities, which can be used as an indicator of the technical condition.

According to the results of research, it is established that models of neural networks with a training period of 4 months have relatively better characteristics, while the number of hidden layers and neurons is recommended to be determined experimentally.

Such approach to diagnostics of energy equipment can make energy industries more ecological, because concentrations of different pollutants in exhaust gases of thermal power facilities can use as information parameter of the process. So, continuous diagnostics on the bases of exhaust gases analysis can reduce the concentration of pollutants which throw out into the atmosphere.

References

1. S. Li, J. Li, Condition monitoring and diagnosis of power equipment: review and prospective. *High Voltage*, 2(2), 82-91 (2017). [doi:10.1049/hve.2017.0026](https://doi.org/10.1049/hve.2017.0026)
2. L. Ma, Y. Ma, K.Y. Lee, An Intelligent Power Plant Fault Diagnostics for Varying Degree of Severity and Loading Conditions. *IEEE Transactions on Energy Conversion*, 25(2), 546-554 (2010). [doi:10.1109/TEC.2009.2037435](https://doi.org/10.1109/TEC.2009.2037435).
3. V.P. Babak, S.V. Babak, M.V. Myslovych, A.O. Zaporozhets, V.M. Zvaritch. *Principles of Construction of Systems for Diagnosing the Energy Equipment*. In: Diagnostic Systems For Energy Equipments. Studies in Systems, Decision and Control, 281, 1-22 Springer, Cham (2020). [doi:10.1007/978-3-030-44443-3_1](https://doi.org/10.1007/978-3-030-44443-3_1)
4. V.P. Babak, S.V. Babak, M.V. Myslovych, A.O. Zaporozhets, V.M. Zvaritch, *Technical Provision of Diagnostic Systems*. In: Diagnostic Systems For Energy Equipments. Studies in Systems, Decision and

- Control, 281, 91-133 Springer, Cham (2020). [doi:10.1007/978-3-030-44443-3_4](https://doi.org/10.1007/978-3-030-44443-3_4)
5. G. Shavachych, B. Moroz, I. Pobocii, D. Kozenkov, V. Bysygin, *Automated Control Parameters Systems of Technological Process Based on Multiprocessor Computing Systems*. In: Arai K., Kapoor S. (eds) *Advances in Computer Vision. CVC 2019. Advances in Intelligent Systems and Computing*, 944, 666-688. Springer, Cham (2020). [doi:10.1007/978-3-030-17798-0_53](https://doi.org/10.1007/978-3-030-17798-0_53)
 6. I. Korobiichuk, A. Ladanyuk, D. Shumyhai, R. Boyko, V. Reshetiuk, M. Kamiski, *How to Increase Efficiency of Automatic Control of Complex Plants by Development and Implementation of Coordination Control System*. In: Szewczyk R., Kaliczyńska M. (eds) *Recent Advances in Systems, Control and Information Technology. SCIT 2016. Advances in Intelligent Systems and Computing*, 543, 189-195. Springer, Cham (2017). [doi:10.1007/978-3-319-48923-0_23](https://doi.org/10.1007/978-3-319-48923-0_23)
 7. A.O. Zaporozhets, *Hardware and Software Implementation of Modules of the System of the Fuel Combustion Control Process*. In: *Control of Fuel Combustion in Boilers. Studies in Systems, Decision and Control*, 287, 61-87. Springer, Cham. [doi:10.1007/978-3-030-46299-4_3](https://doi.org/10.1007/978-3-030-46299-4_3)
 8. L.A. Rusinov, I.V. Rudakova, V.V. Kurkina, *Real time diagnostics of technological processes and field equipment*. *Chemometrics and Intelligent Laboratory Systems*, 88(1), 18-25 (2007). [doi:10.1016/j.chemolab.2006.11.007](https://doi.org/10.1016/j.chemolab.2006.11.007)
 9. S. Babak, V. Babak, A. Zaporozhets, A. Sverdlova, *Method of statistical spline functions for solving problems of data approximation and prediction of objects state*. In: *CEUR Workshop Proceedings*, 2353, 810-821 (2019). Online available: <http://ceur-ws.org/Vol-2353/paper64.pdf>
 10. A. Zaporozhets, *Development of Software for Fuel Combustion Control System Based on Frequency Regulator*. In: *CEUR Workshop Proceedings*, 2387, 223-230. Online available: <http://ceur-ws.org/Vol-2387/20190223.pdf>
 11. S.M. Gertsyk, Y.I. Gyzhko, V.M. Zvarich, M.V. Myslovych, L.B. Ospapchuk, R.M. Sysak, *Use of autonomous measuring systems for diagnosing of electrical equipment with regard to its operating modes*. *Tekhnichna elektrodynamika*, 5, 116-120 (2018). [doi:10.15407/techmed2018.05.116](https://doi.org/10.15407/techmed2018.05.116)
 12. S. Hertshyk, Y. Gyzhko, V. Zvaritch, M. Myslovych, L. Ostapchuk, *Features of the Construction of Autonomous Diagnostic Systems of Electrical Equipment, Taking into Account the Modes of Its Operation*. In: 2019 IEEE 20th International Conference on Computational Problems of Electrical Engineering (CPEE), Lviv-Slavske, Ukraine, 1-3 (2019). [doi:10.1109/CPEE47179.2019.8949148](https://doi.org/10.1109/CPEE47179.2019.8949148).
 13. M.V. Myslovych, R.M. Sysak, L.B. Ostapchuk, Yu.I. Gyzhko, S.M. Hertsyk, *Algorithms of operation and software of multilevel system for monitoring and technical diagnostics of electrical power facilities equipment*. *Tekhnichna elektrodynamika*, 4, 86-88 (2016). [doi:10.15407/techmed2016.04.086](https://doi.org/10.15407/techmed2016.04.086)
 14. A. Zaporozhets, V. Eremenko, R. Serhienko, S. Ivanov, *Methods and Hardware for Diagnosing Thermal Power Equipment Based on Smart Grid Technology*. In: Shakhovska N., Medykovskyy M. (eds) *Advances in Intelligent Systems and Computing III. CSIT 2018. Advances in Intelligent Systems and Computing*, 871, 476-489. Springer, Cham (2020). [doi:10.1007/978-3-030-01069-0_34](https://doi.org/10.1007/978-3-030-01069-0_34)
 15. A.A. Zaporozhets, V.S. Eremenko, R.V. Serhienko, S.A. Ivanov, *Development of an Intelligent System for Diagnosing the Technical Condition of the Heat Power Equipment*. In: 2018 IEEE 13th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT), Lviv, 48-51 (2018). [doi:10.1109/STC-CSIT.2018.8526742](https://doi.org/10.1109/STC-CSIT.2018.8526742).
 16. A. Zaporozhets, *Analysis of Control System of Fuel Combustion in Boilers with Oxygen Sensor*. *Periodica Polytechnica Mechanical Engineering*, 63(4), 241-248 (2019). [doi:10.3311/PPme.12572](https://doi.org/10.3311/PPme.12572)
 17. L. Shkitsa, T. Yatsyshyn, M. Lyakh, O. Sydorenko, *Innovative approaches to the formation of environmental safety at the objects of oil and gas production*. In: *IOP Conf. Ser.: Mater. Sci. Eng.*, 749, 012009 (2020). [doi:10.1088/1757-899X/749/1/012009](https://doi.org/10.1088/1757-899X/749/1/012009)
 18. L. Shkitsa, T. Yatsyshyn, M. Lyakh, O. Sydorenko, *Means of atmospheric air pollution reduction during drilling wells*. In: *IOP Conf. Ser.: Mater. Sci. Eng.*, 144, 012009 (2016). [doi:10.1088/1757-899X/144/1/012009](https://doi.org/10.1088/1757-899X/144/1/012009)
 19. I.O. Popov, A. Iatsyshyn, V. Kovach, V. Artemchuk, I. Kameneva, D. Taraduda, V. Sobyna, D. Sokolov, M. Dement, T. Yatsyshyn, *Risk assessment for the population of Kyiv, Ukraine as a result of atmospheric air pollution*. *J. Health Pollut.* 10, 200303 (2020). [doi:10.5696/2156-9614-10.25.200303](https://doi.org/10.5696/2156-9614-10.25.200303)
 20. A. Zaporozhets, *Analysis of the Air Pollution Monitoring System in Ukraine*. In: Babak V., Isaienko V., Zaporozhets A. (eds) *Systems, Decision and Control in Energy I. Studies in Systems, Decision and Control*, 298, 85-110. Springer, Cham (2020). [doi:10.1007/978-3-030-48583-2_6](https://doi.org/10.1007/978-3-030-48583-2_6)
 21. T. Yatsyshyn, Y. Mykhailiuk, M. Liakh, I. Mykhailiuk, V. Savyk, I. Dobrovolsky, *Establishing the dependence of pollutant concentration on operational conditions at facilities of an oil and gas complex*. *Eastern-European Journal of Enterprise Technologies* 2/10(92), 56-63 (2018). [doi:10.15587/1729-4061.2018.126624](https://doi.org/10.15587/1729-4061.2018.126624)

22. N. Pobihun, Y. Korobeinykova, O. Pobihun, I. Iuras, *Ecological and monitoring studies of oil production territories and possibility of their use in recreation*. In: Proceedings of the XIII International Scientific Conference “Monitoring of Geological Processes and Ecological Condition of the Environment”, 2019, 1-5. (2019). [doi:10.3997/2214-4609.201903183](https://doi.org/10.3997/2214-4609.201903183)
23. A. Paletto, S. Bernardi, E. Pieratti, F. Teston, M. Romangnoli, Assessment of environmental impact of biomass power plants to increase the social acceptance of renewable energy technologies. *Heliyon*, 5(7), e02700. [doi:10.1016/j.heliyon.2019.e02070](https://doi.org/10.1016/j.heliyon.2019.e02070)
24. Z. Payandeh, K. Kheiralipour, M. Karimi, B. Khoshnevisan, Joint data envelopment analysis and life cycle assessment for environmental impact reduction in broiler production systems. *Energy*, 127, 768-774 (2017) [doi:10.1016/j.energy.2017.03.112](https://doi.org/10.1016/j.energy.2017.03.112)
25. A. Kaab, M. Sharifi, H. Mobli, A. Nabavi-Pelesaraci, K.-w. Chau, Combined life cycle assessment and artificial intelligence for prediction of output energy and environmental impacts of sugarcane production. *Science of The Total Environment*, 664, 1005-1019 (2019). [doi:10.1016/j.scitotenv.2019.02.004](https://doi.org/10.1016/j.scitotenv.2019.02.004)
26. T. Yatsyshyn, L. Shkitsa, O. Popov, M. Liakh, *Development of mathematical models of gas leakage and its propagation in atmospheric air at an emergency gas well gushing*. *Eastern-European Journal of Enterprise Technologies* 5/10(101), 49-59 (2019). [doi:10.15587/1729-4061.2019.179097](https://doi.org/10.15587/1729-4061.2019.179097)
27. E. Shaburov, A. Fedyukhin, O. Derevianko, *Analysis of Energy Efficiency of Municipal Solid Waste Usage for Thermal and Electrical Energy Production*. In: IOP Conf. Ser.: Earth Environ. Sci., 272, 022105 (2019). [doi:10.1088/1755-1315/272/2/022105](https://doi.org/10.1088/1755-1315/272/2/022105)
28. Y. Kyrylenko, I. Kameneva, O. Popov, A. Iatsyshyn, V. Artemchuk, V. Kovach, Source Term Modelling for Event with Liquid Radioactive Materials Spill. In: Babak V., Isaienko V., Zaporozhets A. (eds) *Systems, Decision and Control in Energy I. Studies in Systems, Decision and Control*, 298, 261-279. Springer, Cham (2020). [doi:10.1007/978-3-030-48583-2_17](https://doi.org/10.1007/978-3-030-48583-2_17)
29. L. Skitsa, T. Yatsyshyn, M. Liakh, O. Sydorenko, *Ways to improve safety of a pumping-circulatory system of a drilling rig*. *Mining of Mineral Deposits* 12(3), 71-79 (2018). [doi:10.15407/mining12.03.071](https://doi.org/10.15407/mining12.03.071)
30. T. Yatsyshyn, N. Glibovytska, L. Skitsa, M. Liakh, S. Kachala, *Biotechnogenic System Formed by Long-Term Impact of Oil Extraction Objects*. In: Babak V., Isaienko V., Zaporozhets A. (eds) *Systems, Decision and Control in Energy I. Studies in Systems, Decision and Control*, 298, 165-177. Springer, Cham (2020). [doi:10.1007/978-3-030-48583-2_11](https://doi.org/10.1007/978-3-030-48583-2_11)
31. O. Savko, I. Melnychuk, I. Hoby, N. Havadzyn, *Evaluation of the environmental taxation effectiveness in the field of oil and gas production*. *Procedia Environ. Sci. Eng. Manag.* 6(4), 607-617 (2019). Online available: http://www.procedia-emscem.eu/pdf/issues/2019/no4/69_Savko_19.pdf. Accessed 30 Nov 2020
32. O.M. Mandryk, L.M. Arkhypova, O.V. Pobigun, O.R. Maniuk, *Renewable energy sources for sustainable tourism in the Carpathian region*. IOP Conf. Ser.: Mater. Sci. Eng. 144, 012007 (2016). [doi:10.1088/1757-899X/144/1/012007](https://doi.org/10.1088/1757-899X/144/1/012007)
33. O. Mandryk, N. Moskalchuk L. Arkhypova, M. Prykhodk, O. Pobigun, *Prospects of environmentally safe use of renewable energy sources in the sustainable tourism development of the Carpathian region of Ukraine*. E3S Web Conf. 166, 04005 (2020). [doi:10.1051/e3sconf/202016604005](https://doi.org/10.1051/e3sconf/202016604005)
34. O.M. Mandryk, N.R. Moskalchuk, L.M. Arkhypova, M.M. Pryhodko, O.V. Pobigun, *Research quantitative indicators of the potential of solar energy in the Carpathian region of Ukraine*. IOP Conf. Ser.: Mater. Sci. Eng. 749, 012033 (2020). [doi:10.1088/1757-899X/749/1/012033](https://doi.org/10.1088/1757-899X/749/1/012033)
35. A. Zaporozhets, Overview of Quadrocopters for Energy and Ecological Monitoring. In: Babak V., Isaienko V., Zaporozhets A. (eds) *Systems, Decision and Control in Energy I. Studies in Systems, Decision and Control*, 298, 15-36. Springer, Cham (2020). [doi:10.1007/978-3-030-48583-2_2](https://doi.org/10.1007/978-3-030-48583-2_2)
36. M.L. Myrontsov, *Multi-Probe Hardware for Electrometry of Oil and Gas Wells*, *Science and innovation* 14(3), 51-56 (2018). [doi:10.15407/scine14.03.051](https://doi.org/10.15407/scine14.03.051)
37. M.L. Myrontsov, *The method to solve the inverse problem of lateral logging sounding and lateral logging*. In: Proceedings of the XIII International Scientific Conference “Monitoring of Geological Processes and Ecological Condition of the Environment”, 2019, 1-5. (2019). [doi:10.3997/2214-4609.201903244](https://doi.org/10.3997/2214-4609.201903244)
38. M.L. Myrontsov, *Lateral logging sounding and lateral logging complex effective inverse problem solving method*. In: Proceedings of the 19th International Conference Geoinformatics – Theoretical and Applied Aspects 2020, 2020, 1-5. (2020). [doi:10.3997/2214-4609.2020geo092](https://doi.org/10.3997/2214-4609.2020geo092)
39. M.L. Myrontsov, *Electrometry effective inverse problem solving method*. In: Proceedings of the 19th International Conference Geoinformatics – Theoretical and Applied Aspects 2020, 2020, 1-5. (2020). [doi:10.3997/2214-4609.2020geo090](https://doi.org/10.3997/2214-4609.2020geo090)
40. M.L. Myrontsov, *The method to research equivalent solutions zones for inverse problem of well logging electrometry*. In: Proceedings of the XIII International Scientific Conference “Monitoring of Geological Processes and Ecological Condition of the Environment”, 2019, 1-5. (2019). [doi:10.3997/2214-4609.201903217](https://doi.org/10.3997/2214-4609.201903217)

41. M.L. Myrontsov, *A new method and program for multiprobe electric logging quantitative interpretation*. In: Proceedings of the 11th EAGE International Conference on Geoinformatics - Theoretical and Applied Aspects, 00028 (2012).
42. A.V. Iatsyshyn, V.O. Kovach, V.O. Lyubchak, Y.O. Zuban, A.G. Piven, O.M. Sokolyuk, A.V. Iatsyshyn, O.O. Popov, V.O. Artemchuk, M.P. Shyshkina, *Application of augmented reality technologies for education projects preparation*. In: CEUR Workshop Proceedings, 2643, 134–160 (2020). Online available: <http://ceur-ws.org/Vol-2643/paper07.pdf>.
43. A. Iatsyshyn, A. Iatsyshyn, V. Artemchuk, I. Kameneva, V. Kovach, O. Popov, *Software tools for tasks of sustainable development of environmental problems: peculiarities of programming and implementation in the specialists' preparation*. E3S Web Conf. 166, 01001 (2020). doi:10.1051/e3sconf/202016601001
44. A. Iatsyshyn, A. Iatsyshyn, V. Kovach, I. Zinovieva, V. Artemchuk, O. Popov, O. Cholyskhina, O. Radchenko, O. Radchenko, A. Turevych, *Application of Open and Specialized Geoinformation Systems for Computer Modelling Studying by Students and PhD Students*. In: CEUR Workshop Proceedings 2732, 893-908 (2020). Online available: <http://ceur-ws.org/Vol-2732/20200893.pdf>.
45. V. Gurieiev, Yu. Kutsan, A. Iatsyshyn, A. Iatsyshyn, V. Kovach, E. Lysenko, V. Artemchuk, O. Popov, *Simulating Systems for Advanced Training and Professional Development of Energy Specialists in Power Sector*. In: CEUR Workshop Proceedings 2732, 693-708 (2020). Online available: <http://ceur-ws.org/Vol-2732/20200693.pdf>
46. S. Kis, L. Mosora, Y. Mosora, O. Yatsiuk, G. Malynovska, S. Pobihun, *Personnel Certification as a Necessary Condition for Enterprise' Staff Development*, Management Systems in Production Engineering, 28(2), 121-126 (2020). doi:10.2478/mspe-2020-0018
47. S. Hawkins, H. He, G. Williams, R. Baxter, *Outlier detection using replicator neural networks*. In: Kambayashi Y., Winiwarter W., Arikawa M. (eds) Data Warehousing and Knowledge Discovery. DaWaK 2002. Lecture Notes in Computer Science, 2454, 170-180. Springer, Berlin, Heidelberg (2002). doi:10.1007/3-540-46145-0_17
48. W. Yan and L. Yu, *On Accurate and Reliable Anomaly Detection for Gas Turbine Combustors: A Deep Learning Approach* (2019). Online available: <https://arxiv.org/ftp/arxiv/papers/1908/1908.09238.pdf>.
49. M. Agyemang, K. Barker and R. Alhaji, *A comprehensive survey of numeric and symbolic outlier mining techniques*. Intelligent Data Analysis, 10(6), 521-538 (2006). doi: 10.3233/IDA-2006-10604
50. A. Patcha, J.-M. Park, *An overview of anomaly detection techniques: Existing solutions and latest technological trends*. Computer Networks, 51(12), 3448–3470 (2007).
51. M. Markou, S. Singh, *Novelty detection: A review-part 1: Statistical approaches*. Signal Processing, 83(12), 2481-2497 (2003). doi:10.1016/j.sigpro.2003.07.018
52. G. Thatte, U. Mitra, J. Heidemann, *Parametric methods for anomaly detection in aggregate traffic*. IEEE/ACM Transactions on Networking, 19(2), 512-525 (2011). doi: 10.1109/TNET.2010.2070845
53. A. Kind, M. P. Stoecklin, X. Dimitropoulos, *Histogram-based traffic anomaly detection*, IEEE Transactions on Network and Service Management, 6(2), 110-121 (2009). doi: 10.1109/TNSM.2009.090604
54. M. Zhang, B. Xu, J. Gong, *An Anomaly Detection Model Based on One-Class SVM to Detect Network Intrusions*. In: 2015 11th International Conference on Mobile Ad-hoc and Sensor Networks (MSN), Shenzhen, 102-107 (2015). doi: 10.1109/MSN.2015.40
55. M. Sharma, K. Das, M. Bilgic, B. Matthews, D. Nielsen, N. Oza, *Active Learning with Rationales for Identifying Operationally Significant Anomalies in Aviation*. In: Berendt B. et al. (eds) Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2016. Lecture Notes in Computer Science, 9853. Springer, Cham, 209-225 (2016). doi:10.1007/978-3-319-46131-1_25
56. S. Omar, A. Ngadi, H.H. Jebur, *Machine Learning Techniques for Anomaly Detection: An Overview*, International Journal of Computer Applications, 79(2), 33-41 (2013).
57. C. M. Bishop, *Neural networks for pattern recognition*. Oxford University Press (1995).
58. D. Thierens, The Linkage Tree Genetic Algorithm. In: Schaefer R., Cotta C., Kołodziej J., Rudolph G. (eds) Parallel Problem Solving from Nature, PPSN XI. PPSN 2010. Lecture Notes in Computer Science, 6238, 264-273. Springer, Berlin, Heidelberg (2010). doi:10.1007/978-3-642-15844-5_27
59. W.-C. Lin, S.-W. Ke, C.-F. Tsai, CANN: An intrusion detection system based on combining cluster centers and nearest neighbors. Knowledge Based Systems, 78, 13-21 (2015). doi:10.1016/j.knsys.2015.01.009
60. M. Panda, A. Abraham, M.R. Patra, *Hybrid intelligent systems for detecting network intrusions*. Security Comm. Networks, 8(16), 2741-2749 (2012). doi:10.1002/sec.592
61. S. Erfani, M. Baktashmotlagh, S. Rajasegarar, S. Karunasekera and C. Leckie, *A randomised nonlinear approach to large-scale anomaly detection*. Proceedings of the 29th AAAI Conference on

- Artificial Intelligence (AAAI-15). The AAAI Press, United States of America, 432-438 (2015)
62. E. Alpaydin, *Introduction to Machine Learning*. MIT Press (2020)
63. P. J. Werbos, *Applications of Advances in Nonlinear Sensitivity Analysis*. In: Drenick R.F., Kozin F. (eds) System Modeling and Optimization. Lecture Notes in Control and Information Sciences, 38, 762-770. Springer, Berlin, Heidelberg (1982). doi:10.1007/BFb0006203
64. D. P. Kingma, J. Ba, Adam: *A Method for Stochastic Optimization*. International Conference on Learning Representations, San Diego, USA (2015). arXiv:1412.6980
65. J. Duchi, E. Hazan, Y. Singer, *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization*. Journal of Machine Learning Research, 12, 2121-2159 (2011).
66. S. Hochreiter, Y. Bengio, P. Frasconi, J. Schmidhuber, *Gradient Flow in Recurrent Nets: the Difficulty of Learning Long-Term Dependencies*. In: J.F. Kolen, S.C. Kremer (eds). A Field Guide to Dynamical Recurrent Networks. Wiley-IEEE Press, 237-243. (2001)
67. J. A. Perez-Ortiz, F. A. Gers, D. Eck, J. Schmidhuber, *Kalman filters improve LSTM network performance in problems unsolvable by traditional recurrent nets*. Neural Networks, 16(2), 241-250 (2003). doi:10.1016/S0893-6080(02)00219-8
68. F. Gers, J. Schmidhuber, F. Cummins, *Learning to forget: continual prediction with LSTM*. In: 9th International Conference on Artificial Neural Networks: ICANN '99, 850-855 (1999). doi:10.1049/cp:19991218
69. A. Graves, A-r. Mohamed and G. Hinton, *Speech recognition with deep recurrent neural networks*. In: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, 6645-6649 (2013). doi:10.1109/ICASSP.2013.6638947.
70. G. V. Houdt, C. Mosquera, G. Nápoles, *A review on the long short-term memory model*. Artificial Intelligence Review, 53, 5929-5955 (2020). doi:10.1007/s10462-020-09838-1
71. R. Metzler, W. Kinzel, L. Ein-Dor, I. Kanter, *Generation of unpredictable time series by a neural network*. Physical Review E, 63(5), 056126 (2001). doi:10.1103/PhysRevE.63.056126
72. K. Hornik, *Approximation capabilities of multilayer feedforward networks*, Neural Networks, 4(2), 251-257 (1990). doi:10.1016/0893-6080(91)90009-T
73. J. Heaton, *The Number of Hidden Layers*. Heaton Research (2017). Online available: <https://www.heatonresearch.com/2017/06/01/hidden-layers.html>.
74. K. G. Sheela, S. N. Deepa, *Review on Methods to Fix Number of Hidden Neurons in Neural Networks*. Mathematical Problems in Engineering, 2013, 425740 (2013). doi:10.1155/2013/425740
75. J.-Y. Li, T. W. S. Chow, Y.-L. Yu, *Estimation theory and optimization algorithm for the number of hidden units in the higher-order feedforward neural network*. Proceedings of ICNN'95 - International Conference on Neural Networks, Perth, WA, Australia, 3, 1229-1233 (1995). doi:10.1109/ICNN.1995.487330.
76. S. Tamura, M. Tateishi, *Capabilities of a four-layered feedforward neural network: four layers versus three*. IEEE Transactions on Neural Networks, 8(2), 251-255 (1997). doi:10.1109/72.557662
77. O. Fujita, *Statistical estimation of the number of hidden units for feedforward neural networks*. Neural Networks, 11(5), 851-859 (1998). doi:10.1016/S0893-6080(98)00043-4
78. Z. Zhang, X. Ma, Y. Yang, *Bounds on the number of hidden neurons in three-layer binary neural networks*. Neural Networks, 16(7), 995-1002 (2003). doi:10.1016/S0893-6080(03)00006-6
79. J. Ke, X. Liu, *Empirical Analysis of Optimal Hidden Neurons in Neural Network Modeling for Stock Prediction*. In: 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, Wuhan, 828-832. (2008). doi:10.1109/PACIIA.2008.363.
80. S. Xu, L. Chen, *A novel approach for determining the optimal number of hidden layer neurons for FNN's and its application in data mining* In: 5th International Conference on Information Technology and Applications (ICITA 2008), 23-26 June 2008, Cairns, Queensland, Australia, 683-686 (2008). Online available: <https://eprints.utas.edu.au/6995/>
81. K. Shibata, Y. Ikeda, *Effect of number of hidden neurons on learning in large-scale layered neural networks*. In: 2009 ICCAS-SICE, Fukuoka, 5008-5013 (2009).
82. D. Hunter, H. Yu, M. S. Pukish III, J. Kolbusz, B. M. Wilamowski, *Selection of Proper Neural Network Sizes and Architectures—A Comparative Study*. IEEE Transactions on Industrial Informatics, 8(2), 228-240, (2012). doi:10.1109/TII.2012.2187914
83. *Predictive Equipment Failures. Data*. Online available: <https://www.kaggle.com/c/equipfailstest/data>