

DATA PROCESSING IN PREDICTIVE ANALYTICS SYSTEM FOR ENERGY FACILITIES

*Shcherbatov I.A., Tsurikov G.N., Maksimova E.V.,
Bukanyov A.N., Nazarenko A.S.*

Keywords: predictive analytics system, power engineering, neural networks, big data, predictive model, forecasting, diagnostic parameter.

Abstract. The description of methods and algorithms of data processing in the developed system of predictive analytics for energy objects is given.

ОБРАБОТКА ДАННЫХ В СИСТЕМЕ ПРЕДИКТИВНОЙ АНАЛИТИКИ ДЛЯ ОБЪЕКТОВ ЭНЕРГЕТИКИ

*Щербатов И.А., Цуриков Г.Н., Максимова Е.В.,
Буканев А.Н., Назаренко А.С.*

Ключевые слова: система предиктивной аналитики, энергетика, нейронные сети, большие данные, предиктивная модель, прогнозирование, диагностический параметр.

Аннотация. Приведено описание методов и алгоритмов обработки данных в разработанной системе предиктивной аналитики для объектов энергетики.

In [1], a review of the state of the problem of developing predictive analytics systems (PAS) for energy enterprises was carried out, the applied methods of data processing and forecasting of various indicators were identified. The processes of data preprocessing in PAS can be attributed to three stages: creating a model; adaptation of the model; forecasting using a model.

To build an adequate model for predicting the technical condition (TC) and reliability of energy equipment (predictive models), data should be prepared accordingly. The attribution of the same processing processes to the stage of adaptation of the predictive model is due to the fact that often the structure of the model must be adjusted as new, previously unused data are received. It can be argued about the continuous adaptation of the predictive model to improve the accuracy and quality of forecasting parameter values, their trends, equipment status and its reliability characteristics.

The data entering the PAS can be either measured (or obtained from accounting systems) or expert (expert knowledge that requires additional formalization before processing in the system). Virtually all stages of data preprocessing are aimed at eliminating the uncertainty of technological parameters, calculated indicators and expert knowledge.

Initially, it is necessary to determine those parameters and indicators that will be diagnostic (DP) and used to predict the technical condition and reliability of power equipment. The diagnostic parameters will be those technological parameters and calculated indicators that allow us to establish a functional relationship between them and the reliability of a piece of equipment, which provides a reliability

forecast with some predetermined accuracy (Fig. 1, where R , $R_{\max}$ – are the current and maximum correlation coefficients, AM , $AM_{\max}$ – are expert scores).

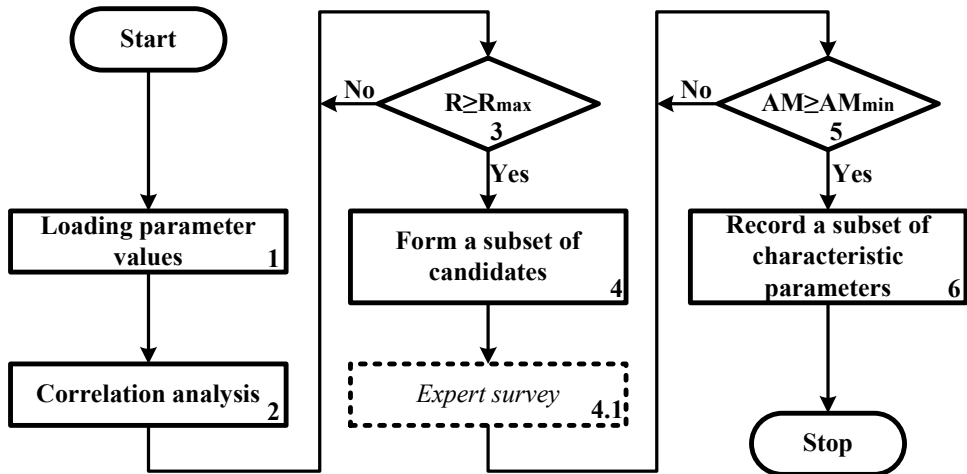


Fig. 1. Algorithm for selecting diagnostic parameters

When studying the methods of gaps in the data, it was found that for the application of PAS components as part of the mathematical support, the nearest neighbors method should be used, and if there is a sufficient amount of the training sample, the regression model method should be used. And for the classification of possible anomalies in data samples, it is proposed to use artificial neural networks (ANNs). The verification of the adequacy of outlier identification models was carried out by comparison with the calculation methods given in ISO GOST R 16269-4-2017.

To forecast defects, trends are applied. Getting a trend – performing the operation of fitting the time series of the DP. Fitting can be performed using the moving average method or exponential fitting. The moment of time corresponding to the intersection point of the trend line and the line corresponding to the limit value of the parameter is the time moment of the improper TC, for example, it can characterize a defect.

To predict the values of a single parameter, ANN technology is used. A specific parameter is a one-dimensional time series. Each defect of a piece of equipment is characterized by a specific set of DP. The construction of the functional dependence of emerging defects on changing DPs allows us to solve the problem of timely planning of measures to bring equipment into an appropriate TC [2].

Thus, the use of the above methods and algorithms for processing data and DP values provides the construction of an adequate mathematical model of the object and the solution of the problems of forecasting parameters, defects, failures and TC, which is not affected by anomalous values in the data, which reduce the quality and accuracy of the forecasts obtained.

References / Список литературы

1. Shcherbatov I.A. Current state of predictive analytics systems development in the energy sector // Journal of Advanced Research in Technical Science. 2019. № 14-2. P. 118-123.
2. Protalinsky O.M., Shcherbatov I.A., Stepanov P.V. Identification of the actual state and entity availability forecasting in power engineering using neural-network technologies // J. Physics: International Conference Series "Problems of Thermal Physics and Power Engineering" PTPPE 2017". 2017. P. 012289.

Щербатов Иван Анатольевич – кандидат технических наук, доцент, доцент кафедры «Автоматизированные системы управления тепловыми процессами», shcherbatovia@mpei.ru.	Shcherbatov Ivan Anatolyevich – Candidate of Technical Sciences, Associate Professor, Associate Professor of the Department of Automated Control Systems of Thermal Processes, shcherbatovia@mpei.ru
Цуриков Григорий Николаевич – магистрант, grishatsurikov9826@yandex.ru	Tsurikov Grigory Nikolaevich – undergraduate, grishatsurikov9826@yandex.ru
Максимова Екатерина Владимировна – аспирант, maxim1995ka@mail.ru	Maksimova Ekaterina Vladimirovna – postgraduate, maxim1995ka@mail.ru
Буканев Александр Николаевич – бакалавр, santa200207@mail.ru	Bukanyov Alexander Nikolaevich – bachelor, santa200207@mail.ru
Назаренко Анастасия Станиславовна – бакалавр, nastg0@mail.ru	Nazarenko Anastasia Stanislavovna – бакалавр, nastg0@mail.ru
Национальный исследовательский университет «МЭИ», Москва, Россия	National Research University Moscow Power Engineering Institute, Moscow, Russia

Received 19.09.2019