ISSN 2658-5014 (Print)
2021. Т. 3, № 2 (6) С. 26–34

СИИТ
СИСТЕМНАЯ ИНЖЕНЕРИЯ И ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ

ISSN 2686-7044 (Online)
http://siit.ugatu.su

УДК 004.896

APPLICATION OF LSTM NEURAL NETWORKS FOR DIAGNOSTICS OF THE STATE OF NODES OF METAL-CUTTING MACHINES

K. A. MASALIMOV[1], R. A. MUNASYPOV[2]

[1] masalimov.k.a@gmail.com, [2]rust40@mail.ru

Ufa State Aviation Technical University, Russia

Поступила в редакцию 24 февраля 2021 г.

Abstract. This article presents a description of approach for development the predictive diagnostic models of computer numerical control (CNC) metal-cutting machine nodes. This approach based on using of bidirectional long short-term memory (BiLSTM) neural networks. The architecture of such neural networks, the method of preprocessing the data recorded during the operation of the machine tool are described. Examples of the application of the approach for diagnostics of the state of the cutting tool and bearings of electric motors on the machine tool are presented. To estimate the remaining lifetime of the cutting tool, the proposed BiLSTM model uses indirect information - vibration values and dynamometry along three axes. The article presents a comparison of the data obtained from the diagnostic model for assessing the maximum wear of the tool edge and real data for the values from the test sample. A diagnostic model has been developed to assess the state of the bearings of an electric motor by vibration values, which determines the presence of a malfunction and one of four classes - a defect in the cage, ball, inner ring or outer ring of the bearing. The optimization of the classifier model was carried out taking into account the specifics of vibration signals in the presence of defects in the bearings of the electric motor of a metal-cutting machine.

Key word: metal-cutting machining; diagnostic; cutting tool state; cutting tool wear; bearings state; deep neural networks; long short-term memory; neural networks; monitoring system; predictive diagnostic.

INTRODUCTION

Timely determination of the condition of the equipment and its maintenance has been and remains an important task in any production. Moving to the concept of Industry 4.0 and digital factories leads to the fact that during the operation of equipment, huge amounts of data are stored [1].

It should be noted that the operation of the same type of equipment allows using the accumulated information not only for analyzing the technological and economic indicators of the efficiency of a particular type of machine tool [2], but also for the automatic determination of pre-defect states and generation of such software blocks for integration into the control system of a specific machine tool.

This work was supported by the Ministry of Education of the Russian Federation, project 0838-2020-0006. "Fundamental study of new principles for the creation of promising electromechanical energy converters with characteristics above the world level, with increased efficiency and minimum specific indicators, using new highly efficient electrotechnical materials."

There are a number of works devoted to the development of diagnostic models based on various types of neural networks for determining the state of nodes of CNC machines.

Most of the approaches are based on determining the state of any machine tool unit based on operational vibrometric information and other time signals.

For example, in a number of works examples of determining the state of a cutting tool of a CNC machine tool are presented using neural networks with long short-term memory [3–5], convolutional neural networks [6–7], fuzzy neural networks [8].

There are also works devoted to the determination of bearing defects in electric motors of CNC machine tools.

In this work we show approach to use deep learning to develop diagnostic models of a CNC machine tool based on bi-directional LSTM.

APPROACH BASED ON BILSTM NEURAL NETWORKS. BIDIRECTIONAL LSTM

Recurrent neural networks with long short-term memory (LSTM) [9] are well suited for solving the task of classifying and predicting the time series in cases where the boundaries of events in the studied system are not strictly defined and the dependence of some events on others is separated by a certain time interval.

LSTM is able to memorize dependencies, both for a short period and for a long period due to the inclusion of "forgetting gates" in its composition.

One node of such a neural network is a unit that includes several gates, each of which performs its function.

A graphical representation of the architecture of the one unit of LSTM presented in Fig. 1.

At each time step t in the LSTM, the hidden state h^t is updated with current data, the hidden state at the previous time step h^{t-1}, the input gate i^t, the forgetting gate f^t, output gate o^t and memory cell c^t.

The update process is performed according to the system of equations (1).

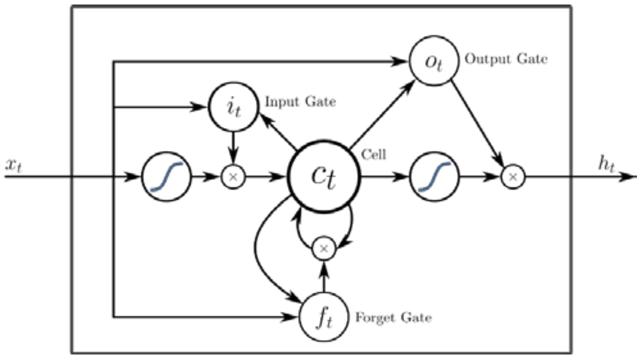

Fig. 1. Architecture of the one unit of LSTM

$$i^t = \sigma\left(W^i x^t + V^i h^{t-1} + b^i\right),$$
$$f^t = \sigma\left(W^f x^t + V^f h^{t-1} + b^f\right),$$
$$o^t = \sigma\left(W^o x^t + V^o h^{t-1} + b^o\right), \qquad (1)$$
$$c^t = f^t \odot c^{t-1} + i^t \odot \tanh\left(W^c x^t + V^c h^{t-1} + b^c\right),$$
$$h^t = o^t \odot \tanh(c^t),$$

where the model parameters, including $W \in R^{d\times k}$, $V \in R^{d\times d}$, as well as $b \in R^d$ are constant for all stages of the model's work and get values during model training, σ is the sigmoidal activation function, $\odot$ stands for the Hadamard product, and k is the hyperparameter representing the dimension of hidden vectors.

Thus, the system of equations (1) determines the function of the hidden layer H.

LSTM is designed directly for processing serial data expressed as a time series.

In addition, the output signal at the final time stage is used to predict the output signal on the linear regression layer, as shown in the equation (2).

$$\bar{y}_i = W^r h_i^T \qquad (2)$$

where $W^r \in R^{k\times z}$ and z is the dimension of the output of the entire model. To train the model, the predicted value of the target value $\bar{y}$ is compared with the true value of the target value y, and the mean-squared error is calculated as the loss function (3).

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\bar{y}_i - y_i)^2 \qquad (3)$$

where n is the size of the training sample.

The disadvantage of simple LSTM is the fixation of the considered values of time series only in the opposite direction - from the current state to the initial one.

When diagnosing the state of complex objects, serial data from sensors have stable time dependences and it makes sense when building models to take into account not only the direct dependence of the value, but also the inverse.

To solve this problem, a bidirectional LSTM is used. Bidirectional LSTM [10] are able to process input time series data in two directions - forward and reverse (each direction has its own hidden layer of a neural network), and then transmit their output layer of linear regression.

The following system of equations (4) and (5) determine the corresponding function of the hidden layer, the signs → and ← denote the direct and reverse processes.

Thus, the system of equations (4) characterizes the hidden layer for the direct LSTM.

$$\begin{gathered}\vec{i}^t = \sigma\left(\overrightarrow{W}^i \vec{x}^t + \overrightarrow{V}^i \vec{h}^{t-1} + \vec{b}^i\right),\\ \vec{f}^t = \sigma\left(\overrightarrow{W}^f \vec{x}^t + \overrightarrow{V}^f \vec{h}^{t-1} + \vec{b}^f\right),\\ \vec{o}^t = \sigma\left(\overrightarrow{W}^o \vec{x}^t + \overrightarrow{V}^o \vec{h}^{t-1} + \vec{b}^o\right),\\ \vec{c}^t = \vec{f}^t \odot \vec{c}^{t-1} + \vec{i}^t \odot \tanh\left(\overrightarrow{W}^c \vec{x}^t + \overrightarrow{V}^c \vec{h}^{t-1} + \vec{b}^c\right),\\ \vec{h}^t = \vec{o}^t \odot \tanh(\vec{c}^t)\end{gathered} \quad (4)$$

The system of equations (5) characterizes the hidden layer for the inverse LSTM.

$$\begin{gathered}\overleftarrow{i}^t = \sigma\left(\overleftarrow{W}^i \overleftarrow{x}^t + \overleftarrow{V}^i \overleftarrow{h}^{t-1} + \overleftarrow{b}^i\right),\\ \overleftarrow{f}^t = \sigma\left(\overleftarrow{W}^f \overleftarrow{x}^t + \overleftarrow{V}^f \overleftarrow{h}^{t-1} + \overleftarrow{b}^f\right),\\ \overleftarrow{o}^t = \sigma\left(\overleftarrow{W}^o \overleftarrow{x}^t + \overleftarrow{V}^o \overleftarrow{h}^{t-1} + \overleftarrow{b}^o\right),\\ \overleftarrow{c}^t = \overleftarrow{f}^t \odot \overleftarrow{c}^{t-1} + \overleftarrow{i}^t \odot \tanh\left(\overleftarrow{W}^c \overleftarrow{x}^t + \overleftarrow{V}^c \overleftarrow{h}^{t-1} + \overleftarrow{b}^c\right),\\ \overleftarrow{h}^t = \overleftarrow{o}^t \odot \tanh(\overleftarrow{c}^t).\end{gathered} \quad (5)$$

Then the full representation of the hidden layer of bidirectional LSTM is the concatenation of the vectors of the direct and reverse processes (6).

$$h^t = \vec{h}^t \cdot \overleftarrow{h}^t \quad (6)$$

DATA COLLECTION AND PREPROCESSING

In the course of technological processing by CNC machines, information about the operating modes of the machine is accumulated - the workpiece being processed, the frequency of rotation of the machine units, the processing time and time series value information - vibrometry, electrical parameters, etc. The amount of time value information depends on the equipment of the machine with sensors. All received information is accumulated in a single enterprise-level database. Also in this database, protocols are recorded for the failure of machine units, replacement of consumables and repairs. All recorded time value information from the sensors is stored in two forms - raw, that is, directly the values recorded from the sensors and normalized [11].

Data normalization is performed according to equation (7).

$$v_{ni} = \frac{v_i}{H_{max} - H_{min}} \quad (7)$$

where v_{ni} is normalized value, v_i – real value from the sensor, H_{max} and H_{min} – upper and lower measurement limits of the sensor.

DIAGNOSTIC OF CUTTING TOOL STATE

We used the results of metal processing on a high-speed CNC machine RodersTech RFM760 [12] (Fig. 2) as input data for the development of diagnostic models of the state of the cutting tool.

Fig. 2. CNC machine Roders Tech RFM760

The machine is additionally equipped with vibration acceleration sensors for recording vibration values along three axes and a three-axis platform dynamometer installed between the processing table and the workpiece to measure the values of cutting forces.

The cutting tool is a tungsten carbide cutter with three flutes. The amount of tool wear was measured using a LEICA MZ12 microscope [13]. Stainless steel of hardness 52HRC [14] was used as the material to be processed. The spindle rotation speed is 10400 min^{-1}; feed speed – 1555 mm / min; radial cutting depth (Y-axis) – 0.125 mm; axial cutting depth (Z-axis) – 0.2 mm. Data from vibration acceleration sensors and a dynamometer were recorded with a sampling rate of 50 kHz. CNC machine processing was carried out using 6 cutting tools. During processing, for each tool many times (more than 300 times for each tool), data from vibration and force sensors and the total amount of wear of each cutting edge of the tool after the next milling transition are recorded. Milling transitions were carried out with the same cutting length, that is, the cutting path was a constant.

Fig. 3 shows the dependence of the wear value of each cutting edge for one of the cutting tools with an increase of milling transitions.

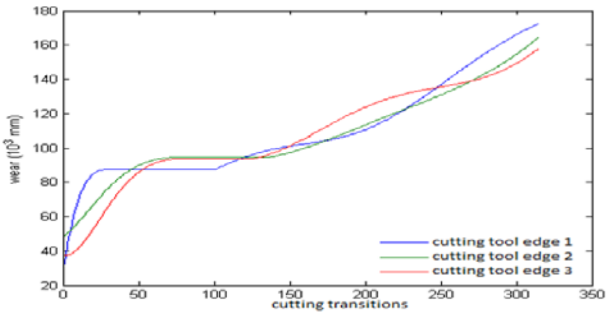

Fig. 3. Dependence of the wear value of each cutting edge for one of the cutting tools with an increase of milling transitions

DEVELOPMENT OF A DIAGNOSTIC MODEL TO PREDICT THE AMOUNT OF TOOL WEAR

The task of assessing wear is formulated as follows: from the values from the vibration and force sensors, it is necessary to estimate how many more processing transitions it is possible to carry out using the cutter – the remaining lifetime. The restrictions – the cutter is considered suitable for further processing if the wear of any of the cutting edges does not exceed the value of 165*10^{-3} mm.

Equation (8) is the function of evaluation the predicted value of diagnostic models.

$$S(d) = \begin{cases} e^{-d/10} - 1, \; if\ d < 0 \\ e^{d/4.5} - 1, \; if\ d \geq 0 \end{cases} \qquad (8)$$

$$d = c_M - c_F$$

where c_M is the residual life value predicted by the model, c_F is the actual residual life value, d is the model prediction error. It should also be noted that the evaluation function is set in such a way that an overestimated estimate of the residual life has a greater exponential penalty [8].

The experimental results are divided into two main sets – the data obtained during processing by three cutting tools were used in the training of the diagnostic model; data obtained during processing by the other three cutting tools were used to test the final model. Training and testing of diagnostic models were carried out using the Keras package [15].

The bidirectional LSTM includes two LSTMs (100 neurons in the hidden layer and 1 neuron in the output), differing in the direction of the input time series and a fully connected layer to combine the results with one output neuron.

Fig. 4 shows the resulting dependences of the wear (maximum wear value of any edge) of the cutting on the number of milling transitions, obtained from the actual data (real value) and wear values predicted by the diagnostic model based on the bidirectional LSTM for one of the test cutters. Dependence deviations are within 2.5%.

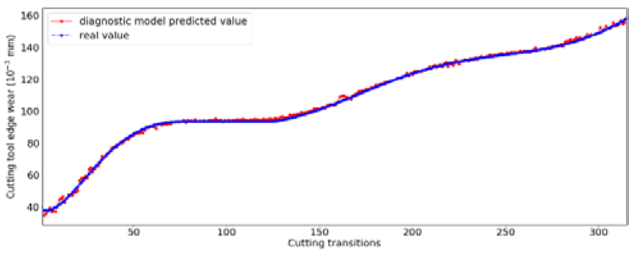

Fig. 4. Real wear value of cutting edge and the value predicted by the bi-directional LSTM diagnostic model

Therefore, the proposed diagnostic model based on the bidirectional LSTM quite accurately determines the value of wear according to information obtained from vibration and force sensors and can be used to assess the wear of a cutting tool for operational use of the CNC machine.

DEVELOPMENT OF A DIAGNOSTIC MODEL FOR DETERMINING A DEFECT IN THE BEARING OF THE MACHINE SPINDLE MOTOR

Bearing defects occur at characteristic frequencies, some of which are associated with its geometric parameters, while others are purely random. High frequency bearing vibrations are irregular or random in nature. The dependences of such vibrations obtained using a vibration analyzer make it possible to evaluate fluctuations or "jumps" in amplitude and frequency. Such phenomena can be explained by examining the forces generated by bearing defects.

A defective bearing can generate vibration of different types of frequencies, including at natural frequencies [16]. Impact interaction between bodies and raceways of a bearing excites vibrations of machine elements and bearing elements at natural frequencies. Each element during shock exposure is excited at its own frequency. Bearing defects act by shock pulses on various parts of the bearing, causing them to vibrate in their own modes of vibration.

The manifestation of the eigenfrequencies of the bearing elements is closely related to rotor frequencies. But unlike frequencies that are multiples of the rotational frequency, vibration at natural frequencies is almost always generated by several different bearing elements that generate several different frequencies of different amplitudes. Bearing defects can be divided into several types, depending on the type of wear [17]:

defect of the cage of bearing

$$F = \frac{D_i}{D_i + D_o} \times RPM;$$

defect of the ball

$$F = \frac{D_o}{D_b} \times \frac{D_i}{D_i + D_o} \times RPM;$$

defect of the inner ring

$$F = \frac{D_o}{D_i + D_o} \times M \times RPM;$$

outer ring defect

$$F = \frac{D_i}{D_i + D_o} \times M \times RPM.$$

Where: *Di* is the diameter of the inner ring, *Do* is the diameter of the outer ring, *Db* is the diameter of the ball, *M* is the number of rolling bodies, *RPM* is the shaft rotation frequency, *F* is the frequency of the defect.

The amplitude-frequency characteristics are obtained in the presence of various types of defects (5 classes). The initial sample is divided in a 60/40 ratio for each type of defect, respectively, into the training and verification ones: 10 different cases of presence of each of the defects were recorded, 6 of them were trained in the neural network, and 4 were tested. This data set allowed us to obtain more accurate results and avoid random coincidences. For the classification of these sequences of frequency and amplitude, a recurrent neural network with long short-term memory was used. To speed up learning on large data sets, we used the distribution of calculations and data among the processor cores and the graphics processor.

The LSTM performs additive interactions that can help improve the gradient flow over long sequences during training. Each defect is represented as an array of one line, in which the obtained values of the amplitudes are recorded in a row. The result is an even larger array of 50 such lines. During training, by default, the software breaks the training data into mini-lots and completes the sequences so that they have the same length. Too much filling can have a negative impact on network performance. In order to avoid adding too many indents during the training process, it is necessary to sort the training data by the length of the sequence so that they have the same length.The initial data of the frequency response array were sorted by the length of the sequence. A histogram of sorted data lengths is shown in the Fig. 5.

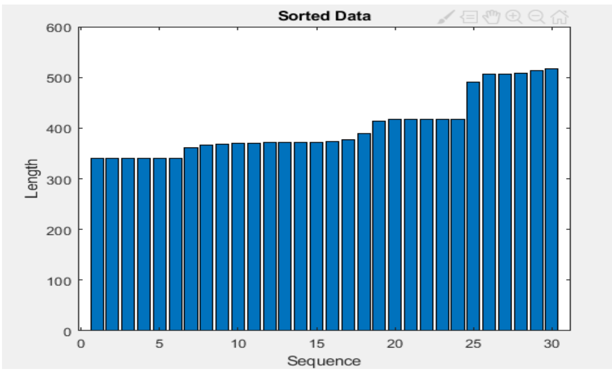

Fig. 5. The histogram of sorted data

The architecture of the BiLSTM is defined. The size of the input data is set, which will be a sequence of size 1 (the size of the input data by the number of features). The bidirectional layer

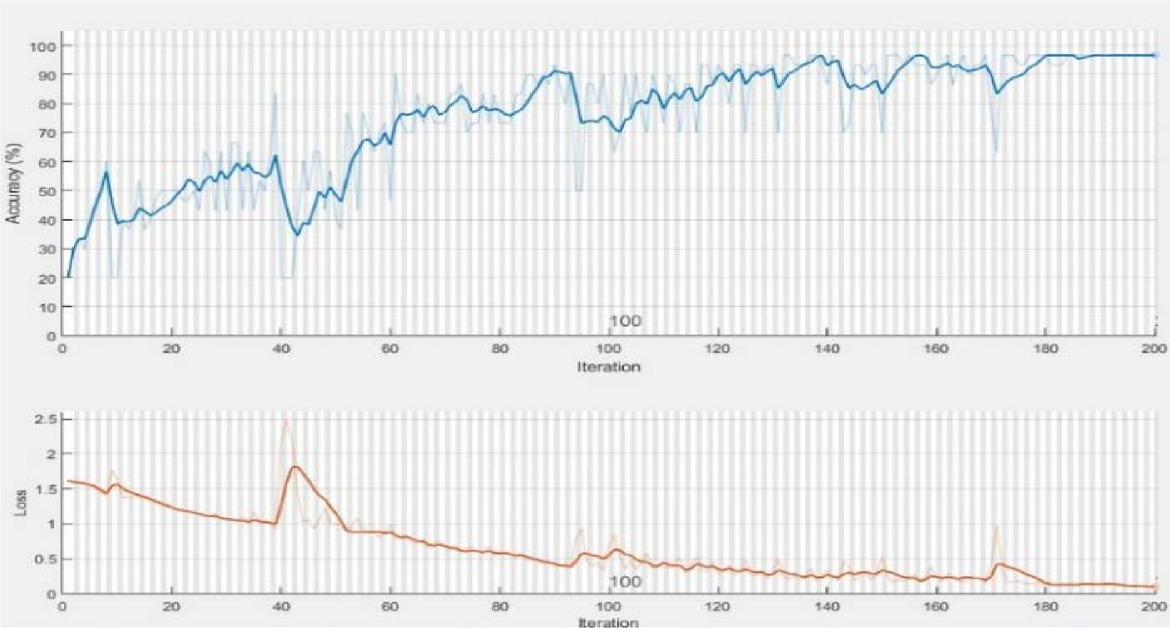

Fig. 6. Learning curve and loss-classification errors curve

of the BiLSTM with 100 hidden nodes and an out-put representing the last element of the sequence is specified. The network also includes a fully bonded size 4 layer, followed by a softmax layer and a classification layer

The neural network for detecting and classifying bearing defects as a result has 5 layers.

The first layer is the sequence input layer. The second Bi-Directional LSTM Level (BiLSTM) studies the bi-directional long-term relationships between time series time steps or sequence data.

These dependencies can be useful when it is necessary for the network to learn from the complete time series at each time step. The third fully connected layer multiplies the input signal with a weight matrix, and then adds a dis-placement vector. The fourth layer of softmax is an output block activation function. For classification tasks, the softmax layer should follow the fully connected layer. Finally, the fifth output layer is the classification layer, which computes the cross-entropy loss for multiclass classification problems with mutually exclusive classes. This loss is a measure of the discrepancy between the true value of the estimated parameter and the parameter estimate. The standard method for training neural networks is the method of stochastic gradient descent (SGD). However, it can diverge or converge very slowly if the training step is not tuned accurately enough. Therefore, there are many alternative methods to accelerate the convergence of learning and save the user from the need to carefully configure hyperparameters.

These methods often calculate gradients more efficiently and adaptively change the iteration step. One such method is the adaptive inertia method (Adam). The Adam solver was used as a learning function, the gradient threshold was set to 1, and the maximum number of epochs was 200. The sequence was added to the maximum length. To keep the data sorted by the length of the sequence, their mixing is prohibited.Neural network training completed. In the learning process, a so-called learning curve is obtained (expresses the accuracy of the network), which tends to 100% with an increase in the number of iterations. A curve of the number of loss-classification errors is also obtained. Graphs of both curves are shown in the fig. 6.

From the graph in figure 6 it can be seen that after about the 180th iteration, accuracy has an unchanged value, which suggests that 200 iterations are enough for learning. Thus, the neural network is trained at 60% of the sample.

To determine the classification accuracy value, 40% of the sample was used. Sequence control data is loaded as an *XTest* array. The *YTest* categorical tag vector contains 4 instances of each defect.The neural network model was trained using mini-series of sequences of the same length. The control data must be organized in the same way. Sorted test data by sequence length. The classification of defects on the test data set and the calculation of the classification accuracy value are performed. *YPred* is a categorical label vector predicted by the neural network model. By comparing the *YPred* vector with the *YTest* vector, the accuracy of neural network classification is calculated. After obtaining the classification accuracy values, the values of *YPred* tags are analyzed. Their comparison with the categorical vector of labels *YTest*, which was set during the verification of the neural network, showed which defects were identified correctly, and which were difficult to identify. The table 1 shows the results of comparing the given control values *YTest* and the obtained *YPred*. The table 1 shows that 14 out of 20 *YPred* values coincide with the *YTest* control values, which ultimately gives about 70% accuracy.

The defect of the separator is determined with 100% accuracy, the defect of the ball with 75% accuracy, and the defect of the inner ring and the outer with 50% accuracy. The results for defects of the separator and the ball can be considered satisfactory, since in most cases the predicted data are the same. 50% coincidence of defects of the outer and inner rings can be explained by the similarity of the types of these defects, which is also reflected in the nature of the noise introduced into the signal by these defects.

Table 1

Accuracy of initial bearing state diagnostic model

№	*YTest*	*YPred*
1	1	1
2	1	1
3	1	1
4	1	1
5	2	2
6	2	2
7	2	3
8	2	3
9	3	3
10	3	2
11	3	3
12	3	3
13	4	3
14	4	4
15	4	4
16	4	4
17	5	2
18	5	5
19	5	5
20	5	5

As a result, it was found that the designed bidirectional recurrent neural network with long short-term memory is a classifier capable of determining the presence and class of a bearing defect by its amplitude-frequency characteristics of vibration signals with an accuracy of about 70%.

It is noted that the greatest probability of a diagnostic model error occurs when determining whether a defect belongs to the inner or outer bearing ring. With this in mind, when introducing a diagnostic model into the machine control and monitoring system, it was decided to form a single defect class - “defect of the inner or outer bearing ring”. Combining these defects into a single class is possible because from the technological point of view, the consequences of the transition from a pre-defect state to an accident are the same - destruction of the bearing body. The combination of these two classes of defects in 1 made it possible to increase the accuracy of the model by 10%.

The result of the development and subsequent optimization of the structure of the diagnostic model of bearings is presented in the table 2.

Table 2

Accuracy of bearing state diagnostic models after optimization

Iteration	*Defect classes count*	*Accuracy*
1	5	75%
40	5	75.4%
43	4	85.7%
102	4	91.9%

CONCLUSION

The approach for diagnosing the condition of a nodes of a metal cutting machine presented in the article allows one to determine node state by indirect indicators obtained on the basis of information from vibration and force sensors. The article describes the architecture of the proposed diagnostic model based on the bidirectional recurrent neural networks with long short-term memory. It is showed practical examples of application bidirectional LSTM for estimate the expected wear of the cutting tool and define electrical motor bearing state. The use of such diagnostic systems as part of machine control and monitoring systems will optimize technological processing and reduce the duration of downtime.

REFERENCES

1. **Mohammadi N., Taylor J. E.** Smart city digital twins // 2017 IEEE Symposium Series on Computational Intelligence (SSCI). 2017. **[** N. Mohammadi, J. E. Taylor, "Smart city digital twins", in *IEEE Symposium Series on Computational Intelligence (SSCI)*, 2017. **]**

2. **Minaev V.** Maintenance and repair strategies // Economics and life. 2013. **[** V. Minaev, "Maintenance and repair strategies" in *Economics and life*. 2013. **]**

3. **Martínez-Arellano G., Terrazas G., Ratchev S.** Tool wear classification using time series imaging and deep learning // Int. J. Adv. Manuf. Technol. 2019. 104. Pp. 3647-3662. DOI: 10.1007/s00170-019-04090-6. **[** G. Martínez-Arellano, G. Terrazas, S. Ratchev "Tool wear classification using time series imaging and deep learning", in *Int. J. Adv. Manuf. Technol.* 104, Pp. 3647-3662, 2019. DOI: 10.1007/s00170-019-04090-6. **]**

4. **Zhou J., Zhao X., Gao J.** Tool remaining useful life prediction method based on LSTM under variable working conditions // Int. J. Adv Manuf. Technol. 2019. 104. Pp. 4715-4726. DOI: 10.1007/s00170-019-04349-y. **[** J. Zhou, X. Zhao, J. Gao "Tool remaining useful life prediction method based on LSTM under variable working conditions", in *Int. J. Adv Manuf. Technol*. 104, Pp. 4715-4726, 2019. DOI: 10.1007/s00170-019-04349-y. **]**

5. **Remaining** Life Prediction Method for Rolling Bearing Based on the Long Short-Term Memory Network / F. Wang, *et al* // Neural Process Lett. 2019. 50, Pp. 2437–2454. DOI: 10.1007/s11063-019-10016-w. **[** F. Wang, *et al .,* "Remaining Life Prediction Method for Rolling Bearing Based on the Long Short-Term Memory Network", in *Neural Process Lett.* 50, Pp. 2437-2454. 2019. DOI: 10.1007/s11063-019-10016-w. **]**

6. **Ambadekar P. K., Choudhari C. M.** CNN based tool monitoring system to predict life of cutting tool // SN Appl. Sci. 2020. 2. 860. DOI: 10.1007/s42452-020-2598-2. **[** P. K. Ambadekar, C. M. Choudhari, "CNN based tool monitoring system to predict life of cutting tool", in *SN Appl. Sci.* 2. 860. 2020. DOI: 10.1007/s42452-020-2598-2. **]**

7. **Review** of tool condition monitoring in machining and opportunities for deep learning / G. Serin et al // Int J. Adv. Manuf. Technol. 2020. 109. Pp. 953-974. DOI: 10.1007/s00170-020-05449-w. **[** G. Serin, *et al.,* "Review of tool condition monitoring in machining and opportunities for deep learning", in *Int J. Adv. Manuf. Technol.* 109. Pp. 953-974. 2020. DOI: 10.1007/s00170-020-05449-w. **]**

8. **Fuzzy** Neural Network Modelling for Tool Wear Estimation in Dry Milling Operatio / X. Li, et al. // Annual Conference of the Prognostics and Health Management Society. 2009. **[** X. Li, et al.**,** "Fuzzy Neural Network Modelling for Tool Wear Estimation in Dry Milling Operatio" in *Annual Conference of the Prognostics and Health Management Society*. 2009. **]**

9. **Hochreiter S., Schmidhuber J.** Long short-term memory // Neural Computation. 1997. Vol. 9, no. 8 Pp. 1735-1780. DOI: 10.1162/neco.1997.9.8.1735. **[** S. Hochreiter, J. Schmidhuber, "Long short-term memory", in *Neural Computation*. Vol. 9, no. 8, Pp. 1735-1780. 1997. DOI: 10.1162/neco.1997.9.8.1735. **]**

10. **Graves A., Schmidhuber J.,** Framewise phoneme classification with bidirectional LSTM and other neural network architectures // Neural Networks. 2005. No. 18, Pp. 602-610. **[** A. Graves, J. Schmidhuber, "Framewise phoneme classification with bidirectional LSTM and other neural network architectures", in *Neural Networks*. No. 18, Pp. 602-610 **.** 2005. **]**

11. **Datum** unit optimization for robustness of a journal bearing diagnosis system / B. C. Jeon, et al. // Int. J. Precis. Eng. Manuf. 2015. 16. Pp. 2411-2425. DOI: 10.1007/s12541-015-0311-y. **[** B. C. Jeon, et al. "Datum unit optimization for robustness of a journal bearing diagnosis system", in *Int. J. Precis. Eng. Manuf.* 16. Pp.2411-2425. 2015. DOI: 10.1007/s12541-015-0311-y. **]**

12. **Roders** Tech RFM760. [Электронный ресурс]. URL: http://www.tritechtooling.com/ equipment/ roeders-rfm760/ (дата обращения 20.09.2020). **[** Roders Tech RFM760. [Электронный ресурс]. URL: http://www.tritechtooling.com/ equipment/ roeders-rfm760/ (дата обращения 20.09.2020). **]**

13. **Leica** MZ12. [Электронный ресурс]. URL: http://www.lightglassoptics.com/ Leica-MZ125-Stereomicroscope-on-A-Stand_p_103. html (дата обращения 20.09.2020). **[** Leica MZ12. [Электронный ресурс]. URL: http://www.lightglassoptics.com/Leica-MZ125-Stereomicroscope-on-A-Stand_p_103.html (дата обращения 20.09.2020). **]**

14. **William S., Javad H.** Foundations of Material Science and Engineering (4th ed.). McGraw-Hill. P. 229. **[** S. William, H. Javad. In *Foundations of Material Science and Engineering (4th ed.). McGraw-Hill*. P. 229. **]**

15. **Keras**. [Электронный ресурс]. URL: https://keras.io/ (дата обращения 20.09.2020). **[** Keras. [Электронный ресурс]. URL: https://keras.io/ (дата обращения 20.09.2020). **]**

16. **Method** of Operational Diagnostics of Metal Cutting Machine Modules / K. A. Masalimov, et al // Proceedings - 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency, SUMMA 2019. **[** K. A. Masalimov, *et al***,** "Method of Operational Diagnostics of Metal Cutting Machine Modules", in *Proceedings 2019 1st International Conference on Control Systems, Mathematical Modelling, Automation and Energy Efficiency*, SUMMA. 2019. **]**

17. Detection of defects in rolling bearings. [Электронный ресурс]. URL: http://www.vibration.ru/obnar_defekt.shtml (дата обращения 20.09.2020).
[Detection of defects in rolling bearings. [Электронный ресурс]. URL: http://www.vibration.ru/obnar_defekt.shtml (дата обращения 20.09.2020). **]**

ABOUT AUTHORS

MASALIMOV, Kamil Adipovich, Postgrad. (PhD) Student, Dept. of Industrial Automation, Institute of Aerospace Technology and Material Science of Ufa State Aviation Technical University.
MUNASYPOV, Rustem Anvarovich, Prof., Head of Dept. of Industrial Automation, Institute of Aerospace Technology and Material Science of Ufa State Aviation Technical University.

МЕТАДАННЫЕ

Заголовок: Применение нейронных сетей LSTM для диагностики состояния узлов металлорежущих станков.

Авторы: К. А. Масалимов [1], Р. А. Мунасыпов [2]

Принадлежность: Уфимский государственный авиационный технический университет, г. Уфа, Россия.

Почта: [1]masalimov.k.a@gmail.com, [2]rust40@mail.ru.

Язык: Английский.

Источник: СИИТ (научный журнал Уфимского государственного авиационного технического университета), т. 3, № 2 (6), стр. 26–34, 2021. ISSN 2686-7044 (онлайн), ISSN 2658-5014 (печатный вариант).

Аннотация. Представлено описание подхода к разработке прогностических диагностических моделей узлов металлорежущих станков с числовым программным управлением (ЧПУ). Этот подход основан на использовании нейронных сетей с двунаправленной долгосрочной краткосрочной памятью (BiLSTM). Описана архитектура таких нейронных сетей, метод предварительной обработки данных, записанных в процессе работы станка. Приведены примеры применения методики диагностики состояния режущего инструмента и подшипников электродвигателей на станке. Для оценки оставшегося срока службы режущего инструмента в предлагаемой модели BiLSTM используется косвенная информация – значения вибрации и динамометрия по трем осям. Представлено сравнение данных, полученных из диагностической модели для оценки максимального износа режущей кромки инструмента, с реальными данными для значений из тестового образца. Разработана диагностическая модель для оценки состояния подшипников электродвигателя по значениям вибрации, которая определяет наличие неисправности и одного из четырех классов – дефект сепаратора, шара, внутреннего кольца или наружное кольцо подшипника. Оптимизация модели классификатора проводилась с учетом специфики сигналов вибрации при наличии дефектов подшипников электродвигателя металлорежущего станка.

Ключевые слова: металлорежущая обработка; диагностический; состояние режущего инструмента; износ режущего инструмента; состояние подшипников; глубокие нейронные сети; долговременная кратковременная память; нейронные сети; Система наблюдения; прогностическая диагностика.

Об авторах:

МАСАЛИМОВ Камиль Адипович, аспирант кафедры промышленной автоматизации Института аэрокосмических технологий и материаловедения, Уфимский государственный авиационный технический университет.

МУНАСЫПОВ Рустем Анварович, проф., зав. каф. промышленной автоматизации Института аэрокосмических технологий и материаловедения Уфимский государственный авиационный технический университет.