IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 21, 2019, accepted February 5, 2019, date of publication February 13, 2019, date of current version March 5, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2899074

Genetic Optimization Method of Pantograph and Catenary Comprehensive Monitor Status Prediction Model Based on Adadelta Deep Neural Network

ZHIJIAN QU [1], SHENGAO YUAN [1], RUI CHI [1], LIUCHEN CHANG [2], AND LIANG ZHAO [3]

[1] School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China
[2] School of Electrical and Computer Engineering, University of New Brunswick, Fredericton, NB E3B 5A3, Canada
[3] Changzhou Rail Transit Development Co., Ltd., Changzhou 213000, China

Corresponding author: Zhijian Qu (08117324@bjtu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 51567008 and Grant 51867009, in part by the Foundation Plan for Distinguished Young Scholars in Jiangxi Province under Grant 20162BCB23045, in part by the Natural Science Foundation of Jiangxi Province under Grant 20171BAB206044, and in part by the Applied and Cultivation Program of Science and Technology Department of Jiangxi Province under Grant 20181BBE58010.

ABSTRACT The status of the Pantograph and Catenary is the guarantee for the safe operation of the railway. However, the traditional Pantograph and Catenary status judgment efficiency is not satisfactory, and it is impossible to timely repair the catenary, which may lead to the greater economic loss. In this paper, a new GA-ADNN-based (genetic algorithm-Adadelta deep neural network-based) optimization method for the prediction model for catenary comprehensive pantograph and catenary monitor (CPCM) status is proposed. According to the status values of the CPCM parameters such as height, stagger, hard point, contact force, and height difference within span, the status value of the pillars in the catenary has been calculated by the analytic hierarchy process, and then the prediction model for predicting catenary CPCM status has been established and then optimized by genetic algorithm to avoid prediction model falling into local optimum. Finally, the CPCM test parameters of each pillar of the catenary in the actual example are input and the CPCM status value of the corresponding pillar is predicted. With the smallest prediction error found, the genetic algorithm is used for optimization, the optimal learning rate of the prediction model is 0.0559, and the optimal number of the hidden layer of the CPCM status prediction model is determined to be 14. The experimental results show the feasibility of GA-ADNN-based prediction model for predicting the catenary CPCM status, and that compared with the support vector machine and traditional artificial neural network prediction methods, the GA-ADNN-based prediction model has higher prediction precision and better generalization ability.

INDEX TERMS CPCM, analytic hierarchy process, catenary, genetic algorithm, GA-ADNN.

I. INTRODUCTION

As the important traction power equipment for electrified high-speed railway transportation system, the catenary in sound status is a guarantee of stable current collection and on-time running of high-speed EMUs (Electric motor train Units) [1], [2]. In order to ensure the normal operation of the catenary, overhaul is necessary. However, the traditional overhaul is time consuming with very low work efficiency as it generally takes half a month or one month and many labor forces. The 6C system (detection and monitoring system of the pantograph-catenary in railway) detects the relevant parameters of the catenary and so provides data basis forthe maintenance of the catenary status [3]–[5]. The detected data are mainly divided into two categories: numerical data and image data. The image data focus on the identification of abnormalities and faults. At present, limited by the less obvious image anomaly characteristics and the insufficient training samples in library, manual work is required for checking images in the dispatching data center. Further, the CPCM

The associate editor coordinating the review of this manuscript and approving it for publication was Luca Chiaraviglio.

(Comprehensive Pantograph and Catenary Monitor) has large sample size in numerical data, satisfying the requirement of data amount for prediction model training, and therefore, current researches of catenary status prediction focus on the operating status of the pantograph and catenary. CPCM detection is an important part of the 6C system, and the failure of the pantograph and catenary, as one of the most common faults in railway power supply system, has the characteristics of long power outage, wide range of damage and great difficulty in repair. For example, on March 10-17, 2015, a number of consecutive pantograph and catenary failures incurred in the Wuhan-Chengdu section of the Shanghai-Wuhan-Chengdu high-speed railway, resulted in damage by striking to the carbon contact strip of 14 CRH trains. Therefore, how to accurately and efficiently predict the status of the pantograph and catenary based on the CPCM detecting parameters is the basis for the decision making for catenary maintenance and overhaul.

There are two main methods for determining the operating status of the pantograph and catenary system: the first is to arrange the staff to determine the CPCM status of all catenary pillars of the entire line according to each catenary parameter. This method requires great workload, but is prone to human error, and so the work efficiency is not satisfying; and the second is to conduct the predictive status identification with the machine learning technology [6]–[9]: firstly, determine the status values of the pillars numbered with an even number multiplied with 2 in the catenary, and then establish a training model to predict CPCM status values of the remaining pillars numbered with an odd number multiplied with 2. According to the sample sizes they can process, status prediction models are divided into two main categories: the first one is the support vector machine-based (SVM) prediction model. In literature [10], the SVM model is used to predict the transient stability status after faults. This method is mainly for small sample size data prediction, and it is difficult to meet the prediction accuracy requirements of a large number of pillar samples in the catenary. The second category is the status prediction model for processing large samples. It is mainly divided into shallow network model and deep neural network model. The shallow model mainly includes random forest [11], Bayesian model, ensemble model [12] and artificial neural network, etc, literature [13] proposed a driving risk status prediction method based on Bayesian network prediction model; in literature [14], neural network model is used to predict the nitrogen status of wheat plants. Compared with the conventional gray-world and scale-by-max approaches, the performance of neural network model is better. But the hidden layer of the shallow model is rarely able to dig out more essential features of the data and the processing effect on the complex function relationship is not ideal, and sometimes even produces a large prediction error. Deep neural network is another direction of machine learning [15]–[16]. According to the type of data processed, deep neural networks are mainly divided into two categories: image type and numerical type. The convolutional neural network model is a deep neural network model for processing image data [17]. For the processing of numerical models, there are mainly deep feedforward neural networks, memory neural networks and recurrent neural networks. In the literature [18], the recurrent neural network model is used to predict the health status of the hard disk drive, compared to the simple fault prediction method, the prediction performance of the recurrent neural network model is better, but the model mainly deals with the time series cyclic data; the literature [19] proposes a method based on the long and short time memory neural network model to predict the operating state of the equipment. The experimental results show that the error is small, thus the method is suitable for analysis of time sequence data. In the field of catenary, the research on status prediction is still not reported in the literature. Considering that the data of the comprehensive detection of the catenary is non-time series numerical data, a catenary CPCM status prediction model based on Adadelta deep feedforward neural network is proposed. Compared with the traditional neural network model, the catenary CPCM status prediction model has a hidden layer of nonlinear transformation, which can handle more complex functional relationships. The global optimization ability of genetic algorithms is better than other evolutionary algorithms[20]–[22]. Therefore, in order to avoid the prediction model falling into local optimum, the global search method of genetic algorithm is used to optimize the catenary CPCM status prediction model to achieve higher prediction accuracy.

Aiming at the problem that the traditional Pantograph and Catenary status judgment efficiency is not satisfactory, which leads to be unable to perform targeted maintenance in time, this paper proposes a new GA-ADNN-based prediction optimization method for deep neural network for determining the CPCM status. Firstly, the status value of the pillars numbered with an even number multiplied with 2 in the catenary is worked out by the eigenvalue analysis method, and then catenary status prediction model built based on the deep neural network model is trained and optimized, and finally each parameter of the remaining pillars numbered with an odd number multiplied with 2 in the catenary is input to the model to predict the status value of the those pillars, and the GA-ADNN prediction optimization is verified by the comparison experiments between various prediction models based on the actual engineering example data.

II. OPERATING STATUS OF PANTOGRAPH AND CATENARY SYSTEM

A. LEVEL OF OPERATING STATUS OF PANTOGRAPH AND CATENARY SYSTEM

The catenary is one of the important facilities of the railway transportation system and its pantograph and catenary system is the guarantee for the safe and stable operation of the railway. The schematic diagram of the catenary is shown in Fig.1.

As shown in Fig.1, the status of the catenary includes the height, the stagger, the hard point, the contact voltage and

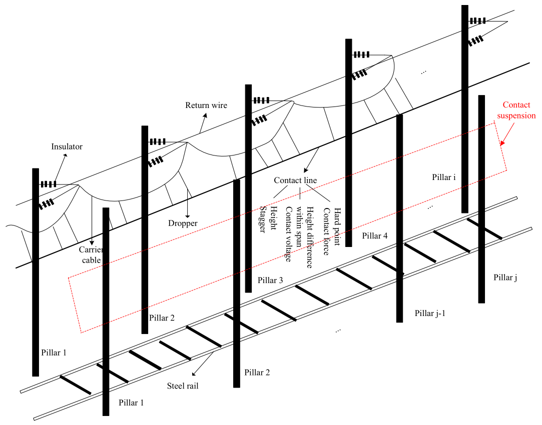


FIGURE 1. Main status parameters of pantograph and catenary system.

the height difference within span. As these parameters are related to each other and interact with each other, they should be taken into account in a comprehensive manner for the calculation of the status value of the pantograph and catenary system. The operating status of the pantograph and catenary system is mainly divided into five grades : excellent, good, average, minor fault and fault, corresponding to values 1, 0.8, 0.6, 0.4 and 0.2. As for other grades : 0.9, 0.7, 0.5, 0.3 and 0.1, the status is between the above main grade. ''Excellent'' refers to that the status of the pantograph and catenary system is very good, and so no maintenance is needed; ''Good''refers to that the status is good, and so no maintenance but precaution is needed; ''Average'' refers to that the status is not very good, so proper maintenance is needed; ''Minor fault'' refers to that there is a slight fault in the pantograph and catenary system, and maintenance and minor overhaul are needed; ''Fault'' refers to that the pantograph and catenary system is in fault and overhaul is needed.

B. ANALYSIS AND CALCULATION OF THE STATUS OF THE PANTOGRAPH AND CATENARY SYSTEM BY ANALYTIC HIERARCHY PROCESS

AHP (Analytic hierarchy process) is a qualitative and quantitative analysis method for determining the weight of each indicator in decision-making according to the importance and influence of each indicator on the target[23]–[24]. By comparing the indicators in pairs, the judgment matrix A is constructed based on the relative order of excellence of each index.

$$A = \begin{bmatrix} 1 & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & 1 & a_{23} & \dots & a_{2n} \\ a_{31} & a_{32} & 1 & \dots & a_{3n} \\ \dots & \dots & \dots & \dots & .. \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{bmatrix}$$

The above A indicates the judgment matrix, where a_{ij} is the result of importance comparison between Factor i with Factor j; and aij = 1/aij. The importance comparison between the factors is shown in Tab. 1.

TABLE 1. Comparison of importance of different factors.

a_{ij}	Importance comparison between Factor i with Factor j
1	Equally important
3	i slightly more important than j
5	i more important than j
2,4,6	Median of comparison results

TABLE 2. Random consistency indicator of matrix of each order.

n	1	2	3	4	5	6	7	8	9
RI	0	0	0.52	0.89	1.12	1.26	1.36	1.41	1.46

Following the construction of the judgment matrix, it is to judge whether it satisfies the consistency condition or indicator [25], and then check the degree of consistency. The consistency indicator expression is as shown in formula (1).

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{1}$$

where, $\lambda_{\max}$ is the maximum eigenvalue of the judgment matrix A, n is the order of the judgment matrix A.

In order to measure the size of CI, a random consistency indicator(RI)is introduced. The random consistency standard value of each order matrix, as shown in Table2.

When it is checked whether the judgment matrix satisfies the consistency, the consistency ratio (CR) needs to be calculated. If $CR < 0.1$, the matrix is judged to be consistent, otherwise it is not satisfied. The expression of the consistency ratio (CR) is as shown in formula (2).

$$CR = \frac{CI}{RI} \tag{2}$$

where, CI is the consistency indicator of the judgment matrix A, RI is the random consistency indicator of the judgment matrix A.

The calculation of the status of the pantograph and catenary system incorporates the parameters such as height, stagger and hard point of contact line, contact voltage, height difference within span and contact force, and according to the Procedures of Overhaul of the Catenary and engineering empirical value, the level of each parameter status determined is shown in Tab. 3.

As having different failure probability and failure significance, each parameter has different weights. And such weights of the parameters in the pantograph and catenary system are defined by the analytic hierarchy process. Taking the detection data of the uplink catenary of the Quzhou-Jiujiang railway as an example and referring to the status values of the parameters in Table 3, the height is easier to go wrong than the height difference within span, the height difference within span is easier to malfunction than the stagger, and the contact

TABLE 3. Status values of each parameter of the pantograph and catenary system. (a) Main status values of each parameter. (b) Other status values of each parameter.

(a)

Indicator/Grade	Excellent (1)	Good (0.8)	Average (0.6)	Poor (0.4)	Extremely poor (0.2)
Height of contact line/mm	6000	6000±30	6000±60	6000±100	6000±150
The absolute of stagger of contact line/mm	300	320	340	380	400
Height difference within span/mm	0	25	50	100	150
Hard point/mm	30	40	50	100	150
Contact force/N	120	120±30	120±60	120±80	120±100
Contact voltage/kV	25	26.5	27.5	28.5	29.5

(b)

Indicator/Grade	0.9	0.7	0.5	0.3	0.1
Height of contact line/mm	(6000,6030) or (5070,6000)	(6030,6060) or (5040,5070)	(6060,6100) or (5000,5040)	(6100,6150) or (4850,5000)	>6150 or <4850
The absolute of the stagger of contact line/mm	(300,320)	(320,340)	(340, 380)	(380,400)	>400
Height difference within span/mm	(0,25)	(25,50)	(50,100)	(100,150)	>150
Hard point/mm	(30,40)	(40,50)	(50,100)	(100,150)	>150
Contact force/N	(120,150) or (120,150)	(150,180) or (120,150)	(180,200) or (120,150)	(200,220) or (120,150)	>220 or <20
Contact voltage/kV	(25,26.5)	(26.5,27.5)	(27.5,28.5)	(28.5,29.5)	>29.5

voltage is easier to fail than contact force and hard point, but the contact force has a slightly greater influence on the CPCM status than the contact voltage. So from Tab.1 we can obtain that $a_{11} = 1$; $a_{12} = 2$; $a_{13} = 2$; $a_{14} = 3$; $a_{15} = 4$; $a_{16} = 5$, so its judgment matrix A is as follows:

$$A = \begin{bmatrix} 1 & 2 & 2 & 3 & 4 & 5 \\ 1/2 & 1 & 1 & 2/3 & 1/2 & 2/5 \\ 1/2 & 1 & 1 & 2/3 & 1/2 & 2/5 \\ 1/3 & 3/2 & 3/2 & 1 & 3/4 & 3/5 \\ 1/4 & 2 & 2 & 4/3 & 1 & 4/5 \\ 1/5 & 5/2 & 5/2 & 5/3 & 5/4 & 1 \end{bmatrix}$$

According to the calculations, in the matrix A, the maximum eigenvalue λ_{max} is 6.4001, the normalized eigenvector W_1 corresponding to the maximum eigenvalue λ_{max} is $W_1 = [0.392, 0.093, 0.093, 0.114, 0.140, 0.168]$, and the consistency indicator CI is 0.08.

The consistency ratio (CR) is calculated as:

$$CR = \frac{CI}{RI} = 0.063 < 0.1,$$

indicating that the consistency meets the requirements. Based on the normalized eigenvector corresponding to the maximum eigenvalue of the judgment matrix by analytic hierarchy process as the weight of each factor, the weight of each parameter of the catenary is calculated as W = [0.392, 0.093, 0.093, 0.114, 0.140, 0.168]. Therefore, the CPCM status value of each pillar can be calculated according to the status value and weight of each parameter.

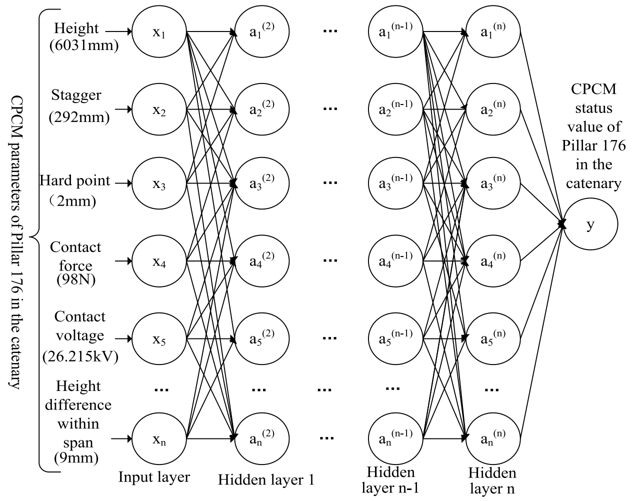


FIGURE 2. Prediction model for CPCM status.

III. DEEP FEEDFORWARD NEURAL NETWORK MODEL FOR CPCM STATUS PREDICTION

A. STRUCTURE OF PREDICTION MODEL FOR PANTOGRAPH AND CATENARY SYSTEM OPERATING STATUS

The prediction model for pantograph and catenary system status is mainly divided into three parts: input layer, hidden layer and output layer. The training model realizes the fitting of the predicted value and the actual value of the catenary CPCM status via the hidden layer and finally predicts the catenary CPCM status. A large number of hidden layers in the prediction model perform multiple nonlinear transformations, feature extraction and learning on the input data to obtain more essential features of the data. In the prediction model, the height, stagger, hard point, the contact force,the contact voltage and height difference are the input data, and the CPCM status value is output data. The hidden layer in the middle has a large number of neurons. Taking pillar No. 176 as an example, the schematic diagram of the CPCM status prediction model is as shown in Fig.2.

The functions corresponding to the CPCM status prediction model are as shown in formula (3):

$$\begin{cases} a_1^{(2)} = f(W_{11}^{(1)}x_1 + W_{12}^{(1)}x_2 + \ldots + W_{1n}^{(1)}x_n + b_1^{(1)}) \\ a_2^{(2)} = f(W_{21}^{(1)}x_1 + W_{22}^{(1)}x_2 + \ldots + W_{2n}^{(1)}x_n + b_2^{(1)}) \\ \ldots \\ a_n^{(2)} = f(W_{n1}^{(1)}x_1 + W_{n2}^{(1)}x_2 + \ldots + W_{nn}^{(1)}x_n + b_n^{(1)}) \\ \ldots \\ y = f(W_{1n}^{(n-1)}a_1^{(n-1)} + \ldots + W_{nn}^{(n-1)}a_n^{(n-1)} + b_n^{(n-1)}) \end{cases} \tag{3}$$

where, y indicates the predicted value of the CPCM status; W the weight between the neurons of each layer; x_1 the height of the CPCM status; x_2 is the stagger of the CPCM status; x_n the height difference within span; f the activation function of the model; and b the threshold.

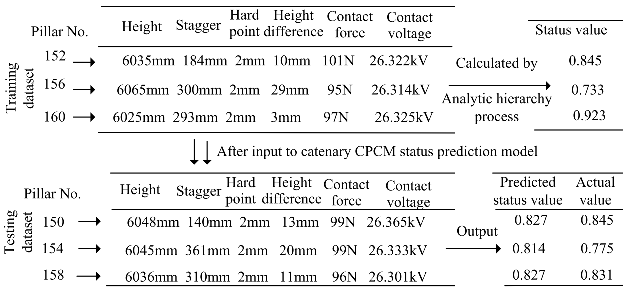


FIGURE 3. Training dataset and Testing dataset of CPCM status prediction model.

Since the catenary is generally divided into uplink and downlink, the pillar numbers of the catenary are even numbers, and the pillar numbers of the catenary are all odd. This paper takes the monitoring data of the uplink catenary of the Quzhou-Jiujiang line as an example, and the pillar numbers of the uplink catenary are 2, 4, 6, 8, 10, 12..., of which 2, 6, 10... is the odd multiple of 2, and 4, 8, 12... is an even multiple of 2. For example, for the pillars No. 150-160, the training dataset and testing dataset of data of each parameter and pillar status in the model are shown in Fig.3.

As shown in Fig. 3, the parameter data such as the height, stagger, hard point and contact voltage of the pillars No. 152, 156 and 160 in the catenary are taken as the input data set for training and optimization, and then the data of pillars No. 150, 154 and 158 are input to predict the status value of the corresponding pillar.

B. ACTIVATION FUNCTION AND LOSS FUNCTION FOR CPCM STATUS PREDICTION MODEL

The activation function and the loss function are two important parameters of the CPCM status prediction model. Compared with the Sigmoid function, the Tanh function has an output mean of 0 and fewer iterations, and its convergence speed is faster than that of the Sigmoid function. The expression of the Tanh activation function is shown in formulas (4) and (5).

$$\tanh(x) = \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{4}$$

$$\tanh(x) = 2sigmoid(2x) - 1 \tag{5}$$

By substituting with the activation function $\tanh(x)$ in formula (3), a functional expression of the CPCM status prediction model is obtained, as shown in formula (6).

$$\mathrm{y} = \mathrm{tanh(x)}(W_{1\mathrm{n}}^{(\mathrm{n-1})}\mathrm{a}_1^{(\mathrm{n-1})} + \ldots + W_{\mathrm{nn}}^{(\mathrm{n-1})}\mathrm{a}_\mathrm{n}^{(\mathrm{n-1})} + \mathrm{b}_\mathrm{n}^{(\mathrm{n-1})}) \tag{6}$$

Since the crossing entropy loss function is generally used for solving classification problems and the quadratic cost function is suitable for prediction problems, the loss function of the prediction model is a quadratic cost function, as shown in formula (7).

$$= \frac{1}{2n}\sum_x \left\| y(x) - a^L(x) \right\|^2 \tag{7}$$

where, n is the number of samples; x is the input sample of CPCM status parameter; $\mathrm{a}^L(x)$ is the output value of CPCM status; and L is the number of layers.

C. TRAINING OF PARAMETERS OF CATENARY CPCM STATUS PREDICTION MODEL

In order to reduce the prediction error of the prediction model, the training is needed to update the model parameters. While the traditional gradient descent method first derivates the training parameters of the loss function of all samples in the training set and calculates all derivative sums, and then updates all the training parameters, the stochastic gradient descent method is to derivate the loss function of each small batch of samples[26], and then sum up the derivatives and update the parameters. Each update by the stochastic gradient method may not proceed in the correct direction, which may cause fluctuations. However, the fluctuation may jump from the current local minimum to another better local extremum, and the error curved surface of the deep neural network is a non-convex function that eventually converges to a global extremum point. The stochastic gradient calculation process is as follows:

$$J(\theta) = \frac{1}{2}\sum_{i=1}^{m}(Y_\theta(x^{(i)} - y^{(i)})^2 \tag{8}$$

$$\theta_\mathrm{t} = \theta_{\mathrm{t}1} - \alpha\frac{\partial}{\partial\theta}J(\theta) \tag{9}$$

$$\begin{cases} \frac{\partial}{\partial\theta}J(\theta) = \frac{\partial}{\partial\theta}\frac{1}{2}(Y_\theta(x) - y)^2 \\ = (Y_\theta(x) - y)\frac{\partial}{\partial\theta}(Y_\theta(x) - y) \\ = (Y_\theta(x) - y)x_i, (m = 1) \end{cases} \tag{10}$$

$$\theta_\mathrm{t} = \theta_{\mathrm{t}1} - \alpha(Y_\theta(x^{(j)} - y^{(j)})x_{t-1}^{(j)}(1 \le j \le m) \tag{11}$$

where, θ_t is the parameter of the catenary CPCM status prediction model; $J(\theta)$ the loss function; $Y_\theta(x)$ the predicted value of the CPCM status; $y$the actual value of the CPCM status; m the number of training samples; and α the learning rate of CPCM status prediction model.

The stochastic gradient descent method minimizes the value of $J(\theta)$ by continuously adjusting the parameters of the CPCM status prediction model. First, the partial derivative of $J(\theta)$ is worked out so to determine the direction in which the gradient decreases the fastest, and then the parameters of the prediction model are updated in this direction. Since the stochastic gradient descent algorithm updates the parameters each time, instead of working on all the training sets, and only a parameter for updating the prediction model in batch is utilize, the training time of the catenary prediction model is reduced.

IV. REALIZATION OF OPTIMIZING CATENARY CPCM STATUS PREDICTION MODEL

A. ADADELTA-BASED OPTIMIZATION FOR STOCHASTIC GRADIENT DESCENT METHOD

Since the stochastic gradient descent algorithm requires manual setting of learning rate, the improper learning rate selected may lead to low prediction precision. However, The Adadelta (adaptive learning rate algorithm) is able to automatically adjust the learning rate as an optimization of the stochastic gradient descent algorithm, and improves the prediction precision. The Adadelta is calculated as follows:

$$g_t = \frac{\partial J(\theta)}{\partial(\theta)}, \quad \Delta\theta_t = \theta_t - \theta_{t-1} \tag{12}$$

$$n_t = mn_{t-1} + (1-m)g_t^2 \tag{13}$$

$$\Delta\theta_t = -\frac{\alpha}{\sqrt{n_t+\varepsilon}}g_t \tag{14}$$

$$E\left|g^2\right|_t = 0.5E\left|g^2\right|_{t-1} + 0.5g_t^2 \tag{15}$$

$$\Delta\theta_t = -\frac{\alpha}{\sqrt{E\left|g^2\right|t+\varepsilon}}g_t \tag{16}$$

where, g_t is the gradient of the stochastic gradient descent algorithm; m the coefficient; θ_t the updated parameter variation; α is the learning rate of the prediction model; $E|g^2|_t$ the expectation on the absolute squared value of the gradient; ε a constant to prevent the denominator from being zero.

B. GENETIC OPTIMIZATION OF THE CATENARY CPCM STATUS PREDICTION MODEL

In order to prevent the local optimum, optimization is performed by genetic algorithm. The learning rate is a key parameter of the prediction model as it determines the rate of the update of parameters such as the weight between neurons in the model. The learning rate α can be obtained by formula (9), as shown in formula (17).

$$\alpha=(\theta_{t1}-\theta_t)\frac{\partial(\theta)}{\partial J(\theta)} \tag{17}$$

where, θ_t is the parameter of the CPCM status prediction model at time t; $J(\theta)$ the error function of the prediction model; and α the learning rate of the model.

The learning rate α determines the speed at which the parameters of the model move to the optimal value. If the learning rate is too high, the optimal value may be exceeded; or if the learning rate is too low, the update parameter works inefficiently and so the parameters such as weights cannot be updated to their optimal value in a long time. Therefore, genetic algorithm, an algorithm of global searching of optimal solution that simulates the natural selection of biological evolution theory and the genetic mechanism, works to find the optimal learning rate of the catenary status prediction model, and the learning rate genetic optimization process of the prediction model is as shown in Fig. 4.

As shown in Fig. 4, the learning rate genetic optimization process is mainly divided as follows: first, encoding the model learning rates $\alpha1$, $\alpha2$, $\alpha3$ and $\alpha4$; then calculating the fitness of these learning rates, doing selection, crossover and mutation operations, and determining whether the learning rates satisfy the error requirement for the catenary status prediction; and finally decoding the new learning rates $\alpha1$, $\alpha2$, $\alpha3$ and $\alpha4$.

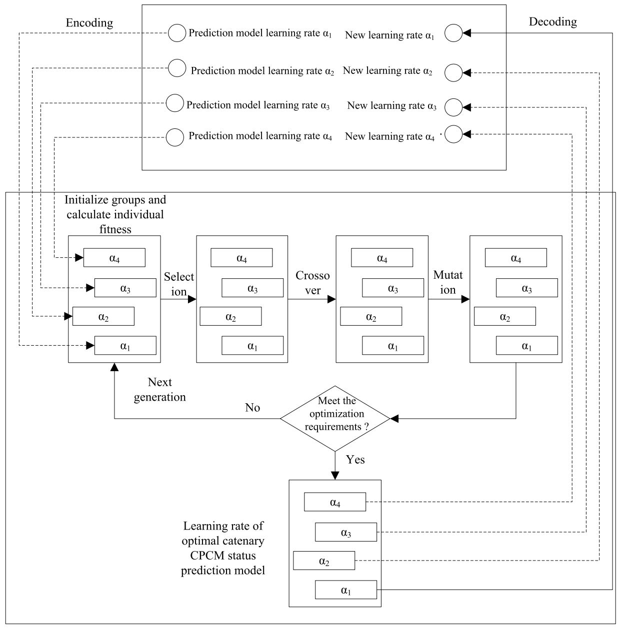


FIGURE 4. Genetic optimization of the learning rate of the catenary CPCM status prediction model.

The basic steps of the genetic algorithm to optimize the learning rate of the prediction model are as follows:

(1) Encoding learning rate

First, the learning rate parameter α of the spatial catenary prediction model is expressed as genotype string structure data of the genetic space, and the different combinations of these string structure data constitute different points.

(2) Initializing the learning rate group

N initialized string structure data such as $\alpha_1, \alpha_2, \alpha_3 ... \alpha_n$ are randomly generated, each datum indicating a learning rate and N data constituting a group, and they are used as initial points for the iteration in the genetic algorithm.

(3) Calculate the fitness value of each learning rate in the group

The fitness value of the new learning rate of the prediction model is calculated. The fitness function is the loss function of the catenary prediction model, and the fitness value is an indicator for evaluating the merits and demerits of the individual or solution.

(4) Selecting

It is to select a good learning rate from the current group of $\alpha_1, \alpha_2, \alpha_3 ... \alpha_n$, for probably reproduce offspring and generate new learning rates.

(5) Crossing

Crossing is the most important genetic operation in genetic algorithm. The new generation learning rate α_3 and α_9 can

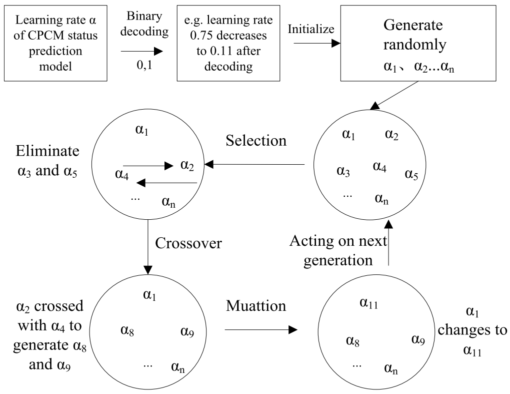


FIGURE 5. Main flow of the genetic algorithm for the catenary CPCM status prediction model.

be obtained by crossing α_2 and α_4. The new generation learning rate has combined the learning rate of the previous generation.

(6) Mutating

It is to randomly select a learning rate α_1 from the group of $\alpha_1, \alpha_2, \alpha_3 \ldots \alpha_n$, and randomly changes the value of each string structure data with a certain probability for the selected α_1, and then it is mutated to the new learning rate α_{11}. The main flow of the genetic algorithm for the catenary CPCM status prediction model is shown in Fig. 5.

(7) Judging the learning rate of prediction model

It is to judge whether the learning rate of the prediction model satisfies the optimization condition, if yes, proceeding from (3) and otherwise, proceeding from (7).

(8) Outputting the optimal learning rate of prediction model

After the operations of selecting, crossing and mutating, the learning rate α with optimal fitness in the prediction model is output.

C. IMPLEMENTATION PROCESS OF THE CATENARY CPCM STATUS PREDICTION MODEL OPTIMIZATION

The optimization process of the prediction model mainly involves the optimization of the stochastic gradient descent algorithm for model learning and the optimization of the prediction model parameters. Taking pillars 150, 152, 154... 160 as an example, the implementation flow of the optimization of the catenary CPCM status prediction model is shown in Figure. 6.

The optimization flow of prediction model for catenary CPCM status is as follows:

Step1: Input the CPCM status parameters such as height, stagger, height difference within span, hard point, contact force and contact voltage and corresponding status value of pillars No. 150, 152, 154... 160 to the model as its training samples. For example, the values of the parameters of pillar No. 152 are 6,035mm, 184mm and 2mm, 10mm, 101N and 26.322kV, and the CPCM state value is 0.845.

Step 2: Determine the structure of the CPCM status prediction model.

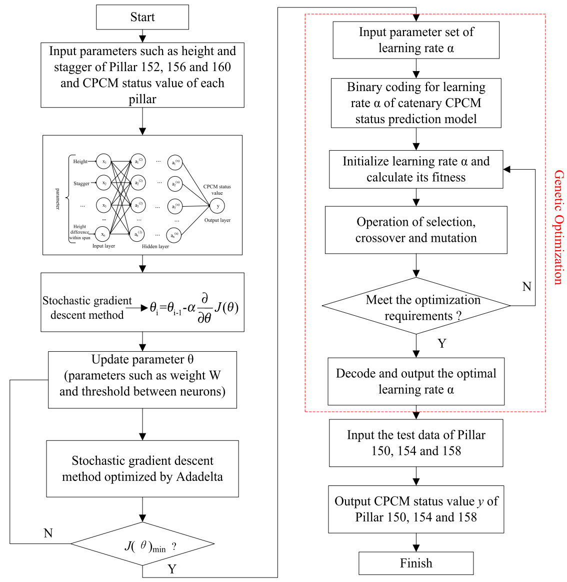


FIGURE 6. Optimization flow of prediction model for catenary CPCM status.

Step3: Use stochastic gradient descent algorithm to update the parameters such as weight W and threshold b between neurons; finally judge whether the parameter satisfies the condition of minimum loss function $J(\theta)$: if yes, the training is ended; otherwise, the training continues.

Step 4: Utilize the adaptive learning rate algorithm to automatically adjust the learning rate and optimize the stochastic gradient descent algorithm and finally improve the prediction precision of the prediction model as high as possible.

Step 5: Input the model learning rate α as the parameter to be optimized.

Step 6: Binary 01 encode the learning rate α of the mode.

Step 7: Initialize the learning rate to generate $\alpha_1, \alpha_2 \ldots \alpha_n$, calculate the fitness of each learning rate, and carry out selecting, crossing and mutating operations for $\alpha_1, \alpha_2 \ldots \alpha_n$ to generate a new optimized learning rate.

Step 8: Determine whether the learning rate satisfies the optimal condition: if yes, output the optimal value; otherwise, repeat from step 4.

Step 9: Input the CPCM status parameters such as the height, the stagger, the hard point, height difference within span, contact force and contact voltage of pillars No. 150, 154 and 158 of the catenary, for example: the value of those parameters of Pillar No. 158 is 6,036mm, 310mm, 2mm, 11mm, 96N and 26.30kV, to predict the CPCM status values of the respective pillars, and finally output the CPCM status value y of the pillars.

V. ANALYSIS OF EXAMPLES

A. ANALYSIS OF EXPERIMENTAL DATA

Taking the measured data of the uplink catenary in a station area of the Quzhou-Jiujiang railway as a test example, the deep neural network is used to establish the CPCM status prediction model for the catenary.

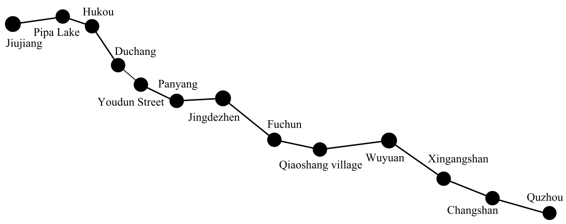


FIGURE 7. Route map of Quzhou-Jiujiang railway.

TABLE 4. Detection data of uplink catenary along Quzhou-Jiujiang railway.

Station	Pillar No.	Height/mm	Stagger/mm	Hard point/mm	Height difference within span/mm	Contact Force/N	Contact voltage/kV	Speed/km/h
Hukou - Du chang	1186	6041	-243	3	8	99	27.06	152
	1184	6033	-247	3	8	95	27.101	152
	1182	6037	-235	4	4	93	27.09	152
	1180	6048	-251	4	11	94	27.103	152
	1178	6033	-55	4	15	98	27.092	153
	1176	6034	-251	4	1	99	27.102	153
	1174	6029	-226	4	5	93	27.078	153
	1172	6033	-161	3	4	96	27.076	153

Tab.4 shows examples of the test data of pillar No. 162-178 of the catenary along Hukou-Duchang, including the pillar number, the kilometer mark, the contact line height (/mm), the contact line stagger (/mm), the height difference with span (/mm), the hard point (/mm), the contact force (/N) and other key CPCM status parameters. The measured test values in the project are input into the prediction model for learning, training and optimization, and for experimental analysis and comparison.

B. EVALUATION INDICTOR FOR THE PREDICTION MODEL OF THE SUSPENSION STATUS OF THE CATENARY

1) MEAN ABSOLUTE PERCENTAGE ERROR AND ROOT MEAN SQUARE ERROR OF THE MODEL

In order to evaluate the prediction precision of the prediction model, the mean absolute percentage error (MAPE) and the root mean square error (RMSE) are selected as two indicators, and the calculation is as shown in formulas (18) and (19).

$$MAPE = \frac{1}{N}\sum_{i=1}^{n}\frac{|y-y*|}{y} \times 100\% \tag{18}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left|\frac{y-y*}{y}\right|^2} \tag{19}$$

where, N is the number of samples predicted; y∗ the predicted value of the catenary CPCM status; y the actual value of the CPCM status.

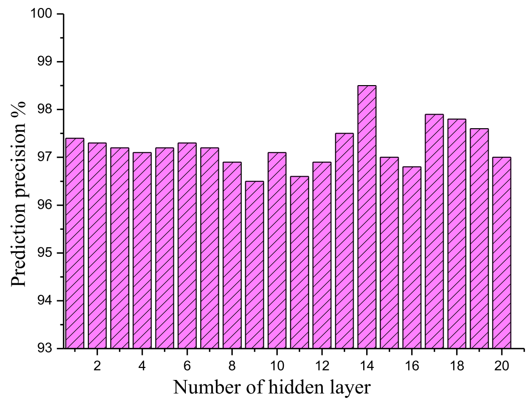


FIGURE 8. Curve of prediction precision in different hidden layers.

2) GENERALIZATION ABILITY OF PREDICTIVE MODEL

Generalization ability refers to the ability of the learning model to predict unknown data, and generalization error is an indicator of generalization ability. The generalization error indicates the prediction precision of the prediction model for different data sets. The generalization error formula is shown in formulas (20) and (21).

$$R_1 = \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(x_i)) \tag{20}$$

$$R = \frac{1}{n}\sum_{j=1}^{n} R_j \tag{21}$$

where, N is the number of samples; $L(y_i, f(x_i))$ the loss function; y_i the actual value; and $f(x_i)$ is the predicted value.

C. EXPERIMENTAL ANALYSIS

1) INFLUENCE OF DIFFERENT HIDDEN LAYERS ON MODEL PREDICTION PRECISION

In order to find the optimal hidden layer number of the CPCM status prediction model, with other parameters unchanged, the comparison experiment of the prediction precision is carried out by changing the hidden layer number. The prediction precision curves of different hidden layers are shown in Fig. 8.

It can be seen from Fig. 8 that the model has the maximum prediction precision of 98.5% when the number of hidden layers is 14, and so 14 hidden layers are optimal configuration.

2) COMPARISON OF PREDICTION PRECISION OF DIFFERENT MODELS

With the same data amount, the comparison experiment of the CPCM status prediction model for the catenary was carried out in different prediction models. The experimental results are shown in Fig. 9 and Tab. 5.

Both Tab. 5 and Fig. 9 reveal the prediction errors of the three CPCM status prediction models are small, but the prediction error of the GA-ADNN model is smaller.

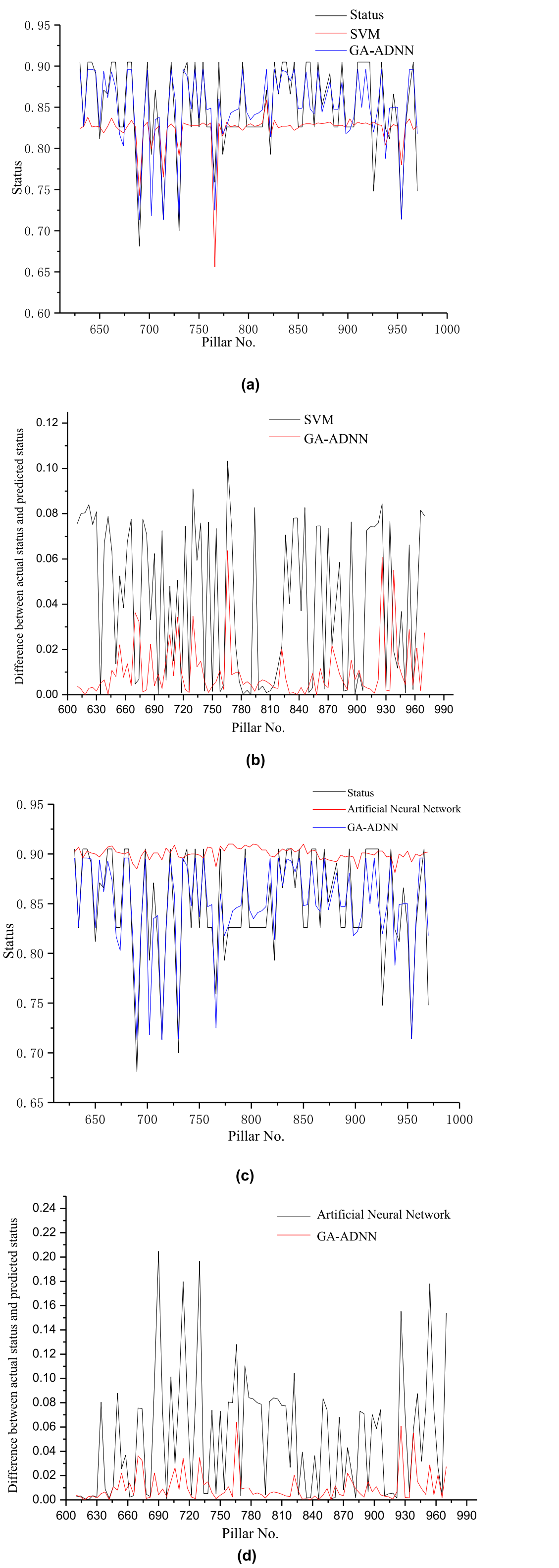


FIGURE 9. Comparison curves of predicted results by different models. (a) Comparison curve of results of SVM prediction and GA-ADNN-based prediction. (b) Comparison curve of result difference between SVM prediction and GA-ADNN-based prediction. (c) Comparison curve of predicted results by artificial neural network and GA-ADNN-based model. (d) Comparison curve of differences of predicted results by artificial neural network and GA-ADNN-based model.

TABLE 5. Prediction results by different CPCM status prediction models.

Pillar No.	Status	SVM	Artificial Neural Network	GA-ADNN
178	0.853	0.845	0.904	0.895
174	0.909	0.840	0.904	0.895
170	0.909	0.841	0.903	0.895
162	0.923	0.840	0.905	0.896
158	0.831	0.850	0.904	0.827
154	0.775	0.845	0.903	0.814
...	...	...	...	...
94	0.923	0.837	0.906	0.896

TABLE 6. Prediction results of different CPCM status prediction models.

Model/indicator	RMSE	MAPE/%
Random Forest	0.512	72.4
Ensemble Model	0.514	72.77
Bayesian Model	0.445	72.47
ANN	0.076	92.5
SVM	0.271	93.2
ADNN	0.032	97
GA-DNN	0.038	95
GA-ADNN	0.029	98.5

The RMSE and MAPE of the prediction results of the three CPCM status prediction models are respectively calculated by the formulas (18) and (19). The evaluation indicator results of the different prediction models are shown in Tab. 6.

Tab. 6 shows that under the same conditions, the mean absolute percentage error and root mean square error of the GA-ADNN model are the smallest compared with other prediction models, and the prediction precision is 6% higher than that of the artificial neural network, 5.3% higher than that of the support vector machine.

3) GENERALIZATION ABILITY OF DIFFERENT MODELS

In order to study the generalization characteristics of the prediction algorithm in different model datasets, the five datasets composed of five different segments of the uplink Quzhou-Jiujiang railway are predicted and compared, and the prediction precision of different datasets is obtained, as shown in Fig. 10.

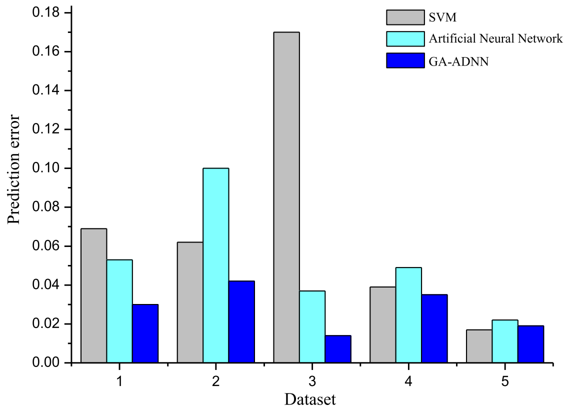
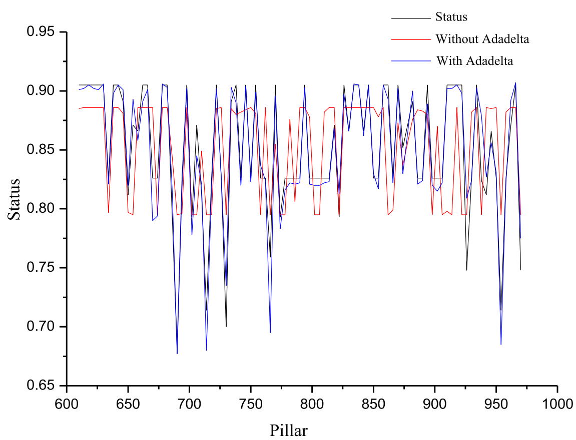
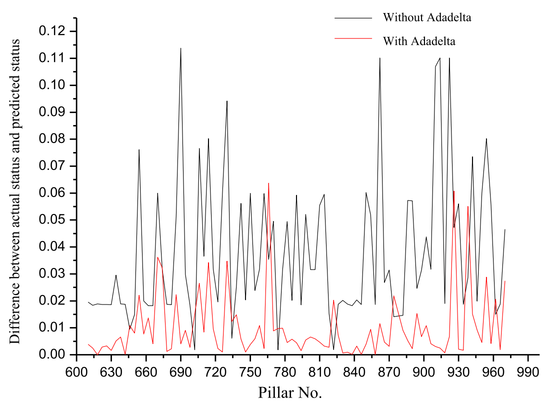


FIGURE 10. Prediction precision of different models for datasets of Quzhou-Jiujiang railway segments.



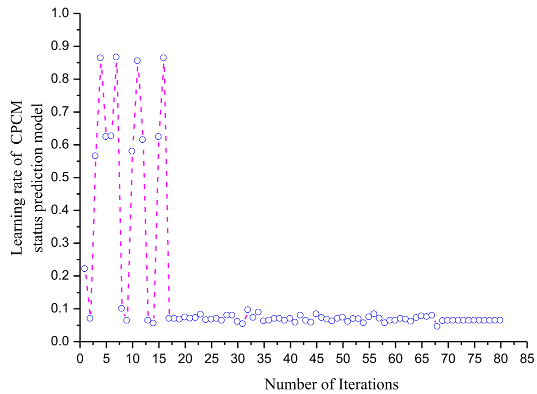
(a)



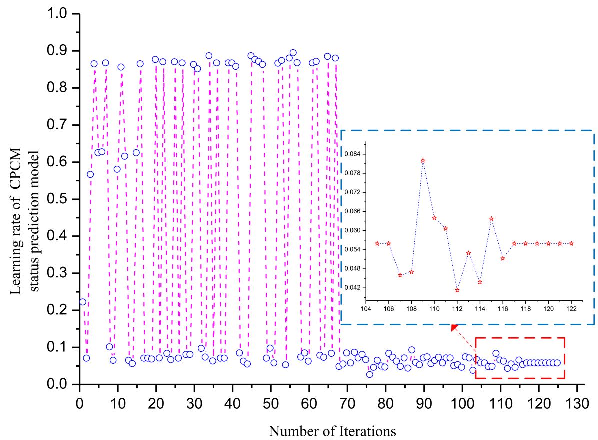
(b)

FIGURE 11. Influence of the Adadelta CPCM status prediction model. (a) Comparison curve of predicted results of the prediction models with and without Adadelta. (b) Comparison curve of the difference of predicted results of the prediction models with and without Adadelta.

It can be seen from Fig.10 that the GA-ADNN model has higher prediction precision for datasets of Quzhou-Jiujiang railway segments, compared with the support vector machine and the artificial neural network models. Based on the



(a)



(b)

FIGURE 12. Iteration and convergence curve of genetic algorithm. (a) Iteration and convergence curve of genetic algorithm without Adadelta. (b) Iteration and convergence curve of genetic algorithm with Adadelta.

formulas (20) and (21), the generalization error of each model is obtained as follows:

$$R_{\mathrm{ANN}} = 1/5(0.053+0.10+0.037+0.049+0.022)=0.052$$
$$R_{\mathrm{SVM}} = 1/5(0.07+0.06+0.017+0.039+0.017)=0.071$$
$$R_{\mathrm{GA-ADNN}} = 1/5(0.03+0.04+0.014+0.035+0.019)=0.028$$

Therefore, the GA-ADNN model has the smallest generalization error compared with the SVM and the traditional ANN models, and its value is 0.028.

4) INFLUENCE OF ADAPTIVE LEARNING RATE ALGORITHM ON THE PREDICTION PRECISION OF CPCM STATUS PREDICTION MODEL

Under the condition of genetic algorithm optimization, the comparison curve of prediction precision of the prediction models for the CPCM status with and without Adadelta is shown in Fig. 11.

Fig.11 shows that compared with the CPCM status prediction model without Adadelta, the prediction curve of the CPCM status prediction model with Adadelta has better fitness with the actual status value.

5) CONVERGENCE EXPERIMENT FOR GENETIC OPTIMIZATION OF THE CATENARY CPCM STATUS PREDICTION MODEL

The genetic algorithm optimizes the learning rate of the prediction model for the catenary CPCM status. By continuously iterating to find the optimal learning rate, the iteration and convergence curve of the genetic algorithm is as shown in Fig. 12.

By comparing the convergence graph of the genetic algorithm with or without Adadelta, it is known that the learning rate of CPCM status predition model is mainly adjusted by the Adadelta before 70 times, and the learning rate fluctuates between 0.047 and 0.89. The optimal learning rate is found in the range of $0.05 \sim 0.08$ at about 70 times. Finally, the genetic algorithm is used to optimize the the learning rate after 70 times and find the optimal learning rate. It can be seen from Fig. 12 that the genetic optimization without Adadelta converges faster than the genetic optimization with Adadelta, but the accuracy of CPCM status prediction model with Adadelta is higher than that without Adadelta, which can be seen from the Tab. 6. However, the main purpose of this paper is to pursue high precision, so the CPCM status prediction model with Adadelta is chosen.

VI. CONCLUSIONS

In this paper, a new GA-ADNN-based optimization method for the prediction model for catenary CPCM status is proposed. The aim is to solve the problem of low efficiency of traditional periodic maintenance of CPCM of railway catenary. Firstly, the deep neural network model is applied to the status prediction of the catenary CPCM status. Secondly, the Adadelta and stochastic gradient descent algorithm are used to update the parameters of the model. Finally, the genetic algorithm is used to optimize the model globally.

The key findings of this paper are the following:

1) Based on the measured data of the catenary project in a station area of the Quzhou-Jiujiang Railway, a deep neural network optimization prediction model is established. Under the condition of the smallest prediction error, an optimal learning rate of 0.0559 of the prediction model is found by the genetic algorithm, and it has been experimentally verified that 14 hidden layers are optimal configuration.

2) The GA-ADNN prediction optimization method involving the deep neural network for catenary CPCM status prediction helps to improve the prediction precision of the catenary CPCM status from 92.5% to 98.5%, and reduce the generalization error of prediction model from 0.071 to 0.028.

REFERENCES

[1] D. Feng, Z. He, S. Lin, Z. Wang, and X. Sun, "Risk index system for catenary lines of high-speed railway considering the characteristics of time-space differences," *IEEE Trans. Transp. Electrific.*, vol. 3, no. 3, pp. 739–749, Sep. 2017.

[2] L. Guo, X. Gao, Q. Li, W. Huang, and Z. Shu, "Online antiicing technique for the catenary of the high-speed electric railway," *IEEE Trans. Power Del.*, vol. 30, no. 3, pp. 1569–1576, Jun. 2015.

[3] S. Barmada, M. Raugi, M. Tucci, and F. Romano, "Arc detection in pantograph-catenary systems by the use of support vector machines-based classification," *IET Elect. Syst. Transp.*, vol. 4, no. 2, pp. 45–52, Jun. 2014.

[4] S. B. Gao, Z. G. Liu, and L. Yu, "Detection and monitoring system of the pantograph-catenary in high-speed railway (6C)," in *Proc. 7th Int. Conf. Power Electron. Syst. Appl.*, Hong Kong, Dec. 2017, pp. 1–7.

[5] Z. Liu, H. Wang, R. Dollevoet, Y. Song, A. Núñez, and J. Zhang, "Ensemble EMD-based automatic extraction of the catenary structure wavelength from the pantograph-catenary contact force," *IEEE Trans. Instrum. Meas.*, vol. 65, no. 10, pp. 2272–2283, Oct. 2016.

[6] P. Jiang, F. Liu, and Y. Song, "A hybrid forecasting model based on date-framework strategy and improved feature selection technology for short-term load forecasting," *Energy*, vol. 119, pp. 694–709, Jan. 2017.

[7] F. Liu, J. Lu, and G. Zhang, "Unsupervised heterogeneous domain adaptation via shared fuzzy equivalence relations," *IEEE Trans. Fuzzy Syst.*, vol. 26, no. 6, pp. 3555–3568, Dec. 2018.

[8] P. Louridas and C. Ebert, "Machine learning," *IEEE Softw.*, vol. 33, no. 5, pp. 110–115, Sep./Oct. 2016.

[9] A. J. Stimpson and M. L. Cummings, "Assessing intervention timing in computer-based education using machine learning algorithms," *IEEE Access*, vol. 2, pp. 78–87, 2014.

[10] F. R. Gomez, A. Rajapakse, U. Annakkage, and I. Fernando, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1474–1483, Aug. 2011.

[11] Y. Wang, S. Xia, Q. Tang, J. Wu, and X. Zhu, "A novel consistent random forest framework: Bernoulli random forests," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 8, pp. 3510–3523, Aug. 2018.

[12] E. Kabir, S. Guikema, and B. Kane, "Statistical modeling of tree failures during storms," *Rel. Eng. Syst. Saf.*, vol. 177, pp. 68–79, Sep. 2018.

[13] L. Yan, Z. Huang, Y. Zhang, L. Zhang, D. Zhu, and B. Ran, "Driving risk status prediction using Bayesian networks and logistic regression," *IET Intell. Transport Syst.*, vol. 11, no. 7, pp. 431–439, Sep. 2017.

[14] S. B. Sulistyo, W. L. Woo, and S. S. Dlay, "Regularized neural networks fusion and genetic algorithm based on-field nitrogen status estimation of wheat plants," *IEEE Trans. Ind. Informat.*, vol. 13, no. 1, pp. 103–114, Feb. 2017.

[15] R. Sahraeian and D. van Compernolle, "Cross-entropy training of DNN ensemble acoustic models for low-resource ASR," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 26, no. 11, pp. 1991–2001, Nov. 2018.

[16] C. Xu *et al.*, "Multi-loss regularized deep neural network," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 26, no. 12, pp. 2273–2283, Dec. 2016.

[17] A. H. Abdulnabi, G. Wang, J. Lu, and K. Jia, "Multi-task CNN model for attribute prediction," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 1949–1959, Nov. 2015.

[18] C. Xu, G. Wang, X. Liu, D. Guo, and T. Liu, "Health status assessment and failure prediction for hard drives with recurrent neural networks," *IEEE Trans. Comput.*, vol. 65, no. 11, pp. 3502–3508, Nov. 2016.

[19] W. Zhang *et al.*, "LSTM-based analysis of industrial IoT equipment," *IEEE Access*, vol. 6, pp. 23551–23560, 2018.

[20] T. Kari *et al.*, "Hybrid feature selection approach for power transformer fault diagnosis based on support vector machine and genetic algorithm," *IET Gener., Transmiss. Distrib.*, vol. 12, no. 21, pp. 5672–5680, Nov. 2018.

[21] K. Choi, D. Jang, S. Kang, J. Lee, T. Chung, and H. Kim, "Hybrid algorithm combing genetic algorithm with evolution strategy for antenna design," *IEEE Trans. Magn.*, vol. 52, no. 3, Mar. 2016, Art. no. 7209004.

[22] S. Piersanti and A. Orlandi, "Genetic algorithm optimization for the total radiated power of a meandered line by using an artificial neural network," *IEEE Trans. Electromagn. Compat.*, vol. 60, no. 4, pp. 1014–1017, Aug. 2018.

[23] G. Zhang, L. Wang, A. P. Duffy, D. D. Febo, A. Orlandi, and H. Sasse, "Applying the analytic hierarchy process (AHP) to an FSV-based comparison of multiple datasets," *IEEE Trans. Electromagn. Compat.*, vol. 57, no. 3, pp. 477–483, Jun. 2015.

[24] N. Yaraghi, P. Tabesh, P. Guan, and J. Zhuang, "Comparison of AHP and monte carlo AHP under different levels of uncertainty," *IEEE Trans. Eng. Manage.*, vol. 62, no. 1, pp. 122–132, Feb. 2015.

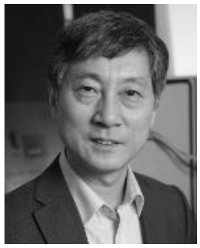
[25] T. Nguyen and S. Nahavandi, "Modified AHP for gene selection and cancer classification using type-2 fuzzy logic," *IEEE Trans. Fuzzy Syst.*, vol. 24, no. 2, pp. 273–287, Apr. 2016.

[26] J. Konečný, J. Liu, P. Richárik, and M. Takàć, "Mini-batch Semi-stochastic gradient descent in the proximal setting," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 242–255, Mar. 2016.

ZHIJIAN QU was born in China, in 1978. He received the M.S. degree from the School of Electrical Engineering, East China Jiaotong University, Nanchang, China, in 2004, and the Ph.D. degree from the School of Electrical Engineering, Beijing Jiaotong University, Beijing, China, in 2012. He has participated a special research in the University of New Brunswick, Canada, as a Visiting Scholar. He is currently a Professor with East China Jiaotong University, China. His current research interests include smart distribution grid information networks, big data sets information systems, and intelligent monitoring systems.

SHENGAO YUAN was born in China, in 1993. He is currently pursuing the master's degree with the School of Electrical and Automation Engineering, East China Jiaotong University. His research interests include intelligent algorithms and big data analysis and mining.

RUI CHI was born in China, in 1985. She received the M.S. degree in computational mathematics from Northern University for Nationalities, Yinchuan, China, in 2012, and the Ph.D. degree in transportation information engineering and control from the School of Automation, Wuhan University of Technology, Wuhan, China, in 2017. She is currently a Lecturer with East China Jiaotong University, China. Her current research interests include intelligent optimization algorithms and neural networks.



LIUCHEN CHANG received the B.Sc. degree from Beijing Jiaotong University, in 1982, the M.Sc. degree from the China Academy of Railway Sciences, in 1984, and the Ph.D. degree from Queen's University, in 1991, all in electrical engineering. He joined the Faculty of University of New Brunswick, in 1992, where he is currently a Professor in electrical and computer engineering. His current interests include power converters, direct load controls, and distributed generation systems. He is a Fellow of the Canadian Academy of Engineering and the Vice President of the IEEE Power Electronics Society. He was the General Chair of the 2015 7th IEEE Energy Conversion Congress and Exposition (ECCE), in Montreal, Canada, and 2016 8th IEEE International Power Electronics and Motion Control Conference–ECCE Asia, Hefei, China.

LIANG ZHAO was born in China, in 1992. He is currently pursuing the School of Electrical and Automation Engineering, East China Jiaotong University. His research interest includes distribution network dispatch automation monitoring data processing.