Article

Prediction of the First Weighting from the Working Face Roof in a Coal Mine Based on a GA-BP Neural Network

Tingjiang Tan, Zhen Yang *, Feng Chang and Ke Zhao

School of Mines, Key Laboratory of Deep Coal Resource Mining, Ministry of Education of China, China University of Mining & Technology, Xuzhou 221116, China; xiaotan@cumt.edu.cn (T.T.); changfeng@cumt.edu.cn (F.C.); zhaoke@cumt.edu.cn (K.Z.)

* Correspondence: yangzhen@cumt.edu.cn; Tel.: +86-186-5179-8874

Received: 1 September 2019; Accepted: 30 September 2019; Published: 3 October 2019

Abstract: The accidents caused by roof pressure seriously restrict the improvement of mines and threaten production safety. At present, most coal mine pressure forecasting methods still rely on expert experience and engineering analogies. Artificial neural network prediction technology has been widely used in coal mines. This new approach can predict the surface pressure on the roof, which is of great significance in coal mine production safety. In this paper, the mining pressure mechanism of coal seam roofs is summarized and studied, and 60 sets of initial pressure data from multiple working surfaces in the Datong mining area are collected for gray correlation analysis. Finally, 12 parameters are selected as the input parameters of the model. Suitable back propagation (BP) and GA(genetic algorithm)-BP initial roof pressure prediction models are established for the Datong mining area and trained with MATLAB programming. By comparing the training results, we found that the optimized GA-BP model has a larger determination coefficient, smaller error, and greater stability. The research shows that the prediction method based on the GA-BP neural network model is relatively reliable and has broad engineering application prospects as an auxiliary decision-making tool for coal mine production safety.

Keywords: first weighting strength; first weighting interval; coal mine; gray correlation analysis; GA-BP

1. Introduction

Roof accidents are one of the main disasters during coal mining, which seriously restricts the yield of mines and threatens the safety of production [1]. The most direct cause of roof accidents is excessive roof pressure on the working face, which leads to roof caving, sheet walls, roof support collapse, and impact ground pressure [2,3]. Therefore, improving roof pressure monitoring and prediction is an effective method to reduce roof accidents, and so is of great significance in coal mine production safety. In the process of roof pressure monitoring and prediction, it is very important to calculate the strength and interval of the first weighting. The first weighting indicates that the working face support forces generally increase when the roof of the coal mine face collapses for the first time. At present, there are two methods to determine the first weighting strength and interval. One is the estimation method, which is calculated according to the existing mine pressure research results and uses an empirical formula. Due to the numerous and complex factors affecting the first weighting, various estimation methods are used in different conditions, but they are not sufficiently accurate. Another method is practical measurement, which determines the first weighting strength and interval of the adjacent working face according to the measured data of the mined working face [4–6]. How to determine the strength and interval more accurately is an urgent problem to be solved in production. Because the first

weighting of the coal mine roof is mainly affected by the geological conditions and mining parameters, and the relationship between them is nonlinear for the same or similar mining geological conditions and mining parameters, we can make use of artificial neural network technology in terms of the analysis of the measured training data for one mining area to predict another one.

In recent years, with the rapid development of artificial intelligence technologies, artificial neural networks have been widely used to deal with nonlinear problems in coal mining and rock mechanics [6–10]. Zhang et al. (2010) used a coupled fault tree analysis (FTA) and artificial neural network (ANN) model to improve the prediction of the potential coal and gas outburst events during the underground mining of thick and deep coal seams in China [11]. A gray model that combines the finite element method (FEM) and an artificial neural network (ANN) was developed for a more precise prediction of pore pressure changes [12]. The recorded nonlinear behavior of the mudrock is modeled using an artificial neural network (ANN). Using this method facilitates the prediction of the swelling pressure of the argillaceous rocks with different swelling potentials and site conditions [13]. An artificial neural network is used to predict the surface subsidence caused by coal mining [14]. Inspired by these studies, we considered applying an artificial intelligence algorithm to the study of coal face roofs. This paper tries to use a BP neural network to predict the first weighting of the working face roof. The training results show that, although the BP model can predict the first strike step and strength, it has the disadvantages of easily falling into a local minimum and slow convergence speed. Other researchers have tried to solve these problems by combining the ANN with other optimization methods such as a genetic algorithm (GA), response surface methodology (RSM), particle swarm optimization (PSO), and ant colony optimization (ACO). For example, Ashena and Moghadasi have proven that both GA and ACO are highly effective at optimizing the performance of the neural networks to estimate the bottom hole pressure [15]. Winiczenko et al. introduced an efficient optimization method, combining RSM and GA to find the optimal topology of artificial neural networks (ANNs) for predicting color changes in rehydrated apple cubes [16]. Bashir and El-Hawary utilized PSO for the optimization of the ANN's weights and obtained improved results compared with the traditional BP algorithm [17]. Because GA theory is more mature and widely used, this paper uses GA to optimize the BP neural network model. GA is a kind of bionic optimization algorithm first proposed by John Holland of the University of Michigan in 1975; it has the advantages of global search ability, parallelism, and randomness. These characteristics of GA can make up for the shortcomings of a neural network [18]. Thus, this study aims to establish a neural network model, optimized by a genetic algorithm that can be used to predict the initial pressure of coal mine working face, which is of great significance for guiding the safe and efficient production of coal mines.

2. The BP Neural Network Model

2.1. Fundamentals of the BP Neural Network

The BP neural network is a multilayer feedforward neural network trained in accordance with the error backpropagation algorithm. The three-layer network structure based on the BP algorithm includes an input layer, a hidden layer, and an output layer, which are composed of neurons. The input layer is responsible for receiving and transmitting the external information. The middle layer is responsible for information processing and transformation. The output layer outputs the results of the information processing. In the neural network process, information is transmitted from the input layer to the output layer (information is transmitted forward) through the hidden layer. If the current output is the same as expected, the training ends. Otherwise, the error is fed back into the network (error backpropagation). By calculating the error signal of the output and expectation in reverse, according to the original connection path, new weights and threshold values of each layer are obtained and propagated to the input layer one by one. The memory training of the neural network refers to the repeated alternating of information forward propagation and error backpropagation. When the global error of the network is less than the given error value, the convergence of the network and the

corresponding stable weights can be obtained, and the learning stops. The flowchart of the BP neural network algorithm is shown in Figure 1 [19–21].

Figure 1. Flowchart of the BP neural network.

The general steps of the BP neural network model construction include sample selection, sample data normalization, and model establishment.

2.2. Selection of the Sample Data

The accuracy and reliability of the first weighting prediction depend on the understanding of the roof pressure theory. Before mining, the coal seam is balanced with the strata in all directions. After excavating the cut hole, the rock will be deformed due to the destruction of the original stress balance, and a new stress balance will be generated, resulting in the formation of a temporarily balanced rock loosening circle above the roof. The roof weighting refers to the pressure of the hydraulic support supporting the circle. Two main factors affect the roof pressure of the working face: the geological conditions and mining process parameters. The geological conditions mainly include the lithology and thickness of the direct roof and basic roof of the working face, the inclination, and the depth of coal seam and other parameters. The mining technological parameters mainly include the support pattern, the length, width, and mining height of the working face, and the advance speed [22–26].

There is a nondeterministic relationship between the first weighting and the various influencing factors in the coal mine roof, and the influence degree of each factor on the output results is also different. Too many input parameters will affect the speed of the neural network training and bring about some difficulties in data acquisition. Therefore, a correlation analysis of various factors is needed to select the appropriate sample parameters. The gray correlation analysis method is adopted in this paper to analyze the relationships between various factors, which applies to nonlinear problems with complex structures and can be used to conduct a quantitative comparative analysis of a series of dynamic data [27]. The initial sample input parameters chosen are the depth and dip of the coal seam and their varying rates, the length and width of the working face, the thickness of the coal

seam, main roof and immediate roof, mining height, advance speed, and roof condition. The output parameters are selected as the initial pressure and periodic pressure. Because the data for the mining height, dip, and coal thickness are only average values and are changing, a simple average value cannot reflect the change of the values, so this paper introduces the rate of change of the mining height ΔS , the rate of change of the dip $\Delta\beta$, and the rate of change of the coal thickness ΔM .

$$\Delta S = \frac{S_2 - S_1}{\bar{S}} \quad (1)$$

$$\Delta\beta = \frac{\beta_2 - \beta_1}{\bar{\beta}} \quad (2)$$

$$\Delta M = \frac{M_2 - M_1}{\bar{M}}, \quad (3)$$

where S_1 , S_2 , and $\bar{S}$ are the minimum, maximum, and average values, respectively, of the buried depth; β_1 , β_2 , and $\bar{\beta}$ are the minimum, maximum, and average values, respectively, of the coal seam pitch; and M_1 , M_2 , and $\bar{M}$ are the minimum, maximum, and average values, respectively, of the coal thickness. Table 1 represents a portion of the original data.

Because the dimensions and orders of magnitude of each input parameter are different, the input parameters must be dimensionless before conducting the gray correlation analysis [28], and the formula used is Equation (4):

$$x'_i(k) = \frac{x_i(k)}{\frac{1}{m} \sum_{k=1}^m x_i(k)}, \quad (4)$$

where $x'_i(k)$ is the k th dimensionless value of the i th input parameter, $x_i(k)$ is the k th value of the i th input parameter, and m is the total of the i th input parameter. The following formulas are used to analyze the gray correlation degree:

$$\xi_i(k) = \frac{\min_i(\Delta_i(\min)) + 0.5\max_i(\Delta_i(\max))}{|x_0(k) - x_i(k)| + 0.5\max_i(\Delta_i(\max))} \quad (5)$$

$$\min_i(\Delta_i(\min)) = \min_i(\min_k |x_0(k) - x_k(k)|) \quad (6)$$

$$\max_i(\Delta_i(\max)) = \max_i(\max_k |x_0(k) - x_i(k)|) \quad (7)$$

$$\gamma_i = \frac{1}{N} \sum_{k=1}^N \xi_i(k), \quad (8)$$

where $\xi_i(k)$ is the relative difference between the comparison curve x_i and the reference curve x_0 at time k , and γ_i is the correlation degree of x_i . Table 2 shows the correlation degree table of the first weighting and related influencing factors.

Table 1. The original data.

Buried Depth (m)	Change Rate of Buried Depth	Coal Seam Pitch (°)	Change Rate of Coal Seam Pitch	Working Face Width (m)	Thickness of Immediate Roof (m)	Thickness of Main Roof (m)	Mining Height (m)	Coal Thickness (m)	Change Rate of Coal Thickness	Advance Speed (m/d)	Roof Condition of Coal Seam	First Weighting Strength (MPa)	First Weighting Interval (m)
225.18	0.119	1.4	0.615	184	12.35	23.08	4	4.1	0.4	5.8	3	46.9	59.8
155.71	0.303	2.7	0.6	253	8.08	17.35	3.4	3.4	0.15	10.5	2	46.0	45.1
190.83	0.046	6.6	0.35	295	4.09	30.00	4.1	4.3	0.98	7.0	1	35.9	46.6
217.87	0.081	1.0	0.67	267	6.20	20.60	3.5	3.7	0.27	9.2	2	42.4	49.1
244.73	0.055	7.5	0.875	205	4.84	14.29	3.6	3.5	0.34	11.8	3	39.2	43.3
254.30	0.556	6.1	0.5	228	6.37	34.89	4	4.0	0.17	8.0	2	42.7	54.3
196.61	0.469	4.3	0.3	240	8.28	20.04	4.2	4.2	0.22	9.9	2	40.4	49.2
139.41	0.353	5.0	0.8	308	9.33	25.83	4.4	4.5	0.75	10.3	3	40.6	59.7
140.90	0.4	2.5	1.5	200	8.49	32.95	2	2.0	0.08	8.0	3	36.9	49.2
156.05	0.66	4.1	0.8	291	13.50	23.42	4.3	4.2	0.23	9.6	2	38.1	53.2
237.70	0.182	7.7	0.8	264	9.22	34.35	3.4	3.5	0.18	5.8	1	34.3	59.8
155.60	0.46	4.7	0.33	229	14.32	24.45	4.5	4.5	0.41	11.5	2	48.8	45.1
234.00	0.28	4.6	1.25	205	10.65	32.82	3.3	3.3	0.13	6.3	2	44.5	46.6
154.09	0.294	2.5	0.22	236	14.49	33.42	3.1	3.2	0.45	6.9	3	48.9	49.1
250.10	0.079	4.3	0.8	243	5.89	25.83	3.9	4.0	0.63	10.6	1	36.6	43.3
169.00	0.31	5.3	0.8	196	11.11	13.57	2.5	2.6	0.04	8.4	2	48.7	54.3
147.52	0.23	5.7	0.157	257	6.47	22.59	3.1	3.2	0.13	10.4	2	46.7	49.2
155.15	0.063	3.6	0.72	209	11.06	25.30	4.2	4.3	0.27	7.8	2	43.2	59.7
206.25	0.381	3.4	0.67	230	11.34	20.84	3.3	3.4	0.04	6.9	2	41.0	49.2
186.26	0.43	7.9	0.8	256	3.82	18.71	4	4.1	0.41	5.3	2	38.1	53.1

Table 2. The correlations of the first weighting and related influencing factors.

Influencing Factors	γ_i	Influencing Factors	γ_i
Buried depth	0.7149	Thickness of main roof	0.8155
Change rate of buried depth	0.7094	Mining height	0.8468
Coal seam pitch	0.6952	Coal thickness	0.8371
Change rate of coal seam pitch	0.8343	Change rate of coal thickness	0.7017
Working face length	0.4523	Advance speed	0.858
Working face width	0.8604	Roof condition of coal seam	0.8518
Thickness of immediate roof	0.8174		

By analyzing the correlation degree of the factors in the above table, it can be concluded that, apart from the working face length, which has little impact on the mine pressure, the other factors have a significant impact. Therefore, the width of the working face, mining height, advance speed, roof condition of the coal seam, burial depth, thickness and dip of the coal seam, change rate of the inclination angle, burial depth and coal thickness, direct top thickness, and thickness of the main roof are selected as the input parameters for the model training and prediction. The first weighting strength and interval are taken as the output values of the training and prediction of the neural network model.

2.3. Model Parameters

According to the relevant theoretical research on the BP neural network and roof pressure of the working face, we determine the BP neural network model parameters in this paper, including the number of hidden layers, number of hidden layer neurons, initial weights, activation function, expected error, learning rate, and learning times.

(1) The number of hidden layers

Robert Hecht-Nielson proved that a three-layer BP neural network containing only one hidden layer can perform any nonlinear mapping, which can be a mapping from the n dimension to the m dimension. Therefore, the BP neural network prediction model established in this paper is a three-layer BP neural network with only a single hidden layer.

(2) The number of hidden layer neurons

The suitable empirical formula for a BP neural network with a single hidden layer to determine the number of neurons in the hidden layer is shown as follows:

$$z = \sqrt{p + q} + a, \quad (9)$$

where p and q represent the number of nodes in the input layer and output layer, respectively, and a is a constant whose value range is (1,10).

According to the above analysis of the gray correlation degree, the number of input data characteristic values is 12, and the number of output results is 1. Therefore, $z = \sqrt{12 + 1} + 1 \approx 5$, and the initial number of neurons in the hidden layer of the BP neural network model established in this paper is 5. After repeated tests, the graph of the relationship between the number of hidden layer nodes and absolute relative errors is shown in Figure 2. The absolute relative error can be obtained from formula 10. It can be seen from the figure that when the number of hidden layer nodes reaches 15, the network output error is minimized.

$$\varphi = \frac{|y - y'|}{y}, \quad (10)$$

where φ is the absolute relative error; y and y' represent the measured results and training results, respectively.

Figure 2. The errors corresponding to the number of hidden layer nodes.

(3) The initial weights

To avoid the saturation region of the activation function, the weighted sum output value will be near to 0 as much as possible, which can increase the weight adjustment range. Therefore, the initial connection weights are usually random numbers between $(-1,1)$ or $(0,2)$, so that the network will not be greatly affected.

(4) The activation function

The activation function of the hidden layer in this paper is the tansig function, which is found in the MATLAB toolbox, whose value range is $(-1,1)$, which also conforms to the normalized data falling between $(-1,1)$. The output layer selects the purelin function from the MATLAB toolbox, which is linear to increase the value range of the output value. The tansig function is shown below:

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}, \quad (11)$$

where e is the base of the natural log function.

(5) The learning rate

According to practical experience, the value range of the learning rate in the BP neural network model is generally between 0.01 and 0.08. In this paper, based on the sample data collected at present, it is concluded by repeated training that when the learning rate is adopted, the performance of the established roof pressure prediction model based on the BP neural network is better.

(6) The expected error

In this paper, the mean square error function is selected in the construction of the BP neural network, and the formula is as follows:

$$MSE = \frac{1}{mp} \sum_{p=1}^p \sum_{j=1}^m (y'_{pj} - y_{pj})^2, \quad (12)$$

where m represents the number of neurons in the output layer, p is the number of samples, y'_{pj} is the true value, and y_{pj} represents the output value of each training iteration.

2.4. Training Results of the BP Model

It can be seen from the above analysis that the number of nodes in the input layer of the model is 12. The input parameters are the inclined length of the working face, advance speed, roof condition of the coal seam, mining height, coal thickness, change rate of the inclination angle, direct top thickness, basic top thickness, burial depth, change rate of the burial depth, coal thickness change rate, and coal seam inclination angle; through repeated tests in this paper, the number of nodes in the hidden layer of the model was determined to be 15; and according to the requirements, the number of nodes in the output layer of the model is set to 1. Sixty sets of data collected from the Datong mining area were selected to establish the model, among which 50 sets were used as training samples and 10 sets were used as prediction samples. The BP neural network model is trained and evaluated by predicting the initial compressive strength and periodic compressive strength of the roof. MATLAB is used in this paper to train the sample data and validate the model. Due to the instability of the BP neural network, the results of each training session are different, so the pressure of the initial weighting and periodic weighting are taken as the target output for 30 training sessions, and the optimal training results are selected as the model obtained by training.

(1) The first roof weighting strength

Figure 3 shows the comparison between the predicted value and the real value of the first roof weighting strength predicted by the BP prediction model, which is the best result of the curve fitting between the predicted value and true value in 30 training sessions, with a maximum determination coefficient of 0.8807. The determination coefficient is derived from Equation (13) [29]. The larger the determination coefficient, the better the fit is.

$$R^2 = \frac{\left(N \sum_{i=1}^N y'_i y_i - \sum_{i=1}^N y'_i \sum_{i=1}^N y_i \right)^2}{\left(N \sum_{i=1}^N y_i'^2 - \left(\sum_{i=1}^N y'_i \right)^2 \right) \left(N \sum_{i=1}^N y_i^2 - \left(\sum_{i=1}^N y_i \right)^2 \right)}, \quad (13)$$

where N is the number of predicted samples, y'_i is the predicted value, y_i is the true value, and R^2 is the determinant.

Figure 3. Comparison of the first roof weighting strength.

(2) Predicting the first roof weighting interval

Figure 4 shows the comparison between the predicted value and the real value of the first roof weighting interval predicted by the BP prediction model, which is the best result of the curve fitting between the predicted value and true value in 30 training sessions, with a maximum determination coefficient of 0.8756.

Figure 4. Comparison of the first roof weighting interval.

It can be concluded from the above diagram that, although the BP neural network model can predict the first weighting strength and first weighting interval of the coal mine roof, the training determination coefficient is small, which indicates that this model is not stable and has a large error. Therefore, this prediction model needs to be optimized.

3. BP Prediction Model of First Weighting Based on the Genetic Algorithm (GA)

It is necessary to optimize the BP neural network because the volatility of the network is relatively large and it is easy to fall into a locally optimal solution. GA is a domain-independent problem-solving technique, which encodes the problem into a chromosome composed of genes and generates a chromosome representing the solution to the problem through a continuous process of selecting, crossing, and mutating the chromosome. The GA operates on the chromosome code and isolates the characteristics of the problem itself, so it has wide adaptability. At the same time, since the algorithm operates in parallel from multiple initial points, the search process is free from convergence to a local extremum. The GA implements the search process through groups, making it different from a single-point search and easy to parallelize, thus improving the efficiency of the algorithm [30,31]. Therefore, this paper uses a genetic algorithm to optimize the BP network and establish a GA-BP network model. The operational flow of the GA-BP algorithm is shown in Figure 5.

The optimization of the BP neural network by GA mainly includes determining the structure of the BP neural network and optimizing the weights, thresholds, training, and prediction of the model. Since the optimization of the genetic algorithm is the initial weight and threshold of the BP neural network, as long as we know the topological structure of a neural network, we can determine the number of optimization parameters of the genetic algorithm, to determine the coding length of population individuals. The process of genetic algorithm includes the population initialization, fitness function, selection operator, crossover operator, and mutation operator [32,33].

Figure 5. The flowchart of the GA-BP algorithm.

(1) Population initialization

Through individual binary coding, each individual is represented by a binary string. The network structure of this paper is 12-15-1, so the number of weights and thresholds is listed in Table 3.

Table 3. The number of weights and thresholds.

Link Weights between the Input Layer and the Hidden Layer	Thresholds of the Hidden Layer	Link Weights between the Hidden Layer and the Output Layer	Thresholds of the Output Layer
180	15	15	1

(2) Fitness function

In this paper, to make the residual difference between the predicted and actual value as small as possible, the sum of the absolute errors between the predicted and actual value of the predicted sample is selected as the output of the objective function.

$$F = \sum_{i=1}^n \text{abs}(y'_i - y_i), \quad (14)$$

where n is the number of network output nodes, y'_i is the predicted value, and y_i is the true value.

(3) Operator selection

Random ergodic sampling is adopted as the selection operator. A roulette wheel is used to select a pair of individuals; the two individuals compete and the one with the highest fitness is selected.

$$p_i = \frac{F_i}{\sum_{i=1}^n F_i}, \quad (15)$$

where n is the number of individuals in the population, F_i is the fitness function value of the i th individual, and p_i is the probability that the i th individual will be chosen.

(4) Crossover operator

The simplest single-point crossover is adopted as the crossover operator. In the individual code string, only one intersection point is randomly selected, and the coding parts of two pairs of individuals are exchanged at that point.

(5) Mutation operator

The mutation operator produces some mutated genes with a certain probability, and the mutated genes are selected randomly. If the selected gene code is 1, it becomes 0; otherwise, it becomes 1. Table 4 shows the parameters of the genetic algorithm.

Table 4. Parameters of the genetic algorithm.

Population Size	Maximum Genetic Algebra	Binary Digits	Generation Gap	Crossover Probability	Mutation Probability
40	50	10	0.95	0.9	0.01

4. Results and Discussion

The improved GA-BP model is trained and used to predict the initial compressive strength and interval of the working face roof, and then the predicted results are compared with the predicted results obtained by the BP model above, as shown in Figures 6 and 7.

4.1. Comparison of the First Roof Weighting Strength

It can be seen from Figure 6 that the GA-BP model is used to train and predict the initial compressive strength results with a smaller error than that of the BP model, and the GA-BP model has a larger and more stable determination coefficient during the 30 training sessions. The iteration times in Table 5 refers to the number of iterations when the model converges. The iteration times of the GA-BP model and BP model are 21 and 47, respectively, which indicates that the GA-BP model has a faster convergence rate. These results show that the optimized model can better predict the initial compressive strength.

Table 5. Comparison of the BP and GA-BP models.

Methods	R^2	Iteration Times	Computation Time (s)
BP	0.8557	47	24.81
GA-BP	0.9609	21	61.39

Figure 6. Comparison of the first roof weighting strength results. (a) The value of the first roof weighting strength predicted by the BP and GA-BP prediction models; (b) a comparison of the determination coefficients predicted by the BP and GA-BP models.

4.2. Comparison of the First Roof Weighting Interval

It can be seen from Figure 7 that the GA-BP model is used to train and predict the results of the first weighting interval with a smaller error than the BP model, and the GA-BP model has a larger and more stable determination coefficient during the 30 training sessions. It can be seen from Table 6 that the iteration times of the GA-BP model and BP model are 25 and 53 respectively. This indicates that the GA-BP model has a faster convergence rate. These results indicate that the optimized model can better predict the initial compressive interval.

Figure 7. Comparison of the first roof weighting interval results: (a) the value of the first roof weighting interval predicted by the BP and GA-BP prediction models; (b) a comparison of the determination coefficients predicted by the BP and GA-BP models.

Table 6. Comparison of the BP and GA-BP models.

Methods	R^2	Iteration Times	Computation Time (s)
BP	0.8561	53	27.32
GA-BP	0.9605	25	63.87

5. Conclusions

In this paper, we used 60 data points from the Datong mining area for the model training and verification, and established a GA-BP model that is suitable for the Datong mining area to calculate the initial pressure strength and a step distance of the roof. The following conclusions can be obtained:

1. A lot of influencing factors can be determined in the course of prediction of the first weighting of the working face roof utilizing artificial neural network methods, which can map the complicated nonlinear relation between them.
2. The gray correlation degree is used to calculate the relational degrees between the first weighting of the working face roof and various influencing factors in the Datong mining area. The input parameters of the BP prediction model include the width of the working face, mining height, advance speed, roof condition of the coal seam, burial depth, thickness and dip of the coal

- seam, change rate of the inclination angle, burial depth, coal thickness, direct top thickness, and thickness of the main roof.
3. Compared with traditional BP, the BP-GA model has stronger robustness, is available for parallel global searching, and can improve the accuracy of prediction results and the stability of the prediction model.
 4. We can extend this method of calculating the initial compression strength and step distance to other mining areas. As long as we can collect enough input parameter data, we can train a prediction model that is suitable for each mining area.

Author Contributions: Conceptualization, T.T. and Z.Y.; methodology, T.T. and K.Z.; investigation, F.C.; data curation, F.C. and K.Z.; writing—original draft preparation, T.T.; writing—review and editing, Z.Y.

Funding: This work was supported by the Fundamental Research Funds for the Central Universities [Grant No. 2017CXNL01].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Yuan, X. The Characters and Trend of Accidents in the Coal Mining in China. *Dis. Adv.* **2012**, *5*, 866–869.
2. Duzgun, H.S.B.; Einstein, H.H. Assessment and management of roof fall risks in underground coal mines. *Saf. Sci.* **2004**, *42*, 23–41. [[CrossRef](#)]
3. Palei, S.K.; Das, S.K. Logistic regression model for prediction of roof fall risks in bord and pillar workings in coal mines: An approach. *Saf. Sci.* **2009**, *47*, 88–96. [[CrossRef](#)]
4. Razani, M.; Yazdani-Chamzini, A.; Yakhchali, S.H. A novel fuzzy inference system for predicting roof fall rate in underground coal mines. *Saf. Sci.* **2013**, *55*, 26–33. [[CrossRef](#)]
5. Maiti, J.; Khanzode, V.V. Development of a relative risk model for roof and side fall fatal accidents in underground coal mines in India. *Saf. Sci.* **2009**, *47*, 1068–1076. [[CrossRef](#)]
6. Qian, M.G.; Shi, P.W. *Mining Pressure and Strata Control*; Xu, J.L., Ed.; China University of Mining and Technology Press: Xuzhou, China, 2010; pp. 100–119.
7. Khandelwal, M.; Kumar, D.L.; Yellishetty, M. Application of soft computing to predict blast-induced ground vibration. *Eng. Comput.* **2011**, *27*, 117–125. [[CrossRef](#)]
8. Khandelwal, M.; Singh, T.N. Prediction of blast-induced ground vibration using artificial neural network. *Int. J. Rock Mech. Min. Sci.* **2009**, *46*, 1214–1222. [[CrossRef](#)]
9. Monjezi, M.; Ahmadi, Z.; Varjani, A.Y.; Khandelwal, M. Backbreak prediction in the Chadormalu iron mine using artificial neural network. *Neural Comput. Appl.* **2013**, *23*, 1101–1107. [[CrossRef](#)]
10. Zhang, H.; Song, J.; Su, C.; He, M. Human attitudes in environmental management: Fuzzy Cognitive Maps and policy option simulations analysis for a coal-mine ecosystem in China. *J. Environ. Manag.* **2013**, *115*, 227–234. [[CrossRef](#)]
11. Zhang, R.; Lowndes, I.S. The application of a coupled artificial neural network and fault tree analysis model to predict coal and gas outbursts. *Int. J. Coal Geol.* **2010**, *84*, 141–152.
12. Mohammed, M.; Watanabe, K.; Takeuchi, S. Grey model for prediction of pore pressure change. *Environ. Earth Sci.* **2010**, *60*, 1523–1534. [[CrossRef](#)]
13. Doostmohammadi, R.; Moosavi, M.; Araabi, B.N. A model for determining the cyclic swell-shrink behavior of argillaceous rock. *Appl. Clay Sci.* **2008**, *42*, 81–89. [[CrossRef](#)]
14. Ambrozic, T.; Turk, G. Prediction of subsidence due to underground mining by artificial neural networks. *Comput. Geosci.* **2003**, *29*, 627–637. [[CrossRef](#)]
15. Ashena, R.; Moghadasi, J. Bottom hole pressure estimation using evolved neural networks by real coded ant colony optimization and genetic algorithm. *J. Pet. Sci. Eng.* **2011**, *77*, 375–385. [[CrossRef](#)]
16. Winiczenko, R.; Górnicki, K.; Kaleta, A.; Mańkowska-Janaszek, M. Optimisation of ANN topology for predicting the rehydrated apple cubes colour change using RSM and GA. *Neural Comput. Appl.* **2018**, *30*, 1795–1809. [[CrossRef](#)] [[PubMed](#)]
17. Bashir, Z.A.; El-Hawary, M.E. Applying wavelets to short-term load forecasting using PSO-based neural networks. *IEEE Trans. Power Syst.* **2009**, *24*, 20–27. [[CrossRef](#)]

18. Liu, K.; Guo, W.; Shen, X.; Tan, Z. Research on the forecast model of electricity power industry loan based on GA-BP neural network. *Energy Procedia* **2012**, *14*, 1918–1924.
19. Bahadir, E. Prediction of prospective mathematics teachers' academic success in entering graduate education by using back-propagation neural network. *J. Educ. Train. Stud.* **2016**, *4*, 113–122. [[CrossRef](#)]
20. He, Y.; Li, X.L.; Shao, Y.N. Discrimination of varieties of apple using near infrared spectra based on principal component analysis and artificial neural network model. *Spectrosc. Spectr. Anal.* **2006**, *26*, 850–853.
21. Cao, J.C.; Cao, S.H. Study of forecasting solar irradiance using neural networks with preprocessing sample data by wavelet analysis. *Energy* **2006**, *31*, 3435–3445. [[CrossRef](#)]
22. Gu, S.; Yao, B. Study on strata behavior regularity of 1301 face in thick bedrock of Wei—Qiang coal mine. In *IOP Conference Series: Materials Science and Engineering*; IOP Publishing: Bristol, UK, 2017; Volume 231, p. 12108.
23. Li, J.; Huang, Y.; Zhang, J.; Li, M.; Qiao, M.; Wang, F. The influences of Key Strata compound breakage on the overlying strata movement and strata pressure behavior in fully mechanized caving mining of shallow and extremely thick seams: A case study. *Adv. Civil Eng.* **2019**, *2019*, 1–11. [[CrossRef](#)]
24. Lin, G.-Q.; Song, E.-B. The measurement of underground pressure regulation in Ulan Moron coal mine. In *2014 International Conference on Mechanics and Civil Engineering (icmce-14)*; Atlantis Press: Paris, France, 2014; Volume 7, pp. 287–291.
25. Zhao, J.; Liu, L.; Zheng, Z. Law of strata pressure behavior in shallow coal seam. In *IOP Conference Series: Earth and Environmental Science*; IOP Publishing: Bristol, UK, 2018; Volume 113.
26. Zhu, B.; Li, G.; Kou, W.; Li, H. Simulation of strata behaviour laws of a coal mine in Jungar Coalfield. In *Proceedings of the 3rd International Symposium on Mine Safety Science and Engineering*, Montreal, Canada; 2016; pp. 651–654.
27. Liu, S.F.; Xie, N.M. *Grey System Theory and Application*; Science Press: Beijing, China, 2010; pp. 42–132.
28. Yang, B.; He, L.B. Establishment and analysis of two normalization methods for backscattering intensity data. *Acta Oceanol. Sin.* **2018**, *40*, 145–151.
29. Gocic, M.; Motamedi, S.; Shamshirband, S.; Petkovic, D.; Sudheer, C.; Hashim, R.; Arif, M. Soft computing approaches for forecasting reference evapotranspiration. *Comput. Electron. Agric.* **2015**, *113*, 164–173. [[CrossRef](#)]
30. Whitley, D. A genetic algorithm tutorial. *Stat. Comput.* **1994**, *4*, 65–85. [[CrossRef](#)]
31. Leung, Y.W.; Wang, Y.P. An orthogonal genetic algorithm with quantization for global numerical optimization. *IEEE Trans. Evol. Comput.* **2001**, *5*, 41–53. [[CrossRef](#)]
32. Liu, C. Price forecast for gold futures based on GA-BP neural network. In *Proceedings of the 2009 International Conference on Management and Service Science*, Wuhan, China, 20–22 September 2009; pp. 1–4.
33. Tahmasebi, P.; Hezarkhani, A. A hybrid neural networks-fuzzy logic-genetic algorithm for grade estimation. *Comput. Geosci.* **2012**, *42*, 18–27. [[CrossRef](#)] [[PubMed](#)]