

## Article

# Source Discrimination of Mine Water by Applying the Multilayer Perceptron Neural Network (MLP) Method—A Case Study in the Pingdingshan Coalfield

Man Wang <sup>1,2</sup> , Jianguo Zhang <sup>1,2</sup>, Xinyi Wang <sup>3</sup>, Bo Zhang <sup>1,2</sup> and Zhenwei Yang <sup>1,2,3,\*</sup> 

<sup>1</sup> China Pingmei Shenma Holding Group Co., Ltd., Pingdingshan 467000, China; wangman.w@gmail.com (M.W.); zhangjg\_z@126.com (J.Z.); zb2757@163.com (B.Z.)

<sup>2</sup> State Key laboratory of Coking Coal Resources Green Exploitation, China Pingmei Shenma Group, Pingdingshan 467000, China

<sup>3</sup> Institute of Resources & Environment, Henan Polytechnic University, Jiaozuo 454000, China; wangxy@hpu.edu.cn

\* Correspondence: yangzhenwei2006@126.com

**Abstract:** In a complex multiaquifer mine, discriminant approaches that have previously been presented cannot successfully locate water sources. With multiple processing layers, computing models may learn representations of data at various levels of abstraction. Modern technology in many domains has significantly enhanced these models. The problem of distinguishing the source of mine water in mines and tunnels has been addressed by studying the hydrochemical components of the Pingdingshan coalfield and applying the multilayer perceptron neural network (MLP) method to discriminate the source of the mine water. There were five types of mine water in the Pingdingshan coalfield. Each type of water was encoded with the numbers 0 to 4. The one-hot code method was used to encode the numbers, which is the output set. On the basis of hydrochemical data processing, the MLP model was developed using a characteristic ion contrast on aquifers with distinctive chemical properties. The research results show that two hidden layers (with 10 neurons in each hidden layer) for the model were fit for completing the prediction process with a better performance. This approach enabled us to discriminate water sources for the Pingdingshan coalfield and could be tried for other coalfields with similar hydrogeological conditions.

**Keywords:** source discrimination; mine water; artificial neural network; hydrochemical composition; model



**Citation:** Wang, M.; Zhang, J.; Wang, X.; Zhang, B.; Yang, Z. Source Discrimination of Mine Water by Applying the Multilayer Perceptron Neural Network (MLP) Method—A Case Study in the Pingdingshan Coalfield. *Water* **2023**, *15*, 3398. <https://doi.org/10.3390/w15193398>

Academic Editors: Yijun Xu and Reza Derakhshani

Received: 27 July 2023

Revised: 13 September 2023

Accepted: 20 September 2023

Published: 28 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

With the increasing depth of coal mining, the source of mine water inrush becomes increasingly complex. Water inrush in mines can lead to serious disasters all around the world. With rates of 2053, 1270, 1167, and 1678 m<sup>3</sup>/min, respectively, the Dongpang Coal Mine in Xingtai, the Luotuoshan Coal Mine in Wuhai, the Fangezhuang Coal Mine in Tangshan, the Fangezhuang Tunnel on the Yiwan Railway, and the Yeshanguan Tunnel on the Yiwan Railway all had the highest recorded rates of water inrush. Particularly in northern China, the complicated hydrogeological conditions are unique [1,2]. Generally, there are several confined aquifers in the roof or floor of coal seams. When the mining destroys the balance of the original stress of the rock, the water contained in these aquifers might enter the working face or roadway through fracturing, affecting the efficient production of the coal mine or causing serious water inrush accidents [3,4]. This not only reduces the production efficiency, causing great economic loss, but also threatens the safety of miners [5–7]. In order to resume production and save miners, it is crucial to quickly and accurately identify the source of the water inrush.

Between rocks and groundwater, a number of physical and chemical reactions occur regularly, including the water–rock reaction, redox reaction, and others. However, these are

in a dynamic balance [8–10]. Chemical characteristics of the water are distinct in different aquifers. Considering the chemical characteristics of the water in aquifers, hydrochemistry and mathematical methods are widely used to identify water sources in hydrogeology, such as gray relational analysis, principal component analysis, Fisher's discriminant analysis, the distance discrimination method, Bayes' discriminant analysis, etc. [11–14]. It is easy to establish a set of discriminant models that focus on one or several average values. However, the influence of extreme values on average values are usually ignored, and the implicit information of the water chemistry cannot be mined intuitively [7,8]. Given the development of artificial intelligence (AI) in parameter identification, both today and in the foreseeable future, it is important to take advantage of these new technological developments and innovations in the source discrimination of mine water inrush.

Because of its powerful ability to automatically extract high-level representations from complex data, and because it has been widely utilized in the domains of natural science, social science, and engineering, artificial neural network modeling is a hot topic in many fields [15,16]. The multilayer perceptron neural network (MLP), which is widely used in many domains, including computer vision, intelligent robots, natural language processing, and data mining, has recently taken the globe by storm [17–19]. The MLP model enables computational models with numerous processing layers to learn data representations at various degrees of abstraction. It uses the back propagation algorithm to find complex structures in big data sets, suggests how a machine should modify its internal settings, and computes the representation in each layer from the representation in the preceding layer. Among the various options, the MLP is one of the representative models of the artificial neural network model. It exhibits a high prediction accuracy and a strong function approximation ability [20]. This network type is rapidly transforming many industries, including healthcare, energy, finance, transportation, etc. [21,22].

The contribution of this study is to (1) introduce a multilayer perceptron neural network to build a discriminant model for the source discrimination of mine water under the Keras framework based on the Python 3.6 platform and (2) to train the model parameters and to apply them to water discrimination in the Pingdingshan coalfield. The objective of the study is to develop a new idea for the discrimination of water inrush sources and to improve the prevention of water inrush in mines.

## 2. Study Area

### 2.1. Outline of the Coalfield

The third-largest coal producer in China is the Pingdingshan coalfield, which is situated in the central and western regions of Henan Province (113°00'–114° E, 33°30'–34°00' N) (Figure 1). The coalfield is roughly 20 km wide in the north and 20 km long in the east. The Guodishan fault divides the low-lying terrain into eastern and western regions, where the Pingdingshan coalfield is situated. It is a substantial syncline structurally, with proportionally dipping limbs. The coal-bearing sediments are mostly Permian in age and are composed of sandstone, siltstone, and carbonaceous shale, which are overlain by Neogene, Paleogene and Quaternary deposits [23]. The entire sequence is underlain by Cambrian karstic limestone (Figure 1), and the comprehensive histogram of the strata in the Pingdingshan coalfield mine is shown in Figure 2.

The Pingdingshan coalfield was constructed in the 1950s, and the mining technology used for excavation is fully mechanized sublevel caving. The coal seams that are mined in the coalfield are V<sub>2</sub>, IV<sub>3</sub>, III<sub>2</sub>, IV<sub>1</sub>, II<sub>2</sub>, and II<sub>1</sub>. Currently, the clamping gangue is less than 0.3 m, the inclination angle is not greater than 25 degrees, and the mining height ranges from 1.8 to 3.8 m. The first level is currently in the residual mining stage and is now primarily focused on the second level's output.

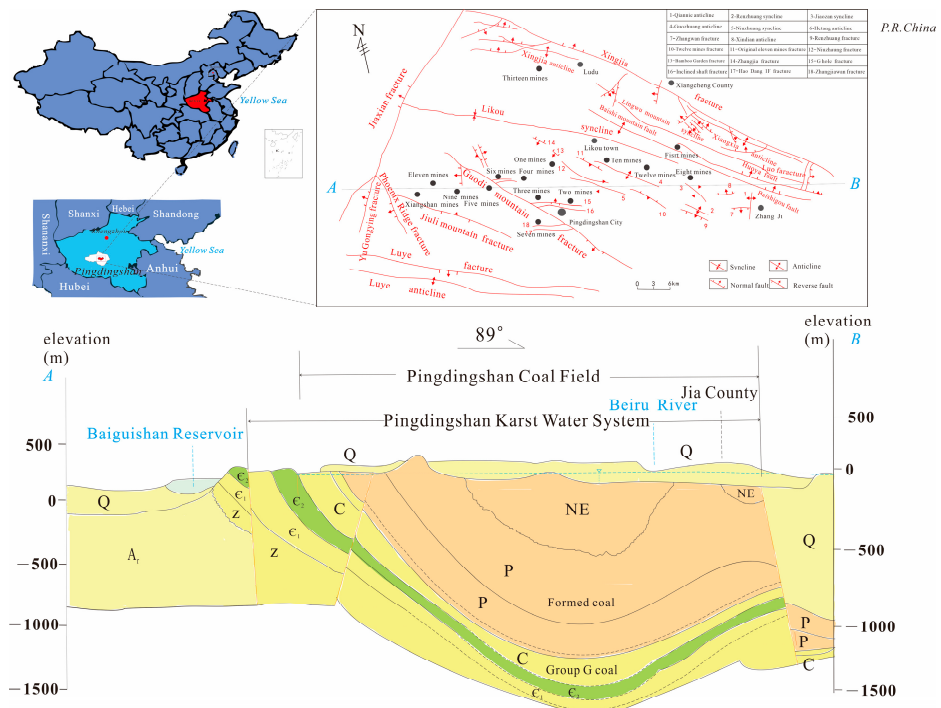
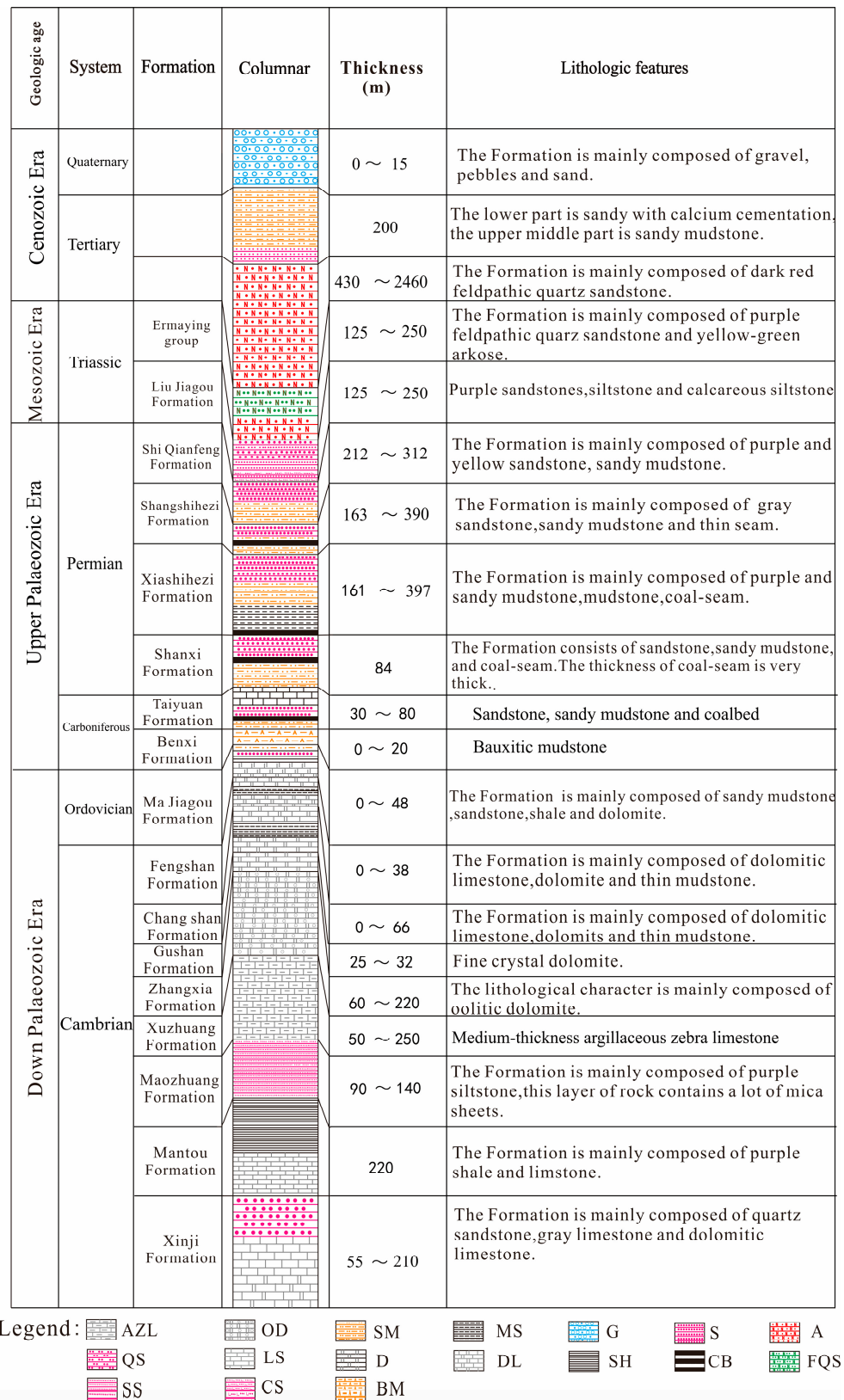


Figure 1. General map of the study area and the cross section along the line A–B.

### 2.2. Hydrogeological Background

The research area is situated in a transitional zone from a warm temperate zone to a subtropical zone, with a long-term average precipitation of 747.4 mm/year, mainly concentrated between July and September. The geomorphology in the east and south is an alluvial plain with a layer of 200 m~500 m thickness. The ground elevation is +75~80 m. With a surface elevation varying from 900 m to 1040 m, the topography is low in the southeast and high in the northwest. Influenced by the topographic features, the surface water is mainly distributed in the south and north of the mining area, that is, the Shahe River, Ruhe River, Zhanhe River, and the Baiguishan Reservoir. The Ruhe River and the Shahe River are perennial rivers that lie on the northern and southern margins of the study area. There are some seasonal rivers and man-made ditches, such as Zhanhe, the Beigan Canal, and the Xigan Canal. The riverbed inserts into Cambrian limestone or Neogene marl, which has a certain replenishment effect on the groundwater of limestone in the Qikuang mine in the southwest of the Pingdingshan coalfield [24].

There are several aquifers from the bottom to the upper part of the study area. The first category is the Cambrian limestone aquifer, which is the indirect water-filled aquifer in the upper coalbed. It is also the main water inrush aquifer at the coal-seam floor because of its high hydraulic pressure and abundant recharge sources. The second category is the Taiyuan Formation of the Carboniferous system. Seven layers of limestone are detected from the upper to the bottom part, where the L<sub>7</sub> limestone is dominated by the corrosion fissure and has poor supplementation conditions. The osmotic coefficient of the aquifer is 0.0076–3.047 m/d, and the water inflow per unit is 0.00018~0.3569 L/sm. The third category is the Dyas sandstone aquifer, which is mainly composed of medium-sized and large sandstones and has generally poor water yield. The fourth category is the Quaternary sand–gravel pore aquifer, which is made up of coarse and medium- and fine-sized sands. The aquifer covers the coal strata and only contacts the mineable seam at the outcrop position. The Pingdingshan coal-mining area mainly contains Cambrian limestone aquifer groundwater, while the Carboniferous limestone aquifer in the Taiyuan Formation has Dyas sandstone aquifer groundwater and Quaternary sand–gravel pore aquifer groundwater. The Cambrian karst aquifer is developed with karst and has the strongest water among the others [25,26].



**Figure 2.** Comprehensive histogram of the strata in the Pingdingshan coalfield mine. AZL, argillaceous zebra limestone; QS, quartz sandstone; SS, siltstone; OD, oolitic dolomite; LS, limestone; CS, calcareous siltstone; SM, sandy mudstone; D, dolomite; BM, bauxitic mudstone; MS, mudstone; DL, dolomitic limestone; G, glutenite; SH, shale; S, sand-shale; CB, coalbed; A, arkose; FQS, feldspar-quartz sandstone.

### 3. Multilayer Perceptron Neural Network and Data

#### 3.1. Artificial Neural Network (ANN)

The idea of an artificial neural network (ANN) is the mathematical equivalent of the human biological neural system [27]. A comparison of a simplified biological neuron system and a multilayer neural network model is demonstrated in Figure 3. The input layer ( $x$ ) of the network is linked to the hidden layer by a defined weight factor ( $W$ ). The hidden layer is characterized by several computation connectionist neurons ( $\sigma$ ). The numbers of hidden layers and neurons are decided by trial and error [28]. The output layer ( $y$ ) is associated with a “purelin” transfer function, and the hidden layers are assigned with an activation function. Both the hidden layers and the output layer of the network are all connected together in a forward direction. The MLP network is solely dependent on the adjustment of the weight factors between the layers [29].

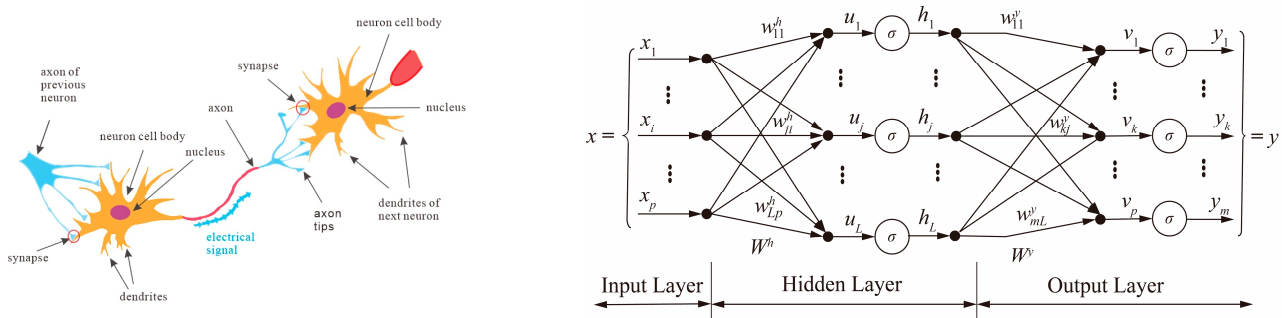


Figure 3. Analogy of a biological neuron (left) and a multilayer neural network (right).

To lower the computational time and avoid model network complexity, no more than two hidden layers in an ANN structure should be employed. The training, validation, and test data are generated by a probability distribution over the datasets. The learning data subsets (70% of total samples) are employed to adjust the weights of the trained neural network [30]. Moreover, the testing data subsets (15%) are utilized for testing the final solution to assess the performance of the deterministic neural network. The remaining 15% validation data subsets are used to minimize the overfitting of the model. When training the MLP model, we can compute some error measure on the training set; thus far, what we have described is an optimization problem [31].

#### 3.2. The Architecture of a Multilayer Perceptron Neural Network (MLP)

The multilayer perceptron neural network is a subset of machine learning, where the artificial neural network comes in relation to it. It solves all the complex problems with the help of algorithms and their processes. The idea is that the additional level of abstraction improves the capability of the network to be generalized to unseen data and hence outperforms traditional ANNs on data outside of the network training set. The learning process is deep because the structure of artificial neural networks consists of multiple input, output, and hidden layers. Each layer contains units that transform the input data into information that the next layer can use for a certain predictive task [32].

In this work, the multilayer perceptron neural network is applied to the problems of source discrimination of mine water inrush. The multilayer perceptron neural network further exploits the power of ANNs by relying on the network itself to identify, extract, and combine the inputs into abstract features, which contain much more pertinent information of the problem, that is, predicting the output, as illustrated in Figure 4. Every neuron accepts inputs from neurons on the previous layer based on linear or nonlinear activation functions (e.g., ReLU). A total of 149 sets of data with six ions ( $\text{Na}^+ + \text{K}^+, \text{Ca}^{2+}, \text{Mg}^{2+}, \text{Cl}^-, \text{SO}_4^{2-}, \text{and HCO}_3^-$ ) are delivered from the input layer to the output layer, where the output layer corresponds to the expectations to be predicted, which are the surface water,

pore water of Quaternary limestone, sandstone water of Permian limestone, karst water of Carboniferous limestone and karst water of Cambrian limestone (Figure 4).

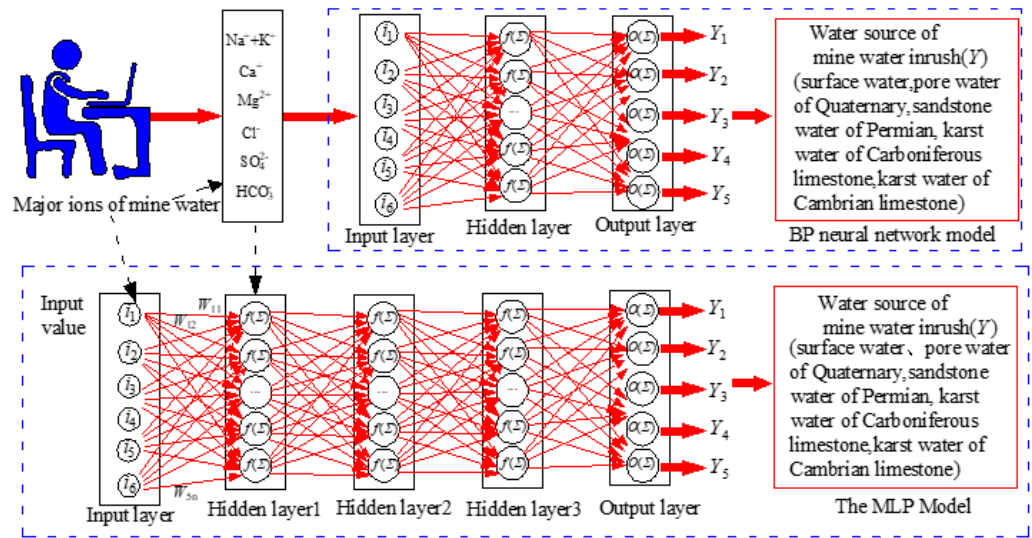


Figure 4. The BP neural network compared with the MLP neural network.

The multilayer perceptron neural network with two hidden layers and one output layer is shown in Figure 5. Every layer constitutes a module through which one can back-propagate gradients. At every layer, we compute the total input  $i$  to every unit first, which is a weighted sum of the outputs of the units in the layer below. Then, a nonlinear function  $f$  is applied to  $i$  to obtain the output of the unit. For the sake of simplicity, the bias terms are omitted. The nonlinear functions in the hidden layer using the ANN include the rectified linear unit (ReLU)  $f(z) = \max(0, z)$ . At the output layer, softmax is used to calculate the probability of the water source. This has been commonly used in recent years.

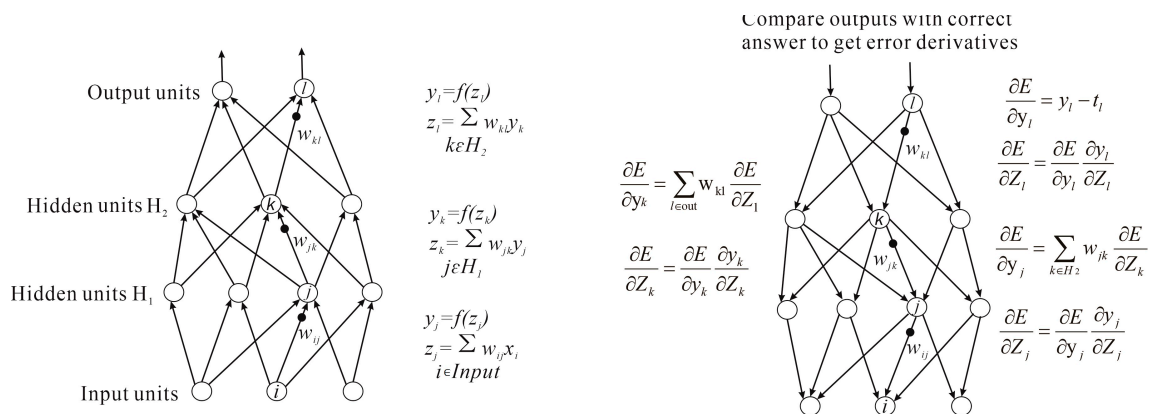


Figure 5. Multilayer neural networks and backpropagation (Left: The equations used for computing the forward pass. Right: The equations used for calculating the backward pass). Compare outputs with correct answers to obtain error derivatives.

At every hidden layer, we calculate the error derivative with respect to the output of every unit, which is a weighted sum of the error derivative with respect to the total inputs to the units in the layer above. Then, we convert the error derivative with respect to the output into the error derivative with respect to the input by multiplying it by the gradient of  $f$ . At the output layer, the error derivative with respect to the output of a unit is calculated by differentiating the cost function. This gives  $y_l - t_l$  if the cost function for unit  $l$  is  $1/2 \times (y_l - t_l)^2$ , where  $t_l$  is the target value. Once  $\partial E/\partial z_k$  is known, the error derivative for the weight  $w_{jk}$  on the connection from unit  $j$  in the layer below is just  $y_j \partial E/\partial z_k$ .

The Python machine-learning library Keras, with a TensorFlow backend and GPU acceleration, is used to train the MLP. TensorFlow 2.0 is an end-to-end open-source platform for machine learning. It has a comprehensive, flexible ecosystem of tools, libraries, and community resources that allows researchers to push the state of the art in MLP for developers to easily build and deploy deep-learning-powered applications. The model parameters of the intelligent evaluation of the MLP model are shown in Table 1.

**Table 1.** Model parameters of the intelligent evaluation of the MLP model.

Number	Parameter	Value
1	Type of model	Sequential model
2	The number of neurons in the input layer	6
3	The number of hidden layers and neurons	3.5
4	The number of neurons in the output layer	5
5	Activation function of the hidden layer	ReLU
7	Activation function of the output layer	Softmax
8	Epoch	200
9	Learning rate	0.01
10	Optimizer function	Adam
11	Batch size	10
12	Dropout rate	0.5
13	Error limitation	$1 \times 10^{-4}$
14	Momentum coefficient	$\eta = 0.8$

### 3.3. Data

In the Pingdingshan coal mine, 149 water samples were collected during the period of mining from January 2006 to December 2020. All samples were sent to the laboratory as soon as possible for further analysis. Each sample was collected into a pre-cleaned 550 mL high-density polyethylene plastic bottle, and all the bottles and covers were repeatedly washed 3–5 times with the original water before sampling. After the water samples were taken back to the laboratory, they were processed chemically to inhibit the redox reaction and any biochemical action. They were filtered through a 0.45  $\mu\text{m}$  microporous membrane. The water sample for cation analysis was acidified with 1:4  $\text{HNO}_3^-$  to adjust the pH to 2, and cations were measured using inductively coupled plasma atomic emission spectrometry (ICP-AES; IRIS Intrepid IIXSP, Thermo Fisher Scientific, Waltham, MA, USA). Ion chromatography (ICS-1500, Dion, Long Beach, CA, USA) was used to measure the anion ( $\text{Cl}^-$ ,  $\text{SO}_4^-$ , and  $\text{HCO}_3^-$ ) concentrations by double-indicator titration. Ultrapure water was used for the experiment, and the vessels were soaked in 6 mol/L  $\text{HNO}_3^-$  solution for 24 h, then washed and dried with deionized water [33]. To ensure accuracy, conservation calculations (positive and negative charges) were performed for each experimental test.

Due to the large amount of data, the statistical information of the data is given in Table 2. Table 2 clearly shows that there is a large difference between the max. and min. values. Therefore, data standardization is about ensuring that the data are internally consistent, that is, that each data type has the same content and format. Standardized values are useful for tracking data that is not easy to compare. The raw data is normalized individually according to Equation (1):

$$Z_{ij} = (x_{ij} - \text{mean}(x_j)) / \text{std}(x_j) \quad (1)$$

where the subscript  $i$  represents the row of the data matrix, the subscript  $j$  represents the column of the data matrix,  $Z_{ij}$  represents the data after standardization,  $x_{ij}$  represents the source data, and the symbol  $\text{std}$  represents the standard deviation of the related data [10].

**Table 2.** Statistics of the chemical composition of different aquifers' water in the Pingdingshan coalfield.

	PH	TDS	Na <sup>+</sup> + K <sup>+</sup>	Ca <sup>2+</sup>	Mg <sup>2+</sup>	Cl <sup>-</sup>	SO <sub>4</sub> <sup>2-</sup>	HCO <sub>3</sub> <sup>-</sup>
<b>Surface water</b>								
Average value	7.60	257.47	46.36	79.94	15.41	30.67	112.98	242.98
Standard deviation	0.23	121.44	44.15	35.22	7.48	26.31	92.67	91.05
Coefficient of variation	0.03	0.47	0.95	0.44	0.49	0.86	0.82	0.37
<b>Quaternary pore water</b>								
Average value	7.44	477.54	28.61	154.22	22.54	53.77	172.58	339.98
Standard deviation	0.28	246.22	26.64	75.19	16.86	41.17	145.09	128.02
Coefficient of variation	0.04	0.52	0.93	0.49	0.75	0.77	0.84	0.38
<b>Permian sandstone water</b>								
Average value	7.98	208.04	365.73	48.87	21.69	51.22	177.54	718.75
Standard deviation	0.58	338.16	315.92	79.17	38.58	32.79	297.86	593.50
Coefficient of variation	0.07	1.63	0.86	1.62	1.78	0.64	1.68	0.83
<b>Carboniferous limestone karst water</b>								
Average value	7.49	275.33	120.58	78.44	23.09	49.87	117.17	414.10
Standard deviation	0.39	153.40	147.22	45.18	15.35	32.28	111.63	217.18
Coefficient of variation	0.05	0.56	1.22	0.58	0.66	0.65	0.95	0.52
<b>Cambrian limestone karst water</b>								
Average value	7.32	311.75	114.47	86.17	27.53	64.01	158.24	370.90
Standard deviation	0.58	229.74	110.33	81.43	15.09	42.60	178.25	140.06
Coefficient of variation	0.08	0.74	0.96	0.95	0.55	0.67	1.13	0.38

In the datasets, the label column is categorical data (string values). These labels have no specific order of preference, and since the data are string labels, the artificial neural network model cannot work on such data directly [34]. One approach to solving this problem is label encoding, where we assign a numerical value to these labels [35]. For example, the surface water and pore water of the Quaternary were mapped to 0 and 1. However, this can add bias in our model, as it will start giving greater preference to the pore water of the Quaternary parameter as  $1 > 0$ , while ideally, both labels are equally important in the datasets [36–38]. To address this issue, we will use the one-hot encoding technique, which will create a binary vector of length 5. For example, the surface water, labeled as “0”, is encoded with a binary vector of [0,0,0,0,1], as shown in Table 3.

**Table 3.** One-hot encoding (the first column is the groundwater type (label column); 0 represents the surface water; 1 represents the pore water of the Quaternary limestone; 2 represents the karst water of the Carboniferous limestone; 3 represents the sandstone water of the Permian limestone; and 4 represents the karst water of the Cambrian limestone. The second column is the one-hot code).

Natural Number	One-Hot Encoding
0	0,0,0,0,1
1	0,0,0,1,0
2	0,0,1,0,0
3	0,1,0,0,0
4	1,0,0,0,0

### 3.4. Hydrochemical Analysis

The chemical characteristics of each aquifer were ascertained from the 149 water samples by analysis. The average pH of the Cambrian limestone water is 7.32, which is a neutral value. The average pH values of the Carboniferous limestone, Permian sandstone, Quaternary pore water, and surface water samples were 7.49, 7.57, 7.98, and 7.60, respectively.



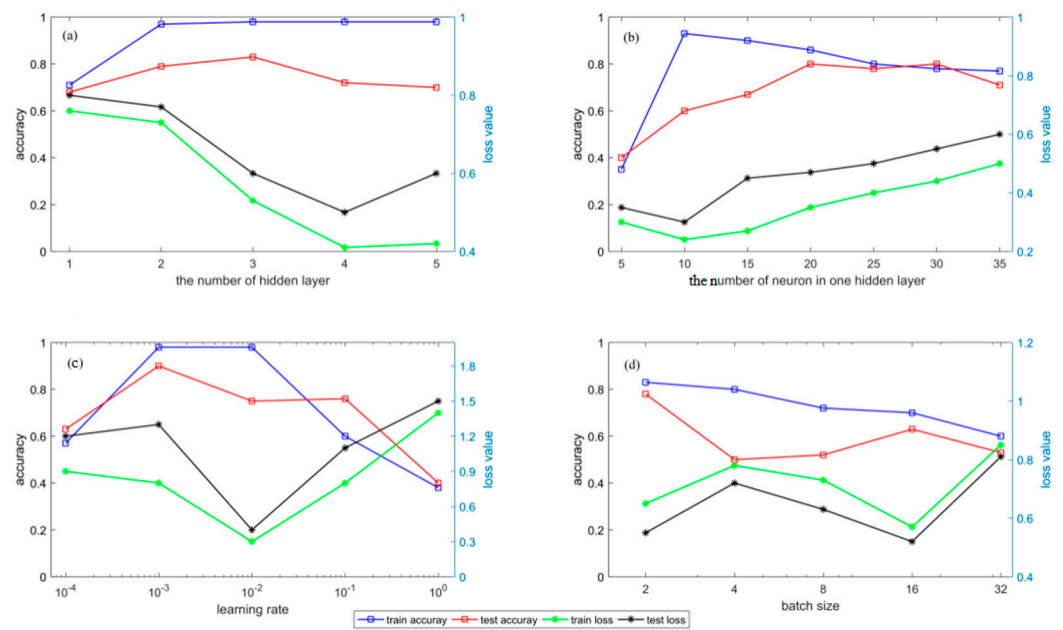
The average total dissolved solid (TDS) value of the surface water was 257.47 mol/L. The ions  $\text{SO}_4^{2-}$  and  $\text{Cl}^-$  had low mass concentrations and  $\text{HCO}_3^-$  had the greatest mass concentration among the anions. The mass concentrations of the cations were in the following order:  $\text{Ca}^{2+} > \text{Na}^+ + \text{K}^+ > \text{Mg}^{2+}$ . Except for  $\text{Mg}^{2+}$ , the coefficients of variation for all the ions in different aquifers ranged significantly. The average concentration of  $\text{Ca}^{2+}$  was larger than that of  $\text{Na}^+ + \text{K}^+$  and  $\text{Mg}^{2+}$  in Quaternary water, and that of  $\text{HCO}_3^-$  was twice greater than that of  $\text{SO}_4^{2-}$  and three times greater than that of  $\text{Cl}^-$ . Except for  $\text{Mg}^{2+}$ , all of the aquifer's standard deviations were significant. The mean TDS value of Permian sandstone water, Carboniferous limestone, and Cambrian limestone water were 208.04 mol/L, 275.33 mol/L, and 311.75 mol/L, respectively. The cations were mainly  $\text{Na}^+ + \text{K}^+$  and the anions were mainly  $\text{SO}_4^{2-}$  and  $\text{HCO}_3^-$  in Carboniferous limestone and Cambrian limestone water. The standard deviation of  $\text{Na}^+ + \text{K}^+$ ,  $\text{SO}_4^{2-}$ , and  $\text{HCO}_3^-$  were greater than the other ions in Permian sandstone, Carboniferous limestone, and Cambrian limestone water. Most of the coefficients of variation in all the aquifers were less than 1.

#### 4. Results and Discussion

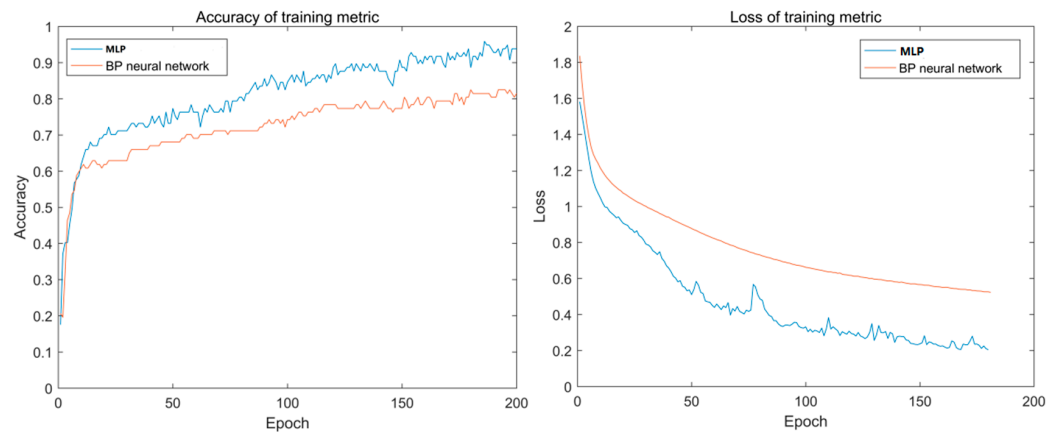
The learning rate, batch size, and learning rate are three important parameters of the established MLP model. Among these parameters, the learning rate, which controls the learning progress of the model, is used to scale the magnitude of parameter updates during a gradient descent. The batch size refers to the number of training examples utilized in one iteration. It can also have a significant impact on the model's performance. The process is programmed to continue until the model error is suitably minimized. The model output depends on these input parameters.

The parametric sensitivity analysis was conducted in this study. As shown in Figure 6, one of these parameters remain changed while the other parameters remain fixed, and the model accuracy and loss change with these parameters. When the number of hidden layers is set to three, the model test sample accuracy is the highest, and the loss value decreases and continues to decrease (Figure 6a). When the number of neurons in one hidden layer is set to 10, the accuracy of the model is the highest, and the loss value is reduced to the lowest value (Figure 6b). The learning rate is an important parameter for the MLP model training. It is also sensitive for accuracy and loss. As shown in Figure 6c, if it continually increases, the model experiences difficulty in converging, and the loss value rises sharply. When the learning rate is set to  $10^{-2}$ , the accuracy of the training and test samples reaches the maximum value, and the loss value reaches the minimum value. The batch size should also be set properly for the training. The accuracy of the training sample exhibits a downward trend, and the trend of the loss changes becomes complicated. Therefore, it is a better option for the batch size to be set to 16 (Figure 6d). In short, based on the output results above, training duration, and other factors, the number of hidden layers, the number of neurons in one hidden layer, the learning rate, and the batch size were set to 3, 25,  $10^{-2}$ , and 16, respectively.

The accuracy and loss of the training metrics of the multilayer perceptron neural network (MLP) and the BP neural network have been compared, which are drawn by Matplotlib (Figure 7). In Figure 7, the abscissa axis represents the times of forward calculation and backpropagation, and the vertical axis represents the accuracy or loss. The blue curve represents the accuracy and loss of the training metric of the deep neural network, and the red curve represents the accuracy and loss of the training metric of the BP neural network. It can be seen that the accuracy of the training metric of MLP is higher than that of the BP neural network, and the loss of the training metric of the deep neural network is lower than that of the BP neural network. More precisely, the accuracy of the MLP model is 94.03%. It means that the MLP does a better job than the BP neural network in the source discrimination of mine water.



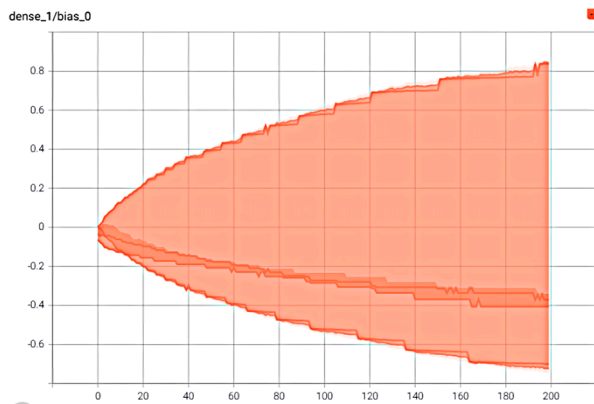
**Figure 6.** Influence of the different parameters on the MLP model ((a) the accuracy and loss varies with different number of hidden layer; (b) the accuracy and loss varies with different number of neuron in hiddern layer; (c) the accuracy and loss varies with different learning rate; (d) the accuracy and loss varies with different batch size).



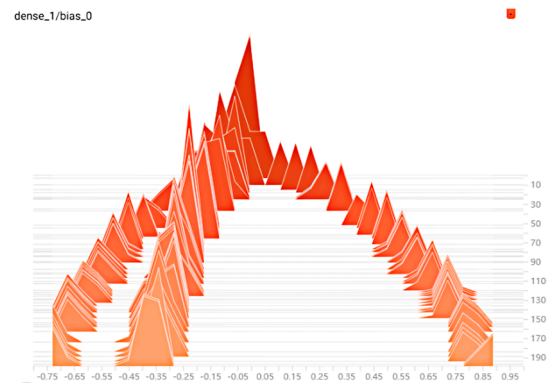
**Figure 7.** Accuracy and loss of the training metrics.

The data source of distributions is the same as that of the histogram, which is shown in different formats (Figure 8). The distribution of weights and bias of the first layer are shown in Figure 8a,c. The abscissa represents the training times, and the ordinate represents the range of weights. It shows the range of weight values in the training process as a whole, which is constrained to learn of the layer by optimizing its weights so that the layer truly “eats up” many errors. Almost the same number of weights have values between  $-0.8$  to  $0.8$ . Some weights have slightly smaller or higher values, but they might not be using their full potential. In comparison, this simply looks like the weights have been initialized using a uniform distribution with a zero mean and a value range of  $-0.8$  to  $0.8$  (Figure 8c,d). The histogram of the layer forms a bell-curve-like shape. The values are centered around a specific value, but they may also be greater or smaller than that. Each slice in the histogram visualizer displays a single histogram. The slices are organized by steps; older slices are further “back” and darker, while newer slices are close to the foreground and lighter in color. The  $y$ -axis on the right shows the step number. Most values appear close to the mean of  $0$ , but values do range from  $-1.3$  to  $1.2$ . With increasing training times, the color of the

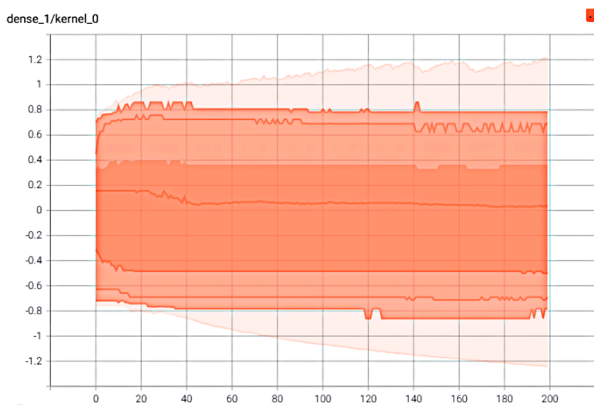
curves gradually becomes lighter from back to front. There are many slices in Figure 8b,d, and each slice represents the frequency of the weight in the distribution of weights.



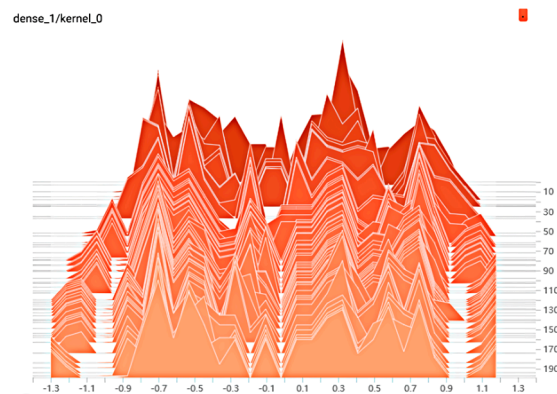
(a) The distribution of the bias



(b) The histogram of the bias



(c) The distribution of the weights



(d) The histogram of the weights

**Figure 8.** The distribution and histogram of the weights and bias.

The unitless numbers’ accuracy and loss show how well the classifier fits the validation training data. A perfect fit is represented by a loss value of 0. The fit is more precise the further the accuracy is from 0. For the batches finished inside each epoch as well as for the finished epochs, separate loss charts are supplied. The loss values shown in the epoch\_loss and epoch\_val\_loss plots should, in theory, rapidly decline during the first few epochs and then converge toward zero as the number of epochs rises. The “loss” number shown in the training model dialog as a model is being trained is matched by the epoch\_val\_loss plot.

The probability is based on the fraction of correctly predicted values to the total number of values predicted to be in a class, which is calculated by the softmax function (Figure 9). Ten newly collected mine water samples are inputted into the trained MLP model to test its accuracy of prediction. In comparison, the data are also inputted into the BP model. The predicted and the real values are shown in Table 4. From this table, we can see that nine water samples have been predicted correctly with the MLP model, and one water sample has been predicted incorrectly. With the BP model, four water samples were predicted correctly, and six water samples were predicted incorrectly. The reason for the different prediction accuracy is that two hidden layers can improve the performance of the ANN model.

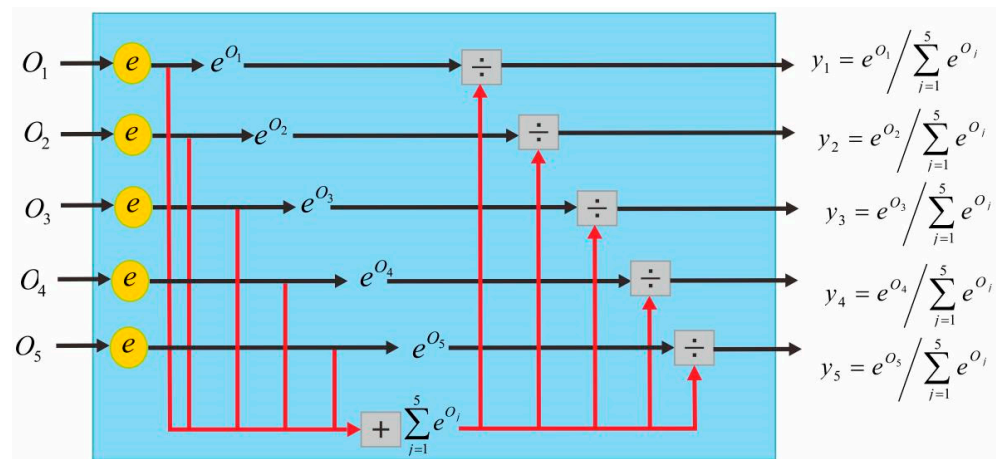


Figure 9. Principle of softmax function for multiclass classification.

Table 4. Probability of sources of mine water by the BP and MLP methods.

Sample	Probability of Surface Water (I)	Probability of Pore Water of Quaternary (II)	Probability of Sandstone Water of Permian (III)	Probability of Karst Water of Carboniferous Limestone (IV)	Probability of Karst Water of Cambrian Limestone (V)	Predicted by MLP	Predicted by BP	Real Source of Mine Water
NO.1	0% (0%)	0% (0%)	0% (0%)	1% (99%)	99% (1%)	V	IV	V
NO.2	0% (0%)	0.5% (0.05%)	99% (99.8%)	0.5% (0.05%)	0% (0.1%)	III	III	III
NO.3	0% (0%)	0% (0%)	0.5% (63%)	99% (36%)	0.5% (1%)	IV	III	IV
NO.4	0% (0.6%)	0% (1.09%)	99% (57.7%)	0.7% (0%)	0.3% (40.61%)	III	III	III
NO.5	0% (1.53%)	0% (3.3%)	99% (89.1%)	0.7% (0.03%)	0.3% (6.04%)	III	III	III
NO.6	0% (0.185%)	0.8% (4.45%)	99% (46.12%)	0.2% (0%)	0% (49.245%)	III	V	III
NO.7	0% (4.22%)	0% (1.38%)	0% (62.3%)	95.8% (0.23%)	4.2% (19.45%)	IV	III	IV
NO.8	84.7% (0%)	5.8% (0%)	2.34% (44.76%)	1.15% (54.68%)	5.96% (0.56%)	I	IV	I
NO.9	0% (0%)	0.5% (0%)	99% (34.12%)	0% (65.81%)	0.5% (0.07%)	III	IV	IV
NO.10	0.03% (0.14%)	52.43% (3.16%)	41.39% (84.58%)	5.89% (0%)	0.26% (1.21%)	II	III	II

### 5. Conclusions and Outlook

In the research reported here, on the basis of the water chemistry characteristics, an MLP model for mine-water source discrimination is established under the Keras framework in the Pingdingshan mining area. To avoid the complexity structure and to obtain accurate results, two hidden layers (with 10 neurons in each hidden layer) for MLP models are fit for completing the prediction process with a better performance. Based on the trained model, 10 newly collected mine water samples are tested to determine the precision of the model, and 9 samples of mine water were predicted correctly. The MLP model presented here provides significant guidance for the discrimination of mine water.

It is also recommended that the high predictive accuracy, combined with very low computational costs in the execution of the full framework, makes this very well suited for discriminating the source of mine water. It appears that the research approach in data-driven models of water source discrimination can be useful for geologists and engineers in the mining industry.

The limitation of the research is that the amount of data is not large enough, and the number of hidden layers is shallow. Therefore, compared with the deep-learning model, the learning ability of the trained model would be a little weaker. In the end, such a research approach can be implemented using more data points for various purposes, such as for predictive model development, the optimization of problems that can have various applications in science, engineering, and geohazard forecasting.

**Author Contributions:** Z.Y. designed the model. M.W. and J.Z. collected the mine water samples. B.Z. and X.W. performed model training and data analysis. M.W. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Open Project of Key Laboratory from the State Key Laboratory of Development and Comprehensive Utilization of Coking Coal Resources (grant number 41040220201308).

**Data Availability Statement:** The data used in the research is available upon request.

**Acknowledgments:** We thank China Pingmei Shenma Holding Group Co., Ltd., for providing the platform of data sampling. The authors are also grateful to the five anonymous reviewers for their constructive comments and suggestions.

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

## References

1. Jiang, C.; An, Y.; Zheng, L.; Huang, W. Water source discrimination in a multiaquifer mine using a comprehensive stepwise discriminant method. *Mine Water Environ.* **2021**, *40*, 442–455. [[CrossRef](#)]
2. Zuo, R.; Xiong, Y.; Wang, J.; Carranza, E.J.M. Deep learning and its application in geochemical mapping. *Earth-Sci. Rev.* **2019**, *192*, 1–14. [[CrossRef](#)]
3. Shah, S.A.; Jehanzab, M.; Lee, J.-H.; Kim, T.-W. Exploring the factors affecting streamflow conditions in the Han River Basin from a regional perspective. *KSCE J. Civ. Eng.* **2021**, *25*, 4931–4941. [[CrossRef](#)]
4. Shah, S.A.; Lakho, G.M.; Keerio, H.A.; Sattar, M.N.; Hussain, G.; Mehdi, M.; Vistro, R.B.; Mahmoud, E.A.; Elansary, H.O. Application of drone surveillance for advance agriculture monitoring by Android application using convolution neural network. *Agronomy* **2023**, *13*, 1764. [[CrossRef](#)]
5. Wu, Q.; Mu, W.; Xing, Y.; Qian, C.; Shen, J.; Wang, Y.; Zhao, D. Source discrimination of mine water inrush using multiple methods: A case study from the Beiyangzhuang Mine, Northern China. *Bull. Eng. Geol. Environ.* **2017**, *78*, 469–482. [[CrossRef](#)]
6. LeCun, Y.; Bengio, Y. Hinton Deep learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
7. Zhou, Z.Q.; Li, S.C.; Li, L.P.; Shi, S.S.; Xu, Z.H. An optimal classification method for risk assessment of water inrush in karst tunnels based on gray system theory. *Geomech. Eng.* **2015**, *8*, 631–647. [[CrossRef](#)]
8. Barral, N.; Husillos, R.; Castillo, E.; Cánovas, M.; Lam, E. Hydrochemical evolution of the Reocín Mine filling water (Spain). *Environ. Geochem. Health* **2021**, *43*, 5119–5134. [[CrossRef](#)]
9. Barral, N.; Husillos, R.; Castillo, E.; Cánovas, M.; Lam, E.J.; Calvo, L. Volumetric quantification and quality of water stored in a mining lake: A case study at Reocín mine (Spain). *Minerals* **2021**, *11*, 212. [[CrossRef](#)]
10. Barral, N.; Maleki, M.; Madani, N.; Cánovas, M.; Husillos, R.; Castillo, E. Spatio-temporal geostatistical modelling of sulphate concentration in the area of the Reocín Mine (Spain) as an indicator of water quality. *Environ. Sci. Pollut. Res.* **2021**, *29*, 86077–86091. [[CrossRef](#)]
11. Yang, Y.Y.; Xu, Y.S.; Shen, S.L. Mining-induced geo-hazards with environmental protection measures in Yunnan, China: An overview. *Bull. Eng. Geol. Environ.* **2015**, *74*, 141–150. [[CrossRef](#)]
12. Yin, S.; Zhang, J.; Liu, D. A study of mine water inrushes by measurements of in situ stress and rock failures. *Nat. Hazards* **2016**, *79*, 1961–1979. [[CrossRef](#)]
13. Juncosa, R.; Delgado, J.; Cereijo, J.L.; Muñoz, A. Analysis of the reduction processes at the bottom of Lake Meirama: A singular case of lake formation. *Environ. Monit. Assess.* **2023**, *195*, 1004. [[CrossRef](#)] [[PubMed](#)]
14. Juncosa, R.; Delgado, J.; Cereijo, J.L.; García, D.; Muñoz, A. Comparative hydrochemical analysis of the formation of the mining lakes of As Pontes and Meirama (Spain). *Environ. Monit. Assess.* **2018**, *190*, 526. [[CrossRef](#)] [[PubMed](#)]
15. Qian, J.; Wang, L.; Ma, L.; Lu, Y.; Zhao, W.; Zhang, Y. Multivariate statistical analysis of water chemistry in evaluating groundwater geochemical evolution and aquifer connectivity near a large coal mine, Anhui, China. *Environ. Earth Sci.* **2016**, *75*, 747. [[CrossRef](#)]
16. Ma, D.; Miao, X.; Bai, H.; Huang, J.; Pu, H.; Wu, Y.; Zhang, G.; Li, J. Effect of mining on shear sidewall groundwater inrush hazard caused by seepage instability of the penetrated karst collapse pillar. *Nat. Hazards* **2016**, *82*, 73–93. [[CrossRef](#)]
17. Gu, H.; Ma, F.; Guo, J.; Li, K.; Lu, R. Assessment of water sources and mixing of groundwater in a coastal mine: the Sanshandao gold mine, China. *Mine Water Environ.* **2017**, *37*, 351–365. [[CrossRef](#)]
18. Richards, B.A.; Lillicrap, T.P.; Beaudoin, P.; Bengio, Y.; Bogacz, R.; Christensen, A.; Clopath, C.; Costa, R.P.; de Berker, A.; Ganguli, S.; et al. A deep learning framework for neuroscience. *Nat. Neurosci.* **2019**, *22*, 1761–1770. [[CrossRef](#)]
19. Dauphin, Y.N.; Pascanu, R.; Gulcehre, C.; Cho, K.; Ganguli, S.; Bengio, Y. Identifying and attacking the saddle point problem in high-dimensional non-convex optimization. *Adv. Neural Inf. Process. Syst.* **2014**, *2*, 2933–2941.
20. Tompson, J.; Goroshin, R.; Jain, A.; LeCun, Y.; Bregler, C. Efficient Object Localization Using Convolutional Networks. In Proceedings of the 2015 Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 648–656.
21. Srivastava, N.; Hinton, G.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R. Dropout: A simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* **2014**, *15*, 1929–1958.
22. McCulloch, W.S.; Pitts, W. A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biol.* **1990**, *5*, 99–115. [[CrossRef](#)]
23. Huang, P.; Wang, X. Piper-PCA-Fisher recognition model of water inrush source: A case study of the Jiaozuo mining area. *Geofluids* **2018**, *2018*, 9205025. [[CrossRef](#)]
24. Miah, M.I.; Zendejboudi, S.; Ahmed, S. Log data-driven model and feature ranking for water saturation prediction using machine learning approach. *J. Pet. Sci. Eng.* **2020**, *194*, 107291. [[CrossRef](#)]

25. Jiang, C.; Zhu, S.; Hu, H.; An, S.; Su, W.; Chen, X.; Li, C.; Zheng, L. Deep learning model based on big data for water source discrimination in an underground multiaquifer coal mine. *Bull. Eng. Geol. Environ.* **2021**, *81*, 26. [[CrossRef](#)]
26. Ji, Y.; Dong, D.L.; Gao, J.; Wei, Z.L.; Ding, J.; Hu, Z.Q. Source discrimination of mine water inrush based on spectral data and EGA-PNN model: A case study of Huangyuchuan mine. *Mine Water Environ.* **2022**, *41*, 583–593. [[CrossRef](#)]
27. Zendejboudi, S.; Rezaei, N.; Lohi, A. Applications of hybrid models in chemical, petroleum, and energy systems: A systematic review. *Appl. Energy* **2018**, *228*, 2539–2566. [[CrossRef](#)]
28. Yang, Y.; Yue, J.; Li, J.; Yang, Z. Mine water inrush sources online discrimination model using fluorescence spectrum and CNN. *IEEE Access* **2018**, *6*, 47828–47835. [[CrossRef](#)]
29. Wang, Y.; Shi, L.; Wang, M.; Liu, T. Hydrochemical analysis and discrimination of mine water source of the Jiaojia gold mine area, China. *Environ. Earth Sci.* **2020**, *79*, 123. [[CrossRef](#)]
30. Yan, B.; Ren, F.; Cai, M.; Qiao, C. Bayesian model based on Markov chain Monte Carlo for identifying mine water sources in Submarine Gold Mining. *J. Clean. Prod.* **2020**, *253*, 120008. [[CrossRef](#)]
31. Zheng, Q.S.; Wang, C.F.; Liu, W.T.; Pang, L.F. Evaluation on development height of water-conducted fractures on overburden roof based on nonlinear algorithm. *Water* **2022**, *14*, 3853. [[CrossRef](#)]
32. Li, G.; Wang, Z.; Ma, F.; Guo, J.; Liu, J.; Song, Y. A case study on deformation failure characteristics of overlying strata and critical mining upper limit in submarine mining. *Water* **2022**, *14*, 2465. [[CrossRef](#)]
33. Duan, X.L.; Ma, F.S.; Gu, H.Y.; Guo, J.; Zhao, H.J.; Liu, G.W.; Liu, S.Q. Identification of mine water sources based on the spatial and chemical characteristics of Bedrock Brines: A case study of the Xinli gold mine. *Mine Water Environ.* **2022**, *41*, 126–142. [[CrossRef](#)]
34. Yang, Z.; Lv, H.; Wang, X.; Yan, H.; Xu, Z. Classification of Water Source in Coal Mine Based on PCA-GA-ET. *Water* **2023**, *15*, 1945. [[CrossRef](#)]
35. Qiu, M.; Shi, L.; Teng, C.; Zhou, Y. Assessment of water inrush risk using the fuzzy delphi analytic hierarchy process and grey relational analysis in the Liangzhuang coal mine, China. *Mine Water Environ.* **2017**, *36*, 39–50. [[CrossRef](#)]
36. Liu, Q.; Sun, Y.J.; Xu, Z.M.; Xu, G. Application of the comprehensive identification model in analyzing the source of water inrush. *Arab. J. Geosci.* **2018**, *11*, 189. [[CrossRef](#)]
37. Yang, W.F.; Shen, D.Y.; Ji, Y.B.; Wang, Y. Discrimination of Mine Water Bursting Source Based on Fuzzy System. *Appl. Mech. Mater.* **2012**, *13*, 873. [[CrossRef](#)]
38. Yang, B.; Yuan, J.; Duan, L. Development of a system to assess vulnerability of flooding from water in karst aquifers induced by mining. *Environ. Earth Sci.* **2018**, *77*, 91. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.