

Article



A Comparative Study of Machine Learning Models with Hyperparameter Optimization Algorithm for Mapping Mineral Prospectivity

Nan Lin^{1,2,*}, Yongliang Chen², Haiqi Liu³ and Hanlin Liu¹

- ¹ College of Surveying and Exploration Engineering, Jilin Jianzhu University, Changchun 130018, China; 952647339@163.com
- ² College of Earth Science, Jilin University, Changchun 130026, China; chenyl@jlu.edu.cn
- ³ College of Resources and Civil Engineering, Northeastern University, Shenyang 110819, China; 1900999@stu.neu.edu.cn
- * Correspondence: linnan@jlju.edu.cn; Tel.: +86-431-8456-6206

Abstract: Selecting internal hyperparameters, which can be set by the automatic search algorithm, is important to improve the generalization performance of machine learning models. In this study, the geological, remote sensing and geochemical data of the Lalingzaohuo area in Qinghai province were researched. A multi-source metallogenic information spatial data set was constructed by calculating the Youden index for selecting potential evidence layers. The model for mapping mineral prospectivity of the study area was established by combining two swarm intelligence optimization algorithms, namely the bat algorithm (BA) and the firefly algorithm (FA), with different machine learning models. The receiver operating characteristic (ROC) and prediction-area (P-A) curves were used for performance evaluation and showed that the two algorithms had an obvious optimization effect. The BA and FA differentiated in improving multilayer perceptron (MLP), AdaBoost and one-class support vector machine (OCSVM) models; thus, there was no optimization algorithm that was consistently superior to the other. However, the accuracy of the machine learning models was significantly enhanced after optimizing the hyperparameters. The area under curve (AUC) values of the ROC curve of the optimized machine learning models were all higher than 0.8, indicating that the hyperparameter optimization calculation was effective. In terms of individual model improvement, the accuracy of the FA-AdaBoost model was improved the most significantly, with the AUC value increasing from 0.8173 to 0.9597 and the prediction/area (P/A) value increasing from 3.156 to 10.765, where the mineral targets predicted by the model occupied 8.63% of the study area and contained 92.86% of the known mineral deposits. The targets predicted by the improved machine learning models are consistent with the metallogenic geological characteristics, indicating that the swarm intelligence optimization algorithm combined with the machine learning model is an efficient method for mineral prospectivity mapping.

Keywords: machine learning; bat algorithm; firefly algorithm; Youden index; P-A curve; ROC curve analysis; mineral prospectivity mapping

1. Introduction

In the information era, mineral resources information is becoming increasingly abundant. It is significant enough to construct accurate and efficient mineral prediction models and carry out quantitative mineral prospectivity mapping by data mining and artificial intelligence to exploit mineral resources. Data-driven models are commonly used in mineral prospectivity mapping, and specific mathematical models are used to quantitatively describe the statistics of potential evidence or spatial distribution to predict mineral targets [1–4]. With the rapid development of machine learning theory and technology, the toolset based on the data-driven models has been increasingly enriched. Machine learning technology can adaptively simulate the relationship between input and output and



Citation: Lin, N.; Chen, Y.; Liu, H.; Liu, H. A Comparative Study of Machine Learning Models with Hyperparameter Optimization Algorithm for Mapping Mineral Prospectivity. *Minerals* **2021**, *11*, 159. https://doi.org/10.3390/min11020159

Academic Editor: Martiya Sadeghi Received: 30 December 2020 Accepted: 30 January 2021 Published: 3 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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/). is effective in solving nonlinear mineral resource prediction. According to the expected results of the algorithm, machine learning can be divided into supervised learning [5], unsupervised learning [6,7], semi-supervised learning [8–10], and reinforcement learning [11]. Supervised learning uses known training samples to adjust the hyperparameters of the classification or regression to achieve the desired performance, including artificial neural network (ANN) [12,13], decision tree [14], adaptive neural fuzzy system [15], constrained boltzmann machine [16], support vector machine (SVM) [17–20], random forest (RF) [21,22], extreme learning machine (ELM) [23], and MaxEnt models [24].

When using a machine learning model to predict mineral targets, it is important to effectively determine the internal key hyperparameters of models, such as the network weight of ANN, fault tolerance penalty factor and kernel hyperparameter of SVM, and the number of growing trees of RF. In the practical application of the machine learning model, it is very difficult to identify the optimal value of model hyperparameters effectively. Some previous studies used the grid search or random optimization approaches to find the optimal value in predefined hyperparameter values, but frequently failed to find the global optimal hyperparameter, which was not likely to be contained in the predefined hyperparameter value [25,26]. In recent years, many optimization algorithms have been put forward and gradually used to optimize the internal hyperparameters of the machine learning model, for example, genetic algorithm (GA) [27,28], and particle swarm optimization (PSO) [29–31]. However, these classical optimization algorithms are sensitive to the initial hyperparameters setting, making it easy to fall into local optimal solutions and causing the convergence rate of the algorithm to slow down later. In addition, bat algorithm (BA) [32], firefly algorithm (FA) [33], ant colony algorithm (ACA) [34] and other emerging swarm intelligence algorithms have also been widely used to balance the global and local search ability of the algorithm and improve the model hyperparameter optimization performance.

In this study, multilayer perceptron (MLP), one-class support vector machine (OCSVM), and AdaBoost were applied to establish mineral prospectivity mapping models. In the modeling process, the key hyperparameters of different machine learning models were optimized by BA and FA. Meanwhile, the area under curve (AUC) value of the receiver operating characteristic (ROC) curve was selected as the fitness value of the optimal objective function, optimizing the number of hidden neurons of MLP and the fault tolerance penalty coefficient and kernel hyperparameter of OCSVM, and updating learning times and the weight reduction coefficient of the weak regression of AdaBoost. The AUC value of the ROC curve and the prediction/area (P/A) value of the prediction-area (P-A) curve were applied to evaluate the model performance by analyzing the prediction results of machine learning methods before and after hyperparameter optimization. The main contribution of this paper is the proposal that the performance of machine learning models combined with BA or FA can be improved in mineral prospectivity mapping.

2. Methods (Machine Learning Models and Hyperparameter Optimization)

2.1. MLP Model

MLP is a network structure commonly used in artificial neural network (ANN), which is a multilayer composite of perceptrons consisting of input layer, hidden layer, and output layer. Besides the input layer, each node of the other layers is a neuron with a nonlinear activation function [35,36]. The MLP network learning process consists of building a feature vector and passing it to the hidden layer. The result is then calculated by weight and activation function and passed to the next layer. In this process, continuous learning and adjustment only occur during the training feedback signal. In the classification process, it must be passed forward until reaching the output layer [37–39]. The output steps of MLP are as follows [40]:

The weighted sum of the input was calculated by

$$s_j = \sum_{i=1}^n (W_{ij} X_i) - \theta_j, j = 1, 2, \dots, h$$
(1)

where *n* is the number of input nodes, W_{ij} is the connection weight from *i*-th node of the input layer to *j*-th node of the hidden layer, θ_j is the bias of the *j*-th hidden node, and X_i is the *i*-th input. The output information of each hidden node was calculated by

$$S_j = sigmoid(s_j) = \frac{1}{(1 + \exp(-s_j))}, j = 1, 2, \dots, h$$
 (2)

The output function of the hidden node and the final output information passing activation function were calculated respectively by Equations (3) and (4):

$$o_k = \sum_{j=1}^h W_{jk} \cdot S_j - \theta_k, k = 1, 2, \dots, m$$
 (3)

$$O_k = sigmoid(o_k) = \frac{1}{(1 + \exp(-o_k))}, k = 1, 2, \dots, m$$
 (4)

where W_{jk} is the connection weight from the *j*-th hidden node to the *k*-th output node and θ_k is the bias of the *k*-th output node. Because the weight and bias affect the final output value, it is necessary to find the best connection weight and bias to train the MLP for equaling the actual output to the expected output as much as possible.

2.2. AdaBoost Model

In AdaBoost, the weights of the same samples are constantly updated, and the weak regressions with different weights are gathered together to form a final strong regression [41]. The weight of each weak regression is calculated according to the predicting accuracy of each sample in each training set and the overall predicting accuracy of the last training set. In addition, the distribution weight of each sample is updated and the regression results obtained from each training are weighted and summed as the final output result of the strong regression [42,43]. The specific processes of the AdaBoost algorithm are as follows [44,45]:

A weak regression algorithm and training set were defined by $\{(x_1, y_1), \dots, (x_N, y_N)\}$, and the weight vector of the training data was initialized by

$$w^{1} = \left[w_{1}^{(1)}, w_{2}^{(1)}, \cdots , w_{N}^{(1)}\right]^{T}, w_{i}^{(1)} = 1/N, i = 1, 2, \cdots, N$$
(5)

The weak regression $G_{m(x)}$ was obtained with the training set of weight vector $w^{(m)}$ distribution, and the training error of each sample x_i in the weak regression was calculated by

$$e_{i,m} = |G_m(x_i) - y_i| \tag{6}$$

The prediction error rate E_m of the weak regression $G_{m(x)}$ on the training data set was calculated by

$$E_m = p(e_{i,m} > \theta) = \sum_{i=1}^N w_{mi} I(e_{i,m} > \theta)$$
(7)

where E_m is the sum of weights of training samples when $e_{i,m} > \theta$, and θ was the predetermined threshold.

The weight coefficient α_m of $G_{m(x)}$ was calculated by

$$\alpha_m = \frac{1}{2e^{E_m}} \tag{8}$$

which represented the importance of $G_{m(x)}$ in the final prediction.

Finally, the following formula was used to update the weight distribution of each sample in the training set for the next iteration.

$$w_{m+1,i} = \begin{cases} \frac{w_{mi}}{Z_m} \times 1, e_i(t) > \theta \\ \frac{w_{mi}}{Z_m}, e_i(t) \le \theta \end{cases}$$
(9)

By iteration, the weight of the training samples with larger error in the weak regression $G_{m(x)}$ increased, while that of the training samples with smaller error remains unchanged. After weight conversion, the AdaBoost model focused on the training samples with low predictive accuracy.

The linear combination was constructed by normalization factor Z_m , calculated by

$$Z_m = \sum_{i=1}^N w_{mi} \tag{10}$$

The final strong regression was obtained by

$$G(x) = \sum_{m=1}^{M} a_m G_m(x) \tag{11}$$

2.3. OCSVM Model

The OCSVM model, a special case of the SVM model, can process unidentified data. A subset of the input space is estimated as the supporting set of the high dimension probability distribution of the input data. The samples not belonging to the supporting set and extracted from the high dimension probability distribution are identified as multivariate abnormal samples [46]. OCSVM is used to determine the minimum region boundary, Γ , making the decision function, f(x), satisfy the boundary conditions [47–49]. Γ contains at least (1 - v)m normal samples, where *m* is the number of training samples and *v* is the proportionality coefficient to control the abnormal samples in the training samples.

The algorithm first nonlinearly transforms the training sample from the input space to the regenerative Hilbert space, $\Phi : \mathbb{R}^d \to H$, where the dimension *H* is infinite, and the function f(x) is easily determined. In this space, f(x) can map the inner product function and compute the inner product of an infinite dimensional mapping space by kernel functions in the input space, $K : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$.

$$K(x,y) = \left\langle \Phi(x), \Phi(y) \right\rangle_{H} = \exp\left(-\|x-y\|^{2}/2\sigma^{2}\right)$$
(12)

The SVM model is in condition $\langle w, \Phi(x_i) \rangle_H \ge b - \xi_i$ and $\xi_i \le 0, i = 1, 2, ..., m$, maximizing the following formula:

$$\left[\frac{1}{2}\mathbf{w}_{\mathrm{H}}^{2}\right] + \left[\frac{1}{vm}\sum_{i=1}^{m}\xi_{i} - b\right]$$
(13)

where ξ_i is a slack variable that can prevent the model from overfitting. This equation is a conditional optimization, and the solution is as follows:

$$f(x) = \sum_{i=1}^{m} \alpha_i K(x_i, x) - b$$
(14)

meanwhile,

$$\sum_{i=1}^{m} \alpha_i K(x_i, x_j), j \in [1, 2, \dots, m]$$
(15)

where α_i is the Lagrange multiplier.

2.4. Bat Algorithm

BA is a heuristic search algorithm, which optimizes the search process by simulating the usage of sonar to detect prey and avoid obstacles. The fitness value is used to select the location of bats, and the iterative search process, where the less feasible solution is replaced, is simulated based on survival of the fittest. After initializing each hyperparameter, the heuristic search starts from a random position, z_i , in the *d*-dimension search space [50,51]. The bats search for prey at a fixed frequency, at different wavelengths and sounds. During prey search, the bats automatically adjust the wavelength based on the distance to the prey. In each iteration, (0 < t < T), the global search is conducted to update the flight speed and space position of each bat. The space positions of each bat are used to calculate the fitness value of the objective function, and the one corresponding to the maximum fitness value is selected as the current optimal position [52–55]. The updating formula of speed and space position is as follows:

$$f_t = f_{\min} + (f_{\max} - f_{\min}) \cdot \beta , \ v_i^{t+1} = v_i^t + (z_i^t - z^*) \cdot f_i , \ z_i^{t+1} = z_i^t + v_i^{t+1}$$
(16)

where v_i^t , v_i^{t+1} represent the flight speed of the bat *i* at *t* and *t* + 1, respectively; z_i^t , z_i^{t+1} represent the location of the bat *i* at *t* and *t* + 1, respectively; z^* represents the global optimal position; f_i is the pulse frequency of the bat *i*; β is a random number between [0,1]; (f_{\min}, f_{\max}) is the range of pulse frequency. After each iteration, the intensity and frequency of the sound are updated according to the attenuation coefficient of pulse loudness and the increased coefficient of pulse frequency.

2.5. Firefly Algorithm

FA is a novel bionic swarm intelligent optimization algorithm proposed by Krishnanand and Ghose in 2005 [56]. *n* fireflies with different initial brightness values are randomly distributed. The brightness is related to the function value of the current position. The better the position, the higher the brightness [57,58]. Each firefly looks for others that are brighter through a line of sight (known as dynamic decision domain *RDI*), forming a neighbor collection. The firefly with the highest relative brightness is chosen with the roulette probability method. The brightness, position, and dynamic decision domain are updated and reiterated to find the next suitable firefly [59]. The iterative process of the algorithm is divided into brightness update, position update, and dynamic decision domain update [60].

Brightness update depends on the fitness value, f(xi(t)), of the corresponding objective function at different positions:

$$l_i(t) = (1 - \rho)l_i(t - 1) + \gamma f(x_i(t))$$
(17)

where ρ is fluorescein volatility and γ is the fluorescein replacement rate.

Within dynamic decision domain radius, $\gamma_d^i(t)$, firefly *i* selects other fireflies with higher brightness to constitute the domain set, $N_i(t)$, and moves to firefly *j* in the domain with probability $p_{ij}(t) = \frac{l_j(t) - l_i(t)}{\sum\limits_{k \in N_i(t)} l_k(t) - l_i(t)}$.

The position updates are as follows:

$$x_i(t+1) = x_i(t) + \alpha(\frac{x_j(t) - x_i(t)}{\|x_j(t) - x_i(t)\|})$$
(18)

where, α is mobile step and $||x_i(t) - x_i(t)||$ is the European distance between fireflies.

The dynamic decision domain is equivalent to the field of vision of fireflies, which will reduce if there are too many companions. The formula is

$$r_d^i(t+1) = \min\left\{r_s, \max\{0, r_d^i(t) + \beta(n_t - |N_i(t)|)\right\}\right\}$$
(19)

where β is the updated rate of dynamic decision domain, $|N_i(t)|$ is the number of neighboring fireflies, and r_s is the radius of perception.

2.6. Construction of Mapping Mineral Prospectivity Model

The regional stratigraphy, geological formation, magmatic activity, and geochemical and remote sensing geological data were extracted after studying the basic geological data of the study area and the metallogenic condition of known mineral deposits, establishing the multi-source metallogenic information spatial database. MLP, AdaBoost, and OCSVM methods were applied and combined with BA and FA, constructing mineral resource prediction models. The ROC and P-A curves were used to evaluate the metallogenic prediction effect of the different combined models and analyze their metallogenic prediction potential. The specific process is shown in the Figure 1.



Figure 1. Calculation process of mineral potential.

3. Study Area and Information Extraction

3.1. Geological Setting and Mineral Deposits

The study area was located in the west of Qinghai province on the East Kunlun orogenic belt, which is a part of the polymetallic metallogenic belt in the East Kunlun Mountains. It has experienced a complex geological formation evolution and has good potential for polymetallic prospecting. The tectonic position in the area belongs to the Qinling-Qilian-Kulun orogenic system. It is adjacent to the Qaidam Basin in the north, and the Kunbei Deep fault zone runs through the central part of the region. The southern part is bounded by the Kunzhong fault zone and connected with the subduction complex zone of the south slope of East Kunlun [61]. The stratigraphic distribution in the area is relatively complete, with the terrestrial volcanic sedimentary formation of Late Paleozoic-Mesozoic in the north, the marine volcanic sedimentary series of Early Paleozoic Qimantage Group in the middle, and the ancient metamorphic series of the Paleoproterozoic Jinshuikou Group in the south [62]. The regional strata from new to old are: Quaternary Holocene system, Middle Neogene system and Late Pleistocene system, Triassic Elashan Formation, Carboniferous Shiguizi and Dagangou Formation, Devonian Maoniushan Formation, Ordovician-Silurian Tanjianshan Group, Jinshuikou Group and Baishahe Formation of archaeozoic era, as shown in Figure 2.



93°10'E

Figure 2. Simplified geologic map of the study area.

Regional geological formation, intrusions, and polymetallic mineralization are controlled by the NWW-EW trending deep fault, which constitutes the Kunbei fault zone throughout the study area [63]. There were strong tectono-magmatic activities in the Kunbei fault zone during the Silurian and Late Permian. The earth crust stretches depleted mantle of the environment to form the parent basaltic magma after the collision. The magmatic activities are mainly developed in different stages of the Variscan-Indosinian orogenic cycle, which are composed of Early Permian, Late Triassic, and a small amount of Early Cretaceous intermediate-acid intrusive rocks. In particular, granite intrusive rocks expose in large quantities, which are distributed in the form of batholith and stock. The intermediate-acid rock mass is mainly distributed on the north side of the Kunbei fault in the form of small scale batholith and stock, and the distribution direction is consistent with the regional tectonic direction. The base-neutral rock mass is mainly distributed in the North Kunlun magmatic rock zone in the form of large-scale batholith. Regional mineralization occurs near the contact zone between different strata and intrusive rocks of different periods [64,65]. The metallogenic geological conditions in the study area are superior, and more than 10 mineral deposits have been discovered so far, including the Xiarihamu large magmatic Cu-Ni sulfide deposit, the upper Lalinggaoli River skarn-porphyty Mo-polymetallic deposit, and the lower Lalinggaoli River skarn Fe-Cu deposit. It is one of the important mineral exploration areas in the west of East Kunlun and is also a key area of concern for geological researchers.

It is concluded that the controlling factors of polymetallic mineralization in this area are determined mainly by analyzing the metallogenic characteristics of the existing mineral deposits, combined with previous research results [23,66]: (1) NWW-EW trending deep faults provide sufficient active conditions for diagenesis and mineralization. (2) Different types of skarn, hornfels, and keratinized rocks are formed by the metasomatism of the Indosinian and Yanshanian intrusions and strata contact zones of different periods, among which the skarn type mineralization related to contact metasomatism is developed and provides favorable conditions for Cu-Ni sulfide mineralization and skarn-porphyry polymetallic mineralization. (3) Magmatic intrusive complexes are formed in Indosinian, and their parent basaltic magma provides metal sources for regional polymetallic mineralization occurring between the parent basaltic magma and acidic crustal components and/or partially melted crust, provide metal and heat sources for skarn and skarn-porphyry polymetallic mineralization.

3.2. Metallogenic Information Extraction

The metallogenic information data obtained included mainly regional stratigraphy, geological formation, and geochemical and remote sensing geological data through collection and field investigation; the regional stratigraphy and geological formation data were from the regional geological report completed by the Qinghai Geological Survey Institute. The geochemical data were derived from a recent digital geological project in the study area, and remote sensing geological information was derived from tectonic and mineralization alteration information extracted from remote sensing images. Analyzing the results of controlling factors of polymetallic mineralization, the existing metallogenic information of the study area was combined. Then the metallogenic information data were divided into three categories, namely, metallogenic geological background, remote sensing geological, and geochemical element anomaly information. These data were extracted from geological and mineral data, remote sensing images, and stream sediments, respectively. As a result, a total of 17 different metallogenic information were extracted. This information included the key and favorable factors for mineralization in the study area. With the GIS software platform, the study area was divided into 18,600 grid cells, each with a size of 0.3039×0.3041 km², considering the scale of geological background data, sampling density of 1:50,000 water system sediments of geochemical elements, and the cell area of remote sensing geological anomaly information, meanwhile guaranteeing that the grid numbers of the horizontal and vertical direction are integers. Seventeen areas of multi-source metallogenic information were converted to the same projection coordinate system and rasterized on the same scale as the grid cell, constructing the multi-source metallogenic information spatial database.

3.2.1. Metallogenic Geological Background Information

The magmatic intrusions in the region are intense and frequent, and the intermediateacid intrusions related to mineralization can be preliminarily divided into the Middle Triassic, the Late Triassic, and the Early Cretaceous. The rock types are monzogranite, granodiorite, tonalite, quartz diorite, syenogranite, and moyite. Analyzing the regional geological data and the distribution of known minerals, three intermediate-acid intrusions and the lithological information of four geological masses, including carbonatite of the Tanjianshan Group and Dagangou Formation, monzogranite, granodiorite and quartz diorite of Middle and Late Triassic, were stored as vector information and selected as metallogenic geological background information (see Figure 2).

3.2.2. Element Concentration Anomaly Information

Geochemical data were from four 1:50,000 stream sediment geochemical measurement results in the study area, with a 1170 km² total sampling area and 8282 samples collected, with an average sampling density of 7.1/km². During the field sampling, some sample points were added, and unreasonable points (especially the eolian sand cover area) were eliminated to better control the catchment area. The sampling points were evenly distributed in the high-level water system. In addition, the samples were tested by Qinghai Rock and Mineral Analysis and Research Center, where the concentration values of 16 geochemical indexes were tested in total. The processing, testing and analysis of the samples all met the relevant specification requirements and quality standards.

The evolution of the rock composition was closely related to the diagenetic age, and there were significant differences between the diagenetic compositions in different periods. The evolution of the rock composition in the study area was from basic, to neutral, to acidic, and the corresponding trace elements enriched in different rocks were combined, which provided metallogenic materials for the tectonic accumulation. The metallogenic characteristics of magmatic Cu-Ni sulfide deposits and skarn-porphyry Mo-polymetallic deposits were analyzed, and the 1:200,000 regional geological results that were conducted by Qinghai Geology and Mineral Bureau in 1986 were combined. The regional abundance characteristics of elements and the distribution characteristics of elements in different strata are summarized as follows: The known metal mineralization and mineralization clues were closely related to the internal and external contact zone of the intrusions, and multiple periods of intermediate-acid magmatic activities brought abundant medium- and low-temperature hydrothermal metallogenic elements and hydrothermal conditions to this area. The regional background contents and discrete eigenvalues of Ni, Pb, W, Au, Zn, and Ag elements were large, and the local enrichment trend was obvious. The spatial distribution of Pb, Cu, Ni, Cr, and W element anomalies was controlled by faulted structure and magmatic activities. The known small iron deposit as well as copper, lead, zinc, and iron mineral occurrences were exposed in the higher anomaly area of Cr, Ni, Pb, and W elements. Based on the above analysis results, Ag, Cu, Cr, Pb, Ni, Zn, Mo, and W elements were selected as metallogenic geochemical information.

The background and anomaly information of geochemical elements was needed to separate effectively for enhancing the mineralization indication of geochemical information. The content of trace elements in various natural substances should generally obey the lognormal or normal distribution, which can be used as the basis for determining the background value and the anomaly threshold by the quantitative statistical method. In this study, the mean iteration method was used to process geochemical elements. The analysis of arithmetic mean, variance, and variation coefficient of eight geochemical elements, including Ag, Cr, Cu, Mo, Ni, Pb, W, and Zn, indicated that both the raw and logarithmic values of each element were not likely to obey normal distribution because each element had the extreme value, and the standard deviation and variation coefficient were also large. Therefore, the iterative method was applied to process the extreme outliers of element data

as follows: the mean value, X, and corresponding standard deviation, σ , of each element

were calculated; extreme values were eliminated with $X \pm 3\sigma$ as the limits; the mean and standard deviation of the data were recalculated. The calculation was done when there was no extreme value, and all data was distributed between $\left[\bar{X} - 3\sigma, \bar{X} + 3\sigma\right]$. It can be

seen from Table 1 that after iterative elimination calculation, the distribution of the original data was greatly improved. The logarithmic frequency histograms of each element after iteration are shown in Figure 3; it can be seen that although the standard deviation and variation coefficient of individual elements are still very large, almost all elements obey normal distribution.

0

-0.4 -0.3 -0.2 -0.1 0

0.1

ω(LogW)/10⁻⁶

(g)



Table 1. The eigenvalues of each element after iteration.

Ag

Cr

Cu

Mo

Ni

Element

Figure 3. Frequency histograms of geochemical elements. (a) frequency histogram of Ag; (b) frequency histogram of Cr; (c) frequency histogram of Cu; (d) frequency histogram of Mo; (e) frequency histogram of Ni; (f) frequency histogram of Pb; (g) frequency histogram of W; (h) frequency histogram of Zn.

1

0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9

0

0.8 1 1.2 1.4 1.6 1.8 2

The observed values of water sediments of each geochemical index were gridded in Surfer software (Surfer 8.0, Golden Software, Golden, CO, USA) using the reciprocal interpolation method of distance, and the grid space accuracy was consistent with the

Zn

W

Pb

2.2 2.4

 $\omega(\text{LogZn})/10^{-6}$ (h)

evaluation cell grid. Additionally, the mean and logarithmic values of the standard deviation of each element were converted into true values, and the sum of the mean and two times of the standard deviation $(\bar{X} \pm 2\sigma)$ were used as the anomaly lower limit to extract geochemical anomaly information. The anomaly distribution maps of each element, which were made by Surfer software, overlayed and compared with the known mineral deposits (Figure 4). It can be seen that the anomaly of each element was closely related to the known metallogenic cells, and the delineated anomaly focusing area corresponded to the spatial location of known mineral occurrences, indicating that calculation of the anomaly lower limit was valid and reliable. Vector information extracted from these four geological bodies was stored in the MapGis software spatial database (MapGis 7.0, China University of Geosciences, Wuhan, China) as geochemical mineralization information in the study area.



Figure 4. Cont.



Figure 4. Geochemical anomalies of Ag, Cr, Cu, Mo, Ni, Pb, W, and Zn. (**a**) Geochemical anomaly of Ag; (**b**) Geochemical anomaly of Cr; (**c**) Geochemical anomaly of Cu; (**d**) Geochemical anomaly of Mo; (**e**) Geochemical anomaly of Ni; (**f**) Geochemical anomaly of W; (**h**) Geochemical anomaly of Zn.

3.2.3. Remote Sensing Evidence Information

The study area was located at the junction of Qaidam and Kunlun blocks, with well-developed faults, which provided sufficient conditions for regional diagenesis and mineralization. In this study, manual visual interpretation was used to extract the structural information of fault and contact zone based on texture and shape features of remote sensing images. The ETM+ and "ZY-1" 02C data were selected, where ETM+ was the Landsat 7 satellite image data, with 30 m resolution, and "ZY-1" 02C were the first high-resolution remote sensing data in China, with 2 m resolution. The middle and large scale linear structures in the region were extracted with ETM+ remote sensing images, and "ZY-1" 02C data were applied to extract the small secondary linear structures in the key metallogenic area (Figure 5). In data-driven mineral predicting, the target variable is usually point entity (known deposit and mineral occurrence) and the predicting variable is surface entity. When wired or point entities are in the predicting variables, buffer analysis in GIS is needed to convert them into surface entities. The linear structures were buffered with a distance of 200 m, and the results were expressed as polygon shapes (Figure 6), because the average collection density of geochemical sample points was no less than 7.1/km². Additionally, the regional magmatic activities were frequent, which was closely related to polymetallic mineralization, and the rock exposed well. Therefore, combined with the spectral characteristics of polymetallic mineralization alteration, iron-stained and hydroxyl alteration information caused by the ferritization and jarpsite mineralization characteristics were extracted from the ASTER remote sensing images with principal component analysis. Because chloritization mineralization characteristics cause hydroxyl alteration anomaly, the linear structures and alteration zones that contained Fe⁺ and OH⁻ were used as remote sensing evidence information (Figure 7).



Figure 5. Linear construct information.



Figure 6. Linear construct buffer.



Figure 7. Alteration information distribution.

4. Mapping Mineral Prospectivity

4.1. Map Layer Selection

In mineral predicting, the metallogenic control conditions or the altered states of conditional combination were assigned by binary assignment. If the existence of a space entity could be used as a favorable marker for mineralization, the evaluation grid cell where the sign was located was assigned a 1 value, otherwise, it was assigned a 0 value. The evaluation grid cells after assignment were equal to the discrete random variables in probability theory. In addition, the statistical method was applied to quantitatively study the relationship of the comprehensive information variables to one another, and the relationship with the output state of mineral resources (deposit-bearing and nondeposit-bearing), so as to realize the quantitative selecting of comprehensive information variables. Therefore the information variables that were not related to mineralization or redundant with other information variables were eliminated. In this study, the Youden index, commonly used in medical statistical analysis, was introduced into the optimization of metallogenic evidence layers. The correlation between metallogenic information and mineral deposit output state was measured, and the comprehensive metallogenic prediction variables were optimized by using the Youden index. The Youden index is defined as the difference between True Positive Rate (TPR) and False Positive Rate (FPR).

$$yd = TPR - FPR \tag{20}$$

$$TPR = tp/(tp+fn) \tag{21}$$

$$FPR = fn/(tn+fp) \tag{22}$$

In medical statistics, one class can be marked as positive and the other as negative when solving problems. Assuming that the training samples are composed of *P* positive and N negative samples, the classification results can be expressed as in the following four situations: *tp* represents the true positives in predicted positive classes; *tn* represents the true negatives in predicted negative classes; *fp* represents the false positives in predicted negative classes (the actual negatives recognized as positives); fn represents the false negatives in predicted positive classes (the actual positives recognized as negatives). TPR represents the proportion of true positives in predicted positive classes to the total actual positives, and FPR represents the proportion of false positives in predicted positive classes to the total actual negatives. In metallogenic information prediction, TPR represents the proportion of the number of deposit-bearing cells recognized as abnormal cells to the total number of abnormal cells in predicted results, and FPR represents the proportion of the number of non-deposit-bearing cells recognized as background cells to the total number of background cells in predicted results. The Youden index was between -1 and +1, representing the difference between the probability of a positive sample predicted and a negative actual predicted value. In addition, it means the percentage of the benefits that exceed the costs in prediction, and it is a statistic that describes the comprehensive characteristics of benefit and cost in metallogenic prediction. The Youden index of each evidence layer was calculated by training cells. When the Youden index is between 0 and 1, it indicates that the probability that the sample predicted value is positive is higher than the probability that the actual predicted value is negative, and the evidence layer is highly correlated with the known mineral deposits. When the Youden index is between -1 and 0, it indicates that the probability that the sample predicted value is positive is lower than the probability that the actual predicted value is negative, namely the number of nondeposit-bearing cells identified as background cells accounts for a high proportion of the total predicted background cells. This result means that the cost of metallogenic prediction exceeds the benefit and has a negative effect on metallogenic prediction. Therefore, the layers with a Youden index higher than 0 have a spatial correlation with the known mineral deposits, which can predict exploration targets as predictor variables. The calculation results are shown in Table 2.

Layer	Correlation Intension	Layer	Correlation Intension	
Ag	0.3884	Iron oxide alteration	0.1342	
Cr	0.3194	Hydroxyl alteration	0.0801	
Cu	0.3942	Linear structure	0.1405	
Mo	0.2111	Monzonitic granite	-0.0094	
Ni	0.3349	Granodiorite	-0.2158	
Pb	0.2073	Quartz diorite	-0.0058	
W	0.2114	Carbonate	0.0669	
Zn	0.2359			

Table 2. Youden index for metallogenic evidence layers.

It can be seen from Table 2 that 12 evidence layers had a Youden index higher than 0, including Ag, Cr, Cu, Mo, Ni, Pb, W, Zn, iron oxide alteration, hydroxyl alteration, linear structure, and carbonate. The 12 evidence layers selected by the Youden index and one target layer that contains known mineral deposits were spatially overlayed and used to train the learning data during the machine learning modeling process. Using 12 evidence layers and one target layer as input and output data for modeling, 18,600 grid cells in the study area were trained, where the deposit-bearing and non-deposit-bearing cells were assigned values of 1 and 0, respectively. These numbers represent that the grid contained or did not contain known polymetallic deposits (mineral occurrences), respectively.

4.2. Machine Learning Model Combined with Hyperparameter Optimization

BA and FA were applied to optimize the hyperparameters of different machine learning models. These two optimization algorithms were needed to define the fitness value of the objective function as the standard for hyperparameter optimization calculation. When the fitness value reached the optimal value, the hyperparameters of the machine learning models obtained the optimal solution. In this study, the AUC value of the ROC curve was selected as the fitness value of the optimal objective function. The AUC value can be considered as the probability that the number of correctly predicted grid cells is higher than that of incorrectly predicted grid cells [67]. The higher the AUC value, the better the performance of the model in the mineral potential prediction [68,69]. The AUC value is applied to evaluate the classification performance when multiple classifiers are compared. According to the Wilcoxon Mann–Whitney statistic [70,71], the AUC value can be expressed as

$$AUC = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} \varphi(x_i, y_j)$$
(23)

$$p(x_i, y_j) = \begin{cases} 1 & x_i > y_j \\ 0.5 & x_i = y_j \\ 0 & x_i < y_j \end{cases}$$
(24)

When using BA and FA to optimize the hyperparameters of machine learning models, the search space of the algorithm was a one-dimensional or two-dimensional space with model hyperparameters as coordinate axes. Artificially setting the space range of hyperparameters, the iterative search process started from the random position in the search space. Thus, the model hyperparameter corresponding to the maximum fitness function value obtained in the iterative search process was the best for modeling hyperparameters. In this study, the number of hidden layer neurons in MLP, the updated learning times and the weight reduction coefficient of the weak regression in AdaBoost, and the fault tolerance penalty coefficient and kernel hyperparameter in OCSVM were optimized respectively by using the two algorithms, and the AUC value of the ROC curve was selected as the fitness function. The hyperparameter optimization process is shown in Figures 8 and 9, and the metallogenic prediction program of machine learning models combined with hyperparameter optimization was written on Python platform.



Figure 8. Bat algorithm (BA) optimization process.

4.2.1. MLP Model Combined with Hyperparameter Optimization

The metallogenic probability of each grid cell was calculated using the MLP model to map the mineral prospectivity of the study area. The greater the value of the grid cell, the more likely the grid cell was to be a deposit-bearing cell. In the regression prediction process, the setting of the number of hidden layer neurons, N_{hid} , would influence the prediction results of the MLP model. Therefore, the number of hidden layer neurons, N_{hid} , was an important hyperparameter. BA and FA were used to optimize the N_{hid} in the MLP model. For determining the search space of BA and FA populations, the value range of N_{hid} was defined as (0,300], according to experience.

The initial hyperparameters of BA include population size *L*, iterations *T*, upper and lower pulse frequency f_{\min} and f_{\max} , upper and lower limits of pulse intensity A_{\min} and A_{\max} , the attenuation coefficient of pulse loudness α , and pulse frequency increase coefficient γ . BA is not sensitive to other initial hyperparameters, except iterations *T*. In this study, the population size *L* = 20, iterations *T* = 50, pulse frequency range $f_{\min} = 0$, $f_{\max} = 1$, pulse intensity range $A_{\min} = 0$, $A_{\max} = 1$, attenuation coefficient of pulse loudness



 α = 0.9, pulse frequency increase coefficient γ = 0.9, and the BA was initialized as the default hyperparameters.

Figure 9. Firefly algorithm (FA) optimization process.

In this study, the initial hyperparameters of FA were set as follows: number of fireflies n = 60, iterations T = 50, fluorescein volatility $\rho = 0.4$, fluorescein update rate $\gamma = 0.6$, initial value of brightness $l_0 = 10$, moving step value $\alpha = 0.6$, dynamic decision domain updating rate $\beta = 0.8$, perceived radius $r_s = 5$, the number of fireflies in the neighborhood N(t) = 5, and the FA was initialized as the default hyperparameters.

According to the setting search range of N_{hid} , BA and FA performed optimization calculation in the defined search space, and the fitness function AUC value of different iterations is shown in Figure 10. It can be seen that the AUC value gradually increased with the increase of iterations *T*. When iterations *T* = 27, the AUC value reached the maximum, as seen in Figure 10a, and the corresponding N_{hid} = 96. When iterations *T* = 36, the curve reached a stable state, as seen in Figure 10b, and the corresponding N_{hid} = 39.



Figure 10. Curves of area under curve (AUC) value: (**a**) bat algorithm multilayer perceptron (BA-MLP); (**b**) firefly algorithm multilayer perceptron (FA-MLP).

4.2.2. AdaBoost Model Combined with Hyperparameter Optimization

The AdaBoost model was applied to the metallogenic prediction of the Lalingzaohuo area, and the CART decision tree was selected as the weak regression of the AdaBoost model. During regression prediction, the update learning times, *t*, and the weight reduction coefficient, v (0 < v < 1), of the weak regression were two important hyperparameters. Small t was challenging to fit, while large t was easy to overfit. For the same training set, smaller v meant more times were needed to select the weak regression. In addition, the two hyperparameters t and v should be optimized and adjusted together. Therefore, BA and FA were respectively used to optimize and calculate t and v of the AdaBoost model. For determining the search space of BA and FA populations, the value range of update learning times t and the weight reduction coefficient v were defined as (0,200] and (0,1], respectively. In the process of calculation, T = 20, and the other hyperparameters were set the same as those in Section 4.2. During the optimization process, the change curve of the AUC value was as shown in Figure 11. As seen from Figure 11, in the BA-AdaBoost model, when iteration reached 12, the AUC value reached the maximum, and the update learning times t and the weight reduction coefficient v were 98 and 0.1332, respectively. In the FA-AdaBoost model, when iteration reached 16, the AUC value reached the maximum, and the update learning times t and the weight reduction coefficient vwere 75 and 0.1335, respectively.



Figure 11. Curves of AUC value: (a) BA-AdaBoost; (b) FA-AdaBoost.

4.2.3. OCSVM Model Combined with Hyperparameter Optimization

The OCSVM model was used to calculate the value of the decision function f(x) of each geological statistic cell. If f(x) > 0, geological statistical cells belonged to a nondeposit-bearing area; otherwise, they belonged to a deposit-bearing area (exploration target area). For the convenience of application, the negative number of the statistical cell decision function was used. The larger it was, the more possible it was to be a deposit-bearing cell. During the modeling process, the selection of kernel functions influences the results of the OCSVM model. The commonly used kernel functions include polynomial kernel function, radial basis kernel function, and sigmoid kernel function. In this study, the Gaussian radial basis kernel function was selected as the kernel function. The fault tolerance penalty coefficient γ and kernel hyperparameter σ^2 , two necessary adjustment hyperparameters, were optimized by BA and FA. For determining the search space of BA and FA populations, the value range of fault tolerance penalty coefficient γ and kernel hyperparameter σ^2 were defined as (0,10] and (0,1], respectively. In the calculation process, the iterations T = 30 and the remaining initial hyperparameters were set the same as those in Section 4.2. During the optimization process, the change curve of AUC value was as shown in Figure 12. As seen from Figure 12, in BA-OCSVM, when the iteration reached 14, the AUC value reached the maximum, and the output fault tolerance penalty coefficient γ and kernel hyperparameter σ^2 were 0.00845 and 0.88328, respectively. In FA-OCSVM, when the iteration reached 19, the AUC value reached the maximum, and the output fault tolerance penalty coefficient γ and kernel hyperparameter value σ^2 were 6.82289 and 0.36972, respectively.



Figure 12. Curves of AUC value: (**a**) bat algorithm one-class support vector machine (BA-OCSVM); (**b**) firefly algorithm one-class support vector machine (FA-OCSVM).

4.3. Mineral Potential Mapping

The mineral prospectivity mapping model was established using the optimized hyperparameters; it was applied to calculate the metallogenic potential value of each evaluation cell. The value range of the mineral potential of each evaluation cell was different because of the different principles of the three machine learning models. The results for the MLP and AdaBoost models represented the metallogenic probability of the evaluation grid cell, while the results of the OCSVM model were the decision function values corresponding to the selected kernel function. Although the result ranges of the three machine learning models were different, it did not affect the prediction of mineral resources, because the mineral potential was evaluated according to the prediction value of the evaluation grid cell. The high value area of mineral potential was delineated with the optimal threshold method. This method specified that the area delineated by the optimal threshold and the known mineral deposits have the maximum spatial correlation, where optimal threshold was calculated by the Youden index. Meanwhile, TPR represented the proportion of the number of deposit-bearing cells recognized correctly to the total number of deposit-bearing cells in the predicted result, and FPR represented the proportion of the number of nondeposit-bearing cells recognized correctly to the total number of non-deposit-bearing cells in the predicted result. The maximum and minimum of the calculated metallogenic advantage index were taken as the endpoints of the continuous interval, which was divided into 1000 equal sub-intervals, each of which was regarded as a threshold. The model was used to calculate the TPR and FPR corresponding to the thresholds, and the TPR minus the FPR was used to calculate the Youden index. The threshold corresponding to the maximum Youden index was selected as the optimal threshold for delineating the exploration targets, and the maximum Youden indices and optimal thresholds of different models were as shown in Table 3. With the GIS software platform and the optimal thresholds, the mineral potential maps of different machine learning models after optimizing were established (Figure 13). In order to compare the effect of metallogenic prediction, the low and high value areas of mineral potential were represented by green and blue, respectively, the known mineral deposits were represented by red dots, and the layer of known mineral deposits was overlayed with the mineral potential map.

 Table 3. Maximum Youden indices and optimal thresholds for different models.

Model	BA-MLP	FA-MLP	BA-AdaBoost	FA-AdaBoost	BA-OCSVM	FA-OCSVM
Maximum Youden index	0.6649	0.55489	0.807039	0.815599	0.70648	0.52872
Threshold	0.0028	0.0014	0.0452	0.0167	109.7745	96.05151



Figure 13. Cont.



Figure 13. Mineral potential maps: (a) BA-MLP; (b) FA-MLP; (c) BA-AdaBoost; (d) FA-AdaBoost; (e) BA-OCSVM; (f) FA-OCSVM.

5. Results

For comparing and analyzing the metallogenic prediction effect of machine learning models before and after parameter optimization, the original MLP, AdaBoost, and OCSVM models were applied to study the modeling and prediction of mineral potential. Setting the default value of hyperparameters in Python's scikit-learn module as the hyperparameter value of the three models, the hidden layer neurons, N_{hid} , of the MLP model was set as 50, the update learning times, t, and the weight reduction coefficient, v, of the AdaBoost model were set as 10 and 0.1, respectively, and the fault tolerance penalty coefficient, γ , and kernel hyperparameter value, σ^2 , of the OCSVM model were set as 1 and 0.1, respectively. In this section, the mineral prospectivity mapping results were statistically evaluated using ROC and P-A curves. ROC curve analysis has been increasingly applied in the field of machine learning. It has insensitivities in category distribution and cost as well as good intuition and strong understanding [72,73]. The coordinate system for the ROC curve is formed by taking the FPR and TPR as the X-axis and Y-axis, respectively. After having been trained, the discrete and binary output points correspond to the points in the coordinate system, i.e., a set of single points is obtained by setting different thresholds for the same classifier. These points are connected in a curve from left to right in the ROC coordinate system. The ROC curve has a monotonically decreasing slope. The optimal classifier can be selected from a set of classifiers according to their ROC curves [74,75] The better a classifier is, the closer its ROC curve is to the upper left corner of the ROC space [76,77]. The AUC value represents the area under the ROC curve, and its calculation method is shown in Equations (23) and (24), using the AUC value to evaluate classification performance. The range of the AUC value is between (0,1). The closer the ROC curve is to the upper left corner of the ROC space, the closer the AUC value calculated by the model is to 1.0. The TPR and FPR corresponding to different thresholds within the range of the metallogenic probability of different models could be obtained by calculating mineral potential, and the ROC curves are drawn as shown in Figure 14. It can be seen that although the ROC curves are intersecting, the ROC curves (blue) corresponding to the original machine learning models without hyperparameter optimization are all lower than that of the machine learning models after hyperparameter optimization, indicating the optimized machine learning model is better than the corresponding original model in metallogenic prediction.

In addition, in a study area, the model predicting accuracy, *P*, and predicted area percentage, *A*, can be calculated according to the mineral potential maps and the distribution of known mineral occurrences. *P* represents the benefits of the model prediction, and *A* represents the cost of the model prediction, using *P* and *A* values to draw a P-A curve [78,79]. In this study, according to the mineral potential value of the evaluation grid cell calculated by the three models, the *P* and *A* of corresponding models were calculated when different metallogenic potential was taken as the threshold. In addition, P-A curves of different machine learning models were drawn. In Figure 15, the horizontal axis of the coordinate system is the value of metallogenic potential calculated by the model. In addition, the ratio of *P* value to *A* value can be used as statistics. When the ratio becomes larger, the scale of the predicted targets will be larger; the target area will be smaller; manpower, material resources, and financial input will be less; the effect of the prediction model will be better [80]. Therefore, the height of the intersection point of the predictive accuracy and the predicted area percentage curves can be used to evaluate the performance of the model. Meanwhile, the intersection point of the two curves corresponds to the optimal threshold of the mineral prediction results, which can be used to further delineate the prospecting target area (Figure 15).



Figure 14. Receiver operating characteristic (ROC) curves of the prediction models: (**a**) MLP; (**b**) AdaBoost; (**c**) OCSVM.

The AUC value and P/A value with different hyperparameter optimization algorithms were calculated and compared with results without hyperparameter optimization (Table 4) to analyze the metallogenic prediction effects of machine learning models.



Figure 15. Predication-area (P-A) curves of the prediction models: (**a**) BA-MLP; (**b**) FA-MLP; (**c**) BA-AdaBoost; (**d**) FA-AdaBoost; (**e**) BA-OCSVM; (**f**) FA-OCSVM.

By analyzing ROC curves (Figure 14), P-A curves (Figure 15) and evaluation indices (Table 4) of machine learning models, the following conclusions can be obtained:

- 1. The AUC values obtained by the machine learning models were all above 0.7, indicating that the metallogenic prediction results are consistent with the distribution regularity of known mineral deposits and mineralization points in the area. The AdaBoost model had the best effect, where the AUC value reached 0.9579 after using the FA algorithm, while the AUC value of the BA-OCSVM model was 0.8758, and that of the BA-MLP model was 0.8712.
- 2. The BA and FA hyperparameter optimization algorithms have an obvious effect, and the accuracy of the MLP, AdaBoost, and OCSVM models was improved after optimization. The accuracy of the FA-AdaBoost model was improved most significantly, with the accuracy 17.42% higher after optimization. For the MLP and OCSVM models, the accuracy of the BA algorithm after optimization was higher than that of the FA algorithm, and the predicting accuracy of BA-MLP increased 11.29%, while that of the FA-MLP model only increased 2.41%.

Model	AUC	P/A
MLP	0.7828	2.449
BA-MLP	0.8712	5.533
FA-MLP	0.8017	3.745
AdaBoost	0.8173	3.156
BA-AdaBoost	0.9588	8.182
FA-AdaBoost	0.9597	10.765
OCSVM	0.8056	3.954
BA-OCSVM	0.8758	7.417
FA-OCSVM	0.8262	6.769

Table 4. Evaluation indices of different prediction models.

6. Discussion

- 1. The BA and FA, which have the same structural characteristics, belong to the global optimization algorithm based on population random optimization. In the search domain, each bat or firefly represents a solution of the optimization function; the fitness value evaluates their position and finds the optimal individual by adjusting its own population hyperparameters. As the difference of machine learning model structures, different iteration times are defined in the test. The optimization hyperparameter of the MLP model is only one, as its structure is simple, while the optimization hyperparameter number of the AdaBoost and OCSVM models is two. The later defined iteration times are 20 and 30, and the computational speed is slow. The results show that BA and FA can converge a global maximum in the number of iterations set, and at the same time, the AdaBoost and OCSVM models spare more time for searching optimal hyperparameters.
- 2. The evaluation indices results show that the characteristic regularity of the ROC curve is basically consistent with the AUC curve, while the P-A curve and the P/A index can provide another accuracy evaluation standard for the mineral potential prediction model in terms of prospecting benefit. For example, the AUC value of the BA-MLP is 0.8712, while the AUC value of the BA-OCSVM is 0.8758. The calculated values of the ROC curve evaluation indices of the two models are very close, meaning that the evaluation effect of the model cannot be compared. However, it can be found that the P-A curve intersection of the OCSVM model is obviously higher than that of the BA-MLP model, where the P/A value of BA-OCSVM is 7.417, while the P/A value of BA-MLP is 5.533, which indicates the superiority of the BA-OCSVM model in terms of prospecting benefit.
- 3. When comparing the three kinds of mineral potential maps (Figure 13) after hyperparameter optimization, the range and space distribution trend of the predicted high potential prospecting area have a high similarity. Besides the known mineral zone, the models delineate a potential prospecting area in the northeast, northwest, and southwest of the study area. The mineralized geological environment of the northeast is relatively superior, where the Ordovician-Silurian Tanjianshan Group and the carbonate formation of Early Carboniferous Dagangou Group are the main formations, and the monzonitic granite of Late Triassic is the main intrusive mass. Skarn mineral deposit is more likely to be formed in this high-value distribution region, which is the same with the forecast results of the model, and will be a key area for detailed geological and mineral investigation in the future. The north of Kunbei fault is in the northwest, Tanjianshan Group and the carbonate formation of Early Carboniferous Dagangou Group distribution as linear along the structural line direction under the control of fault. Monzogranite of the early Jurassic distributed in the north of sedimentary strata, where the contact sites are more likely to form a skarn belt. Stacking the metallogenic potential and geological map, the partial high value area is in the contact zone of carbonate rock strata and intrusive body. The metallogenic potential in the southwest is distributed in the intrusive body of different periods, and the high

value area is distributed sporadically, with the possibility of mineralization being very small.

4. BA and FA perform excellently in hyperparameter optimization calculation; however, in the modeling calculation, it is found that there are still some problems in the combination of this kind of swarm intelligence optimization algorithm and machine learning model, which need to be further analyzed. One problem is that both algorithms need to set a large number of initial hyperparameters. However, in this study, only the corresponding relationship between the iterations *T* and the fitness function value is tested, and others use the default values. The other problem is that the machine learning model optimized by this kind of algorithm takes a long time to run. In the BA-MLP model, the calculation time of the machine learning model with optimized hyperparameters is about eight times that of the original model, which affects the test efficiency to a certain extent.

7. Conclusions

- 1. The mineral prospectivity mapping models are constructed under conditions where the search space needs to be set and the optimal hyperparameters are automatically searched by combining bat and firefly swarm intelligence optimization algorithms with different machine learning models. Compared with the traditional optimization algorithms, the BA and FA are free to switch between global and local optimization processes and have more opportunities to find the global optimal hyperparameters of machine learning models.
- The BA and FA have different improvement effects on MLP, AdaBoost, and OCSVM models. The accuracy of the machine learning models is greatly improved after hyperparameter optimization, indicating that the model hyperparameter optimization is effective and reliable in the application of machine learning methods.
- 3. ROC and P-A curves are applied to quantitatively evaluate the prediction performance of the mineral prospectivity mapping models. The evaluation results show that the AUC value can effectively measure the accuracy of the different models, but it is not the only index. The P/A value of the curve intersection is calculated by the P-A curve. The higher the value is, the more accurate the metallogenic prediction. The P-A curve represents the prospecting benefit under limited manpower, material resources, and financial resources, and the P/A value can be used as the accuracy evaluation standard for another mineral potential prediction.

Author Contributions: N.L. processed data and wrote the text; Y.C. developed the code; H.L. (Haiqi Liu) plotted the figures; H.L. (Hanlin Liu) reviewed the original draft. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (Grant Nos. 41702357 and 41672322).

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy.

Acknowledgments: The authors are grateful to the three anonymous reviewers for their constructive comments that greatly improved the manuscript.

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

References

- 1. Sun, T.; Li, H.; Wu, K.X.; Chen, F.; Zhu, Z.; Hu, Z. Data-driven predictive modelling of mineral prospectivity using machine learning and deep learning methods: A case study from southern Jiangxi province, China. *Minerals* **2020**, *10*, 102. [CrossRef]
- Xiong, Y.H.; Zuo, R.G. GIS-based rare events logistic regression for mineral prospectivity mapping. *Comput. Geosci.* 2018, 111, 18–25. [CrossRef]
- Wang, J.; Zuo, R.G.; Xiong, Y.H. Mapping mineral prospectivity via semi-supervised random forest. *Nat. Resour. Res.* 2020, 29, 189–202. [CrossRef]

- Liu, Y.; Cheng, Q.M.; Xia, Q.L.; Wang, X.Q. Application of singularity analysis for mineral potential identification using geochemical data—A case study: Nanling W-Sn-Mo polymetallic metallogenic belt, South China. J. Geochem. Explor. 2013, 134, 61–72. [CrossRef]
- 5. Moller, M.F. A scaled conjugate-gradient algorithm for fast supervised learning. Neural Netw. 1993, 6, 525–533. [CrossRef]
- 6. Figueiredo, M.A.T.; Jain, A.K. Unsupervised learning of finite mixture models. *IEEE Trans. Pattern Anal. Mach. Intell.* 2002, 24, 361–396. [CrossRef]
- 7. Barlow, H.B. Unsupervised learning. Neural Comput. 1989, 1, 295–311. [CrossRef]
- 8. Ashfaq, R.A.R.; Wang, X.Z.; Huang, J.Z.; Abbas, H.; He, Y.L. Fuzziness based semi-supervised learning approach for intrusion detection system. *Inf. Sci.* 2017, *378*, 484–497. [CrossRef]
- 9. Mallapragada, P.K.; Jin, R.; Jain, A.K.; Liu, Y. SemiBoost: Boosting for semi-supervised learning. *IEEE Trans. Pattern Anal. Mach. Intell.* **2009**, *31*, 2000–2014. [CrossRef]
- Wu, H.; Prasad, S. Semi-supervised deep learning using pseudo labels for hyperspectral image classification. *IEEE Trans. Image Process.* 2018, 27, 1259–1270. [CrossRef]
- 11. Kaelbling, L.P.; Littman, M.L.; Moore, A.W. Reinforcement learning: A survey. J. Artif. Intell. Res. 1996, 4, 237–285. [CrossRef]
- 12. Zhang, G.Q.; Patuwo, B.E.; Hu, M.Y. Forecasting with artificial neural networks: The state of the art. *Int. J. Forecast.* **1998**, *14*, 35–62. [CrossRef]
- 13. Basheer, I.A.; Hajmeer, M. Artificial neural networks: Fundamentals, computing, design, and application. *J. Microbiol. Methods* **2000**, *43*, 3–31. [CrossRef]
- 14. Safavian, S.R.; Landgrebe, D. A survey of decision tree classifier methodology. *IEEE Trans. Syst. Man Cybern.* **1991**, *21*, 660–674. [CrossRef]
- 15. Chen, C.L.P.; Liu, Y.J.; Wen, G.X. Fuzzy neural network-based adaptive control for a class of uncertain nonlinear stochastic systems. *IEEE Trans. Cybern.* 2014, 44, 583–593. [CrossRef]
- 16. Degloria, A.; Faraboschi, P.; Olivieri, M. Clustered boltzmann machines: Massively parallel architectures for constrained optimization problems. *Parallel Comput.* **1993**, *19*, 163–175. [CrossRef]
- 17. Chang, C.C.; Lin, C.J. LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Technol. 2011, 2, 1–27. [CrossRef]
- 18. Hearst, M.A. Support vector machines. *IEEE Intell. Syst. Their Appl.* **1998**, *13*, 18–21. [CrossRef]
- 19. Lin, C.F.; Wang, S.D. Fuzzy support vector machines. IEEE Trans. Neural Netw. 2002, 13, 464–471. [CrossRef]
- 20. Amari, S.; Wu, S. Improving support vector machine classifiers by modifying kernel functions. *Neural Netw.* **1999**, *12*, 783–789. [CrossRef]
- 21. Breiman, L. Random forests. Mach. Learn. 2001, 45, 5–32. [CrossRef]
- Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol. Rev.* 2015, 71, 804–818. [CrossRef]
- Chen, Y.L.; Wu, W. Mapping mineral prospectivity using an extreme learning machine regression. Ore Geol. Rev. 2017, 80, 200–213. [CrossRef]
- 24. Liu, Y.; Zhou, K.F.; Qia, Q.L. A MaxEnt Model for Mineral Prospectivity Mapping. Nat. Resour. Res. 2018, 27, 299–313. [CrossRef]
- Tan, Y.Y.; Li, L.L.; Zhang, X.; He, C. An improved method for microseismic source location based on grid search. *Chin. J. Geophys.* 2017, 60, 293–304. [CrossRef]
- 26. Yao, Z.X.; Wang, C.Y.; Lou, H. Ascertaining the structure parameters of Kunlun fault zone using the grid searching method based on trapped wave correlation. *Chin. J. Geophys.* **2010**, *53*, 1167–1172. [CrossRef]
- 27. Deh, K.; Pratap, A.; Agarwal, S.; Meyarivan, T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evolut. Comput.* **2002**, *6*, 182–197. [CrossRef]
- 28. Holland, J.H. Genetic algorithms. Sci. Am. 1992, 267, 66–72. [CrossRef]
- 29. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the IEEE International Conference on Neural Networks, Perth, Australia, 27 November–1 December 1995; pp. 1942–1948. [CrossRef]
- 30. Coello, C.A.C.; Pulido, G.T.; Lechuga, M.S. Handling multiple objectives with particle swarm optimization. *IEEE Trans. Evolut. Comput.* **2004**, *8*, 256–579. [CrossRef]
- Eberhart, R.C.; Shi, Y.H. Particle swarm optimization: Developments, applications and resources. In Proceedings of the 2001 Congress on Evolutionary Computation, Seoul, Korea, 27–30 May 2001; pp. 81–86.
- Yang, X.S. A new metaheuristic bat-inspired algorithm. In Proceedings of the Nature Inspired Cooperative Strategies for Optimization (NICSO 2010), Granada, Spain, 12–14 May 2010; González, J.R., Pelta, D.A., Cruz, C., Terrazas, G., Krasnogor, N., Eds.; Springer: Berlin, Germany, 2010; pp. 65–74. [CrossRef]
- 33. Yang, X.S.; Gandomi, A.H. Bat algorithm: A novel approach for global engineering optimization. *Eng. Comput.* **2012**, *29*, 464–483. [CrossRef]
- 34. Dorigo, M.; Maniezzo, V.; Colomi, A. Ant system: Optimization by a colony of cooperating agents. *IEEE Trans. Syst. Man Cybern. Part B-Cybern.* **1996**, *26*, 29–41. [CrossRef] [PubMed]
- 35. Lakshminarayana, P.; SureshKumar, T.V. Automatic Generation and Optimization of Test case using Hybrid Cuckoo Search and Bee Colony Algorithm. *J. Intell. Syst.* **2020**, *30*, 59–72. [CrossRef]

- 36. Bae, J.; Ahn, J.; Lee, S.J. Comparison of Multilayer Perceptron and Long Short-Term Memory for Plant Parameter Trend Prediction. *Nucl. Technol.* **2020**, *206*, 951–961. [CrossRef]
- 37. Delgoshaei, A.; Gomes, C. A multi-layer perceptron for scheduling cellular manufacturing systems in the presence of unreliable machines and uncertain cost. *Appl. Soft Comput.* **2016**, *49*, 27–55. [CrossRef]
- Chiappini, F.A.; Allegrini, F.; Goicoechea, H.C.; Olivieri, A.C. Sensitivity for Multivariate Calibration Based on Multilayer Perceptron Artificial Neural Networks. *Anal. Chem.* 2020, 92, 12265–12272. [CrossRef] [PubMed]
- 39. Araujo, R.D.; Oliveira, A.L.I.; Meira, S. A class of hybrid multilayer perceptrons for software development effort estimation problems. *Expert Syst. Appl.* **2017**, *90*, 1–12. [CrossRef]
- 40. Kang, F.; Li, J.J.; Xu, Q. System reliability analysis of slopes using multilayer perceptron and radial basis function networks. *Int. J. Numer. Anal. Methods Geomech.* **2017**, *41*, 1962–1978. [CrossRef]
- 41. Raghu, S.; Sriraam, N. Optimal configuration of multilayer perceptron neural network classifier for recognition of intracranial epileptic seizures. *Expert Syst. Appl.* **2017**, *89*, 205–221. [CrossRef]
- 42. Ke, J.; Zhu, Z.X.; Zhe, Y.; Yu, F.; Bi, T.S.; Zhang, J.K. Intelligent islanding detection method for photovoltaic power system based on Adaboost algorithm. *IET Gener. Transm. Distrib.* **2020**, *14*, 3630–3640. [CrossRef]
- 43. Hu, G.S.; Yin, C.J.; Wan, M.Z.; Zhang, Y.; Fang, Y. Recognition of diseased Pinus trees in UAV images using deep learning and AdaBoost classifier. *Biosyst. Eng.* 2020, *194*, 138–151. [CrossRef]
- 44. Li, L.; Huang, Y.X.; Tao, J.F.; Liu, C.L.; Li, K. Featured temporal segmentation method and AdaBoost-BP detector for internal leakage evaluation of a hydraulic cylinder. *Measurement* **2018**, *130*, 279–289. [CrossRef]
- 45. Tharwat, A.; Gaber, T.; Hassanien, A.E.; Elhoseny, M. Automated toxicity test model based on a bio-inspired technique and AdaBoost classifier. *Comput. Electr. Eng.* **2018**, *71*, 346–358. [CrossRef]
- 46. Wang, L.; Lv, S.X.; Zeng, Y.R. Effective sparse adaboost method with ESN and FOA for industrial electricity consumption forecasting in China. *Energy* **2018**, *155*, 1013–1031. [CrossRef]
- 47. Chen, Y.; Wu, W. Mapping mineral prospectivity by using one-class support vector machine to identify multivariate geological anomalies from digital geological survey data. *Aust. J. Earth Sci.* **2017**, *64*, 639–651. [CrossRef]
- 48. Jove, E.; Casteleiro-Roca, J.L.; Casado-Vara, R.; Quintian, H.; Perez, J.A.M.; Mohamad, M.S.; Calvo-Rolle, J.L. Comparative Study of One-Class Based Anomaly Detection Techniques for a Bicomponent Mixing Machine Monitoring. *Cybern. Syst.* **2020**, *51*, 649–667. [CrossRef]
- 49. Al Shorman, A.; Faris, H.; Aljarah, I. Unsupervised intelligent system based on one class support vector machine and Grey Wolf optimization for IoT botnet detection. *J. Ambient Intell. Humaniz. Comput.* **2019**, *11*, 1–17. [CrossRef]
- 50. Gonzalez-Abril, L.; Angulo, C.; Nunez, H.; Leal, Y. Handling binary classification problems with a priority class by using Support Vector Machines. *Appl. Soft Comput.* **2017**, *61*, 661–669. [CrossRef]
- Chen, Y.L.; Wu, W.; Zhao, Q.Y. A Bat Algorithm-Based Data-Driven Model for Mineral Prospectivity Mapping. *Nat. Resour. Res.* 2020, 29, 247–265. [CrossRef]
- Chen, Y.L.; Wu, W.; Zhao, Q.Y. A Bat-Optimized One-Class Support Vector Machine for Mineral Prospectivity Mapping. *Minerals* 2019, 9, 317. [CrossRef]
- 53. Gu, Y.; Budati, C. Energy-aware workflow scheduling and optimization in clouds using bat algorithm. *Fut. Gener. Comput. Syst.* **2020**, *113*, 106–112. [CrossRef]
- 54. Tawhid, M.A.; Ali, A.F. Multi-directional bat algorithm for solving unconstrained optimization problems. *Opsearch J. Oper. Res. Soc. India* **2017**, *54*, 684–705. [CrossRef]
- 55. Kaced, K.; Larbes, C.; Ramzan, N.; Bounabi, M.; Dahmane, Z.E. Bat algorithm based maximum power point tracking for photovoltaic system under partial shading conditions. *Sol. Energy* **2017**, *158*, 490–503. [CrossRef]
- 56. Riffi, M.E.; Saji, Y.; Barkatou, M. Incorporating a modified uniform crossover and 2-exchange neighborhood mechanism in a discrete bat algorithm to solve the quadratic assignment problem. *Egypt. Inform. J.* **2017**, *18*, 221–232. [CrossRef]
- 57. Krishnanand, K.N.; Debasish, G. Formations of minimalist mobile robots using local-templates and spatially distributed interactions. *Robot. Auton. Syst.* 2005, 53, 194–213. [CrossRef]
- 58. Kaveh, A.; Seddighian, M.R. Simultaneously multi-material layout, and connectivity optimization of truss structures via an Enriched Firefly Algorithm. *Structures* **2020**, *27*, 2217–2231. [CrossRef]
- 59. Kumar, R.; Talukdar, F.A.; Rajan, A.; Devi, A.; Raja, R. Parameter optimization of 5.5 GHz low noise amplifier using multi-objective Firefly Algorithm. *Microsyst. Technol.* **2018**, *26*, 3289–3297. [CrossRef]
- Dekhici, L.; Belkadi, K. A Firefly Algorithm for the Mono-Processors Hybrid Flow Shop Problem. Int. J. Adv. Comput. Sci. Appl. 2017, 8, 424–433. [CrossRef]
- Kora, P. ECG based Myocardial Infarction detection using Hybrid Firefly Algorithm. *Comput. Methods Progr. Biomed.* 2017, 152, 141–148. [CrossRef]
- 62. Zhao, Z.M.; Ma, H.D.; Wang, B.Z.; Bai, Y.S.; Li, R.S.; Ji, W.H. The evidence of intrusive rocks about collision-orogeny during early Devonian in eastern Kunlun area. *Geol. Rev.* 2008, 54, 47–56. [CrossRef]
- 63. Du, Y.L.; Jia, Q.Z.; Han, S.F. Mesozoic tectono-magmatic-mineralization and copper-gold polymetallic ore prospecting research in east Kunlun metallogenic belt in Qinghai. *Northwest. Geol.* **2012**, *45*, 69–75.
- 64. Zheng, J.K. Regional tectonic evolution of east Kunlun. *Qinghai Geol.* 1992, 1, 15–25.

- 65. Dai, M.F.; Wang, S.L. Metallogenic background and prospect analysis of Lalingzaohuo region in Qinghai Province. *Qinghai Sci. Technol.* **2011**, *18*, 11–15. (In Chinese)
- 66. Chen, J.; Xie, Z.Y.; Li, B.; Tan, S.X.; Ren, H.; Zhang, Q.M.; Li, Y. Petrogenesis of Devonian intrusive rocks in theLalingzaohuo area, eastern Kunlun, and its geological significance. *J. Mineral. Petrol.* **2013**, *33*, 26–34. [CrossRef]
- 67. Du, W.; Ling, J.L.; Zhou, W.; Wang, Z.X.; Xia, Z.D.; Xia, M.Z.; Fan, Y.Z.; Jiang, C.Y. Geological characteristics and genesis of Xiarihamu nickel deposit in east Kunlun. *Miner. Depos.* **2014**, *33*, 713–726. [CrossRef]
- 68. Lin, N.; Chen, Y.L.; Lu, L.J. Mineral potential mapping using a conjugate gradient logistic regression model. *Nat. Res.* **2020**, 29, 173–188. [CrossRef]
- 69. Bradley, A.P. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognit.* **1997**, 30, 1145–1159. [CrossRef]
- 70. Bergmann, R.; Ludbrook, J.; Spooren, W.P.J.M. Different outcomes of the wilcoxon-mann-whitney test from different statistics packages. *Am. Stat.* 2000, *54*, 72–77. [CrossRef]
- 71. Pinsky, P.F. Scaling of true and apparent ROC AUC with number of observations and number of variables. *Commun. Stat.-Simul. Comput.* **2005**, *34*, 771–781. [CrossRef]
- 72. Kurum, E.; Yildirak, K.; Weber, G.W. A classification problem of credit risk rating investigated and solved by optimisation of the ROC curve. *Cent. Eur. J. Oper. Res.* **2012**, *20*, 529–557. [CrossRef]
- 73. Yu, W.; Chang, Y.C.I.; Park, E. A modified area under the ROC curve and its application to marker selection and classification. *J. Korean Stat. Soc.* **2014**, *43*, 161–175. [CrossRef]
- 74. Dragos, C. ROC curve for discrete choice models an application to the Romanian car market. *Appl. Econ. Lett.* **2010**, *17*, 75–79. [CrossRef]
- Obuchowski, N.A.; Bullen, J.A. Receiver operating characteristic (ROC) curves: Review of methods with applications in diagnostic medicine. *Phys. Med. Biol.* 2018, 63, 07TR01. [CrossRef] [PubMed]
- Chen, Y.L.; Wu, W. A prospecting cost-benefit strategy for mineral potential mapping based on ROC curve analysis. *Ore Geol. Rev.* 2016, 74, 26–38. [CrossRef]
- 77. Kottas, M.; Kuss, O.; Zapf, A. A modified Wald interval for the area under the ROC curve (AUC) in diagnostic case-control studies. *BMC Med. Res. Methodol.* 2014, 14, 26–30. [CrossRef] [PubMed]
- Yousefi, M.; Carranza, E.J.M. Prediction-area (P-A) plot and C-A fractal analysis to classify and evaluate evidential maps for mineral prospectivity modeling. *Comput. Geosci.* 2015, 79, 69–81. [CrossRef]
- 79. Yousefi, M.; Carranza, E.J.M. Fuzzification of continuousvalue spatial evidence for mineral prospectivity mapping. *Comput. Geosci.* 2015, 74, 97–109. [CrossRef]
- 80. Yousefi, M.; Carranza, E.J.M. Data-driven index overlay and boolean logic mineral prospectivity modeling in greenfields exploration. *Nat. Resour. Res.* **2016**, *25*, 3–18. [CrossRef]