Research Article

Feature Selection Based on BP Neural Network and Adaptive Particle Swarm Algorithm

Bingsheng Chen , Huijie Chen, and Mengshan Li

College of Physics and Electronic Information, Gannan Normal University, Ganzhou 341000, Jiangxi, China

Correspondence should be addressed to Bingsheng Chen; chenbingsheng@gnnu.edu.cn

Received 14 April 2021; Accepted 24 May 2021; Published 1 June 2021

Academic Editor: Sang-Bing Tsai

Copyright © 2021 Bingsheng Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Feature selection can classify the data with irrelevant features and improve the accuracy of data classification in pattern classification. At present, back propagation (BP) neural network and particle swarm optimization algorithm can be well combined with feature selection. On this basis, this paper adds interference factors to BP neural network and particle swarm optimization algorithm to improve the accuracy and practicability of feature selection. This paper summarizes the basic methods and requirements for feature selection and combines the benefits of global optimization with the feedback mechanism of BP neural networks to feature based on backpropagation and particle swarm optimization (BP-PSO). Firstly, a chaotic model is introduced to increase the diversity of particles in the initial process of particle swarm optimization, and an adaptive factor is introduced to enhance the global search ability of the algorithm. Then, the number of features is optimized to reduce the number of features on the basis of ensuring the accuracy of feature selection. Finally, different data sets are introduced to test the accuracy of feature selection, and the evaluation mechanisms of encapsulation mode and filtering mode are used to verify the practicability of the model. The results show that the average accuracy of BP-PSO is 8.65% higher than the suboptimal NDFs model in different data sets, and the performance of BP-PSO is 2.31% to 18.62% higher than the benchmark method in all data sets. It shows that BP-PSO can select more distinguishing feature subsets, which verifies the accuracy and practicability of this model.

1. Introduction

Image retrieval and other application fields are emerging [1]. In these problems, the data is often cumbersome, and the number of features is large. Therefore, there are higher requirements for feature selection, and feature selection methods for complex data emerge as the times require [2]. Although these algorithms have certain search ability, the efficiency is not high, and the waste of resources is serious, so we need to have a more efficient search strategy applied to feature selection.

In the research of search algorithm, Wang et al. combine genetic algorithm (GA) and feature subset selection. Compared with heuristic search strategy, genetic evolutionary algorithm has stronger search ability [3]. Yu et al. use the classification performance of support vector machine (SVM) to evaluate the selected feature subset, which is

similar to the feature selection method based on decision tree [4]. Wang and Cai proposed a stochastic subspace ensemble learning algorithm, whose principle is to train multiple individual samples by using different feature subsets generated randomly, so as to enhance individual differences [5]. Wang et al. suggested that training sets for BP neural networks can be basically divided into two categories according to different evaluation mechanisms: wrappers and filters [6]. In order to better improve the utilization rate of features, Aguirre et al. proposed to use information theory measure to select related features. Compared with the previous threshold method, information theory measure can improve the quality of feature subset [7].

For the research on the update mechanism of particle swarm optimization [8], Ali et al. introduce gene expression data into the update mechanism of particle swarm optimization in the feature selection method of particle swarm

optimization based on gene expression data to improve the efficiency of the algorithm [9]. Jia et al. proposed a feature selection method of wrapper class based on binary particle swarm optimization algorithm and adopted the evaluation mechanism of encapsulation mode based on support vector machine [10]. Chen et al. proposed that by simulating the group behavior of bird foraging and fish swarm activities, a single particle uses the information of other particles in the population to search the target solution [11]. Al-Musaylh et al. find the optimal solution of the objective function by iteratively searching the feasibility space. Similar to other evolutionary algorithms (genetic algorithms), particle swarm optimization also belongs to swarm intelligence algorithm, but particle swarm optimization does not use the group competition mechanism to find the optimal solution of the objective function; instead, it uses the group cooperation mechanism [12]. Cao et al. proposed a feature selection method of particle swarm optimization based on new initialization and update mechanism, which imitates the typical forward and backward search process. In the update mechanism of local and global optimal position of particle swarm optimization, the new update mechanism improves the efficiency of particle swarm optimization [13]. The above studies put a lot of effort into controlling the mechanism of the initial and update process of particle swarm optimization and effectively improve the efficiency of feature selection, but when combined with the BP neural network, the neural network converges. It reduces the overall performance of the model to ensure accuracy [14].

This paper uses BP neural network to format the feature data set and then establishes the search strategy for the formatted data set. In this paper, interference factors are added to BP neural network and particle swarm optimization algorithm to improve the accuracy and practicability of feature selection. This paper summarizes the basic methods and requirements of feature selection and proposes a BP-PSO feature selection model combining the advantages of BP neural network feedback mechanism and global optimization. On the basis of ensuring the accuracy of element selection, the number of elements is optimized and the number of elements in the data is reduced. Finally, a variety of data sets are introduced to test the accuracy of model feature selection, and encapsulation pattern and filter pattern evaluation mechanism are used to verify the effectiveness of the model.

2. Feature Selection and Particle Swarm Intelligence Algorithm

2.1. Feature Selection Strategy. All kinds of search strategies have their own advantages and disadvantages. Swarm intelligence optimization algorithm has become a research hotspot because of its good trial and search ability. According to the evaluation function, several features are selected as feature subsets, which are regarded as a search strategy with limited length to evaluate whether a search strategy is effective. The criterion is whether the strategy can improve the ability and speed of finding the optimal feature subset [15, 16].

2.1.1. Fast Feature Selection Method. The feature selection method of encapsulation model uses the actual training model to measure the usefulness of feature subset (also called supervised learning), and the obtained model has better utilization value [17]. In the screening mode, the evaluation of feature subset is independent of a specific learning algorithm. Feature quality is measured by analyzing the internal properties of a feature subset. The number of one or more thresholds to select is set, and each feature is weighted. This represents the importance of dimensional features, which are sorted according to weight [18]. The filtering pattern is generally used for preprocessing, which has nothing to do with the selection of classifier, so the classification performance of the feature subset obtained is worse than that of the encapsulation pattern, but it is faster and more robust [19, 20]. Robust matrix factorization is as follows:

$$\Delta L^{l3} = -\eta \delta^{l3} X^{l3} = \eta(t - y) f'(X^{l3} W^{l3}) X^{l3}. \quad (1)$$

If the element in L is Q , the conversion relation is

$$Q(Y) = \frac{\sigma(Y) - \text{avg}(\sigma(Y'), \sigma(Y'))}{\sigma(Y)}. \quad (2)$$

Q is the dimension of the feature and the number of samples. Each sample point can be expressed as

$$X = (i, j)^{l3} = (t - y) f'(X^{l3} W^{l3}). \quad (3)$$

The matrix X is decomposed into

$$M_{ik} = \sum_a^n \tau_1 X_{ik} + \sum_b^C \tau_2 U(Y_{ik}) * Q_{ik}. \quad (4)$$

M is the potential feature matrix (feature selection center), and a and b are the class indicator matrix (C is the number of feature selection). The matrix U with orthogonal and nonnegative constraints can be regarded as a scaled feature selection index matrix [21]. Orthogonal constraints and nonnegative constraints exist at the same time to improve the sparsity of the matrix, and nonnegative constraints are more in line with the actual situation. Since the norm is used as the loss function, it is easy to be affected by noise points and outliers [22]. In this paper, the norm is used to replace the common norm to avoid large loss, which makes the proposed model more robust. It is expressed as follows:

$$U = \arg \max \sum_{a_n \in W_K(a)} |(b_n = c_n)M|. \quad (5)$$

The feature selection index matrix can be obtained by optimizing formula (5), and then the feature selection index indication matrix is used to guide feature selection. However, this is not enough. More information needs to be added to improve the quality of the selected feature subset [23]. The local geometry of data is very important in BP-PSO feature selection task. In this paper, adaptive graph learning is used to extract the sample similarity based on the internal geometric structure of the data, which makes the feature selection index matrix closer to the real category of the data,

achieves the ideal feature selection effect, and finally improves the performance of feature selection [24].

2.1.2. Adaptive Graph Construction. The traditional similarity between samples needs prelearning, and because the similarity between samples is learned from the original feature space, it is easily affected by noise and redundant features [25]. Therefore, in this paper, you can use the adaptive graph method to repeatedly update the similarity graph and learn the local structure while performing feature selection. The methods proposed in the literature are used for data processing. The similarity probability between sample points is defined as S_{ij} , which is an element of the similarity matrix R [26]. For simplicity, the square of Euclidean distance is used; that is, the distance between two sample points is calculated. The similarity matrix s of samples can be obtained by the following formula:

$$S_{ij}(v, e) = [\zeta_1 c_1(t) + \zeta_2 c_2(k) + \zeta_3 c_3(k) + \zeta_4 c_4(k) + \zeta_5 c_5(k) + \zeta_6 w_{ik}], \quad (6)$$

where w is the regularization parameter to avoid invalid solutions. Obviously, the regular term c can be used to adjust the number of neighbor nodes. The optimal value of c should be such that most vectors v contain only k nonzero elements, where k is the number of neighbors connected to $c(t)$.

$$\begin{aligned} c_1(t) &\geq 0, \\ c_2(k) &\geq 0, \\ c_3(k) &\geq 0, \\ c_4(k) &\geq 0, \\ c_5(k) &\geq 0. \end{aligned} \quad (7)$$

After the similarity matrix is obtained, a priori constraint can be added to improve its reliability. In the next section, an adaptive graph optimization based on ideal local structure is proposed to further improve the correctness of adaptive graph [27]. For the similarity matrix, the ideal state is to contain only a connected component, that is, a cluster. However, the similarity matrix learned by local structure is almost impossible to be in this state, that is, aware:

$$A_{\lim} = f(x) = \sum_{j \in Q} c_j \frac{x_j}{\sigma(X'_j)} - p - B * L. \quad (8)$$

In formula B, l is the Laplacian matrix obtained from the similarity matrix. a is a Laplacian matrix, and the degree matrix $f(x)$ is a diagonal matrix. It can be proved that if the characteristic variables are unchanged, then the similarity matrix will have p connected components. Therefore, we can add a priori constraint to local structure learning to obtain ideal local structure information [28]. Since the a priori constraint depends on the similarity matrix s , equation (8) is difficult to solve. The constraint can be simplified to

$$\begin{aligned} D &= N(d_i, w_j) = P(d_i)P(w_j|d_i); P(w_j|d_i) \\ &= \sum_{k=1}^K P(w_j|z_k)P(z_k|d_i), \\ V &= \{(a_1, b_1), (a_2, b_2), \dots, (a_n, b_n)\}. \end{aligned} \quad (9)$$

As long as V is a large enough value, it tends to zero. By changing the problem from the original rank constraint to the solution of trace, the complexity of the original problem is greatly simplified [29]. Through this method, we can learn the ideal local structure information, make the result of feature selection more accurate, and finally improve the performance of feature selection.

2.2. Particle Swarm Optimization Intelligent Algorithm and Feature Selection Mutation Control. Although the particle swarm optimization algorithm has shown good performance, the feature selection method based on particle swarm optimization still has a lot of room for improvement: in the existing feature selection method based on particle swarm optimization algorithm, the optimization algorithm and feature selection are independent of each other, and the particle swarm optimization algorithm is only used as a tool to rely on its own search ability for the evaluation process. The verification function provides a subset of features, which is not substantially combined with feature selection [30]. Feature selection is a complex issue. This is to select a better subset in the original data set (by evaluating the subset using classification performance) and ensure the simplification of the subset reduces the number of features in the subset. The benefits of optimization algorithms are reflected not only in good search capabilities, but also in combination with the target problem. The existing research on feature selection based on optimization algorithm often ignores this problem.

The cluster structure of feature selection method is determined by Laplacian graph and adaptive discriminant regularization, which is dynamic. Through the feedback of feature selection results, this kind of method can get better feature selection analysis. In order to improve the accuracy of calculation and control efficiency of practical engineering problems, RNG $K - \varepsilon$ model is selected as turbulence control model after comparison. The two basic equations, K equation and ε equation, are as follows:

$$\begin{aligned} \text{RNG } k - \varepsilon &= \frac{e^x - e^{-x}}{e^x + e^{-x}}, \\ \varepsilon &= \frac{x}{1 + |x|} * K. \end{aligned} \quad (10)$$

Based on the above feature selection methods, a good feature selection method needs to extract reliable data geometric structure and reliable cluster structure to guide the feature selection process. However, the existing algorithms either focus on the extraction of ideal data geometric structure or focus on robust feature selection to guide better feature selection (rufsm) and rarely take these two aspects into account at the same time. Based on the above algorithm, this paper proposes to establish a unified

framework of special selection method, which integrates adaptive local structure learning, discriminant information extraction, and feature selection into the framework. Therefore, before information fusion, the original signal G must be preprocessed strictly:

$$G = \frac{1}{2}(t - \gamma)^2 = \frac{1}{2}[t - f(WX)]^2. \quad (11)$$

In this method, the discriminant information is obtained by robust matrix decomposition, and the local structure information in the data set is learned by adaptive graph embedding, and a priori constraint is added to the learned similarity matrix without preconstructing the similarity matrix between samples. The influence of noise and redundant features is reduced; thus, the robustness of the model is improved and the learned local structure information is improved more accurate. Finally, the feature selection function is realized by applying row sparse norm to the feature selection matrix. The feature vectors obtained by multisensor synthesis are as follows:

$$L = \Delta N^l = -\eta \frac{\partial E}{\partial W^l} = \eta (X^l)^T \delta^l + B, \quad (12)$$

where n is the number of extracted eigenvalues and W is multiple samples. The dimension of each element in the eigenvector matrix L is still not uniform, which does not conform to the data format processed by neural network. Because the neural network focuses on capturing the differences between different types of data, and the feature vectors only compare the elements in the same position, it is necessary to convert the values of each element from absolute value to relative value to meet the needs of neural network. Let the set of elements in each column be a . The normalized transformed characteristic matrix is n . The relationship between N and B is shown in the following formula:

$$\Delta Q_i = N_j - \eta E' = \eta x_i(t - \gamma)f'(WX) = \eta x_i \delta. \quad (13)$$

The transformed eigenvalues are all greater than 0, and the maximum value of each column element is still uncertain. It is necessary to further normalize the elements in B before they can be input into the neural network. The traditional normalization method usually converts all the data to a specific interval by linear or nonlinear processing. The premise of this method is that you know the numerical range of the data so that you can compare the data and perform global normalization, but the data is collected by two types of sensors that consist of three characteristic values, and the numeric range of each column of data does not change. Here, the value of Q is uncertain; that is, the numerical range of different columns of data is different, so it cannot be compared with each other. The data processed column by column can still reflect the differences between different rows of data. Therefore, the eigenvalues are further converted to the same range by column normalization.

3. Feature Selection Research Design

3.1. Content. In this paper, interference factors are added to BP neural network and particle swarm optimization algorithm to improve the accuracy and practicability of feature selection. This paper summarizes the basic methods and requirements of feature selection and then proposes a feature selection model based on BP-PSO combining the feedback mechanism of BP neural network and the data global optimization advantage of particle swarm optimization. Firstly, chaotic model is introduced to increase the diversity of particles in the initial process of particle swarm optimization, and adaptive factor is introduced to enhance the global search ability of the algorithm. Then, the number of features is optimized to reduce the number of features on the basis of ensuring the accuracy of feature selection. Finally, different data sets are introduced to test the accuracy of feature selection, and the evaluation mechanisms of encapsulation mode and filtering mode are used to verify the practicability of the model.

3.2. Design. In order to have an objective understanding of the performance of BP-PSO model, all features will be used as the benchmark in the experiment and compared with other five related adaptive feature selection algorithms. The control models involved in the experiment were baseline, lapscore, UDFs, NDFs, fsasl, and sogfs. In order to ensure the fairness and effectiveness of the comparative experiment, when the nearest neighbor parameter needs to be set in advance in the experiment, it is set to $k=5$. And the parameter size in the Gaussian heat kernel function is set to 1.

The feature data set is formatted by BP neural network, and then the search strategy is established for the formatted data set. Six data sets (PalmData25, Ecoli, Isolet, Jaffe, Yale, and Coil20) are introduced to evaluate the accuracy and practicability of BP-PSO model. The simplified model of BP model introduced in this paper is shown in Figure 1.

In addition, the dimension space of the feature selection matrix (projection matrix) and the number of potential clusters are set to C (C is the number of real categories of the data set). BP-PSO model also needs to adjust three parameters. According to the different substeps involved in the feature selection process, these embedded methods can be further divided into the following different types. The first type of embedded method first detects the structure of the data and then directly selects those features that can be used to best retain the closed structure. Typical methods include trace ratio criterion and unsupervised discriminant feature selection (UDFs). The second method of embedded feature selection is to first create various Laplacian graphs to obtain sample similarity information and then to detect the cluster structure of the data by spectral analysis. Then sparse spectral regression gets the sparse feature selection matrix. Finally, the feature selection matrix selects the features with the highest scores, that is, the features that best fit the cluster

FIGURE 1: BP model simplified mode.

structure. These clustering structures, which are mined by graph embedding or other feature selection methods, can be regarded as the approximate process of real data labels.

4. Results and Discussion

4.1. Feature Selection Accuracy of BP-PSO for Different Data Sets. In this section, comparative experiments are used to verify the effectiveness of the features selected by BP-PSO model. A total of 6 data sets are used for feature selection analysis.

As shown in Figure 2, most of the adaptive feature selection algorithms have better performance than baseline. It shows that there are a lot of redundancy and noise features in the original data. The performance of the learner can be improved by feature selection. The importance of feature selection is explained. Lapscore's feature selection results are not satisfactory, because its feature selection strategy is to select features one by one. This feature selection method ignores the relationship between features, and the selected features are more redundant than other models.

As shown in Table 1, the average accuracy of BP-PSO is 8.65% higher than that of the suboptimal NDFs model. In addition, it is worth noting that the performance of BP-PSO algorithm on all data sets is improved by 2.31% to 18.62% compared with the benchmark method. This shows that BP-PSO can select a more discriminating feature subset. From the experimental results, the relationship between the number of features and the results of feature selection can be further analyzed. The best result on PalmData25 is slightly better than other data sets because of the low feature dimension and low feature redundancy. The difficulty of feature selection is relatively high.

Through repeated use of *K*-means feature selection 20 times, take the average value, and take the dimension with the best results to make a three-dimensional map. The results of feature selection of BP-PSO corresponding to different parameters on six public data sets are shown in Figure 3. From the feature selection NMI of Ecoli and isolet data sets, it can be concluded that it is necessary to select parameters for data, and different parameter combinations may have a huge impact on the results. The performance of ACC and NMI under the same parameter combination may be greatly different. This is because ACC and NMI are two completely different evaluation indexes. Different evaluation indexes result in different evaluation results. In the experiment, BP-

PSO needs to set some parameters in advance, mainly focusing on the influence of two main regularization parameters on the experimental results, that is, to ensure the local similarity of data and to control the sparse constraint of feature selection matrix.

The feature decomposition of 3D model is shown in Figure 4. On six data sets, ACC is selected for the feature corresponding to different values. Different combinations of parameters have different results. The results of the algorithm on PalmData25 and Coil20 are stable. This may be due to the large number of samples in PalmData25 and Coil20 data sets and the low difficulty of feature selection.

As shown in Figure 5, for LIHC, the feature selected by Chi2 and DTree classifier can achieve the maximum accuracy of 6.32%. For the features selected by RF or SVM-RFE and DTree classification model, the maximum accuracy of each data set is improved by at least 1.25%. If the best model of any feature selection algorithm and classification algorithm is used as the final model of the data set, the six cancer data sets can be improved by BP-PSO coding features. The accuracy of the final model PRAD and THCA has increased by 0.23%. The final model of the data set BRCA has an improved maximum accuracy of 0.36%. At the same time, from the final results, the four combinations of feature selection algorithms *t*-test and RF-fs (representing feature selection algorithm RF) and the classifiers DTree and RF can give the best experimental results. Therefore, researchers can consider them as further experimental schemes.

As shown in Table 2, compared with the original methylation features of all six data sets, BP-PSO encoded features have better feature selection accuracy in general. PRAD data set is more difficult in the problem of binary classification, and the feature selection accuracy of BP-PSO coding or original feature is not higher than 0.9500. The *t*-test *p* value of THCA with only one BP-PSO coding feature is more significant than the original feature. Only the top BP-PSO coding features can achieve the feature selection accuracy of 0.9734, which is better than the best feature selection accuracy of 0.9698 obtained by using 8 original features.

As shown in Figure 6, significant improvement can be achieved in LIHC, PRAD, and THCA data sets. Without using feature selection algorithm, the best results of the three data sets are 0.96, 0.98, and 0.99. Therefore, in the feature selection step, these three data sets are not much improved. The accuracy of LIHC, PRAD, and THCA in the other three data sets was improved by 3.5129%, 4.3478%, and 6.7496%, respectively. Therefore, in many cases, the feature selection of BP-PSO coding can be used to improve the feature selection model.

The BP-PSO coding features are shown in Figure 7. Further experiments are carried out to evaluate whether the BP-PSO coding model can get useful information from the lower ranking features. Three groups of 500 features with *t*-test ranking are selected for BP-PSO coding. Then, the three BP-PSO coded feature sets are compared with the top 500 *t*-test ranked original features.

As shown in Figure 8, when detecting the image feature data set, the BP-PSO encoded feature is superior to the

FIGURE 2: Unsupervised feature selection algorithm. (a) Structure. (b) Particle swarm algorithm simulation.

TABLE 1: Average accuracy of different data sets.

Item	PalmData25	LIHC	PRAD	<i>t</i> -test	Wilcoxon test	Original feature
BP-PSO	1.09	1.91	1.55	1.78	1.05	1.99
DTree	1.72	1.09	2.33	1.55	3.57	2.76
NDFS	4.43	4.22	5.01	4.37	5.33	5.9
RF	1.21	3.41	2.59	4.61	2.14	5.44
BP-PSO	1.42	2.64	2.54	3.42	4.75	1.81
THCA	6.24	3.35	5.85	1.12	5.51	3.33

FIGURE 3: The result of feature selection for different parameters.

original methylation feature. Using any of the six feature selection algorithms, the two classifiers logistic R and SVM have not improved the accuracy of feature selection in image feature data set. Moreover, the features encoded by BP-PSO cannot be used to improve NBayes model. However, the best NBayes model using original features can only achieve the accuracy of 0.75, while the feature selection algorithm using t -test or Wilcoxon test for BP-PSO coding features can improve the accuracy of NBayes model to 0.8053. The model using lr-rfe and DTree even achieves the maximum improvement of 16.4%. The BP-PSO coding features obtained

by RF classifier and rf-fs feature selection algorithm achieve the best feature selection accuracy of 0.86.

As shown in Table 3, BP-PSO model has a good improvement effect when using t -test and RF feature selection algorithms. Therefore, the experiment in this section uses these two feature selection algorithms. Compared with other experiments, the classification effect of the original feature and the BP-PSO coding feature is compared.

As shown in Figure 9, ACC and NMI do not simply increase with the increase of feature dimension, but start to stabilize or even fluctuate after entering a certain feature

FIGURE 4: Feature decomposition of 3D model.

FIGURE 5: Absolute value of sensitivity varies with eccentricity.

TABLE 2: Original methylation characteristics of data sets.

Item	Accuracy	Improved NBayes	LIHC	PRAD	<i>t</i> -test	Wilcoxon test
Original	1.24	0.6	1.1	0.91	0.61	1.66
DTree	3.2	2.54	2.97	3.96	1.92	1.86
SVM-RFE	6	4.27	5.38	2.93	2.98	5.85
RF	5.34	1.63	4.66	2.54	3.19	1.33
BP-PSO	1.82	1.06	1.63	1.23	4.47	3.95
THCA	3.55	2.68	6.74	1.23	6.62	3.41

dimension. This shows that there are few effective features in the data, and most of them are redundant and invalid. Too many features do not only bring computational burden, but also may bring negative effects. This also shows the necessity of feature selection. Sogfs, fsasl, and BP-PSO, which are local manifold models with adaptive learning data, perform better than other models. The importance of local geometric information in adaptive feature selection is illustrated. The curve of BP-PSO is basically above other algorithms, which shows that the full use of local geometric structure and discriminant information is helpful to select a good feature subset.

As shown in Figure 10, the features of BP-PSO coding are not significantly improved in the pems-sf data set, and the overall effect is similar to or slightly smaller than the feature selection accuracy of the original features. However, when using the logistic *r* classifier, the features of BP-PSO coding are 6.95% higher than the original features. Overall, the feature selection accuracy of BP-PSO coding is 0.96. The feature of BP-PSO coding has been greatly improved in the image data set, and in many cases, it has been improved by more than 10%. The best feature selection accuracy rate is 0.94, which is much higher than the best feature selection accuracy rate of 0.89 of the original feature.

FIGURE 6: Live-action image special conquest effect.

FIGURE 7: Unsupervised feature selection algorithm.

FIGURE 8: Sudden change of control sensitivity.

TABLE 3: Classification effect of the original features on the data set.

Item	BP-PSO	THCA	SAE	LIHC	PRAD	Chi2
DTree	0.86	1.51	5.08	1.4	3.17	4.8
SVM-RFE	0.8	3.43	4.89	3.06	2.05	1.73
RF	1.32	3.91	5.49	3.42	1.01	1.86
<i>t</i> -test	0.31	1.28	2.88	3.89	1.08	3.01
Wilcoxon test	0.01	1.39	2.5	4.17	1.12	5.89
Original feature	1.08	2.27	3.59	3.28	3.24	6.61

FIGURE 9: Really valid features in the data.

FIGURE 10: Optimize ship operation management.

As shown in Figure 11, the swarm behavior data set using BP-PSO also has a good improvement effect, especially when using NBayes classifier, the improvement effect even reaches 32.5%. Overall, the feature selection accuracy of BP-PSO coding is 0.99. A positive value indicates that the features encoded by BP-PSO outperform the original features with the same rank. BP-PSO can get more useful features from BP coding.

4.2. Discussion. This paper studies the feature selection performance of BP-PSO coding features in different feature selection problems. The feature data set has four

development stages, I/II/III/IV, which can be retrieved from TCGA database. In this paper, a feature engineering algorithm BP-PSO based on sparse self-encoder is proposed to enrich the original methylation features, so as to distinguish TCGA cancer samples from normal samples. BP-PSO can even obtain useful information from the lower-ranking features and outperform the feature selection results of the highest-ranking original features. Compared with the original methylation features, BP-PSO coding features show better performance in difference analysis, feature selection, and classification. In addition to the conventional problems, the experimental results show that the model also has a good effect for BRCA and other heterologous cancers, and the BP-PSO coding features can prove the accuracy of feature selection used to improve the image feature data set (THCA). At the same time, you can apply the BP-PSO model to some engineering problems for excellent feature selection performance. Experimental data show that the combination of BP-PSO coding capabilities and traditional feature selection algorithms can generate feature selection models with excellent performance. The classical rough set theory is based on the equivalence relation, and the values it deals with are discrete values. In reality, there are many attributes with continuous values. The classical rough set theory usually discretizes the continuous values, which will lead to the loss of information. The real field rough set model can solve this problem, and the real field rough set model is based on the generalized attribute the extension under the importance measure criterion. Rough set is not suitable to deal with

FIGURE 11: Operation management during ship navigation.

large data set because it needs to traverse the whole space to get the optimal subset. The calculation of generalized importance can aim at the continuous attribute value and solve the problem of information loss caused by the classical rough set to the data discretization processing, which affects the processing results.

5. Conclusions

Feature selection is a mature data preprocessing method, which can effectively remove redundant features and noise and then achieve the purpose of reducing the data dimension and retaining the important information of the data at the same time. The feature selection process is guided by the feature selection information obtained by matrix decomposition and the local structure information obtained by adaptive graph. In order to prevent overfitting, the redundant features in the data are further removed by row sparse norm to improve the effect of feature selection. Compared with other unsupervised feature selection models, BP-PSO model improves ACC and NMI to a certain extent. The model has higher accuracy and higher robustness and is insensitive to parameters. BP-PSO, like NDFs, can obtain discriminative information by feature selection. However, NDFs spectral feature selection is in the original space, and there are a lot of noise and redundant features in the original space, which will affect the accuracy of the features learned. BP-PSO obtains feature selection label through robust matrix decomposition, which can effectively reduce the influence of noise and outliers on feature selection. Both BP-PSO and fsasl consider the influence of noise and outliers, and the models are robust. But the difference is that BP-PSO has more detailed prior constraints on the similarity matrix, so the structure information is more accurate. The matrix factorization model is used to select the features of a data point that is a linear model. However, the data is often embedded in low-dimensional manifolds, and there are complex nonlinear relationships between the data that cannot be processed by a linear matrix factorization model. The focus of this paper is on how to apply more advanced

techniques to feature selection such as tensor calculus and decomposing machines.

Data Availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant nos. 51663001, 52063002, and 42061067) and the Science and Technology Research Project of the Education Department of Jiangxi Province (Grant nos. GJJ180773 and GJJ180754).

References

- [1] N. Krishnaraj, M. Elhoseny, E. L. Lydia, K. Shankar, and O. ALDabbas, "An efficient radix trie-based semantic visual indexing model for large-scale image retrieval in cloud environment," *Software Practice and Experience*, vol. 51, no. 9, 2020.
- [2] B. Cao, J. Zhao, P. Yang et al., "Multiobjective feature selection for microarray data via distributed parallel algorithms," *Future Generation Computer Systems*, vol. 100, pp. 952–981, 2019.
- [3] D. Wang, D. Tan, and L. Liu, "Particle swarm optimization algorithm: an overview," *Soft Computing*, vol. 22, no. 2, pp. 387–408, 2018.
- [4] H. Yu, Y. Tan, J. Zeng, C. Sun, and Y. Jin, "Surrogate-assisted hierarchical particle swarm optimization," *Information Sciences*, vol. 454–455, no. 4, pp. 59–72, 2018.
- [5] Z. Wang and J. Cai, "The path-planning in radioactive environment of nuclear facilities using an improved particle

- swarm optimization algorithm," *Nuclear Engineering and Design*, vol. 326, no. 1, pp. 79–86, 2018.
- [6] Wang, Hongman, Li et al., "Multi-objective test case prioritization based on multi-population cooperative particle swarm optimization," *The Journal of China Universities of Posts and Telecommunications*, vol. 27, no. 1, pp. 42–54, 2020.
 - [7] N. F. Aguirre, A. Morgenstern, M. J. Cawkwell, E. R Batista, and P Yang, "Development of density functional tight-binding parameters using relative energy fitting and particle swarm optimization," *Journal of Chemical Theory and Computation*, vol. 16, no. 3, pp. 1469–1481, 2020.
 - [8] M. H. Ali, B. Mohammed, M. Ismail et al., "A new intrusion detection system based on Fast Learning Network and Particle swarm optimization," *IEEE Access*, vol. 5, no. 9, pp. 11–16, 2018.
 - [9] M. A. El Aziz, A. M. Hemdan, A. A. Ewees et al., "Prediction of biochar yield using adaptive neuro-fuzzy inference system with particle swarm optimization," in *Proceedings of the 2017 IEEE PES PowerAfrica*, pp. 115–120, IEEE, Accra, Ghana, June 2017.
 - [10] Y.-H. Jia, W.-N. Chen, T. Gu et al., "A dynamic logistic dispatching system with set-based particle swarm optimization," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 48, no. 9, pp. 1607–1621, 2018.
 - [11] Y. Chen, L. Li, J. Xiao, Y. Yang, J. Liang, and T. Li, "Particle swarm optimizer with crossover operation," *Engineering Applications of Artificial Intelligence*, vol. 70, no. 4, pp. 159–169, 2018.
 - [12] M. S. Al-Musaylh, R. C. Deo, Y. Li, and J. F. Adamowski, "Two-phase particle swarm optimized-support vector regression hybrid model integrated with improved empirical mode decomposition with adaptive noise for multiple-horizon electricity demand forecasting," *Applied Energy*, vol. 217, no. 4, pp. 422–439, 2018.
 - [13] L. Cao, L. Xu, and E. D. Goodman, "A neighbor-based learning particle swarm optimizer with short-term and long-term memory for dynamic optimization problems," *Information Sciences*, vol. 453, no. 4, pp. 463–485, 2018.
 - [14] Y. Tang and M. Elhoseny, "Computer network security evaluation simulation model based on neural network," *Journal of Intelligent and Fuzzy Systems*, vol. 37, no. 3, pp. 3197–3204, 2019.
 - [15] A. Jp, A. Bp, C. Bkb et al., "Reliability analysis of pumping station for sewage network using hybrid neural networks - genetic algorithm and method of moment - ScienceDirect," *Process Safety and Environmental Protection*, vol. 145, no. 9, pp. 39–51, 2021.
 - [16] A. Ag, B. Pm, and B. Bp, "Graph neural network: current state of Art, challenges and applications," *Materials Today: Proceedings*, vol. 7, no. 4, pp. 96–102, 2021.
 - [17] S. Wang, J. Jiang, and X. Lu, "Study on the classification of pulse signal based on the BP neural network," *Journal of Biosciences and Medicines*, vol. 08, no. 5, pp. 104–112, 2020.
 - [18] J. Alnas, G. Giddings, and N. Jeong, "Bandwidth improvement of an inverted-F antenna using dynamic hybrid binary particle swarm optimization," *Applied Sciences*, vol. 11, no. 6, pp. 55–59, 2021.
 - [19] W. Liu, G. Niu, Q. Cao, M.-O. Pun, and J. Chen, "Particle swarm optimization for interference-limited unmanned aerial vehicle-assisted networks," *IEEE Access*, vol. 8, no. 4, pp. 174342–174352, 2020.
 - [20] M. Elhoseny and K. Shankar, "Optimal bilateral filter and convolutional neural network based denoising method of medical image measurements," *Measurement*, vol. 143, pp. 125–135, 2019.
 - [21] A. Msa, A. Gn, B. Ss et al., "Finite-time stability analysis of fractional-order memristive fuzzy cellular neural networks with time delay and leakage term," *Mathematics and Computers in Simulation*, vol. 185, no. 5, pp. 468–485, 2021.
 - [22] F. Zhao, S. Qin, G. Yang, W. Ma, C. Zhang, and H. Song, "A factorial based particle swarm optimization with a population adaptation mechanism for the no-wait flow shop scheduling problem with the makespan objective," *Expert Systems with Applications*, vol. 126, no. 1, pp. 41–53, 2019.
 - [23] A. Gd, B. Bp, B. Jpd et al., "Perceptually-guided deep neural networks for ego-action prediction: object grasping," *Pattern Recognition*, vol. 88, no. 52, pp. 223–235, 2019.
 - [24] Z. Tang and M. Gong, "Adaptive multifactorial particle swarm optimisation," *CAAI Transactions on Intelligence Technology*, vol. 4, no. 1, pp. 37–46, 2019.
 - [25] A. P. Engelbrecht, J. Grobler, and J. Langeveld, "Set based particle swarm optimization for the feature selection problem," *Engineering Applications of Artificial Intelligence*, vol. 85, no. 4, pp. 324–336, 2019.
 - [26] J.-P. Zhang, P.-F. Gao, and F. Fang, "An ATPSO-BP neural network modeling and its application in mechanical property prediction," *Computational Materials Science*, vol. 163, no. 6, pp. 262–266, 2019.
 - [27] Y. Mei, J. Yang, Y. Lu et al., "BP-ANN model coupled with particle swarm optimization for the efficient prediction of 2-chlorophenol removal in an electro-oxidation system," *International Journal of Environmental Research and Public Health*, vol. 16, no. 14, pp. 44–54, 2019.
 - [28] C. Guan and Y. Yang, "Adoption of back propagation network improved by particle swarm optimization in network intrusion detection," *Journal of Physics: Conference Series*, vol. 1650, no. 3, pp. 96–98, 2020.
 - [29] A. Eto, B. Ape, and C. Cwc, "Movement patterns of a particle swarm in high dimensional spaces," *Information Sciences*, vol. 512, no. 2, pp. 1043–1062, 2020.
 - [30] L. Yao, L. Ren, and G. Gong, "Evaluation of chloride diffusion in concrete using PSO-BP and BP neural network," *IOP Conference Series: Earth and Environmental Science*, vol. 687, no. 1, pp. 37–45, 2021.