



Applying particle swarm optimization-based decision tree classifier for wart treatment selection

Junhua Hu¹ · Xiangzhu Ou¹ · Pei Liang¹ · Bo Li¹

Received: 18 November 2020 / Accepted: 17 March 2021 / Published online: 12 April 2021
© The Author(s) 2021

Abstract

Wart is a disease caused by human papillomavirus with common and plantar warts as general forms. Commonly used methods to treat warts are immunotherapy and cryotherapy. The selection of proper treatment is vital to cure warts. This paper establishes a classification and regression tree (CART) model based on particle swarm optimisation to help patients choose between immunotherapy and cryotherapy. The proposed model can accurately predict the response of patients to the two methods. Using an improved particle swarm algorithm (PSO) to optimise the parameters of the model instead of the traditional pruning algorithm, a more concise and more accurate model is obtained. Two experiments are conducted to verify the feasibility of the proposed model. On the hand, five benchmarks are used to verify the performance of the improved PSO algorithm. On the other hand, the experiment on two wart datasets is conducted. Results show that the proposed model is effective. The proposed method classifies better than k-nearest neighbour, C4.5 and logistic regression. It also performs better than the conventional optimisation method for the CART algorithm. Moreover, the decision tree model established in this study is interpretable and understandable. Therefore, the proposed model can help patients and doctors reduce the medical cost and improve the quality of healing operation.

Keywords Wart treatment method · Classification · Decision tree · Particle swarm optimisation

Introduction

Wart is a skin disease caused by human papillomavirus (HPV) that infects the outer layer of the skin and induces excessive proliferation of skin cells [1]. Treatment is urgently needed by patients with warts because of the effect of this disease on human appearance and daily life. Methods clinically available for the effective treatment of warts include immunotherapy, antimitotic therapy and destructive methods. However, existing methods only eliminate or ease the symptoms and signs of warts [2]. In addition, patients

who catch the same wart but under different symptom conditions have different responses to the same treatment [3]. The different methods induce different levels of side effects, cost and pain. Therefore, the selection of method for different patients is important to maximise the elimination of wart symptoms, reduce side effects and pain and minimise the treatment cost.

In the traditional clinical treatment of warts, doctors choose a method for patients in a subjective manner. This method is not reliable especially when the doctor lacks clinical experience. In some cases, physicians are obliged to test each method individually, adding financial burden and pain to patients. Machine learning can effectively solve these problems. It can automatically learn from previous cases and assist doctors in choosing the appropriate treatment for patients. It effectively eliminates the symptoms of patients and avoids problems caused by multiple treatments.

In recent years, the machine learning and data mining algorithms employed in medical science have grown rapidly because they are intelligent in analysis and discovering knowledge from datasets [5–9]. Many studies used machine learning methods to treat skin diseases, such as melanoma

✉ Pei Liang
shirley_lp@csu.edu.cn

Junhua Hu
hujunhua@csu.edu.cn

Xiangzhu Ou
867817496@qq.com

Bo Li
151601008@csu.edu.cn

¹ School of Business, Central South University, Changsha 410083, China

[4–6], which is a type of skin cancer developing from melanocytes. Machine learning has also been used for skin cancer diagnosis [10–13], skin lesion classification [11, 14, 15] and skin sensitisation [16, 17]. A number of studies have compared the two treatment methods in terms of medical science [18–24], but a few studies have used machine-learning methods to select the two methods. Khozeimeh et al. [31] proposed an expert system to predict the responses of patients with warts to cryotherapy and immunotherapy but only achieved 83.33% accuracy. Guimaraes et al. [25] also developed a system based on fuzzy logic to predict the adaptability of a patient to immunotherapy. Given their high prediction accuracy, decision tree (DT) algorithms including J48, ID3 and classification and regression tree (CART) were built as classifiers to select the optimal treatment method. Ghiasi et al. [26] employed the CART algorithm to develop accurate predictive models for analysing the effectiveness of cryotherapy and immunotherapy. Akben [27] proposed ID3 classification model-based interpretable fuzzy informative images to predict the success of wart treatment methods. Khatri et al. [28] used the genetic programming-based J48 algorithm to enhance the predictive accuracy with the same dataset.

However, these methods are only 70–90% accurate and cannot give physical meaning. Ghiasi [29] proposed a model with sufficient accuracy, but their DT structure is complicated and difficult to understand. Through experiments, authors found that compared with other classifiers, such as k-nearest neighbour (K-NN), support vector machine (SVM), etc., the prediction accuracy of the decision tree classifier on immunotherapy dataset and cryotherapy dataset is relatively high, reaching 82–93% [9, 37]. And as a classical machine-learning algorithm, the DT is known to be computationally inexpensive and no assumptions are needed concerning the predictors' distribution. Therefore, this paper considers further improvements to the decision tree, and proposes an improved decision tree model using PSO for parameter optimization. This paper uses PSO algorithm to optimise the parameters of the decision tree instead of the traditional pruning optimization method, and then uses the optimal parameter combination to establish an improved decision tree classification model. The final result shows that the structure of the proposed model is simpler compared with other decision tree model, and the prediction accuracy of the proposed model reaches 100%. Besides, it has a better interpretability.

This paper has the following contributions:

1. Combining the linearly decreased inertia weight with time-varying acceleration coefficients, the optimization capability of PSO has been improved.
2. The parameters of DT were optimised using the PSO algorithm instead of the pruning algorithm, making

training process of the DT model simpler and structure more concise.

3. The classification accuracy of the proposed model is 100%, indicating that it can effectively select the best treatment for warts.

This paper includes five sections. Section “[Related work](#)” briefly introduces the essential earlier works and findings about the WART and some related works about the CART and PSO algorithm. In section “[Model development](#)” the process of model construction is explained. First, the dataset was introduced, and then the imbalanced dataset was processed using the Synthetic Minority Oversampling Technique (SMOTE) algorithm. Finally, the DT model was built based on PSO using the CART algorithm. The experiment was carried out in section “[Experiments and analysis](#)”, and the proposed method was compared with the traditional method to evaluate its performance. Finally, section “[Conclusion](#)” summarises this paper.

Related work

To improve the diagnosis of medical science and reduce the subjective interference of doctors' decision-making, several researchers have studied the application of predictive algorithms to the diagnosis and treatment of wart skin diseases. For example, Khozeimeh et al. [31] proposed an expert system based on fuzzy logic rules to study the therapeutic effects of immunotherapy and cryotherapy on warts. On the basis of the expert system proposed by Khozeimeh et al. [31], Guimaraes et al. [38] used the fuzzy neural network method to improve the prediction accuracy of the model. On other study, based on the dataset provided by Khozeimeh et al. [31], Akben [10] established an ID3 decision tree classifier to predict the choice of treatment options for warts. Similarly, Khatri et al. [9] also use the same dataset to establish a J48 decision tree classification model to help the selection the treatment of warts. Because of the excellent performance of DTs in these models, this article considers further improvements to the decision tree, which is of great significance in improving the accuracy of model predictions and increasing model interpretability.

As a kind of classical supervised machine learning and data mining approaches, DTs have been widely used in classification and data mining problems because of its fast classification speed, high accuracy and easy understanding. The optimization and improvement of decision tree algorithms has always been a research hotspot in the field. There are several forms of DT to establish a DT-based model, including CART, C4.5 and ID3. CART was proposed by Breiman [7]. The classification tree is used to analyse classification outcomes, and the regression tree is utilised to

analyse continuous outcome. CART divides the sample into subgroups repeatedly based on certain criteria. The CART algorithm is easy to implement and interpret. Therefore, it has many applications in medical diagnosis and prognosis.

In the diagnosis of skin diseases, tree-based approaches, such as CART, [8], J48 [9] and ID3 [10], are used in wart treatment selection. Non-tree-based approaches, including SVM [11, 12], K-NN [13, 14], convolutional neural network [15] and Bayesian Network [13], are utilised for skin disease diagnosis. However, tree-based methods are generally superior to their counterparts in terms of performance. Especially in terms of accuracy, the DT model is always higher than other models. It offers distinct advantages. On the one hand, the CART algorithm does not need to make any assumptions about features neither do they need to transform the variables. It can solve classification and continuous variables simultaneously. On the other hand, the CART model is easy to visualise and has strong interpretability. Compared with black-box models such as neural networks, DT models can be displayed in the form of graph and, therefore, have stronger interpretability.

However, the CART classifier does not always produce ideal results. It can be associated with some disadvantages. One of the main drawbacks is that its structure is susceptible to sample variables and sample size and may become complicated due to many branches. In this case, the performance of CART can be improved in two ways. One is the ensemble method, such as random forest. It combines the results of several classifiers to improve the performance over a single classifier. Another approach to improve the performance of CARTs is to optimise the parameters of the classifier. In this paper, the PSO algorithm is used to optimise the parameters of the decision tree, which can better control the generation scale of the tree while ensuring good classification accuracy, making the structure of the generated decision tree simpler.

The meta-heuristic technique is an intelligent optimisation algorithm for classifiers, such as ant colony optimisation, PSO and genetic algorithm (GA). In this paper, the PSO algorithm is employed to optimise the parameters of CART. The PSO algorithm is an evolutionary computation algorithm proposed by Kennedy and Eberhart [16, 17] in 1995. The basic concept of PSO stems from a metaphor of social interaction, and the PSO algorithm is constructed to simulate the social behaviour of birds flocking or fish schooling. Each particle represents a candidate solution of the research problem. Suppose that all potential solutions to the optimisation problem are particles in the search space. All particles have an adaptation value determined by the optimised function. Each particle has a velocity vector that determines the direction and distance of their flight. The particle follows the current optimal particle to search in the solution space, traces the optimal solution p_{best} currently found by itself and the optimal solution p_{gbest} currently found

by the population and then searches through generations until the optimal solution is obtained. Each particle updates its speed and position, respectively [18], as follows:

$$v_{\text{id}}^{k+1} = wv_{\text{id}}^k + c_1r_1(p_{\text{id}} - x_{\text{id}}^k) + c_2r_2(p_{\text{gd}} - x_{\text{id}}^k), \quad (1)$$

$$x_{\text{id}}^{k+1} = x_{\text{id}}^k + v_{\text{id}}^{k+1}, \quad (2)$$

where d indicates the d th dimension of the particle; i indicates the particle i ; k indicates the k th generation; c_1, c_2 are the learning factors, also called the acceleration constant; r_1, r_2 are random numbers between [0, 1] and w is the inertia weight function.

Although PSO is relatively simple to apply and has a high convergence rate, it is easy to fall into local optimum and premature convergence. In the past few years, many improved PSO algorithms have been proposed to improve the search capability of the PSO algorithm and reduce the probability of PSO falling into a local optimum [19–22, 40–42]. As an important parameter of PSO, the improvement of inertia weight can greatly improve the performance of PSO. A proper w can help PSO achieve a balance between global exploration and local exploitation. In general, a larger w is better for global exploration, and a smaller w is better for local exploitation [23]. In [22], linearly decreased inertia weight is introduced in PSO (PSO-LDIW), where w is defined as follows:

$$w = w_{\text{max}} - (w_{\text{max}} - w_{\text{min}}) \times \text{iter}/\text{maxiter}, \quad (3)$$

where w_{max} and w_{min} indicate the maximum and minimum values of the inertia weight, respectively; iter represents the number of the current iteration; and maxiter denotes the maximum iteration number. Moreover, Asanga [23] has used the time-varying inertia weight factor in PSO and introduced time-varying acceleration coefficients to control the local search and convergence to the global optimum. The time-varying acceleration coefficients [23] are calculated as follows:

$$c_1 = (c_{1f} - c_{1i}) \times \text{iter}/\text{maxiter} + c_{1i}, \quad (4)$$

$$c_2 = (c_{2f} - c_{2i}) \times \text{iter}/\text{maxiter} + c_{2i}, \quad (5)$$

where c_{1i} and c_{1f} denote the initial and the final values of the acceleration coefficient c_1 , $c_{1i} = 2.5$, $c_{1f} = 0.5$; c_{2i} and c_{2f} represent the initial and the final values of c_2 , $c_{2i} = 0.5$, $c_{2f} = 2.5$.

In the field of medical science, the PSO algorithm has been used in many aspects, such as brain tumour [24–26], heart disease [27] and skin and blood cancer [21, 28, 29]. For example, PSO has been combined with K-means clustering

for enhanced skin lesion segmentation and classification [30]. Tan [21] utilised the enhanced PSO algorithm during feature selection to conduct intelligent skin cancer diagnosis, and results showed that this method is superior to those presented in other studies. In this paper, we utilise the PSO algorithm to optimise the parameters of the CART tree and to select proper treatment methods for warts. Given that it makes no hypothesis to optimise the problem and can search large spaces of a candidate solution, the PSO algorithm is applied to optimise the DT parameters among large search spaces. The linearly decreasing strategy is adopted for inertia weight with (3) to achieve the balance of the global and local searching performance of PSO. The time-varying strategy is also adopted in this paper to improve the search ability of conventional PSO algorithms.

Model development

Dataset

The datasets used in this paper are from the literature [31]. The dataset collected 180 data from patients with common warts or/and plantar warts treated using immunotherapy or cryotherapy in the Dermatology Department of Ghaem Hospital in Mashhad from September 2013 to February 2015. The two types of warts and treatment approaches were used because they are the common.

Among the 180 data, 90 records were collected when immunotherapy was used to treat patients. These records contained eight features, namely, gender, age, time elapsed before treatment, number of warts, types of warts, surface area of the warts, induration diameter of initial test and response to treatment. Another 90 records were collected when cryotherapy was used to treat patients. These records contained seven features, namely, gender, age, time elapsed before treatment, number of warts, types of warts, surface

Table 1 Features in the immunotherapy dataset

S/N	Feature name	Value/type
1	Response to treatment	1—Yes 2—No
2	Gender	1—Man (41) 2—Woman (49)
3	Age (years)	15–56
4	Time elapsed before treatment (months)	0–12
5	Number of warts	1–19
6	Types of wart (count)	1—Common (47) 2—Plantar (22) 3—Both (21)
7	Surface area of the warts (mm ²)	6–900
8	Induration diameter of initial test (mm)	5–70

area of the warts and response to treatment. Data collection was performed as previously described [31]. Tables 1 and 2 list the attributes and related values or types in the immunotherapy and cryotherapy datasets, respectively. The class attribute in these datasets is the Response to treatment feature.

Classifier development

We introduce an extended DT method to construct the model. DT is a classifier with simple structure and high search efficiency, and the classification system is a branch of PSO algorithm research. Maximising the advantages of the PSO algorithm, a DT is constructed using the structure and node operations of the DT. Many DT algorithms exist, but the CART algorithm is utilised in study because of its capacity to deal with discrete and continuous variables.

Data preprocessing

A class imbalance exists in the dataset used in this article. The so-called class imbalance problem refers to the imbalanced distribution of the categories of the training sets used in the training classifier. As shown in Table 3, category 1 indicates that the patient responded well to the treatment, and 0 indicates that the patient did not adapt to the treatment. The statistical results in Table 3 show that the class imbalance problems in the two datasets have different levels. The number of data with category 1 in the immunotherapy dataset is 71, and that with category 0 is 19; meanwhile, the number of data with category 1 in the cryotherapy dataset

Table 2 Features in the cryotherapy dataset

S/N	Feature name	Value/type
1	Response to treatment	1—Yes 2—No
2	Gender	1—Man (47) 2—Woman (43)
3	Age (years)	15–67
4	Time elapsed before treatment (months)	0–12
5	Number of warts	1–12
6	Types of wart (count)	1—Common (54) 2—Plantar (9) 3—Both (27)
7	Surface area of the wart (mm ²)	4–750

Table 3 Category distribution in dataset

	Immuno-therapy	Cryotherapy
1	71	48
0	19	42

is 48, and that with category 0 is 42. The category imbalance in the immunotherapy dataset is relatively large. Thus, we chose the SMOTE algorithm to handle the imbalance problem.

SMOTE is a common means of dealing with imbalanced data. The SMOTE algorithm analyses and simulates a small number of samples and adds new samples of artificial simulation to the dataset to solve the serious imbalance in the categories in the original data. The generation of a new sample by SMOTE is shown in Fig. 1 [32].

According to Fig. 1, in a binary dataset, the classification of samples is not always balanced. According to the classification criteria, the samples can be divided into two categories. In Fig. 1, in other words, the sample can be divided into majority class and minority class according to the decision boundary. To generate a new sample for the imbalanced dataset, for each sample x in the minority class, the distance $d(x, y)$ from all samples in the minority sample set to x is calculated as follows using the Euclidean distance standard to obtain its K-NN [33]:

$$\begin{aligned} d(x, y) &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2} \\ &= \sqrt{\sum_{i=1}^n (x_i - y_i)^2}. \end{aligned} \quad (6)$$

Secondly, a sampling ratio is set according to the sample imbalance ratio to determine the sampling magnification N , which is calculated as follows:

$$N = N_{\text{maj}} - N_{\text{min}} \quad (7)$$

where N_{maj} is the majority sample number and N_{min} is the minority sample number.

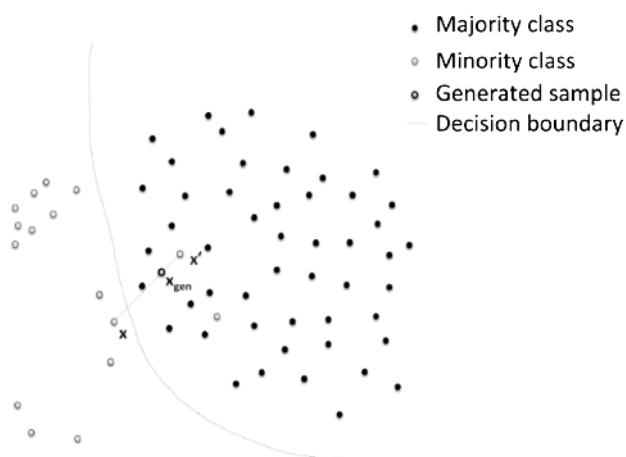


Fig. 1 Generation of a new sample by SMOTE

Finally, for each minority sample x , several samples are randomly selected from its K-NNs, assuming that the selected neighbour is x_n . For each randomly selected neighbour x_n , a new sample is constructed with the original sample in accordance with the following formula [34]:

$$x_{\text{new}} = x + \text{rand}(0, 1) \times |x - x_n|, \quad (8)$$

where δ is a random number between 0 and 1. Therefore, a generated sample is obtained as is showed in the Fig. 1.

After data preprocessing, a model is constructed to select a suitable treatment alternative.

Parameter optimisation algorithm

There are some traditional optimization methods for DTs, including postpruning, prepruning, and error-based pruning. In this paper, an improved particle swarm optimization method is proposed to optimise the parameters of the decision tree to achieve the purpose of model optimization. Based on the contribution of inertia weights and learning factors to particle swarm optimization, this paper combines the linearly decreased inertia weight and the time-varying acceleration coefficients to propose an improved particle swarm algorithm to optimise the parameters of the decision tree.

In the particle swarm algorithm, the inertia weight w represents the ability of the particle to maintain the state of motion at the previous moment, which can be used to balance the global and local search capabilities of the particle, so it is very important in the particle swarm algorithm. Many scholars believe that the value of inertia weight should be large when the particle swarm is in the exploration state, and small in the exploitation state. This paper employees the linearly decreased inertia weight [22] to change inertia weight with the iterations, it can be calculated by formula (3). Moreover, in different states of the particle swarm, the requirements for the local and global exploration capabilities of the particle swarm are different. The learning factor c_1 represents "self-cognition", which helps to explore the local area and maintain the diversity of the group, and the learning factor c_2 represents "social cognition", which can help the group quickly converge, so the learning mechanism requirements for the learning factors c_1 and c_2 are also different. Therefore, the time-varying acceleration coefficients [23] are used in this paper. It can be computed by formula (4) and (5).

In other words, the evaluation function is the only interface between the PSO algorithm and the specific application problem. It is a quantitative reflection of the individual's advantages and disadvantages in the population. Its structure directly affects the efficiency of problem solving.

The accuracy of DT is an important criterion for evaluating DTs. Moreover, the shape of the DT and the number of nodes may greatly vary from one individual to another. The number of attributes and the number of attribute values are unrestricted quantities. Hence, using a fixed-length bit string to represent the DT is inappropriate. In this study, PSO is used to iterate the parameter values continuously, giving different values to the parameters. Optimising the parameter values, we can find a DT with a simple structure and high accuracy. Let us determine the adaptation function as:

$$\text{fitness}_{\text{function}} = w_{\text{acc}} \times f(\text{feature}_{\text{num}}, \text{depth}, \text{simple}_{\text{split}}, \text{simple}_{\text{leaf}}, \text{leaf}_{\text{node}}) - c \times (w_{\text{depth}} \times \text{depth} + w_{\text{num}} \times \text{num}_{\text{node}}). \quad (9)$$

The structure and accuracy of the DT can be affected by several parameters, including max_features, max_depth, min_simple_split, min_simple_leaf and max_leaf_node, represented in the formula as feature_num, depth, simple_split, simple_leaf and leaf_node, respectively. ‘max_features’ is the number of features considered when searching for the best partition, ‘max_depth’ is the maximum depth of the specified tree, ‘min_simples_split’ represents the minimum number of samples needed to split an internal node, ‘min_simples_leaf’ specifies the minimum number of samples required for each leaf node and ‘max_leaf_nodes’ specifies the leaves’ maximum number of nodes.

In the adaptation function, $f(\text{feature}_{\text{num}}, \text{depth}, \text{simple}_{\text{split}}, \text{simple}_{\text{leaf}}, \text{leaf}_{\text{node}})$ indicates the accuracy of the DT, which is determined by five parameters of DT. Moreover, w_{acc} , w_{depth} and w_{num} are weights corresponding to the accuracy, depth and number of nodes. Since the accuracy and depth and the number of nodes are different magnitude units, a balance factor c is required. Let $w_{\text{acc}} = 0.95$, $w_{\text{depth}} = 0.025$, $w_{\text{num}} = 0.025$, and the balance coefficient c be 0.1.

Classifier development

The tree-based classifier CART is used because of its explicit meaning and simple properties. Our model utilises the CART algorithm for the selection of wart treatment. The CART model is fast to build, and it is applicable to qualitative and quantitative data. Recursive binary splitting is also applied.

Given training set D and continuous attribute a . Assuming that at the node i , attribute a has n different values in D_i . First, these values are sorted from small to large, denoted as $a = \{a^1, a^2, \dots, a^n\}$, where $a^1 \leq a^2 \leq \dots \leq a^n$.

Then, D_i is divided into subsets D_{it}^- and D_{it}^+ based on the dividing point t . D_{it}^- is the sample in dataset D_i of which the value of attribute a is less than the value of t . D_{it}^+ is the sample in dataset D_i of which the value of attribute a is greater

than that of t . Then, D_i is denoted as $D_i = \{D_{it}^-, D_{it}^+\}$, where a^i in D_{it}^- is less than t , and a^{i+1} in D_{it}^+ is more than t .

For the adjacent attribute values a^i and a^{i+1} , the result of the division of t by taking any value in the interval $[a^i, a^{i+1}]$ is the same. Therefore, for the continuous attribute a , the median point $\frac{a^i + a^{i+1}}{2}$ of the interval $[a^i, a^{i+1}]$ is taken as the candidate segmentation point, and then the Gini value is calculated:

$$\text{candidate_segmentation_point} = (a^i + a^{i+1})/2. \quad (10)$$

When we used the CART algorithm to build our model, the process of dividing the nodes was repeated to classify the samples as accurately as possible. However, this step produced too many branches and too complicated DT structure. Considering that the PSO algorithm makes no hypothesis to optimise the problem and can search large spaces of a candidate solution, we introduced it to iterate the DT parameters. It needs short search time and can consider as many combination schemes as possible. Therefore, the proposed model combining PSO with CART can realise a DT with a simple structure and high classification accuracy.

The DT is constructed using the PSO algorithm to generate the next-generation group from the previous generation DT group through the particle motion operation and gradually evolves until the termination condition of the PSO algorithm is satisfied. The flowchart of our model is shown in Fig. 2. The steps of our model are as follows:

Step 1: Employ the SMOTE algorithm to generate new samples from the initial class-imbalanced dataset D , and obtain the class-balanced dataset D' ;

Step 2: Divide the class-balanced dataset D' into training and testing sets;

Step 3: Initialise a group of particles (group size is m), and randomly generate parameter values within the range of values;

Step 4: Generate a DT group based on randomly generated parameter values and training sets, i.e. particle group m ;

Step 5: Use the validation set to calculate the accuracy of each DT, and evaluate the fitness of each particle;

Step 6: For each particle, compare its current fitness value with the best position p_{best} it has experienced, and if it is better, replace it with the current best position p_{best} ;

Step 7: For each particle, compare its fitness value with the best position g_{best} that has been experienced globally, and if it is better, reset g_{best} ;

Step 8: Update the particles’ velocity and position using formulas (3), (4) and (5);

Step 9: Repeat step 3 until the termination condition is reached.

The optimal DT refers to the DT with the largest adaptation function value. In the evolution of particle motion, the last generation of groups does not necessarily contain the

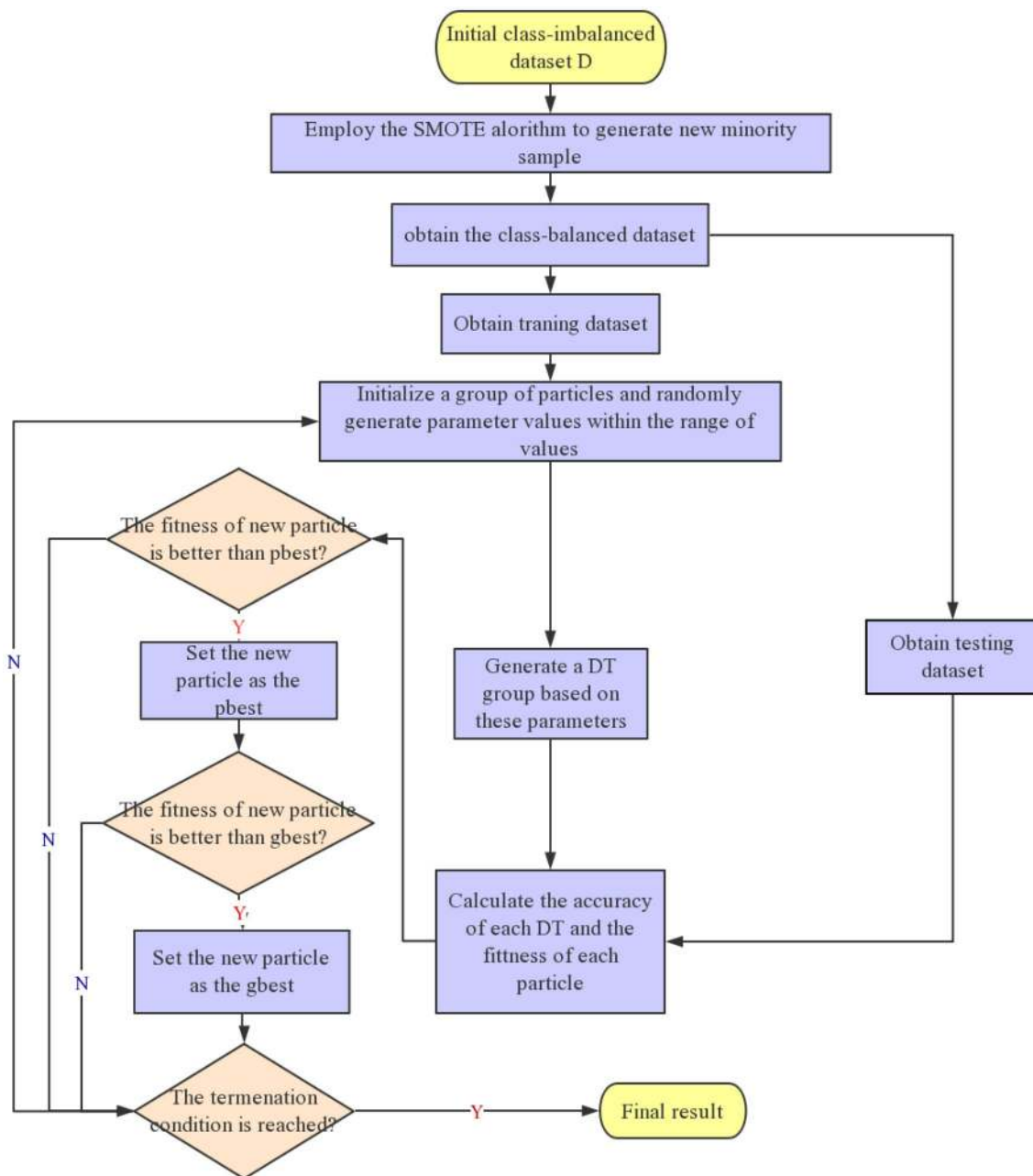


Fig. 2 Flowchart of our model

optimal DT. Therefore, the best DT must be saved in the whole process. Whenever the next generation is produced, compare it with the saved one and eliminate the relatively poor one. Thus, until the end of evolution, the saved DT is basically the best, that is, the result to be saved.

Performance criteria

The two most commonly used performance metrics in classification tasks are error rate and accuracy. They are

suitable for two-category and multi-category tasks. The accuracy rate is the ratio of the number of correctly classified samples to the total number of samples, and the error rate is the proportion of the number of samples with the wrong classification to the total number of samples. In addition, precision and sensitivity can be used to measure the performance metrics of the model. They are all available from the confusion matrix.

The confusion matrix is a basic tool for evaluating classifier credibility. Taking a two-category problem as the

		Actual Value	
		True	False
Predicted Value	Positive	TP True Positive	FP False Positive
	Negative	FN False Negative	TN True Negative

Fig. 3 Confusion matrix

Table 4 The details of the benchmark functions

Number	Name	Dimension	Search space	Threshold	Minimum
$f_1(x)$	Sphere	50	$[-100, 100]$	0.01	0
$f_2(x)$	Rosenbrock	50	$[-30, 30]$	100	0
$f_3(x)$	Ackley	50	$[-32, 32]$	0.01	0
$f_4(x)$	Rastrigin	50	$[-5.12, 5.12]$	50	0
$f_5(x)$	Schwefel	50	$[-10, 10]$	0.01	0

research object, the confusion matrix of Fig. 3 shows all the cases that a classifier may encounter, where the row of Fig. 3 corresponds to the category to which the instance actually belongs, and the column of Fig. 3 represents the result of the prediction classification.

With these four basic indicators, multiple classifier evaluation indicators can be derived. Accuracy of the model generated by the classification algorithm according to Table 4 is given by Eq. (10) [8]:

$$\text{Accuracy : ACC} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}}. \quad (11)$$

The proximity of the observed value to the true value is called precision, and its equation is shown in the following equation [8]:

$$\text{Precision : PRE} = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (12)$$

Estimating the validity of the classifier's positive classification label is called sensitivity. The equation of sensitivity is shown in the following equation [8]:

$$\text{Sensitivity : SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (13)$$

The formula of F-measure is shown in the following equation [8]:

$$\text{F-measure : F-meas} = \frac{2 \times \text{PRE} \times \text{SEN}}{\text{PRE} + \text{SEN}}. \quad (14)$$

The proportion of negative classes in the sample that are predicted to be negative classes is called specificity. The equation of specificity is shown in the following equation:

$$\text{Specificity : SPE} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \quad (15)$$

Also, ROC was used to evaluate the performance.

In this section, a brief introduction of the dataset has been given, and the development of the classifier has been introduced. In the classifier development part, the dataset was preprocessed using the SMOTE method first, and then a parameter optimization method based on PSO was proposed to optimise the parameters of DT, and the optimal parameter combination obtained from the parameter optimization results was used to establish the DT model. Finally, the performance criteria of the model have been introduced.

Experiments and analysis

The experiment in this article consists of two parts. On the one hand, the improved particle swarm algorithm used in this article is verified experimentally using several benchmark functions, and compared with other improved particle swarm algorithms. On the other hand, the wart dataset and the model developed in the third section were used, and experiments were carried out to verify the performance of the CART model improved by the particle swarm algorithm proposed in this paper. The main experiment is as follows.

Estimation on PSO

Benchmark functions

Mathematical optimization problems need to use some functions to test the performance of the optimization algorithm. Similarly, to verify the improved performance of the particle swarm algorithm in terms of parameters, a test function is required. This paper selects five classical standard test functions to evaluate the performance of the improved PSO algorithm and compare them with the standard PSO algorithm, PSO-CK [39], PSO-TVAC [23] and PSO-LDIW [22].

(1) Sphere function

$$f_1(x) = \sum_{i=1}^D x_i^2. \quad (16)$$

The Sphere function is a unimodal function, which is often used to test the convergence rate of optimization problems.

(2) Rosenbrock function

$$f_2(x) = \sum_{i=1}^{D-1} \left(100(x_{i+1} - x_i)^2 + (x_i - 1)^2 \right). \quad (17)$$

The Rosenbrock function is a unimodal non-convex function, and the correlation between the function variables is strong and difficult to solve.

(3) Ackley function

$$f_3(x) = -20e^{-0.2\sqrt{\frac{1}{D}\sum_{i=1}^D x_i^2}} + 20 + e - e^{\frac{1}{D}\sum_{i=1}^D \cos 2\pi x_i}. \quad (18)$$

The Ackley function is widely used to test optimization algorithms, but because of the many minimum points of the function, it is very difficult to find the optimal solution.

(4) Rastrigin function

$$f_4(x) = \sum_{i=1}^D (x_i^2 - 10 \cos 2\pi x_i + 10). \quad (19)$$

The Rastrigin function is a very typical nonlinear multimodal function. Because of its large search interval and a large number of minimum points, it is difficult for this function to find the optimal solution.

(5) Schwefel function

$$f_5(x) = \sum_{i=1}^D |x_i| + \prod_{i=1}^D |x_i|. \quad (20)$$

Schwefel function is also a very typical unimodal function. There are accumulative terms in this function, which makes calculation difficult.

More details such as the dimension, the search space of the dimension, the threshold and etc. can be showed in Table 4.

Experiments and results

As discussed above, five benchmark functions are employed to evaluate the introduced PSO algorithm. The superiority of the introduced PSO algorithm is demonstrated over four popular PSO algorithms, including the standard PSO algorithm, PSO-CK [39], PSO-TVAC [23] and PSO-LDIW [22]. The experimental parameters are set as follows: the dimension of the particle swarm is $D=50$, and the population of the particle swarm is $S=20$. It should be noted that to avoid the influence of random factors, each experiment has been repeated 20 times

independently, and the average value of the 20 experiments is taken as the experimental result. The detailed information of the experiment result are shown in Table 5, including the average, maximum and minimum fitness values of different PSO algorithms under each benchmark function.

It can be seen from the results in Table 5 that on the five benchmark functions, the optimal solution obtained by the PSO-LT algorithm used in this paper is the closest to the minimum value of 0. Among them, the performance of the PSO-LT on $f_1(x)$ is the best, which is the closest to 0. The performance of the PSO-LT on $f_2(x)$ is slightly worse, but it is still better than other PSO algorithms. It is worth noting that in $f_4(x)$, the minimum fitness value obtained by PSO-LT used in this article is 0. In summary, compared with other particle swarm algorithms, the PSO-LT algorithm combining PSO-LDIW and PSO-TVAC in this paper performs best, and the average fitness value obtained is closest to the minimum value, which proves the performance and feasibility of the PSO-LT algorithm.

Experiment on proposed classifier

Results

The initial datasets of immunotherapy and cryotherapy show different levels of class imbalance. To avoid bringing large bias of the classifier, we applied the SMOTE algorithm to solve this class-imbalanced problem.

In the immunotherapy dataset, the number of minority samples is 19, and the number of majority samples is 71.

Table 5 The mean fitness of the four PSO algorithms on five benchmark functions

		PSO	PSO-LDIW	PSO-TVAC	PSO-LT
$f_1(x)$	Mean	0.0005	3.52×10^{-23}	34.0877	3.20×10^{-41}
	Max	0.0006	2.51×10^{-22}	47.5005	3.20×10^{-40}
	Min	0.0002	9.44×10^{-25}	12.7414	1.07×10^{-53}
$f_2(x)$	Mean	7.3461	1.3588	3.9815	0.4528
	Max	7.3661	1.6511	5.4042	0.6941
	Min	7.3270	0.8918	2.7387	0.0956
$f_3(x)$	Mean	0.0064	4.62×10^{-14}	0.0893	5.68×10^{-15}
	Max	0.0074	1.10×10^{-13}	0.1040	7.11×10^{-15}
	Min	0.0054	7.11×10^{-15}	0.0649	3.55×10^{-15}
$f_4(x)$	Mean	0.0316	4.43×10^{-11}	3.6549	3.18×10^{-13}
	Max	0.0423	1.22×10^{-10}	4.6394	8.81×10^{-13}
	Min	0.0249	4.26×10^{-12}	2.8332	0.0000
$f_5(x)$	Mean	0.0123	9.59×10^{-15}	0.1473	3.94×10^{-18}
	Max	0.0148	5.52×10^{-14}	0.1652	1.97×10^{-17}
	Min	0.0102	1.00×10^{-16}	0.1233	5.08×10^{-23}

Bold values represent the results of the proposed model for better comparison

Therefore, the sampling ratio N is 52, which is the number of new samples generated by the SMOTE algorithm. The final number of data in the immunotherapy dataset is 142. Similarly, the final number of data in the cryotherapy dataset is 96.

Normally, 80–90% of the data points are used for model training [19, 35, 36]. In our paper, we set the training set to 80% of the total dataset and the verification set to 20% of the total dataset. Using 80% of the wart dataset, our model can provide satisfactory results. The process in section “Classifier development” is then used to train the model. During optimisation, the PSO structure parameters are set as follows: the particle swarm initial population is 200, the search space dimension is 5 and the maximum iteration number $k=200$. Considering the data characteristics of the immunotherapy and cryotherapy datasets, the parameter range of the DT is set as shown in Table 6.

Comparisons

This section mainly includes two parts. One is the comparison with the traditional optimisation method for the CART algorithm, and the other is the comparison with the existed method for wart treatment selection, to prove the validity of the proposed model. Before the comparison parts, the evaluation index of the model is defined.

Comparison with the traditional optimisation method The traditional optimisation method of the DT is to use pruning methods, including prepruning, backpruning and reduced-error pruning. In this paper, three pruning methods were

Table 6 Range of parameters

Parameter	max_fea- tures	max_ depth	min_sim- ples_split	min_sim- ples_leaf	max_leaf_ nodes
Range	[1, 7]	[4, 10]	[3, 20]	[3, 20]	[10, 300]

Table 7 Comparison of the traditional optimisation method with the CART algorithm

Dataset	Method	Evaluating system				
		ACC	PRE	SEN	F-meas	SPE
Cryotherapy dataset	Original CART	0.867	1.00	0.786	0.880	0.900
	Prepruning	0.967	0.933	1.000	0.965	0.967
	Backpruning	0.967	1.000	0.929	0.963	0.967
	Reduced-error pruning	0.957	1.000	0.933	0.965	0.944
	Proposed method	1.000	1.000	1.000	1.000	1.000
Immunotherapy dataset	Original CART	0.900	0.833	0.714	0.769	0.875
	Prepruning	0.867	0.714	0.714	0.965	0.814
	Backpruning	0.900	0.833	0.714	0.769	0.875
	Reduced-error pruning	0.869	0.667	0.500	0.572	0.783
	Proposed method	1.000	1.000	1.000	1.000	1.000

Bold values represent the results of the proposed model for better comparison

used to evaluate the effect of the proposed method, as shown in Table 7, and the results were calculated. The performance criteria include the accuracy, precision, sensitivity, F-measure and specificity.

In Table 7, the proposed method has the best performance. The accuracy, precision and sensitivity of the proposed have reached 100%, and its AUC is the largest in two of the datasets. Combining the performance between the two datasets, backpruning optimisation is in the second place. Thus, we can conclude that the proposed method has the best ability in optimising the performance of the CART algorithm.

Figures 4 and 5 show the performance of the five optimisation methods in the immunotherapy and cryotherapy datasets. The AUC has been calculated in two pictures. Figure 4 shows that the lines of backpruning and prepruning in the cryotherapy dataset have coincided, and the two of them share the same AUC of 0.9667, which is the second largest value among the five results. In addition, the line of the proposed method has contained the four other lines in the axis. Figure 5 shows that the line of the original CART and backpruning has also coincided. Similarly, the line of the proposed method has contained the four other lines in the axis in the immunotherapy dataset, which means that the effect of the proposed method is better than those of the four other methods.

The above results show that the PSO-optimised CART method is superior to the other optimisation methods in accuracy, precision, sensitivity and other evaluation metrics.

Comparison with existing models In the literature [37], the Bayesian, C4.5 DT, logistic regression and K-NN classification algorithms were used to construct the model for the immunotherapy and cryotherapy datasets, and the accuracy of the classification results was calculated. The model in this paper is compared with the model in [37]. The results are shown in Table 8.

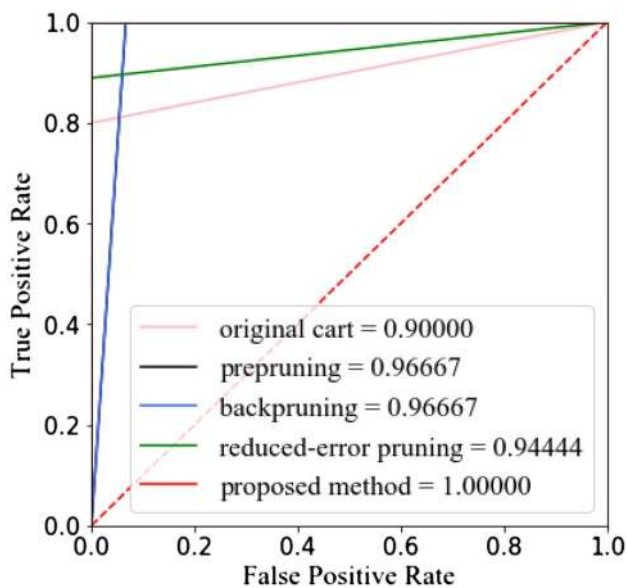


Fig. 4 ROC of the cryotherapy dataset

Compared with the four models in [37], the proposed PSO-based CART DT model is optimal in terms of accuracy and sensitivity, both of which are 1. The accuracy of the C4.5 DT model is suboptimal, and that of the Bayesian classifier is the worst. The accuracy, precision, sensitivity, F-measure and specificity of the final classification result of the classifier are optimal. Therefore, the proposed model performs better than the four models in [37]. In addition to the CART model proposed in this article, we can see from Table 8 that the second best performing model is C4.5. C4.5 is also a kind of decision tree algorithm. It can be seen from this that the accuracy of the model based on the decision

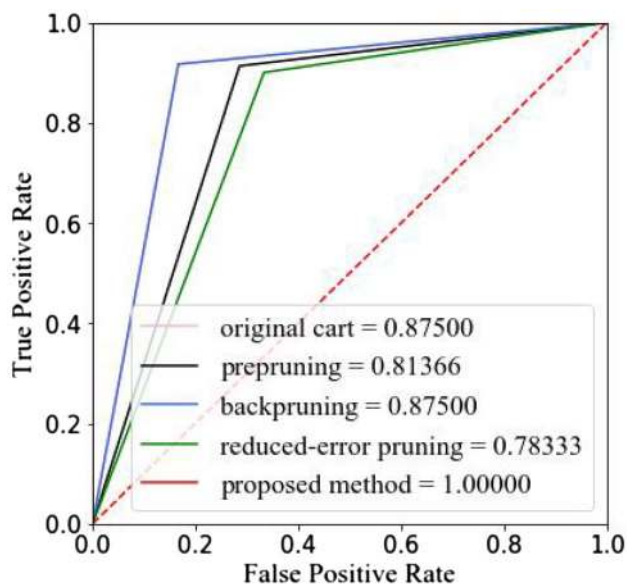


Fig. 5 ROC of the immunotherapy dataset

tree is higher than other traditional models in the application of classification and prediction problems, which once again proves the superior performance of the improved algorithm based on the CART decision tree in this paper.

Discussion

Wart is a skin disease caused by HPV, where common and plantar warts are their general types. Immunotherapy and cryotherapy are the two most utilised treatments for warts. In consideration that different warts under different circumstances have different responses to different treatment methods, the treatment of warts needs to be selected according to each individual's condition. Therefore, this paper establishes a CART classifier for the selection of immunotherapy or cryotherapy for plantar and common warts to help patients choose the appropriate treatment.

CART tree is a classic classifier method useful in many types of classifier and regression situations. However, the tree is easy to overfit. To solve this problem, the PSO algorithm is used to discover the advantages of the optimal region in the complex search space through the interaction between particles. The algorithm model combining PSO and DT is constructed. The optimal DT is approximated by the idea of gradual approximation. The linearly decreased inertia weight and the time-varying acceleration coefficients are adopted to improve its optimisation performance. Compared with the traditional optimisation methods, such as prepruning, backpruning and error-reduced pruning, the proposed method is the most effective.

This study analysed the original immunotherapy and cryotherapy datasets, used the SMOTE algorithm to classify the original dataset and expanded the size of the original dataset. Then, the PSO algorithm is used to optimise the parameters of DT, and the optimal parameter combination obtained from the results is used to build a DT classifier. To verify the performance of the PSO algorithm used in this paper, five benchmark functions were used to test the convergence of the PSO algorithm in the first part of the experiment and compared with other improved PSO algorithms. The experimental results prove that the PSO algorithm used in this paper has the performance advantages of effectively finding the optimal solution and avoiding premature maturity. In the second part of the experiment, this article uses the wart data set to establish a DT classifier to select the optimal treatment plan for warts, and compares the latest selection method of wart treatment plans with traditional classifiers. The final experimental results prove that after the optimisation of parameters the structure of the selected DT model is simpler, with an accuracy rate of 100%. The resulting model has 100% accuracy and sensitivity on both datasets, and its performance is superior over those of other conventional algorithms, such as K-NN, NB and LR. Moreover, the DT

Table 8 Comparison of the proposed model with the models constructed in [37]

Dataset	Algorithms	Evaluating system				
		ACC	PRE	SEN	F-mea	SPE
Cryotherapy dataset	PSO-based CART	1.000	1.000	1.000	1.000	1.000
	Navie Bayes [37]	0.844	0.845	0.896	0.829	0.883
	k-NN [37]	0.933	0.930	0.929	0.915	0.927
	C4.5 [37]	0.954	0.937	0.976	0.933	0.964
	LR [37]	0.920	0.893	0.889	0.894	0.912
Immunotherapy dataset	PSO-based CART	1.000	1.000	1.000	1.000	1.000
	Navie Bayes [37]	0.780	0.848	0.820	0.805	0.825
	k-NN [37]	0.810	0.830	0.828	0.812	0.822
	C4.5 [37]	0.840	0.901	0.832	0.851	0.864
	LR [37]	0.800	0.775	0.800	0.837	0.817

model established in this paper is visual and easy to understand. It can assist doctors and patients with sputum skin diseases to choose a suitable treatment, save doctors' diagnosis time and reduce patients' medical costs. In future studies, if further information is obtained from different groups of patients, other efficient DT-based models that can be implied to different cases may be established.

Conclusion

This study analysed the original immunotherapy and cryotherapy datasets, and established a mixed classification model through the combination of PSO and CART DT. The PSO algorithm used in this article uses a linearly decreased inertia weight and time-varying acceleration coefficients to avoid premature and getting trapped into local optima. The superior performance of the PSO algorithm is verified through experiments. At the same time, the PSO algorithm is used to optimise the parameters of the decision tree instead of the traditional pruning algorithm. Because of the simple operation, few parameters and excellent search ability to find

the optimal solution, the PSO algorithm used in this paper can quickly find the optimal parameter combination of the CART DT. And then a more concise and more accurate classification model is established. The classification accuracy of this model has achieved 100% on both datasets, and its performance is superior over those of other conventional algorithms. Compared with the conventional optimisation methods, such as prepruning, backpruning and reduced-error pruning, this new PSO-optimised model tree is superior over other optimisation methods. Compared with existing methods for the selection of wart treatment, the proposed method also performs better than the other methods. Moreover, the proposed DT model established is visual and easy to understand. It can assist doctors and patients with sputum skin diseases to choose a suitable treatment, save doctors' diagnosis time and reduce patients' medical costs.

Appendix

See Table 9.

Table 9 The pseudo-code of the improved PSO

```

procedure Improved PSO

  for each particle  $i$ 

    Initialize velocity  $V_i$  and position  $X_i$  for particle  $i$ 

    Evaluate particle  $i$  and set  $pBest_i = X_i$ 

  end for

   $gBest = \min\{pBest_i\}$ 

  while not stop

    for  $i = 1$  to  $N$ 

      Compute the inertia weight  $w$  by formula (3)

      Compute the learning factors  $c_1$  and  $c_2$  by formula (4) and (5)

      Update the velocity of particle  $i$  according to formula (1)

      Update the position of particle  $i$  according to formula (2)

      Calculate the fitness of particle  $i$  according to formula (9) as  $fit(X_i)$ 

      if  $fit(X_i) < fit(pBest_i)$ 

         $pBest_i = X_i$ ;

      if  $fit(pBest_i) < fit(gBest)$ 

         $gBest = pBest_i$ ;

    end for

  end while

  print  $gBest$ 

end procedure

```

Acknowledgements The authors would like to thank Editor-in-Chief, editor, and anonymous reviewers for their valuable comments and helpful suggestions.

Author contributions Conceptualization: JH and XO; methodology: XO and PL; software: XO; validation: XO; formal analysis: JH and XO; investigation: BL; resources: BL; data curation: BL; writing-original draft preparation: XO; writing-review and editing: XO, PL; visualization: JH; supervision: JH; project administration: JH and PL; funding acquisition: JH.

Funding This work is supported by National Natural Science Foundation of China under Grant Number 71871229.

Declarations

Conflict of interest The authors declare no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Hausen H (1996) Papillomavirus infections—a major cause of human cancer. *Biochim Biophys Acta (BBA) Rev Cancer* 1288(2):F55–F78
- Vlahovic TG, Khan MT (2016) The human papillomavirus and its role in plantar warts: a comprehensive review of diagnosis and management. *Clin Podiatr Med Sur* 33:337–353
- Bruggink SC, Gussekloo J, Egberts PF et al (2015) Monochloroacetic acid application is an effective alternative to cryotherapy for common and plantar warts in primary care: a randomized controlled trial. *J Invest Dermatol* 135:1261–1267
- Ruiz D, Berenguer V, Soriano A et al (2011) A decision support system for the diagnosis of melanoma: a comparative approach. *Expert Syst Appl* 38:15217–15223
- Mirzaalian H, Lee TK, Hamarneh G (2016) Skin lesion tracking using structured graphical models. *Med Image Anal* 27:84–92
- Sumithra R, Suhil M, Guru DS (2015) Segmentation and classification of skin lesions for disease diagnosis. *Proc Comput Sci* 45:76–85
- Breiman L, Friedman J, Olshen R et al (1984) Classification and regression trees (cart). Wadsworth, Pacific Grove
- Ghiassi MM, Zendeheboudi S (2019) Decision tree-based methodology to select a proper approach for wart treatment. *Comput Biol Med* 108:400–409
- Khatri S, Arora D, Kumar A (2018) Enhancing decision tree classification accuracy through genetically programmed attributes for wart treatment method identification. *Proc Comput Sci* 132:1685–1694
- Akben SB (2018) Predicting the success of wart treatment methods using decision tree based fuzzy informative images. *Biocybern Biomed Eng* 38:819–827
- Chen B, Lu Y, Pan W et al (2019) Support vector machine classification of nonmelanoma skin lesions based on fluorescence lifetime imaging microscopy. *Anal Chem* 91:10640–10647
- Seeja RD, Suresh A (2019) Deep learning based skin lesion segmentation and classification of melanoma using support vector machine (svm). *Asian Pac J Cancer Prev* 20:1555–1561
- Mirarchi D, Vizza P, Tradigo G et al (2018) On the use of mining techniques to analyse human papilloma virus dataset. In: 2018 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE, pp 2041–2048
- Maciel VH, Correr WR, Kurachi C et al (2017) Fluorescence spectroscopy as a tool to in vivo discrimination of distinctive skin disorders. *Photodiagn Photodyn* 19:45–50
- Hekler A, Utikal JS, Enk AH et al (2019) Superior skin cancer classification by the combination of human and artificial intelligence. *Eur J Cancer* 120:114–121
- Kennedy J, Eberhart R (1995) Particle swarm optimization. In: Proceedings of ICNN'95-international conference on neural networks, vol 4. IEEE, pp 1942–1948
- Eberhart R, Kennedy J (1995) A new optimizer using particle swarm theory. In: MHS'95. Proceedings of the sixth international symposium on micro machine and human science. IEEE, pp 39–43
- Trelea IC (2003) The particle swarm optimization algorithm: convergence analysis and parameter selection. *Inform Process Lett* 85:317–325
- Chamkalani A, Zendeheboudi S, Bahadori A et al (2014) Integration of lssvm technique with pso to determine asphaltene deposition. *J Petrol Sci Eng* 124:243–253
- Ma T, Wang C, Wang J et al (2019) Particle-swarm optimization of ensemble neural networks with negative correlation learning for forecasting short-term wind speed of wind farms in western china. *Inf Sci* 505:157–182
- Tan TY, Zhang L, Lim CP (2019) Intelligent skin cancer diagnosis using improved particle swarm optimization and deep learning models. *Appl Soft Comput* 84:105725
- Shi Y, Eberhart RC (1998) Parameter selection in particle swarm optimization. In: International conference on evolutionary programming. Springer, Berlin, pp 591–600
- Ratnaweera A, Halgamuge SK, Watson HC (2004) Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. *IEEE Trans Evolut Comput* 8:240–255
- Mekhmoukh A, Mokrani K (2015) Improved fuzzy c-means based particle swarm optimization (pso) initialization and outlier rejection with level set methods for mr brain image segmentation. *Comput Methods Progr Biol* 122:266–281
- Sharif M, Amin J, Raza M et al (2020) An integrated design of particle swarm optimization (pso) with fusion of features for detection of brain tumor. *Pattern Recogn Lett* 129:150–157
- Sharif MI, Li JP, Khan MA et al (2020) Active deep neural network features selection for segmentation and recognition of brain tumors using mri images. *Pattern Recogn Lett* 129:181–189
- Pandit D, Zhang L, Liu C et al (2017) A lightweight qrs detector for single lead ecg signals using a max-min difference algorithm. *Comput Methods Progr Biol* 144:61–75
- Srisukkhham W, Zhang L, Neoh SC et al (2017) Intelligent leukaemia diagnosis with bare-bones pso based feature optimization. *Appl Soft Comput* 56:405–419
- Tan TY, Zhang L, Neoh SC et al (2018) Intelligent skin cancer detection using enhanced particle swarm optimization. *Knowl Based Syst* 158:118–135
- Tan TY, Zhang L, Lim CP (2020) Adaptive melanoma diagnosis using evolving clustering, ensemble and deep neural networks. *Knowl Based Syst* 187:104807
- Khozeimeh F, Alizadehsani R, Roshanzamir M et al (2017) An expert system for selecting wart treatment method. *Comput Biol Med* 81:167–175
- Douzas G, Bacao F (2019) Geometric smote a geometrically enhanced drop-in replacement for smote. *Inf Sci* 501:118–135
- Merigó JM, Casanovas M (2011) Induced aggregation operators in the euclidean distance and its application in financial decision making. *Expert Syst Appl* 38:7603–7608
- Chawia N (2002) Smote: Synthetic minority over-sampling technique. *J Artif Intell Res* 16:321–357
- Kamari A, Mohammadi AH, Bahadori A et al (2014) Prediction of air specific heat ratios at elevated pressures using a novel modeling approach. *Chem Eng Technol* 37:2047–2055
- Arabloo M, Bahadori A, Ghiassi MM et al (2015) A novel modeling approach to optimize oxygen–steam ratios in coal gasification process. *Fuel* 153:1–5
- Basarslan MS, Kayaalp F (2018) A hybrid classification example in the diagnosis of skin disease with cryotherapy and immunotherapy treatment. In: 2018 2nd international symposium on multidisciplinary studies and innovative technologies (ISMSIT). IEEE, pp 1–5
- Guimarães AJ, Silva Araujo VJ, de Campos Souza PV et al (2018) Using fuzzy neural networks to the prediction of improvement in expert systems for treatment of immunotherapy. Springer International Publishing, Cham, pp 229–240
- Clerc M, Kennedy J (2002) The particle swarm: explosion, stability, and convergence in a multi-dimensional complex space. *IEEE Trans Evol Comput* 6(1):58–73
- Geng N, Chen Z, Quang A et al (2021) Particle swarm optimization algorithm for the optimization of rescue task allocation with uncertain time constraints. *Complex Intell Syst*. <https://doi.org/10.1007/s40747-020-00252-2>

41. Liu F, Zhang J, Liu T (2020) A PSO-algorithm-based consensus model with the application to large-scale group decision-making. *Complex Intell Syst* 6:287–298
42. Qin S, Sun C, Zhang G et al (2020) A modified particle swarm optimization based on decomposition with different ideal points for many-objective optimization problems. *Complex Intell Syst* 6:263–274

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.