

## Research Article

Ravinder Reddy Baireddy\* and R. Nagaraja

# A hybrid particle swarm optimization with multi-objective clustering for dermatologic diseases diagnosis

<https://doi.org/10.1515/jisys-2022-0028>

received September 02, 2021; accepted October 23, 2021

**Abstract:** Effective and personalized treatment relies heavily on skin disease categorization. In the stratification of skin disorders, it is crucial to identify the subtypes of illnesses to provide an efficient therapy. To attain this aim, researchers have focused their attention on cluster algorithms for the stratification of skin disorders in recent decades. But, cluster algorithms have real-world drawbacks, including experimental noises, a large number of dimensions, and a poor ability to comprehend. Cluster algorithms, in particular, determine the quality of clusters using a single internal evaluation operation in the majority of cases. A single internal assessment procedure is difficult to design and robust for all datasets, which is a problem. The multi-objective particle swarm obtained high sensitivity in the existing work, but it is not able to anticipate all kinds of classes. An optimized cluster distance parameter for  $K$ -means clustering is determined using a hybrid particle swarm and moth flame optimization. Multi-objective is guided by two cluster value indices, including the  $K$ -means clustering misclassification rate and neural network classification rate. Hybrid PSO will solve the multi-objective problem to identify the optimal cluster for clustering. On the dermatological dataset from the UCI repository, MATLAB R2020a will be used to evaluate the proposed method. This will be followed by an evaluation of the proposed method's performance using the cluster evaluation indices.

**Keywords:** skin diseases, clustering, multi-objective,  $K$ -means, moth flame optimization, particle swarm optimization

## 1 Introduction

Most dermatological infections fall under the erythematous-squamous disease (ESD) category. Psoriasis, seborrheic dermatitis, lichen planus, pityriasis rosea, chronic dermatitis, and rubra pilaris are among the six types of ESD. Erythemato-squamous illnesses are crusty inflammatory skin conditions that usually affect the scalp and cheeks. There are six different types of erythema and scale, but they all have had very similar clinical symptoms. Genetic and environmental factors can contribute to dermatological disorders, which tend to manifest themselves at certain life stages, such as late childhood/early adolescence. Skincare differential data analysis is a tough subject since one illness may show early signs of another disease and then develop its own characteristics as the disease progresses through its life cycle.

\* **Corresponding author: Ravinder Reddy Baireddy**, Department of ISE, Bangalore Institute of Technology, Bangalore, Karnataka, India, e-mail: ravibaireddy@gmail.com

**R. Nagaraja:** Department of ISE, Bangalore Institute of Technology, Bangalore, Karnataka, India, e-mail: profnagaraja@gmail.com

A huge dataset is mined for hidden, previously unknown, and potentially useful information using a data mining technique. Data mining is an interdisciplinary approach for finding and interpreting important information. On this basis, the following two parts described the data mining techniques utilized for categorizing ESD illnesses. Large amounts of medical information are now available as a result of recent technological advances. These records are full of useful information. As a result, data mining techniques are employed to uncover important patterns. This paper will provide an overview of data analysis and its many methodologies, as well as a review of the medical data mining research. The importance of diverse data mining approaches is highlighted. In medical analytics, classification is a reliable method.

A technique presented in this paper [1] can be used to identify skin disorders. In a two-stage technique for accurate disease diagnosis, computer vision and machine learning are combined with clinically evaluated histological features. When the system was trained and tested on six diseases, it obtained a 95% accuracy rate.

When it comes to preparing effective and efficient data for data mining, data pre-processing is a key component. When it comes to developing complete predictive models, feature selection is crucial since it provides significant features [2]. The use of FSMs such as forward selection and backward elimination to pick the optimal subset of characteristics reduces training time and improves system performance employed in this prediction model.

In this work [3], a predictive framework for ESDs is constructed using a mix of classifiers. It is suggested that the model be broken down into two stages. A simple set-based feature selection approach is used to choose an initial feature subset. In the second phase, an aggregation of three classifiers consisting of MLP, KNN, and SVM is created. Finally, each classifier arrives at a diagnosis result using the majority voting approach. The recommended model achieves 97.78% accuracy in ensemble categorization.

Self-organization maps (SOM) of clinical and pathological findings will be used in the present investigation to explore clusters associated with different illnesses [4]. A SOM is an encompass of neurons that adapts to a dataset's topology structure, enabling us to view massive datasets and find probable groupings. By constantly shifting its neurons directly to the data points, a SOM understands the form of a dataset. SOM is a size reduction method that uses multidimensional input to reduce the issue. As a consequence, the artificial intelligence program SOM was used to classify the six kinds of erythematous-squamous skin conditions. The classification and regression tree (CART) was used to predict the ESD diagnosis [5]. A CART is a computational intelligence forecasting technique. It illustrates how the quantities of a target attribute can be anticipated using other variables. It is a tree structure with each fork separated into a predictor variable and predictions for the target attribute at the end of each node [6]. The proposed model was 94.84% accurate.

Bayesian network, support vector machine (svm), logic boost, multilayer perceptron (MLP), Random Forest, and J48 data mining classifier were utilized in the current research to predict the illness [7]. They were able to diagnose illnesses with an accuracy of 83.6 and 84.5%, respectively, by using Bayesian Network and Logic Boost methods.

For skin disease detection, a hybrid model combining Case-Based Reasoning and Artificial Neural Networks methods is suggested in ref. [8]. There are nine input factors (attributes) that have a significant impact on the skin diagnosis process in the suggested model. To implement the concept, a user-friendly and interactive computer program has been built.

To identify erythematous-squamous illness, ChiGA (Chi-Square and Genetic Algorithm) is a unique feature selection approach [9]. ChiGA selects the optimum feature subset by combining the benefits of filter and wrapper techniques. Ninety-eight percent of the characteristics are optimal, with 18 of them being particularly accurate.

Dermatology disease processing technique is defined in this study. Each of the three stages of the suggested work paradigm is described in more detail [10]. Statistical parameters are used in the first step of data processing and analysis. To create the predictive cell structure, these factors are monitored with probabilistic measurements in the second step of the process. Bayesian networks are used for illness categorization in the final stage. As a result of its implementation, the method's outcomes have been very accurate.

Filter technique and Partial Forward Search (F PFS algorithm) is a novel weighted probabilistic strategy for feature selection that determines the best models of support vector machines to detect various skin disorders [11]. Overfitting is reduced and accuracy is improved, especially when using a classifier like a support vector machine.

This paper is structured as follows: In Section 2, we will look at some of the most recent methods for classifying skin disorders. Section 3 explains the shortcomings of the current and the originality of the suggested technique. Section 4 describes the clustered and optimized classifier. Based on the suggested technique, results and discussion are presented in Section 5. Lastly, the paper concludes with a summary, which is expanded in Sections 6 and 7.

## 2 Literature survey

With the WEKA data mining tool, this article [12] seeks to forecast skin disorders by combining the following algorithms Neural Network, Decision Tree, and Ensemble Method based on Majority Voting. The prediction algorithm was evaluated using a dermatology dataset. Based on accuracy and precision as well as true positive and false positive, the performance evaluation was conducted to determine the effectiveness of the test. It creates a superior skin disease prediction model with significantly greater accuracy and generalization, which avoids both over-fitting and under-fitting.

A combination Dynamic ant colony model with 3 updating stages, wavelets transform, and singular value decomposition incorporating svm was suggested in this article [13]. The suggested method's intended goal is to reduce a subset of characteristics to provide an acceptable illness diagnosis with the best accuracy, sensitivity, and specificity across a wide variety of disorders. Five benchmark medical datasets of diverse illnesses will be used to assess the proposed approach.

In the categorization of Erythematous-squamous Dataset, Naive Bayes has a 100% success rate [14]. As a consequence of the integration of correlation-based filter methods and different heuristic search methods, the number of characteristics has been decreased from 34 to 19. To locate important class characteristics, the original data are subjected to the correlation-based feature selection (CFS) process. Continuous and discrete problems are filtered using a fast correlation-based filtering technique. To reduce dimension, features are considered to use the relief method.

They [15] represent a new strategy in this paper that combines six distinct data mining categorization techniques. The most important 15 characteristics that will play a key influence in prediction are chosen using a correlation-based feature technique. Then, combining Bagging, AdaBoost, and Gradient Boosting classifier methods, they create an ensemble approach. Gradient boosting is a sort of boosting used in deep learning. It is based on the assumption that whenever the best potential future model is coupled with prior models, the prediction accuracy error is minimal. If a slight modification in a case's forecast generates no modification in error, the case's subsequent targeted result is nil [16]. When applied to dermatological datasets, the ensemble approach and feature selection produce better results than individual classifier methods. The ensemble technique seems to be more efficient and precise in predicting skin diseases.

When utilizing data mining algorithms for the prediction of skin disorders, this article [17] explores the effect of feature selection approaches. PCA, GA, and Chi-square were the feature selection methods used to minimize the number of features in a dataset of skin disorders. Feature selection methods did not increase the precision and sensitivities of these methods, according to empirical observations.

Psoriasis – Cronic Dermatitis and Seborreic Dermatitis Pitriasis Rosea were identified in the same cluster, and the diagnoses are confounded as a result of this procedure [4]. Clinical and histological results of the illnesses were grouped by the SOM technique to avoid this misunderstanding [18]. SVM was used to classify clinical and histological results that clustered together. This led to the categorization of psoriasis and chronic dermatitis at 0.89 and 0.93, respectively, and seborreic dermatitis with pitriasis rosea at 0.79 and 0.80.

When applied to dermatological datasets, the ensemble technique and feature selection produce better results than individual classifier approaches [19]. A more accurate and efficient way to predict skin disorders appears to be using the ensemble approach. As a novel method, they [20,21] integrate data mining classification techniques from six different sources.

NB, K-Nearest Neighbor (KNN), SVM, Random Forest (RF), and Multilayer Perception are some of the machine learning algorithms they utilized for skin diseases classification [22,23]. To improve the performance of the trained model, they employed three fundamental ensemble learning: Bagging, Boosting, and Stacking. A model's utility is measured by its accuracy, precision, sensitivity, and f1-score.

Using Artificial Neural Networks, this research aims to categorize the six types of ESD with great accuracy. Using a backpropagation feed-forward approach, they classified ESDs using the JustNN program [6]. They used the pre-process dataset to train and validate the suggested model. The proposed model has a 98.36% accuracy rate.

Calinski–Harabasz and Davies–Bouldin cluster value indices are used to assist the development of multi-objectives [16]. The clustering is done with the assistance of a  $k$ -means cluster and particle optimization to identify the optimal clusters. The multi-objective problem will be solved by PSO to find the optimal cluster.

### 3 Existing method

In the categorization of Erythematous-squamous Dataset, Naive Bayes has a 100% success rate [14]. As a consequence of the integration of correlation-based filter methods and different heuristic search methods, the number of characteristics has been decreased from 34 to 19. When utilizing data mining algorithms for the prediction of skin disorders, this article explores the effect of feature selection approaches [18]. PCA, GA, and Chi-square were the feature selection methods used to minimize the number of features in a dataset of skin disorders [24]. Feature selection methods did not increase the precision and sensitivities of these methods, according to empirical observations.

The above two algorithms utilized the feature selection approach for skin diseases prediction using supervised classifiers. Due to that, it indicates that feature selection has a lesser role in classification.

Calinski–Harabasz and Davies–Bouldin cluster value indices are used to assist the development of multi-objectives [16]. The clustering is done with the assistance of a  $k$ -means cluster and particle optimization to identify the optimal clusters [25]. The multi-objective problem will be solved by PSO to find the optimal cluster. Yet, this paper also suffers from poor classification.

#### 3.1 Research gap identifications

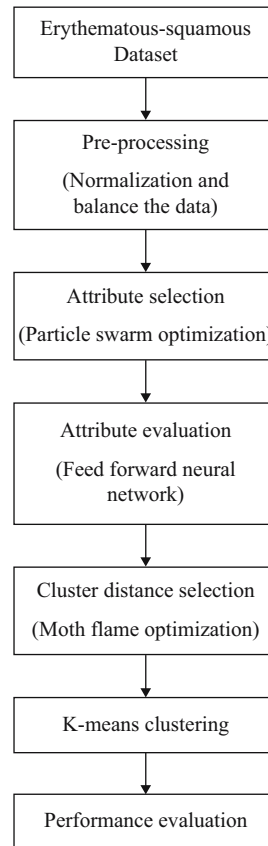
Hence, in this paper, the following objectives were framed to improve the prediction using unsupervised classification.

- Minimal attribute for classification.
- Achieve good prediction using unsupervised classification.
- Utilize minimal attributes for both supervised and unsupervised approaches.

### 4 Proposed method

With the aid of machine learning methods, an optimized data mining strategy is provided in this work to forecast skin disorders. A hybrid particles swarm and moth flame optimization are used to get an optimal

clustered distance value for  $K$ -means grouping. The  $K$ -means clustering misclassification rates and the neural network classification rate are two cluster value indexes that govern multi-objectives. The multi-objective problem would be solved using hybrid PSO to find the best clusters for cluster analysis [26]. The process in the proposed method is shown in Figure 1.



**Figure 1:** Proposed method process flowchart.

Figure 1 depicts the fundamental data mining method for predicting skin disorders, which includes the following steps:

1. The UCI machine learning repository website provided the input skin diseases dataset.
2. To balance the dataset and its values, pre-processing processes such as normalization and rows addition are performed.
3. Next, particle swarm optimization is used to identify the best attribute for predicting diseases.
4. A feed-forward neural network is used to verify the specified characteristics.
5. The best distance metric for clustering is then calculated using the optimal attribute and moth flame optimization.
6. After determining the best attribute and distance, the data are clustered using  $k$ -means.
7. Finally, accuracy is used to assess the suggested method's performance.

The working of each block is explained as follows.

## 4.1 Input dataset

The UCI Machine Learning Repository is a library of datasets, domains ideas, and data sources that the machine learning group uses to test machine learning techniques empirically [27]. The UCI machine learning repository website provided the input skin diseases dataset [26]. There are 366 examples in the dataset, with 34 inputs and six distinct ESD classes. After removing eight samples with missing data, the total number of samples is decreased to 358. Patients are initially clinically evaluated with 12 distinct tests during clinical diagnosis. Dermatologists, on the other hand, seek a biopsy to extract 22 pathological characteristics from each sample.

With the exception of age and family history, all clinical and histological characteristics are graded on a scale of 0–3. When the greatest possible value happens, it takes 3 (maximum), but if the feature does not exist, it takes 0 (NON). It corresponds to Low or Medium, respectively, for relative intermediate values 1 and 2. If any of these illnesses are seen in a relative, value 1 is assigned to the family history characteristic. The record distribution for the six classes is shown in Table 1.

**Table 1:** Record distribution in the dataset

Classes	No of records
Psoriasis	111
Seborrheic dermatitis	60
Lichen planus	71
Pityriasis rosea	48
Chronic dermatitis	48
Rubra pilaris	20

Table 1 reveals that, in comparison to the other classes, class six has the fewest records. As a result, the *K*-means cluster does not attain a high predictive value in the existing work. As a result, data balancing is done in this study.

## 4.2 Pre-processing

There are two pre-processing phases in this work. The first step is to balance the data for class six by summing the mean and median attribute values for 28 records. The balanced data is then normalized to avoid the overfitting problem in the classifier using *Z*-score normalization as in equation (1).

$$Z \text{ score} = \frac{\text{Score} - \text{Mean}}{\text{Standard deviation}}. \quad (1)$$

The notion and process of placing multiple parameters on the very same scale are known as *Z*-score normalization. The *Z*-score is an extremely important data statistic since it enables a data manager to determine the likelihood of a score happening inside the data normally distributed. It also allows the administrator to evaluate two distinct ratings derived from various information standard normal distribution.

## 4.3 Attribute selection using particle swarm optimization

In this stage, the pre-processed data is put through an optimization procedure to find the best characteristics out of a total of 34. It is determined by solving equation (2) objective function for optimization.

$$\text{PSO}_{\text{of}} = \min(1 - \text{FFNN}_{\text{Accuracy}}). \quad (2)$$

The particle swarm uses the following procedure to find the two optimum characteristics based on this equation and the accompanying Table 2 initialization values.

**Table 2:** PSO Initialization parameters

Variable name	Value
Swarm count	50
Iterations (it)	100
Lower bound	[11]
Upper bound	[34 34]
Learning coefficients for swarms S1 and S2	1.5 and 2.0

Individuals, known as a particle, were “capable of flying” through a hyper-dimensional solution space in PSO. Variations in the location of particles inside the search area are predicated on people’s social-psychological propensity to imitate other people’s opinions. A particle’s search behavior is thus influenced by the activity of many other particles in the swarm. The selection procedure seems to be that particles probabilistically returned to previously popular places in the searching space as a result of replicating these behavioral patterns.

The use of particle swarms to identify prey based on bird movements has been developed [28]. The particles are used to figure out which features are most useful in predicting skin diseases. Here, two lower and higher boundary conditions are used, encompassing the range [11 to 34 34]. A hundred times were performed on every one of the 50 particles or birds. Particle mobility is limited by the S1 and S2 learning factors, which have values of 1.5 and 2. Overall iteration best solution (OIBS) and current iteration best solution (CIBS) are the two speed-up options for particle movement. The speed is updated using both OIBS and CIBS. These components result in the next position (NP) and velocity (V) update equations.

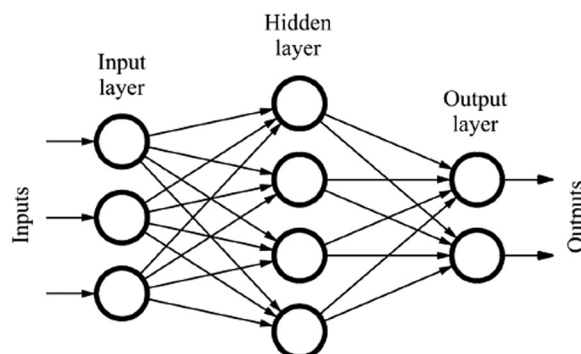
$$\begin{aligned} \text{swarm}_V(it + 1) = & \text{swarm}_V(it) + S1 * \text{rand } 1, 0 * \text{CIBS}(it) \\ & - \text{swarm}_{NP}(it) * S2 * \text{rand } 1, 0 * \text{OIBS}[it] - \text{swarm}_{NP}(it), \end{aligned} \quad (3)$$

$$\text{swarm}_{NP}(it + 1) = \text{swarm}_{NP}(it) + \text{swarm}_V(it). \quad (4)$$

The swarm updates its locations and finds the two best characteristics from the pre-processed dataset using the aforementioned equations.

#### 4.4 Attribute evaluation using feed-forward neural network

As illustrated in Figure 2, the feed-forward neural network is trained with 70% of the best records from the specified characteristics in this stage. A feedforward network’s purpose is to estimate a mapping function



**Figure 2:** FFFN process flow.



and input to a classification. A feedforward system creates a mapping and trains the values of parameters which create great estimations of the functionality. The data flows from the inputs to the functional getting assessed, then to the intermediate calculations necessary to define it, and finally to the result. That there were no feedback links; therefore, the model's outputs do not respond to the valuable feedback into it [29].

The optimum characteristics from the previous phase are used as inputs. Each input has its own weights based on the Levenberg–Marquardt function, which are modified in the hidden layer by multiplying with the Gaussian function. Finally, a single hidden layer with 10 hidden neurons is used to connect the input and output classes. The network is trained with the best attribute using this method. The trained network is then tested using 30% of the dataset to assess the particle swarm optimization-selected characteristic. Calculating the accuracy of the testing dataset is used to do the attribute assessment.

After that, using multi-objective moth flame optimization, the best cluster distance is found.

## 4.5 Cluster distance selection using Moth flame optimization

In this stage, the optimal attribute data is put through an optimization procedure to find the best cluster distance for  $k$ -means clustering. It is determined by solving equation (7) objective function for optimization.

$$\text{MFO}_{\text{of1}} = 1 - K\text{means}_{\text{Accuracy}}, \quad (5)$$

$$\text{MFO}_{\text{of2}} = \text{Silhouette value}, \quad (6)$$

$$\text{MFO}_{\text{of}} = \min(\text{equation 5, equation 6}). \quad (7)$$

MFO is a metaheuristics paradigm that replicates the cognition of moth, known as transverse orientation, which is employed to travel in an environment, and the effectiveness of MFO has been proven to be rather good in existing works to tackle clustering-based difficulties. The MFO's advantages, such as convergence speed to the global optimal solution as well as the ability to ignore local optima, are taken advantage of. Moths and their locations are assumed to be potential answers and issue variables, accordingly, by the algorithm. Moths may change their location vectors and fly in multiple dimensions search domains. As with other optimization techniques, MFO starts with a randomized set of feasible solutions within the boundary's region [30].

The moth flames use the following procedure to find the optimal  $K$ -means clustering distance based on these equations and the accompanying Table 2 initialization values.

Here, the MFO process to select the optimal cluster distance from three distance metrics such as Euclidean, squared Euclidean, and city block. The actual straight-line difference between adjacent places in Euclidean space is measured by Euclidean distance that could be squared to given things being further apart increasingly more value; this isn't a metric, but that's a good way to compare distances. Moreover, the absolute discrepancies in the locations of the two particles are investigated using city block lengths. The traditional distance function or metrics of Euclidean geometry is substituted by a new measurement wherein the distance between two is equivalent to the total of their Cartesian variable's absolute differences (Table 3).

**Table 3:** MFO Initialization parameters

Variable name	Value
Search agents	50
Iterations (it)	100
Lower bound	[1]
Upper bound	[3]



However, it obtains two solutions in MFO for moth and light. Moth denotes the search agent, which roams the search area in quest of the best solution [31]. The light shows the solution's best cluster distance. As a result, both of these strategies were utilized to update the moth position. The moth travels in a logarithmic spiral movement (MLSM) in this case, as shown in equation (8).

$$\text{MLSM}(M, L) = d_m e^{bt} \cos 2\pi(t) + L_l. \quad (8)$$

The letters  $M$  and  $L$  stand for moth and light, respectively. The letters  $I$  and  $j$  stand for the moth and light numbers, respectively. The constant values for  $b$  and  $t$  should be between  $-1$  and  $1$ . In equation (8), the quantity  $d$  denotes the spacing between the moth and the light. The spacing between moth and light is given in equation (9).

$$d_m = |M_m - L_l|. \quad (9)$$

Because moths are drawn to artificial light, the following condition was utilized to update the position of the insect.

$$\text{LN} = \text{round}(\text{LN}_{\max} - it) \quad (10)$$

The terms  $\text{LN}$  and  $\text{LN}_{\max}$  refer to the light number and the maximum light number utilized in the optimization process, respectively.

The moth updates its locations and its flames and finds the optimal cluster distance for the clustering process using the aforementioned equations.

## 4.6 K-means clustering

K-Means is a centroid-based technique that assigns a median to each cluster. The main goal of this technique is to reduce the number of ranges among data points and the clusters that they belong to. It is an iterative technique that splits unprocessed information into  $k$  groups, for each dataset belonging to only one group of similar qualities.

The ideal distance metric from MFO, as well as records holding the optimal attributes, was sent into the K-means clustering at this stage. The distance between the point and the centroid is then used to allocate the data points to each category. It will be paired with the data point with the shortest distance cluster centroid. It is based on equation (11) following function.

$$K \text{ means} = \sum_{i=1}^{\text{NR}} \sum_{k=1}^{\text{NC}} CP_{ik} \|d_i - \text{mean}_k\|^2. \quad (11)$$

The term  $CP$  in equation (11) shows where the data point belongs in the clusters. If the data point is part of that cluster, it will be 1; otherwise, it will be 0. Number of classes ( $\text{NC}$ ) = 6 and Number of records ( $\text{NR}$ ) = 375 in this case. Mean  $k$  is used to signify the cluster's centroid, which is provided in equation 12.

$$\text{Mean}_k = \frac{1}{\text{NR}_k} \sum_{d_i \in C_k} d_i. \quad (12)$$

Using these equations, the optimized dataset is classified using K-means clustering and then subjected to the performance evaluation.

## 4.7 Performance evaluation

The accuracy of unsupervised categorization by K-means clustering is assessed using the Accuracy measure.

Three measures are used to evaluate attribute selection using FFNN: accuracy, sensitivity, and specificity.

#### 4.7.1 Research gap identifications

The proportion of correct predictions (including true positives and true negatives) among the total number of instances studied is known as accuracy.

#### 4.7.2 Sensitivity

The capacity of a test to accurately discover records that do have a particular illness is referred to as sensitivity.

#### 4.7.3 Specificity

The capacity of a test to accurately reject a record without a particular disease is referred to as specificity.

## 5 Implementation and discussion

The proposed approach was implemented in this study using the MATLAB R2020b program in a Windows 10 environment. Any three-variable optimization issue with two inequity style restrictions can be solved using particle swarm optimization algorithms. In MATLAB scripts which can be added to the Linked data, the programs can simply be modified to include more variables and restrictions. MATLAB can examine larger amounts of data and scale it up to cluster quickly.

The input for the processing in MATLAB is taken in the form of Excel data as shown in Figure 3.

	A	B	C	D	E	F	G	H	I	J	K
	erythema	scaling	definite_borders	itching	koebner_phenomenon	polygonal_papules	follicular_papules	oral_mucosal_involvement	knee_and_elbow_involvement	scalp_involvement	family_history
1											
2		2	2	0	3	0	0	0	0	1	0
3		3	3	3	2	1	0	0	0	1	1
4		2	1	2	3	1	3	0	3	0	0
5		2	2	2	0	0	0	0	0	3	2
6		2	3	2	2	2	0	0	2	0	0
7		2	3	2	0	0	0	0	0	0	0
8		2	1	0	2	0	0	0	0	0	0
9		2	2	3	3	3	3	0	2	0	0
10		2	2	1	0	2	0	0	0	0	0
11		2	2	1	0	1	0	0	0	0	0
12		3	3	2	1	1	0	0	0	2	2
13		2	2	0	3	0	0	0	0	0	0
14		3	3	1	2	0	0	0	0	0	1
15		2	3	3	0	0	0	0	0	1	1
16		2	2	3	3	0	3	0	2	0	0
17		1	1	0	1	3	0	0	0	0	0
18		2	2	1	3	0	0	0	0	0	0

Figure 3: Input dataset format.

Pre-processing is applied to the input dataset, and the values are normalized using Z-score normalization, which transforms the values in the dataset from 0 to 3 to different values as seen in Figure 4.

	1	2	3	4	5	6	7
1	-0.1128	0.2883	-1.6798	1.4590	-0.6857	-0.4612	-0.334
2	1.3978	1.7297	1.6182	0.5708	0.4226	-0.4612	-0.334
3	-0.1128	-1.1532	0.5189	1.4590	0.4226	2.7026	-0.334
4	-0.1128	0.2883	0.5189	-1.2055	-0.6857	-0.4612	-0.334
5	-0.1128	1.7297	0.5189	0.5708	1.5310	1.6480	-0.334
6	-0.1128	1.7297	0.5189	-1.2055	-0.6857	-0.4612	-0.334
7	-0.1128	-1.1532	-1.6798	0.5708	-0.6857	-0.4612	-0.334
8	-0.1128	0.2883	1.6182	1.4590	2.6393	2.7026	-0.334
9	-0.1128	0.2883	-0.5805	-1.2055	1.5310	-0.4612	-0.334
10	-0.1128	0.2883	-0.5805	-1.2055	0.4226	-0.4612	-0.334
11	1.3978	1.7297	0.5189	-0.3174	0.4226	-0.4612	-0.334
12	-0.1128	0.2883	-1.6798	1.4590	-0.6857	-0.4612	-0.334
13	1.3978	1.7297	-0.5805	0.5708	-0.6857	-0.4612	-0.334
14	-0.1128	1.7297	1.6182	-1.2055	-0.6857	-0.4612	-0.334
15	-0.1128	0.2883	1.6182	1.4590	-0.6857	2.7026	-0.334
16	-1.6233	-1.1532	-1.6798	-0.3174	2.6393	-0.4612	-0.334

Figure 4: Pre-process output.

The pre-processed data is then submitted to particle swarm optimization to pick the best attribute, with the associated convergence curve for attribute selection illustrated in Figure 5.

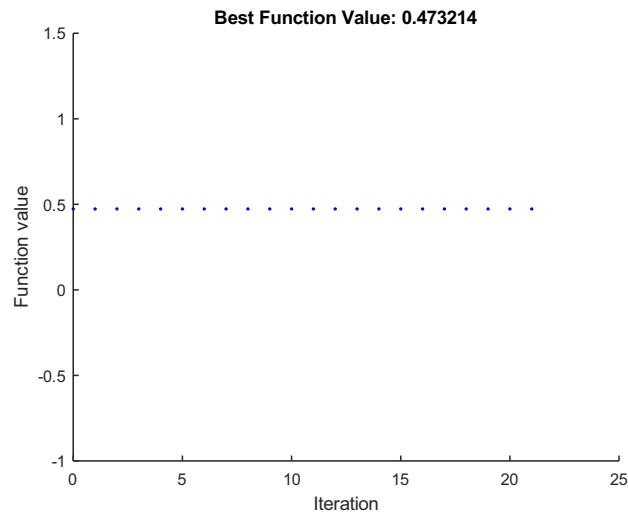


Figure 5: PSO convergence curve.

From Figure 5, it is observed that the optimal attribute is determined after the 20th iteration itself and the corresponding objective function value is 0.47 at the end of the process. From the PSO, it was observed that the scalp involvement and clubbing of the rete edges were important attributes of the classification process.

The records with the best attributes are then put through a neural network to see how well they choose their attributes. Figure 6 depicts the network's training procedure.

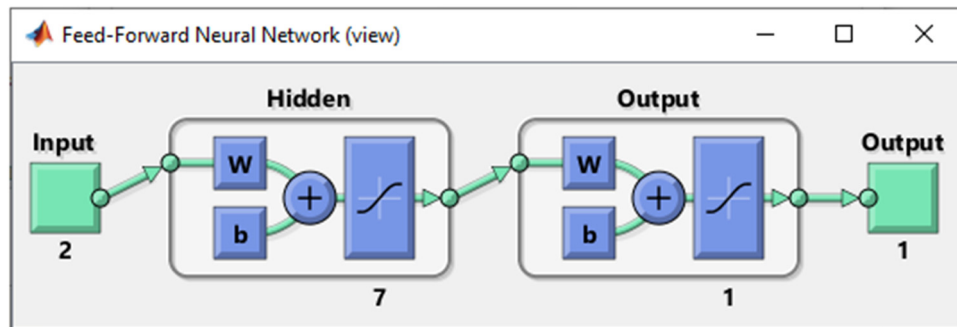


Figure 6: FFNN training.

Then, the FFNN is tested with the testing dataset, then it is evaluated using evaluation metrics and is shown in Figure 7.

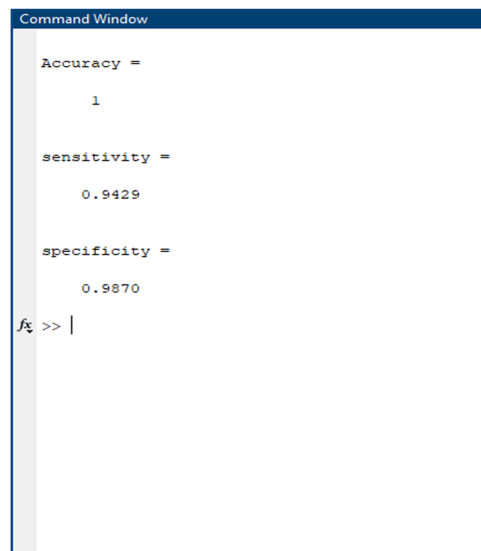


Figure 7: Attribute selection evaluation.

Then, the optimal cluster distance for the optimal records is determined using Moth flame optimization and the corresponding convergence curve is shown in Figure 8.

Finally, from MFO, the optimal distance for  $K$ -means clustering is selected as a city-block, and the optimal dataset is classified using an unsupervised algorithm called  $K$ -means clustering, and its performance is evaluated in terms of accuracy as shown in Figure 9.

The comparison between the existing and proposed approaches in terms of performance evaluation is listed in Tables 4 and 5.

Table 4 compares the performance of the testing dataset with original and adjusted data for attribute selection using an improved feed-forward neural network. The proposed normalized method increased the accuracy to 100%. It shows that it can categorize all six classes; however, identifying the specific classes is decreased by 6% when compared to prior studies.

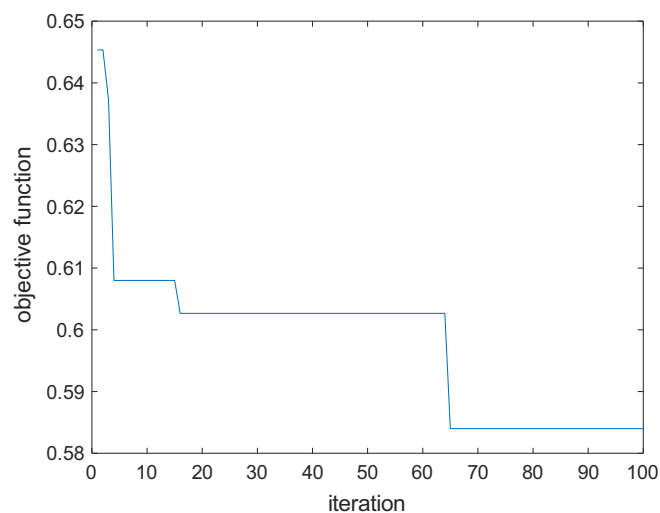


Figure 8: MFO convergence curve.



Figure 9: Evaluated in terms of accuracy.

Table 4: Attribute selection comparison

Metric	PSO-original data	PSO-normalized data
Accuracy	75	100
Sensitivity	100	94.29
Specificity	100	98.70

Table 5: Overall performance comparison

Metric	PSO	PSO-MFO
Accuracy	75	85

The purpose of this article, on the other hand, is to undertake an unsupervised classification using only a few parameters to discover sickness subgroups and give effective treatment. As a result, Table 5 compares the performance of the current and proposed work.

Table 5 demonstrates that when compared to previous work, the suggested increased its unsupervised classification accuracy by 10%. It was accomplished utilizing MFO and a normalized dataset, as well as cluster distance selection. As a result, the suggested hybrid PSO-MFO is ideal for data mining in the categorization of skin disorders.

## 6 Conclusion

Multi-objective-based Hybrid optimal clustering was used in this work. The findings of their experiments were compared to the existing particle swarm-based clustering approach. Six types of ESDs are included in the dataset, all of which have redness (erythema) as a symptom that causes cell death (squamous). The UCI skin diseases undergo various process steps to determine the optimal attribute and optimal cluster distance for clustering. The research provided comprehensive results for all tests, as well as tables and illustrations for various comparisons. Three distinct performance measures generated from the multiclass confusion matrix were used to make the comparisons. They're accuracy, sensitivity, and specificity. Based on the results, it was observed that the proposed method improved its performance over the existing work.

- The normalization and data balance improved the prediction process.
- The attribute selection reduces the computational time for processing the testing data.
- The cluster distance selection improved the unsupervised prediction accuracy.

## 7 Future work

In the future, an ideal clustering technique will be chosen to enhance all of the assessment criteria for the prediction process.

**Conflict of interest:** Authors state no conflict of interest.

## References

- [1] Kumar VB, Kumar SS, Saboo V. Dermatological disease detection using image processing and machine learning. In 2016 Third International Conference on Artificial Intelligence and Pattern Recognition (AIPR). Xiamen, China: IEEE; 2016, September. p. 1–6.
- [2] Raghavendra S, Indiramma M. Hybrid data mining model for the classification and prediction of medical datasets. *Int J Knowl Eng Soft Data Paradig.* 2016;5(3–4):262–84.
- [3] Lahijanian B, Farahani FV, Zarandi MF. A new multiple classifier system for diagnosis of erythemato-squamous diseases based on rough set feature selection. In 2016 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE). Vancouver, BC, Canada: IEEE; 2016, July. p. 2309–16.
- [4] Fidan U, Ozkan N, Calikusu I. Clustering and classification of dermatologic data with Self Organization Map (SOM) method. In 2016 Medical Technologies National Congress (TIPTEKNO). Antalya, Turkey: IEEE; 2016, October. p. 1–4.
- [5] Maghooli K, Langarizadeh M, Shahmoradi L, Habibi-Koolaei M, Jebraeili M, Bouraghi H. Differential diagnosis of erythmato-squamous diseases using classification and regression tree. *Acta Inform Med.* 2016;24(5):338–42.
- [6] Shakeel P, Manogaran G. Prostate cancer classification from prostate biomedical data using ant rough set algorithm with radial trained extreme learning neural network. *Health Technol.* 2018;10(1):157–65. doi: 10.1007/s12553-018-0279-6.
- [7] Geetha BK. Diagnosis of erythemato techniques. *Aust J Basic Appl Sci.* 2017;11(5):45–51.
- [8] Dabowsa NIA, Amaitik NM, Maatuk AM, Aljawarneh SA. A hybrid intelligent system for skin disease diagnosis. In 2017 International Conference on Engineering and Technology (ICET). Antalya, Turkey: IEEE; 2017, August. p. 1–6.

- [9] Maryam Setiawan NA, Wahyunggoro O. A hybrid feature selection method using multiclass SVM for diagnosis of erythematous-squamous disease. In *AIP Conference Proceedings*. vol. 1867, no. 1. AIP Publishing LLC; 2017, August. p. 020048.
- [10] Rani S. A dual phase probabilistic model for dermatology classification. In *Computer communication, networking and internet security*. Singapore, Australia: Springer; 2017. p. 443–50.
- [11] Parikh KS, Shah TP. Feature selection paradigm using weighted probabilistic approach. *Int J Adv Sci Technol*. 2017;100(3):1–14.
- [12] Oladele TO, Aro TO, Segun AS. Skin diseases predictive model using individual base and ensemble base approach. *Comput Inf Syst*. 2018;22(3):143–52.
- [13] Al-Tashi Q, Rais H, Abdulkadir SJ. Hybrid swarm intelligence algorithms with ensemble machine learning for medical diagnosis. In *2018 4th International Conference on Computer and Information Sciences (ICCOINS)*. Kuala Lumpur, Malaysia: IEEE; 2018, August: p. 1–6.
- [14] Çetin A, Gökhan T. Differential diagnosis of erythematous squamous diseases with feature selection and classification algorithms. In *Nature-Inspired Intelligent Techniques for Solving Biomedical Engineering Problems*. Newziland: IGI Global; 2018; p. 103–29.
- [15] Verma AK, Pal S, Kumar S. Comparison of skin disease prediction by feature selection using ensemble data mining techniques. *Inform Med Unlocked*. 2019;16:100202.
- [16] Ahilan A, Babu GC, Murugan NS, Parthasarathy A, Manogaran G, Raja C, et al. Segmentation by fractional order darwinian particle swarm optimization based multilevel thresholding and improved lossless prediction based compression algorithm for medical images. *IEEE Access*. 2019;7:89570–80. doi: 10.1109/access.2019.2891632.
- [17] Abikoye OC, Komolafe RG, Aro TO. Performance evaluation of feature selection algorithms on skin disease prediction.
- [18] Fidan U, Uzunhisarcikli E, Çalikuşu İ. Classification of dermatological data with self organizing maps and support vector machine. *Afyon Kocatepe Üniver Fen Ve Mühendislik Bilimleri Derg*. 2019;19(3):894–901.
- [19] Marini F, Walczak B. Particle swarm optimization (PSO). A tutorial. *Chemometrics Intell Lab Syst*. 2015;149:153–65.
- [20] Verma AK, Pal S, Kumar S. Prediction of different classes of skin disease using machine learning techniques. In *Smart innovations in communication and computational sciences*. Singapore: Springer; 2021. p. 91–100.
- [21] Verma AK, Pal S, Tiwari BB. Skin disease prediction using ensemble methods and a new hybrid feature selection technique. *Iran J Computer Sci*. 2020;3(4):207–16.
- [22] Pal M, Roy BR. Evaluating and enhancing the performance of skin disease classification based on ensemble methods. In *2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)*. UIU, Dhaka, Bangladesh: IEEE; 2020, November. p. 439–43.
- [23] Verma AK, Pal S, Kumar S. Prediction of skin disease using ensemble data mining techniques and feature selection method – a comparative study. *Appl Biochem Biotechnol*. 2020;190(2):341–59.
- [24] Saravanan V, Pralhaddas KD, Kothari DP, Isaac Woungang. An optimizing pipeline stall reduction algorithm for power and performance on multi-core CPUs. *Hum Cent Comput Inf Sci*. 2015;5:2. doi: 10.1186/s13673-014-0016-8.
- [25] <https://archive.ics.uci.edu/ml/datasets/dermatology#::~:~:text=The%20diseases%20in%20this%20group,many%20histopathological%20features%20as%20well>.
- [26] Mathan K, Kumar PM, Panchatcharam P, Manogaran G, Varadharajan R. A novel Gini index decision tree data mining method with neural network classifiers for prediction of heart disease. *Autom Embedded Syst*. 2018;22(3):225–42. doi: 10.1007/s10617-018-9205-4.
- [27] Sathishkumar VE, Park J, Cho Y. Using data mining techniques for bike sharing demand prediction in metropolitan city. *Computer Commun*. 2020;153:353–66. doi: 10.1016/j.comcom.2020.02.007.
- [28] Al-Kahlout BI, Naeem MM, Shepherd MJ. ANN for the classification of erythematous-squamous disease. Gaza, Palestine; 2021
- [29] Mirjalili S. Moth-flame optimization algorithm: A novel nature-inspired heuristic paradigm. *Knowl Syst*. 2015;89:228–49.
- [30] Kumar PM, Hong CS, Afghah F, Manogaran G, Yu K, Hua Q, et al. Clouds proportionate medical data stream analytics for internet of things-based healthcare systems. *IEEE J Biomed Health Inform*. 2022 Mar;26:973–82. doi: 10.1109/JBHI.2021.3106387.
- [31] Reddy BR. A novel approach for diagnosis of dermatologic diseases based on multi-objective clustering using particle swarm optimization. *Turkish J Computer Math Educ (TURCOMAT)*. 2021;12(9):1161–72.