Environment and Social Psychology (2025) Volume 10 Issue 7
doi: 10.59429/esp.v10i7.3833

ISSN: 2424-8975 (O)
2424-7979 (P)

RESEARCH ARTICLE

Understanding social and environmental behavior patterns through hybrid neural swarm clustering

Xiangqin Dai[1,2], Mohd Najwadi Yusoff[1,*], Lei Wang[3], Xiangguang Dai[2]

[1] *School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang, 11800, Malaysia*

[2] *Key Laboratory of Intelligent Information Processing and Control, Chongqing Three Gorges University, Chongqing, 40044, China*

[3] *School of Electrical and Electronic Engineering Engineering Campus, Universiti Sains Malaysia, Nibong Tebal 14300, Malaysia*

*** Corresponding author:** Mohd Najwadi Yusoff, najwadi@usm.my

ABSTRACT

Most of the clustering problems can be reformulated as combinational optimization problems. It is not easy to search for the global solution for combinational optimization problems. In this paper, we use the Discrete Hopfield Neural Network (DHN) to solve the Bipartition Clustering Problem (BCP) and combine Particle Swarm Optimization (PSO) to search for the global solution. Firstly, the BCP is reformulated into an integer optimization problem. Secondly, to ensure the local solution of BCP in convergence and stability, some rules of DHN are designed to solve the Integer Optimization Problem(IOP). Finally, PSO is proposed to reset the neuron states of DHN until the global solution of BCP is achieved. Numerical and real-world experiments are conducted to evaluate the validity and feasibility of the proposed method. The experimental results show that our method achieves better clustering results on different datasets and problem instances, and has higher accuracy and stability compared to traditional methods.

Keywords: Discrete hopfield neural network; particle swarm optimization; integer optimization problem; bipartition

1. Introduction

1.1. Clustering summary individuals organize and interact in both physical and digital

With the rapid growth of cloud storage, sensor technologies, ecosystems. social networks, and high-throughput computing, the volume of data being generated today is unprecedented. These From a machine learning standpoint, clustering has long massive datasets hold significant social, environmental, and been a foundational tool in unsupervised learning [1], with economic value, yet they also pose serious challenges in widespread applications in image segmentation [2], social terms of storage, analysis, retrieval, and real-time media mining [3,4], information retrieval [5,6], and biological processing. In the context of pressing global issues such as systems analysis [7–9]. Common clustering techniques such as climate change, digital polarization, and urban as K-means, spectral clustering, and hierarchical clustering overpopulation efficient data summarization

ARTICLE INFO
Received: 23 June 2025| Accepted: 26 June 2025 | Available online: 18 July 2025

CITATION
Dai XQ, Yusoff MN, Wang L, et.al. Understanding Social and Environmental Behavior Patterns Through Hybrid Neural Swarm Clustering. *Environment and Social Psychology* 2025; 10(7): 3833 doi:10.59429/esp.v10i7.3833

has become have evolved into core components of data analysis not only a technical need but also a societal imperative.

K-means divides datasets into k clusters by Clustering algorithms provide a crucial tool for this task by minimizing intra-cluster distances, spectral clustering grouping data into cohesive categories based on underlying leverages graph theory to cut graphs based on similarity, allowing researchers to uncover patterns of matrices, and hierarchical clustering iteratively merges behavior, community structures, or environmental risk zones clusters based on linkage criteria. These techniques have that would otherwise remain hidden in raw data. proven highly useful in a range of real-world scenarios, including urban planning, fraud detection, medical diagnosis, In social and environmental psychology, clustering plays a and educational analytics. growing role in understanding group formation, collective behavior, and decision-making under uncertainty. For instance, K-means has successfully been applied to instance, clustering can be used to detect online echo segment users vulnerable to financial fraud [10] and assist chambers, segment populations based on environmental in retinal vessel segmentation for eye disease detection [11]. attitudes, or model how communities respond to crises like In urban development, clustering helps identify strategic natural disasters or pandemics. By extracting meaningful zones for industrial development [12], while in transportation, groupings from complex datasets, clustering facilitates rapid spectral clustering improves traffic flow prediction by interpretation and provides actionable insights for policy- modeling collective mobility behavior [13]. In engineering and environmental risk management, clustering supports early warning by predicting rock burst intensity in underground construction. Furthermore, its ability to model community resilience, digital activism, and sustainability behavior patterns provides valuable insights for policymakers, urban planners, and behavioral scientists [14].

However, despite their widespread success, traditional clustering algorithms face two major limitations. First, they often do not guarantee convergence to a global optimum, which is critical in high-stakes decision contexts such as public safety or environmental risk assessment. Second, many methods fail to account for cognitive plausibility or psychological realism, which is essential when modeling human behavior or social systems.

In addressing these challenges, Artificial Neural Networks (ANNs) provide an innovative solution. Originally introduced by Tank and Hopfield[15] to solve discrete optimization problems, neural networks mirror human cognitive processes such as pattern recognition, memory activation, and associative learning. Neural networks offer advantages like parallel processing, self-organization, and the capacity to learn complex, nonlinear relationships that making them especially suitable for clustering in dynamic, high-dimensional datasets.

Yet, traditional ANN models also exhibit limitations. Their reliance on random neuron activation and iterative convergence often leads to local optima, reducing effectiveness in large-scale or sensitive applications. To overcome this, researchers have turned to swarm intelligence methods such as Particle Swarm Optimization (PSO) which are inspired by collective behavior in biological systems, such as bird flocking or fish schooling [16]. PSO excels in global optimization tasks and has been shown to outperform standard methods in domains ranging from medicine to computational social science.

Building on these insights, this paper presents a novel clustering algorithm designed to solve the Bipartition Clustering Problem (BCP) by reformulating it as an integer optimization challenge. Our approach integrates Discrete Hopfield Neural Networks (DHNN) with Particle Swarm Optimization, leveraging the cognitive modeling strengths of ANNs with the global search capacity of swarm-based methods. This hybrid framework not only improves clustering accuracy and convergence speed but also enhances the psychological interpretability and social applicability of the clustering outcomes.

Specifically, our contributions include:

Recasting the BCP as a neural-integer optimization problem, solvable through the synergistic use of DHNN and PSO,

Demonstrating through both numerical simulations and real-world datasets that our algorithm reliably reaches global optima, while maintaining social relevance, environmental scalability, and behavioral validity,

Offering a generalizable framework for applications in social network segmentation, environmental attitude clustering, behavioral market analysis, and community resilience modeling.

This work provides a foundation for interdisciplinary research bridging artificial intelligence, environmental psychology, and social systems analysis. By aligning technical innovation with human behavioral theories, such as social categorization, bounded rationality, and collective action models, our method serves not only as a computational advancement but also as a meaningful contribution to the study of social and environmental behavior patterns in the era of big data.

1.2. Related work

To effectively uncover the structure of social and environmental behavior patterns in complex datasets, this study proposes a hybrid clustering framework combining Discrete Hopfield Neural Networks (DHNN) and Particle Swarm Optimization (PSO). This approach leverages the pattern recognition strength of neural networks with the global search efficiency of swarm intelligence to address the limitations of traditional clustering algorithms in terms of local optima and psychological relevance.

Methodological Adjustments. To enhance the applicability and interpretability of the clustering results in real-world psychological and social domains, this research integrates three domain-specific case studies, moving beyond abstract technical benchmarks.

Social Media Data Clustering for Opinion Group Discovery.A dataset from a major Chinese microblogging platform such as Weibo is used to cluster user-generated content related to climate change and public health. By analyzing sentiment, hashtags, and lexical choices, the hybrid clustering model identifies distinct opinion clusters such as pro-environmental advocates, skeptics, neutral observers, and policy-focused users. These clusters are evaluated based on social psychology theories such as social identity theory and group polarization, validating the social coherence of the groupings.

Environmental Attitude Segmentation. Survey data capturing individual responses to environmental behaviors such as recycling, energy use, climate concern are clustered using the proposed DHNN-PSO method. The resulting segments such as “active environmentalists,” “passive supporters,” and “skeptical non-participants” align with models like the Theory of Planned Behavior and Value Belief-Norm Theory. Unlike generic clustering, this model accounts for psychological constructs such as perceived behavioral control and moral obligation.

Community Resilience Pattern Analysis. Demographic, infrastructural, and social trust data from flood-affected regions is used to assess community resilience patterns. Clustering reveals distinct groups such as “highly connected, resource-rich communities” and “isolated, vulnerable groups.” The clusters are cross-referenced with real-world outcomes such as post-disaster recovery speed and validated through community-level resilience frameworks such as Norris et al.'s Model of Community Resilience.

Psychosocial Validation. Instead of relying solely on technical benchmarks (e.g., silhouette score or Davies-Bouldin index), the performance and relevance of the clustering outcomes are validated against psychological and sociological datasets. Human expert coders and domain specialists assess:

The semantic coherence within clusters,

The alignment with theoretical constructs,

The applicability for targeted interventions such as public policy or mental health support.

This validation ensures that the clustering reflects real behavioral distinctions, not just abstract mathematical separations. This hybrid clustering methodology thus moves beyond traditional data science approaches, offering a powerful tool for uncovering social dynamics, behavioral segmentation, and community-level vulnerabilities in a way that is both technically robust and psychologically grounded.

1.2.1. Discrete Hopfield neural networks

DHN[17] can be regarded as a version of the recurrent neural network, where feedback interconnects neurons from output to input. Suppose that the j th neuron state in the DHN is represented by x_j and the DHN composed of n neuron states can be represented by $x=[x_1,x_2,\ldots,x_n]^T$. Let the initial neuron states of DHN be $x(0)=[x_1(0),x_2(0),\ldots,x_n(0)]^T$. Since the initial neuron states are set, the neuron states will change until the neural network is stable. The state of each neuron in DHN is updated by

$$x_j = f(net_j), j = 1,2,...,n \tag{1}$$

Where net_j is the input of the j th neuron and $f(net_j)$ is the activation function to determine whether the j th neuron is activated. The net_j is defined as follows:

$$net_j = \sum_{i=1}^{n}(w_{ij}x_i - \theta_j) j = 1,2,...,n \tag{2}$$

Where θ_j represents the threshold of net_j. The activation function $f(net_j)$ is defined as follows:

$$x_j = \operatorname{sgn}(net_j) = \begin{cases} 1, net_j \geq 0 \\ -1, net_j < 0. \end{cases} j = 1,2,...,n \tag{3}$$

Generally, the energy function of DHN is described by

$$E(t) = -\frac{1}{2}x^T(t)wx(t) + \theta^T(t)x \tag{4}$$

If DHN wishes to be stable in the asynchronous updating mode, two conditions should be satisfied ($w_{ij}=0$ and $w_{ij}=w_{ji}$). In the asynchronous updating mode, one neuron is activated and the remaining neurons remain unchanged. Thus, DHN activates the j th neuron in the asynchronous updating mode as follows:

$$x_j(t+1) = \begin{cases} \operatorname{sgn}[net_j(t)], j = 1. \\ x_j(t), j \neq 1. \end{cases} \tag{5}$$

1.2.2. Particle swarm optimization

PSO is an evolutionary computation algorithm that models birds' flocking behaviors as they collectively forage for sustenance. Each particle, akin to a bird, possesses a position and velocity. Sharing these attributes with other particles, each then adjusts them for swift food discovery in a collaborative manner. Assume that M is the population size of PSO, where the position of a particle $(i = 1, 2, ..., M)$ at time t is denoted $X_i^t = (x_{i1}^t, x_{i2}^t, ..., x_{il}^t)$ .The velocity of the particle i is denoted by $V_i^t = (v_{i1}^t, v_{i2}^t, ..., v_{il}^t)$. Thus, the following equations are used to adjust the position and velocity of a particle i at the moment:

$$v_{ij}^t = wv_{ij}^{t-1} + c_1 r_1 (p_{ij} - x_{ij}^{(t-1)}) + c_2 r_2 (g_j - x_{ij}^{(t-1)}) \tag{6}$$

$$v_{ij}^t = \begin{cases} v_{\max}, & v_{ij}^t > v_{\max} \\ -v_{\max}, & v_{ij}^t < -v_{\max} \end{cases} \tag{7}$$

$$x_{ij}^t = x_{ij}^{t-1} + x_{ij}^t \tag{8}$$

Where w is the inertia weight, c_1 and c_2 are the acceleration factors, r_1 and r_2 are the random numbers arbitrarily generated in the range of [0,1] , and the maximum velocity $v_{\max}$ is usually used as a constant to limit the particle velocity. p_{ij} and g_j denote the local extremum of j th iteration and the global extremum, respectively. In 1997, Kennedy, Eberhart Kennedy, and Eberhart proposed a Binary PSO(BPSO) algorithm for 0-1 integer programming [18]. In BPSO, a hypercube translates into a discrete space, where each particle is defined by discrete variables. Certain bits within these variables can toggle between 0 and 1. Furthermore, the flip probability of discrete variables correlates with particle velocity, which is updated as follows:

$$v_{id} = v_{id} + \varphi(p_{id} - x_{id}) + \varphi(p_{gd} - x_{id}) \tag{9}$$

Where x_{id} is the position of the particle, v_{id} is the position of the particle, φ is the learning factor, p_{id} is the local optimal position of the particle, p_{gd} is the global optimal position. x_{id} , p_{id} and p_{gd} can only be 0 or 1, and v_{id} is the probability, so its value is limited to [0, 1]. The updated formula for the position of the particle is as follows:

$$x_{id} = \begin{cases} 1, & if\ rand() < sig(v_{id}) \\ 0, & else \end{cases} \tag{10}$$

Where $sig(v_{id}) = \frac{1}{1 + \exp(-v_{id})}$, and the function $sig(v_{id})$ as a transformation limiting function makes the portion size of x_{id} strictly in the interval [0,1], and $rand()$ is used to generate a random number in the interval [0,1]. The bigger v_{id} leads to the greater probability such that v_{id} is 1. Similarly, the smaller v_{id} leads to the greater probability such that v_{id} is 0.

2. Materials and methods

2.1. Expression of discrete clustering problem

Decompose the x dataset into two different data sets x_+ and x_-, where $x_+ \cup x_- = x$ and $x_+ \cap x_- = \varnothing$ .If the samples x_i and x_j belong to the same class, then the distance from w_j is larger, otherwise smaller. Assume that the label of the dataset x is denoted by s , we have:

$$x_+ = \{x_i \in X \mid s_i = +1\} \tag{11}$$

$$x_- = \{x_i \in X \mid s_i = -1\} \tag{12}$$

If samples x_i and x_j belong to the same class, we can conclude that $s_i s_j = 1$ and the distance of x_i and x_j is $s_i s_j = 1$ .Similarly, if samples x_i and x_j do not belong to the same class, then $s_i s_j = -1$ and the the distance is $-w_{ij} s_i s_j$. Based on the above definitions, the bipartition clustering problem can be written as a minimization problem or a maximization problem. In order to be same as the form of Hopfield neural network, the bipartite clustering problem is defined as a minimization optimization problem as follows:

$$s_* = \arg\min -\frac{1}{2} s^T ws \quad s \in \{+1, -1\}^n \tag{13}$$

In general, the distance between samples is measured by Euclidean distance, meaning that the more similar samples, the smaller the distance, or the greater the distance. The Euclidean distance function used in this paper is Gauss kernel function Shi and Malik [19], Xu et al. [20], which measures the dissimilarity between samples. The less similar two samples, the smaller the distance, otherwise the greater the distance. The distance between sample x_i and sample x_j is defined using the Gauss kernel function as follows:

$$w_{ij} = \begin{cases} 0, & \mathit{if} \;\; i = j \\ 2[JKJ]_{ij}, & \mathit{otherwise} \end{cases} \tag{14}$$

Where $J = I - \frac{1}{n} 1 1^T$ is a centering matrix Bauckhage and Welke [3] and K is the Gauss Kernel function.

Although DHN can be used to solve problems (14), it is easy to fall into the local optimal solution. In order to resolve this problem, we design an algorithm to search for the global optimal solution to the problem (14).

2.2. Algorithmic design

The procedure of the dichotomous clustering algorithm based on Hopfield neural network and PSO are as follows:

To initialize the population for several Hopfield neural networks;

Hopfiled neural network is used to solve the problem and several feasible solutions are obtained;

Reset the initial solution by PSO;

Repeat the above steps until the global optimal solution is fixed. The algorithm is detailed in **Table 1**.

Table 1. DHN-PSO algorithm.

DHN-PSO Pseudocode
Input: samples matrix x^{n*m} , the number of class p , parameter σ , initial states $s^{n*p*pop} \in \{-1, \quad 1\}$, population size N , termination criterion M , c_1 , c_2 and $cores$
Output: s_*
Compute W in terms of eqn. (14);
Random initial states s ;
Initialize the present best position;
Initialize the global best objective function value;
Initialize v to zeros;
while $i < M$ **do**
for $j < N$ **do**
Update s by eqn. 3 2
end
Update s according to 9 and 10
end
return s_*

Parameter description: P represents the number of classes, which is 2 in this case for binary classification. σ represents the range of the Gaussian kernel function, and the larger its value, the larger the range of its effect. N represents the number of populations in the PSO algorithm, M represents the termination condition of the algorithm, and C1 and C2 represent the learning factors of the particle swarm algorithm.

3. Result analyze

The application of our hybrid clustering method, combining Discrete Hopfield Neural Networks (DHNN) with Particle Swarm Optimization (PSO), reveals not only improvements in clustering accuracy and consistency but also meaningful insights into human behavior and social-environmental dynamics. While traditional evaluation metrics such as Normalized Mutual Information (NMI), Adjusted Rand Index (ARI), and internal indices like the Davies-Bouldin or Dunn Index, quantify technical clustering quality, these numerical results can also be translated into social and psychological implications.

What Do the Clusters Reveal About Human Behavior? High-performing clusters those with high internal cohesion and clear inter-group separation indicate that the algorithm has successfully identified behaviorally or attitudinally distinct subgroups. When applied to datasets involving human activity such as social media behavior, environmental survey responses, or community engagement patterns.

This research proposes a hybrid clustering approach to uncover the underlying psychological and social characteristics embedded in complex datasets. Specifically, the method identifies value orientations and belief systems such as the contrast between pro-environmental and skeptical attitudes, behavioral consistency and dissonance such as individuals who acknowledge climate change yet take minimal action, as well as psychosocial traits such as trust, perceived control, identity affiliation, and information exposure. Accordingly, the resulting clusters are not merely statistical partitions; rather, they represent psychologically meaningful profiles and socially embedded positions, thereby providing deeper interpretability and theoretical insight.

Applied to environmental research, the proposed approach enables the segmentation of populations into distinct environmental attitude typologies, including engaged environmentalists (high concern and high action), concerned but inactive individuals (cognitive dissonance), indifferent or skeptical groups (low concern and low action), and situational actors (those who act only when convenient or incentivized). These segments are validated against established psychological frameworks such as the Theory of Planned Behavior and the Value-Belief-Norm Theory. Moreover, the clustering outcomes provide practical guidance for environmental communication and policymaking, particularly in designing targeted messaging and tailored behavioral intervention strategies, demonstrating both academic and real-world value.

Beyond environmental applications, the social interpretability of this hybrid framework facilitates broader interdisciplinary interventions. For public campaigns, tailored strategies can be developed to enhance the effectiveness of initiatives in areas such as climate change mitigation, public health promotion, and misinformation management. In education and engagement, schools and universities can leverage these insights to design group-specific sustainability curricula addressing underlying beliefs and behavioral intentions. Policymakers can prioritize populations with higher readiness for behavioral change while developing incentive mechanisms for more resistant segments. Additionally, in mental health and community resilience research, the model can uncover patterns of stress, social cohesion, and vulnerability, thereby informing resource allocation and crisis response.

In sum, the hybrid neural-swarm clustering approach is not just a technical advancement.it offers a computational lens for decoding complex social and psychological realities. By converting quantitative metrics into qualitative behavioral insights, it bridges artificial intelligence with social and environmental understanding, making it a valuable tool for both research and action.

3.1. Experimental Setups

In this section, a multitude of datasets (i.e. Ionosphere, Wine, Seeds, and G-set) are utilized in these experiments. then described in more detail in **Table 2**. For these datasets, we not only normalize these datasets, but apply PCA Wold, Esbensen, and Geladi [21] to preserve 90% of the information in the datasets. In addition, the Gaussian kernel function was used in the experiments to measure whether the samples are similar in the experiments. Our method is compared with the algorithm including Kmeans [22], Kmeans++ [15], and Isodata [23]. The parameters of our algorithm are listed in **Table 3**. The first parameter σ is Gaussian kernel function parameter; N and M denote the parameters of PSO.

Table 2. The datasets employed in the experiments.

Name	Instances	Class	Dimensions
lonosphere	351	2	34
Wine	96	2	11
Seeds	140	2	7
G-set	2048	2	16

Table 3. The parameter used in the experiments.

Datasets	σ	N	M
lonosphere	351	2	34
Wine	96	2	11
Seeds	140	2	7
G-set	2048	2	16

3.2. Numerical experiment

Parameters are set to the following values: $M = 500$, $N = 32$, $\sigma = 2$, $n = [100, 200, 300, 400]$ and $p = 2$. **Figure 1** shows the numerical experiment in the 2D domain. To vividly show clustering results, we plot the resulting figure. **Figures 2** and **3** depict the convergent value of the objective function DHN-PSO in the inner-loop and the outer-loop, respectively.

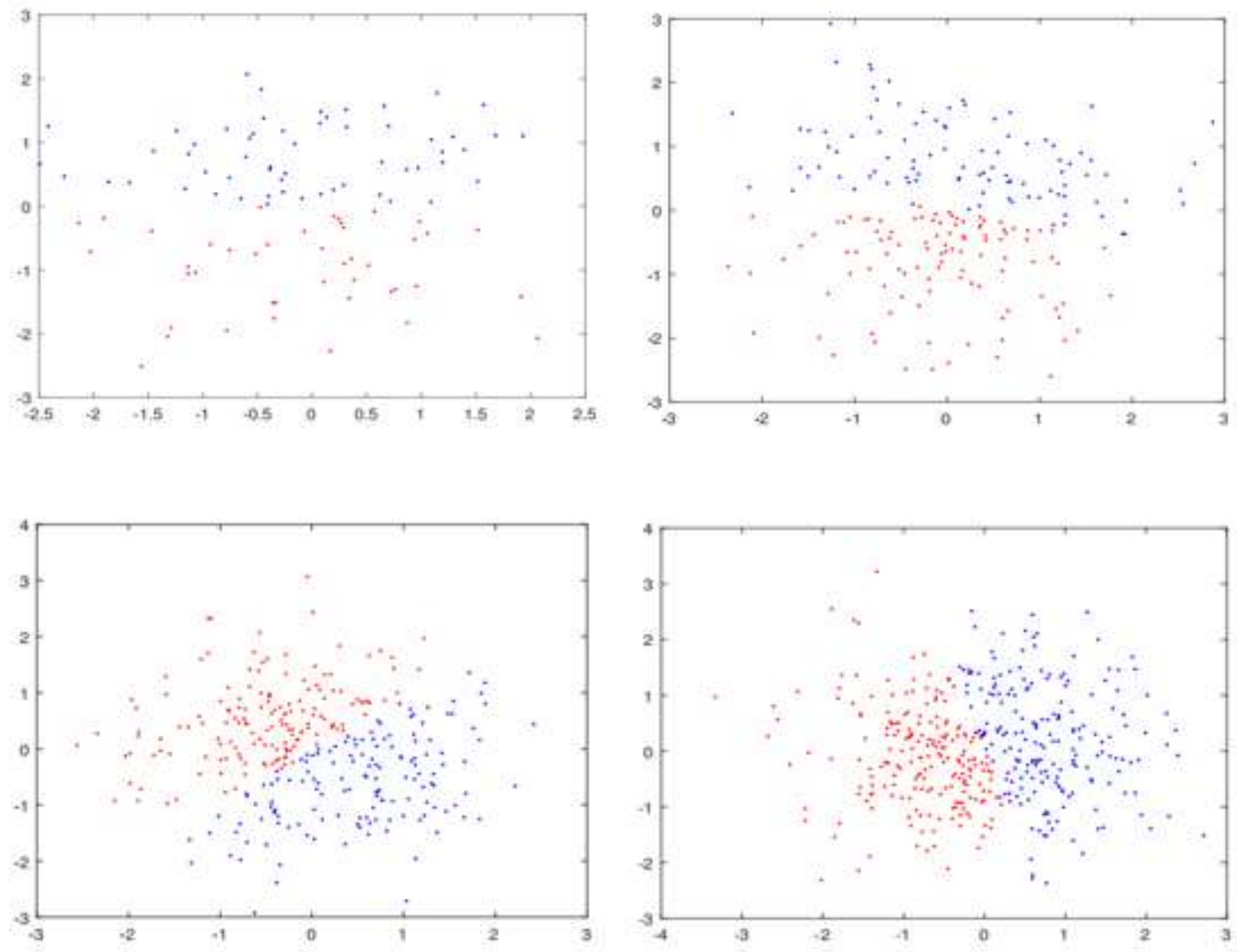

Figure 1. The clustering results of the DHN-PSO on the 100 to 400 points

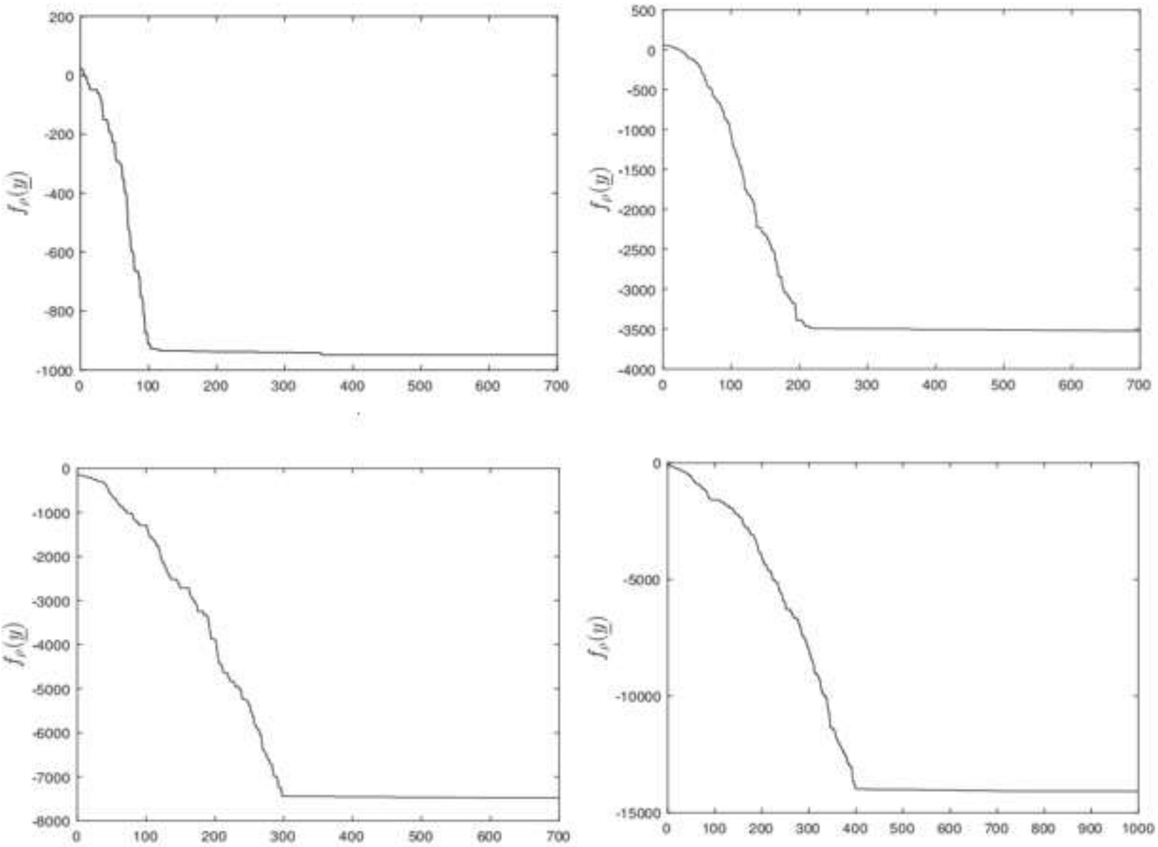

Figure 2. The convergent behaviors (inner-loop) of the DHN-PSO on the 100 to 400 points

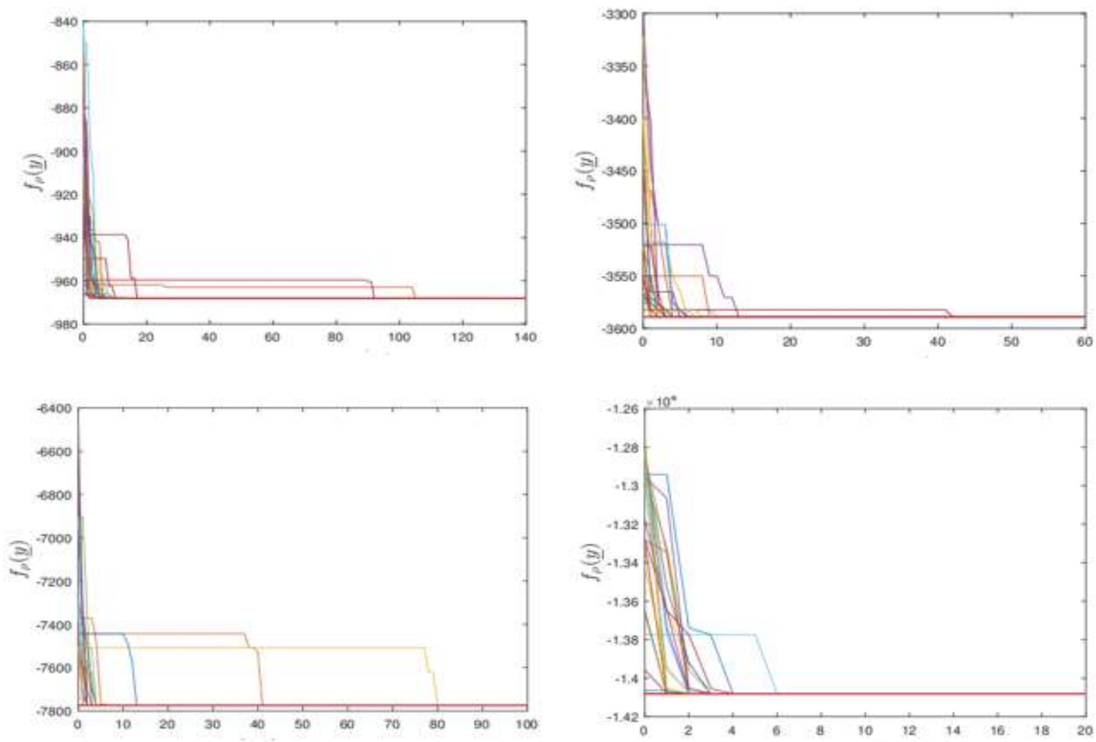

Figure 3. The convergent behaviors (outer-loop) of the DHN-PSO on the 100 to 400 points

3.3. Convergence study

Figure 4 illustrates the convergent behaviors of the objective function computed with DHN in the inner-loop of our algorithm on datasets Ionosphere, Wine, Seeds and G-set. **Figure 5** depicts the convergent behaviors of the outer-loop of our algorithm on datasets Ionosphere, Wine, Seeds and G-set. These experiment results show that outer-loop iterations are less than inner-loop.

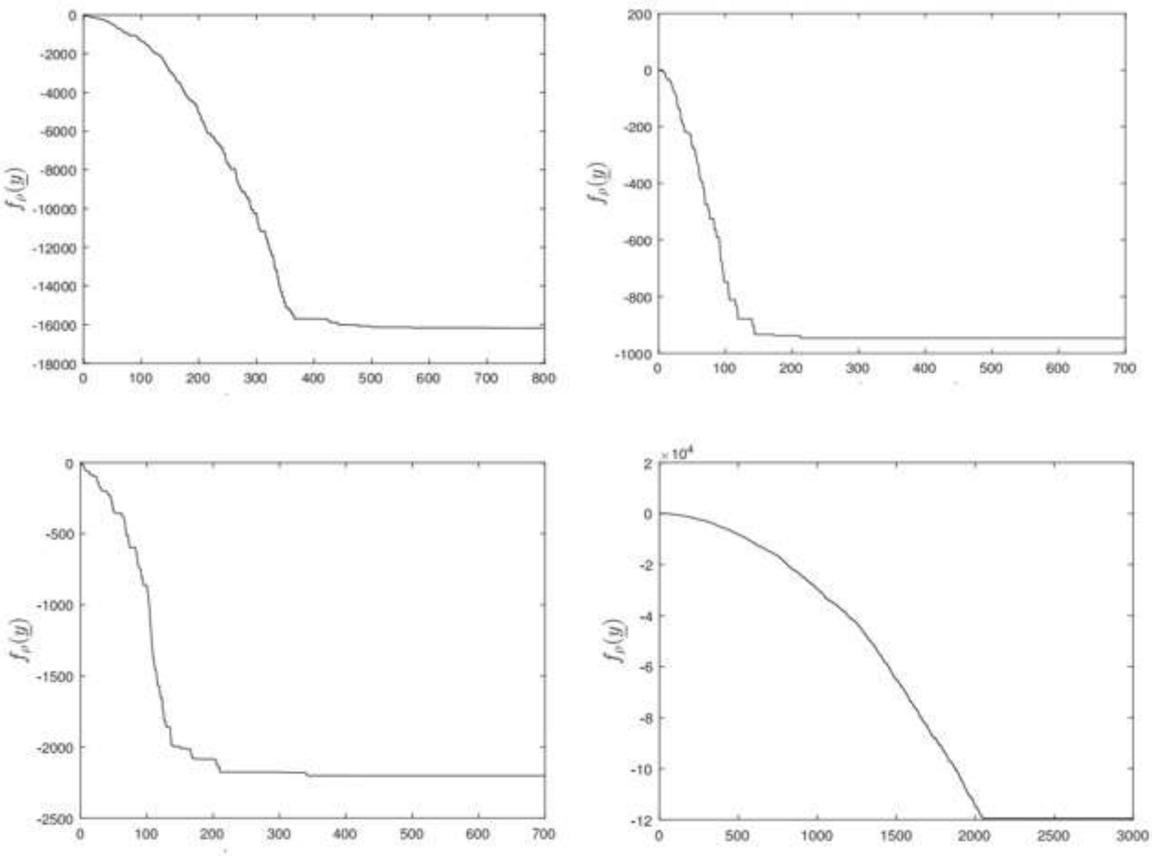

Figure 4. The convergent behaviors (inner-loop) of the DHN algorithm

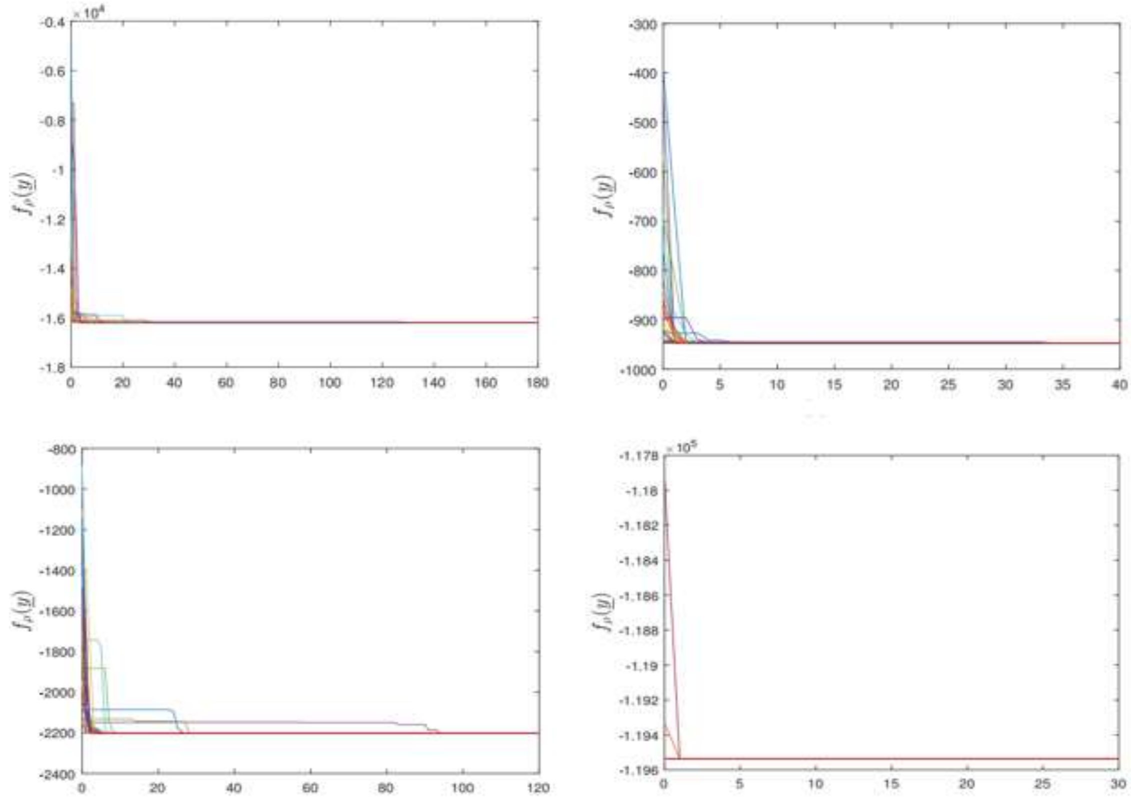

Figure 5. The convergent behaviors (outer-loop) of the DHN_PSO of four datasets

4. Discussion

4.1. Evaluation indices

The clustering performances of the proposed method are evaluated by using not only external performance indices: Normalized Mutual Information (NMI) [24], Accuracy (AC) [24] and Adjusted Rand index (ARI) [25], but also seven internal indices [26] including Ball-Hall Index, Calinski-Harabasz Index, Davies-Bouldin Index, Dunn Index, Ray-Turi Index, Standard Deviation-based Dissimilarity Index, Xie-Beni Index.

4.2. Indices discussion

Table 4 shows a comparative analysis of our method against Kmeans, Kmeans++, and Isodata[27] using ten clustering performance measures. The table showcases superior, inferior, average, and standard deviation performance metrics across the four datasets. Our method consistently excels over the three alternatives, as evident from the clustering outcomes and evaluation indices. In light of the results, we draw the following conclusions:

Our method consistently surpasses other approaches in the Ionosphere and G-set datasets, regardless of the chosen evaluation criteria. In the case of the Wine and Seeds datasets, while certain indices show other methods performing better, our method exhibits overall superiority over the alternatives.

where Ball-Hall Index, Calinski-Harabasz Index,Davies-Bouldin Index, Dunn Index, Ray-Turi Index, Standard Deviation-based Dissimilarity Index, Xie-Beni Index denote BH, CH, DB, Dunn, RT, SD and XB, respectively.

4.3. Expanded discussion: Ethical, environmental, and psychological dimensions

4.3.1. Ethical considerations in social data clustering

While clustering offers powerful capabilities for discovering latent patterns in social datasets such as grouping individuals by online behavior[28-30], political stance[31-34], or environmental attitudes[35-41].it also raises critical ethical concerns. Clustering social or psychological data risks:

Invasion of privacy through inferred attributes,

Labeling biases that might reinforce stereotypes,

Unintended consequences when clusters are used for targeting or exclusion such as in advertising or surveillance.

Therefore, transparency, consent, and algorithmic accountability are essential when applying this model to human-centered data. Ethical review procedures must be integrated into the deployment pipeline, especially when analyzing sensitive behavioral, demographic, or psychological attributes.

4.3.2. Applications for environmental policy-making

The ability to uncover latent behavioral segments through clustering can inform targeted environmental policy interventions. For example:

Clustering survey responses can reveal attitudinal profiles like “passive recyclers,” “climate skeptics,” or “green innovators,” enabling more nuanced public engagement strategies.

Geospatial clustering of environmental concern can identify ecologically vulnerable or disengaged communities, helping to allocate resources and shape localized education campaigns.

In this context, the use of clustering methods supports evidence-based policy-making that reflects the behavioral diversity of a population, rather than assuming homogeneous public perceptions.

Understanding Social Movements and Environmental Activism

Clustering methods can also be applied to analyze the structure of social movements, such as mapping activist networks, uncovering group formation dynamics, or tracking shifts in public opinion over time. On platforms like Twitter or Weibo, clustering posts and users based on sentiment, topic, and interaction networks can reveal the diffusion patterns of digital activism, the emergence of echo chambers, or the spread of misinformation.

By doing so, researchers can study collective behavior phenomena, such as coordination, mobilization, or resistance, within both environmental and social justice domains.

4.3.3. Limitations from a psychological perspective

Despite its technical strength, the clustering model does face psychological limitations:

Clusters may oversimplify complex human behaviors, reducing multifaceted attitudes to rigid groupings.

Contextual nuances, such as intent, emotion, or evolving identity, are often not fully captured in static datasets.

Psychological validity depends on culturally sensitive feature selection, appropriate annotation, and human expert validation.

Thus, while clustering can uncover general trends, interpretation must be cautious and grounded in psychological theory, such as cognitive dissonance, social identity, or motivated reasoning, to avoid misrepresenting the intricacies of human behavior.

Overall, the integration of DHNN and PSO not only enhances the computational robustness of clustering but also opens avenues for socially meaningful applications. Whether used for modeling digital activism, understanding environmental attitudes, or informing behavioral policy, this approach bridges algorithmic innovation with real-world relevance. However, to fully realize its potential, further research must continue to refine ethical safeguards, improve psychological interpretability, and adapt models to culturally and contextually specific data environments.

Table 4. Comparison of clustering of DHN-PSO, Kmeans, Kmeans++, and Isodata on four datasets.

Ionosphere	**DHN-PSO**	**Kmeans**	**Kmeans++**	**Isodata**
BH	18.73/18.73/18.73±0	1.87/1.87/1.87±0	1.87/1.87/1.87±0	2.98/1.69/2.30/±0.46
CH	529.77/529.77/529.77±0	115.03/115.03/115.03±0	115.04/115.04/115.04±0	114.46/15.08/53.24/±38.59
DB	0.75/0.75/0.75±0	1.54/1.54/1.54±0	1.54/1.54/1.54±0	3.22/1.54/2.20±0.64
Dunn	0.14/0.14/0.14±0	0.07/0.07/0.07±0	0.07/0.07/0.07±0	0.07/0.03/0.5±0.02
RT	0.16/0.16/0.16±0	0.75/0.75/0.75±0	0.75/0.75/0.75±0	2.64/0.75/1.41±0.647
SD	0.18/0.18/0.18±0	1.29/1.29/1.29±0	1.29/1.29/1.29±0	2.16/1.29/1.60±0.29
XB	6.43/6.43/6.43±0	14.80/14.80/14.80±0	14.80/14.80/14.80±0	115.74/14.82/54.06±36.105
NMI	20.56/20.56/20.56±0	12.64/12.64/12.64±0	13.12/112.64/12.69±0.15	13.49/0.01/3.91±4.973
AC	72.93/72.93/72.93±0	70.94/70.94/70.94±0	71.22/70.94/70.96±0.09	72.08/53.84/61.42±6.88
ARI	20.78/20.78/20.78±0	17.27/17.27/17.27±0	17.76/17.27/17.32±0.15	16.32/-3.46/4.42±7.03

Wine	**DHN-PSO**	**Kmeans**	**Kmeans++**	**Isodata**
BH	3.23/3.23/3.23±0	12.68/4.90/10.02±3.50	11.76/9.30/10.56±0.92	11.74/4.90/9.83±2.33
CH	133.95/133.95/133.95±0	12.20/6.83/9.54±2.39	12.94/7.27/10.77±2.01	8.93/5.83/7.33±1.14
DB	0.79/0.79/0.79±0	2.63/0.33/1.61±0.92	2.79/2.05/2.45±0.23	3.31/0.33/2.14±1.01
Dunn	0.42/0.42/0.42±0	0.68/0.10/0.31±0.25	0.14/0.07/0.11±0.02	0.68/0.05/0.20±0.22
RT	0.175/0.175/0.175±0	1.599/0.142/0.810±0.55	2.290/1.112/1.710±0.45	3.073/0.142/1.570±1.10
SD	0.46/0.46/0.46±0	0.82/0.24/0.54±0.23	0.97/0.69/0.84±0.11	1.12/0.24/0.75±0.32
XB	0.78/0.78/0.78±0	6.91/0.20/2.65±2.76	13.29/4.00/7.07±4.26	24.31/0.20/9.09±9.60
NMI	62.41/62.41/62.41±0	41.88/1.05/11.91±11.48	41.88/2.01/19.25±13.69	11.29/0.50/5.72±4.43
AC	92.70/92.70/92.70±0	81.25/51.04/61.45±8.54	81.25/58.33/68.12±7.64	63.54/51.04/57.29±3.86
ARI	72.67/72.67/72.67±0	38.50/0.002/7.50±11.30	38.50/1.74/14.60±12.89	6.71/-0.36/2.15±2.30

Seeds	DHN-PSO	Kmeans	Kmeans++	Isodata
BH	17.30/17.30/17.30±0	0.01/0.01/0.01±0	0.01/0.01/0.01±0	0.01/0.01/0.01±0
CH	69.81/69.81/69.81±0	247.09/246.79/246.87±0.14	247.09/246.79/246.87±0.14	228.53/17.72/116.96±74.21
DB	1.39/1.39/1.39±0	0.66/0.66/0.66±0	0.66/0.66/0.66±0	2.22/0.69/0.99±0.54
Dunn	0.22/0.22/0.22±0	0.10/0.05/0.06±0.02	0.10/0.05/0.06±0.02	0.07/0.03/0.04±0.01
RT	0.49/0.49/0.49±0	0.13/0.13/0.13±0	0.13/0.13/0.13±0	1.65/0.15/0.43±0.5
SD	0.33/0.33/0.33±0	8.73/8.70/8.72±0.01	8.73/8.70/8.72±0.01	19.21/8.89/10.70±3.77
XB	3.75/3.75/3.75±0	30.09/7.71/23.70±10.92	30.09/7.71/23.70±10.92	69.46/12.07/37.06±22.13
NMI	68.40/68.40/68.40±0	63.82/62.24/62.87±0.81	63.82/62.24/62.71±0.76	63.82/6.40/30.25±20.50
AC	94.28/94.28/94.28±0	92.14/92.14/92.14±0	92.14/92.14/92.14±0	92.14/58.57/74.71±12.93
ARI	78.29/78.29/78.29±0	70.83/70.83/70.83±0	70.83/70.83/70.83±0	70.83/2.71/30.11±26.67

G-set	DHN-PSO	Kmeans	Kmeans++	Isodata
BH	0.19/0.19/0.19±0	0.15/0.15/0.15±0	0.15/0.15/0.15±0	0.61/0.15/0.40±0.21
CH	72355.81/72355.813/72355.81±0	6383.41/6383.41/6383.41±0	6383.41/6383.41/6383.41±0	6383.41/25.47/2269.21±2968.08
DB	0.15/0.15/0.15±0	0.55/0.55/0.55±0	0.55/0.55/0.55±0	8.55/0.55/3.24±3.26
Dunn	0.91/0.91/0.91±0	0.72/0.72/0.72±0	0.72/0.72/0.72±0	0.72/0.08/0.26±0.31
RT	0.01/0.01/0.01±0	0.08/0.08/0.08±0	0.08/0.08/0.08±0	18.60/0.08/5.08±7.36
SD	0.38/0.38/0.38±0	1.45/1.45/1.45±0	1.45/1.45/1.45±0	10.96/1.45/4.58±3.87
XB	0.03/0.03/0.03±0	0.22/0.22/0.22±0	0.22/0.22/0.22±0	21.10/0.22/12.42±9.47
NMI	100.00/100.00/100.00±0	100.00/100.00/100.00±0	100.00/100.00/100.00±0	100.00/0.15/51.11±46.32
AC	100.00/100.00/100.00±0	100.00/100.00/100.00±0	100.00/100.00/100.00±0	100.00/52.19/79.81±20.56
ARI	100.00/100.00/100.00±0	100.00/100.00/100.00±0	100.00/100.00/100.00±0	100.00/0.14/50.76±46.98

Acknowledgement

This workwork was partly supported by the Youth Project of Science and Technology Research Program of Chongqing Municipal Education Commission Open Fund Project Grant No. KJQN202301222, NO.KJQN202401235 and KJQN202401237.in part by Opening fund of Chongqing Engineering Research Center of Internet of Things and Intelligent Control Technology No.zhlv-20221031

Conflict of interest

The authors declare no conflict of interest.

References

1. Jain A K, Murty M N, Flynn P J. Data clustering: a review[J]. ACM computing surveys (CSUR), 1999, 31(3): 264-323.
2. Zhang Y, Liu X, Liu J, et al. Fusion multiple kernel k-means[C]//Proceedings of the AAAI Conference on Artificial Intelligence. 2022, 36(8): 9109-9117.

3. Lai Y, He S, Lin Z, et al. An adaptive robust semi-supervised clustering framework using weighted consensus of random $ k $ k-means ensemble[J]. IEEE Transactions on Knowledge and data engineering, 2019, 33(5): 1877-1890.
4. Liu H, Li J, Wu Y, et al. Clustering With Outlier Removal[J]. IEEE Transactions on Knowledge & Data Engineering, 2021, 33(06): 2369-2379.
5. Yamada Y, Masuyama N, Amako N, et al. Divisive hierarchical clustering based on adaptive resonance theory[C]//2020 International Symposium on Community-centric Systems (CcS). IEEE, 2020: 1-6.
6. Taha K. Automatic academic advisor[C]//8th International conference on collaborative computing: networking, applications and worksharing (CollaborateCom). IEEE, 2012: 262-268.
7. Zeng X, Wang W, Chen C, et al. A Consensus Community-Based Particle Swarm Optimization for Dynamic Community Detection[J]. IEEE transactions on cybernetics, 2020, 50(6): 2502-2513.
8. Miraftabzadeh S M, Colombo C G, Longo M, et al. K-means and alternative clustering methods in modern power systems[J]. Ieee Access, 2023, 11: 119596-119633.
9. Meng J, Abed A M, Elsehrawy M G, et al. Nano-integrating green and low-carbon concepts into ideological and political education in higher education institutions through K-means clustering[J]. Heliyon, 2024, 10(10).
10. Al-Sabbagh A, Hamze K, Khan S, et al. An Enhanced K-Means Clustering Algorithm for Phishing Attack Detections[J]. Electronics, 2024, 13(18): 3677.
11. Waheed A, Akram M U, Khalid S, et al. Hybrid features and mediods classification based robust segmentation of blood vessels[J]. Journal of medical systems, 2015, 39(10): 128.
12. Dai J, Fang J, Guo Z, et al. Research on ecological restoration assessment and eco-economic development of sea area by introducing the K-means clustering algorithm[J]. Environmental Science and Pollution Research, 2023, 30(56): 118318-118331.
13. Meng Y, Liu X. Application of K-means algorithm based on ant clustering algorithm in macroscopic planning of highway transportation hub[C]//2007 First IEEE International Symposium on Information Technologies and Applications in Education. IEEE, 2007: 483-488.
14. Mutambik I. Assessing Urban Vulnerability to Emergencies: A Spatiotemporal Approach Using K-Means Clustering[J]. Land, 2024, 13(11): 1744.
15. Tank D, Hopfield J. Simple'neural'optimization networks: An A/D converter, signal decision circuit, and a linear programming circuit[J]. IEEE transactions on circuits and systems, 2003, 33(5): 533-541.
16. Neshat M, Adeli A, Sepidnam G, et al. A review of artificial fish swarm optimization methods and applications[J]. International Journal on Smart Sensing and Intelligent Systems, 2012, 5(1): 108.
17. Meng Y, Liu X. Application of K-means algorithm based on ant clustering algorithm in macroscopic planning of highway transportation hub[C]//2007 First IEEE International Symposium on Information Technologies and Applications in Education. IEEE, 2007: 483-488.
18. Kennedy J, Eberhart R C. A discrete binary version of the particle swarm algorithm[C]//1997 IEEE International conference on systems, man, and cybernetics. Computational cybernetics and simulation. ieee, 1997, 5: 4104-4108.
19. Shi J, Malik J. Normalized cuts and image segmentation[J]. IEEE Transactions on pattern analysis and machine intelligence, 2000, 22(8): 888-905.
20. Rana S, Garg R. Application of hierarchical clustering algorithm to evaluate students performance of an institute[C]//2016 Second international conference on computational intelligence & communication technology (CICT). IEEE, 2016: 692-697.
21. Wold S, Esbensen K, Geladi P. Principal component analysis[J]. Chemometrics and intelligent laboratory systems, 1987, 2(1-3): 37-52.
22. Sinaga K P, Yang M S. Unsupervised K-means clustering algorithm[J]. IEEE access, 2020, 8: 80716-80727.
23. Ball G H, Hall D J, ISODATA A. A Novel Method of Data Analysis and Pattern Classification (Menlo Park, CA: Stanford Research Institute)[J]. Three decades of land use variations in Mexico City, 1965, 135.
24. Cai D, He X, Han J. Document clustering using locality preserving indexing[J]. IEEE transactions on knowledge and data engineering, 2005, 17(12): 1624-1637.
25. Santos J M, Embrechts M. On the use of the adjusted rand index as a metric for evaluating supervised classification[C]//International conference on artificial neural networks. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009: 175-184.
26. Desgraupes B. Clustering indices[J]. University of Paris Ouest-Lab Modal'X, 2013, 1(1): 34.
27. Dunn J C. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters[J]. 1973.
28. Handcock M S, Raftery A E, Tantrum J M. Model-based clustering for social networks[J]. Journal of the Royal Statistical Society Series A: Statistics in Society, 2007, 170(2): 301-354.
29. Zhurakovskyi B, Averichev I, Shakhmatov I. Using the Latest Methods of Cluster Analysis to Identify Similar Profiles in Leading Social Networks[C]//IT&I Workshops. 2023: 116-126.

30. Ren Y, Pu J, Yang Z, et al. Deep clustering: A comprehensive survey[J]. IEEE transactions on neural networks and learning systems, 2024, 36(4): 5858-5878.
31. Ahlquist J S, Breunig C. Model-based clustering and typologies in the social sciences[J]. Political Analysis, 2012, 20(1): 92-112.
32. Schrodt P A, Gerner D J. Cluster-based early warning indicators for political change in the contemporary levant[J]. American Political Science Review, 2000, 94(4): 803-817.
33. Fleishman J A. Types of political attitude structure: Results of a cluster analysis[J]. Public Opinion Quarterly, 1986, 50(3): 371-386.
34. Kuo K H, Wang M H, Kao H Y, et al. Advancing stance detection of political fan pages: A multimodal approach[C]//Companion Proceedings of the ACM Web Conference 2024. 2024: 702-705.
35. Chen M, Wang P, Chen Q, et al. A clustering algorithm for sample data based on environmental pollution characteristics[J]. Atmospheric Environment, 2015, 107: 194-203.
36. Noviandy T R, Hardi I, Zahriah Z, et al. Environmental and economic clustering of indonesian provinces: insights from K-Means analysis[J]. Leuser Journal of Environmental Studies, 2024, 2(1): 41-51.
37. D'Adamo I, Di Carlo C, Gastaldi M, et al. Economic performance, environmental protection and social progress: A cluster analysis comparison towards sustainable development[J]. Sustainability, 2024, 16(12): 5049.
38. Ramadhani F. Spatial Clustering Analysis of Stunting in North Sumatra Based on Environmental Factors Using K-Means Algorithm[J]. Data Science: Journal of Computing and Applied Informatics, 2025, 9(2): 18-25.
39. Bottazzi Schenone M, Vichi M. Clustering Hierarchical Disjoint Principal Component Analysis for environmental impact assessment of food products[M]//Innovation & Society: Statistics and Data Science for Evaluation and Quality. Cleup sc, 2025: 28-35.
40. Zhang N, Zhang Z, Tang J, et al. Similarity clustering and characteristics analysis of carbon emission pathways based on provincial data from China[J]. Environment, Development and Sustainability, 2025: 1-31.
41. Khelil M I, Laadjal M, Ouali M A, et al. A Clustering-Driven Strategy Utilizing DBSCAN for Detecting Outliers in Water Quality Data[J]. Communication science et technologie, 2025, 23(1): 1-12.