

Retraction

Retracted: Text Knowledge Acquisition Method of Collaborative Product Design Based on Genetic Algorithm

Journal of Mathematics

Received 19 December 2023; Accepted 19 December 2023; Published 20 December 2023

Copyright © 2023 Journal of Mathematics. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Manipulated or compromised peer review

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

- [1] C. Li and M. Yang, "Text Knowledge Acquisition Method of Collaborative Product Design Based on Genetic Algorithm," *Journal of Mathematics*, vol. 2022, Article ID 3661477, 9 pages, 2022.

Research Article

Text Knowledge Acquisition Method of Collaborative Product Design Based on Genetic Algorithm

Chongyuan Li ¹ and Mengmeng Yang^{1,2}

¹Haikou University of Economics, Haikou, Hainan 570000, China

²HaiNan Normal University, Hainan 570000, China

Correspondence should be addressed to Chongyuan Li; 171843237@masu.edu.cn

Received 23 December 2021; Revised 31 January 2022; Accepted 4 February 2022; Published 27 February 2022

Academic Editor: Naeem Jan

Copyright © 2022 Chongyuan Li and Mengmeng Yang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In order to overcome the problems of poor clustering effect and large error of text knowledge acquisition in traditional text knowledge acquisition methods, a new text knowledge acquisition method of collaborative product design based on genetic algorithm is proposed in this paper. The definition of collaborative product design text knowledge clustering is given. According to the operation process of the genetic algorithm, the chromosomes of clustered text are constructed and encoded and the initial population is obtained. The fitness function of clustering is constructed by the DB index evaluation method; the selection, crossover, and mutation operators in the genetic algorithm are determined; and the objective function of collaborative product design text knowledge clustering is constructed. After the text knowledge clustering is completed, the text knowledge data of collaborative product design are obtained in an all-around way by using the method of rough set and neural network. The experimental results show that compared with the traditional text knowledge acquisition methods, the clustering effect of the proposed method is better and the text knowledge error is reduced up to 0.02.

1. Introduction

The product design process is an important stage of the whole product development. Modern product design is becoming more and more complex. It is difficult for a single unit to have a design team with a professional configuration. Many product design units adopt the networked collaborative product design method to unite product designers scattered in local alliances to jointly complete product design [1]. In the process of collaborative product design, knowledge exchange and sharing are important to guarantee successful product design. At present, most of the commonly used knowledge acquisition methods are that product designers obtain the knowledge they need from the product design knowledge base of each unit through a query. It requires collaborative product designers to have a deep understanding of the product design knowledge base of cooperative units [2–4]. With the continuous increase of product design knowledge, the query takes longer and

longer, and it is even difficult to find the knowledge that meets the requirements of product design, which affects the efficiency of product design. Therefore, the traditional query method has been difficult to meet the requirements of networked product collaborative design.

In the process of collaborative product design, multidisciplinary design teams in distributed, remote, and dynamic environments involve many complex interactive tasks [5]. The practice results show that the sharing and reuse of design knowledge can reduce unnecessary repeated labor and shorten the product development cycle. In particular, the variant design of products can reuse the design experience and design knowledge accumulated by the design unit, so as to solve a large number of problems in many fields encountered in design [6]. Therefore, it is very important to study an effective text knowledge acquisition method for collaborative product design.

Reference [7] proposes a text knowledge acquisition method based on multigranularity. Multigranularity

cognitive ability is a common strategy for human beings to analyze complex data. As one of the complex data types, multisource data make data analysis complex because of its many data sources. Inspired by the idea of multigranularity, based on multisource information system and pessimistic decision-making strategy, the definition of multisource partition reduction set is proposed. The relationship between the multisource partition reduction set and partition reduction set is discussed, and the corresponding discrimination method of attribute characteristics is given. Finally, for a multisource decision information system, based on an optimistic decision strategy, multisource decision rules are proposed. Based on the multigranularity model, the multisource data analysis method proposed from a new perspective further enriches the method of knowledge acquisition. Reference [8] proposes a text knowledge acquisition method based on a fuzzy hyper network. Combined with the relevant knowledge of fuzzy rough set theory and hyper network, the fuzzy hyper network model adopts fuzzy equivalent relationship to replace the clear equivalent relationship in hyper network and improves the generation and evolution of hyperedges on this basis. According to the distribution of samples, the sample set is divided into three regions, namely positive domain, boundary domain, and negative domain. Samples in different regions generate super edges in different ways. According to the classification effect, the super edge set is also divided into three regions, and the super edges in different regions are replaced accordingly, so as to complete the acquisition of text knowledge. Reference [9] proposes a text knowledge acquisition method based on word meaning disambiguation of knowledge map, uses TF-IDF model to obtain the set of text feature words, uses the word meaning sequence relationship expressed by knowledge map to determine the unique semantics of polysemy in a specific semantic environment, completes the vectorial representation of text at the level of word meaning concept, and realizes text knowledge acquisition.

Based on the above research results, in order to further improve the clustering and acquisition effect of collaborative product design text knowledge, a collaborative product design text knowledge acquisition method based on a genetic algorithm is proposed. A genetic algorithm is used to optimize the text knowledge clustering algorithm to improve the effectiveness of clustering, so as to obtain more effective collaborative product design text knowledge.

The research is organized as follows: the introduction and literature review are presented in Section 1. Section 2 analyzes the text knowledge acquisition of collaborative product design based on genetic algorithm. Section 3 discusses the experimental verification of the proposed concepts. Finally, in Section 4, the research work is concluded.

2. Text Knowledge Acquisition of Collaborative Product Design Based on Genetic Algorithm

In this section, we described the text knowledge clustering of collaborative product design based on genetic algorithm and collaborative product design text knowledge acquisition with detail.

2.1. Text Knowledge Clustering of Collaborative Product Design Based on Genetic Algorithm. Collaborative product design text knowledge clustering is a fully automatic process of grouping knowledge text sets. It is a typical unsupervised machine learning process. The goal of knowledge text clustering is to find a set of such classes. The similarity between classes is as small as possible, and the similarity within classes is as large as possible [10–13]. As an unsupervised machine learning method, clustering does not need a training process and manual labeling of collaborative product design text knowledge in advance. Therefore, clustering technology is very flexible and has a high automatic processing ability. At present, it has become an important means to effectively organize, summarize, and navigate collaborative product design text knowledge.

In order to improve the effect of text knowledge clustering in collaborative product design, a genetic algorithm is used to optimize a k-means algorithm for text knowledge clustering. The specific process is as follows.

Collaborative product design text knowledge clustering is based on the famous clustering hypothesis: the similarity of similar text knowledge is large, while the similarity of different text knowledge is small [14]. Text knowledge clustering is to divide the text knowledge set into several clusters without any predictive information. It is required that the similarity of text knowledge content in the same cluster should be as large as possible, and the similarity of objects between different clusters should be as small as possible.

Definition 1. Collaborative product design text knowledge clustering is to divide a given text knowledge set $F = \{f_1, f_2, \dots, f_m\}$ to obtain a class set $V = \{v_1, v_2, \dots, v_l\}$, where $v_k \subset F (k = 1, 2, \dots, l)$, making $\forall f_o (f_o \in F)$, $\exists v_l (v_l \in V)$, and $f_o \in v_l$, and minimizing the cost function $g(V)$.

The grouping in clustering is also called a cluster. The definition of a cluster is as follows: (1) the set of similar objects, and the objects in different clusters are not similar; (2) the distance between two objects in the cluster is less than the distance between any object in the cluster and any object outside the cluster.

Clustering is similar to classification in that it groups data. However, different from classification, the input set of clustering analysis is a group of unidentified records. Therefore, this is an unsupervised learning process; that is, the clustering algorithm does not need supervision and does not need to provide training data. It tends to the natural division of data. A classification algorithm is a supervised learning process, which needs to train the labeled data set.

Genetic algorithm is similar to natural evolution. It finds a good chromosome to solve the problem by acting on the genes on the chromosome. Similar to nature, the genetic algorithm knows nothing about the problem itself. All it needs is to evaluate each chromosome generated by the algorithm and select chromosomes based on fitness value so that the chromosomes with good adaptability have more reproductive opportunities. In a genetic algorithm, several numerical codes of the problem are generated randomly,

that is, chromosomes, to form the initial population; A numerical evaluation is given to each individual through the fitness function, the chromosomes with low fitness are eliminated, and the individuals with high fitness are selected to participate in the genetic operation. The individual set after genetic operation forms the next generation of the new population, the next round of evolution of the new species.

The biggest difference between a genetic algorithm and a traditional algorithm is that the calculation does not terminate automatically. The termination conditions are often artificially limiting the evolutionary algebra or setting the termination conditions by controlling the consistency of evolutionary results.

The genetic algorithm differs from other search methods in that

- (1) The encoded variable value is utilized instead of the variable itself when employing the genetic algorithm to search
- (2) The search method for a genetic algorithm is to iterate from one set of points to the next set of points and to iterate from one point to another, unlike most other search algorithms, but search by population
- (3) Rather than using a deterministic operation process, the genetic algorithm employs a random operation process
- (4) The qualities of the search space (such as connectiveness, convexity) are unimportant to the genetic algorithm, which only requires the value of each point of the objective function and no additional auxiliary information

2.1.1. Chromosome Construction and Coding of Clustering Problems. According to the objective function $K(I, W)$ of collaborative product design text knowledge clustering, the ultimate goal of collaborative product design text knowledge clustering is to obtain l partition matrices I of sample set C and the prototype matrix W of clustering, and I and W are related. The partition matrix I can also be encoded, and the prototype matrix W can also be encoded.

Let M samples be divided into l classes. $M = (s_1, s_2, \dots, s_m)$ represents the chromosome structure of the genetic algorithm, s is an $1 * N$ -dimensional row vector, s_o represents the o -th gene, but $s_o \in \{1, 2, \dots, l\}$, $i = 1, 2, \dots, M$. When $s_o = v$, it means that the o -th sample belongs to class v . Generally, standard genetic operations can be used to support this expression. Therefore, for any scheme that divides the sample M into l clusters, it is its corresponding coding matrix:

$$I = (\mu_{ok})l * M, \mu_{ok} = \begin{cases} 1, & \text{if } o \in k, \\ 0, & \text{else.} \end{cases} \quad (1)$$

2.1.2. Initial Population. The initial population required by the genetic algorithm should have diversity. Therefore, when operating a genetic algorithm, first determine the population

size M and then randomly generate M random numbers. According to the o -th random number l_o , a chromosome with length l_o is generated. The chromosome represents the centroid set containing l_o genes, and each gene in the chromosome represents an initial cluster center. In this way, the initial population with the population size of M and different chromosome lengths is obtained.

2.1.3. Fitness Function. The genetic algorithm basically does not use external information in the process of evolutionary search, and only takes the fitness function as the basis for optimization. The fitness function in the genetic algorithm is used to evaluate the fitness of individuals and distinguish the advantages and disadvantages of individuals in the population. The higher the fitness, the greater the probability of being inherited, and the better the clustering effect [15]. Therefore, the selection of fitness function is very important, which directly affects the convergence speed of the genetic algorithm and the ability to find the optimal solution. DB index evaluation method is adopted.

The DB evaluation method is described as follows:

$$D_i = \frac{1}{N_i} \sum_{G \in C_i} |G - H_i|, H_i = \frac{1}{N_i} \sum_{G \in C_i} G. \quad (2)$$

Here, H_i represents the center of the set, $|G - H_i|$ represents the distance from the collaborative product design text knowledge vector P to the center H_i , and D_i represents the average cohesion of the cluster set.

$$f_{ij} = |H_i - H_j|. \quad (3)$$

Here, f_{ij} represents the distance between two clustering centers.

$$S_i = \max \left\{ \frac{d_i + d_j}{f_{ij}} \right\} i \neq j, \quad (4)$$

$$DB = \frac{1}{Z} \sum_{i=1}^Z S_i.$$

Because the smaller the DB is, the better the clustering effect is, the reciprocal $1/DB$ of the DB can be used as the fitness function, that is:

$$F = \frac{1}{DB}. \quad (5)$$

From the above calculation, it can be seen that the smaller the DB, the greater the fitness value F , and the greater the probability of individual inheritance; the smaller the DB, the better the clustering effect. Therefore, it shows that the fitness value F is directly proportional to the clustering effect.

2.1.4. Genetic Operator. The main task of genetic operation is to operate the individuals of the population according to their degree of adaptation to the environment, so as to realize the evolutionary process of survival of the fittest. From the perspective of optimization search, genetic operations can optimize

the problem from generation to generation, constantly approaching the optimal solution.

The genetic algorithm includes three basic operators: selection, crossover, and mutation.

- (1) *Selection Operator.* In the process of biological evolution, species with strong adaptability to the living environment will have more opportunities to inherit to the next generation; the probability of inheritance to the next generation is relatively small. The selection operation in genetic algorithm embodies the principle of "survival of the fittest": the higher the fitness, the higher the probability of participating in offspring reproduction [16–18]. A selection operation is a genetic operation used to determine how to select which individuals from a parental population to inherit to a descendant population. It is based on the evaluation of individual fitness: individuals with higher fitness have a higher probability of genes being inherited into the offspring population; in individuals with low fitness, the probability of genes being inherited into the offspring population is small.
- (2) *Crossover Operator.* Crossover operation, also known as recombination, is the most important genetic operation in genetic algorithm. The crossover operation can not only maintain the excellent characteristics of the parent population to a certain extent but also make the algorithm explore the new gene space, so as to maintain the individual diversity in the new population. Crossover operation in genetic algorithm means that two paired individuals exchange some genes with each other in some way, so as to form two new individuals. The crossover operation is carried out according to a certain probability (called crossover probability). Crossover operation is an important feature that distinguishes genetic algorithm from other evolutionary operations. It plays a key role in genetic algorithm and is the main method to generate new individuals. The crossover operator simulates the mutation of natural organisms and embodies the idea of information exchange. The higher the crossover probability, the faster the convergence to the global optimal solution, but it may also lead to premature convergence; the crossover probability is too low, which may cause the search to stagnate, generally 0.4–0.9.
- (3) *Mutation Operator.* Mutation operation is to imitate the variation link in the process of biological inheritance and evolution. It refers to replacing the gene value of some gene positions in the individual chromosome coding string with other alleles of the gene position, so as to form a new individual. Variation takes place with a very small probability (called variation probability). It is a random algorithm, and combined with selection and crossover operators, it can avoid the loss of some information caused by selection and crossover operations and ensure the

effectiveness of genetic algorithm. Mutation is an auxiliary method to generate new individuals, but it is also an essential step of genetic algorithm, because it determines the local search ability of genetic algorithm. Mutation operation and crossover operation cooperate with each other to complete the global search and local search of the search space so that the genetic algorithm can complete the optimization process of the optimization problem with good search performance. It can increase the diversity of groups. If the mutation probability is too large, the genetic algorithm may degenerate into a random search. If the mutation probability is too small, it may not be able to produce new genes. Generally, it is 0.001–0.1.

2.1.5. Construction of Clustering Objective Function. In this study, the problem of clustering analysis of collaborative product design text knowledge is described from a mathematical point of view. Let $C = \{C_1, C_2, \dots, C_m\}$ be the whole of m objects to be clustered. Each C_k ($k = 1, \dots, m$) objects in C are often described by a finite number of attributes S_o ($o = 1, \dots, n$). Each attribute value depicts a feature of C_k , which can be expressed as $C_k = [c_{k,1}, c_{k,2}, \dots, c_{k,n}]$, and $C_k = [c_{k,1}, c_{k,2}, \dots, c_{k,n}]$ is the feature vector of C_k . Cluster analysis is to analyze the similarity of multicorresponding eigenvectors of m objects in all C objects, divide C_1, C_2, \dots, C_m into multiple disjoint subsets V_1, V_2, \dots, V_l according to the distance relationship between each object, and meet the following conditions:

$$\begin{aligned} V_1 \cup V_2 \cup \dots \cup V_k &= C, \\ V_o \cap V_k &= \Phi, \\ V_o &\neq \Phi, \\ V_o &\neq C, \\ 1 &\leq o \neq k \leq l. \end{aligned} \quad (6)$$

The membership relationship of m samples C_o to the subset can be expressed as follows:

$$\mu_{v_i}(C_k) = \mu_{ok} = \begin{cases} 1, & C_k \in V_o, \\ 0, & C_k \notin V_o. \end{cases} \quad (7)$$

The membership function must meet $\mu_{ok} \in RI$; that is, each sample can only belong to a certain class, and each subclass is required to be nonempty. For l partition space Q_l of C , there are

$$\left\{ \begin{aligned} Q_l &= \{I \in T^{lm} \mid I = [\mu_{ok}]_{l \times m}; \mu_{ok} \in \{0, 1\}\} \\ \sum_{o=1}^l \mu_{ok} &= 1, \forall k; 0 < \sum_{k=1}^n \mu_{ok} < n, \forall o \end{aligned} \right\}. \quad (8)$$

Let $T = \{T_1, T_2, \dots, T_l\}$ represent l clustering models, and then, the objective function of clustering analysis can be expressed as follows:

$$\left\{ \begin{aligned} K_1(I, W) &= \sum_{o=1}^l \left(\sum_{C_k \in V_o} (f_{ok})^2 \right), \\ \text{s.t. } I &\in Q_l. \end{aligned} \right. \quad (9)$$

Here, $f_{ok} = f(C_k, W_o)$ represents the dissimilarity measure between sample C_k and prototype V_o of W_o in the o -th type V_o , and $K_1(I, W)$ represents the sum of squares of errors between various samples and their prototypes. Using μ_{ok} , $K_1(I, W)$ can also be expressed as follows:

$$\begin{cases} K_1(I, W) = \sum_{k=1}^m \sum_{o=1}^l \mu_{ok} (f_{ok})^2, \\ \text{s.t. } I \in Q_l, \end{cases} \quad (10)$$

where $K_1(I, W)$ is also called clustering objective function. The clustering criterion is to find the best group pair (I, W) so that when the constraint $\mu_{ok} \in Q_l$ is satisfied, $K_1(I, W)$ is the minimum, that is, $\min\{K_1(I, W)\}$.

By solving the objective function of clustering, the clustering results of collaborative product design text knowledge can be obtained for subsequent text knowledge acquisition.

The process of collaborative product design text knowledge clustering based on genetic algorithm is shown in Figure 1.

2.2. Collaborative Product Design Text Knowledge Acquisition. Based on the above clustering results of collaborative product design text knowledge, text knowledge is obtained according to rough set theory.

Rough set theory gives the dependency relationship between conditional attribute set and decision attributes; that is, the mapping relationship between input data and output data can be obtained by decision table simplification.

The text knowledge acquisition methods proposed in this study can be divided into the following steps: using self-organizing mapping (SOM) neural network method, based on the above clustering text data, the data are discretized, and the data interval is divided; rough set theory is used to extract rules to determine the input text data. According to the maximum matching rule, the optimal rule is selected to obtain the rough set evidence distribution; the evidence distribution is obtained according to the output results of the neural network, and the text data fusion results are obtained by Murphy average method combined with decision rules [19, 20].

The self-organizing mapping network is used to realize the discretization of clustering text knowledge data. There is a SOM network corresponding to each attribute a_i , which is composed of subnet. The single-attribute data are one-dimensional change, and the two-dimensional rough set decision table is obtained after network training. In order to prevent excessive discretization of the decision table, the initial number of neurons in the output layer is set to 2. The dependency among attributes in the data information system is obtained by rough set theory. The basic results of the neural network are shown in Figure 2.

The rough membership function of the relative set X of attribute values under relationship U is expressed as

$$\sigma_X^U(\alpha) = \frac{|X \cap [\alpha]_U|}{|[\alpha]_U|}. \quad (11)$$

Here, $|\cdot|$ represents the number of elements in the set, and α represents the attribute value. It can be seen from the

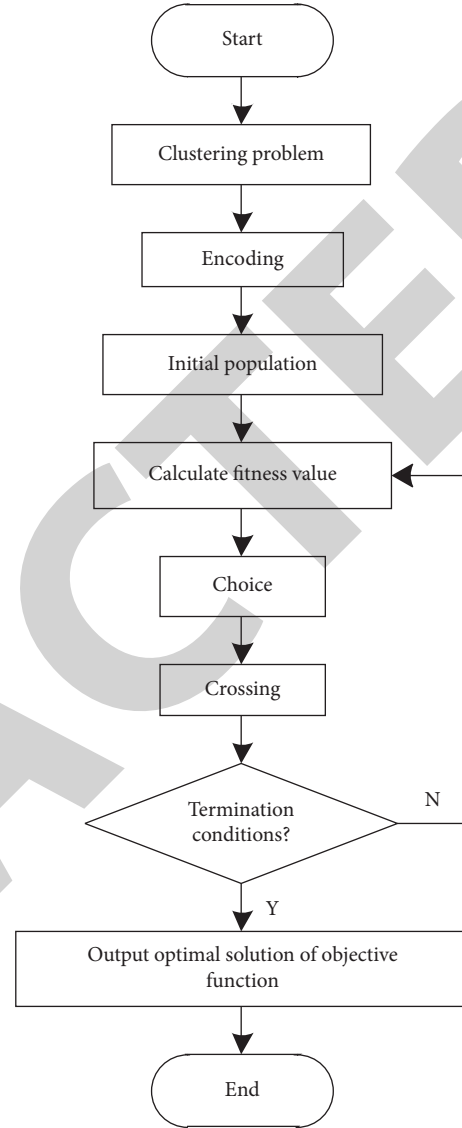


FIGURE 1: Collaborative product design text knowledge clustering process.

above formula that the greater the value of $\sigma_X^U(\alpha)$, the higher the possibility of obtaining the X set according to the attribute value α . Using rough set theory, iQ (condition) $\rightarrow h$ (decision) rules are extracted from the original data of discrete collaborative product design text knowledge. The n -th rule U^i in the description rules is as follows.

If $f(x, q_1) = t_1^i, \dots, f(x, q_n) = t_n^i$, then exist $f(x, q_1) = r_1^i, \dots, f(r, p_r) = r_n^i$.

The matching degree of a set of input IN and the i -th rule is described as follows:

$$B_i = \frac{\sum_{k=1}^n IN(k)}{n}, \quad i = 1, 2, \dots, n. \quad (12)$$

Here, $IN(k)$ represents the input mode function. Formula (12) represents the matching degree between the input pattern and the rule. The rules are obtained from the text knowledge data of collaborative product design, and the reliability of each rule is different. The rough membership

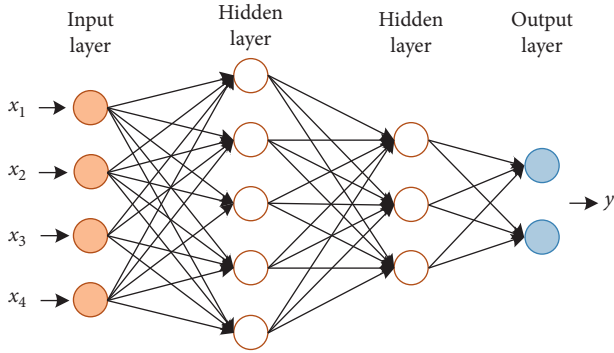


FIGURE 2: Structure of the neural network.

function is used to describe the reliability of the rules, and the applicability is obtained combined with the matching degree [21].

The applicability of rule i corresponding to input in is described as

$$\eta_i = \max\{\eta_X^{qj}(t_1^i)r_i(1), \eta_X^{qj}(t_2^i)r_i(2), \eta_X^{qj}(t_n^i)r_i(n)\}. \quad (13)$$

The evidence P_{RS} is obtained by rough set theory analysis, and the weight value ω is obtained by extracting rules and processing the input. According to the principle of maximum matching degree, appropriate rules are selected. It is assumed that the subscript set of rules is $\vartheta = \{e_1, e_2, \dots, e_n\}$, the corresponding decision expression is $\{Le_1, Le_2, \dots, Le_n\}$, and the applicability is $\{\varepsilon e_1, \varepsilon e_2, \dots, \varepsilon e_n\}$. Assuming that the output of the neural network is $y = [Y_1, Y_2, \dots, Y_n]$, the output $\lambda_i = [\gamma_1, \gamma_2, \dots, \gamma_n]^T$ of the collaborative product design text knowledge acquisition function is solved according to the Murphy average method; then,

$$\lambda_i = \frac{(P_{RS}(i) + P_{NN}(i)/2)^2}{\sum_{i=1}^n (P_{RS}(i) + P_{NN}(i)/2)^2}, \quad i = 1, 2, \dots, n. \quad (14)$$

Here, P_{NN} represents collaborative product design text knowledge data input into neural network to obtain evidence. According to the decision rules of text knowledge acquisition, the text data output decision category is obtained as follows:

$$w_n = \begin{cases} \text{ture}, & \lambda_i = \max(\gamma_1, \gamma_2, \dots, \gamma_n) \\ \text{false}, & \text{else} \end{cases}. \quad (15)$$

The K-means clustering is optimized by genetic algorithm, the collaborative product design text knowledge is clustered, and the rough set combined with neural network method is used to obtain the collaborative product design text knowledge data in an all-round way [22].

3. Experimental Verification

In order to verify the practical application effect of the proposed collaborative product design text knowledge

acquisition method based on genetic algorithm, simulation and comparative verification experiments are carried out.

3.1. Experimental Data. The experimental research data come from the developed collaborative product design project, which is jointly completed by 100 people. Therefore, it has high requirements for the acquisition frequency and speed of text knowledge, which meets the requirements of experimental verification. The structure of the collaborative product synchronous design pattern is shown in Figure 3.

Due to the large scale of the project, the amount of collaborative product design text knowledge data generated is large, with a total of 100 GB. And because the text knowledge data generated in the design process may be used, there is no need to preprocess the experimental data.

3.2. Experimental Environment Deployment. The experimental environment parameters are shown in Table 1.

3.3. Analysis of Experimental Results. The experimental verification scheme is as follows: taking the clustering performance, clustering time, algorithm convergence, and text knowledge acquisition error as the experimental comparison indexes, the text knowledge acquisition method based on genetic algorithm proposed in this paper is compared with the multigranularity method proposed in reference [7] and the method based on fuzzy hyper network proposed in reference [8].

3.3.1. Clustering Performance Judgment. The commonly used indicators for judging clustering performance are accuracy, recall, and F1 test value. Accuracy (PR) represents the correctness of clustering, that is, the proportion of texts in the document that really meet the search intention. It is defined as follows:

$$\text{Pr}_i = \frac{TP_i}{TP_i + FP_i}. \quad (16)$$

Recall (Re) represents the integrity of clustering, that is, the ratio of all texts that meet the search intention to the retrieved text. It is defined as follows:

$$\text{Re}_i = \frac{TP_i}{TP_i + FN_i}. \quad (17)$$

Here, FP_i represents the number of text incorrectly assigned to a certain cluster and belonging to other clusters, TP_i represents the number of text correctly assigned to a certain cluster, and FN_i represents the number of text belonging to a certain cluster and assigned to other clusters.

Accuracy and recall reflect two different aspects of clustering quality, which must be considered comprehensively. In some cases, some new evaluation indexes, F1 test value, will be used, which is defined as follows:

$$F1 = \frac{\text{Pr} \times \text{Re} \times 2}{\text{Pr} + \text{Re}}. \quad (18)$$

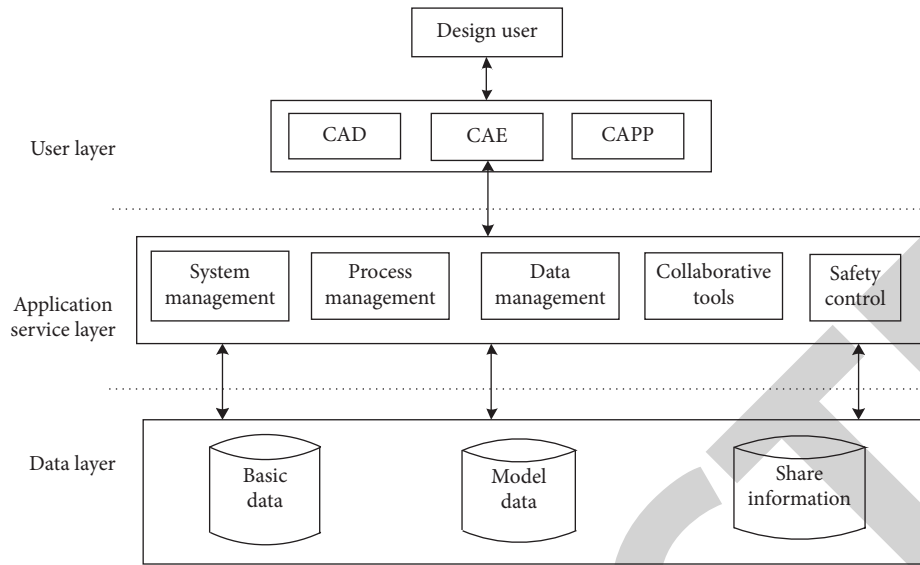


FIGURE 3: Structure of the collaborative product synchronous design pattern.

TABLE 1: Experimental environment parameters.

Parameter name	Parameter index
Equipment model	Dell OptiPlex
Operating system	Windows 10
CPU	Inter(R) core(TM) i5
Memory	16 GB
Network environment	300 GB

It can be seen from the above formula that the larger the F1 value, the better the clustering effect.

The clustering performance comparison results of the three methods are shown in Table 2.

It can be seen from the experimental results in Table 2 that the F1 values of the proposed genetic algorithm are higher than the two comparison methods, so it shows that the clustering effect of this method is better than that of the traditional method.

3.3.2. Clustering Time. In order to more fully verify the performance of the proposed method, considering the acquisition time of the overall text knowledge, clustering time is the key index. Therefore, the clustering time of the proposed method is verified, and the proposed method is compared with two traditional methods. The comparison results of clustering time of the three methods are shown in Figure 4.

By observing the clustering time comparison results shown in Figure 4, it can be seen that under multiple experiments, the proposed method based on genetic algorithm has a shorter clustering time, the maximum clustering time is no more than 1 min, the maximum clustering time based on multigranularity method is 2.5 min, and the clustering time based on fuzzy hyper network method is even 5.8 min. Therefore, it shows that the text knowledge clustering form optimized by genetic algorithm can effectively shorten the clustering time.

3.3.3. Convergence of Clustering Algorithm. The convergence of clustering algorithm has a key impact on the clustering results of collaborative product design text knowledge. Therefore, taking the convergence of clustering algorithm as the experimental comparison index, the convergence of this clustering algorithm is compared with the multigranularity method proposed in reference [7] and the fuzzy hypernetwork method proposed in reference [8]. The convergence comparison results of clustering algorithms in the three methods are shown in Figure 5.

From the convergence comparison results shown in Figure 5, it can be seen that the convergence speed based on fuzzy hyper network method is faster (20 iterations), but it is earlier than a larger objective function value, while the convergence speed based on multigranularity method is slower (45 iterations). Only the method based on genetic algorithm proposed in this paper converges to a minimum value under the requirement of convergence speed (33 iterations). Therefore, it shows that the convergence performance of the proposed method is good.

3.3.4. Collaborative Product Design Text Knowledge Acquisition Error. In the process of collaborative product design, once the text knowledge acquisition error is large, it will affect the final design result of the product. Therefore, the accuracy of text data acquisition in collaborative product design is required. Therefore, it is necessary to verify the text acquisition error of collaborative product design based on text method. The comparison results of collaborative product design text knowledge acquisition errors are shown in Figure 6.

From the comparison results of text knowledge acquisition errors shown in Figure 6, compared with the two comparison methods, the collaborative product design text knowledge acquisition error of this method is the smallest, and the error of this method does not change greatly with the increase of the number of experiments.

TABLE 2: Comparison results of clustering performance.

Test group	Amount of text knowledge data (GB)	F1 value		
		Method based on genetic algorithm	Method based on multigranularity	Method based on fuzzy hyper network
1	20	0.890	0.652	0.637
2	20	0.902	0.686	0.554
3	20	0.939	0.737	0.611
4	20	0.868	0.584	0.617
5	20	0.895	0.687	0.560

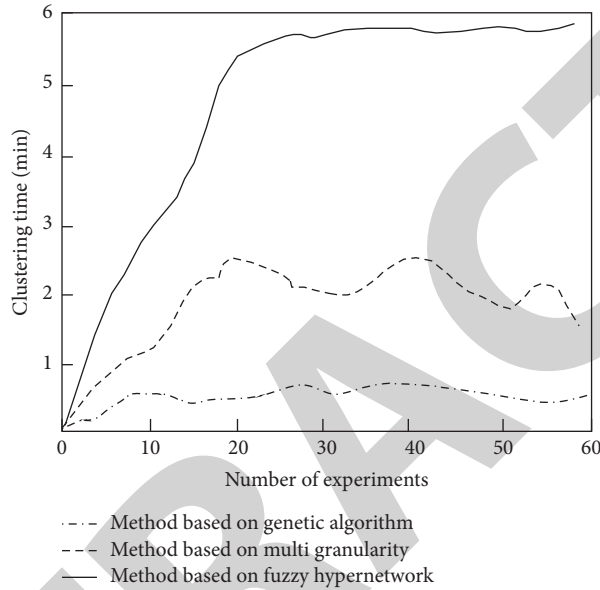


FIGURE 4: Comparison results of clustering time.

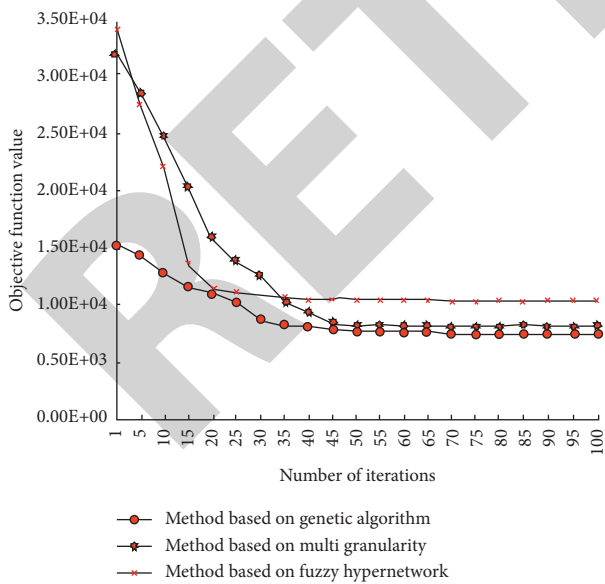


FIGURE 5: Comparison results of convergence of clustering algorithm.

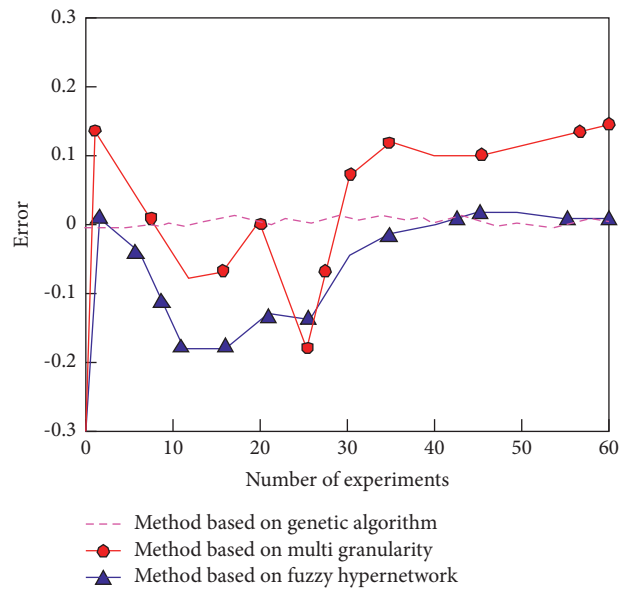


FIGURE 6: Comparison results of text knowledge acquisition errors.

4. Conclusion

In order to improve the effect of collaborative product design text knowledge acquisition, a collaborative product design text knowledge acquisition method based on genetic algorithm is proposed, and the performance of the method is verified from both theory and experiment. This method has shorter clustering time and lower text knowledge acquisition error in collaborative product design. Specifically, compared with the method based on multigranularity, the clustering time of this method is significantly shortened, up to no more than 1 min. Compared with the method based on fuzzy hyper network, the text knowledge acquisition difference of this method is significantly reduced, up to 0.02. Therefore, it shows that the proposed text knowledge acquisition method based on genetic algorithm can better meet the requirements of collaborative design product text knowledge acquisition.

Data Availability

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

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This project was supported by the Key R&D Program of Hainan Province, Research on the Development Technology of Hainan Characteristic Tourism Commodity, under ZDYF2018177, and the Social Science Programming of Philosophy in Hainan Province, Research on Hainan Tourism Commodity Development Path Based on Virtual Reality Technology, under JD(ZC) 21-57.

References

- [1] H. Zhao, K. Liu, and T. Wang, "Simulation of abnormal information acquisition for network text implication relationship recognition," *Computer Simulation*, vol. 37, no. 08, pp. 256–260, 2020.
- [2] dH. E. Bradfor, S. Susanne, and B. Michael, "A knowledge acquisition method for event extraction and coding based on deep patterns," *Springer*, vol. 19, no. 3, pp. 116–119, 2018.
- [3] H. K. McGinty, "KNnowledge acquisition and representation methodology," (*KNARM*) and Its Applications, vol. 79, no. 33, pp. 561–565, 2018.
- [4] A. Fesseha, S. Xiong, and E. D. Emiru, "Text Classification of News Articles Using Machine Learning on Low-Resourced Language: Tigrigna," in *Proceedings of the 2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, Chengdu, China, May 2020.
- [5] N. Cs, "Data science based research on sentiment analysis," *Machine Learning and Text Classification of Social Media and Digital Forensic Data*, vol. 13, no. 4, pp. 71–77, 2018.
- [6] J. J. Mcarthur, N. Shahbazi, R. Fok, C. Raghubar, B. Bortoluzzi, and A. An, "Machine learning and BIM visualization for maintenance issue classification and enhanced data collection," *Advanced Engineering Informatics*, vol. 38, no. 10, pp. 101–112, 2018.
- [7] Q. Wan, Y. Ma, and L. Wei, "Knowledge acquisition of multi-source data based on multigranularity," *Journal of Shandong University*, vol. 55, no. 01, pp. 41–50, 2020.
- [8] L. Cheng and F. Hu, "Fuzzy hypernetwork-based knowledge acquisition method," *CAAI Transactions on Intelligent Systems*, vol. 14, no. 03, pp. 479–490, 2019.
- [9] Y. Zhang, G. Wang, and Z. Zhu, "Text clustering method based on knowledge graph word sense disambiguation," *Journal of Hebei United university natural science edition*, vol. 41, no. 04, pp. 89–94, 2019.
- [10] V. S. Pillutla and P. J. Giabbanelli, "Iterative generation of insight from text collections through mutually reinforcing visualizations and fuzzy cognitive maps," *Applied Soft Computing*, vol. 76, no. 31, pp. 459–472, 2019.
- [11] V. Alekseev, E. Egorov, and K. Vorontsov, "TopicBank: collection of coherent topics using multiple model training with their further use for topic model validation," *Data & Knowledge Engineering*, vol. 21, no. 4, pp. 101–108, 2021.
- [12] M. Das, J. Li, and E. Fosler-Lussier, "Sequence-to-Set semantic tagging: end-to-end multi-label prediction using neural attention for complex query reformulation and automated text categorization," *Machine Learning*, vol. 36, no. 8, pp. 614–618, 2019.
- [13] N. Alami, M. Meknassi, and N. En-Nahnahi, "Unsupervised neural networks for automatic Arabic text summarization using document clustering and topic modeling," *Expert Systems with Applications*, vol. 19, no. 7, pp. 646–652, 2021.
- [14] A. K. Abasi, A. T. Khader, and M. A. Al-Betar, "A novel hybrid multi-verse optimizer with K-means for text documents clustering[J]," *Neural Computing & Applications*, vol. 32, no. 3, pp. 112–118, 2020.
- [15] J. Pareek, D. Singhania, R. R. Kumari, and S. Purohit, "Gujarati handwritten character recognition from text images," *Procedia Computer Science*, vol. 171, pp. 514–523, 2020.
- [16] Q. Lu and W. Chen, "Automatically building a large scale dictionary of Chinese entity sentiment expressions," *Journal of Chinese Information Processing*, vol. 32, no. 08, pp. 32–41, 2018.
- [17] M. M. Mironczuk and J. Protasiewicz, "A recent overview of the state-of-the-art elements of text classification," *Expert Systems with Applications*, vol. 106, no. 19, pp. 36–54, 2018.
- [18] J. Currie and E. Zwiers, "Technology and big data are changing economics: mining text to track methods," *NBER Working Papers*, vol. 65, no. 17, pp. 244–249, 2020.
- [19] F. Ertam, "Deep Learning Based Text Classification with Web Scraping Methods," in *Proceedings of the 2018 International Conference on Artificial Intelligence and Data Processing (IDAP)*, Malatya, Turkey, September 2019.
- [20] S. K. Kang, L. Path, and A. Rangarajan, "Automated feedback generation for formal manufacturing rule extraction," *Artificial Intelligence for Engineering Design Analysis and Manufacturing*, vol. 33, no. 3, pp. 1–13, 2019.
- [21] Q. Xu, Q. Guo, and C. X. Wang, "Network differentiation: a computational method of pathogenesis diagnosis in traditional Chinese medicine based on systems science," *Artificial Intelligence in Medicine*, vol. 2021, no. 7724, Article ID 102134, 2017.
- [22] L. Minmin, J. Huajuan, H. Yule et al., "A systematic review on botany, processing, application, phytochemistry and pharmacological action of Radix Rehmanniae," *Journal of Ethnopharmacology*, vol. 285, Article ID 114820, 2021.