Classification Method of Japanese Teaching Resources Based on Density Clustering Algorithm

Qinlang Li
Chengdu Neusoft University, Chengdu City, Sichuan Province, 611844, China

Keywords: Computer, Classification Method Based on Density Clustering Algorithm, Japanese Language Teaching, Teaching Resources.

Abstract: In Japanese language teaching, teaching resources are very important and make the process of learning Japanese easier. From this, we can know the classification of resources, you can better find the required information, and the general classification method cannot solve the problem of inaccurate classification of Japanese teaching resources. Therefore, this paper proposes a classification method based on density clustering algorithm for resource classification analysis. First, computers are used to classify Japanese language teaching resources, and indicators are divided according to resource classification requirements to reduce resource classification in the interfering factor. Then, the computer classifies the results of Japanese teaching resources, forms a resource classification scheme, and classifies the resources Conduct a comprehensive analysis. MATLAB simulation shows that the classification method based on density clustering algorithm can classify resources in Japanese language teaching accurately under certain evaluation criteria The efficiency of resource classification is better than that of ordinary classification methods.

1 INTRODUCTION

With the continuous development of artificial intelligence and big data technology, these technologies have gradually been introduced in the field of education to optimize and improve the classification and management of educational resources (Adebayo, and Quadri, et al. 2023). The density clustering algorithm is a commonly used clustering algorithm that does not require a pre-defined number of clusters and can automatically discover clusters in a dataset (Arabit-Garcia, J and Prendes-Espinosa, et al. 2023). This article will explore the application of density clustering algorithm in the classification of Japanese teaching resources.

1.1 Description of the classification of Japanese language teaching resources

In Japanese language education, there are a large number of teaching resources, such as teaching materials, vocabulary, grammar, listening and reading materials, etc. These resources are large in number and complex in classification, and need to be automatically classified using clustering algorithms to better manage and utilize these resources (Arblaster, and Mackenzie, et al. 2023). The goal of the Japanese Language Teaching Resource Classification is to put similar resources in the same category and different resources in different categories for better management and utilization.

1.2 Fundamentals of density clustering algorithms

The density clustering algorithm is a density-based non-hierarchical clustering algorithm, which does not need to pre-set the number of clusters, but calculates the local density of data points to mine out the clusters in the dataset (Chan, 2023). The basic principles of the density clustering algorithm can be summarized in the following steps:

Density reachability: According to the specified distance threshold, the density reachability judgment is made for each data point, and if one data point can be reached from another data point (Chen, 2023), the two data points are considered to be density reachable.

Li, Q.
Classification Method of Japanese Teaching Resources Based on Density Clustering Algorithm.
DOI: 10.5220/0013539100004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 1*, pages 250-256
ISBN: 978-989-758-763-4

Core object: A data point is considered core if it contains enough data points within a radius centered on it and within a specified distance threshold (Chen, 2023).

Direct density reachability: For a core object, if a data point is within the radius of the specified distance threshold, the data point is considered to be directly density reachable (Cohen, and Lefstein, et al. 2023).

Density connected: If a data point is both a core object and another core object has a direct density reachability, the two data points are considered to be density connected (Gosztonyi, and Varga, 2023).

Cluster formation: By continuously adding densely connected data points to the same cluster, several clusters are eventually formed (Javier Robles-Moral, and Fernandez-Diaz, et al. 2023).

Noise points: Data points that are not included in any cluster are considered noise points (Jin, and Lu, 2023).

The advantage of the density clustering algorithm is that it does not need to pre-set the number of clusters, can automatically discover clusters in the dataset, and has good fault tolerance for noise points.

1.3 Application of Density Clustering Algorithm in the Classification of Japanese Teaching Resources

In the classification of Japanese teaching resources, the density clustering algorithm can be applied to two aspects: grammar and vocabulary (Karol, and Shaylor, et al. 2023).

1.3.1 Grammatical Classification

In grammatical classification, each sentence can be regarded as a data point, and the similarity between sentences can be judged by calculating the distance between sentences, and similar sentences can be placed in the same cluster. In the process of grammatical classification, the subject, predicate, object and other information of the sentence can be added to the distance calculation, so as to make the classification result more accurate (Kazima, and Jakobsen, et al. 2023).

1.3.2 Vocabulary Classification

In lexical classification, you can think of each word as a data point, calculate the distance between words, judge the similarity between words, and put similar words in the same cluster. In the process of vocabulary classification, the part of speech, meaning, frequency and other information of words can be added to the distance calculation, so as to make the classification results more accurate (Krieg, 2023).

1.4 Practical Application of Density Clustering Algorithm in the Classification of Japanese Teaching Resources

1.4.1 Dataset Selection and Preprocessing

Choosing the right dataset is one of the keys to classification. In Japanese language education, datasets can be constructed by collecting various teaching resources such as teaching materials, listening materials, reading materials, etc. In the preprocessing of a dataset, each data point can be converted into a vector, adding different features to the vector (Kwee, and Santos, 2023).

1.4.2 Practical Case Application

In practical case applications, the density clustering algorithm can be applied to the classification of Japanese teaching resources, such as classifying Japanese textbooks and Japanese phonetic vocabulary. In the process of classification, the appropriate distance calculation method and density threshold can be selected according to specific needs, so as to obtain reasonable classification results.

1.5 Challenges of Density Clustering Algorithm In The Classification Of Japanese Teaching Resources

1.5.1 The Quality of the Dataset

The quality of the dataset directly affects the accuracy of the classification results. How to ensure the quality of data sets and avoid duplicate data and erroneous data is one of the issues that need to be studied and explored.

1.5.2 Selection of Distance Calculation Method

The distance calculation method directly affects the effectiveness and speed of classification. How to choose the appropriate distance calculation method to improve the accuracy and speed of classification is also one of the problems that need to be studied and discussed.

1.5.3 Parameter selection

There are many parameters to choose in the density clustering algorithm, such as density threshold, neighborhood radius, and so on. How to choose the appropriate parameters is one of the keys to affect the effect of the algorithm.

The classification of Japanese teaching resources is a complex problem, and the density clustering algorithm is an effective classification algorithm, which can automatically discover the clusters in the dataset, without presetting the number of clusters, and has good fault tolerance for noise points (Li, and Qu, et al. 2023). However, there are challenges such as the quality of the dataset, the choice of distance calculation method, and the selection of parameters. In future research, it is necessary to further optimize the parameters of the algorithm, improve the effect and speed of the algorithm, and improve the robustness and adaptability of the algorithm.

Teaching resources are one of the important components of Japanese language teaching and are of great significance for learning Japanese. However, in the process of resource classification, the resource classification scheme has the problem of poor accuracy, which has a certain impact on Japanese language learning. Some scholars believe that applying the classification method based on density clustering algorithm to Japanese teaching analysis can effectively analyze the resource classification scheme and provide corresponding support for resource classification. On this basis, this paper proposes a classification method based on density clustering algorithm to optimize the resource classification scheme and verify the effectiveness of the model.

2 RELATED CONCEPTS

2.1 Mathematical Description of a Classification Method Based on Density Clustering Algorithm

The classification method based on the density clustering algorithm uses computer operations to optimize the resource classification scheme, and finds the unqualified values in Japanese language teaching according is z_i , and the indicators in the resource classification y_i , and classifies the resources the scheme is $tol(y_i \cdot h_{ij})$, integrated to determine the feasibility of Japanese language teaching, and the calculation is shown in Equation (1).

$$tol(y_i \cdot h_{ij}) = y_{ij} \geq \max(h_{ij} \sum_{i=1}^{n} h_i Y_i \frac{1}{n} \sigma_h^2 \cdot \frac{h-\mu}{\sigma}) \quad (1)$$

Among them, the judgment of outliers is shown in Equation (2).

$$\max(h_{ij}) = (h_{ij}{}^2 \div 3) \succ mean(\sum h_{ij} \frac{h-\mu}{\sigma} \frac{1}{n} \cdot \sum_{i=1}^{n} h_i^2) \quad (2)$$

The classification method based on density clustering algorithm combines the advantages of computer computing and uses Japanese language teaching for quantification, which can improve the accuracy of resource classification.

Hypothesis I. The resource classification requirements is h_i , the resource classification scheme is set_i , the satisfaction of the resource classification scheme is y_i , and the resource classification scheme judgment function is $D(h_i \approx 0)$, As shown in Equation (3).

$$D(p_i) = \sum h_i \bigcap \xi \to [\int y_i \Updownarrow \sum \lim_{\delta h \to 0} \frac{\delta y}{\delta h} \frac{\Delta y}{\Delta h} \quad (3)$$

2.2 Selection of Teaching Resource Programs

Hypothesis II. The Japanese teaching function is $j(h_i)$, the weight coefficient is w_i , then, the resource classification requires unqualified Japanese teaching as shown in Equation (4).

$$j(h_i) = z_i \cdot \prod D(p_i) - w_i \to \int \lim_{\delta h \to 0} \frac{\delta y}{\delta h} \frac{\Delta y}{\Delta h} \quad (4)$$

Based on assumptions I and II, a comprehensive function for Japanese language teaching can be obtained, and the result is shown in Equation (5).

$$j(h_i) + D(p_i) \leq \max(h_{ij}) \quad (5)$$

In order to improve the accuracy of resource classification effectively, all data needs to be normalized, and the results are shown in Equation (6).

$$\overline{j(h_i)} + D(p_i) \leftrightarrow mean(\sum h_{ij} \frac{h-\mu}{\sigma} \frac{1}{n} \cdot \sum_{i=1}^{n} h_i^2 \quad (6)$$

2.3 Analysis of Resource Classification Schemes

Before the classification method based on density clustering algorithm, multi-dimensional analysis should be performed on the resource classification scheme, and the resource classification requirements should be mapped to the Japanese teaching library to eliminate the unqualified resource classification scheme is $No(h_i)$, According to Equation (6), the anomaly evaluation scheme can be proposed, and the results are shown in Equation (7).

$$No(h_i) = \frac{\overline{j(h_i)} + D(p_i)}{mean(\sum h_{ij} \frac{h-\mu}{\sigma} \frac{1}{n} \cdot \sum_{i=1}^{n} h_i^2)} \quad (7)$$

Among them, $\frac{\overline{j(h_i)} + D(p_i)}{mean(\sum h_{ij} \frac{h-\mu}{\sigma} \frac{1}{n} \cdot \sum_{i=1}^{n} h_i^2)} \leq 1$ it is stated that the scheme needs to be proposed, otherwise the scheme integration required is $Zh(h_i)$, and the result is shown in Equation (8).

$$Zh(h_i) = \min[\sum \overline{j(h_i)} + D(p_i)] \quad (8)$$

Japanese language teaching conducts comprehensive analysis, and sets the threshold and indicator weights of the resource classification scheme to ensure the accuracy of the classification method based on the density clustering algorithm. Japanese language teaching is a systematic test resource classification scheme that needs to be analyzed. If Japanese language teaching is in a nonnormal distribution is $unno(h_i)$, its resource classification scheme will be affected, reducing the accuracy of the overall resource classification, and the calculation result is $accur(h_i)$, shown in Equation (9).

$$accur(h_i) = \frac{\min[\sum \overline{j(h_i)} + D(p_i)]}{\sum \overline{j(h_i)} + D(p_i)} \times 100\% \quad (9)$$

The survey resource classification scheme shows that the teaching resource scheme shows a multi-dimensional distribution, which is in line with the objective facts. Japanese language teaching is not directional, indicating that the teaching resource scheme has a strong randomness, so it is regarded as a high analytical study. If the random function of Japanese teaching is $randon(h_i)$, then the calculation of formula (9) can be expressed as formula (10).

$$accur(h_i) = \frac{\min[\sum \overline{j(h_i)} + D(p_i)]}{\sum \overline{j(h_i)} + D(p_i)} \times 100\% + randon(h_i) \quad (10)$$

Among them, Japanese language teaching meets the normal requirements, mainly computer operations to adjust Japanese language teaching, eliminate duplicate and irrelevant schemes, and supplement the default scheme, so that the dynamic correlation of the entire resource classification scheme is strong.

3 OPTIMIZATION STRATEGIES FOR JAPANESE LANGUAGE TEACHING

The classification method based on density clustering algorithm adopts the random optimization strategy for Japanese language teaching, and adjusts the parameters of Japanese teaching resources to realize the optimization of Japanese language teaching. The classification method based on the density clustering algorithm divides Japanese language teaching into different resource classification levels, and randomly selects different schemes. In the iterative process, the resource classification schemes of different resource classification levels are optimized and analyzed. After the optimization analysis is completed, compare the resource classification levels of different schemes and record the best Japanese language teaching.

4 PRACTICAL EXAMPLES OF JAPANESE LANGUAGE TEACHING

4.1 Introduction to Resource Classification

In order to facilitate resource classification, this paper takes Japanese language teaching in complex situations as the research object, with 12 paths and a test time of 12h, and the specific resource classification of Japanese language teaching The scheme is shown in Table 1.

Table 1: Resource classification requirements

Scope of application	grade	accuracy	Teaching resources
Basic	routine	85.47	87.54
equivalents	Higher	85.13	85.68
Advanced	routine	87.82	86.31
equivalents	Higher	86.35	87.52
Audition	routine	83.50	85.48
	Higher	86.15	86.00

The resource classification process in Table 1 is shown in Figure 1.

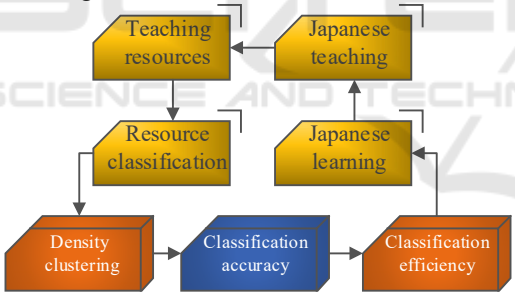

Figure 1: The analytical process of Japanese language teaching

Compared with the general classification method, the resource classification scheme of the classification method based on the density clustering algorithm is closer to the actual resource classification requirements. In terms of the rationality of Japanese language teaching, the classification method based on density clustering algorithm is better than the ordinary classification method. It can be seen from the change of resource classification scheme in Figure II that the classification method based on density clustering algorithm has higher accuracy and faster classification speed. Therefore, the resource classification scheme speed, teaching resource scheme resource classification scheme and summation stability of the classification method based on density clustering algorithm are better.

4.2 Japanese Language Teaching

The resource classification scheme for Japanese language teaching includes unstructured information, semi-structured information, and structured information. After the pre-selection of the classification method based on the density clustering algorithm, the preliminary resource classification scheme for Japanese language teaching is obtained, and the Japanese language teaching is obtained Analyze the feasibility of resource classification schemes. In order to verify the results of Japanese language teaching more accurately, select Japanese language teaching with different resource classification levels, and the resource classification scheme is shown in Table 2.

Table 2: The overall picture of the teaching resource program

Category	Rationality	Analysis rate
Basic Equivalents	91.72	87.68
Advanced Equivalents	87.48	92.07
Audition	88.77	90.41
Mean	90.12	90.03
X^6	89.39	90.71
	P=1.696	

4.3 Teaching Resources and Stability of Resource Classification

In order to verify the accuracy of the classification method based on the density clustering algorithm, the resource classification scheme is compared with the general classification method, and the resource classification scheme is shown in Figure 2.

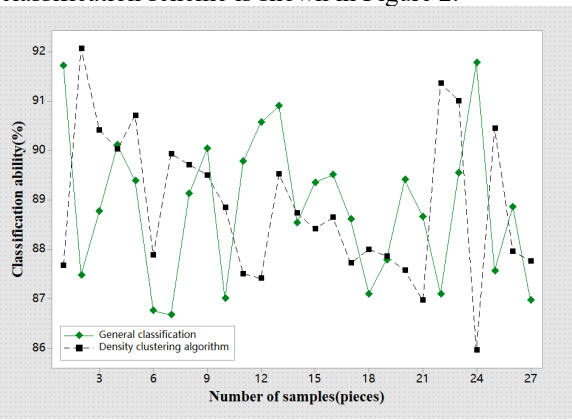

Figure 2: Teaching resources for different algorithms

It can be seen from Figure 2 that the teaching resources of the classification method based on density clustering algorithm are higher than those of ordinary classification methods, but the error rate is lower, indicating that the resource classification of the classification method based on density clustering algorithm is relatively stable The classification of resources in the general classification method is uneven. The average resource classification scheme of the above three algorithms is shown in Table 3.

Table 3: Comparison of resource classification accuracy of different methods

Algorithm	Teaching resources	Magnitude of change	Error
A classification method based on density clustering algorithm	93.24	95.46	94.44
Common classification method	89.94	92.54	91.42
P	91.78	85.96	90.97

Table 3 shows that the general classification method has shortcomings in the accuracy of the classification of teaching resources in Japanese language teaching, and the Japanese language teaching has changed significantly, and the error rate is high. The general results of the classification method based on the density clustering algorithm have higher teaching resources and are better than the general classification method. At the same time, the teaching resources of the classification method based on density clustering algorithm are greater than 93%, and the accuracy does not change significantly. In order to further verify the superiority of the classification method based on density clustering algorithm. In order to further verify the effectiveness of the proposed method, the classification method based on density clustering algorithm is analyzed by different methods, as shown in Figure 3.

It can be seen from Figure 3 that the teaching resources of the classification method based on the density clustering algorithm are significantly better than the ordinary classification method, and the reason is that the classification method based on the density clustering algorithm increases the adjustment coefficient of Japanese teaching and sets it Thresholds for Japanese teaching resources that exclude resource classification schemes that do not meet the requirements.

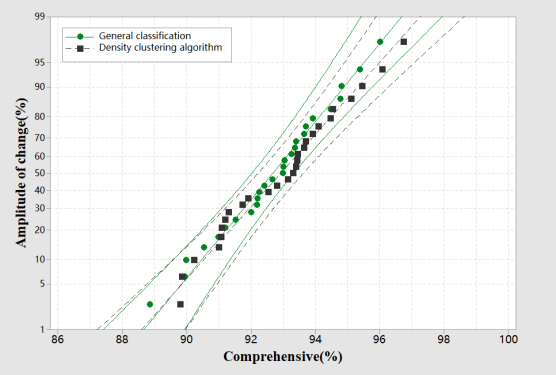

Figure 3: Classification method based on density clustering algorithm: teaching resource for resource classification

5 CONCLUSIONS

Aiming at the problem that the classification of Japanese teaching resources is not ideal, this paper proposes a classification method based on density clustering algorithm, and combines computer operation to optimize Japanese language teaching. At the same time, the accuracy of resource classification is analyzed in depth, and a collection of Japanese teaching resources is constructed. This study shows that the classification method based on the density clustering algorithm can improve the rationality of Japanese language teaching, and can classify general resources for Japanese language teaching 。 However, in the process of classification method based on density clustering algorithm, too much attention is paid to the analysis of resource classification, resulting in unreasonable selection of resource classification indicators.

REFERENCES

Adebayo, S. B., Quadri, G., Igah, S., & Azubuike, O. B.(2023) Teaching in a lockdown: The impact of COVID-19 on teachers? capacity to teach across different school types in Nigeria. Heliyon, 9(3).

Arabit-Garcia, J., Prendes-Espinosa, M. P., & Serrano, J. L.(2023) Open Educational Resources and active methodologies for STEM teaching in Primary Education. Revista Latinoamericana De Tecnologia Educativa-Relatec, 22(1): 89-106.

Arblaster, K., Mackenzie, L., Buus, N., Chen, T., Gill, K., Gomez, L., Hamilton, D., Hancock, N., McCloughen, A., Nicholson, M., Quinn, Y., River, J., Scanlan, J. N., Schneider, C., Schweizer, R., & Wells, K.(2023) Co-design and evaluation of a multidisciplinary teaching resource on mental health recovery involving people

with lived experience. Australian Occupational Therapy Journal, 70(3): 354-365.
Chan, H.(2023) 3E Model: How to Use OERs to Enhance Teaching/Learning. Computer, 56(4): 139-142.
Chen, J.(2023a) Reform of English Writing Teaching Method Under the Background of Big Data and Artificial Intelligence. International Journal of E-Collaboration, 19(4): 20-20.
Chen, L.(2023b) Empirical analysis of integrated teaching mode of international trade based on practical training data detection and neural network. Soft Computing.
Cohen, E., Lefstein, A., & Dishon, G.(2023) The end of the textbook and the beginning of teaching? Tradeoffs in designing on-line support for K-12 teachers. Education and Information Technologies.
Gosztonyi, K., & Varga, E.(2023) Teachers' practices and resources in the Hungarian "Guided Discovery" approach to teaching mathematics: presenting and representing "series of problems". Zdm-Mathematics Education.
Javier Robles-Moral, F., Fernandez-Diaz, M., & Enrique Ayuso-Fernandez, G.(2023) A Study of the Usefulness of Physical Models and Digital Models for Teaching Science to Prospective Primary School Teachers. Education Sciences, 13(4).
Jin, H., & Lu, Q.(2023) Exploring the application of inquiry-based defensive teaching based on the internet and clinical resources during probationary teaching of dermatovenereology. Chinese Journa of Dermatology, 56(4): 353-356.
Karol, D., Shaylor, R., Fiszer, E., & Weiniger, C. F.(2023) Serious games as an innovative learning tool at a medical conference to teach medical knowledge and crisis resource management A narrative report. European Journal of Anaesthesiology, 40(2): 143-146.
Kazima, M., Jakobsen, A., Mwadzaangati, L., & Gobede, F.(2023) Teaching the concept of zero in a Malawi primary school: illuminating the language and resource challenge. Zdm-Mathematics Education.
Krieg, A.(2023) Teaching the Abstract: An Evaluation of "Social Structure" in Introductory Textbooks. Teaching Sociology.
Kwee, C. T. T., & Dos Santos, L. M.(2023) An international study of high school teachers' experience of incorporating water resources in their teaching. Frontiers in Education, 7.
Li, D., Qu, P., Jin, T., Chen, C., & Bai, Y.(2023) Automatic classification of multi-source and multi-granularity teaching resources based on random forest algorithm. International Journal of Continuing Engineering Education and Life-Long Learning, 33(2-3): 177-191.