Retraction

Retracted: Prediction of Customer Demand for Hands-On Inquiry-Based Learning (HIBL) Product Based on Big Data Clustering Algorithm

Security and Communication Networks

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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) Peer-review manipulation

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] X. Bai, "Prediction of Customer Demand for Hands-On Inquiry-Based Learning (HIBL) Product Based on Big data Clustering Algorithm," *Security and Communication Networks*, vol. 2022, Article ID 7828378, 7 pages, 2022.

WILEY | Hindawi

Research Article

Prediction of Customer Demand for Hands-On Inquiry-Based Learning (HIBL) Product Based on Big Data Clustering Algorithm

Xingxing Bai [ID]

Ideological and Political Department, Lvliang University, Lvliang, Shanxi 033000, China

Correspondence should be addressed to Xingxing Bai; 20131100@llu.edu.cn

The objective was to identify previously unknown groups in a dataset using various techniques. Significant progress has been made in this field in recent years, resulting in the development of novel and promising clustering algorithms. With the constant advancement of big data technology, research on study tours has also become crucial. Clustering can unearth the potential hidden information in large datasets, thereby facilitating more efficient work. Diverse measures have been proposed to quantify similarity, including the Euclidean distance and data space density. As a result, clustering becomes a multi-objective optimization problem. Clustering algorithms are extensively utilized in data preprocessing, data classification, and big data prediction. In this study, we examine clustering methods for big data from a theoretical perspective to comprehend their correlations across a large number of datasets. In addition, we predicted customer demand for research products using fabricated metrics.

1. Introduction

Various metrics, including edit distance, density in Euclidean or non-Euclidean data space, distance computed using the Minkowski metric, proximity metric, and probability distribution, can be used to objectively characterize data clustering [1]. All metrics concur that a threshold should be specified for the grouping of items within a cluster and that anything exceeding this threshold should be regarded as distinct and separated from the remainder of the cluster. Clustering may be used to better describe data because the characteristics of all items within a cluster are less variable and can be summarized more effectively than when they are not clustered. Clustering is applicable in a variety of other domains, such as predicting missing data values and identifying data outliers. As a result, clustering is a meta-learning technique applicable to numerous datasets and fields, including market research, e-commerce, social network analysis, and the aggregation of search results. It is possible to organize data into clusters using a variety of techniques, but there is no universally applicable method. Due to the fact that each algorithm is based on particular assumptions and has its own biases, there is no consensus

regarding the optimal algorithm. In unsupervised clustering, there is no widely accepted objective criterion for accuracy or effectiveness, and each of these methods has its own set of flaws and accomplishments when it comes to addressing the complex challenges of unsupervised clustering [2].

Clustering is one of the most frequently employed techniques for unsupervised learning and data analysis in data mining, as well as one of the most frequently employed techniques in machine learning. Clustering can be used to uncover and exploit the internal relationships that are concealed within a dataset. Consumer data for research goods are evaluated using clustering, and its hidden correlations are efficiently mined to provide relevant departments with useful information. Clustering can be divided into four categories based on the various clustering principles [3–5]: hierarchical clustering, partitioned clustering, density-based clustering, and model-based techniques. Hierarchical clustering is the most common type of clustering. The partitioned clustering method is one of the most extensively utilized of these algorithms, owing to the simplicity of its premise and ease of implementation. Application: following the concepts of similarity and dissimilarity, the partition clustering method separates the dataset into k

classes, each of which contains a single value. For data analysis utilizing the partition clustering technique, it is important to know the number of classes k of the dataset in advance of starting the analysis. To begin, k data points are randomly picked as the initial clustering center. The remaining data points are then separated into distinct groups based on the idea of comparable rows, which is illustrated in Figure 1. New cluster centers are selected by the optimization goal as long as the specified requirements are satisfied.

Observational research [6–8]: the term study tour is sometimes used to refer to tourism. It began earlier in Europe and the United States, as well as Japan and South Korea in Asia, and has progressed more rapidly since then. It is a tourist initiative in the shape of a team that will go to various locations to study and practice [9–11]. As a form of education for the children of nobility, research travel in Europe and the United States began in the early 1800s and has persisted. In later years, it was gradually promoted within the educational system, such as through graduation trips and extracurricular activities. Outdoor education and educational tourism are the two most frequently encountered types of educational tourism. Japan was the first nation to introduce the concept of study tourism, which was adopted into the elementary and secondary education curriculum in the 1960s and implemented throughout the nation. Since then, the idea has evolved and advanced, becoming more complex. South Korea, which has benefited from the Japanese study model since the 1980s, has begun to promote study tourism in its rapidly developing nation. The success of research and study travel initiatives abroad is measured by their capacity to produce positive educational outcomes.

Even though our country has utilized visits to famous mountains and rivers to cultivate temperament and emotion since antiquity, the development of research tourism in the modern sense is a relatively recent phenomenon. Since the economic reform and liberalization, research tourism has predominantly taken the form of winter and summer camps. Relevant national policies are introduced on a regular basis. It is an off-campus educational activity organized and arranged by the school's education department, and it combines travel experience and research study through centralized board and lodging. It is an off-campus educational activity that combines travel and research through group travel and centralized board and lodging. Tourism research and practice in my country, as well as tourism research and study, grew rapidly to a crest.

2. Related Work

2.1. Clustering Algorithm. Dividing the dataset of a given data item into many categories and assigning a division category label to each of these categories are the goal of the division-clustering method, which is designed to maximize an objective function. K-means and k-proximity algorithms, which are among the most widely used partitioning and clustering algorithms, have been used to generate an enormous number of deformation algorithms [12–16]. The

K-means method, the K-medoids algorithm, and the CLARANS algorithm are the three most important representative algorithms.

While K-means can detect data distributions for aspheric classes, density-based corner clustering can detect irregular shapes in aspheric classes. At the same time, it is capable of detecting clusters of varying forms and sizes in noisy data and of dealing well with noisy data. This approach, in contrast to partitioning and hierarchical clustering methods, defines clusters as a collection of points that are closely linked. Collection: the DBSCAN method, the DENCLUE algorithm, and other density clustering techniques are among the most widely used [17–20]. The DBSCAN method, for example, finds clusters by linking high-density neighborhood points in a given area.

Grid-based clustering is a technique in which the data space is separated into a grid structure, and the bottom-up grid division method may divide the parameters entered by the user into grid cells that are all the same size. The term high-density grid cell refers to a grid cell that contains far more data points than usual. As a result, there are only a limited number of cells in the grid, and all processing takes place in only one cell. The key advantage of the network-based technology is the rapidity with which it can be processed. STING [21], WaveCluster, CLIQLE, and other grid clustering methods are among the most widely used in practice.

The goal of automated multi-objective clustering is to alleviate the difficulty of having to identify the ultimate number of clusters or divisions in advance, which is a common occurrence. Multi-objective clustering with automatic k-determination [22] is a multi-objective clustering algorithm based on the PESA-II [23] algorithm that is used to optimize two complementary objective functions: population bias and connectivity. The MOCK method outperforms typical single-objective clustering algorithms on a variety of benchmark datasets, as demonstrated in previous research. Suitable for hyperspherical or split clusters, the approach can produce superior clustering results than previous algorithms. However, the clustering results obtained by MOCK on overlapping clusters are unsatisfactory in nature. The VAMOSA clustering algorithm is presented in [24], and it is a multi-objective clustering technique. When stringifying, this approach makes use of the multi-objective optimization method based on simulated annealing as the underlying optimization strategy, and it makes use of the encoding rules of clustering centers as the underlying optimization strategy. In clustering, the point-symmetric distance, which was recently developed in [25], is employed instead of the conventional Euclidean distance, as previously stated. Overall, the experimental findings demonstrate that the VAMOSA method outperforms the MOCK algorithm in terms of overall performance, but that the approach does not exhibit high resilience when clustering data from overlapping datasets. The gene expression programming method [26] has also been used to develop a class of clustering algorithms known as GEP cluster, which all make use of biological properties that occur throughout the process of genetic evolution, such as niches and other similar
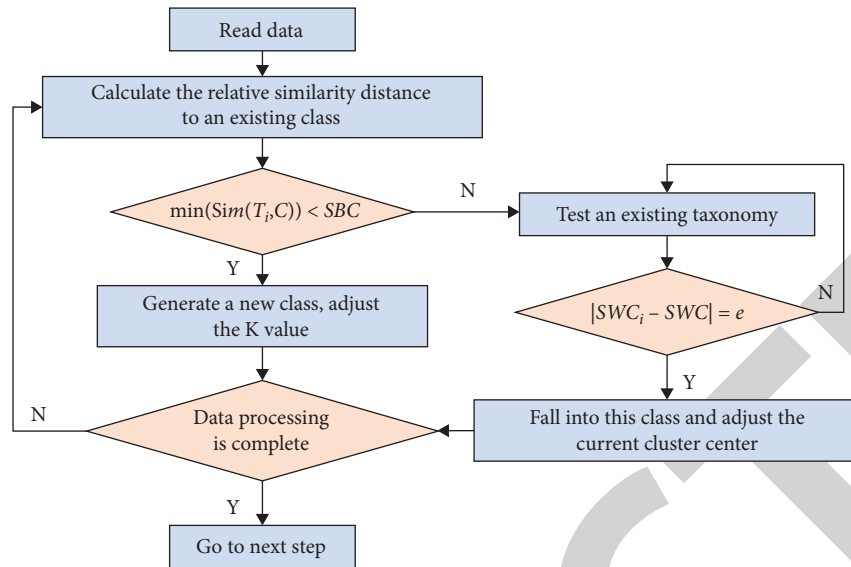
Figure 1: Algorithm flowchart.

traits. The relevant experimental results are listed below. According to the results, the performance of this type of algorithm is superior to the performance of the original GEP-based clustering method.

Multiple researchers [27] have studied existing clustering algorithms in a variety of categories. In the course of their research, they constructed a comparison between five categories and their most representative algorithms to identify the algorithm that performed the best with large amounts of data. By examining various problems and characteristics, the authors of a recent paper [28] provide a comprehensive overview of data mining techniques and platforms that may be employed in the field of big data. They conclude by suggesting directions for future research. The study [29] examines a variety of large-scale data mining techniques and, after conducting a thorough comparison, determines which method is most suitable. Certain researchers are particularly interested in the study of classification algorithms that may be employed in statistics, as well as their application to classification algorithms for specific databases [30]. Some academics [31] have studied the closest neighbor search, decision trees, and neural networks, among other methods, because they are capable of handling enormous datasets. Certain scientists [32] studied the use of MapReduce for parallel categorization and discussed several clustering algorithms, including MapReduce. They present an overview of the various data mining clustering techniques.

2.2. Research Trips.

The practice of learning outside has a long history. It is rooted in the ideals of British educator John Locke's gentleman's education, and it is backed by contemporary theories such as Dewey's learning by doing, Piaget's constructivism, and Kolb's experiential learning cycle. Since the turn of the twentieth century, outdoor education has emerged as a significant kind of education in Western countries, with promising study outcomes. Typical research themes [33] include the environment, outdoor sports and adventure, and environmental-related values. Other common topics include history, culture, energy, and the natural world. When it comes to objective outcomes, outdoor education has traditionally been focused on instilling environmental awareness and values in students, experiencing the terrain and the wild, and learning about natural resources and sustainability [34, 35]. Furthermore, there are several research studies that have been conducted that have focused on techniques of outdoor education. Some researchers have argued that a place-responsive pedagogy might enable teachers to work in school-based outdoor education in new and novel ways, hence facilitating the implementation of more cross-curricular teaching and learning activities on a local level [36]. A comparative study of two approaches of outdoor sports instruction has been carried out by a number of scholars in the field. The researchers discovered that a hybrid school-based approach to outdoor education combined with Moodle outperformed traditional learning environments when they analyzed grades and results. This is due to the fact that outdoor activities bring children into contact with the natural world outside of the school grounds, where they can gain physical knowledge and abilities. Aspects of research methods include conducting empirical research using qualitative and quantitative methods, conducting case studies with one or more outdoor education programs, and collecting data from questionnaires and interviews with stakeholders such as program directors, students, and teachers.

The long-term viability of tour studies is a significant development to monitor in international research. Concerned that the tour study will have a negative impact on sustainable development for local society and the environment, a number of researchers have advocated for the inclusion of sustainability in the tour study project's mission statement and the training of project leaders in the

economic, environmental, and sociocultural aspects of achieving sustainability assessment measures. The findings of a separate study that investigated the relationship between perceived value and travel experience and educational travel enjoyment were mirrored. According to research, the greater an individual's level of engagement, the more likely it is that he or she will comprehend knowledge and society. Furthermore, it has been demonstrated that negotiating the specifics of how to travel, acquiring information about different cultures, visiting new places, and interacting with other travelers are all important factors in facilitating experiential learning through the educational travel process [37].

The goal of research travel courses [38–40] is to assist students in getting out from school, becoming more acquainted with nature and society, and becoming more knowledgeable about the local culture. In contrast to research bases, such as university campuses, museums, wetland parks, and other such establishments, research bases have a specialized kind or specific topic of research resources and do not necessarily contain lodging facilities. Generally speaking, research camps [41, 42] are bigger in scale, have a certain amount of outdoor area, and are sites that may give both research experience and lodging. It is essential to provide assistance for research-related supply services, which are frequently overlooked. The research and study travel base, an emerging product that combines education and tourism, serves the purposes of providing supplementary extracurricular information for local young students and scholars, transmitting local distinctive culture, and fostering patriotic feelings.

This study predicts client demand for research items for product research using a large data clustering approach, which is useful for product research.

3. Research Design

3.1. K-Means Algorithm. The big data-based clustering algorithm proposed in this study is the K value clustering algorithm, referred to as the K-means clustering algorithm. The application of demand forecasting mainly includes data type preprocessing, initial center selection, and K value determination.

In the study product consumer demand database D, it is defined that there are n record sets $T = \{T_1, T_2, \ldots, T_n\} (n \geq 1)$ in total. When it comes to the characteristic attributes of the object, it has two sets of attributes, one called $X = X_c \cup X_d$ and another called $X_c \cap X_d = \phi$. The numerical attribute subset X_d and the character attribute subset X_c are both parts of the attribute set. Each record T consists of m-dimensional properties that are unique to it.

$$T_i = \left(x_1^i, x_2^i, \ldots, x_m^i\right). \tag{1}$$

The numerical attribute similarity distance between records T_i and T_j is defined as X, while the numerical attribute similarity distance between records T_i and T_j is defined as follows:

$$\text{Sim}_d\left(T_i, T_j\right) = \left(\sum_{h=1}^{p} \left|x_1^i - x_1^j\right|^2\right)^{1/2}. \tag{2}$$

The number of numerical characteristics is represented by the letter p among them. $1 \leq p \leq m, i \neq j$.

If T_i and T_j are defined as any two data points, then

$$\text{Sim}_c\left(T_i, T_j\right) = \sum_{h=1}^{q} S\left(x_h^i, x_h^j\right), \tag{3}$$

where q is the number of character qualities that are available. $1 \leq q \leq m, i \neq j$. The similarity distance of its hth character property is represented by the symbol $S(x_h^i, x_h^j)$.

The similarity distance between two points T_i and T_j is denoted by the symbol

$$\text{Sim}\left(T_i, T_j\right) = \text{Sim}_c\left(T_i, T_j\right) + \text{Sim}_d\left(T_i, T_j\right)(i \neq j). \tag{4}$$

Cluster sets are being defined as follows:

$$C = \{C_i\}(i = 1, 2, \ldots, K), \tag{5}$$

where $C_i = \left\{T_f, T_r, \ldots, T_g\right\}$ is the ith cluster containing r entries, where r is the row number.

The cluster center m_i of the C_i clustering may be represented as $m_i = m_i^d + m_i^c$, where the cluster center m_i^d of the numerical attribute takes the value of the average value of the corresponding attribute of the data record as the value of the cluster center m_i^d of the numerical attribute.

$$m_i^d = \frac{1}{r} \sum_{i=f}^{g} \text{Sim}_d\left(T_i, T_j\right). \tag{6}$$

The character attribute cluster center value is the maximum value of the associated attribute frame rate that has been recorded in the cluster for that character attribute.

$$m_i^c = \text{Max}\left(\text{Sim}_c\left(T_i, T_j\right)\right). \tag{7}$$

The similarity between T_i and the current cluster j is given as the distance between the remaining cluster centers m_j, where m_j is the distance between the remaining cluster centers.

$$\text{Sim}\left(T_i, m_j\right) = \text{Sim}_d\left(T_i, m_j\right) + \text{Sim}_c\left(T_i, m_j\right). \tag{8}$$

The shortest possible distance is as follows:

$$\text{Min}\left(\text{Sim}\left(T_i, C\right)\right) = \text{Min}\left(\text{Sim}\left(T_i, C_j\right)\right). \tag{9}$$

The lowest similarity distance between any two clusters C_i and C_j is defined as follows:

$$\text{SBC} = \text{Min}\left(\text{Sim}\left(m_i, m_j\right)\right). \tag{10}$$

Alternatively, the average SWC_i of data item similarity in the C_ith class comprising r data objects may be written as follows:

$$\text{SWC}_i = \text{Avg}\left(\sum_{h=1}^{j-1} \sum_{l=h+1}^{j} \text{Sim}\left(T_h, T_j\right)\right)(h \neq j). \tag{11}$$

As a standard, the highest degree of similarity is determined between the received record and a specific class, and then, the record is categorized. According to the principle of the closest distance, the greatest similarity distance between two classes of objects is as follows:

$$SWC = Max(SWC_i)(i = 1, 2, \ldots, k). \quad (12)$$

The record distribution density function is defined as follows:

$$d_i = \frac{z_i}{\sum_{j=1}^{l} z_j}, \quad (13)$$

$z_i = \sum_{i=1, l \neq 1}^{l} 1/\text{Sim}(T_i, T_j)$. With increasing d_i, the density of sample points increases, and the influence on categorization increases proportionally as well.

New clusters are created, and their respective K values are determined. Through a series of iterations, the minimal similarity distance between classes and the maximum similarity distance within a class are determined. Using a dynamic K value that is modified throughout the clustering process, it is possible to classify objects according to the criterion of being closest within a class and farthest between classes. The algorithm's flow is depicted in Figure 1.

3.2. Data Sources. The research products are based on information gathered from a Xi'an, Shaanxi Province, Tourism Company. To produce research products, we focus primarily on analyzing the study tour's data. From a research standpoint, the study of research travel is primarily based on pedagogical and tourism perspectives, among others. Currently, educational academics who are examining the design of research trip courses for elementary and secondary schools are the primary source of travel research. Not enough research has been conducted on the product, the market, and the operation of research trips. According to tourism researchers, the applied research of research travel should be strengthened and integrated with the perspective of educational research to encourage the growth of research travel practice. Consequently, using big data clustering techniques, we are able to estimate client demand for products related to research.

4. Results

To determine whether the optimization of anticipating the degree of client demand for research items based on the clustering method for big data is effective, the optimal K value of K-means is used for subdivision, or the optimal number of clusters is determined to divide the data into two. It is possible to analyze the clustering effect of a dataset using the silhouette coefficient as an index. Now, beginning with 2, the K value can be evaluated gradually, and after selecting a K value from the set, the K value with the largest silhouette coefficient can be identified. When determining the upper limit of the maximum number of clusters, it is usually unnecessary to select a very high number. The outcome of selecting the number 7 in this scenario is depicted in Figure 2.
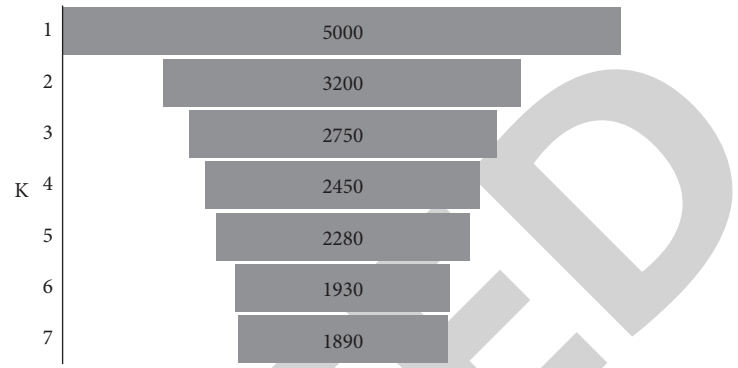

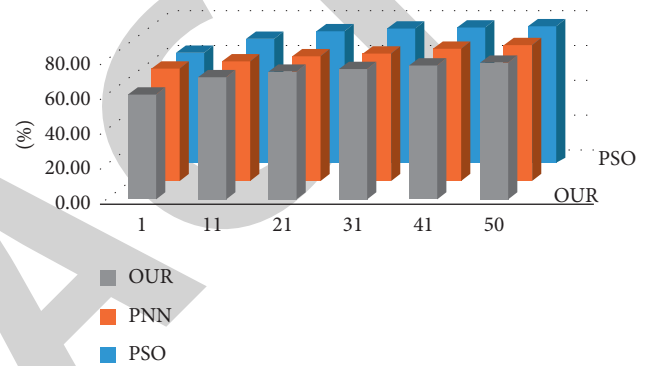Figure 2: Relationship between K and silhouette coefficient.
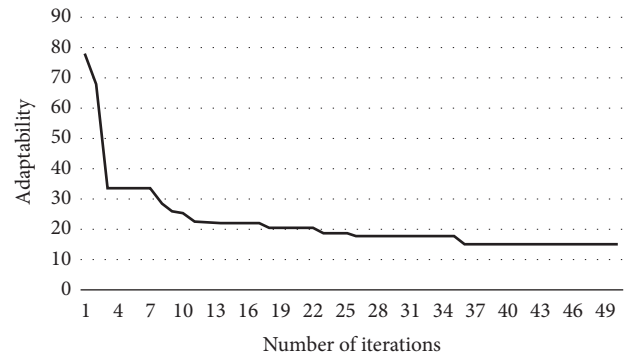

Figure 3: Test set accuracy.

Figure 4: Algorithm optimization process.

As shown in Figure 3, our K-means clustering method has the highest accuracy when compared to particle swarm optimization (PSO) and probabilistic neural networks (PNNs). Similarly, the accuracy rate on the test set is the highest of all tested methods, demonstrating that our model has a high accuracy rate and is able to accurately predict the results. However, the test set accuracy is lower than the training set accuracy, indicating that the model may have been overfitted during training, which is cause for concern.

Figure 4 demonstrates that the model proposed by this research is continuously optimized during the iterative process, resulting in a reasonably satisfactory outcome.

5. Conclusion

Our clustering algorithm is capable of predicting client demand with 78.56 percent degree of precision. Although all existing clustering techniques involve parameter adjustment, this may lead to undesirable overfitting of the data and inadequate generalization of the model. As a result, it fails to identify clusters in the benchmark dataset, and its flaws include a high sensitivity to noise and outliers, an inability to recognize clusters that are not well separated and that are of any form or density, and a lengthy calculation time and high computational complexity. It should be noted that the majority of parallel clustering algorithms presented by certain researchers do not deal with real-time data, but instead focus on a particular type of data, thereby limiting their ability to deal with large amounts of data in general. Consequently, real-time and heterogeneous data processing remains a difficult problem to resolve.

Using our algorithm, it is possible to make optimal predictions about customer demand for research products, to effectively solve the problem that the calculation results of other dynamic clustering methods are dependent on the initial cluster center and the sample input order, and to rapidly converge to the global optimal solution. In the era of big data, it provides a novel approach to multidimensional and massive data analysis.

Data Availability

The data used to support the findings of this study are included in the article.

Conflicts of Interest

The author states that the publishing of this article does not include any conflicts of interest.

Acknowledgments

References

[1] O. Chapelle, B. Scholkopf, and A Zien Eds, "Semi-Supervised Learning (Chapelle, O. et al., Eds.; 2006) [Book reviews]," *IEEE Transactions on Neural Networks*, vol. 20, no. 3, p. 542, 2009.

[2] V. Estivill-Castro, "Why so many clustering algorithms," *ACM SIGKDD explorations newsletter*, vol. 4, no. 1, pp. 65–75, 2002.

[3] E. Aldana-Bobadilla and A. Kuri-Morales, "A clustering method based on the maximum entropy principle," *Entropy*, vol. 17, no. 1, pp. 151–180, 2015.

[4] H. Zha, "Generic summarization and keyphrase extraction using mutual reinforcement principle and sentence clustering [C]," in *Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval*, pp. 113–120, August 2002.

[5] S. Gordon, H. Greenspan, and J. Goldberger, "Applying the information bottleneck principle to unsupervised clustering of discrete and continuous image representations[C]," in *Proceedings of the Computer Vision, IEEE International Conference on. IEEE Computer Society*, vol. 2, p. 370, IEEE, San Juan, PR, USA, November 2003.

[6] S. J. Porth, "Management education goes international: a model for designing and teaching a study tour course," *Journal of Management Education*, vol. 21, no. 2, pp. 190–199, 1997.

[7] S. Orpett Long, Y. S. Akande, R. W. Purdy, and K. Nakano, "Deepening learning and inspiring rigor," *Journal of Studies in International Education*, vol. 14, no. 1, pp. 89–111, 2010.

[8] J. Evans, J. Finch, M. F. Toncar, and J. S. Reid, "Student perceptions of and preferences for A short overseas study tour," *Contemporary Issues In Education Research*, vol. 1, no. 3, p. 11, 2011.

[9] G. Aliyeva, "Impacts of educational tourism on local community: the case of gazimagusa, north Cyprus[D]," *Eastern Mediterranean University (EMU)-Doğu Akdeniz Üniversitesi (DAÜ)*, vol. 5, 2015.

[10] A. A. Samah and M. Ahmadian, "Educational tourism in Malaysia: implications for community development practice," *Asian Social Science*, vol. 9, no. 11, p. 17, 2013.

[11] B. W. Ritchie, *Managing Educational Tourism*, Channel View Publications, Bristol, UK, 2003.

[12] S. Ghosh and S. Kumar, "Comparative analysis of K-means and fuzzy C-means algorithms," *International Journal of Advanced Computer Science and Applications*, vol. 4, no. 4, 2013.

[13] S. S. Yu, S. W. Chu, C. M. Wang, Y. K. Chan, and T. C. Chang, "Two improved k-means algorithms," *Applied Soft Computing*, vol. 68, pp. 747–755, 2018.

[14] L. Bottou and Y. Bengio, "Convergence properties of the k-means algorithms[J]," *Advances in Neural Information Processing Systems*, vol. 7, 1994.

[15] H. S. Park and C. H. Jun, "A simple and fast algorithm for K-medoids clustering," *Expert Systems with Applications*, vol. 36, no. 2, pp. 3336–3341, 2009.

[16] P. Arora, S. Deepali, and S. Varshney, "Analysis of K-means and K-medoids algorithm for big data," *Procedia Computer Science*, vol. 78, pp. 507–512, 2016.

[17] E. Schubert, J. Sander, M. Ester, H. P. Kriegel, and X. Xu, "DBSCAN revisited, revisited," *ACM Transactions on Database Systems*, vol. 42, no. 3, pp. 1–21, 2017.

[18] D. Birant and A. Kut, "ST-DBSCAN: an algorithm for clustering spatial-temporal data," *Data & Knowledge Engineering*, vol. 60, no. 1, pp. 208–221, 2007.

[19] A. Idrissi, H. Rehioui, and A. Laghrissi, "An improvement of DENCLUE algorithm for the data clustering[C]," in *Proceedings of the 2015 5th International Conference on Information & Communication Technology and Accessibility (ICTA)*, pp. 1–6, IEEE, December 2015.

[20] M. Khader and G. Al-Naymat, "An overview of various enhancements of DENCLUE algorithm[C]," in *Proceedings of the Second International Conference on Data Science, E-Learning and Information Systems*, pp. 1–7, Dubai United Arab Emirates, December 2019.

[21] N. D. Dat, V. N. Phu, V. T. N. Tran, V. T. N. Chau, and T. A. Nguyen, "STING algorithm used English sentiment classification in a parallel environment," *International Journal*

of Pattern Recognition and Artificial Intelligence, vol. 31, no. 07, Article ID 1750021, 2017.

[22] J. Handl and J. Knowles, "An evolutionary approach to multiobjective clustering," *IEEE Transactions on Evolutionary Computation*, vol. 11, no. 1, pp. 56–76, 2007.

[23] D. W. Corne, N. R. Jerram, J. D. Knowles, and M. J. Oates, "PESA-II: region-based selection in evolutionary multi-objective optimization[C]," in *Proceedings of the 3rd annual conference on genetic and evolutionary computation*, pp. 283–290, San Francisco California, January 2001.

[24] S. Saha and S. Bandyopadhyay, "A symmetry based multi-objective clustering technique for automatic evolution of clusters," *Pattern Recognition*, vol. 43, no. 3, pp. 738–751, 2010.

[25] S. Bandyopadhyay and S. Saha, "GAPS: a clustering method using a new point symmetry-based distance measure," *Pattern Recognition*, vol. 40, no. 12, pp. 3430–3451, 2007.

[26] C. Ferreira, "Gene expression programming: a new adaptive algorithm for solving problems[J]," 2001, http://arXiv.org/abs/preprintcs/0102027.

[27] A. Fahad, N. Alshatri, Z. Tari et al., "A survey of clustering algorithms for big data: taxonomy and empirical analysis," *IEEE transactions on emerging topics in computing*, vol. 2, no. 3, pp. 267–279, 2014.

[28] A. Sherin, S. Uma, and K. Saranya, "Survey on big data mining Platforms, algorithms and Challenges[J]," *International Journal of Computer Science & Engineering Technology*, vol. 5, pp. 59–65, 2014.

[29] S. Arora and I. Chana, "A survey of clustering techniques for big data analysis[C]," in *Proceedings of the 2014 5th International Conference-Confluence The Next Generation Information Technology Summit (Confluence)*, pp. 59–65, IEEE, Noida, India, Septembe2014.

[30] R. Xu and D. WunschII, "Survey of clustering algorithms," *IEEE Transactions on Neural Networks*, vol. 16, no. 3, pp. 645–678, 2005.

[31] C. Yadav, S. Wang, and M. Kumar, "Algorithm and approaches to handle large Data-A Survey[J]," 2013, http://arXiv.org/abs/1307.5437.

[32] A. S. Shirkhorshidi, S. Aghabozorgi, and T. Y. Wah, "Big data clustering: a review[C]," in *Proceedings of the International conference on computational science and its applications*, pp. 707–720, Springer, Cham, 2014.

[33] P. T. Maher, "International perspectives in outdoor education research editors' notes[J]," *Journal of Outdoor Recreation, Education, and Leadership*, vol. 10, no. 1, pp. 1-2, 2018.

[34] S. Buldur, M. Bursal, N. Yalcin Erik, and E. Yucel, "The impact of an outdoor education project on middle school students' perceptions and awareness of the renewable energy," *Renewable and Sustainable Energy Reviews*, vol. 134, Article ID 110364, 2020.

[35] M. L. Gómez-Ruiz, F. J. Morales-Yago, and M. L. de Lázaro-Torres, "Outdoor education, the enhancement and sustainability of cultural heritage: medieval madrid," *Sustainability*, vol. 13, no. 3, p. 1106, 2021.

[36] J. Mikaels, "Becoming a place-responsive practitioner: exploration of an alternative conception of friluftsliv in the Swedish physical," *Journal of Outdoor Recreation, Education, and Leadership*, vol. 10, no. 1, pp. 3–19, 2018.

[37] D. N. Roberson Jr, "Learning while traveling: the school of travel," *Journal of Hospitality, Leisure, Sports and Tourism Education*, vol. 22, pp. 14–18, 2018.

[38] R. Hatch, "May term travel courses offer faculty a vivid classroom[J]," vol. 6, 2010.

[39] L. E. McGhee, "Influence of race, gender, travel, geography courses, and geography content in other courses upon geographic literacy of high school students[D]," *Memphis State University*, vol. 5, 1991.

[40] V. T. C. Middleton and J. R. Clarke, *Marketing in Travel and tourism[M]*, Routledge, Taylor & Francis, 2012.

[41] K. A. Henderson, M. D. Bialeschki, and P. A. James, "Overview of camp research," *Child and adolescent psychiatric clinics of North America*, vol. 16, no. 4, pp. 755–767, 2007.

[42] S. Thomas, S. Karnik, R. S. Barai, V. K. Jayaraman, and S. Idicula-Thomas, "CAMP: a useful resource for research on antimicrobial peptides," *Nucleic Acids Research*, vol. 38, no. suppl_1, pp. D774–D780, 2010.