Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control
Journal homepage: http://kinetik.umm.ac.id
ISSN: 2503-2267
Vol. 10, No. 2, May, Pp. 207-214

Clustering of high school quality using Fuzzy C-means in the Special Region of Yogyakarta Province

Lilin Rofiqotul Ilmi*[1], Haryanto[2], Maria Sumunaringtyas[1]
Educational Research and Evaluation, Graduate School Program, Yogyakarta State University, Yogyakarta, Indonesia[1]
Electrical Engineering Education, Faculty of Engineering, Yogyakarta State University, Yogyakarta, Indonesia[2]

Article Info
Keywords:
Clustering, Education Report Card, Fuzzy c-means, School Quality

Article history:
Received: December 11, 2024
Accepted: February 25, 2025
Published: May 31, 2025

Cite:
L. R. Ilmi, Haryanto, and M. Sumunaringtyas, "Clustering of High School Quality Using Fuzzy C-Means in the Special Region of Yogyakarta Province", *KINETIK*, vol. 10, no. 2, May 2025.
https://doi.org/10.22219/kinetik.v10i2.2187

*Corresponding author.
Lilin Rofiqotul Ilmi
E-mail address:
lilinrofiqotul.2022@student.uny.ac.id

Abstract
This research aims to reveal the results of clustering high school quality using fuzzy c-means in the Special Region of Yogyakarta Province. This research is quantitative and descriptive. Data collection was conducted through documentation. The research data are secondary data from the 2023 high school education report card. The sample consisted of 51 schools, which were determined using the proportional stratified random sampling. Data analysis was performed using the quantitative descriptive method and fuzzy c-means. The results of the study are clustering on the main indicator data producing three clusters: cluster 1 consists of 11 private schools accredited A and B, cluster 2 consists of 22 public and private schools accredited A, and cluster 3 consists of 18 schools accredited A, B, and C. Cluster 2 excels with the overall best performance, cluster 1 has moderate performance with several areas needing improvement, such as instructional leadership, the use of information technology for budget management, and inclusiveness, and cluster 3 shows the lowest performance, requiring significant attention and improvement in almost all aspects, especially literacy, numeracy, instructional leadership, and the use of information technology for budget management. Cluster 3, which had the lowest performance, showed an urgent need for improvement in almost all aspects.

1. Introduction

The quality of education is one of the main pillars in the progress of a country as it plays a significant role in social, economic and cultural development. The quality of education is the main indicator in determining the quality of human resources (HR) that will drive the country's development [1], [2]. A good quality education promotes personal growth, critical thinking and creativity, all of which are essential for self-development and meaningful contributions to the society [3], [4], [5]. A good quality education is also recognized as a fundamental human right, making it a key focus of global sustainable development efforts [6].

However, despite the government's continuous efforts to improve the quality of education in Indonesia, there are inequalities on education quality in various regions. Education inequality in Indonesia is classified into low education level (50%) and good education level (50%) [7]. A research by Harwanti & Rumiati [8] reveals that only about 54.86% of cities in Indonesia have met the National Education Standards (SNP). This inequality creates a huge gap in the quality of education received by the students, and in turn leads to inequality in opportunities for good quality education.

Although the Special Region of Yogyakarta is known for its great educational tradition and many outstanding schools, inequalities in the quality of education, especially at the senior high school level, remain. Some schools in the Special Region of Yogyakarta have managed to meet the education standards very well, while others are still struggling to achieve these standards. Data from the Department of Education, Youth and Sports (Dikpora) of the Special Region of Yogyakarta in 2023 shows that from a total of 173 senior high schools, 69.24% of schools are accredited A, 22.4% of schools are accredited B, and 3.5% of schools are accredited C. Macro factors affecting the quality of education in Yogyakarta include layout, public education facilities, regulation, and image [9].

One of the ways to assess the quality of education is through accurate and comprehensive data management. Education report card data, which includes indicators such as literacy, numeracy, learning quality, school facilities, and teacher training and experience, among others, provides a clearer picture of the education quality of each school [10], [11]. However, the management and analysis of education data in the Special Region of Yogyakarta is still done manually, which is not only time-consuming but also prone to errors. This manual process interferes clearer identification on quality differences between schools and hinders effective data-driven decision-making.

The issues mentioned above cannot be ignored as they will affect the improvement of education quality. Technology and decision support systems play an important role to ensure that the education data obtained can be effectively analyzed and correctly interpreted to support better and data-driven decisions [12]. Therefore, it is necessary to group the quality of schools based on the Education Report Card which allows a more holistic analysis by utilizing a decision support system algorithm in the form of clustering [13]. Clustering allows for clearer identification of patterns and differences between the schools [14], [15]. Clustering can be used to label unknown data classes, as clustering is often classified as an unsupervised learning method [16]. However, classifications can also be used for datasets with known classes [17]. This is accordance with the Education Report Card, which already has a visible label on 15 main indicators, namely literacy, numeracy, character, teacher training experience, learning quality, reflection and improvement of learning by teachers, instructional leadership, school safety, gender equality, diversity climate, inclusiveness, school community participation, proportion of school resource utilization for quality improvement, utilization of information technology for budget management, and school policies. One clustering technique that is relevant in this context is fuzzy c-means (FCM).

Clustering has several methods, while the most suitable method for clustering education quality is fuzzy clustering. Anbu et al. [18] revealed that fuzzy clustering is one of the widely adopted methods for fault diagnosis due to its advantage in handling its independence from uncertainty. There are several data clustering algorithms that belong to fuzzy groups, one of which is fuzzy c-means (FCM). The advantage of fuzzy c-means over other algorithms is that data is not fully a member of a cluster; the data can be a member of several clusters at the same time and will be assigned to a cluster based on the highest degree of membership of all clusters [19]). In addition, fuzzy c-means has a high level of accuracy and a low error rate in calculations outperforming the k-means algorithm [20], [21], so it is widely applied in various studies for clustering data.

Li's research [22] revealed that FCM was applied in the analysis of teacher allocation in different provinces in China. By clustering the data based on student-teacher ratios, educational backgrounds, and professional degrees of teachers, this study provides recommendations to improve the equity of teacher allocation. Vernanda's research [23] revealed that the fuzzy c-means method effectively grouped 40 schools into three groups, helping the Subang State Polytechnic in determining potential schools for socialization activities and increasing new student admissions. Meanwhile, other research conducted by Az-Zahra [24] revealed that fuzzy c-means helps prospective students in choosing majors at SMK Muhammadiyah 3 Yogyakarta based on the evaluation of their skills, talents, and interests.

Based on the existing research, it appears that although the application of fuzzy c-means (FCM) in education clustering has been widely done, there are still some areas needing further research. One of them is the focus on the application of FCM for clustering high school quality, especially those that utilize a more comprehensive education report card. In this regard, this study aims to cluster the quality of senior high schools (SMA) using fuzzy c-means in the Special Region of Yogyakarta using report cards as the main data. It is expected that this research will result in the clustering of schools based on their education quality level. The results of this study are expected to provide valuable input for the government and other related parties in designing more appropriate and effective policies or programs to improve the quality of education in the Special Region of Yogyakarta. In addition, this research is also expected to open up opportunities for the application of the FCM method in other educational research and become a foundation for further studies that aim to improve the quality of education with data-driven approaches and more in-depth analysis.

2. Research Method

This type of research is a quantitative descriptive that aims to identify, describe, and group the quality of high school based on available data. The data used in the study is secondary data retrieved from the 2023 senior high school education report card, which consists of 15 main indicators, namely literacy, numeracy, character, training experience, learning quality, reflection and improvement of learning by teachers, instructional leadership, school safety, gender equality, diversity climate, inclusiveness, school community participation, proportion of school resource utilization for quality improvement, use of information technology for budget management, and school policies. The research sample consisted of 51 senior high schools (SMA) in the Special Region of Yogyakarta Province which was determined using the proportional stratified random sampling. The data analysis techniques used in this study are quantitative descriptive method and fuzzy c-means. Quantitative descriptive method is used to understand the variation in education quality and explain the deeper context behind the education report card data. Meanwhile, fuzzy c-means was chosen because of its ability to flexibly group the data by taking into account the membership level of each data in the cluster.

The data were analyzed using Rstudio software version 2023.12.1. The process began by loading various *R* packages required for the analysis, such as *pacman::p_load()* which contains packages for clustering, visualization, and statistics. The optimal number of clusters is determined using three methods including the Silhouette method, the Elbow (WSS) method, and the Gap Statistic method, which is displayed with fviz_nbclust(). After obtaining the optimal number of clusters, the clustering is analyzed by Fuzzy c-means (Fanny) using the upfc() function, where clusters were formed with various numbers of clusters. Important parameters in this function include *m* = according to the optimal number of clusters (fuzzy value that controls the membership of data points in the cluster), centers (number of clusters), membership (random membership matrix generated with runif()), iter.max = 100 (maximum iteration limit), and con.val = 1e-05 (convergence condition). The clustering is evaluated by calculating the validation indices such as Silhouette Index, Dunn Index, Partition Coefficient, Modified Partition Coefficient, and Xie-Beni Index, which help assess the quality of the clusters formed. The visualization of the clustering results is performed using clusplot() and fviz_cluster() which show the distribution and cluster structure in the data.

3. Results and Discussion

3.1 Determining the Optimal Number of Clusters

The ideal number of clusters for data depends on the optimal cluster selection process. The methods employed for determining the optimal number of clusters are Silhouette, Elbow, and Statistical Gap methods. Silhouette is an evaluation method that measures how well each piece of data fits into a given cluster and how poorly it fits into other clusters. The silhouette value ranges from -1 to 1, where a positive value indicates that the data matches the selected cluster rather than with other clusters, and a negative value indicates the opposite. The optimal number of clusters is the one that gives the highest silhouette value overall. The Elbow method is used to determine the optimal number of clusters in clustering analysis by plotting the number of clusters against the total value of Within-Cluster Sum of Square (WCSS), also known as Sum of Squared Errors (SSE). This method looks for the point where adding a new cluster is no longer significantly reducing the WCSS, which looks like an "elbow" on the plot. The Gap Statistical method compares the explanation metric of variation for the actual cluster with the expected value of the random distribution. The optimal number of clusters is selected when the Gap Statistics value reaches its peak.

The results of the statistical evaluation to measure the optimal number of clusters based on the Silhouette, Elbow and Statistical Gap methods on the main indicator data are presented in Figure 1.

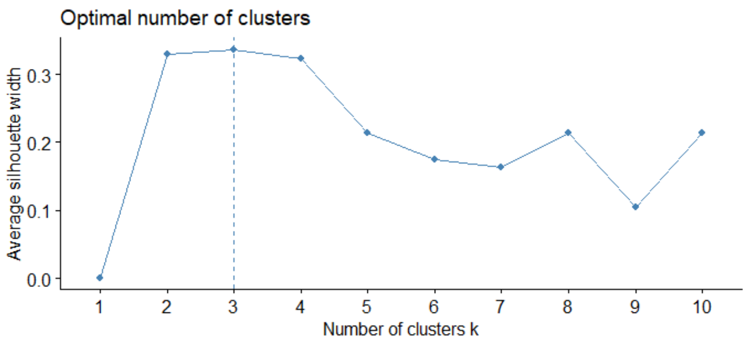

Figure 1. Optimal Cluster Graph of the Silhouette Method

In Figure 1, the graph of the Silhouette method shows that from the experiment on 10 number of clusters, it can be seen that the line that gives the highest silhouette value is at point 3, so the optimal cluster based on the Silhouette method is 3 clusters.

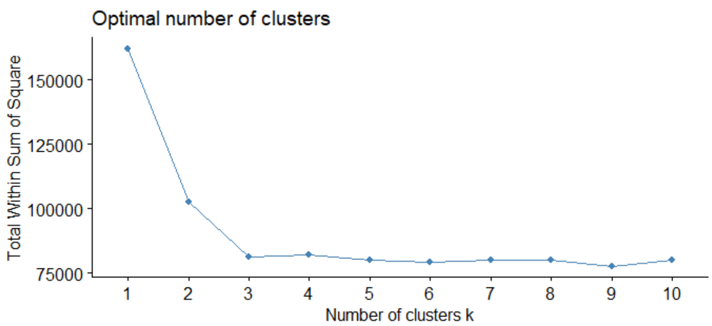

Figure 2. Elbow Method Optimal Cluster Chart

In Figure 2, the graph of the Elbow method shows that from the experiment on 10 number of clusters, it can be seen that the WCSS begins to decrease significantly at point 3, so the optimal cluster based on the Elbow method is 3 clusters.

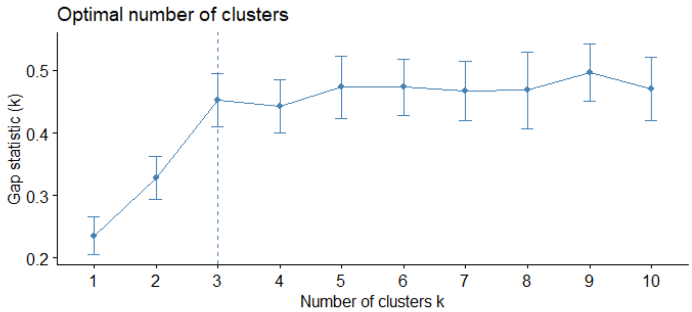

Figure 3. Optimal Cluster Graph Gap Statistical Method

In Figure 3, the graph of the Gap Statistics method from the experiment on 10 number of clusters shows that the highest number of clusters is by using 3 clusters. From these three methods, the optimal number of clusters is equalized, which is as many as 3 clusters.

3.2 Determining the Cluster Center Point (Centroid)

The center point of the cluster (centroid) is the center point or representative of a cluster which is used to describe the average position of all data in the cluster, in this case it is called the center point of school quality from 51 schools. Centroid is used to calculate the variation between clusters which helps in assessing how well the data has been grouped. The analysis of the main indicator data shows that the system divides the center point of the cluster into 3 parts. The three parts have centroids that differ from each other as shown in the Table 1.

Table 1. Center Point of the Key Indicator Cluster

Indicators	Cluster 1	Cluster 2	Cluster 3
Literacy	77.24	94.95	58.46
Numeracy	58.67	86.24	51.99
Character	58.68	62.03	56.40
Teacher Training Experience	72.78	74.20	22.13
Learning Quality	62.40	64.82	61.52
Reflection and Learning Improvement	58.01	59.10	54.79
Instructional Leadership	51.24	55.48	48.34
School Safety	71.86	74.94	69.86
Gender Equality	71.78	73.49	70.28
Diversity Climate	70.86	73.98	69.05
Inclusiveness	56.15	59.19	55.66
School Community Participation	78.25	79.82	75.81
Proportion of Resource Utilization	13.81	24.64	12.47
Utilization of Information Technology	41.83	91.97	50.02
School Policies	74.22	76.08	66.14

The clustering analysis results based on the indicators listed in the table show three clusters with distinct characteristics, which can be described based on the average value of each indicator for each cluster. Cluster 1 has relatively good average scores on most indicators, with the highest values in literacy (77.24) and teacher training experience (72.78). This cluster shows a high level of literacy and good teacher training experience but has lower scores on utilization of information technology (41.83) and proportion of resource utilization (13.81), indicating a deficiency in the use of information technology and resource efficiency. The school policies (74.22) indicator shows a relatively good level of school policy, while school climate-related indicators such as school safety (71.86) and gender equality (71.78) suggest considerable efforts in creating a safe and inclusive environment.

Cluster 2 shows higher scores compared to cluster 1 and cluster 3, particularly on literacy (94.95), numeracy (86.24), and the utilization of information technology (91.97). This suggests that cluster 2 has a very high level of literacy, numeracy, and utilization of information technology, reflecting an environment that strongly supports education and technology. However, the proportion of resource utilization (24.64) is still relatively low, which may indicate an imbalance in resource utilization. Nevertheless, this cluster also scores well on school safety (74.94) and diversity climate (73.98), reflecting a commitment to inclusivity and diversity at school.

Cluster 3 has lower scores than the other two clusters in almost all indicators. The lower scores on literacy (58.46) and numeracy (51.99) indicate challenges in achieving required literacy and numeracy skills. Additionally, the utilization of information technology (50.02) indicates limited use of technology. However, this cluster has better scores on teacher training experience (22.13), suggesting that despite lower literacy and numeracy levels, there may be limited or uneven training opportunities for teacher. Other aspects, such as participation of school community (75.81) and school policies (66.14), show moderate participation and school policies, but there is room for improvement.

Overall, these three clusters represent significant variation in educational indicators and school policies. Cluster 1 tends to have a balance between policy, training, and a safe school environment, but still needs to improve the use of technology and resource efficiency. Cluster 2 shows excellent performance in literacy, numeracy, and the use of information technology, but still requires attention to resource distribution. Meanwhile, Cluster 3 faces significant challenges in literacy, numeracy, and technology although it has positive aspects in the areas of training and simpler school policies. This analysis provides insights that can be used to design more specific and strategic interventions in each cluster to improve the overall quality of education. The method of determining the distance used is square euclidian distance. For example, data number 1 with literacy score data = 95.6, numeracy = 80.0, characters = 59.5, teacher training experience = 43.2, learning quality = 64.4, learning reflection and improvement = 52.1, instructional leadership = 51. 93, school safety = 76.4, gender equality = 69.3, diversity climate = 71.71, inclusiveness = 57.2, proportion of resource utilization for quality improvement = 77.10, IT utilization for budget management = 96.1, and school policy = 63.44 as follows.

D1 = $(95.6 - 77.24)^2 + (80.00 - 58.67)^2 + \ldots + (96.1 - 41.83)^2 + (63.4 - 74.22)^2 = 5401.99$
D2 = $(95.6 - 94.95)^2 + (80.00 - 86.24)^2 + \ldots + (96.1 - 91.97)^2 + (63.4 - 76.08)^2 = 1473.60$
D3 = $(95.6 - 58.46)^2 + (80.00 - 51.99)^2 + \ldots + (96.1 - 50.02)^2 + (63.4 - 66.14)^2 = 5502.23$
the smallest distance value is D2. So that the quality of the first school is included in cluster 2.

3.3 Results of Clustering by Fuzzy C-Means

The results of the fuzzy c-means clustering analysis on the data of 15 main indicators of the education report card from 51 schools conducted using RStudio software with the "ppclust" package showed that the clustering process stopped in the 21st iteration by producing three clusters. Based on the centroid values and membership degrees, cluster 1 includes schools with moderate quality, cluster 2 includes schools of superior quality, and cluster 3 includes schools of low quality. The cluster of superior quality schools (cluster 2) based on the main indicator data are dominated by public schools, while the cluster of moderate quality (cluster 1) and low quality (cluster 3) in the main indicator data are dominated by private schools.

The clustering produces three clusters; cluster 1 consists of 11 schools, cluster 2 consists of 22 schools, and cluster 3 consists of 18 schools. Most of the private schools in cluster 1 are accredited A (8 out of 11 schools), indicating that these schools have high quality standards. In addition to A accreditation, there are also schools with B accreditation (3 schools). This shows that there is a variation in the quality of schools in cluster 1. Unlike the schools in cluster 1, schools in cluster 2 generally have a high quality of education. All schools in cluster 2 have A accreditation, demonstrating a high and consistent quality of education among these schools. Of the 22 schools, 20 are public schools and 2 are private schools. Schools in cluster 2 are spread across various areas in D.I Yogyakarta, including the Cities of Yogyakarta, Bantul, Sleman, Kulon Progo, and Gunungkidul. This shows that the high quality of education is spread evenly in various regions. While the schools in cluster 3 are dominated by private schools with A, B, and C accreditations. Of the 18 schools in cluster 3, 9 schools have A accreditation, while 8 schools have B accreditation, and 1 school has C accreditation.

3.4 Internal Validation of Clustering Results

The internal validation value of the cluster results helps in assessing the quality and effectiveness of the cluster generated by the model. The internal validation of the clustering results in this study uses three types of validation, namely separation index (SI), modified partition coefficient (MPC), and classification entropy (CE) (also known as xie-beni index). Based on the results of the validation of the main indicator data, the cluster model with 3 clusters is quite effective, the SI value of 0.6 shows that the clusters are quite well separated, but not completely isolated. An MPC value of 0.5 indicates that membership in the cluster is moderate, not too assertive or ambiguous. This means that the data in the cluster is not very clear in one particular cluster but is sufficient to provide moderate membership. CE value of 0.1 indicates a very low level of uncertainty, which means clustering is quite good and the data membership in the cluster is clear. Figure 4 are the results of the internal validation of the main indicator data.

```
      Model        SI       MPC        XE
res.upfc3 0.6310364 0.5066077 0.1571503
```

Figure 4. Internal Validation of Key Indicator Data

3.5 Displaying C-Means Fuzzy Visualizations

Fuzzy c-means visualization is used for various important purposes in data analysis and clustering, including to identify how data is divided into different clusters, evaluate how well clusters are by looking at members who are close to each other and far from other clusters, looking at data points that have significant membership in more than one cluster, and other purposes that are closely related to clustering.

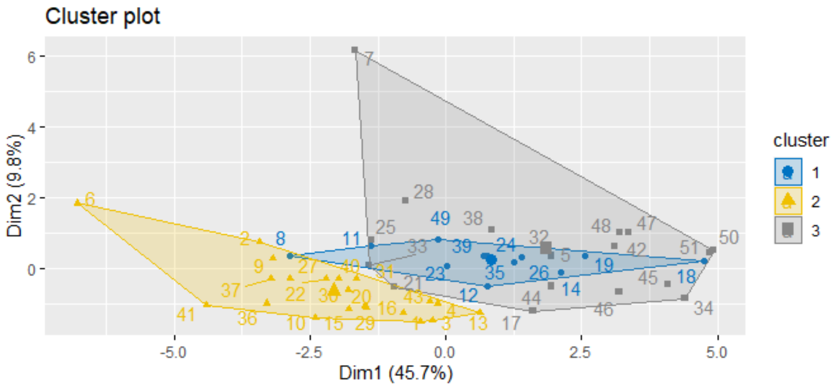

Figure 5. Key Indicator Data Clustering Plot

Figure 5 shows that there are two axes, the first axis (Dim1) explains most of the variance (45.7%), so Dim1 has a significant influence on cluster separation. The second axis (Dim2) explains 9.8% of the variance, so Dim2 makes an additional contribution in the separation but not as much as Dim1.

The data points in cluster 1 represented by the blue color with the circle symbol are mostly to the right of the vertical axis with some spread to the left. The data points in cluster 2 represented by the yellow color with a triangular symbol are spread out on the left side of the plot, with several points adjacent to each other. The data points in cluster 3 are represented by the gray color with the box symbol mostly on the top right of the plot and spread out more vertically than the other clusters. There are several overlapping areas between cluster 1 and cluster 3 around the center of the plot, suggesting similar characteristics in the data, requiring further analysis to ensure the clarity of the cluster separation.

3.6 Discussion

The results of clustering the quality of senior high schools in the Special Region of Yogyakarta using the fuzzy c-means (FCM) method show significant differences between the three clusters formed, reflecting variations in education quality between schools. Cluster 2, which consists of public and private schools with A accreditation, stands out as the most superior with high performance in almost all indicators, including literacy, numeracy and information technology utilization. These findings are in line with previous research by Kumar & Sosnova [25], [26] which states that the use of information technology and adequate teacher training are important factors in improving education quality. Cluster 2, with high scores on the use of technology and good school policies, shows that good management and optimal training play a major role in improving school quality.

Schools in cluster 1, which are dominated by private schools with A and B accreditation, scored adequately on several key indicators such as literacy, supportive school climate and adequate school policies. However, performance on numeracy, character education and teachers' reflection on learning is still at a moderate level. This shows that although the quality of education in cluster 1 is already quite good, there is still room for improvement, especially in numeracy and student character. A study by Jeynes [27] revealed that strong character education in senior high school students is very important to support the development of students' full potential, and this is still a challenge for some schools in Cluster 1.

Cluster 3, which mostly consists of private schools with A and B accreditation, showed the lowest performance, particularly in literacy and numeracy. The lower reading and math skills in this cluster reflect the urgent need for improvement in reading comprehension and math skills. Research by Faiqotusshabrina [28] revealed that low-performing schools often face challenges in improving the quality of students' literacy and numeracy due to low student interest and a lack of supporting infrastructure. Teacher training experience is very low, highlighting the need for significant improvement efforts. While the school climate on safety and gender equality is relatively good, instructional leadership and the use of information technology for budget management still need substantial improvement. School community participation is relatively high but school policies could be further improved. Cluster 3 requires significant attention and improvement in almost all aspects, especially literacy, numeracy, professional teacher training and the use of information technology for budget management. Although some indicators are quite good, this cluster generally shows the lowest performance.

Overall, the clustering results indicate that improvement efforts should be focused on cluster 3 to improve various aspects of education. With a more focused and targeted approach, it is expected that significant improvements can be made in cluster 3, and ultimately improve the quality of education throughout the Special Region of Yogyakarta. This is supported by Lou's research [21] which revealed that the improved fuzzy c-means clustering algorithm significantly improved the evaluation of school quality reaching an accuracy of 95.71% and helped future education development.

4. Conclusion

The conclusion of this discussion shows that clustering the quality of senior high schools in the Special Region of Yogyakarta using the fuzzy c-means (FCM) method produces three clusters that reflect the variation in education quality among schools. Cluster 2 emerged as the leading cluster, with high performance in almost all indicators, reflecting exemplary education standards. Meanwhile, cluster 1, although performing quite well, still has significant room for improvement, especially in the numeracy aspect which is an important element in holistic student development. Cluster 3, which has the lowest performance, shows the urgent need for improvement in almost all aspects, especially in literacy, numeracy and professional teacher training which are crucial for improving the quality of education.

The implications of this study indicate the need to equalize the quality of education in the Special Region of Yogyakarta, with a focus on schools in cluster 3 that show the lowest performance. The government and education office are expected to design policies that strengthen teacher training, especially in literacy and numeracy, increase the use of technology in school management and provide adequate infrastructure. Further research could explore the changing trends of education quality in the Special Region of Yogyakarta by analyzing data from previous years. In addition, to improve accuracy and gain a more diverse perspective, further research should compare Fuzzy c-means with other clustering methods to determine the most effective approach in describing the quality of education in this region. Therefore, the results of this study are not only relevant, but also very important in improving and developing the quality of education in the Special Region of Yogyakarta in a more focused and sustainable manner.

References

[1] Muli Prima Aldi, “Aktualisasi manajemen peningkatan mutu berbasis madrasah (mpmbm),” *Al-Ihda' : Jurnal Pendidikan dan Pemikiran*, vol. 13, no. 2, pp. 178–189, Feb. 2022. https://doi.org/10.55558/alihda.v13i2.16

[2] G. Chiy Ignatius, Pr. R. S. Nkeck Bidias, and Pr. D. Eleanor, “Influence of Human Resource Management on Quality of Education in Public Primary Schools in Cameroon.,” *International Journal of Scientific Research and Management*, vol. 10, no. 06, pp. 2409–2430, Jun. 2022. https://doi.org/10.18535/ijsrm/v10i6.el02

[3] Naldo Janius, Mohammad Aniq Bin Amdan, Mohamad Aidil Hazidi Bin Kasdiah, Nur Izzah Binti Harifin, and Fatin Asyiera Nabila Binti Shamshol Bhari, “Educational opportunity and the purpose of schooling as self-development: A philosophical and sociological inquiry,” *International Journal of Science and Research Archive*, vol. 12, no. 2, pp. 2157–2166, Aug. 2024. https://doi.org/10.30574/ijsra.2024.12.2.1513

[4] K. Liu, “Exploring Responses to Quality Education and Teaching,” *Journal of Education and Educational Research*, vol. 5, no. 2, pp. 44–46, Sep. 2023. https://doi.org/10.54097/jeer.v5i2.12210

[5] H. Guo, “Research on the Construction of the Quality Evaluation Model System for the Teaching Reform of Physical Education Students in Colleges and Universities under the Background of Artificial Intelligence,” *Sci Program*, vol. 2022, 2022. https://doi.org/10.1155/2022/6556631

[6] S. Alifah, “Peningkatan kualitas pendidikan di indonesia untuk mengejar ketertinggalan dari negara lain,” *Jurnal Penelitian*, vol. 5, no. 1, p. 114, Jul. 2021. https://doi.org/10.36841/cermin_unars.v5i1.968

[7] D. A. Suryawan. S and T. Sugiyarto, “Inequality Of Education and Determines of It In Indonesia,” *Seminar Nasional Official Statistics*, vol. 2022, no. 1, pp. 1–10, Oct. 2022. https://doi.org/10.34123/semnasoffstat.v2022i1.1105

[8] N. A. S. Harwanti and A. T. Rumiati, “Pengelompokkan mutu sekolah dasar di indonesia berdasarkan standar nasional pendidikan dengan metode fuzzy c-means,” *Jurnal Sains dan Seni ITS*, vol. 9, no. 2, 2021.

[9] H. Zudianto, “Kearifan Lokal Penguasa Daerah Menghadapi Liberalisasi Pendidikan ‘Upaya Kota Yogyakarta Meningkatkan Kualitas Pendidikan,’” *Unisia*, vol. 29, no. 60, pp. 143–146, Jul. 2006.

[10] W. Wartoni, “The Urgency of Educational Reports for Data Based Planning,” *Jurnal Bahasa dan Sastra*, vol. 12, no. 1, Aug. 2024. https://doi.org/10.60155/jbs.v12i1.449

[11] S. Sundakir, M. Makhdalena, and H. Hambali, “Efforts to Improve the Quality of Schools Based on the Analysis of Educational Reports,” *AL-ISHLAH: Jurnal Pendidikan*, vol. 15, no. 4, Dec. 2023. https://doi.org/10.35445/alishlah.v15i4.3714

[12] K. Nahar, B. I. Shova, T. Ria, H. B. Rashid, and A. H. M. S. Islam, “Mining educational data to predict students performance,” *Educ Inf Technol (Dordr)*, vol. 26, no. 5, pp. 6051–6067, Sep. 2021. https://doi.org/10.1007/s10639-021-10575-3

[13] Fikastiana Cahya, Theresia Wati, and Erly Krisnanik, “Perancangan sistem informasi pengolahan data akademik pada pendidikan anak usia dini berbasis website,” *Journal of Applied Computer Science and Technology*, vol. 2, no. 1, pp. 49–58, Jun. 2021, doi: 10.52158/jacost.v2i1.137.

[14] Mrs. B. J. Et. al., “Comparative Study and Analysis of Students results using Clustering Techniques,” *INFORMATION TECHNOLOGY IN INDUSTRY*, vol. 9, no. 2, pp. 835–842, Apr. 2021. https://doi.org/10.17762/itii.v9i2.421

[15] V. Lakhno, S. Voloshyn, S. Mamchenko, O. Kulynich, and D. Kasatkin, “Cluster Analysis for Researching Digital Footprints of Students in Educational Institutions,” *Cybersecurity: Education, Science, Technique*, vol. 3, no. 23, pp. 31–41, 2024. https://doi.org/10.28925/2663-4023.2024.23.3141

[16] M. A. Muslim *et al.*, *Data Mining Algoritma C4.5*, 1st ed. 2019.

[17] Suyanto, *Data Mining Untuk Klasifikasi Dan Klasterisasi Data*, 1st ed. Bandung: Informatika Bandung, 2019.

[18] Anbu, Thangavelu, and Ashok, “Fuzzy c-means based clustering and rule formation approach for classification of bearing faults using discrete wavelet transform,” *Computation Journal*, vol. 7, no. 4, Sep. 2019. https://doi.org/10.3390/computation7040054

[19] F. Anggoro *et al.*, “Revisiting cluster vulnerabilities towards information and communication technologies in the eastern island of indonesia using fuzzy c means,” *Sustainability*, vol. 14, no. 6, p. 3428, Mar. 2022. https://doi.org/10.3390/su14063428

[20] A. Murugan, D. Gobinath, S. G. Kumar, B. Muruganantham, and S. Velusamy, "A time efficient and accurate retrieval of range aggregate queries using fuzzy clustering means (fcm) approach," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 10, no. 1, p. 415, Feb. 2020. https://doi.org/10.11591/ijece.v10i1.pp415-420

[21] Y. Lou, "Research on Quality Evaluation of Smart Education Schools Based on Fuzzy Clustering Algorithm and Strategies for Future Development," *Applied Mathematics and Nonlinear Sciences*, vol. 9, no. 1, Jan. 2024. https://doi.org/10.2478/amns-2024-1602

[22] Z. Li, S. Chen, L. Pei, J. Chu, and J. Song, "Teacher Allocation and Evaluation Based on Fuzzy C-Means Clustering," *Mobile Information Systems*, vol. 2022, pp. 1–9, Jun. 2022. https://doi.org/10.1155/2022/8465713

[23] D. Vernanda, N. N. Purnawan, and T. H. Apandi, "School Clustering Using Fuzzy C Means Method," *SinkrOn*, vol. 4, no. 1, p. 95, Sep. 2019. https://doi.org/10.33395/sinkron.v4i1.10168

[24] R. R. Az-Zahra, "Analisi Perbandingan Metode Self Organizing Map dan Metode Fuzzy C-Means Pada Pengelompokkan Pemintaan Jurusan Di Sekolah Menengah Kejuruan," *MULTITEK INDONESIA*, vol. 16, no. 2, pp. 65–74, Dec. 2022. https://doi.org/10.24269/mtkind.v16i2.5603

[25] P. Kumar, "Role of ict in teacher training education," *Journal of Teacher Education and Research*, vol. 17, no. 01, pp. 29–30, Jun. 2022. https://doi.org/10.36268/JTER/17106

[26] M. Sosnova, "Information technology as a factor for improving the training of modern teachers," *Dnipro Academy of Continuing Education Herald. Series: Philosophy, Pedagogy*, no. Vol. 2 No. 2 (2023), pp. 40–43, Dec. 2023. https://doi.org/10.54891/2786-7013-2023-2-9

[27] W. H. Jeynes, "A meta-analysis on the relationship between character education and student achievement and behavioral outcomes," *Educ Urban Soc*, vol. 51, no. 1, pp. 33–71, Jan. 2019. https://doi.org/10.1177/0013124517747681

[28] N. Faiqotusshabrina, M. Fita Asri Untari, and B. Ardi Saputro, "Analisis pengetahuan dan kecakapan literasi numerasi siswa dalam pembelajaran operasi hitung bilangan bulat siswa kelas iv sekolah dasar," *Pendas : Jurnal Ilmiah Pendidikan Dasar*, vol. 8, no. 1, pp. 4580–4595, Jun. 2023. https://doi.org/10.23969/jp.v8i1.8043