Published online on: https://journal.uii.ac.id/jurnalsnati/

DOI: 10.20885/snati.v3.i3.40

Jurnal Sains, Nalar, dan Aplikasi Teknologi Informasi

Vol. 3 No. 3 (2024) 121 - 128 ISSN Media Electronic: 2807-5935

Clustering Junior Schools in Implementing Smart School Using The K-Means in Pekanbaru

Aida Nisa[1], M. Khairul Anam[2*], Helda Yenni[3], Parlindungan Kudadiri[4], Gunadi[5]
[1,3,4,5]Department of Informatics, Universitas Sains dan Teknologi Indonesia, Pekanbaru, Indonesia
[2]Department of Informatics, Faculty of Engineering, Universitas Samudra, Langsa, Indonesia
[1]1710031802175@stmik-amik-riau.ac.id, [2]khairula210@gmail.com, [3]heldayenni@sar.ac.id, [4]parlindungan@sar.ac.id, [5]gunadi@sar.c.id

Abstract

The purpose of this research is to determine the readiness of schools in implementing the Smart School system through various stages. One of the concepts of a Smart City involves integrating information and communication technology into the learning process at every school to create Smart Schools. However, not all schools are ready to implement this technology because it requires suitable technology to support the quality of teaching and learning. Another issue is the absence of information systems that can facilitate administrative tasks and the teaching and learning process. The use of the K-Means method is beneficial for clustering schools based on their stages, characteristics, and readiness to implement the Smart School system. This helps identify schools with the highest level of readiness. This research demonstrates that the use of K-Means can identify school readiness based on the established stages related to the Smart School system. It also can pique students' interest in developing and boosting the school's reputation as the best technology-based school.

Keywords: Smart Schools, K-Means, Information Technology, Clustering

1. Introduction

In the last two decades, the concept of the Smart City has gained popularity nearly everywhere, prompting innovative thinking on how to develop Smart Cities [1]. The Institute of Electrical and Electronics Engineers (IEEE) defines a Smart City as a city that integrates technology, government, and the community to build characteristics such as smart governance, smart economy, intelligent mobility, smart environment, smart people, smart living, and smart learning environments (SLE) [2].

A Smart City is one where digital technology is embedded in all services and functions. Developing smart education and intelligent learning environments are key aspects of making a city smart [3]. The advancement of technology in Smart Cities has expanded the concept of being "smart" beyond devices to encompass various systems or structures. The concept of a smart city suggests that such an urban arrangement can facilitate rapid and precise information access for the populace. It is presented as a solution for efficient resource management [4].

Machine Learning is one branch of Artificial Intelligence that focuses on building systems based on data [5]. With the rapid advancement of information technology, especially within the vision of the fourth industrial revolution, it has had an impact on the field of education [6]. Hence, educators need to prioritize introducing their students to current technologies, one of which is Machine Learning [6]. Based on this, building and strengthening a Smart City from its education sector is a way to achieve sustainability in all its structures [7].

Smart Schools are part of these Smart City projects, supported by the government with the aim of improving the education system to meet national educational philosophies and prepare a workforce capable of facing various challenges [8]. Smart Schools have three main dimensions of information: technology, humans, and communities that are interconnected to create an intelligent learning environment. Information technology encompasses hardware, internet network infrastructure, and software services [9].

However, all these efforts use different information technology infrastructure (platforms), ideas, and teaching materials. To drive the success of Smart School development, an Information and Communication Technology infrastructure framework must be developed [10]. Therefore, the success factors of intelligence at each Smart School level can vary and depend on technology components, including intelligence aspects related to all school activities, including software, hardware, network infrastructure,

Received: 19-05-2024 | Revised: 27-06-2024 | Accepted: 28-06-2024 | Published: 05-07-2024

and security facing the Smart School network and in line with educational globalization [11].

The role of technology components also stems from the selection of the appropriate method [12]. Some datasets with specific forms require specialized solving methods adapted to their data conditions [13]. In this research, the method used is K-Means Clustering, which allows data to be grouped into several clusters [14]. K-Means Clustering is a method that processes information or objects by grouping them so that each group contains information that is as similar as possible or potentially different from objects in other groups [15]. Related research on K-Means Clustering conducted by [16] resulted in the identification of three optimal clusters. Subsequently, data from the K-means application, involving 66 students and 9 iterations, resulted in 3 clusters: "ready" with 7 students, "fairly ready" with 30 students, and the remaining 29 students "not ready." Additionally, research by [17] found that K-Means Clustering can determine the grouping of student performance into high, medium, and fair categories. Following this, the grouping of Covid-19 vaccination targets was done using Spectral Clustering. The categories were grouped into three groups: high, medium, and low. The evaluation of cluster results showed nearly identical values, indicating that the members of a cluster in each category were not significantly different [18]. Another study conducted character clustering in the Wild Rift game using K-Modes Clustering. The results of the K-modes calculations were evaluated using the Davies-Bouldin Index validation, which yielded a DBI value of 0.8200, indicating that the K-modes algorithm achieved optimal clustering [19]. K-means clustering has several advantages compared to K-modes and Spectral Clustering. K-means is faster and more efficient for large datasets, suitable for numerical data, and has intuitive interpretation. While K-modes is better for categorical data and Spectral Clustering can handle complex cluster shapes, K-means has lower computational time complexity and is easier to apply to very large datasets.

K-Means is used in this research due to its efficient and straightforward clustering capability in machine learning [20]. his method allows for the grouping of data into several clusters based on certain similarities [21]. In the context of this research, K-Means is used to cluster schools based on the similarity of their readiness criteria in implementing the Smart School system. One of the main advantages of K-Means is that it falls under the category of unsupervised learning, meaning it does not require labeled training data, which is very useful in cases where labeling data is difficult or impractical. Additionally, K-Means can be easily adjusted to various types of data and clustering needs [22]. Researchers can select the desired number of clusters and adjust other parameters to obtain optimal clustering results.

Another advantage of K-Means is that this algorithm is relatively easy to implement and understand, whether using software like Microsoft Excel or programming languages like Python, facilitating the data analysis process and interpretation of results. K-Means is highly scalable for large datasets, allowing the algorithm to handle large amounts of data quite quickly and efficiently compared to other clustering methods. Furthermore, K-Means produces consistent and reliable results with sufficient iterations, making it easier to interpret the results and make decisions based on the formed clusters. This algorithm internally optimizes the position of cluster centroids to minimize the total within-cluster distance, resulting in more accurate and relevant clustering. K-Means also works well on data with many dimensions or features, allowing for the analysis of several attributes such as Utilization, Human Capital, Applications, and Technology Infrastructure simultaneously to determine school readiness. By grouping data into fewer clusters, K-Means helps in data simplification and facilitates the visualization and interpretation of complex information.

Data collection in this research was conducted through direct observation and interviews with school representatives. The collected data includes various attributes relevant to school readiness in implementing the Smart School system, such as the use of information and communication technology, the quality of human resources, the applications used, and the technological infrastructure. This data is then analyzed using the K-Means clustering method.

The use of various analytical tools such as RapidMiner, Microsoft Excel, and Python in this research aims to ensure accurate and reliable results. RapidMiner is used for more interactive and intuitive data visualization and modeling. Microsoft Excel is used for manual calculation processes and clustering result validation, while Python, with various libraries such as Pandas, K-Means, and Matplotlib, is used for automating the calculation processes and data visualization. The use of these three tools allows researchers to validate the consistency of clustering results and ensure that each method produces similar results, which strengthens the validity and reliability of this research's findings.

Overall, the use of K-Means in this research aims to obtain a clear picture of school readiness in implementing the Smart School system. The advantages of K-Means make it the right choice for this clustering analysis.

2. Research Methods

Research Methodology is a technique devised by researchers to gather data or information in conducting research that aligns with the subject or object under investigation. The presence of such data is expected to yield high-quality results. Figure 1 illustrates the flow

of this research methodology:

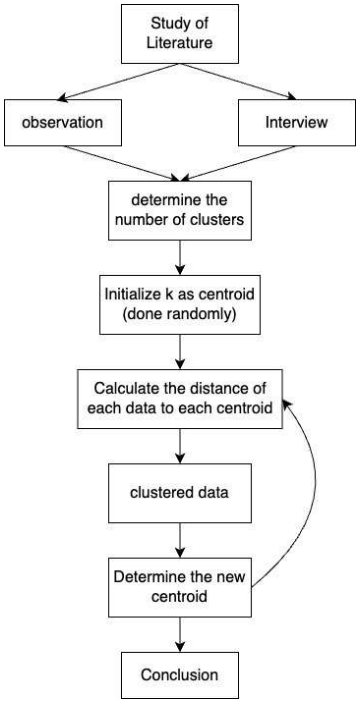

Figure 1. Research methodology

2.1. Literature Review

Literature Review utilized in this study involves the collection of data and information related to technology architecture and facilities, especially within organizations operating in the field of education, from websites, journals, and books related to the research topic. The data collected by the researcher includes concepts or frameworks related to the implementation of Smart Schools. The concept of smart schools has been previously explored in schools in Iran and Malaysia. This research also intends to apply the same concept of smart schools to schools in the city of Pekanbaru. Below is an explanation of the concept or framework from previous research [23].

a. People (Individuals)
   Smart Schools provide full access to students and teachers. Teacher training is a crucial activity that should be carried out at the inception of smart schools.
b. Technology
   Smart schools are expected to produce a workforce that thinks critically and is tech-savvy, thus preparing the workforce for the information age.
c. Process
   In addition to technology, teaching and learning strategies, management and administrative processes, and trained personnel are prerequisites for smart schools.
d. Education Materials
   The curriculum, subject content, and teaching activities play a vital role in Smart Schools.

2.2. Clustering Data

The title should be clear, simple, concise and informative, describe the content of the article, contain a maximum of 12 words, use a font size of 15pt, not be bold and only the first letter of the word should be capitalized. The minimum title contains result, problem and method. Spacing between paragraphs 6pt.

In this stage, data from previous observations and interviews are utilized. The clustering process begins with a manual application of K-Means, which involves initial testing of attributes in the clustering process. These attributes include Information and Communication Technology and possess characteristics such as Utilization, Human Capital, Applications, and Technology Infrastructure. Each attribute is tested individually to observe differences in the results.

As stated by [24], these characteristics are intended to demonstrate not only a school's ability to perform general Information and Communication Technology operations but also its capability to foster creativity, interactivity, collaborative learning, critical thinking, and problem-solving skills. In addition to monitoring school's Information and Communication Technology usage, these characteristics are also used to assess schools based on performance indicators, as follows [25].

a. *Utilization*
   Relevant indicators include: student-to-PC class hours, integration of Information and Communication Technology subject content by teachers for the primary focus of lessons, School Management System updates, usage of School Management System, and completion of self-learning materials by students.
b. *Human Capital*
   Indicators related to this category include: Information and Communication Technology competence of subject teachers, Information and Communication Technology utilization in information dissemination, relationships, multimedia usage in teaching, student competence, and awareness of the availability of educational support equipment.
c. *Applications*
   Indicators encompass: a minimum of 5 modules used for school management, equipment used for teaching, website presence, and maintenance.
d. *Infrastruktur Teknologi*
   Indicators include: PC-to-classroom ratio, PC accessibility, LAN & WAN accessibility.

Automatic calculation processes are carried out using the Python programming language. Subsequently, the data calculations are stored in an Excel format with a portion of the data processed manually. These

processed data are then further analyzed using the RapidMiner application to make informed decisions.

In the next stage, data with defined parameters are visualized using the Python programming language and various libraries or modules, including Pandas, Kmeans, and Matplotlib.pyplot. Pandas is employed to read the data types used as datasets [26]. Kmeans is one of the clustering algorithms used in this process [27]. Matplotlib.pyplot is used for data visualization tasks, such as creating graphs, generating area plots within graphs, adding labels to plots, and more. Pandas is used for reading the data types used as datasets. %matplotlib inline is used to display inline output. Additionally, data is standardized (centered at 0) and scaled to a variance of 1, assuming a normal distribution (Gaussian) for all features.

2.3. Conclusion Stage

Based on the research objectives regarding the implementation of Smart School in schools within the city of Pekanbaru using the K-Means clustering method, we can draw conclusions regarding the clustering results of schools categorized as "Ready" and "Not Ready" to implement Smart School. This categorization is based on the similarity of criteria related to Utilization, Human Capital, Applications, and Technology Infrastructure.

3. Results and Discussions

The results obtained during the conducted research, based on data collection, include obtaining information from each school, which can then be analyzed using K-Means Clustering to determine schools that are ready and not ready to implement Smart School.

3.1. Clustering Results Using Microsoft Application

The application of the K-Means Clustering method in this research was performed using the Microsoft Excel application. The required steps are as follows:

1. Determining the Number of Clusters

The number of clusters used in the school data was determined to be 2 clusters.

2. Determining the Initial Cluster Centroid Values

Cluster centroids, which serve as reference points for calculating the distance between data points and each cluster, were established. In this process, the initial centroid values can be determined randomly at the researcher's discretion, provided that the centroid values fall within the range of data values for each attribute. Alternatively, initial centroid values can also be determined using the average values of each attribute.

In this process, the researcher determined the initial cluster centroid values using the values from the first and fourth variables. These attributes were used to determine the initial cluster centroid values.

3. Calculate the Data Distance to each Cluster

Once the initial cluster centroids have been determined, the next step is to calculate the distance for each cluster. In calculating the distance values for each cluster, the researcher used the Euclidean Distance formula. The formula for Euclidean Distance is as follows:

$$\sqrt{(x_{1i} - x_{1j})^2 + (x_{2i} - x_{2j})^2 + (x_{3i} - x_{3j})^2 + (x_{4i} - x_{4j})^2} \quad (1)$$

Information:

Table 1. Starting Center Point of Each Cluster

Data	Cluster	Utilization	Competence	Application	Technology
2	0	22	27	11	27
4	1	17	22	9	24

Table 2. Cluster Calculating the Data Distance to Each Cluster

School Data	Utilization	Competence	Application	Technology	C0	C1
SMP NEGERI 40 PEKANBARU	27	27	12	27	6	53
SMP NEGERI 25 PEKANBARU	22	27	11	27	0	43
SMP NEGERI 23 PEKANBARU	23	23	12	17	118	65
SMP IT AZIZIYYAH PEKANBARU	17	22	9	24	43	0

a) x_{1i} represents Data point i in attribute 1.

b) x_{1j} represents Centroid j in attribute 1.

4. Grouping Data into Clusters

Following the distance calculation process for each data point to every cluster, the next step is the data assignment process to each formed cluster. This data assignment aims to determine the new cluster centers in the subsequent process. Data is grouped into clusters based on the smallest distance values for each data point.

Table 4. Grouping Data into Clusters 1st Iteration

ITERATION 1							
School Data	Utilization	Competence	Application	Technology	C0	C1	Data Group
SMP Negeri 40 Pekanbaru	27	27	12	27	6	53	Cluster0
SMP Negeri 25 Pekanbaru	22	27	11	27	0	43	Cluster0
SMP Negeri 23 Pekanbaru	23	23	12	17	118	65	Cluster1
SMP IT Aziziyyah Pekanbaru	17	22	9	24	43	0	Cluster1

From the results of the above calculations, a new cluster center and 1st iteration data grouping are obtained as in the table below:

Table 3. New Centroid 1st Iteration

Cluster	Utilization	Competence	Application	Technology
0	24,5	27	11,5	27
1	20	22,5	10,5	20,5

Table 5. Grouping Data into Clusters 2nd iteration

School Data	Utilization	Competence	Application	Technology	C0	C1	Data Group
SMP Negeri 40 Pekanbaru	27	27	12	27	2,75	71,75	Cluster0
SMP Negeri 25 Pekanbaru	22	27	11	27	2,75	64,75	Cluster0
SMP Negeri 23 Pekanbaru	23	23	12	17	117,75	17,75	Cluster1
SMP IT Aziziyyah Pekanbaru	17	22	9	24	47,75	17,75	Cluster1

The next iteration process performs calculations again using the new centroid from the results of the 2nd iteration.

Because in the 1st and 2nd iterations the cluster position did not change, the iteration was stopped.

3.2. Clustering Results Using the Python Programming Language

1. Import Library

The first thing to do is import several Python libraries for dataframe, visualization and clustering needs, with the following steps:

```
[12] from sklearn.cluster import KMeans
     import pandas as pd
     from sklearn.preprocessing import MinMaxScaler
     from matplotlib import pyplot as plt
     %matplotlib inline
```

Figure 2. Importing libraries

2. Import and Reading Data

The next step is to input the dataset. The dataset to be utilized is a school dataset. This dataset comprises a collection of school data obtained through observations and interviews. The following are the steps for inputting the dataset and examining detailed information about the dataset.

	Data Sekolah	Pemanfaatan	Kompetensi	Aplikasi	Teknologi
0	SMP NEGERI 40 PEKANBARU	27	27	12	27
1	SMP NEGERI 25 PEKANBARU	22	27	11	27
2	SMP NEGERI 23 PEKANBARU	23	23	12	17
3	SMP IT AZIZIYYAH PEKANBARU	17	22	9	24

Figure 3. Data results displayed

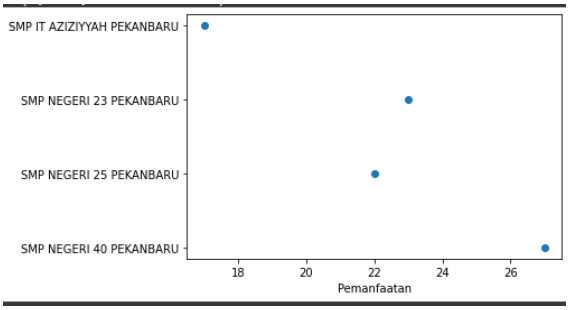

Figure 4. Visualization of Data Distribution in the 1st Attribute

Based on the output above, it can be concluded that the dataset has 5 attributes, including: School Data, Utilization, Competency, Application, Technology.

3. Data Visualization Results for Each Attribute

After successfully importing the dataset file, we now have data where each row contains values. This data is what we will use for clustering. To visualize the data

distribution before performing clustering, we will first create a scatter plot using the following steps:

Based on the visualization results above, the blue color is the point in each data. Apart from that, the visualization of each data looks good, and the data is not close to each other.

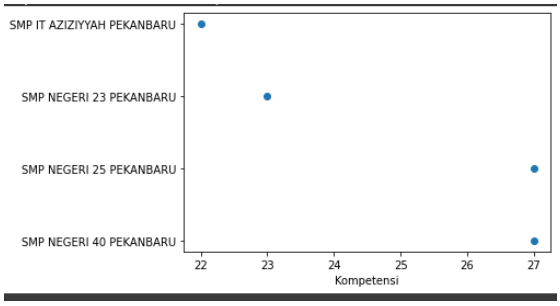

Figure 5. Visualization of data distribution on the 2nd attribute

Next, in the visualization results in Figure 5, the blue color is the point in each data. Apart from that, the visualization of each data looks good, and the distance between the last two data looks close together.

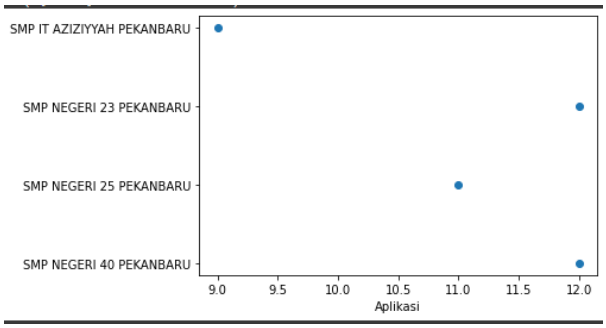

Figure 6. Visualization of data distribution on the 3rd attribute

Next, the visualization results in Figure 6 show that the blue color represents points for each data point. Furthermore, the visualization of each data point appears clear, and there is also a noticeable distance between the data points.

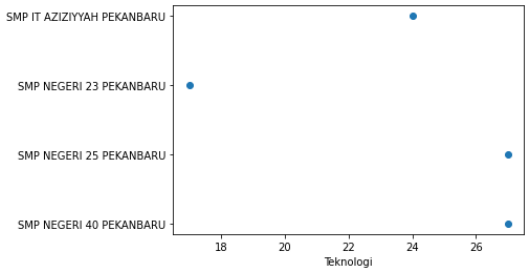

Figure 7. Visualization of data distribution on the 4th attribute

Based on the visualization, it is evident that the data has not yet been categorized into distinct groups. This data distribution will be segmented into multiple categories in our subsequent analysis, enabling researchers to identify which points fall into each of these categories. Subsequently, a clustering process is conducted using the Python programming language.

	Data Sekolah	Pemanfaatan	Kompetensi	Aplikasi	Teknologi	cluster
0	SMP NEGERI 40 PEKANBARU	1.0	1.0	1.000000	1.0	0
1	SMP NEGERI 25 PEKANBARU	0.5	1.0	0.666667	1.0	0
2	SMP NEGERI 23 PEKANBARU	0.6	0.2	1.000000	0.0	1
3	SMP IT AZIZIYYAH PEKANBARU	0.0	0.0	0.000000	0.7	1

Figure 8. Clustering Results Using Python

Based on the results obtained from the clustering experiment with attribute data using the K-Means algorithm, it appears that there is no significant difference between the clustering outcomes. However, the number of iterations required for the K-Means algorithm in Python is fewer compared to using Microsoft Excel. From the grouping results above, it can be observed that there are no further changes in the members of each cluster. Therefore, the iteration process is terminated at this point. It can be concluded that Cluster 0 (1) contains 2 schools, meaning that 2 schools fall into the category of prepared schools, while in Cluster 1 (2), there are also 2 schools, signifying that 2 schools belong to the category of unprepared schools.

Next, to obtain a comparison, we calculate the SSE (Sum of Square Error) for each cluster value. SSE is used to test the ideal number of clusters. As the number of clusters K increases, the SSE value decreases. The formula for SSE in K-Means is as follows.

$$SSE = \sum_{i=1}^{m} (y_i - \hat{y_i})^2 \quad (2)$$

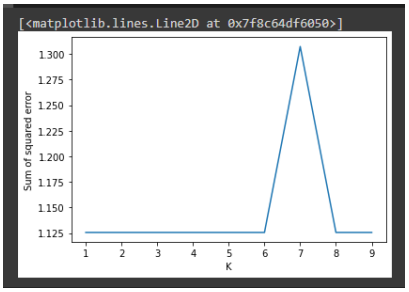

Figure 9. Sum of Squared Error graph

After evaluating the quality of the formed clusters based on the aforementioned approach, the optimal cluster configuration is obtained at k = 7. This is because the value at k = 7 is the highest compared to the others.

3.3. Clustering Results with RapidMiner

The results of school data clustering with RapidMiner are as follows:

Data view is used to display the processed data in its entirety, along with its clusters. The data view can be seen in the image.

Data Sekolah	cluster	Pemanfaatan	Kompetensi	Aplikasi	Teknologi
SMP NEGERI...	cluster_0	27	27	12	27
SMP NEGERI...	cluster_0	22	27	11	27
SMP NEGERI...	cluster_1	23	23	12	17
SMP IT AZIZIY...	cluster_1	17	22	9	24

Figure 10. Overall Results Data that has been processed

Float view to display data that has been processed as a whole complete with its clusters in the form of a Scatter diagram. The appearance can be seen in the following image.

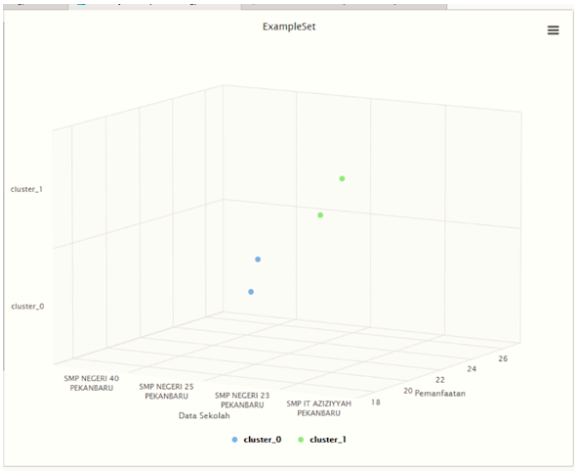

Figure 11. Cluster results in the form of a Scatter diagram

Text view is a sheet to display the database that has been processed as a whole complete with its clusters. The display of the cluster model (clustering) can be seen in the following image.

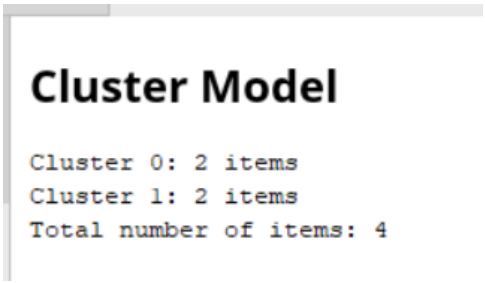

Figure 12. Cluster Model

Within the cluster model, the results of the clustering process stages are assigned to each cluster. Cluster 0 contains 2 data points, while cluster 1 contains 2 items, resulting in a total of 4 data points, as seen in the image above. After conducting an analysis on each data point using various methods, consistent results were obtained in each test. Specifically, SMP Negeri 40 and SMP Negeri 25 were found in cluster 0, indicating that these schools are ready to implement Smart School. On the other hand, SMP Negeri 23 and SMP Aziziyyah were located in cluster 1, signifying that these schools are not prepared to implement Smart School.

3.4. Kesimpulan Hasil Uji Data

The results of the first clustering test using the Microsoft Excel application obtained data results as shown in table 6.

Table 6. Clustering Results Using Excel

School Data	C0	C1	Data Group
SMP NEGERI 40 PEKANBARU	2,75	71,75	Cluster 0
SMP NEGERI 25 PEKANBARU	2,75	64,75	Cluster 0
SMP NEGERI 23 PEKANBARU	117,75	17,75	Cluster 1
SMP IT AZIZIYYAH PEKANBARU	47,75	17,75	Cluster 1

The research results indicate the presence of 2 (two) clusters utilized, namely C0 = 2 items, and C1 = 2 items. Cluster 0 represents schools that are prepared, while Cluster 1 represents schools that are not ready to implement Smart School. Furthermore, when tested using the Python programming language, it revealed the same cluster results as when testing the data using Excel. However, there were changes in the attribute data, as depicted in Figure 13.

	Data Sekolah	Pemanfaatan	Kompetensi	Aplikasi	Teknologi	cluster
0	SMP NEGERI 40 PEKANBARU	1.0	1.0	1.000000	1.0	0
1	SMP NEGERI 25 PEKANBARU	0.5	1.0	0.666667	1.0	0
2	SMP NEGERI 23 PEKANBARU	0.6	0.2	1.000000	0.0	1
3	SMP IT AZIZIYYAH PEKANBARU	0.0	0.0	0.000000	0.7	1

Figure 13. Data test results in python.

Upon closer examination, it is evident that there are several attributes that undergo the most significant changes. This research involved two iterations of cluster count adjustments. To determine the most optimal cluster count, a cluster evaluation was conducted using the Sum of Square Error (SSE). Nevertheless, the results in both clusters consist of 2 items each, with SMP Negeri 40 and SMP Negeri 25 residing in Cluster 0, signifying their readiness to implement Smart School. Meanwhile, SMP Negeri 23 and SMP IT Aziziyyah are situated in Cluster 1, indicating that these schools belong to the category of institutions not yet prepared to implement Smart School.

4. Conclusion

Based on the analysis and discussion of the data, it can be concluded that the research on "Determining School Readiness to Implement Smart School Using the K-Means Clustering Method" has found that the K-Means Clustering Method can be effectively employed to categorize school data as a decision support for school readiness in implementing Smart School. Furthermore, this research has successfully grouped the data into two categories (ready and not ready)

using the K-Means Clustering method. Additionally, it was observed that the initial centroid determination significantly impacts the number of iterations required. Moreover, the results of the cluster validation using Python programming were found to be consistent.

Reference

[1] M. K. Anam, A. Yunianta, H. J. Alyamani, Erlin, A. Zamsuri, and M. B. Firdaus, "Analysis and Identification of Non-Impact Factors on Smart City Readiness Using Technology Acceptance Analysis: A Case Study in Kampar District, Indonesia," Journal of Applied Engineering and Technological Science, vol. 5, no. 1, pp. 1–17, 2023, doi: 10.37385/jaets.v5i1.2401.

[2] N. Chotpittayanon, J. Angsukanjanakul, V. Jintalikhitdee, and N. Phasuk, "The Success Factors For the Sustainable Smart City Development in Thailand," in Proceedings on Engineering Sciences, Faculty of Engineering, University of Kragujevac, 2024, pp. 109–118. doi: 10.24874/PES06.01.013.

[3] Isabella and E. Agustian, "Implementing Digital Literacy Policies and the Challenges of Towards Smart City in Palembang City," Journal of Governance and Local Politics (JGLP), vol. 5, no. 2, pp. 122–132, 2023, doi: 10.47650/jglp.v5i2.936.

[4] A. Hasibuan and oris krianto Sulaiman, "Smart City , Konsep Kota Cerdas Sebagai Alternatif Penyelesaian Masalah Perkotaan Kabupaten / Kota," Buletinutama Teknik, vol. 14, no. 2, pp. 127–135, 2019.

[5] A. I. Khan and S. Al-Habsi, "Machine Learning in Computer Vision," Procedia Comput Sci, vol. 167, no. 2019, pp. 1444–1451, 2020, doi: 10.1016/j.procs.2020.03.355.

[6] A. U. Zailani, A. Perdananto, Nurjaya, and Sholihin, "Pengenalan Sejak Dini Siswa SMP Tentang Machine Learning Untuk Klasifikasi Gambar Dalam Menghadapi Revolusi 4.0," KOMMAS: Jurnal Pengabdian Kepada Masyarakat, vol. 1, no. 1, pp. 7–15, 2020.

[7] Forum New Education, "Smart Education in Smart Cities and Smart Regions," Center for innovative education, 2017.

[8] Z. T. Zhu, M. H. Yu, and P. Riezebos, "A research framework of smart education," Smart Learning Environments, vol. 3, no. 1, 2016, doi: 10.1186/s40561-016-0026-2.

[9] M. Sampebua and S. Mangiwa, "The Design Smart School Application to Increase Education in Junior High School," International Journal of Computer …, vol. 15, no. 10, pp. 154–165, 2017.

[10] J. R. Batmetan, M. Nur, O. S. Turang, M. M. Sumampouw, and G. M. Lahengking, "IT Infrastructure Library Framework Approach to the Measurement of e-Government Maturity," Lahengking International Journal of Information Technology and Education (IJITE), vol. 1, no. 2, pp. 119–128, 2022, doi: 10.62711/ijite.v1i2.51.

[11] Dr. M. R. Taghva, Dr. M. T. T. Fard, Dr. S. M. Taheri, and S. Omidinia, "Success Factors for Smart Schools Emphasizing the Role of Information Technology: A Case Study," International Journal of Engineering and Technology, vol. 11, no. 4, pp. 731–739, 2019, doi: 10.21817/ijet/2019/v11i4/191104065.

[12] Lukman and A. Z. Ibad, "Pemilihan Metode Dan Media Pembelajaran Dalam Blended Learning," Jurnal Ilmiah Promis, vol. 1, no. 2, pp. 86–94, 2020.

[13] Yudi Agusta, "K-Means – Penerapan, Permasalahan dan Metode Terkait," Jurnal Sistem dan Informatika, vol. 3, no. Februari, pp. 47–60, 2007.

[14] K. Fatmawati and A. P. Windarto, "Data Mining: Penerapan Rapidminer Dengan K-Means Cluster Pada Daerah Terjangkit Demam Berdarah Dengue (Dbd) Berdasarkan Provinsi," Computer Engineering, Science and System Journal, vol. 3, no. 2, p. 173, 2018, doi: 10.24114/cess.v3i2.9661.

[15] M. K. Anam, I. Y. Pasa, K. D. kusuma Wardhani, L. Efrizoni, and M. B. Firdaus, "K-Means Clustering to Identity Twitter Build Operate Transfer (BOT) on Influential Accounts," ComTech: Computer, Mathematics and Engineering Applications, vol. 14, no. 2, pp. 143–154, 2023, doi: 10.21512/comtech.v14i2.10620.

[16] N. T. Hartanti, "Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional," Jurnal Nasional Teknologi dan Sistem Informasi, vol. 6, no. 2, pp. 82–89, 2020, doi: 10.25077/teknosi.v6i2.2020.82-89.

[17] M. L. Sibuea and A. Safta, "Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustring," Jurteksi, vol. 4, no. 1, pp. 85–92, 2017, doi: 10.33330/jurteksi.v4i1.28.

[18] M. W. Putri, I. M. Nur, and R. Wasono, "Implementasi Spectral Clustering Algorithm untuk Penelompokan Sasaran Vaksinasi Covid-19 di Indonesia," Jurnal Statistika, vol. 10, no. 1, pp. 26–31, 2022, doi: 10.26714/jsunimus.10.1.2022.26-31.

[19] D. Radixavendra Quinthara, A. Charis Fauzan, and M. Maariful Huda, "Penerapan Algoritma K-Modes Menggunakan Validasi Davies Bouldin Index Untuk Klasterisasi Karakter Pada Game Wild Rift," Journal of System and Computer Engineering (JSCE), vol. 4, no. 2, pp. 123–135, 2023, doi: 10.61628/jsce.v4i2.802.

[20] M. N. Zhafar, K. Usman, and F. Akhyar, "Penerapan Metode Clustering Dengan Algoritma K-Means Untuk Analisa Persebaran Varian Covid-19 (Studi Kasus Kelurahan Antapani Kidul)," in e-Proceeding of Engineering, 2023, pp. 4257–4264.

[21] M. F. J. Muttaqin, "Cluster Analysis Using K-Means Method to Classify Sumatera Regency and City Based on Human Development Index Indicator," in Seminar Nasional Official Statistics, 2022, pp. 967–976.

[22] C. Oktarina, K. Anwar Notodiputro, and Indahwati, "Comparison of K-Means Clustering Method and K-Medoids on Twitter Data," Indonesian Journal of Statistics and Its Applications, vol. 4, no. 1, pp. 189–202, 2020, doi: 10.29244/ijsa.v4i1.599.

[23] M. R. Taghva, M. T. T. Fard, S. M. Taheri, and S. Omidinia, "Success Factors for Smart Schools Emphasizing the Role of Information Technology: A Case Study," International Journal of Engineering and Technology, vol. 11, no. 4, pp. 731–739, Aug. 2019, doi: 10.21817/ijet/2019/v11i4/191104065.

[24] N. A. Razak, H. A. Jalil, S. E. Krauss, and N. A. Ahmad, "Studies in Educational Evaluation Successful implementation of information and communication technology integration in Malaysian public schools : An activity systems analysis approach," Studies in Educational Evaluation, vol. 58, no. April, pp. 17–29, 2018, doi: 10.1016/j.stueduc.2018.05.003.

[25] M. N. N. Lee and S. S. Thah, "Building and Sustaining National ICT/Education Agencies," Building and Sustaining National ICT/Education Agencies, 2016, doi: 10.1596/26265.

[26] R. Snehkunj, K. Vachiyatwala, and C. Author, "Data Analysis Using Pandas Library of Python," 2022.

[27] S. S. Nagari and L. Inayati, "Implementation of clustering using k-means method to Determine nutritional status," Jurnal Biometrika dan Kependudukan, vol. 9, no. 1, p. 62, Jun. 2020, doi: 10.20473/jbk.v9i1.2020.62-68.