JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 249 – 256 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3730
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

K-MEANS CLUSTERING METHOD FEASIBILITY OF SCHOOL BUILDING REHABILITATION IN KABUPATEN ASAHAN

Irma Nurhani[1], Nuriadi Manurung[2*], Elly Rahayu[1]
[1]Information System, Universitas Royal
[2]Computer System, Universitas Royal
email: *irmanurhani0308@gmail.com

Abstract: School building construction is an important part of educational facilities and infrastructure. Many elementary school buildings in Asahan Regency are damaged, unfit for use or lack facilities including classrooms and other supporting infrastructure. The results of the selection that are not transparent in deciding the feasibility of rehabilitation and construction of elementary school buildings are often subjective and time-consuming. This causes decisions to be taken that cannot be made as quickly as possible and inequality, where schools that actually need rehabilitation more do not get priority for rehabilitation. The purpose of this study is to implement data mining in the selection of school building construction projects with the K-Means clustering algorithm in clustering the feasibility of rehabilitation of elementary school building construction. The results of this study found 31 elementary schools that are eligible for school building construction rehabilitation and 4 are not eligible for school building construction rehabilitation. This study is expected to provide a significant contribution in increasing the efficiency of school building construction selection and more transparency towards elementary school buildings to be rehabilitated.

Keywords: education authorities; k-means; rehabilitation and construction of school buildings.

Abstrak: Pembangunan gedung sekolah ialah bagian penting dalam sarana dan prasarana pendidikan. Banyak gedung Sekolah Dasar di Kabupaten Asahan menderita kerusakan, tidak layak pakai atau kekurangan fasilitas meliputi ruang kelas dan infrastruktur pendukung lainnya. Hasil penyeleksian yang bersifat tidak transparan dalam memutuskan kelayakan rehabilitasi dan pembangunan gedung Sekolah Dasar sering kali bersifat subjektif dan memakan waktu lama.Hal tersebut menyebabkan keputusan yang diambil tidak dapat dilakukan secepat mungkin serta ketimpangan, dimana sekolah yang sebenarnya lebih membutuhkan rehabilitasi justru tidak mendapatkan prioritas untuk direhabilitasi. Tujuan penelitian ini ialah mengimplementasi data mining pada pemilihan proyek pembangunan gedung sekolah dengan algoritma K-Means clustering dalam mengcluster kelayakan rehabilitasi pembangunan gedung sekolah dasar. Hasil Penelitian ini terdapat 31 sekolah dasar layak untuk direhabilitasi pembangunan gedung sekolah dan 4 tidak layak untuk direhabalitasi pembangunan gedung sekolah. Penelitian ini diharapkan dapat memberikan kontribusi signifikan dalam menaikkan efisiensi penyeleksian pembangunan gedung sekolah dan lebih transparansi terhadap gedung sekolah dasar yang akan direhabilitasi.

Kata kunci: dinas pendidikan; k-means; rehabilitasi dan pembangunan gedung sekolah.

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 249 – 256 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3730
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

INTRODUCTION

Educational buildings, including elementary school buildings, have a very important role in increasing the human development index. Aiming to realize the goal of education in order to educate the nation's life, which is the main focus of every educational institution [1]. Students will feel comfortable with the infrastructure that supports the teaching and learning process, this will result in better quality of education, especially in Asahan Regency [2].

Asahan Regency is one of the regencies/cities in North Sumatra, Indonesia. Asahan Regency has 25 (twenty five) sub-districts, the total number of elementary schools is around 452 schools. Public elementary schools in 2024 are around 380 (three hundred and eighty) and private elementary schools registered with the Asahan Education Office are around 72 (seventy two) schools spread across the 25 sub-districts. In Asahan Regency, the local government and the local education office continue to strive to improve access and quality of education. However, the condition of education in Asahan Regency is very concerning, especially in terms of infrastructure .

Even though there have been various efforts to develop educational infrastructure, the challenges faced related to the inadequate condition of school buildings are one of the main challenges in efforts to improve the quality of education. Many school buildings are unfit for use, are damaged, or lack basic facilities, such as inadequate classrooms and other supporting infrastructure.

This problem not only hampers the teaching and learning process but also has the potential to endanger the safety of students and teaching staff. Therefore, serious attention from various parties, including the government and related institutions, is needed to ensure effective handling steps are taken. Initial mapping carried out by the Education Department shows that there are a number of schools that require rehabilitation, as well as several schools that require new classrooms to be built in response to the growth in student numbers.

The government, through various proyeks, has allocated a budget for the rehabilitation and construction of school buildings. However, the main challenge in implementation is how to prioritize school buildings that must be repaired or rebuilt efficiently. This decision grouping process is often hampered by limited structured data, causing budget allocation to be less than optimal.The conventional manual approach in determining the feasibility of rehabilitation and construction of school buildings is often subjective and takes a long time.

In this case, evaluating the feasibility of rehabilitation and construction of school buildings needs to be carried out using a more systematic and measurable approach. Apart from considering physical conditions such as damage to buildings and facilities, other aspects such as room capacity, age of the building, light, medium and heavy damage must also be taken into account [3]. With the complexity of the data that needs to be analyzed, conventional manual approaches become less effective and efficient.

The lack of a comprehensive system for determining the appropriateness grouping for school building rehabilitation often results in inequality, where schools that actually need rehabilitation more often do not get it. Therefore, to overcome this problem, decision support is needed to assist the Education Department in determining the feasibility of rehabilitation and construction of school

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 249 – 256 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3730
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

buildings, namely by grouping existing school data using the more structured K-Means Cluster Data Mining method to manage and analyze the data optimally [3].

Data Mining is the process of collecting and processing data to understand interesting information or patterns in data, using certain algorithms, techniques and techniques, with the aim of finding important information in data [4]. The process of grouping records that have useful meaning is called Clustering.

Clustering is grouping a number of data or objects into clusters (groups) so that each cluster will contain data that is as similar as possible and different from objects in other clusters [6]. One of the Clustering algorithms found in data mining is used to obtain groups that have large amounts of data using an efficient and fast separation technique, namely the K-Means Algorithm. The selection of this technique is very easy in the process of application or use and adapts according to the case [7].

Research using the same technique under the title Application of the K-Means Technique on Underprivileged Data in Blitar District, The number of samples used was 22 sub-districts in Blitar Regency and produced 2 clusters, namely cluster 0 and cluster 1 [8].

Next, the research through the title of Grouping of Students at Al-Ma'rifah Islamic Boarding School Based on Region of Origin Using the K-Mean Clustering Algorithma [9]. The purpose of this research can provide a significant contribution in increasing the efficiency of Islamic boarding school management, facilitating monitoring and serving students, and opening up opportunities for developing more targeted educational proyeks according to the characteristics of certain regional groups.

Next, research through the title Application of Clustering Algorithm for Grouping Poverty Levels in Banten Province [10]. In this study, 3 groups of selected areas were obtained based on the level of poverty. The algorithm used in this study has a process and the cluster division vector looks the same, however, based on the DBI results obtained, k-medoid through 0.582, and k-means through a value of 0.602 [10]. Clustering is the process of grouping data into several clusters or groups, where the data in one cluster has high similarity, while the similarity between clusters is minimal [11].

K-means is a non-hierarchical data clustering method that groups data in the form of one or more clusters/groups. Data that has the same characteristics is grouped and data that has different characteristics is grouped with other clusters/groups so that data in one cluster/group has a small level of variation. The aim is to divide the data into one or more clusters.

This study focuses on the application of the K-Means Clustering algorithm to analyze the feasibility of rehabilitation and construction of elementary school buildings in Asahan Regency, while previous studies were manual in determining the feasibility of rehabilitation and construction of school buildings, which were often subjective and time-consuming. It is hoped that the results of this research will not only provide recommendations based on data and facts, but will also serve as a guide for policy makers in planning and developing educational infrastructure in the region.

Apart from that, it is hoped that this effort can support improving the quality of education in Asahan Regency and create a safer and more comfortable learning environment for students. So this re-

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 249 – 256 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3730
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

search want to create a computerized system that can cluster the feasibility of rehabilitation and construction of elementary school buildings.

METHOD

Data mining is the process of finding useful patterns in large data sets. From other sources, data mining is the study of collecting, cleaning, processing, analyzing, and gaining useful insights from data [12]. The K-Means method was chosen because this method uses physical data that is not abstract and clear, this is in accordance with the data that will be used in the problem of clustering school building construction data. The purpose of this study is to determine the application and results of clustering with the K-Means algorithm in clustering school building construction data in Asahan Regency [13].

Data grouping the feasibility of rehabilitation and construction of school buildings by taking samples from the Asahan District Education Office, the following researchers created as an illustration of how researchers process the existing data which will later become a summary per school to then be entered into the calculation formula for the K-Means Clustering Analysis Algorithm. Can be seen in the table 1.

Table 1. Grouping Criteria Scores at the Education Service Kab. Asahan

No	Criteria (C)	Data Source	Weight Value	Terms
1	Study Room Damage (C1)	Dinas Pendidikan Kabupaten Asahan	1	Low
			2	Currently
			3	Heavy
2	Damage to Student Support Facilities (C2)	Dinas Pendidikan Kabupaten Asahan	1	Low
			2	Currently
			3	Heavy
3	Damage to Toilets and Places of Worship (C3)	Dinas Pendidikan Kabupaten Asahan	1	Low
			2	Currently
			3	Heavy

The calculation algorithm using the K-Means method is an algorithm for clustering and attribute-based objects into k partitions, where k < n. In general, K-means clustering is a non-hierarchical data clustering technique that groups data in the form of one or more groups [14].

The aim is to minimize the objective function set in the grouping process, which generally seeks to minimize variation within a group and maximize variation between groups. The steps in the K-Means Clustering Algorithm are by deciding the number of clusters and deciding the centroid value [7].

In deciding the centroid value for the start of the solution, the initial centroid value is done randomly. Meanwhile, if the centroid value is the stage of completion, then the following formula is used [15]:

$$v_{ij} = \frac{1}{Ni} \sum_{k=0}^{Ni} X_{kj} \qquad (1)$$

Description:
Vij : Average centroid of Cluster 1 for variable j
Ni : Number of members of Cluster i
i,k : Imdeks of Cluster
j : Index of Variable

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 249 – 256 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3730
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

Xkj : Value of data k of variable j for the Cluster

Calculates the distance between the centroid points through the points of each object.

$$D_e=\sqrt{(xi-si)^2+(yi-ti)^2} \qquad (2)$$

Description:
De : Euclidean Distance
I : Number of objects
(x,y) : Object coordinates
(s,t) : Centroid coordinates

Grouping objects to determine cluster members is by calculating the minimum distance of objects. Return to stage 2, repeat until the resulting centroid value remains constant and cluster members do not move to another cluster.

RESULTS AND DISCUSSION

In working on the implementation of the K-Means algorithm in this research, the author used the clustering function which is part of data mining which is useful for dividing a set of data into several clusters/groups based on similarities or similarities based on community data in table 1, to be normalized based on the criteria in table 2.

Table 2. Criterion Data Normalization

No	Alternative Name	C1	C2	C3
1	UPTD SDN 017127 Sei Nangka	3	2	3
2	UPTD SDN 010043 Perk. Air Batu III/IX	3	2	3
3	UPTD SDN 010076 Pondok Bungur	3	2	2
4	UPTD SDN 010110 Desa Ambalutu	2	3	3
5	UPTD SDN 013819 Desa Sei Lama	3	3	3
6	UPTD SDN 013830 Desa P. Mahondang	2	3	3
7	UPTD SDN 014649 Desa P. Rakyat Tua	3	3	2
8	UPTD SDN 014667 Danau Sijabut Gardu	3	3	3
9	UPTD SDN 015933 Lobu Rappa	3	2	2
10	UPTD SDN 014671 Sentang	3	2	3
-	-	-	-	-
35	UPTD SDN 013842 Sei Silau Tua	1	1	2

To determine nominal data, it must be initialized first in the form of numbers by sorting the data based on the frequency of occurrence and initializing the data starting from the highest data with a value of 1, then the next data 2, 3 and so on. To decide on the initial center of the cluster, deciding on the centroid value for the start of the iteration is done randomly. In this case the author made an example for 2 elementary schools, namely, PTD SDN 013819 Sei Lama Village and UPTD SDN 015893 Air Putih. Can be seen in the tablel 3.

Table 3. Iteration Centroid Point 1

Kode	Name	C1	C2	C3
A05	SDN 013819 Desa Sei Lama	Heavy	Heavy	Heavy
A32	SDN 015893 Air Putih	Low	Low	Low

After deciding on the initial center of the cluster, deciding on the centroid value for the start of the iteration is done randomly. In this case the author made

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 249 – 256 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3730
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

examples for 2 elementary schools, namely, SDN 013819 Sei Lama Village (Data 5), SDN 015893 Air Putih (Data 32). Calculate the distance between the centroid points through the points of each object using the Euclidian Distance formula, namely;

$$D_e=\sqrt{(xi-si)^2+(yi-ti)^2} \qquad (3)$$

One example is the calculation of Code A01 (SDN 017127 Sei Nangka) to later Code A35 (SDN 013842 Sei Silau Tua):

Because C1 is looking for a new cluster calculation it becomes:

$A1 = \sqrt{(3-3)^2+(2-3)^2+(3-3)^2} = 1{,}0000$

$A2 = \sqrt{(3-3)^2+(2-3)^2+(3-3)^2} = 1{,}0000$

$A3 = \sqrt{(3-3)^2+(2-3)^2+(2-3)^2} = 1{,}4142$

$A4 = \sqrt{(2-3)^2+(3-3)^2+(3-3)^2} = 1{,}0000$

$A5 = \sqrt{(3-3)^2+(3-3)^2+(3-3)^2} = 0{,}0000$ and so on until A35

Because C2 is looking for a new cluster calculation becomes:

$A1 = \sqrt{(3-1)^2+(2-1)^2+(3-1)^2} = 3{,}000$

$A2 = \sqrt{(3-1)^2+(2-1)^2+(3-1)^2} = 3{,}000$

$A3 = \sqrt{(3-1)^2+(2-1)^2+(2-1)^2} = 2{,}45$

$A4 = \sqrt{((2-1)\text{^}2+(3-1)\text{^}2+(3-1)\text{^}2)} = 3{,}00$

$A5 = \sqrt{(3-1)^2+(3-1)^2+(3-1)^2} = 3{,}46$ and so on until A35

Then calculate the distance from data 1 to data 35 to the center of the cluster. So the result of the distance calculation from the first and second iterations is that there are similarities in the clustering results between the two, especially in the names of the elementary schools produced. This shows that there is no need for further iteration or repetition.

So that in Cluster 1, there are 31 elementary schools that are suitable for rehabilitating school buildings, currently available for Cluster 2, which is not suitable for rehabilitating school buildings, are 4 elementary schools, which can be seen in table 4.

Table 4. Grouping Cluster Members

No	Alternative Name	Cluster 1	Cluster 2	Information
1	SDN 017127 Sei Nangka	C1		Worthy
2	SDN 010043 Perk. Air Batu III/IX	C1		Worthy
3	SDN 010076 Pondok Bungur	C1		Worthy
4	SDN 010110 Desa Ambalutu	C1		Worthy
5	SDN 013819 Desa Sei Lama	C1		Worthy
6	SDN 013830 Desa P. Mahondang	C1		Worthy
7	SDN 014649 Desa P. Rakyat Tua	C1		Worthy
8	SDN 014667 Danau Sijabut Gardu	C1		Worthy
9	SDN 015933 Lobu Rappa	C1		Worthy
10	SDN 014671 Sentang	C1		Worthy
-	-	-	-	-
35	UPTD SDN 013842 Sei Silau Tua		C2	Not feasible

JURTEKSI (Jurnal Teknologi dan Sistem Informasi) ISSN 2407-1811 (Print)
Vol. XI No 2, Maret 2025, hlm. 249 – 256 ISSN 2550-0201 (Online)
DOI: http://dx.doi.org/10.33330/jurteksi.v11i2.3730
Available online at http://jurnal.stmikroyal.ac.id/index.php/jurteksi

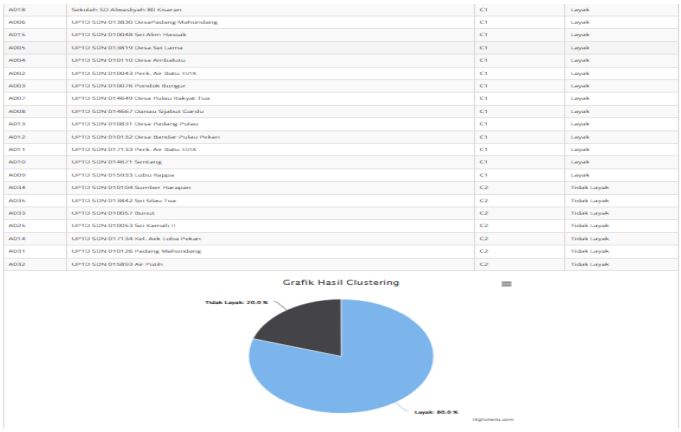

Image 1. Clustering Results Graph

The image above is a graph of the K-Means method in grouping the feasibility of elementary school building rehabilitation, there are 31 elementary schools that are eligible for building rehabilitation and 4 elementary schools that are not eligible for school building rehabilitation. Furthermore, this study obtained a comparison of the results before the implementation of system development using k-means and after implementing the system with k-means, the results are presented in table 5:

Table 5. Comparison Table With and Without K-Means System

Aspect	WITHOUT K-Means Implementing The System	WITH K-Means Implementing The System
Determination Method	Direct observation and subjective assessment	Data-driven with objective cluster analysis
Time Efficiency	It takes a long time due to the comprehensive field survey.	The process is faster because data analysis is done automatically.
Decision Accuracy	The conventional manual approach is becoming less effective and efficient	More accurate because it is based on data patterns
Transparency	Less transparent as it relies on individual assessments	More transparent because it is data-based

CONCLUSION

The clustering data mining system accelerates the grouping of data on the feasibility of rehabilitation and construction of elementary school buildings to rehabilitate elementary school buildings in the Asahan District Education Office. The application carried out by the author includes field studies, literature studies and interviews. From the analysis and observations carried out, it can be seen that the application of data mining in the school building construction proyek with the K-Means clustering algorithm method in clustering the feasibility of elementary school building rehabilitation construction, there are 31 elementary schools that are suitable for building rehabilitation and 4 are not suitable for school building rehabilitation.

BIBLIOGRAPHY

[1] S. R. J. Rajagukguk, S. Tumanggor, J. G. Malau, and H. Turnip, "Pentingnya Pemerhatian Sarana dan Prasarana bagi Pendidikan di Sekolah yang Terpencil," *PEDIAQU J. Pendidik. Sos. dan Hum.*, vol. 2, no. 1, pp. 204–215, 2023.

[2] A. Mosepe, "Analisa Tingkat Kerusakan BangunanGedung Sekolah Dasar Di KecamatanPomona Tenggara Kabupaten Poso," pp. 1–5, 2022.

[3] E. P. A. Yahya and F. A. B. Pradipta, "Asesmen dan analisa gedung eksisting (Studi kasus bangunan Johar Shopping Center Semarang)." Universitas Islam Sultan Agung, 2022.

[4] P. K. Heatubun and M. Fansyuri, "Analisa Data Mining Untuk Prediksi Penjualan Produk Menggunakan Algoritma FP-Growth Berbasis Web Studi Kasus Online Shop Muslim Galeri," *BINER J. Ilmu Komputer, Tek. dan Multimed.*, vol. 1, no. 6, pp. 1376–1382, 2024.

[5] D. S. Angreni and M. Susanti, "Implementasi Data Mining Untuk Rekomendasi Kenaikan Pangkat Pegawai Negeri Sipil Menggunakan Algoritma Naïve Bayes Pada Biro Administrasi Pimpinan Sekretariat Daerah Provinsi Sulawesi Tengah," *Innov. J. Soc. Sci. Res.*, vol. 4, no. 1, pp. 9661–9674, 2024.

[6] A. A. Zulyani, A. S. Y. Irawan, and A. Jamaludin, "Implementasi Data Mining Menggunakan Algoritma K-Means Untuk Memutuskan Tingkat Vaksinasi Pada Kecamatan Tambun Selatan," *Innov. J. Soc. Sci. Res.*, vol. 3, no. 3, pp. 7037–7050, 2023.

[7] R. Harum, F. Fathurrohman, and A. R. Rinaldi, "Pemodelan Pola Penjualan Sepatu Impor Melalui Pendekatan Algoritma K-Means," *JATI (Jurnal Mhs. Tek. Inform.*, vol. 8, no. 3, pp. 4223–4230, 2024.

[8] D. N. Alfiansyah, V. R. S. Nastiti, and N. Hayatin, "Implementasi Metode K-Means pada Data Penduduk Miskin Per Kecamatan Kabupaten Blitar," *J. Repos.*, vol. 4, no. 1, pp. 49–58, 2022.

[9] S. N. Hasanah, N. Suarna, D. Pratama, M. Informatika, and S. I. Cirebon, "Pengelompokan Santri Di Pondok Al- Ma ' Rifah Berdasarkan Asal Wilayah Menggunakan Algoritma K-Mean Clustering," vol. 3, no. 1, pp. 37–48, 2024.

[10] T. A. Munandar, "Implementasi Algoritma Clustering Untuk Pengelompokan Tingkat Kemiskinan Provinsi Banten," *JSiI (Jurnal Sist. Informasi)*, vol. 9, no. 2, pp. 109–114, 2022.

[11] M. Triandini, S. Defit, and G. W. Nurcahyo, "Data Mining dalam Mengukur Tingkat Keaktifan Siswa dalam Mengikuti Proses Belajar pada SMP IT Andalas Cendekia," *J. Inf. dan Teknol.*, vol. 3, pp. 167–173, 2021, doi: 10.37034/jidt.v3i3.120.

[12] W. Handoko and A. Nasution, "Implementation of Data Mining By Using K-Means To Classify Marriage Age," *JURTEKSI (Jurnal Teknol. dan Sist. Informasi)*, vol. 9, no. 2, pp. 161–166, 2023, doi: 10.33330/jurteksi.v9i2.2045.

[13] P. Sari, "Algoritma K-Means Clustering: Sebuah Studi Literatur," *J. Inform.*, vol. 1, no. 1, pp. 1–7, 2024.

[14] R. Supardi and I. Kanedi, "Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis," *J. Teknol. Inf.*, vol. 4, no. 2, pp. 270–277, 2020, doi: 10.36294/jurti.v4i2.1444.

[15] "Implementasi metode k-means clustering untuk pengelompokan kelas berdasarkan pemahaman siswa pada bimbingan belajar matematika sachio banyuwangi skripsi," 2022.