Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

15* February 2025. Vol. 4. No. 2; e-ISSN: 2808-4519

Improving the School Type Clustering Model on the Foundation Using the K-Means Algorithm (Case Study: Kebon Kelapa Al-Ma'rifah, Cirebon Regency)

Hanifah Nur Aulia[1*], Martanto[2], Arif Rinaldi Dikananda[3], Mulyawan[4]

[1,2,3,4] *STMIK IKMI Cirebon*
hanifhaanew@gmail.com[1]*

Abstract

This study aims to improve the school type grouping model at the Kebon Kelapa Al-Ma'rifah Foundation, Cirebon Regency, using the K-Means algorithm. Data-based grouping is very important in supporting efficient education management, especially in environments that have various types of schools such as Madrasah Aliyah (MA), Vocational High School (SMK), Madrasah Tsanawiyah (MTs), and Madrasah Ibtidaiyah (MI). The data used comes from the New Student Registration (PPDB) dataset for the 2023–2024 school year, with demographic attributes such as name, place of birth, gender, and time of school entry. The evaluation of clustering quality was carried out using the Davies-Bouldin Index (DBI) to determine the optimal number of clusters. The results show that the optimal number of clusters is K=5 with the lowest DBI value of 0.201, which results in compact and well-separated clusters. The implementation of the K-Means algorithm helps the foundation understand the distribution pattern of students based on attributes such as gender, region, and entry time. This research provides practical benefits, including more targeted resource allocation, improved quality of education, and efficiency in school management. In addition, this research contributes to the development of data mining models in the education sector and opens up opportunities for the exploration of additional attributes such as academic achievement and socioeconomic conditions. Further research is suggested to use alternative algorithms such as K-Medoids or DBSCAN.

Keywords: *Grouping, K-Means, Davies-Bouldin Index, Kebon Kelapa Al-Ma'rifah Foundation*

1. Introduction

The development of information and communication technology has changed the way we manage data, including in the world of education. In Indonesia, the grouping of school types such as Madrasah Aliyah (MA), Vocational College (SMK), and Madrasah Ibtidaiyah (MI) is becoming increasingly important, considering the increasing number of educational institutions. One approach that can be used to segment school data is the K-Means algorithm, which allows for the identification of patterns in student data, curriculum, and other resources. This study aims to explore how the K-Means algorithm can improve the school type grouping model in the Kebon Kelapa Al-Ma'arifah Foundation, Cirebon Regency.

Challenges in education data management, such as suboptimal resource allocation and incompatibility of educational programs, require a data-driven approach. In addition, evaluating the quality of clustering using techniques such as the Davies-Bouldin Index is important to ensure that the resulting clustering is optimal. This study uses the K-Means algorithm to group school data with the attributes "Gender" and "School Attendance," and evaluate the quality of clustering with the Davies-Bouldin Index. This research aims to provide insight into the management of education data at the Kebon Kelapa Al-Ma'arifah Foundation and fill the literature gap related to the evaluation of clustering in education.

Several previous studies have also applied the K-Means algorithm in various fields. [1] using K-Means to group cases of violence against children and women by age. [2] Use a similar method to segment sales data by region. [3] grouping districts in South Sumatra after Covid-19, while [4] grouping students based on their activeness in learning. Other research such as as [5] also apply K-Means in sales analysis and inventory management. Other research such as [6] also apply K-Means in sales analysis and inventory management. [7] uses K-Means to group students based on IQ scores, while Information [8] applying K-Means to group student achievement based on subject grades and attendance.

This research is expected to improve the efficiency of resource allocation and the development of program based on cluster needs. In addition, these findings can be a reference for educational technology developers to apply clustering algorithms in analyzing large-scale educational data, opening up opportunities for further exploration in student behavior analysis or data-driven curriculum evaluation.

1.1. K-Means

In your thesis, K-means is described as an algorithm used for clustering data into distinct groups based on shared characteristics. The process begins by selecting the number of clusters, followed by the assignment of initial cluster centroids. The algorithm iterates to refine these clusters by adjusting the centroids to minimize intra-cluster variance and maximize inter-cluster separation. This is done through the calculation of Euclidean distances between data points and centroids.

1.2. Data Analysis Techniques

In this study, the Al-goritma K-means clusturing analysis method was used. K-means clusturing is one of the algorithms to determine the classification of objects based on the attributes/characteristics of objects in K clusters/partitions. Remove irrelevant data, eliminate duplicates, and handle missing values. Apply the K-Means algorithm to the processed data to divide schools into clusters.
The formula of the Euclidean Distance is as follows:

$$d = \sqrt{\sum_{N} (X_i - Y_i)^2}$$

1.3. Knowledge Discovery Database

The Knowledge Discovery in Databases (KDD) process is crucial for extracting useful patterns and knowledge from large datasets, especially in fields like education. It involves several stages, starting with data selection, preprocessing, and transformation, followed by data mining techniques and evaluation. In the context of the study presented in the thesis, KDD plays an integral role in analyzing the dataset of new student registrations at the Kebon Kelapa Al-Ma'rifah Foundation. The data is preprocessed to clean and standardize attributes, and then clustering algorithms like K-Means are applied to group similar data points based on demographic features such as gender, age, and geographic location. After data transformation, the K-Means algorithm was applied, and the clustering quality was evaluated using the Davies-Bouldin Index (DBI), which helped determine the optimal number of clusters, ensuring compact and well-separated groupings essential for efficient decision-making in educational management. The KDD process is vital for deriving actionable insights that lead to better resource allocation and program development within the foundation's schools, showing how data mining and KDD can effectively support strategic planning in educational environments.

2. Research Methods

2.1. Methods

The figure below illustrates the process flow *of Knowledge Discovery in Databases* (KDD), which is a systematic step to extract meaningful knowledge from raw data. This process consists of several main stages, namely data selection, pre-processing, transformation, *data mining*, and interpretation and *evaluation*. Each stage plays an important role in ensuring the end result is reliable information.

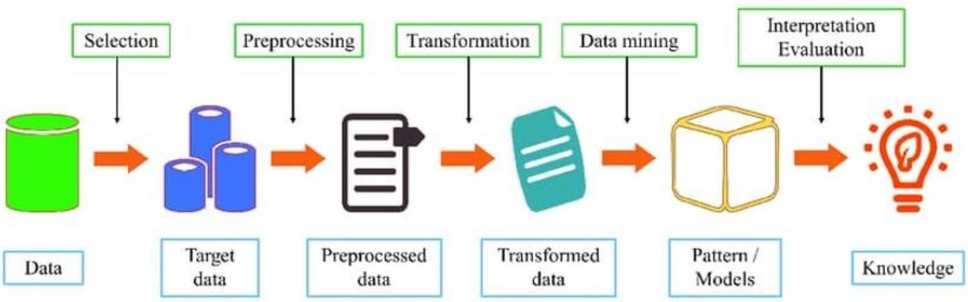


Fig. 1: Research Methods

Through these stages, raw data that was initially unstructured is processed into knowledge that is useful for decision-making. By implementing the KDD process, organizations can maximize the potential of their data to provide strategic insights and create added value. This process is at the core of modern analytics, allowing for the exploration of hidden patterns and models in increasingly complex data.

2.2. Research Supporting Data

This section describes the supporting data used in the research to apply the K-Means algorithm to the grouping of school types in the Kebon Kelapa Al-Ma'rifah Foundation, Cirebon Regency. The data used includes attributes such as student name, gender, region of origin, and type of school entry, which are obtained from the New Student Registration (PPDB) archive for the 2023–2024 school year. This data is the basis for analysis and clustering, supporting the process of pattern identification and validation of research results.

3. Results and Discussion

3.1. Selection

This study uses data on the types of schools under the auspices of the Kebon Kelapa Al-Ma'rifah Foundation of Cirebon Regency, namely MI, MA, MTs, and SMK, which include the attributes of Name, place of birth, date, month, year, gender, village, sub-district, district, school admission.

Table. 1: New Student Admission Dataset for the Academic Year 2023-2024

Name	Place of Birth	Date	Moon	Year	Gender	Village	District	Regency	Attend School
Agniya	Sumedang	15	October	2008	Woman	manggungharja	ciparay	Sumedang	BUT
Cicawarda	Indramayu	15	July	2008	Woman	Totoran	Pasekhan	Indramayu	BUT
Aulia diamond	Bandung	12	December	2008	Woman	cangkuang	Rancaekek	Bandung	BUT
Nailal falah	Cirebon	26	March	2008	Woman	long length	kedawung	Cirebon	BUT
Silvi	Majalengka	17	May	2007	Woman	Baribis	cigasong	majalengka	BUT
Lucky	Karawang	5	June	2012	Man	pasirukem	Cilama kulon	Karawang	ME
M heriawan	Karawang	15	October	2012	Man	ciparage jaya	tempuran	Karawang	ME
Maulida	Cirebon	9	January	2014	Woman	Weu Kidul	Weru	Cirebon	ME
Nabila	Majalengka	7	September	2012	Woman	Sukawera	ligung	Majalengka	ME
Abdul rojak	Cirebon	14	June	2011	Man	Kaliwadas	source	Cirebon	Mts
Adil	Cirebon	31	March	2011	Man	Walahar	gempol	Cirebon	Mts
Aditya	Sumedang	25	August	2010	Man	Cibungur	Rancakalong	Sumedang	Mts
Ahmad	Subang	27	Februarui	2011	Man	kamarung	Paid	subang	Mts
Ahmad	Subang	12	May	2011	Man	Meneng Swamp	lanakan	subang	Mts
Ahmad rafi	North Jakarta	2	September	2011	Man	Ancol	pademangan	North Jakarta	Mts
Aidil	Cirebon	17	November	2010	Man	Ciawi	Palimanan	Cirebon	Mts
Andika	Subang	20	March	2011	Man	Kalentambo	National Heritage	subang	Mts
Arega	Indramayu	27	September	2009	Man	Tempel Kulon	Lelea	Indramayu	Mts
Armahedi	Sumedang	21	February	2011	Man	mekarsari	Sukasari	Sumedang	Mts
It was nothing	Indramayu	10	December	2010	Man	situraja	gantar	Indramayu	Mts

3.2. Preprocessing

Preprocessing is an important stage in ensuring the quality of clustering results. In this study, student data from various types of schools at the Kebon Kelapa Al-Ma'arifah Foundation, Cirebon Regency, were collected in a table that included demographic information and school entry times. The preprocessing stage is carried out to clean the data from missing values, duplicates, and non-uniform formats. This process includes handling missing values, normalizing attributes, and encoding categorical data. With clean data, the K-Means algorithm can generate more accurate clusters, and the use of RapidMiner allows for efficient and structured clustering analysis.

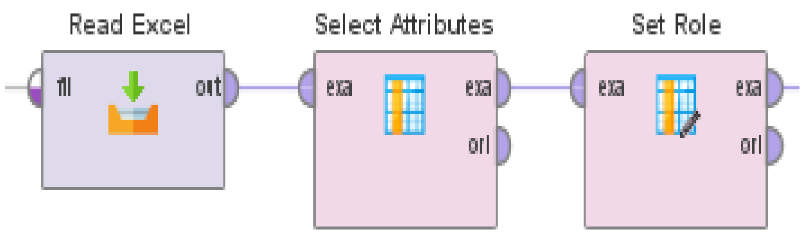
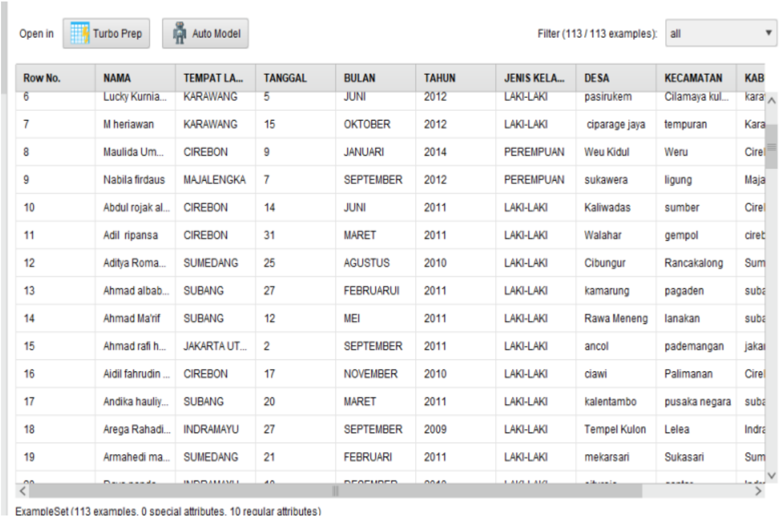


Fig. 2: Operator Read Excel, Select Attrbutes Dan Set Role

Figure 2 shows that Excel's Read Operator is used to import xlsx files into the analysis system, while *Select Attributes* allows the selection of a relevant subset of attributes from the dataset, ignoring unnecessary ones. After that, the Set Role operator is used to assign roles to attributes, such as specifying which are classes (target variables) and which are features (input variables), according to analysis or modeling purposes.

This dataset contains complete information about students, including personal identity and demographic data. The main attributes include the student's name, place and date of birth (date, month, year), gender, and location of residence (village, sub-district, district). This dataset records 113 rows of student data that can be used for further analysis, such as clustering by age or geographic location. This analysis can provide insights to support more effective and efficient education policies, as well as education equity through data segmentation and algorithms such as K-Means.


Open in Turbo Prep Auto Model Filter (113 / 113 examples): all

Row No.	NAMA	TEMPAT LA...	TANGGAL	BULAN	TAHUN	JENIS KELA...	DESA	KECAMATAN	KAB
6	Lucky Kurnia...	KARAWANG	5	JUNI	2012	LAKI-LAKI	pasirukem	Cilamaya kul...	kara
7	M heriawan	KARAWANG	15	OKTOBER	2012	LAKI-LAKI	ciparage jaya	tempuran	Kara
8	Maulida Um...	CIREBON	9	JANUARI	2014	PEREMPUAN	Weu Kidul	Weru	Cire
9	Nabila firdaus	MAJALENGKA	7	SEPTEMBER	2012	PEREMPUAN	sukawera	ligung	Maja
10	Abdul rojak al...	CIREBON	14	JUNI	2011	LAKI-LAKI	Kaliwadas	sumber	Cire
11	Adil ripansa	CIREBON	31	MARET	2011	LAKI-LAKI	Walahar	gempol	cireb
12	Aditya Roma...	SUMEDANG	25	AGUSTUS	2010	LAKI-LAKI	Cibungur	Rancakalong	Sum
13	Ahmad albab...	SUBANG	27	FEBRUARUI	2011	LAKI-LAKI	kamarung	pagaden	suba
14	Ahmad Ma'rif	SUBANG	12	MEI	2011	LAKI-LAKI	Rawa Meneng	lanakan	suba
15	Ahmad rafi h...	JAKARTA UT...	2	SEPTEMBER	2011	LAKI-LAKI	ancol	pademangan	jaka
16	Aidil fahrudin ...	CIREBON	17	NOVEMBER	2010	LAKI-LAKI	ciawi	Palimanan	Cire
17	Andika hauliy...	SUBANG	20	MARET	2011	LAKI-LAKI	kalentambo	pusaka negara	suba
18	Arega Rahadi...	INDRAMAYU	27	SEPTEMBER	2009	LAKI-LAKI	Tempel Kulon	Lelea	Indra
19	Armahedi ma...	SUMEDANG	21	FEBRUARI	2011	LAKI-LAKI	mekarsari	Sukasari	Sum

ExampleSet (113 examples, 0 special attributes, 10 regular attributes)

Fig. 3: Output Operator Read Excel

Figure 3 shows the output of the Excel Read operator that imports data from an Excel file containing information about the student. The table includes several attributes, namely Student Name, Place of Birth, Date, Month, and Year of Birth, Gender, Village, District, Regency, and School Entry. These attributes present detailed information about the student's identity, regional origin, and educational data, which can be used for further analysis

3.3. Transformation

Furthermore, a method is needed to convert nominal data into numerical form. One effective way to perform this transformation is to use the "Nominal to Numerical" operator in RapidMiner, which allows us to convert categorical attributes into a numerical format that is suitable for further modeling. Figure 4 below shows the data transformation process.

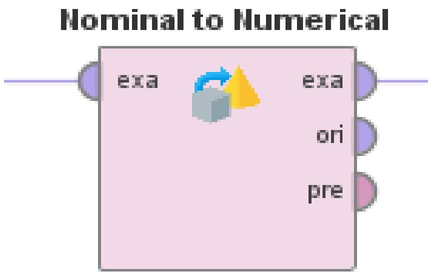


Fig. 4: Operator Nominal to Numerical

Figure 4 Data transformation from nominal to numeric can be performed using the "Nominal to Numerical" operator. This operator converts categorical (nominal) attributes into a numerical format that can be processed by algorithms that require numerical input

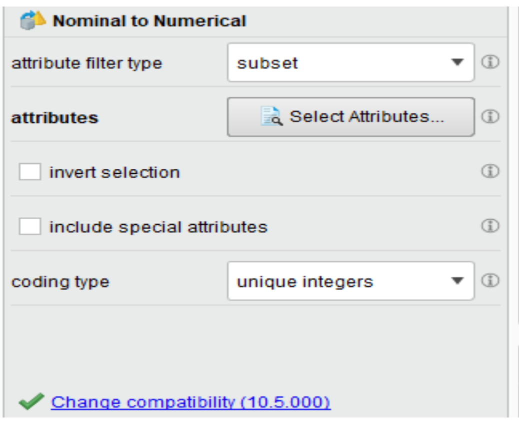


Fig. 5: Parameter Operator Nominal To Numerical

Figure 5 above shows the interface of the "Nominal to Numerical" tool used to convert nominal attributes to numeric in data analysis software. In the *Attribute filter type* (subset) section, users can select the filter type to specify which attributes to convert.

3.3. Data Mining

The next step is a popular technique for grouping data into clusters based on similarities between the data. In this method, the user specifies the desired number of clusters, and then the K-Means algorithm will randomly place the centroid in the data space. Research on the application of the K-Means algorithm to assist foundations in grouping schools based on certain characteristics. With the K-Means algorithm, schools under the auspices of the Kebon Kelapa Al-Ma'rifah Foundation can be grouped.

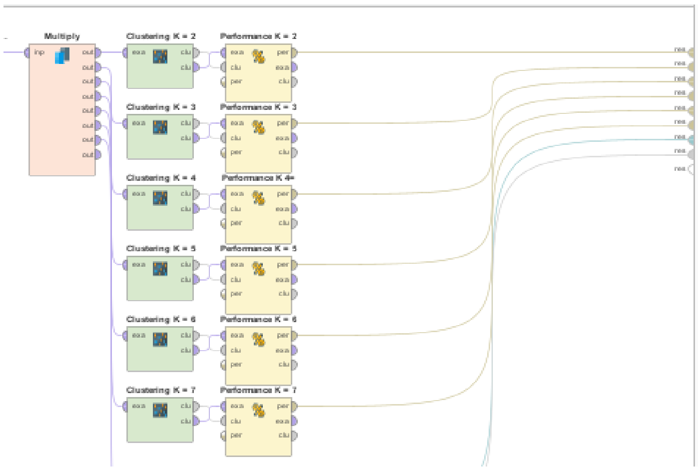


Fig. 6: Evaluate Each Cluster

The figure above shows a flow chart of the clustering process with several different number of clusters (K), ranging from K=2 to K=7. The process begins with a "Multiply" block, which serves as where the input data is processed or prepared before clustering. The data was then divided into several parallel paths to be experimented with with different numbers of clusters. Each "Clustering" block runs a data grouping process for a specific K value, such as K=2, K=3, to K=7. Once the clustering process is complete, the output of each path is evaluated using the "Performance" block, which likely calculates evaluation metrics, such as silhouette score, SSE (Sum of Squared Errors), or other metrics relevant to clustering. The evaluation results of each K value are passed to separate outputs (m1, m2, to m6), allowing performance comparisons to determine the best K value. This diagram illustrates a systematic flow for selecting the optimal number of clusters in the clustering process.

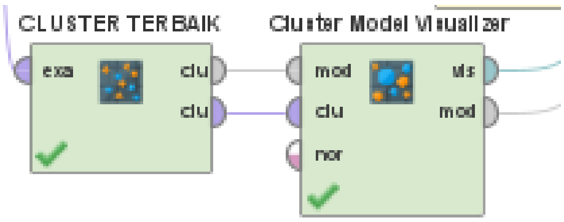


Fig. 7: Best Cluster & Cluster Operator Model Visualizer

Figure 3.6 shows the best cluster results and the Cluster Model Visualizer operator. After carrying out the clustering process, the next step is to visualize the results of the best clusters.

3.4. Evaluate Results

Table 2: ASW Value cluster

K	*Davies Bouldin Index*
2	0.487
3	0.333
4	0.436
5	0.201
6	0.233

Table 2 above shows the Davies-Bouldin Index (DBI) values for the number of clusters (K) from 2 to 7, where a lower DBI value indicates a better and separate cluster. When K = 2, the DBI value is 0.487, then decreases to 0.333, 0.436, 0.201 and 0.233 respectively when K increases to 6. At K = 5, the DBI value reaches 0.201, indicating that the cluster model with K = 5 has excellent cluster separation and compactness, making it the best cluster model based on DBI values.

Cluster Model

```
Cluster 0: 65 items
Cluster 1: 4 items
Cluster 2: 7 items
Cluster 3: 5 items
Cluster 4: 32 items
Total number of items: 113
```

Fig. 8: Cluster Model K = 5

Figure 8 above shows the results of a clustering model that groups data into five different clusters. Cluster 0 has the largest number with 65 items, followed by Cluster 4 with 32 items. Cluster 2 has 7 items, while Cluster 3 contains 5 items, and Cluster 1 is the smallest cluster with only 4 items. In total, there are 113 items that have been successfully grouped in this process. This model is used to identify patterns or groups in data based on specific characteristics.

A. Euclidean Distance Calculation

The formula of the Euclidean Distance is as follows:

$$d = \sqrt{\sum_{N} (X_i - Y_i)^2}$$

Known:
$X_1 = 0$
$X_2 = 2$

So:

1. Distance to Cluster 0 with Centroid = (1,1,969)

$d = \sqrt{(0-1)^2 + (2-1.969)^2}$
$d = \sqrt{1 + 0.0009}$
$d = \sqrt{1.009} = 1.0004$

2. Distance to Cluster 1 with Centroid = (0, 3)

$d = \sqrt{(0-0)^2 + (2-3)^2}$
$d = \sqrt{0+1}$
$d = \sqrt{1} = 1.0$

3. Distance to Cluster 2 with Centroid = (0, 0.286)

$d = \sqrt{(0-0 + (2-0.286)^2}$
$d = \sqrt{0 + 2.937}$
$d = \sqrt{2.937} = 1.713$

4. Distance to Cluster 3 with Centroid = (1, 3)

$d = \sqrt{(0-1)^2 + (2-3)^2}$
$d = \sqrt{1+1}$
$d = \sqrt{2} = 1.414$

5. Distance to Cluster 4 with Centroid = (0, 0)

$d = \sqrt{(0-0)^2 + (0-0)^2}$
$d = \sqrt{0+0}$
$d = \sqrt{0} = 0.0$
number of points in cluster∑

B. Data With Centroid

$$\boldsymbol{Centroid} = \frac{\sum data\ points\ in\ cluster}{data\ points\ in\ cluster}$$

1) Cluster 0 (65 items):
Quantity: 65
Total Number of Attribute Values:
b.1. Number of Values = 65
b.2. Number of Values = 128
Centroid Cluster 0: $Centroid = \frac{65}{65}, \frac{128}{65}, = (1.1.969)$

2) Cluster 1 (4 items):
Quantity: 4
Total Number of Attribute Values:
b.1. Number of Values = 0
b.2. Number of Values = 12
Centroid Cluster 1: $Centroid = \frac{0}{4}, \frac{12}{4}, = (0, 3)$

3) Cluster 2 (7 items):
Number : 7
Total Number of Attribute Values:
b.1. Number of Values = 0
b.2. Number of Values = 2
Centroid Cluster 2: $Centroid = \frac{0}{7}, \frac{2}{7} = (0, 0.286)$

4) Cluster 3 (5 items):
Number : 5
Total Number of Attribute Values: 5
b.1. Total Values = 5
b.2. Total Values = 15
Centroid Cluster 3: $Centroid = \frac{5}{5}, \frac{15}{5} = (1, 3)$

5) Cluster 4 (32 items):
Number : 32
Total Number of Attribute Values:
b.1. Total Values = 0
b.2. Total Values = 64
Centroid Cluster 4: $Centroid = \frac{0}{32}, \frac{64}{32} = (0, 2)$

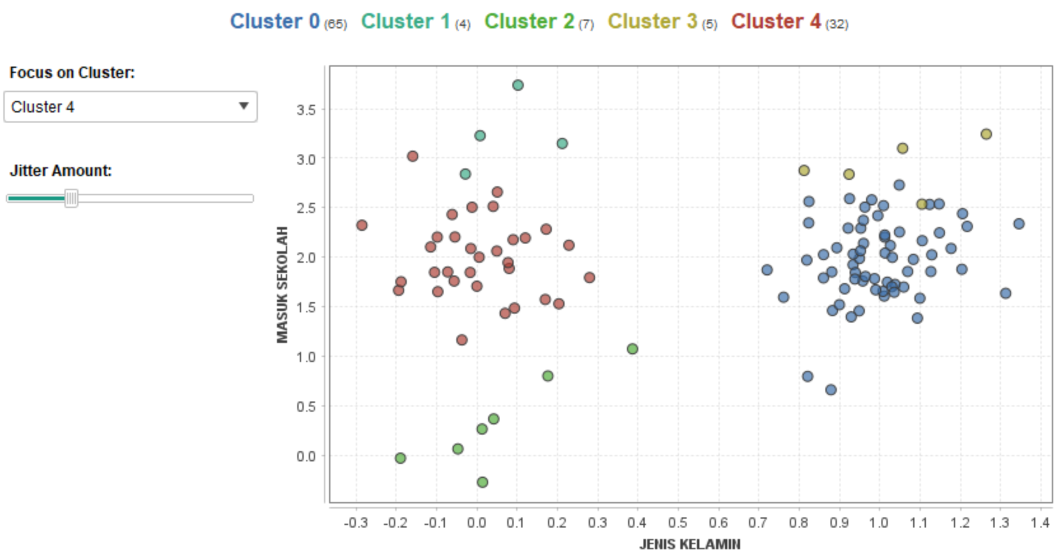


Fig. 9: Scater Plot Visualization

Figure 9 is a scatter plot that displays the results of data clustering, with each point representing one data or observation, and its position is determined by two variables: gender and school attendance status. The X axis (horizontal) represents the gender variable, where the further to the right indicates a higher proportion of males, while the Y axis (vertical) represents the school attendance status, where the higher up indicates the proportion of individuals who are more enrolled in school. Each point on the graph is given a different color and shape to indicate the cluster it represents. Interpretation of the clustering results showed that points with similar characteristics tended to cluster together, indicating that the data had been well divided into homogeneous groups. Cluster 0, for example, appears to have a fairly diverse gender distribution and varied school admission scores, while cluster 1 may exhibit different characteristics, both in terms of gender and school admission status. The relative position between clusters can provide clues about the relationship between gender and school attendance status, such as if one cluster is predominantly male with high school entrance scores, which could indicate a certain relationship between the two variables in that group.

Cluster	JENIS KELAMIN	MASUK SEKOLAH
Cluster 0	1	1.969
Cluster 1	0	3
Cluster 2	0	0.286
Cluster 3	1	3
Cluster 4	0	2

Fig. 10: Centroid Every Cluster

Figure 10 above illustrates the Centroid of each cluster showing the average position of data points in two dimensions. Cluster 0 has a centroid at the point (1, 1.969), which indicates that the data in this cluster is centered around the values X = 1 and Y = 1.969. Cluster 1, with centroid (0, 3), has data centered around the values X = 0 and Y = 3, indicating a higher position on the Y axis. Cluster 2, with centroid (0, 0.286), shows data centered around X = 0 and a lower Y, which is 0.286. Meanwhile, cluster 3 is at point (1, 3), which means that the data in this cluster is spread around X = 1 and Y = 3, higher on the Y axis than other clusters. Finally, cluster 4, with centroid (0, 2), centered on X = 0 and Y = 2, shows a lower position in X but higher on the Y axis than cluster 2. Overall, the position of the centroid on the X-axis varies between 0 and 1, while on the Y-axis, the centroid value ranges from low (0.286) to higher (3).

4. Conclusions

This research makes an important contribution to the management of education data based on the K-Means algorithm. The main findings are summarized as follows:
This study successfully applied the K-Means algorithm to improve the school type grouping model at the Kebon Kelapa Al-Ma'rifah Foundation, Cirebon Regency. With this approach, student data was successfully grouped into five optimal clusters based on demographic attributes such as gender, region, and time of entry.
The evaluation using the Davies-Bouldin Index showed the best results with a DBI value of 0.201 in the configuration of five clusters. This shows a significant level of separation between clusters and good internal cohesiveness.
The findings of the study provide practical benefits in supporting more effective resource allocation, needs-based curriculum development, and improving the efficiency of the management of educational institutions under the foundation.

5. suggestions

To strengthen the results and provide direction for further research, several recommendations were put forward as follows:
For further research, it is recommended to explore alternative clustering algorithms such as K-Medoids or DBSCAN for results validation and see the performance of other algorithms in the context of similar data.
This research can be expanded by adding new attributes, such as academic achievement and socioeconomic conditions of students, to improve the quality of clustering and more in-depth analysis.
It is recommended to utilize the results of this clustering in strategic decision-making, such as planning educational programs, promotions, or adjusting the curriculum based on cluster characteristics.

References

[1] R. Fauziah and A. I. Purnamasari, "Implementation of K-Means Algorithm in Cases of Violence against Children and Women Based on Age," *Hello World J. Computation.*, vol. 2, no. 1, pp. 34–41, 2023, doi: 10.56211/helloworld.v2i1.232.

[2] E. M. Fitri, R. R. Suryono, and A. Wantoro, "Clustering Sales Data by Region Using the K-Means Method on Pt XYZ," *J. Computing*, vol. 11, no. 2, pp. 157–168, 2023, doi: 10.23960/computing.v11i2.12582.

[3] R. D. Romadhona and A. A. Rofiq, "K-Means Clustering Algorithm as a Determination of PIP Scholarship at SMPN 9 Blitar City," *JASIEK (Apl. Science, Information, Electron. and Computers)*, vol. 5, no. 1, pp. 25–30, 2023, doi: 10.26905/jasiek.v5i1.10162.

[4] I. D. Murti Suyoto, T. Rachmadi, and L. T. Parulian, "Determining the Right Cluster with K-Means in the Context of Measuring the Effectiveness of Budget Implementation at the Ministry of Agrarian and Spatial Planning/Land Agency," *Infotech J. Technol. Inf.*, vol. 8, no. 1, pp. 13–22, 2022, doi: 10.37365/jti.v8i1.126.

[5] Pelsri Ramadar Noor Saputra and A. Chusyairi, "Comparison of Clustering Methods in Grouping Puskesmas Data on Complete Basic Immunization Coverage," *J. RESTI (Engineering, Sis. and Technol. Information)*, vol. 4, no. 6, pp. 5–12, 2020, doi: 10.29207/resti.v4i6.2556.

[6] W. W. Kristianto and C. Rudianto, "Application of Data Mining in Product Sales Using the K-Means Clustering Method (Case Study of Footwear Stores)," no. 5, pp. 90–98, 2022.

[7] K. Jukandika, D. Hartama, R. Dewi, S. R. Andani, and Irawan, "Determining Classes Using the K-Means Method Based on Student IQ Scores at Mora College Pematangsiantar Tutoring," *JUKI J. Komput. and Inform.*Vol. 3 No. 2 pp. 69–75, 2021, doi:10.53842/zuki.v3i2.65.

[8] D. Riana *et al.*, "Identification of Pap Smear Images of RepoMedUNM Using K-Means Clustering and GLCM," *J. RESTI (Engineering, Sis. and Technol. Information)*, vol. 6, no. 1, pp. 1–8, 2022, doi: 10.29207/resti.v6i1.3495.