International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V71I7P226

Volume 71 Issue 7, 269-278, July 2023

Original Article

The Application of Data Mining to Information and Computer Technology Skills using the K-Medoid Method

Mirsa Umiyati[1], Yorina An'guna Bansa[2], Hidayatulah Himawan[3], Gusti Ayu Sri Puja Warnis Wijayanti[4], Robbi Rahim[5*]

[1]*Universitas Warmadewa, Bali, Indonesia*
[2]*Universitas Muhammadiyah Jambi, Jambi Indonesia*
[3]*UPN "Veteran" Yogyakarta, Yogyakarta, Indonesia*
[4]*Poltekkes Kemenkes Mataram, Indonesia*
[5]*Sekolah Tinggi Ilmu Manajemen Sukma, Medan, Indonesia*

[*]*Corresponding Author : usurobbi85@zoho.com*

Received: 11 May 2023 Revised: 21 June 2023 Accepted: 29 June 2023 Published: 21 July 2023

Abstract *- The purpose of the study was to classify regions that had a proportionate level of Youth and Adults with Information and Computer Technology Skills (abbreviated as ICT) using data mining algorithms. Data is obtained from the Indonesian Central Statistics Agency (abbreviated BPS-RI), which is processed using the help of RapidMiner software. The data used is data on the proportion of Adolescents and Adults with ICT Skills in Indonesia, which consists of 34 regions ranging from Sabang to Merauke. The settlement method is the K-Medoid method. Data in clustering in two parts, among others: low cluster level (C1) and high cluster level (C2). Results obtained from 34 records there are 3 regions in the low cluster (C1), including East Nusa Tenggara, North Maluku, Papua and 31 regions in the high cluster (C2), among others: Aceh, North Sumatra, West Sumatra, Riau, Jambi, South Sumatra, Bengkulu, Lampung, Bangka Belitung, Riau Islands, DKI Jakarta, West Java, Central Java, East Java, In Yogyakarta, Banten, Bali, West Nusa Tenggara, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Central Sulawesi, South Sulawesi, West Sulawesi, Southeast Sulawesi, Gorontalo, Maluku, West Papua. This can be input to the government in providing information about areas with low Information and Communication Technology skills so that they can be improved so that adolescents, adults who will become the nation's successors, do not become communities that are left behind in Information and Communication Technology.*

Keywords *- Data mining, K-Medoid, Information and Computer Technology, Population, Indonesia.*

1. Introduction

During the Fourth Industrial Revolution, the expansion of information technology has been remarkable, providing us with real-time access to information from Indonesia and around the globe. This growth has had a significant positive impact on education, particularly for adolescents and adults. Information technology, which involves the processing, organization, and storage of data, generates timely and accurate information of the highest quality. By enhancing the standard of education, Information and Communication Technology (ICT) empowers individuals to tackle global challenges effectively.

However, in Indonesia, a developing country, the distribution of communication infrastructure is uneven, particularly in remote areas[1], [2]. This disparity limits access to information and knowledge, hindering educational development for certain segments of the population[3]–[6]. This situation has sparked interest among researchers to explore the potential of computer science techniques in addressing this issue. The primary objective is to classify regions into categories according to the percentage of adolescents and adults who are proficient in ICT[7]–[10].

Complex issues can be tackled by multiple subfields in computer science, each of which has its own set of capabilities as well as limitations. This is seen in experiments employing data mining[11], [12], artificial neural networks[13]–[16], and decision support systems[17]–[21]. Techniques within the field of data mining, such as classification, association, estimate, and forecasting, were utilized to accomplish this study's objectives. Clustering, more specifically the k-medoids method, was the method of mapping that was utilized[22], [23].

Previous studies have shown the effectiveness of k-medoids in problem-solving, such as heart disease prediction[31]. Its easy implementation and short processing time make it an ideal choice for this study. The results of this research, which will be presented to the government, will highlight areas where ICT skills are low and need improvement. This will ensure that adolescents and adults, who are the nation's future leaders, are not left behind in ICT development.

There are some recent studies related to the use of information technology, data mining, and k-medoids in various fields, including education. Dharshinni[32], in his research article, explains the application of the K-Medoids clustering method for grouping food security. The K-Medoids algorithm efficiently handles small datasets, finds the most representative points, and can handle outliers.

The study found that the K-Medoids algorithm resulted in a Davies Bouldin Index (DBI) value of 0.062 and a Silhouette Coefficient value of 0.8980, indicating its effectiveness in the context of food security. Another research by Cahyo Crysdian[26], his research focuses on higher education in Indonesia and how colleges and businesses teach information technology. The study found that higher education in Indonesia needs a more proactive triple helix strategy to foster a knowledge society and economy.

The purpose of this study is to evaluate whether or not K-Medoids are effective in grouping regions in Indonesia according to the percentage of adults and adolescents who have abilities in information and communication technology. Although the initial study used K-Medoids as a clustering method, its primary focus was on food security rather than ICT proficiency, as the current research does. The investigation into the methods used in higher education in Indonesia was carried out without the application of data mining strategies. This research proposes a novel way of using the K-Medoids method for the task of assessing the distribution of information and communication technology (ICT) abilities in Indonesia.

The uneven distribution of information and communication technology (ICT) skills among adolescents and adults in different parts of Indonesia is the subject of our study, which focuses on an aspect of the literature that has not been investigated to this point. This research used data mining, including the K-Medoids method, to categorize geographical locations by ICT skill. This method provides unique insight into the distribution of information and communication technology (ICT) abilities across regions. This technique lets us give the government and other stakeholders insightful information. This will help us identify places that need ICT enhancements.

This study is unique since it examines data mining methodologies and Indonesian ICT distribution. Data mining has been applied to food security and other domains. This study has particular significance since it applies these methodologies to ICT skill distribution. The research also uses these methodologies to gain insights that may influence government policies and actions.

2. Methodology

2.1. Data Mining

Data mining includes examining massive amounts of unprocessed data in repositories to find significant patterns, correlations, and trends. Data analysis reveals patterns, correlations, and trends. Data mining accomplishes this. Data mining covers predictive modeling, association analysis, and classification. Data mining uses semi-automated mathematical, statistical, artificial intelligence, and machine learning approaches to extract useful information from massive databases. This includes machine learning and AI.

Data mining helps uncover predicted insights from massive data warehouses. These insights can aid future projections. This strategy has been used for a while. The technology helps businesses prioritize vital data in data warehouses. Data mining software helps companies predict future trends and take proactive, well-informed action. Businesses can now explore their databases and find hidden patterns thanks to solutions to formerly time-consuming business questions.

Knowledge discovery involves data purification, integration, selection, transformation, data mining, pattern evaluation, and display. This approach involves data mining. The goal is to extract high-quality data from enormous amounts of data to improve data comprehension and corporate decision-making.

2.2. K-Medoids

A dataset consisting of n items can be clustered using the K-Medoids method, which is a type of clustering methodology. A group of data objects comparable to each other within the same cluster but not similar to objects in other clusters is referred to as a cluster. The point in a cluster that is the most centrally placed is called the medoid. This object is defined as the one whose average dissimilarity to all of the other objects in the cluster is the smallest[27,29,33].

2.3. Data

The steps used in this study are:

2.3.1. Data Collection Stage

The data needed in this study is the data Proportion of adolescents and adults with information and communication technology (ICT) skills by region obtained from an official website https://www.bps.go.id. The data collected is in the form of the proportion of adolescents and adults with information and communication technology (ICT) skills by region of Indonesia starting from 2016-2019.

2.3.2. Data Processing Stage

Authors prepare previously collected data for eventual usage in data processing. RapidMiner program will process the data into two clusters with numerous phases.

2.3.3. Clustering Stage

Objects will be grouped into one or more clusters so that objects in one cluster will have a high similarity to one another.

2.3.4. Analysis Phase

At this stage, the results of data analysis were carried out on a proportion of adolescents and adults with information and communication technology (ICT) skills by region using the RapidMiner application.

Table 1. The proportion of adolescents and adults with information and communication technology (ICT) skills

The province	2016	2017	2018	2019
Aceh	23.31	30.56	40.47	46.77
North Sumatra	25.99	35.11	43.65	51.78
West Sumatra	32.53	38.03	47.49	52.85
Riau	32.33	39.78	49.45	55.37
Jambi	27.03	32.8	43.42	50.83
South Sumatra	25.2	32.03	41.33	46.5
Bengkulu	26.34	32.9	40.42	48.7
Lampung	20.87	28.36	40.23	48.37
Kep. Bangka Belitung	28.7	35.31	45.45	54.93
Kep. Riau	50.1	58.87	65.6	77.18
DKI Jakarta	58.4	71.39	77.14	85.17
West Java	34.84	46.09	55.91	65.37
Central Java	29.89	38.75	48.63	58.75
DI Yogyakarta	49.23	57.37	68.82	75.04
East Java	29.59	38.76	48.07	57.23
Banten	37.01	45.49	57.86	66.96
Bali	41.78	48.33	57.71	65.48
West Nusa Tenggara	23.71	30.04	37.11	47.85
East Nusa Tenggara	18.92	25.3	29.65	36.33
West Kalimantan	24.66	30.38	38.92	47.04
Central Kalimantan	28.52	35.43	43.17	54.54
South Borneo	32.61	37.37	49.32	57.82
East Kalimantan	46.11	50.56	60.85	69.44
North Kalimantan	38.5	45.68	58.42	65.36
North Sulawesi	37.2	44.7	51.22	57.48
Central Sulawesi	22.99	31.7	37.02	44.13
South Sulawesi	31.37	38.74	47.07	54.85
Southeast Sulawesi	28.27	35.14	43.94	53.36
Gorontalo	27.3	34.39	42.71	50.62
West Sulawesi	20.86	26.24	33.95	40.95
Maluku	27.55	31.55	39.2	44.02
North Maluku	19.21	25.1	34.24	38.11
West Papua	26.08	34.68	45.41	52.37
Papua	15	21.29	24.23	26.45

3. Results and Discussion

k-medoids work well on tiny datasets. The k-means approach is susceptible to outlier mistakes when an object has a high value. This method addresses this issue. K-medoids first find the dataset point that best represents the entire by computing the distances between each group using every conceivable combination.

Unlike the k-means method, the k-medoids method can overcome noise and outliers. In mapping, data sources were obtained from the Central Statistics Agency (BPS-RI) through https://www.bps.go.id.

3.1. Centroid Data

Before determining clustering, the number of clusters (k) is calculated using the Davies-Bouldin Index (DBI). DBI is the best cluster grouping reference by looking at the minimum value of DBI. The following DBI values for k = 2, k = 3 and k = 4 as shown in the following graph:

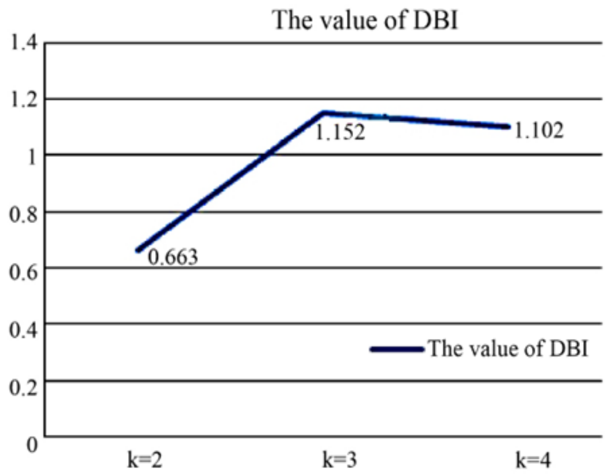


Fig. 1 Graphic comparison of DBI values for each cluster value (k)

Figure 1 shows the lowest DBI value of 0.663 for k = 2. The cluster value used was 2 labels. The k-medoids approach allows the centroid value to be chosen arbitrarily or randomly from the data as long as the number of clusters is 2 (k = 2), namely a cluster of people with low ICT skill (C1) and a cluster of people with high ICT skill (C2).

This yields a 2-point centroid value. Random selection determines cluster points. Initial data will be processed using the Euclidian Distance algorithm and a computed centroid. The first centroids can be seen in Table 2.

Table 2. Initial data centroids (Iteration 1)

C1 (low cluster)	19.21	25.1	29.65	38.11
C2 (high cluster)	27.3	34.39	43.94	52.85

Table 3 below is a tabular representation of the results from the calculation after the first iteration, and the table represents all province names and Euclidean distance results.

Table 3. Data from mapping results in iteration 1

The province	C1	C2	Euclidean distance
Aceh	225.9796	67.6662	67.6662
North Sumatra	489.849	3.0574	3.0574
West Sumatra	716.0381	31.0821	31.0821
Riau	918.57	70.7926	70.7926
Jambi	418.5213	7.1489	7.1489
South Sumatra	260.8294	54.8042	54.8042
Bengkulu	296.111	32.793	32.793
Lampung	229.4916	76.6254	76.6254
Kep. Bangka Belitung	646.2865	8.8529	8.8529
Kep. Riau	3990.17	1683.175	1683.175
DKI Jakarta	6651.898	3546.922	3546.922
West Java	1888.905	444.4613	444.4613
Central Java	983.2525	78.4057	78.4057
DI Yogyakarta	3969.487	1661.421	1661.421
East Java	901.8464	57.6282	57.6282
Banten	2061.679	525.7785	525.7785
Bali	2098.683	557.9334	557.9334
West Nusa Tenggara	179.4228	94.1614	94.1614
East Nusa Tenggara	3.4984	568.1226	3.4984
West Kalimantan	199.0062	77.6766	77.6766
Central Kalimantan	568.7542	5.7506	5.7506
South Borneo	939.3459	67.8357	67.8357
East Kalimantan	2630.121	841.4551	841.4551
North Kalimantan	2013.102	504.8346	504.8346
North Sulawesi	1242.612	190.6314	190.6314
Central Sulawesi	137.8973	135.4709	135.4709
Sulawesi Selatan	781.8936	36.7894	36.7894
Sulawesi Tenggara	546.6282	1.7926	1.7926
Gorontalo	421.4578	6.4858	6.4858
Sulawesi Barat	29.5052	314.2726	29.5052
Maluku	176.0731	108.7521	108.7521
Maluku Utara	21.0681	405.7517	21.0681
Papua Barat	550.3716	3.6954	3.6954
Papua	184.0581	1269.354	184.0581

3.2. Grouping Process

The centroid calculates cluster results by finding the closest distance to each item of processed data. Initial data-based grouping for the two groups in iteration 1. Clusters of mapping results are shown in the table 4.

Table 4. Iteration result cluster 1

C1	C2
3.4984	67.6662
29.5052	3.0574
21.0681	31.0821
184.0581	70.7926
	7.1489
	54.8042
	32.793
	76.6254
	8.8529
	1683.1749
	3546.9224
	444.4613
	78.4057
	1661.4209
	57.6282
	525.7785
	557.9334
	94.1614
	77.6766
	5.7506
	67.8357
	841.4551
	504.8346
	190.6314
	135.4709
	36.7894
	1.7926
	6.4858
	108.7521
	3.6954
238.1298	**10983.8796**
11222.0094	

Table 5. Centroid data (Iteration 2)

C1 (low cluster)	15	21.29	24.23	26.45
C2 (high cluster)	26.08	34.68	45.41	52.37

Table 6. Data from mapping results in iteration 2

The province	C1	C2	Euclidean distance
Aceh	770.8829	75.508	75.508
North Sumatra	1220.7277	3.7206	3.7206
West Sumatra	1535.7452	22.2293	22.2293
Riau	1831.6249	57.5816	57.5816
Jambi	1107.1506	10.8161	10.8161
South Sumatra	819.9601	59.0058	59.0058
Bengkulu	903.3107	41.7974	41.7974
Lampung	792.3413	87.9848	87.9848
Kep. Bangka Belitung	1471.6592	9.5721	9.5721
Kep. Riau	5732.3662	1632.3483	1632.3483
DKI Jakarta	8800.9165	3462.577	3462.577
West Java	3153.2688	418.1981	418.1981
Central Java	1958.3916	71.4477	71.4477
DI Yogyakarta	5685.2526	1599.9431	1599.9431
East Java	1835.5449	50.8516	50.8516
Banten	3379.687	495.6567	495.6567
Bali	3402.1929	525.1846	525.1846
West Nusa Tenggara	709.1269	113.22	113.22
East Nusa Tenggara	146.9909	600.8036	146.9909
West Kalimantan	732.0323	90.439	90.439
Central Kalimantan	1361.2313	12.729	12.729
South Borneo	1889.7614	58.7567	58.7567
East Kalimantan	4077.0074	801.9829	801.9829
North Kalimantan	3301.3163	471.4202	471.4202
North Sulawesi	2261.5491	171.3886	171.3886
Central Sulawesi	592.5246	150.2601	150.2601
South Sulawesi	1649.0981	30.6796	30.6796
Southeast Sulawesi	1317.7247	5.5426	5.5426
Gorontalo	1109.6093	11.6566	11.6566
West Sulawesi	338.2016	335.0909	335.0909
Maluku	650.6234	119.5535	119.5535
North Maluku	254.8818	426.7629	254.8818
West Papua	1310.8109	0	0
Papua	0	1310.8109	0

If the deviation < 0, exchange the object with cluster data to construct a new collection of k medoids. The deviation is calculated by subtracting the entire new distance from the total old distance, which includes the distance of each item in each cluster with new medoids members. Because they do not meet k-medoids, these cluster results continue. Remap and find a new centroid.

Table 6 is a tabular representation of the results from the calculation after the second iteration; the table represents all province names and Euclidean distance results.

If the deviation is more than zero, the clustering process is terminated; however, if the deviation is less than zero, the item is swapped out for the cluster data to produce a new set of k objects that are medoids. The method of determining the deviation consists of adding up the whole value of the new distance and subtracting the total distance of the old, which encompasses the distance between each item in each cluster that contains new medoids members. In this case, the cluster results have not been terminated even though they do not fulfill the conditions of the k-medoids. After that, restart the mapping process and establish a new centroid.

Table 7. Results of iteration cluster 2

C1	C2
146.9909	75.508
254.8818	3.7206
0	22.2293
	57.5816
	10.8161
	59.0058
	41.7974
	87.9848
	9.5721
	1632.3483
	3462.577
	418.1981
	71.4477
	1599.9431
	50.8516
	495.6567
	525.1846
	113.22
	90.439
	12.729
	58.7567
	801.9829
	471.4202
	171.3886
	150.2601
	30.6796
	5.5426
	335.0909
	11.6566
	119.5535
	0
736.9636	**10662.0516**
11399.0152	

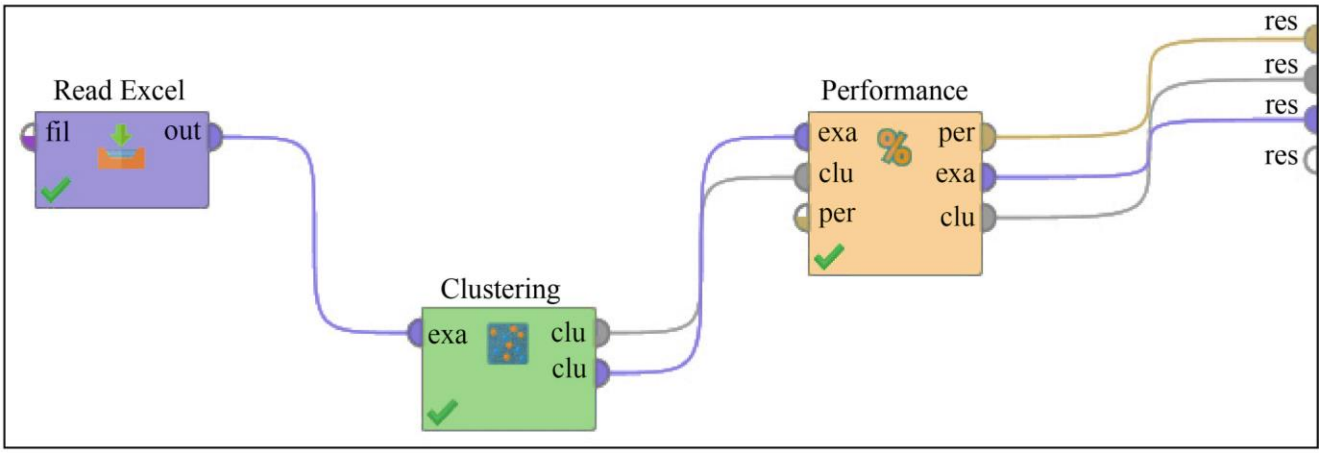


Fig. 2 The Rapid Miner model on ICT mapping

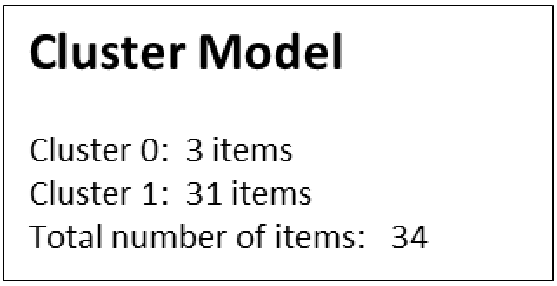
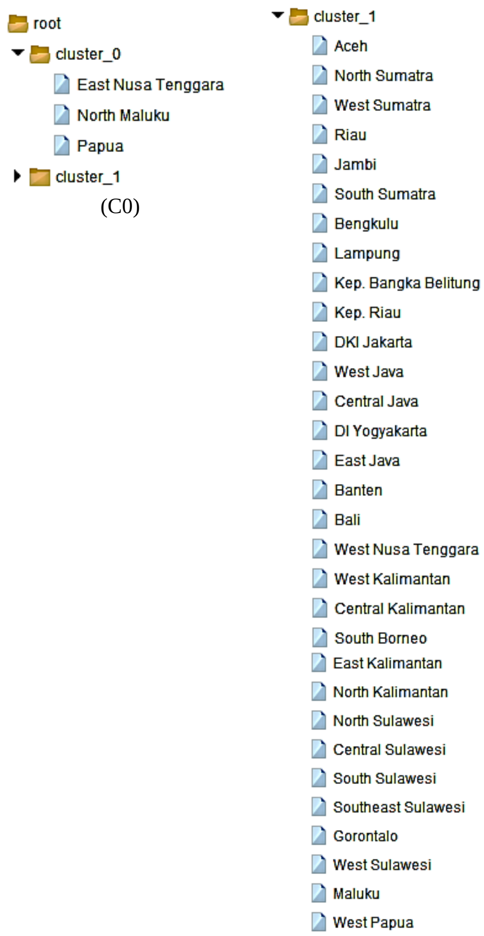


Fig. 3 Results of clustering with k-medoids



(C0)

(C1)

Fig. 4 The low cluster (C0) and high cluster (C1)

Based on the results of the calculation of the k-medoids method and the RapidMiner software assistance trial, the same results are obtained, namely, the low cluster (cluster_0) is 3, and the high cluster (cluster_1) is 31, as shown in the visualization with scatter plotter.

In figure 7, the validity test with the performance vector is performed by looking at the performance value of the Davies-Bouldin Index. The smaller the Davies-Bouldin Index value, the more optimal the cluster results created. In this case, the Davies-Bouldin Index value for k = 2 is 0.663 and is in the good category.

After conducting comprehensive research, it can be confidently concluded that the k-medoids algorithm is an effective strategy for grouping regions based on their proficiency in information technology and computer skills (ICT) among both adolescents and adults. This conclusion was drawn from an exhaustive dataset provided by the Central Statistics Agency (BPS-RI), which covers 34 regions across Indonesia.

The meticulous k-medoids analysis revealed that three regions - East Nusa Tenggara, North Maluku, and Papua - were identified as having the lowest proficiency in ICT skills (C1) throughout Indonesia. 31 other regions were categorized as having the highest level of ICT skills (C2). These regions is Aceh, North Sumatra, West Sumatra, Riau, Jambi, South Sumatra, Bengkulu, Lampung, Bangka Belitung, Riau Islands, DKI Jakarta, West Java, Central Java, East Java, DI Yogyakarta, Banten, Bali, West Nusa Tenggara, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan, North Kalimantan, North Sulawesi, Central Sulawesi, West Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo, Maluku, and West Papua.

The k-medoids algorithm has shown promising results in clustering geographical areas by information and communication technology competencies in young and senior people. This recent breakthrough substantially improves current research procedures and scientific publication conclusions.

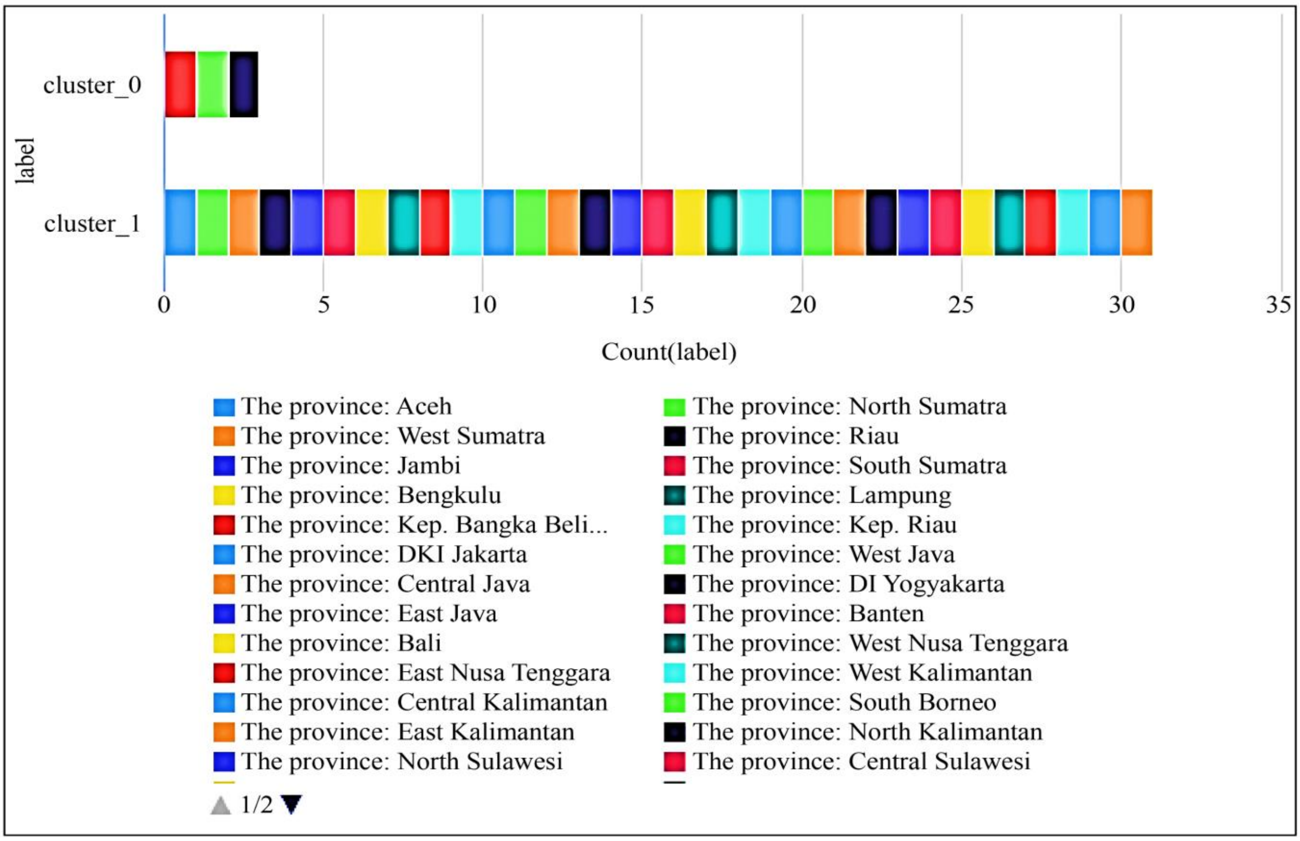


Fig. 5 Scatter Plotter

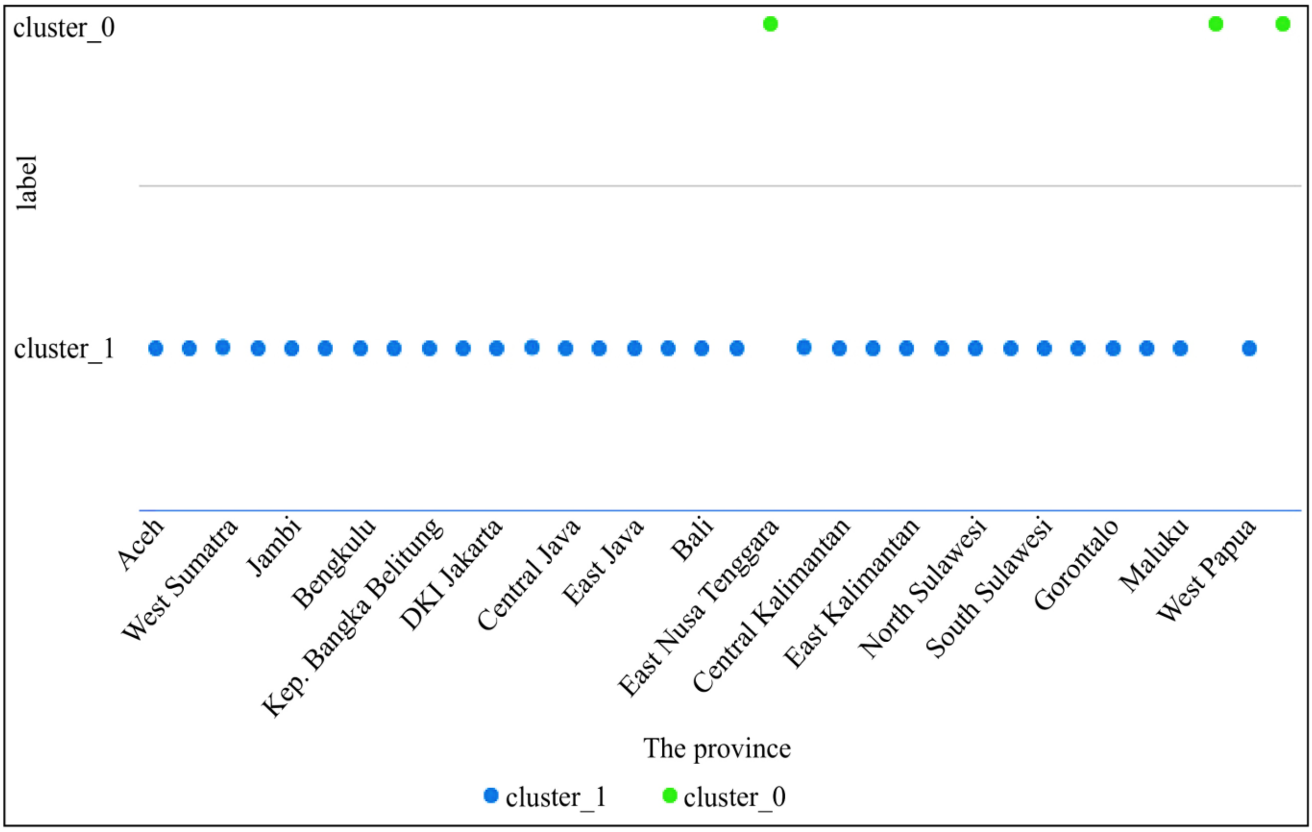


Fig 6. Visualization of clustering results with scatter plotter

Performance Vector

PerformanceVector:
Avg. within centroid distance: -434.243
Avg. within centroid distance_cluster_0: -142.278
Avg. within centroid distance_cluster_1: -462.497
Davies Bouldin: -0.663

Fig. 7 Performance vector results

The k-medoids approach classified Indonesia's 34 regions into the highest and lowest ICT competency groups. This is an advance over prior methodologies, which may not have been precise enough to identify geographical areas by ICT skill.

East Nusa Tenggara, North Maluku, and Papua were identified as having the lowest ICT competency. The research in question presents a valuable contribution to the field due to the uncommon reporting of such a level of detail in the literature.

The k-medoids algorithm employed in this study has exhibited enhanced efficacy in accurately and efficiently clustering regions based on ICT skills, surpassing contemporary methodologies. This phenomenon can be attributed to the algorithm's resilience towards noise and anomalies, alongside its capacity to process extensive datasets.

The study is noteworthy for its practical implications. The findings may be utilized by governmental bodies and pertinent stakeholders to pinpoint regions that necessitate enhancing information and communication technology competencies. The present study represents noteworthy progress compared to prior research endeavors, as the latter may not furnish practical and applicable findings that can be utilized to guide policy-making and intervention strategies.

4. Conclusion

The k-medoids algorithm successfully classified Indonesian adolescents and adults by ICT proficiency. The Central Statistics Agency (BPS-RI) data covered 34 areas, providing a nationwide view of ICT capabilities. The k-medoids research revealed Indonesia's lowest ICT proficiency in East Nusa Tenggara, North Maluku, and Papua. 31 other regions were the most ICT-savvy—the research field benefits from this ICT skill-based region demarcation.

Due to its practical applications, this research stands out from current methods and literature. The government and others can use the findings to identify ICT skill gaps and influence policy and solutions. This research affects academics and has practical applications. Like all studies, this one has limitations. The study relies on BPS-RI data and may not capture all Indonesian adolescents and adults' ICT skills. The k-medoids approach is robust and economical, although it may not be best for all data or research objectives. The dataset should be expanded to include new regions or data sources to understand ICT skills distribution further. To determine the best data mining approach for this research, it would be interesting to compare the results of the k-medoids algorithm with others. This research advances data mining in ICT skills distribution. It takes an innovative approach to a significant issue, perhaps improving ICT skills among Indonesian adolescents and adults and guaranteeing that the nation's future leaders are not left behind in ICT development.

References

[1] Lubna Salsabila, and Eko Priyo Purnomo, “Establishing and Implementing Good Practices E-Government (A Case Study: e-Government Implementation between Korea and Indonesia),” *Journal of Asian Review of Public Affairs and Policy*, vol. 3, no. 3, pp. 36–54, 2018. [Google Scholar] [Publisher Link]

[2] L. Zamzami et al., “Socio-Cultural Impacts of Marine Conservation Areas in Indonesian Fishing Communities,” *IOP Conference Series: Earth and Environmental Science*, vol. 430, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Timothy Ntorukiri Bariu, “Status of ICT Infrastructure used in Teaching and Learning in Secondary Schools in Meru County, Kenya,” *European Journal of Interactive Multimedia and Education*, vol. 1, no. 1, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Abid Haleem et al., “Understanding the Role of Digital Technologies in Education: A review,” *Sustainable Operations and Computers*, vol. 3, pp. 275–285, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] L. Zamzami et al., “Development of Marine Ecotourism in Indonesia: Case of Maligi Nature Reserve, Province of West Sumatra,” *IOP Conference Series: Earth and Environmental Science*, vol. 695, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[6] L. Zamzami et al., “Marine Resource Conservation for Sustainable food security in Indonesia,” *IOP Conference Series: Earth and Environmental Science*, vol. 583, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Mario Liong et al., “Profiles of ICT Identity and Their Associations with Female High School Students’ Intention to Study and Work in ICT: A Mixed-Methods Approach,” *Computers & Education*, vol. 195, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Hairulliza Mohamad Judi et al., “Framework of ICT Impact on Adolescent,” *Procedia Technology*, vol. 11, pp. 1034–1040, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[9] Haiyun Guo et al., "Effectiveness of Information and Communication Technology(ICT) for Addictive Behaviors: An Umbrella Review of Systematic Reviews and Meta-Analysis of Randomized Controlled Trials," *Computers in Human Behaviorvol*, vol. 147, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Dragana Glušac et al., "Adolescents' Informal Computer Usage and Their Expectations of ICT in Teaching – Case Study: Serbia," *Computers & Education*, vol. 81, pp. 133–142, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[11] Shuchita Upadhyaya, and Karanjit Singh "Classification Based Outlier Detection Techniques," *International Journal of Computer Trends and Technology (IJCTT),* vol. 3, no. 2, pp. 294–298, 2012. [Google Scholar] [Publisher Link]

[12] Emmanuel Ahishakiye et al., "Crime Prediction Using Decision Tree (J48) Classification Algorithm," *International Journal of Computer and Information Technology*, vol. 6, no. 3, pp. 188-195, 2017. [Google Scholar] [Publisher Link]

[13] Purwanto C. Eswaran, and R. Logeswaran, "Improved Adaptive Neuro-Fuzzy Inference System for HIV/AIDS Time Series Prediction," *International Conference on Informatics Engineering and Information Science,* vol. 253, pp. 1–13, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[14] Maged M. Elgazzar, and Elsayed E. Hemayed, "Electrical Load Forecasting using Hijri Causal Events," *18th International Middle East Power Systems Conference*, pp. 902-906, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[15] Musa Adebayo Idris et al., "Prediction of Gross Calorific Value of Solid Fuels from their Proximate Analysis using Soft Computing and Regression Analysis," *International Journal of Coal Preparation and Utilization*, vol. 42, no. 4, pp. 1170–1184, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[16] Christian Igel, and Michael Hüsken, "Empirical Evaluation of the Improving the Rprop learning algorithm," *Neurocomputing*, vol. 50, pp. 105-123, 2003. [CrossRef] [Google Scholar] [Publisher Link]

[17] Magdalena Karismariyanti, "Scholarship Recipient Decision Support Simulation using the Composite Performance Index Method," *Jurnal Teknologi Informasi*, vol. 1, no. 2, 2011. [Google Scholar] [Publisher Link]

[18] Yogi Primadasa, and Hengki Juliansa, "Application of the Vikor Method in Selection of Receiving Bonuses for Indihome Salesmen," *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, vol. 10, no. 1, pp. 33–43, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[19] Syed Muzamil Basha et al., "Comparative Study on Performance of Document Classification Using Supervised Machine Learning Algorithms: KNIME," *International Journal on Emerging Technologies*, vol. 10, no. 1, 2019. [Google Scholar] [Publisher Link]

[20] Wahyuni Fithratul Zalmi et al., "Welder Recruitment Decision Support System Using the SMARTER Method," *SAGA Journal of Technology and Information Systems*, vol. 1, no. 2, pp. 44–49, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[21] Ade Yusupa et al., "Decision Support System for Determining the Best PAUD Teacher Using the MOORA Method," SAGA *Journal of Technology and Information Systems,* vol. 1, no. 2, pp. 50–55, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Zulhipni Reno Saputra Elsi et al., "Utilization of Data Mining Techniques in National Food Security during the Covid-19 Pandemic in Indonesia," *Journal of Physics: Conference Series,* vol. 1594, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[23] Robbi Rahim, "Educational Data Mining (EDM) on the Use of the Internet in the World of Indonesian Education," *TEM Journal, ICT Information and Communications Technologies*, vol. 9, no. 3, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] B. P. Ashwini, R. Sumathi, and H. S. Sudhira, "A Dynamic Model for Bus Arrival Time Estimation based on Spatial Patterns using Machine Learning," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 185-193, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[25] Kiran Peddireddy, and Dishant Banga, "Enhancing Customer Experience through Kafka Data Steams for Driven Machine Learning for Complaint Management," *International Journal of Computer Trends and Technology,* vol. 71, no. 3, pp. 7-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[26] Cahyo Crysdian, "The Evaluation of Higher Education Policy to Drive University Entrepreneurial Activities in Information Technology Learning," *Cogent Education*, vol. 9, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Cahyani Oktarina, Khairil Anwar Notodiputro, and Indahwati 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. [CrossRef] [Google Scholar] [Publisher Link]

[28] Johnsymol Joy, "Overview of Different Data Clustering Algorithms for Static and Dynamic Data Sets," *SSRG International Journal of Computer Science and Engineering*, vol. 5, no. 3, pp. 1-3, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[29] Daffa Rafif Agustian, and Budi Arif Darmawan "Analisis Clustering Demam Berdarah Dengue Dengan Algoritma K-Medoids (Studi Kasus Kabupaten Karawang)," *JIKO (Jurnal Informatika dan Komputer)*, vol. 6, no. 1, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[30] D. Arul Selve, K. Kavitha, "Cluster Based Resource Allocation Using K-Medoid Clustering Algorithm," *SSRG International Journal of Computer Science and Engineering*, vol. 3, no. 5, pp. 10-13, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[31] D. Shiny Irene, T. Sethukarasi, and N. Vadivelan, "Heart Disease Prediction Using Hybrid Fuzzy K-Medoids Attribute Weighting Method with DBN-KELM based Regression Model," *Medical Hypotheses*, vol. 143, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[32] N.P. Dharshinni, and Ciok Fandi, “Penerapan Metode K-Medoids Clustering Untuk Mengelompokkan Ketahanan Pangan,” *Jurnal Media Informatika Budidarma*, vol. 6, no. 4, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[33] Leonardo Purba, Saifullah Saifullah, and Rafiqa Dewi, “Clustering of AIDS Cases by Province Using K-Medoids Clustering Data Mining,” *KOMIK (Konferensi Nasional Teknologi Informasi dan Komputer),* vol. 3, no. 1, 2019. [CrossRef] [Google Scholar] [Publisher Link]