

PELATIHAN PENGGUNAAN ALGORITMA *K-MEANS CLUSTERING* UNTUK MENGIDENTIFIKASI KARAKTERISTIK SISWA

Mokhammad Ridwan Yudhanegara¹, Karunia Eka Lestari²

^{1,2}Pendidikan Matematika, Universitas Singaperbangsa Karawang
mridwan.yudhanegara@staff.unsika.ac.id¹, karunia@staff.unsika.ac.id²

ABSTRAK

Salah satu bentuk inovasi di bidang pendidikan adalah pengembangan aplikasi yang dapat digunakan baik dalam pembelajaran maupun administrasi sekolah. Pengelompokan siswa berdasarkan kondisi siswa dengan berbagai macam karakteristiknya sangatlah penting. Informasi dari pengelompokan tersebut dapat dijadikan sebagai acuan bahan pertimbangan bagi guru atau pihak sekolah untuk melakukan langkah secara terstruktur dalam rangka meningkatkan kualitas siswa dan standar mutu sekolah. *Clustering* adalah metode yang dapat digunakan untuk membagi kumpulan data (objek) ke dalam kelompok (*cluster*) berdasarkan kemiripan antar objek. Kegiatan pelatihan yang merupakan bagian dari program pengabdian kepada masyarakat ini dilaksanakan di SDN Karanganyar 01 Karawang. Kegiatan pelatihan ini diikuti oleh 25 peserta yang sebagian besar adalah Guru. Data yang digunakan berupa data hasil nilai mata pelajaran IPA dan IPS siswa. Data dianalisis menggunakan metode *k-means clustering* menggunakan aplikasi XLSTAT. Pelatihan pengklasteran dengan menggunakan metode *k-means* diharapkan mempermudah para guru maupun tenaga pendidik dalam mengidentifikasi karakteristik siswa.

Kata Kunci: *k-means, clustering, karakteristik siswa.*

ABSTRACT

One type of innovation in the field of education is the development of applications that can be used both in learning and school administration. Grouping students based on student conditions with various characteristics is very important. Information from the grouping can be used as a reference for consideration for teachers or schools to take structured steps in order to improve the quality of students and school quality standards. Clustering is a method that can be used to divide a collection of data (objects) into groups (clusters) based on similarities between objects. This training activity which is part of the community service program was held at SDN Karanganyar 01 Karawang. This training activity was attended by 25 participants, most of whom were teachers. The data used are the results of students' science and social studies scores. The data was analyzed using the k-means clustering method using the XLSTAT application. Training in the use of k-means is expected to make it easier for teachers and educators to identify student characteristics.

Keywords: *k-means, clustering, student characteristics.*

1. PENDAHULUAN

Salah satu bentuk inovasi di bidang pendidikan adalah pengembangan aplikasi yang dapat digunakan baik dalam pembelajaran maupun administrasi sekolah. Membentuk kelompok belajar siswa dalam kelas sesuai dengan kriteria yang ada dapat dilakukan sebagai upaya meningkatkan mutu pendidikan melalui layanan pembelajaran (Sulistiyawati & Supriyanto, 2020).

Pengelompokan siswa berdasarkan karakteristiknya sangatlah penting dilakukan, karena informasi yang dihasilkan dapat dijadikan sebagai acuan bahan pertimbangan bagi guru dan pihak sekolah untuk melakukan langkah secara terstruktur dalam rangka meningkatkan kualitas dan standar mutu jenjang sekolah yang ditempati. *Clustering* adalah metode yang dapat digunakan untuk membagi kumpulan data (objek) ke dalam

kelompok (*cluster*) berdasarkan kemiripan antar objek (Yudhanegara & Lestari, 2019) (Yudhanegara et al., 2020a) (Yudhanegara et al., 2020b) (Yudhanegara et al., 2021). Menurut Yudhanegara et al. (2020a) bahwa prinsip utama dari metode *clustering* menggunakan *k-means* adalah menyusun *k* buah partisi atau pusat massa (*centroid*) dari sekumpulan data. Pengelompokan *cluster* yang memiliki kesamaan dan memisahkan perbedaan dapat dilakukan dengan menerapkan algoritma *k-means*. Algoritma *k-means clustering* merupakan bidang kajian dalam analisis *data mining* berdasarkan pada kesamaan data yang tidak terkait (*unsupervised*) (Rencher & Christensen, 2012).

Algoritma *k-means* dapat digunakan untuk mengetahui kelompok siswa berdasarkan aspek kognitif maupun kemampuan ekonomi siswa. Kelompok siswa dalam aspek kognitif seperti mengurutkan nilai siswa dalam mata pelajaran IPA dan IPS disekolah sehingga dapat bermanfaat dalam mengambil kebijakan. Untuk pengelompokan secara kemampuan ekonomi siswa misalnya dapat dilakukan dengan cara mengambil data/informasi penghasilan orang tua siswa, pengeluaran, dan jumlah anggota keluarga.

K-means clustering telah digunakan dalam beberapa penelitian yang memiliki titik berat pada hal pengelompokan. Diantaranya disebutkan oleh Nur et al. (2017) yang menggunakan penerapan algoritma *k-means* pada siswa baru sekolah menengah kejuruan untuk pengelompokan jurusan. Parjito & Permata (2021) menggunakan teknik *k-means clustering* dalam penentuan nilai huruf pada permainan susun kata. Dwitri et al. (2020) menggunakan metode algoritma *k-means* dalam menentukan tingkat penyebaran pandemi Covid-19 di Indonesia.

Pengklasteran dengan metode *k-means* akan memudahkan para pendidik/guru dalam mengidentifikasi karakteristik siswa, walaupun data yang disajikan berukuran besar dan dengan fitur/variabel yang beragam. Metode *k-means* ini juga sangat aplikatif sehingga banyak dijumpai dalam berbagai perangkat lunak pengolahan data/ statistika. Setelah dilakukan observasi langsung terhadap para guru yang ada di desa Karanganyar kecamatan Klari kabupaten Karawang, hampir semua guru tidak mengetahui metode pengklasteran. Begitu juga dengan salah satu metode pengklasteran yaitu *k-means*, semuanya tidak menganal. Untuk itu program pengabdian kepada masyarakat melalui pelatihan penggunaan *k-means clustering* untuk mengidentifikasi karakteristik siswa dengan menggunakan aplikasi XLSTAT sangat diperlukan bagi tenaga pendidik dan kependidikan.

2. METODE PELAKSANAAN PENGABDIAN

Tahapan pelaksanaan dari kegiatan pengabdian masyarakat ini meliputi beberapa tahap yakni: 1. Sosialisasi kegiatan, 2. Analisis Data, 3. Penyajian materi, 4. Pelaksanaan *Workshop* XLSTAT sekaligus pelatihan pengolahan data karakteristik siswa menggunakan *Microsoft Excel* XLSTAT, dan 5. Evaluasi Pelatihan.

2.1. Sosialisasi Kegiatan

Kegiatan sosialisasi diawali dengan memberikan penjelasan mengenai metode *k-means*. *K-means clustering* dalam mengelompokkan karakteristik siswa dengan menggunakan *Add-In* XLSTAT dari *Microsoft Excel* yang mana diharapkan dengan penggunaan pengelompokkan karakteristik ini, guru dapat lebih baik dalam memahami siswa dalam proses belajar mengajar sehingga mendapatkan hasil belajar yang lebih baik. Karakteristik siswa sangat penting untuk diketahui oleh pendidik, karena ini sangat penting untuk dijadikan acuan dalam merumuskan strategi pengajaran (Septianti & Afiani, 2020).

2.2. Analisis Data

Pada tahap ini, hasil observasi yang diambil dilakukan analisis dan dipresentasikan di SDN 1 Karanganyar. Data yang digunakan adalah data primer, yaitu peneliti mengambil 26 sampel nilai. Tahap cluster data nilai IPA dan IPS di SDN 01 Karanganyar menggunakan metode *k-means* menggunakan perangkat lunak *Add-In* dari *Microsoft Excel*.

2.3. Penyajian Materi

Penyajian materi dilakukan dalam rangkaian acara *workshop*, yang diberikan oleh narasumber yaitu Dr. Mokhammad Ridwan Yudhanegara dan Dr. Karunia Eka Lestari dengan materi XLSTAT (www.xlstat.com, 2022). Pada penyajian materi ini diperkenalkan mengenai metode *k-means clustering* dalam pengelompokkan data. *K-means clustering* adalah pendekatan sederhana untuk mempartisi kumpulan data menjadi *k cluster*

yang berbeda dan tidak tumpang tindih. Untuk melakukan *k-means clustering*, pertama-tama kita harus menentukan jumlah *cluster* yang diinginkan. Metode *k-means* akan menetapkan setiap pengamatan tepat satu dari *k cluster* (Semuel & Reba, 2021).

2.4. Pelaksanaan *Workshop*

Pada tahap ini dilaksanakan praktek penggunaan *k-means clustering* dalam mengelompokkan karakteristik siswa menggunakan data yang sudah didapatkan lalu diambil sampel dan dianalisis. Pengambilan sampel menggunakan metode *purposive sampling*.

2.5. Evaluasi Pelatihan

Untuk mengetahui efektif atau tidaknya pelatihan, sebagai evaluasi maka dari itu diperlukan informasi dari peserta sebagai rencana kegiatan pengabdian kepada masyarakat berikutnya. Pada akhir pelaksanaan kegiatan pelatihan, seluruh peserta diminta untuk mengisi angket respon terhadap pelaksanaan pelatihan (Lestari et al., 2021).

3. HASIL DAN PEMBAHASAN

Kegiatan pengabdian kepada masyarakat ini telah dilaksanakan pada hari Sabtu, 3 September 2022. Kegiatan ini dimulai pukul 10.00 – 13.00 WIB yang bertempat di ruang kelas SDN Karanganyar 01 desa Karanganyar, kecamatan Klari, kabupaten Karawang. Kegiatan pelatihan ini diikuti oleh 25 peserta yang sebagian besar adalah Guru. Adapun penyelenggara kegiatan ini adalah Lembaga Penelitian dan Pengabdian kepada Masyarakat (LPPM) Universitas Singaperbangsa Karawang yang bekerja sama dengan SDN Karanganyar 01 dan pemerintah desa Karanganyar.

3.1. Pembahasan Permasalahan Awal tentang Data

Sebelum pelatihan dilaksanakan, ada tahapan persiapan data. Guru diminta untuk menyediakan data tentang siswa berikut permasalahan yang dialami oleh guru dan siswa. Berdasarkan informasi yang diperoleh, bahwa kesulitan guru dalam mengidentifikasi karakteristik siswa disebabkan oleh banyaknya jumlah siswa, dan kondisi data yang sifatnya dinamis atau sering mengalami perubahan (Yudhanegara & Lestari, 2023). Misalnya seorang siswa terkadang nilai matematikanya bagus dan terkadang jelek. Terdapat beberapa mahasiswa yang selalu tinggi di mata pelajaran bahasa, namun pada mata pelajaran lainnya sangat kurang.

Pada tahapan persiapan data ini dibahas pada awal pelatihan. Permasalahan yang sulit untuk dinyatakan dalam pengklasteran, solusinya adalah dilakukan pengklasteran secara subjektif dengan menggunakan data kualitatif. Data kualitatif tersebut menggunakan angket, wawancara, dan lembar observasi. Teknik pengklasterannya dilakukan dengan cara *clustering* hierarki. Namun tidak secara detail dibahas pada pelatihan. Hal ini dikarenakan memerlukan waktu untuk memperoleh data kualitatif tersebut.

3.2. Kegiatan Penjelasan Analisis *Cluster* dengan *K-Means* dan Simulasi

Pada kegiatan ini dijelaskan cara melakukan pengklasteran menggunakan data nilai pengetahuan siswa. Didapat sebanyak 26 data siswa yang berasal dari SDN Karanganyar 01. Data tersebut berupa data hasil nilai mata pelajaran IPA dan IPS siswa. Data yang didapat diproses dengan menggunakan metode *k-means clustering* pada aplikasi XLSTAT (www.xlstat.com, 2022).

Clustering adalah proses membagi kumpulan data ke dalam kelompok/*cluster*. *Cluster* sendiri merupakan kumpulan objek yang dikelompokkan berdasarkan kemiripan antar objek, ditentukan menggunakan jarak (*distance measure*). Suatu objek dikatakan mirip jika mempunyai jarak yang dekat (kecil). Algoritma *clustering* berfungsi untuk mengelompokkan data sesuai dengan karakteristik dan mengukur jarak kemiripan antar data dalam satu kelompok yang memiliki perbedaan satu dengan lainnya (Marlina et al., 2018).

Menurut Yudhanegara et al. (2020a) *k-means* adalah algoritma untuk mengelompokkan *n* objek/data berdasarkan atribut/fitur ke dalam *k cluster*, dengan $k < n$ dan $k > 0$. Algoritma *k-means* merupakan algoritma mudah diimplementasikan dan relatif cepat jika ditinjau dari waktu. *K-means clustering* sendiri merupakan salah satu metode *clustering* non-hirarki yang bertujuan untuk mengelompokkan data dalam bentuk satu atau lebih kelompok.



Gambar 1. Pelaksanaan Pelatihan

Tahap awal dalam menggunakan metode *k-means clustering* pada aplikasi XLSTAT yaitu menentukan pengelompokan. Pada kasus ini jumlah kelas yang diinginkan sebanyak 3 kelas/*cluster*. Kelas 1 dikatakan sebagai kategori *Average*, kelas 2 sebagai kategori *Good*, dan kelas 3 sebagai kategori *Borderline*. Berikut Tabel 1 dan Tabel 2 yang merupakan hasil *clustering* menggunakan aplikasi XLSTAT untuk pengelompokan nilai IPA dan IPS.

Tabel 1. Kelas *Centroid* Nilai IPA

Class	Nilai IPA	Sum of weights	Within-class variance
1	52,778	9,000	13,194
2	72,273	11,000	41,818
3	37,500	6,000	37,500

Tabel 2. Kelas *Centroid* Nilai IPS

Class	Nilai IPS	Sum of weights	Within-class variance
1	72,500	8,000	21,429
2	53,636	11,000	25,455
3	34,286	7,000	61,905

Setelah mendapatkan nilai *centroid* dari setiap nilai IPA dan IPS, maka dapat diperoleh hasil *cluster* untuk setiap kategori mata pelajaran. Berikut hasil nilai *centroid* akhir yang diperoleh dari aplikasi XLSTAT.

Tabel 3. Nilai *Centroid* Akhir Pada Aplikasi XLSTAT Nilai IPA

Class	1 (<i>Average</i>)	2 (<i>Good</i>)	3 (<i>Borderline</i>)
<i>Objects</i>	10	11	5
<i>Sum of weights</i>	10	11	5
<i>Within-class variance</i>	17,778	41,818	30,000
<i>Minimum distance to centroid</i>	2,000	2,273	4,000
<i>Average distance to centroid</i>	3,400	5,207	4,800
<i>Maximum distance to centroid</i>	8,000	12,727	6,000
	S1	S2	S4
	S6	S3	S8
	S7	S5	S13

<i>Class</i>	1 (<i>Average</i>)	2 (<i>Good</i>)	3 (<i>Borderline</i>)
	S10	S9	S15
	S11	S12	S25
	S20	S14	
	S21	S16	
	S22	S17	
	S23	S18	
	S24	S19	
		S26	

Tabel 4. Nilai *Centroid* Akhir Pada Aplikasi XLSTAT Nilai IPS

<i>Class</i>	1 (<i>Average</i>)	2 (<i>Good</i>)	3 (<i>Borderline</i>)
<i>Objects</i>	8	11	7
<i>Sum of weights</i>	8	11	7
<i>Within-class variance</i>	21,429	25,455	61,905
<i>Minimum distance to centroid</i>	2,500	3,636	4,286
<i>Average distance to centroid</i>	3,750	4,628	6,531
<i>Maximum distance to centroid</i>	7,500	6,364	14,286
	S1	S4	S7
	S2	S5	S12
	S3	S6	S13
	S11	S8	S15
	S14	S9	S18
	S16	S10	S25
	S17	S20	S26
	S19	S21	
		S22	
		S23	
		S24	

Setelah didapatkan nilai *centroid* akhir pada tiap nilai IPA dan IPS, maka dapat diperoleh tabel kontingensi nilai IPA dan IPS.

Tabel 5. Hasil *Contingency Table* Kategori Nilai IPA dan IPS

	<i>Average</i>	<i>Borderline</i>	<i>Good</i>
<i>Average</i>	7	1	2
<i>Borderline</i>	2	3	0
<i>Good</i>	2	3	6

Penjelasan Tabel 5 yaitu terdapat 7 siswa dengan nilai IPA dan IPS masuk kedalam kategori *Average*. Selanjutnya erdapat 1 siswa dengan nilai IPA masuk kedalam kategori *Average* serta nilai IPS masuk kedalam kategori *Borderline*. Kemudian terdapat 2 siswa dengan nilai IPA masuk kedalam kategori *Average* serta nilai IPS masuk kedalam kategori *Good*. Lalu terdapat 2 siswa dengan nilai IPA masuk kedalam kategori *Borderline* serta nilai IPS masuk kedalam kategori *Average*. Terdapat 3 siswa dengan nilai IPA dan IPS masuk kedalam kategori *Borderline*. Setelah itu, terdapat 2 siswa dengan nilai IPA masuk kedalam kategori *Good* serta nilai IPS masuk kedalam kategori *Average*. Terdapat 3 siswa dengan nilai IPA masuk kedalam kategori *Good* serta

nilai IPS masuk kedalam kategori *Borderline*. Terakhir, terdapat 6 siswa dengan nilai IPA dan IPS masuk kedalam kategori *Good*.

3.3. Respon Peserta terhadap Pelatihan

Pada umumnya respon peserta terhadap pelatihan ini adalah positif. Hasil angket peserta terhadap pelatihan disajikan pada Tabel 6 sebagai berikut.

Tabel 6. Hasil Angket sebagai Respon terhadap Pelatihan

Pernyataan	Jumlah Peserta yang Menjawab		
	Setuju	Netral	Tidak Setuju
Pelatihan metode pengklasteran/ <i>clustering</i> dengan menggunakan <i>k-means</i> sangat dibutuhkan oleh saya.	20	3	2
Saya tidak bisa mengikuti kegiatan pelatihan dengan fokus.	2	6	17
Pelatihan ini tidak terlalu penting bagi saya.	2	1	22
Saya senang mengikuti pelatihan.	24	1	-
Saya akan mengaplikasikan hasil pelatihan pada proses pembelajaran.	15	5	5
Saya memerlukan pelatihan metode lainnya dalam untuk pengklasteran.	21	3	1

4. KESIMPULAN DAN UCAPAN TERIMA KASIH

4.1. Kesimpulan

Berdasarkan hasil dan pembahasan yang telah diuraikan di atas tentang pengelompokkan siswa dengan metode *k-means clustering* dengan menggunakan perangkat lunak *Add-In XLSTAT* dari *Microsoft Excel* dapat disimpulkan bahwa pengelompokkan siswa menggunakan metode *k-means clustering* sangat cocok digunakan untuk pengelompokkan karakteristik siswa. Program ini membantu guru untuk mengelompokkan siswa berdasarkan karakteristiknya, tujuannya adalah membantu guru dalam menentukan metode pembelajaran yang sesuai dengan kondisi pada tiap kelompok/*cluster* siswa. Perangkat lunak yang digunakan sangat familiar, yaitu berbasis *Microsoft Excel*, sehingga guru dengan mudah dapat mengimplemmentasikannya secara langsung.

4.2. Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada program HIPLA LPPM Universitas Singaperbangsa Karawang tahun 2022 yang sudah membiaya kegiatan pengabdian kepada masyarakat untuk diimplementasikan kegiatan melalui pelatihan. Selanjutnya terima kasih diucapkan kepada Kepala Desa Karanganyar Kecamatan Klari Kabupaten Karawang dan Kepala SDN Karanganyar 01 yang sudah bersedia menjadi mitra. Penulis juga mengucapkan terima kasih kepada Rivqy Ramdhani Putra Dipa, Shafa Khadijah, dan Risky Nugroho yang sudah membantu jalannya pelatihan.

Daftar Pustaka

- Dwitri, N., Tampubolon, J. A., Prayoga, S., R.H Zer, F. I., & Hartama, D. (2020). Penerapan Algoritma K-Means dalam Menentukan Tingkat Penyebaran Pandemi Covid-19 di Indonesia. *Jurnal Teknologi Informasi*, 4(1), 128–132. <https://doi.org/10.36294/jurti.v4i1.1266>
- Lestari, K. E., Utami, M. R., & Yudhanegara, M. R. (2021). Evaluasi Penyelenggaraan Pendidikan di Perguruan Tinggi: Relevansi Bidang Pekerjaan dengan Program Studi. *JUDIKA (Jurnal Pendidikan UNSIKA)*, 9(2), 149–162. <http://journal.unsika.ac.id/index.php/judika>
- Marlina, D., Lina, N., Fernando, A., & Ramadhan, A. (2018). Implementasi Algoritma K-Medoids dan K-Means untuk Pengelompokkan Wilayah Sebaran Cacat pada Anak. *Jurnal CoreIT: Jurnal Hasil Penelitian Ilmu Komputer dan Teknologi Informasi*, 4(2), 64. <https://doi.org/10.24014/coreit.v4i2.4498>
- Nur, F., Zarlis, M., & Nasution, B. B. (2017). Penerapan Algoritma K-Means pada Siswa Baru Sekolahmenengah Kejuruan untuk Clustering Jurusan. *InfoTekJar (Jurnal Nasional Informatika Dan Teknologi Jaringan)*, 1(2), 100–105. <https://doi.org/10.30743/infotekjar.v1i2.70>

- Parjito, & Permata. (2021). Penerapan Data Mining untuk Clustering Data Penduduk Miskin Menggunakan Metode K-Means. *AINET Jurnal Informatika*, 3(1), 31–37.
- Rencher, A. C., & Christensen, W. F. (2012). *Methods of Multivariate Analysis* (2nd ed.). John Wiley and Sons.
- Semuel, I., & Reba, F. (2021). Strategi Penyelesaian Persoalan Perkalian Sederhana Melalui Pendekatan Geometri bagi Siswa SMA Yapis Nimbokrang Jayapura. *Jurnal Pengabdian Papua*, 5(3), 107–111.
- Septianti, N., & Afiani, R. (2020). Pentingnya Memahami Karakteristik Siswa Sekolah Dasar di SDN Cikokol 2. *As-Sabiqun*, 2(1), 7–17. <https://doi.org/10.36088/assabiqun.v2i1.611>
- Sulistiyawati, A., & Supriyanto, E. (2020). Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan. *Jurnal TEKNO KOMPAK*, 15(2), 25–36.
- www.xlstat.com. (2022). *Xlstat 2022*. Addinsoft.
- Yudhanegara, M. R., Indratno, S. W., & Sari, R. K. N. (2020a). Clustering for Item Delivery Using Rule-K-Means. *Journal of the Indonesian Mathematical Society*, 26(2), 185–191. <https://doi.org/10.22342/JIMS.26.2.871.185-191>
- Yudhanegara, M. R., Indratno, S. W., & Sari, R. K. N. (2021). Dynamic Items Delivery Network: Prediction and Clustering. *Heliyon*, 7(5), e06934. <https://doi.org/10.1016/j.heliyon.2021.e06934>
- Yudhanegara, M. R., Indratno, S. W., & Sari, R. R. K. N. (2020b). Clustering for Items Distribution Network. *Journal of Physics: Conference Series*, 1496(1). <https://doi.org/10.1088/1742-6596/1496/1/012019>
- Yudhanegara, M. R., & Lestari, K. E. (2019). Clustering for Multi-dimensional Data set: A case Study on Educational Data. *Journal of Physics: Conference Series*, 1280(4). <https://doi.org/10.1088/1742-6596/1280/4/042025>
- Yudhanegara, M. R., & Lestari, K. E. (2023). Network Clustering Method for Preventing the Spread of Covid-19 in Indonesian Schools. *Communications in Mathematical Biology and Neuroscience*, 2023, 1–15. <https://doi.org/10.28919/cmbn/7922>