

Penerapan K-Means Clustering Dalam Pengelompokan Prestasi Siswa Dengan Optimasi Metode Elbow

Muhammad Qusyairi^{1*}, Zul Hidayatullah², Arnila Sandi³

^{1,3}program Studi Teknik Komputer, Universitas Hamzanwadi

²program Studi Pendidikan IPA, Universitas Hamzanwadi

*erick.rtr@gmail.com

Abstrak

Prestasi siswa merupakan sesuatu yang sangat penting dan bermakna dalam dunia pendidikan, gambaran ini bisa dilihat dari pengelompokan ruangan kelas, penentuan bakat dan minat serta tingkat kemampuan dan kemauan siswa menurut kemampuan yang dimiliki setiap individu. Hal ini yang mendasari mengapa peneliti melakukan penelitian di MIS NW 03 Pancor dengan data kelas 4 sebanyak 23 siswa, kelas 5 sebanyak 27 siswa dan kelas 6 sebanyak 25 siswa yang terkait dengan variabel-variabel dalam prestasi belajar siswa. Dimana dalam penentuan hal tersebut sekolah masih menggunakan metode pengambilan keputusan secara konvensional, sehingga sekolah ini masih sulit untuk menentukan prestasi siswa sesuai dengan kemampuan mereka. Dengan demikian sekolah perlu melakukan perubahan yang dapat membantu pegawai dan institusi untuk meningkatkan mutu pendidikan dengan mengetahui prestasi belajar siswa. Adapun dari permasalahan ini penulis memberikan solusi yang diperlukan sebuah sistem penghitungan dengan konsep data mining. Penguasaan materi pembelajaran salah satu menjadi tolak ukur untuk mendapatkan nilai siswa dan ini merupakan salah satu kebijakan dalam menentukan prestasi siswa. Selain nilai juga terdapat sikap dan keterampilan yang menjadi faktor dalam penentuan prestasi siswa disekolah. Hal ini bisa diselesaikan dengan salah satu teknik dalam data mining yang dapat dilakukan untuk pengelompokan yaitu clustering. Hasil penelitian ini berupa sistem yang dapat mempermudah sekolah dengan cepat dan tepat dalam mengelompokkan prestasi siswa menggunakan k-means yang dioptimasi dengan penggunaan metode elbow dengan perolehan hasil yang didapat yaitu cluster 1 "baik" sebanyak 9 siswa, cluster 2 "sangat baik" sebanyak 6 siswa, cluster 3 "cukup baik" sebanyak 8 siswa dan nilai davies bouldin 0,711 dan avg within centroid distance 18,821 sehingga bias bisa membantu pihak sekolah dalam menentukan pengelompokan prestasi siswa

Kata kunci: Clustering, Data Mining, K-Means, Elbow, Prestasi Siswa, Sekolah

Abstract

Student achievement is very important and meaningful in the world of education, this picture can be seen from the grouping of classrooms, determining talents and interests as well as the level of ability and willingness of students according to the abilities of each individual. This is why researchers conducted research at MIS NW 03 Pancor with data from class 4 as many as 23 students, class 5 as many as 27 students, and class 6 as many as 25 students which were related to variables in student learning achievement. In determining this matter, the school still uses conventional decision-making methods, so it is still difficult to determine student achievement according to their abilities. Thus, schools need to make changes that can help employees and institutions improve education quality by knowing student learning achievements. As for this problem, the author provides a solution that requires a calculation system with the concept of data mining. Mastery of learning material is one of the benchmarks for obtaining student grades and this is one of the policies in determining student achievement. Apart from grades, there are also attitudes and skills which are factors in determining student achievement at school. This can be solved with one of the techniques in data mining that can be used for grouping, namely clustering. The results of this research are a system that can make it easier for schools to quickly and accurately group student achievement using k-means which is optimized using the elbow method with the results obtained, namely cluster 1 "good" with 9 students, cluster 2 "very good" with 6 students, cluster 3 is "fairly good" as many as 8 students and the Davies Bouldin value is 0.711 and the AVG within centroid distance is 18.821 so it can help the school in determining student achievement groups.

Keywords: *Clustering, Data Mining, K-Means, Elbow, Student Achievement, School*

1. Pendahuluan

Perkembangan teknologi saat ini berkembang dengan sangat pesat ini terbukti dengan dipermudahkannya berbagai macam urusan yang bisa terselesaikan dengan cepat dan tepat. Perkembangan ini juga terlihat pada revolusi industri 4.0 yang dimana dampaknya pada dunia sangat besar terlebih di bidang pendidikan indonesia, dimulai sejak munculnya digitalisasi sistem pendidikan yang mengharuskan setiap bidang pendidikan supaya dapat beradaptasi dengan perubahan yang terjadi sehingga membuat segalanya menjadi lebih efektif dan efisien. Dengan erhadirnya revolusi industri 4.0 memberikan pengaruh positif terhadap dunia pendidikan yang dapat dilihat semakin maju dan berkembang dengan sistem pembelajaran yang memberikan pengaruh besar terhadap sekolah. Jarak dan batasan wilayah tidak menjadi penghalang untuk mengetahui dan mengakses informasi dari dunia luar. Industri 4.0 dalam dunia pendidikan juga memberikan dampak positif yang dapat memajukan dan mengembangkan sistem pembelajaran karena mudahnya data yang diolah menjadi informasi yang bermanfaat bagi semua. Akan tetapi jika tidak digunakan dengan baik maka dapat memberikan dampak negatif bagi dunia pendidikan.

Prestasi siswa merupakan sebuah pencapaian serta tolak ukur sekaligus acuan dalam pengetahuan yang diperoleh dari pendidikan di sekolah, adapun diantaranya secara formal yang dijalani melalui nilai tes, penguasaan terhadap mata pelajaran, serta sikap yang ditentukan melalui nilai atau angka yang diberikan oleh guru kepada siswanya[1]. Prestasi siswa menjadi penting untuk ditelaah mengingat prestasi siswa dapat digunakan sebagai acuan dalam pengambilan keputusan kedepan, diantaranya sebagai berikut [2]:

1. Mengetahui sejauh mana siswa mampu menguasai materi pembelajaran yang telah disampaikan guru selama persekolahan.
2. Mengetahui kemampuan, bakat, minat, dan sikap siswa terhadap siswa dalam pembelajaran serta keunggulan dan kekurangan dalam program yang diterapkan pihak sekolah.
3. Mengetahui tingkat kemajuan perkembangan dan kesesuaian hasil belajar atau prestasi belajar siswa dengan standar kompetensi dan kompetensi dasar yang telah ditetapkan.
4. Membantu dalam menyeleksi serta menyaring siswa yang sesuai dengan jenis pendidikan tertentu.

5. Membantu dalam menentukan kenaikan kelas siswa.

6. Membantu dalam menempatkan siswa sesuai dengan potensi yang dimilikinya.

Dari kajian di atas penulis tertarik untuk melakukan penelitian di MIS NW 03 Pancor dengan variabel-variabel yang berkaitan dengan pengaruh prestasi siswa. Dimana dalam penentuan hal tersebut pihak sekolah masih menggunakan metode pengambilan keputusan secara konvensional dimana dengan pengumpulan rekap data serta penyaringan data yang membutuhkan waktu serta ketepatan yang cukup lama membuat sekolah ini masih sulit untuk menentukan prestasi siswa. Dengan demikian sekolah perlu melakukan perubahan yang dapat membantu pegawai dan institusi untuk meningkatkan mutu pendidikan dengan mengetahui prestasi belajar siswa. Sehingga di sekolah ini dapat mengetahui siswa/siswi yang memiliki prestasi baik dan dapat mengetahui prestasi yang kurang baik untuk menghindari potensi ketidاكلulusan serta dapat mengetahui hal-hal dari kajian yang ingin dicapai.

Tujuan peneliti adalah untuk memudahkan pihak sekolah dalam mengelompokkan siswa yang berprestasi dengan cara yang lebih efektif dan efisien. Oleh karena itu dalam melakukan pengujian ini diperlukan sebuah sistem penghitungan dengan konsep data mining.

Data mining adalah proses yang menggunakan teknik perhitungan statistik, matematika, kecerdasan buatan dan machine learning yang digunakan dalam berinteraksi, mengidentifikasi informasi yang bermanfaat serta pengetahuan yang terikat dari berbagai data yang besar[3]. Penguasaan materi pembelajaran salah satu tolak ukur dalam menentukan nilai siswa dan ini merupakan salah satu kebijakan dalam menentukan prestasi siswa yang didalamnya juga terdapat penilaian dari sikap dan keterampilan. Dari beberapa atribut tersebut maka solusi yang dibutuhkan adalah pengelompokan siswa yang berprestasi dan salah satu teknik dalam data mining yang dapat dilakukan untuk pengelompokan yaitu clustering[4].

Penerapan k-means sudah banyak dilakukan oleh peneliti sebelumnya, salah satunya untuk menganalisa data yang besar dan mengetahui tingkat kepuasan siswa dalam pelayanan melalui pengelompokan data hasil pengisian kuesioner dengan cepat dan efisien. Metode data mining yang melakukan proses pemodelan tanpa supervise (unsupervised) dan merupakan salah satu metode yang melakukan pengelompokan data yang ada kedalam beberapa kelompok[5]. Penerapan ilmu data mining dalam pengelompokan data dengan menggunakan metode k-means akan mempermudah sekolah dalam menentukan siswa-siswi yang berprestasi

dan sistem yang dapat dibangun dapat mempermudah sekolah dengan cepat dan tepat dalam pengelompokan prestasi siswa. Penentuan clustering tidak hanya hanya dibuat serta merta melainkan ada beberapa faktor yang mempengaruhi hal tersebut. Untuk mendapatkan nilai yang optimal maka dibutuhkan sebuah metode optimasi yang dikenal dengan *elbow method*. Metode elbow merupakan suatu metode yang digunakan untuk menghasilkan informasi dalam menentukan jumlah cluster yang akan membentuk siku pada suatu titik[6].

2. Tinjauan Pustaka

2.1. Penelitian Terkait

Berikut ini terdapat beberapa penelitian yang terkait dengan penelitian yang akan penulis lakukan.

- Penelitian pertama yang dilakukan oleh (Shefia, dkk) tahun 2022, dengan judul "Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means" penelitian ini dilakukan di sekolah dasar namorambe di mana hasil dari penelitian ini didapatkan hasil pengelompokan cluster siswa dengan total 64 siswa yang didalamnya terdapat cluster dengan "penilaian prestasi siswa baik di pertimbangkan" sebanyak 29 siswa, "penilaian prestasi siswa sangat baik" sebanyak 5 siswa, "penilaian prestasi siswa tidak baik" sebanyak 30 siswa

dengan perhitungan titik centroid pada literasi ke 2 dengan angka bcv 0,108 yang menunjukkan bahwa penerapan algoritma k-means terhadap sistem yang dirancang dan dibangun mampu mempercepat dalam pengelompokan data siswa berprestasi dengan sistem yang dibuat berbasis desktop[7].

- Penelitian kedua dilakukan oleh (Sutanto, dkk) tahun 2023, dengan judul "Analisis Klasterisasi Data Sekolah SMP Di Kabupaten Kotawaringin Timur Menggunakan Metode K-Means" penelitian dilakukan di SMP kabupaten kotawaringin timur dengan hasil 112 data record SMP yang dibuat menjadi 3 cluster pengelompokan diantaranya cluster 0 sebanyak 87 sekolah, cluster 1 sebanyak 4 sekolah, dan cluster 2 sebanyak 21 sekolah dengan titik centroid 4143.863, penelitian ini juga melakukan pengujian sebanyak 8 kali dengan avg. Within centroid distance yang di analisis dengan metode *elbow method* dan menghasilkan cluster yang paling efektif sebanyak 3 pengelompokan[8].
- Penelitian ketiga dilakukan oleh (Syahra) tahun 2023, dengan judul "Penerapan Data Mining Dalam Pengelompokan Data Nilai Siswa Untuk Penentuan Jurusan Siswa Pada SMA Tamora Menggunakan Algoritma K-Means Clustering" dengan hasil analisis

- yang didapat dari jumlah data record 128 siswa dimana pengelompokan dari siswa terdiri dari cluster 1 sebanyak 33 siswa, cluster 2 sebanyak 16 siswa, cluster 3 sebanyak 33 siswa, dan cluster 4 sebanyak 46 siswa dengan perhitungan sampai iterasi ke 4 ratio 0,005. Penerapan data mining dengan algoritma k-means pada sma tamora ini mampu untuk menganalisis permasalahan yang ada yang berkenaan dengan pengelompokan data nilai siswa untuk penentuan jurusan serta aplikasi yang dirancang dapat mengimplementasikan dan dapat dijadikan sebagai solusi pemecahan masalah dalam hal pengelompokan data nilai siswa untuk penentuan jurusan siswa pada sma tamora[9].
- Penelitian keempat dilakukan oleh (Sulistiyawati & Supriyanto) tahun 2020, dengan judul “Implementasi Algoritma K-Means Clustering Dalam Penentuan Siswa Kelas Unggulan” penelitian ini dilakukan di smpn 2 jati agung kabupaten lampung selatan dengan jumlah data record 192 siswa dengan peneglompokan cluster kelas unggulan sebanyak 96 siswa dan cluster kelas tidak unggulan sebanyak 96 siswa dengan perhitungan sampai pada iterasi ke 6 dimana nilai sudah tidak bergeser lagi. Adapun pengujian perangkat lunak dilakukan melalui pengujian blackox untuk pengujian pakar dan pengujian technology acceptance model (tam) untuk pengujian user. Blackbox (blackbox testing) sebagai uji coba fungsionalitas dari system yang sudah dibuat dan mampu untuk memberikan hasil klasifikasi pengelompokan data yang efektif dan proses setiap iterasi perputaran jarak centroid, penentuan titik cluster dibentuk, sebagai acuan object data siswa yang lebih efektif dan efisien dalam menghemat waktu pada saat klasterisasi kelas unggulan sehingga mempermudah pihak sekolah[10].
 - Penelitian terakhir yang dilakukan oleh (Sholihin & Rudiman) tahun 2021 dengan judul “Pemetaan Dan Klasterisasi Sekolah Muhammadiyah Di Kabupaten Ppu Berdasarkan Fasilitas, Pendidik Dan Tenaga Pendidik Menggunakan Metode K-Means Clustering” penelitian ini berhasil membuat pemetaan dan klasterisasi sekolah muhammadiyah di kabupaten Penajam Pasir Utara berdasarkan fasilitas, pendidik dan tenaga pendidik menggunakan metode k-means clustering dengan jumlah data record 14 data sekolah dan cluster yang efisien sebanyak 3 cluster yaitu cluster 2 terdapat 1 sekolah, cluster 1 terdapat 8 sekolah, dan cluster 0 terdapat 5 sekolah sehingga menghasilkan nilai davies bouldin index = 0.446[11].

2.2. Landasan Teori

1. Prestasi

Prestasi adalah apa yang telah dapat diciptakan, hasil pekerjaan, hasil yang menyenangkan hati yang diperoleh dengan jalan keuletan kerja. Untuk itu, dapat dipahami bahwa prestasi adalah hasil dari suatu kegiatan yang telah dikerjakan, diciptakan, yang menyenangkan hati, yang diperoleh dengan jalan keuletan kerja, baik secara individual maupun secara kelompok dalam bidang kegiatan tertentu[12].

2. Data Mining

Data mining adalah salah satu bentuk implementasi yang diterapkan untuk mencari sebuah model dan pola yang mampu melakukan prediksi pada suatu data berdasarkan data sebelumnya di periode waktu tertentu. Data mining adalah bentuk penggalian data yang digunakan untuk menggali pengetahuan dari jumlah data yang besar[6]. Data mining merupakan suatu proses yang dilakukan dengan cara mengekstraksi dasar data menjadi informasi yang mempunyai nilai penting, dengan data pola yang terbentuk dari data, data dasar disebut dengan Penemuan Pengetahuan dalam Database yang dicari menggunakan pembelajaran berbasis berulang untuk menemukan pola yang tetap dan mudah digunakan oleh pengguna dalam pengambilan keputusan[13].

3. Algoritma K-Means

K-means merupakan salah satu algoritma yang bersifat unsupervised learning. K-means memiliki fungsi untuk mengelompokkan data kedalam data cluster[14]. Algoritma ini dapat menerima data tanpa ada label kategori. K-means clustering algoritma juga merupakan metode non-hierarchy. Metode clustering algoritma adalah mengelompokkan beberapa data ke dalam kelompok yang menjelaskan data dalam satu kelompok memiliki karakteristik yang sama dan memiliki karakteristik yang berbeda dengan data yang ada di kelompok lain[9]. Cluster sampling adalah teknik pengambilan sampel di mana unit-unit populasi dipilih secara acak dari kelompok yang sudah ada yang disebut cluster. Clustering atau klasterisasi adalah salah satu teknik penyelesaian masalah dengan teknik unsupervised learning.

4. Elbow Method

Metode *elbow* merupakan salah satu metode yang dapat digunakan untuk menentukan jumlah cluster terbaik, yaitu dengan cara melihat presentase setiap cluster yang akan membentuk siku pada suatu titik tertentu. Metode *elbow* biasa disajikan dalam bentuk grafik untuk mengetahui lebih jelas siku yang terbentuk. Tujuan dari metode *elbow* adalah untuk memilih nilai k yang kecil dan masih memiliki nilai *withinss* yang rendah. Nilai k pada kombinasi siku dengan *k-means* adalah grafik hubungan cluster dengan

penurunan error. Jumlah cluster k yang dihasilkan dari pengujian dengan k -means dievaluasi dengan teknik sse. Sse (*sum of square error*) merupakan rumus yang digunakan untuk mengukur perbedaan antara data yang telah dilakukan sebelumnya[15].

3. Metode Penelitian

Adapun metode pada penelitian yang dilakukan untuk memperoleh informasi yang dibutuhkan diantaranya:

3.1 Metode Pengumpulan Data

Untuk mendapatkan data sebagai pemenuhan penelitian dan data yang diperoleh benar-benar akurat serta dapat digunakan dalam penelitian nantinya, maka dari itu peneliti menggunakan beberapa metode pengumpulan data. Berikut ini beberapa metode yang penulis gunakan dalam penelitian:

1. Observasi Lapangan

Pengumpulan data yang dilakukan dengan pengamatan secara langsung di lokasi penelitian yakni pada sekolah mis nw pancor, guna untuk mendapatkan informasi terkait data siswa untuk mengetahui bagaimana permasalahan yang ada disana.

2. Studi Pustaka

Pengumpulan informasi yang dilakukan dengan cara mencari informasi dan referensi yang berhubungan dengan penelitian, referensi diperoleh dari buku, jurnal / artikel dan internet.

3. Wawancara

Tanya jawab langsung dengan pihak terkait yakni salah satu guru disekolah untuk mendapatkan informasi yang dibutuhkan peneliti agar lebih memahami permasalahan yang terjadi.

3.2 Tahapan Penelitian

Tahapan penelitian merupakan langkah-langkah yang dilakukan oleh peneliti, berikut adalah tahapan yang dilakukan oleh peneliti dapat dilihat pada gambar 1.



Gambar 1. Tahapan Penelitian

Pada tahapan ini dilakukan dengan pengumpulan data dari study literatur yang selanjutnya disesuaikan dengan permasalahan yang terjadi, dan selanjutnya dari kumpulan data yang ada dilakukan pengolahan baik dari cleaning, filter dan pemilihan data selanjutnya dilakukan pengujian di mana dari hasil pengujian jika tidak sesuai dengan data maka akan kembali untuk dilakukan pada tahap Analisa masalah dan pengumpulan serta pengolahan data yang

sesuai, Jika pada proses pengujian hasil yang didapatkan sudah sesuai dengan hasil yang didapatkan maka proses akan berakhir.

3.3 Lokasi Penelitian

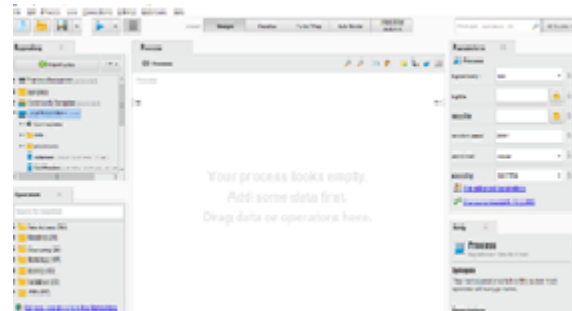
Adapun lokasi/tempat yang menjadi objek penelitian yang penulis ambil yaitu di sekolah MIS NW 03 Pancor, Jl. Dewi Sartika no.21, pancor, kec. Selong, kabupaten lombok timur, nusa tenggara bar. 83611.

4. Hasil dan Pembahasan

4.1 Hasil Penelitian

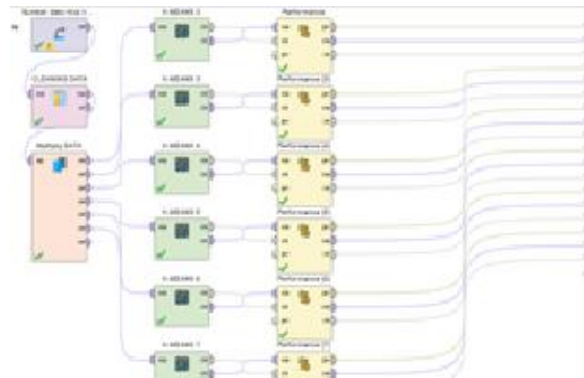
Pada penelitian ini, peneliti menggunakan software rapid miner sebagai alat untuk melakukan pengolahan data siswa. Berikut ini adalah tahapan pengolahan data siswa menggunakan algoritma k-means untuk mengelompokkan data siswa berdasarkan nilai, sikap, dan keterampilan yang dioptimasi dengan penggunaan metode *elbow* untuk penentuan cluster yang paling baik untuk digunakan.

Sumber data yang digunakan adalah excel yang dimana terdapat 18 atribut yang terdiri dari nama, nilai (gabungan dari nilai mata pelajaran), sikap, dan keterampilan dengan jumlah data yang dijadikan pengolahan dari masing-masing kelas yakni kelas 4 sebanyak 23 siswa, kelas 5 sebanyak 27 siswa dan kelas 6 sebanyak 25 siswa.



Gambar 3. Tampilan Dashboard

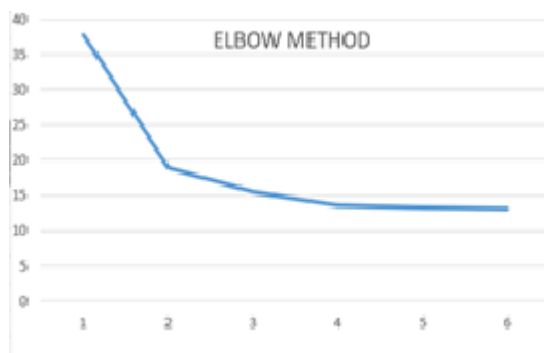
Pada tampilan dashboard ini menampilkan halaman kerja yang akan digunakan termasuk dalam proses import data cleaning, uji validitas, serta tampilan dari hasil dari data yang akan diolah dengan menggunakan algoritme metode yang sesuai kebutuhan.



Gambar 4. Pengolahan Data

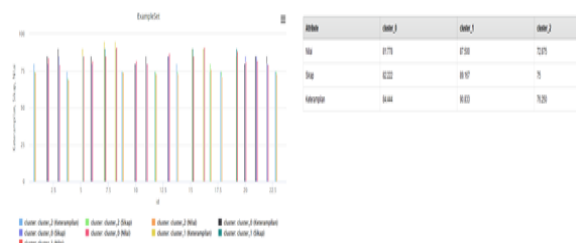
Pada tahapan ini sumber data excel yang yang menjadi input diekstraksi dan dilakukan cleaning data dimana data yang akan digunakan adalah atribut yang terkait dengan prestasi siswa yang terdapat didalamnya : nilai, sikap, dan keterampilan yang kemudian sumber data tersebut dilakukan proses multiply yang bertujuan untuk menduplikat sumber data untuk diuji menggunakan metode klasifikasi k-means dengan pencairian kluster terbaik, untuk

mendapatkan hasil yang diinginkan, selanjutnya setelah proses kluster dilakukan maka proses selanjutnya adalah uji coba dengan pengujian cluster distance performance dengan witin avg centroid distance untuk melihat hasil performance dari data yang telah kita olah dan kemudian di optimasi dengan penggunaan metode elbow.



Gambar 5. Skema Metode *Elbow*

Pada proses *Elbow* ini dilakukan dengan cara menghitung titik Avg Distance dari jumlah kluster yang digunakan sebagai sampling untuk menentukan jumlah cluster yang paling baik untuk digunakan. Sesuai dengan metode elbow / siku maka titik kluster yang terbaik adalah kluster 3 dengan angka avg distance 18.821 atau titik cluster 3.



Gambar 6. Clustering Diagram Chart Bar

Pada tahapan ini data siswa yang telah diolah akan ditampilkan berdasarkan chart bar dimana cluster dengan optimasi elbow yang menunjukkan

performance terbaik pada cluster 3 menampilkan hasil dari pembagian tiap kelompok dari 3 atribut penilaian yaitu: nilai, sikap, dan keterampilan dengan hasil centroid table cluster 1 = 81.77, 82.22, 84.44, cluster 2 = 87.50, 89.16, 90.83, dan cluster 3 = 72.87, 75.00, 76.25.



Gambar 7. Clustering Result Scatter Matrix

Visualisation clustering pada scatter matrix ini memberikan gambaran akan atribut yang menjadi penilaian dari pemoresan data yang diolah kemudian dipecah menjadi 3 cluster dan titik dari keberadaan masing-masing atribut berdasarkan dari pengelompokan dari nilai, sikap, dan keterampilan siswa sehingga mampu untuk mempermudah pihak sekolah dalam melihat data yang sudah dikelompokkan untuk menentukan keputusan kedepannya.

Hasil dari pembuatan cluster pada kelas IV MIS NW Pancor dengan titik optimasi cluster 3 menunjukkan pengelompokan dimana cluster 1 sebagai “baik” terdiri dari atribut 2,3,6,10,11,13,20,21,22 dengan jumlah 9 siswa, selanjutnya cluster 2 sebagai “sangat baik” terdiri dari atribut 5,7,8,15,16,19 dengan jumlah 6

siswa, dan yang terakhir cluster 3 sebagai “cukup baik” terdiri dari atribut 1,4,9,12,14,17,18,23 dengan jumlah 8 siswa dengan hasil performance vector davies bouldin 0.711 dan avg within centroid distance 18.821.

Tabel 1. Cluster Atribut

Set Data	Cluster 1	Cluster 2	Cluster 3
Kelas IV	2.0	5.0	1.0
	3.0	7.0	4.0
	6.0	8.0	9.0
	10.0	15.0	12.0
	11.0	16.0	14.0
	13.0	19.0	17.0
	20.0		18.0
	21.0		23.0
	22.0		
	Total Qty	9	6

5. Kesimpulan

Berdasarkan hasil pengolahan data yang dilakukan di MIS NW 03 Pancor dengan data sampel kelas IV, V, dan IV dengan data set sebanyak 23 siswa pada kelas IV yang dikategorikan berprestasi dengan “sangat baik”, “baik”, dan “cukup baik” dapat dikelompokkan dengan penerapan k-means clustering dalam pengelompokan prestasi siswa dengan optimasi metode elbow dengan sangat baik. Dari pengujian optimasi elbow yang dilakukan dan performance yang dilakukan sebanyak 6 kali pengujian yang dimulai dari 2 cluster sampai dengan 7 cluster yang pada akhirnya ditentukan cluster terbaik yaitu dengan menggunakan 3 cluster yang ditentukan dari metode elbow dengan angka avg distance 18.821, dengan pengelompokan cluster 1

sebagai “baik” terdiri dari atribut 2,3,6,10,11,13,20,21,22 dengan jumlah 9 siswa, selanjutnya cluster 2 sebagai “sangat baik” terdiri dari atribut 5,7,8,15,16,19 dengan jumlah 6 siswa, dan yang terakhir cluster 2 sebagai “cukup baik” terdiri dari atribut 1,4,9,12,14,17,18,23 dengan jumlah 8 siswa dengan hasil performance vector davies bouldin 0.711 dan avg within centroid distance 18.821 sebagai hasil akhir dari pengolahan data tersebut sehingga dengan adanya pengelompokan prestasi siswa pihak sekolah akan dipermudah dalam pengambilan keputusan serta kebijakan-kebijakan yang akan digunakan

6. Daftar Pustaka

- [1] A. Nurjana, A. Perdana Windarto, and H. Qurniawan, “Implementasi Data Mining Dalam Memprediksi Prestasi Siswa Dengan Algoritma C4.5,” *SmartEDU*, vol. 1, pp. 171–180, 2022, [Online]. Available: <https://ejournal.abivasi.id/index.php/SmartEDU>
- [2] A. Sulistiyawati and E. Supriyanto, “Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan,” *Jurnal Tekno Kompak*, vol. 15, no. 2, pp. 25-36, 2021
- [3] A. M. Nur, N. Nurhidayati, and I. Fathurrahman, “Penerapan Metode Naive Bayes Untuk Penentuan Penerima Beasiswa Program Indonesia Pintar (PIP).,” *Infotek: Jurnal Informatika dan Teknologi*, vol. 7, no. 1, pp. 93–102, Jan. 2024, doi: 10.29408/jit.v7i1.23995
- [4] R. Yuliani, “Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di

- Dusun Bagik Endep Sukamulia Timur,” *Jurnal Informatika dan Teknologi*, vol. 4, no. 1, pp. 39-50, 2021, doi: 10.29408/jit.v4i1.2986
- [5] A. E. Wicaksono, “Implementasi Data Mining Dalam Pengelompokan Data Peserta Didik Di Sekolah Untuk Memprediksi Calon Penerima Beasiswa Dengan Menggunakan Algoritma K-Means (Studi Kasus Sman 16 Bekasi),” *Jurnal Teknologi Rekayasa*, vol.21, no.3, pp. 206-216, 2016.
- [6] V. A. Ekasetya and A. Jananto, “Klusterisasi Optimal Dengan Elbow Method Untuk Pengelompokan Data Kecelakaan Lalu Lintas Di Kota Semarang,” *Dinamika Informatika*, vol. 12, no. 1, pp. 20–28, 2020
- [7] S. N. B. Sembiring, H. Winata, and S. Kusnasari, , “Pengelompokan Prestasi Siswa Menggunakan Algoritma K-Means,” *Jurnal Sistem Informasi TGD*, vol.1, no.1, pp. 31-40, 2022.
- [8] H. S. Sutanto and I. Susilawaty, “Analisis Klusterisasi Data Sekolah Smp Di Kabupaten Kotawaringin Timur Menggunakan Metode K-Means,” *JURSISTEKNI (Jurnal Sistem Informasi dan Teknologi Informasi)*, vol.5, no.3, pp. 298-309, 2023.
- [9] Y. Syahra, “Penerapan Data Mining Dalam Pengelompokan Data Nilai Siswa Untuk Penentuan Jurusan Siswa Pada SMA Tamora Menggunakan Algoritma K-Means Clustering,” *Sains dan Komputer*, vol. 17, no. 2, pp. 228–233, 2018.
- [10] A. Sulistiyawati and E. Supriyanto, “Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan,” vol. 15, no. 2, pp. 25-36, 2021.
- [11] M. R. Sholihin and Rudiman, “Pemetaan dan Klusterisasi Sekolah Muhammadiyah di Kabupaten PPU Berdasarkan Fasilitas, Pendidik dan Tenaga Pendidik Menggunakan Metode K-Means Clustering,” *Jurnal Keilmuan dan Aplikasi Teknik Informatika*, vol. 14, no. 2, pp. 45-51, 2021, doi: 10.35891/explorit.
- [12] N. Butarbutar, A. P. Windarto, D. Hartama and Solikhun, “Komparasi Kinerja Algoritma Fuzzy C-Means Dan K-Means Dalam Pengelompokan Data Siswa Berdasarkan Prestasi Nilai akademik Siswa (Studi Kasus : SMP Negeri 2 Pematangsiantar),” *JURASIK (Jurnal Riset Sistem Informasi & Teknik Informatika)* vol. 1, no. 1, pp. 46-55, 2016.
- [13] C. Devi, O. Soleman, N. Pramaita, and M. Sudarma, “Classification of Loyalty Customer Using K-Means Clustering, Studi Case: PT. Sucofindo (Persero) Denpasar Branch,” *International Journal of Engineering and Emerging Technology*, vol. 5, no. 2, pp. 160-167, 2020.
- [14] S. Suhartini and R. Yuliani, “Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur,” *Infotek: Jurnal Informatika dan Teknologi*, vol. 4, no. 1, pp. 39–50, 2021, doi: 10.29408/jit.v4i1.2986
- [15] A. R. Lashiyanti, I. R. Munthe, and F. A. Nasution, “Optimisasi Klusterisasi Nilai Ujian Nasional dengan Pendekatan Algoritma K-Means, Elbow, dan Silhouette,” *Jurnal Ilmu Komputer dan Sistem Informasi (JIKOMSI)*, vol. 6, no. 1, pp. 14–20, 2023