

Klasterisasi Penempatan Siswa yang Optimal untuk Meningkatkan Nilai Rata-Rata Kelas Menggunakan *K-Means*

Yusma Elda^{1✉}, Sarjon Defit², Yuhandri Yunus³, Raemon Syaljumairi⁴

¹SMK Negeri 2 Padang Panjang

^{2,3}Universitas Putra Indonesia YPTK Padang

⁴Politeknik Negeri Padang

yusma.elda@gmail.com

Abstract

The implementation of learning by teachers can measure the quality of schools and students. Schools with diverse student backgrounds need to take strategic steps in managing learning to get optimal learning outcomes. Good learning designs and techniques can motivate students' interest in learning. The teacher's role is very important in managing learning to create an effective teaching and learning process. Data Mining or also known as Knowledge Discovery in Database (KDD) is the process of extracting knowledge from large data to find new patterns to get new knowledge and information. Data Mining technology is used to explore existing knowledge in the database. One of the methods used in data mining is clustering with the K-Means algorithm. This study aims to conduct student clustering to obtain a balanced class composition in order to improve the quality and student learning outcomes as seen in the increasing in the class average score. The data processed in this study came from the main school data as many as 90 students of the XI class of Computer Network Engineering Skills Competency at SMKN Negeri 2 Padang Panjang in the 2020/2021 school year. The variables used in data processing are student scores, parents' income and the distance from where students live to school. The student clustering calculation using K-Means succeeded in grouping 90 students into 3 clusters where cluster 1 totaled 47 students, cluster 2 totaled 10 students and cluster 3 totaled 33 students. Each member of the cluster will be divided evenly into 3 groups studying to get a balanced class composition. This research can be used as a basis for decision making by schools in clustering student placements to improve learning outcomes. By the increasing in the grade point average, the school average score will also increased.

Keywords: KDD, Clustering, K-Means, Principal School Data, Class Average Score.

Abstrak

Pelaksanaan pembelajaran oleh guru dapat mengukur kualitas sekolah dan peserta didik. Sekolah dengan latar belakang siswa yang beragam perlu melakukan langkah-langkah strategis dalam mengelola pembelajaran untuk mendapatkan hasil belajar yang optimal. Desain dan teknik pembelajaran yang baik dapat memotivasi minat belajar siswa. Peranan guru sangat penting dalam mengelola pembelajaran untuk menciptakan proses belajar mengajar yang efektif. Data Mining atau disebut juga *Knowledge Discovery in Database* (KDD) adalah proses pengekstrasian pengetahuan dari data yang berukuran besar untuk menemukan pola-pola baru sehingga mendapatkan sebuah pengetahuan dan informasi baru. Teknologi Data Mining digunakan untuk menggali pengetahuan yang ada didalam database. Salah satu cara yang digunakan dalam Data Mining adalah clustering dengan algoritma *K-Means*. Penelitian ini bertujuan untuk melakukan pengelompokan siswa untuk mendapatkan komposisi kelas yang seimbang guna meningkatkan mutu dan hasil belajar siswa yang dilihat dari meningkatnya nilai rata-rata kelas. Data yang diolah dalam penelitian ini bersumber dari data pokok sekolah sebanyak 90 siswa Kompetensi Keahlian Teknik Komputer Jaringan kelas XI SMKN Negeri 2 Padang Panjang tahun pelajaran 2020/2021. Variabel yang digunakan dalam pengolahan data adalah nilai siswa, penghasilan orang tua dan jarak tempat tinggal siswa ke sekolah. Perhitungan klasterisasi siswa dengan K-Means berhasil mengelompokkan 90 siswa menjadi 3 klaster dimana cluster 1 berjumlah 47 siswa, cluster 2 berjumlah 10 siswa dan cluster 3 berjumlah 33 siswa. Masing-masing anggota klaster akan dibagi merata kedalam 3 kelompok belajar untuk mendapatkan komposisi kelas yang seimbang. Riset ini bisa dijadikan dasar pengambilan keputusan oleh sekolah dalam melakukan klasterisasi penempatan siswa untuk meningkatkan hasil belajar. Dengan meningkatnya nilai rata-rata kelas maka nilai rata-rata sekolah juga akan meningkat.

Kata kunci: KDD, Pengelompokan, *K-Means*, Data Pokok Sekolah, Nilai Rata Kelas.

© 2021 JIDT

1. Pendahuluan

Sekolah merupakan institusi pendidikan yang paling strategis dalam mencapai tujuan pendidikan nasional seperti yang tercantum dalam Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 [1]. Pelaksanaan pembelajaran oleh guru dapat mengukur kualitas sekolah dan peserta didik. Sekolah dengan latar belakang siswa yang beragam perlu melakukan langkah-langkah strategis dalam mengelola pembelajaran untuk mendapatkan hasil belajar yang optimal.

Desain dan teknik pembelajaran yang baik dapat memotivasi minat belajar siswa. Peranan guru sangat penting dalam mengelola pembelajaran untuk menciptakan proses belajar mengajar yang efektif. Pembagian kelompok belajar siswa secara tepat dapat mendukung proses belajar yang kondusif dimana antara siswa dengan kemampuan yang beragam dapat saling berkolaborasi dalam proses pembelajaran.

Data Mining atau disebut juga Knowledge Discovery in Database (KDD) merupakan cara pengekstrasian pengetahuan dari data yang berukuran besar untuk menemukan pola-pola baru sehingga mendapatkan sebuah pengetahuan dan informasi yang baru [2] [3]. Teknologi Data Mining digunakan untuk menggali pengetahuan yang ada didalam database [4] [5].

Clustering dengan algoritma K- Means [6] adalah salah satu cara yang digunakan dalam Data Mining [7]. Cluster merupakan sekumpulan objek data yang sama satu dengan yang lain dalam satu kelompok dan berbeda terhadap objek yang berada dikelompok lain [8]. Clustering termasuk dalam kategori unsupervised learning yang berfungsi untuk mempartisi data tanpa label ke dalam satu kelompok [9].

Clustering melakukan pengelompokan obyek kedalam kelas-kelas obyek yang sama [10] [11]. Perhitungan dengan K-Means clustering adalah proses yang dilakukan secara berulang-ulang [12]. Clustering dengan algoritma K-Means melakukan pengelompokan terhadap data nilai rapor siswa dengan data sampel sebanyak 23 siswa yang menghasilkan 3 cluster tingkat kompetensi siswa untuk pemilihan peserta Olimpiade Sains Nasional [13].

Pembagian kelompok siswa kedalam beberapa kelas dalam satu mata kuliah menggunakan algoritma K-Means berdasarkan pada nilai harian, penilaian tengah semester, penilaian akhir semester, dan akumulasi penilaian setiap semester (IPK) mahasiswa [14]. Pengelompokan data siswa menggunakan metode K-Means Clustering dilakukan berdasarkan nilai ujian semester mata pelajaran Ujian Nasiona [15]. Pengelompokan terhadap data untuk melihat dan menciptakan kelas objek-objek yang mempunyai kemiripan juga dapat dilakukan dengan K-Means Clustering [16].

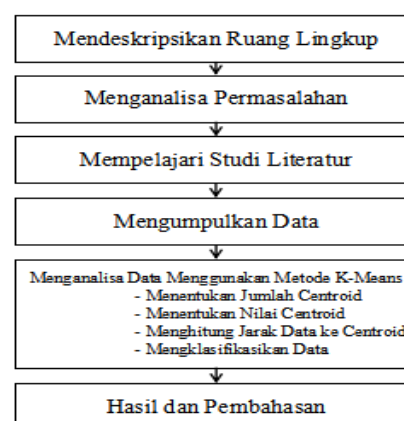
Metode K-Means Clustering and Multiple Linear Regressions (MLR) dengan mengambil data dari UCI (University California Irvine), dimana set data didasarkan dari nilai matematika siswa, nilai ulangan, kuis, dan tugas siswa sebagai faktor utama dalam memprediksi kinerja akademik siswa [17]. Algoritma Clustering K-Means dan Algoritma Greedy, dengan data 37 siswa menghasilkan 5 cluster kualitas siswa, sehingga dari cluster tersebut terbentuk kelompok kerja siswa untuk pemerataan kualitas siswa [18].

Mengklasifikasikan mahasiswa ke dalam kelompok sesuai dengan preferensi gaya belajar dengan metode The Felder dan Silverman's Index Learning Style (ILS) digunakan untuk mengevaluasi preferensi gaya belajar siswa.

Penggalian data dengan metode klasterisasi K-Means dapat diimplementasikan guna melakukan klasterisasi siswa secara sistematis kepada siswa SMK Negeri 2 Padang Panjang yang selama ini masih dilakukan secara manual oleh guru Bimbingan dan Konseling (BK). Penerapan cara kerja Algoritma K-Means dalam klasterisasi penempatan siswa bertujuan untuk mengambil keputusan yang tepat sehingga didapatkan komposisi kelas yang seimbang guna meningkatkan nilai rata-rata kelas.

2. Metodologi Penelitian

Metode penelitian adalah tahapan penelitian yang dilakukan untuk menentukan, mendapatkan, serta mengkaji sehingga dapat memperoleh data yang berpedoman pada tahapan dan waktu yang ditentukan dalam menganalisa data. Kerangka penelitian adalah garis besar dari segala tahapan kegiatan yang dilakukan dalam penelitian sebagai pedoman agar penelitian berjalan terarah. Kerangka penelitian disajikan oleh Gambar 1.



Gambar 1. Kerangka Kerja Penelitian

2.1. Mendeskripsikan Ruang Lingkup Masalah

Mendeskripsikan ruang lingkup masalah adalah untuk memberikan penjelasan ruang lingkup penelitian ini. Data siswa Kompetensi Keahlian Teknik Komputer Jaringan kelas XI Tahun Pelajaran 2020/2021 yang

diperoleh di SMK Negeri 2 Padang Panjang adalah data yang diolah pada penelitian ini.

2.2. Menganalisa Permasalahan

Menganalisa permasalahan adalah langkah selanjutnya agar penelitian yang akan dilakukan dapat memberikan solusi terhadap permasalahan. Permasalahan yang ada selama ini adalah proses penempatan siswa dilakukan secara manual oleh guru Bimbingan dan Konseling (BK). Pada analisa masalah ini digambarkan proses untuk menentukan bagaimana penempatan siswa yang optimal secara sistematis untuk menaikkan nilai rata-rata kelas..

2.3. Mempelajari Literatur

Mempelajari literatur bertujuan untuk memperkuat informasi yang dihasilkan dari suatu analisis atau hipotesa guna menghilangkan keragu-raguan terhadap pengertian suatu masalah. Studi literatur yang dilakukan dalam penelitian ini adalah dengan melakukan kajian jurnal untuk mencari informasi yang berhubungan dengan penelitian dan melakukan diskusi dan wawancara langsung dengan Guru Bimbingan dan Konseling (BK) SMK Negeri 2 Padang Panjang.

2.4. Mengumpulkan Data

Untuk mendapatkan data yang dibutuhkan pada penelitian ini, maka dilakukan pengambilan data melalui observasi langsung peneliti di SMK Negeri 2 Padang Panjang sebagai sumber data dalam penelitian ini.

2.5. Menggunakan Metode *K-Means* untuk Menganalisa Data

Untuk mendapatkan klasterisasi siswa secara optimal dilakukan analisa dengan metode *K-Means*. Langkah-langkah algoritma *K-Means* adalah:

1. Menginputkan data.
2. Menentukan jumlah *cluster*
3. Menentukan titik pusat *cluster* (*centroid*). Ambil data secara acak sebagai pusat *cluster* (*centroid*) awal.
4. Menghitung jarak data ke pusat *cluster* dengan menggunakan rumus persamaan *Euclidean Distance*:

$$D_{(i,j)} = \sqrt{(X_{1i} - X_{1j})^2 + (X_{2i} - X_{2j})^2 + \dots + (X_{ki} - X_{kj})^2} \quad (1)$$

$D_{(i,j)}$ pada rumus diatas adalah nilai jarak data ke i ke pusat *cluster* j . Untuk X_{ki} merupakan data ke- i pada atribut data ke k , dan X_{kj} merupakan jarak titik pusat ke j pada atribut ke k .

5. Mengklasifikasikan Data. Data diklasifikasikan berdasarkan kedekatannya dengan pusat *cluster* (*centroid*). Data dikelompokkan ke dalam *cluster* dengan nilai yang terdekat. Lakukan proses perhitungan pusat *cluster* baru berdasarkan rata-rata anggota yang ada pada *cluster* tersebut. Proses mengelompokkan data (iterasi) dilakukan

sampai hasil iterasi bernilai sama dengan iterasi sebelumnya.

2.6. Hasil dan Pembahasan

Mekanisme pengujian data penelitian dilakukan dengan membandingkan keluaran *output* dari perhitungan manual *K-Means Clustering* dengan menggunakan fungsi SQRT pada *Ms.Excel* dan dengan hasil yang didapatkan dari pengujian dengan menggunakan *software* aplikasi *Rapid Miner Studio Versi 9.2*. Dari hasil pengelompokan ini didapat 3 *cluster* siswa dengan bobot sikap dan kemampuan belajar yang tinggi, sedang dan rendah. Keseluruhan data dengan jumlah 90 siswa yang ada pada 3 *cluster* tersebut akan dibagi merata kedalam 3 kelas belajar untuk mendapatkan komposisi kelas dengan kemampuan siswa yang seimbang.

3. Hasil dan Pembahasan

Sikap dan kemampuan belajar siswa diukur berdasarkan variabel nilai, penghasilan orang tua serta radius tempat tinggal siswa ke sekolah yang menghasilkan 3 *cluster*, yaitu *cluster* 1 dengan bobot sikap dan hasil belajar yang tinggi, *cluster* 2 untuk bobot sikap dan hasil belajar yang sedang dan *cluster* 3 untuk bobot sikap dan hasil belajar yang rendah. Langkah untuk menentukan sikap belajar siswa dengan algoritma *k-means clustering* adalah:

3.1. Menginputkan Data

Melakukan penginputan data kedalam sistem untuk menggali pengetahuan dengan algoritma *K-Means clustering* untuk menghasilkan *cluster* data. Variabel data yang diolah X untuk variabel nilai, Y untuk variabel penghasilan orang tua dan Z untuk variabel jarak dari rumah ke sekolah.

Tabel 1. Data yang Akan di *Cluster*

Data Ke-i	Client	X	Y	Z
1	Adil Saputra	2	1	3
2	Aditya Erlangga	2	1	1
3	Agung Koswara	2	1	1
4	Agung Yandri Fakhri	2	2	1
5	Aidil Fitra Al Munawar	2	2	2
6	Alfi Anshari	2	2	1
7	Alka Vikra	3	1	2
8	Andika Putra Prayuda	3	1	1
9	Andre Pranoval	2	2	1
10	Annisa Aulia Fitri	3	1	1

Pada Tabel 1 ditampilkan sebagian data dari keseluruhan data yang diolah yaitu 90 data siswa Kompetensi Keahlian Teknik Komputer Jaringan kelas XI Tahun Pelajaran 2020/2021 SMK Negeri 2 Padang Panjang.

3.2. Menentukan Jumlah *Cluster*

Jumlah *cluster* ditentukan sebanyak 3 *cluster* yaitu *cluster* 1 dengan bobot sikap dan hasil belajar yang tinggi, *cluster* 2 untuk bobot yang sedang dan *cluster* 3

untuk bobot yang rendah dari variabel data siswa yang diteliti.

3.3. Menentukan Titik Pusat *Cluster* (*centroid*)

Langkah selanjutnya adalah menentukan pusat *cluster* (*centroid*) yang diambil secara sembarang/acak dari data yang ada.

Tabel 2. Nilai *Centroid* Acak

Data ke-i	Centroid	X	Y	Z
23	C1	3	3	1
34	C2	3	2	3
10	C3	3	1	1

3.4. Menghitung Jarak Data ke Pusat *Cluster*

Menghitung jarak masing-masing data ke pusat *cluster*, dilakukan dengan menggunakan rumus *Euclidean Distance* (D) pada rumus (1), dengan perhitungan:

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

$$D_{1,2} = \sqrt{(2-3)^2 + (1-2)^2 + (3-3)^2} = 1,4142135624$$

$$D_{13} = \sqrt{(2-3)^2 + (1-1)^2 + (3-1)^2} = 2,2360679775$$

3.5. Pengelompokan *Data* ke Cluster Terdekat

Dari hasil perhitungan jarak data ke pusat cluster pada iterasi I, maka dapat dikelompokkan masing masing data berdasarkan jarak terdekat seperti terlihat di Tabel 3:

Tabel 3. Data Pengelompokkan Hasil Iterasi I

Data ke-i	Jarak ke Pusat <i>Cluster</i>			Terde kat	C1	C2	C3
	1	2	3				
1	3.000	1.414	2.236	1.414		1	
2	2.236	2.449	1.000	1.000			1
3	2.236	2.449	1.000	1.000			1
4	1.414	2.236	1.414	1.414	1		
5	1.732	1.414	1.732	1.414		1	
6	1.414	2.236	1.414	1.414	1		
7	2.236	1.414	1.000	1.000			1
8	2.000	2.236	0.000	0.000			1
9	1.414	2.236	1.414	1.414	1		
10	2.000	2.236	0.000	0.000			1
		Jumlah			37	20	33

Pada Tabel 3 ditampilkan sebagian data hasil perhitungan iterasi pertama. Didapatkan kelompok data untuk C1, C2 dan C3 yaitu :

- C1 (*cluster 1*) berjumlah 37 siswa yang anggotanya data nomor urut ke: 4, 6, 9, 12, 14, 15, 17, 18, 20, 21, 23, 26, 27, 28, 33, 35, 36, 38, 39, 41, 45, 51, 54, 56, 58, 59, 63, 64, 66, 68, 73, 76, 77, 80, 81, 83, 84
- C2 (*cluster 2*) berjumlah 20 siswa yang anggotanya data nomor urut ke : 1, 5, 19, 31, 32, 34, 43, 46, 47, 48, 49, 50, 55, 62, 71, 72, 82, 85, 88, 90
- C3 (*cluster 3*) berjumlah 33 siswa yang anggotanya data nomor urut ke : 2, 3, 7, 8, 10, 11, 13, 16, 22, 24, 25, 29, 30, 37, 40, 42, 44, 52, 53, 57, 60, 61, 65, 67, 69, 70, 74, 75, 78, 79, 86, 87, 89

3.6. Menentukan Centroid Baru

Lakukan perhitungan untuk mendapatkan nilai pusat *cluster* (*centroid*) baru sebelum melakukan iterasi kedua. Cara menghitungnya dengan merata-ratakan nilai dari data-data yang berada pada *cluster* yang sama. Perhitungan untuk menentukan pusat *cluster* baru dari hasil iterasi pertama adalah:

$$C1(X) = (2+2+2+1+2+2+4+3+3+2+3+3+2+4+5+4+2+2+3+1+2+2+2+2+3+3+2+3+3+2+3+3+1+3+3+3+4)/3 = 2.5946$$

$$C1(Y) = (2+2+2+3+3+3+2+2+3+2+3+2+3+3+2+2+2+3+2+2+2+2+2+3+2+2+2+3+2+2+2+2)/37 = 2.2973$$

$$\text{Cl}(\mathbf{Z}) = (1+1+1+1+2+2+1+1+2+1+1+1+2+1+1+1 \\ +1+1+1+1+1+1+1+1+1+1+1+1+1+1+1+1 \\ +1+1+1+1+1)/37 = 1,1081$$

$$C_2(X) = (2+2+2+2+4+3+3+2+2+2+2+3+3+4+2+2+1+2+3+3)/20 = 2,45$$

$$C2(Y) = (1+2+1+1+2+2+2+2+2+2+2+2+2+1+1+2+1+2+1+2)/20 = 1.65$$

$$C2(Z) = (3+2+4+4+2+3+4+2+2+2+2+2+2+3+4+3+4+2+5+2)/20 = 2,85$$

$$C_3(X) = (2+2+3+3+3+3+3+3+3+3+2+3+2+3+2+2+2+3+3+3+4+3+3+2+2+2+3+2+2+3+3+2+1)/33 = 2.5758$$

$$C_3(Y) = (1+1)/33 = 1$$

$$C_3(Z) = (1+1+2+1+1+1+1+1+1+1+2+2+1+1+1+1+2+1+1+1+2+1+1+2+2+1+1+1+1+1+2+1)/33 = 1.2425$$

Berdasarkan hasil pencarian pusat *cluster* baru, maka didapatkan hasil seperti pada Tabel 4.

Tabel 4. Nilai *Pusat Cluster* untuk Iterasi II

Pusat Cluster	X	Y	Z
C1	2,5946	2,2973	1,1081
C2	2,45	1,65	2,85
C3	2,5758	1	1.2425

Lakukan langkah 3.4, 3.5 dan 3.6 secara terus menerus sampai mendapatkan hasil iterasi dengan nilai yang sama, yaitu jarak data ke *centroid* tidak berubah.

Perhitungan jarak data ke pusat *cluster* pada iterasi ke 4, maka dapat dikelompokkan masing masing data berdasarkan jarak terdekat yang sebagian datanya ditampilkan seperti terlihat di Tabel 5

Tabel 5. Data Pengelompokan Hasil Iterasi III

Data ke-i	Jarak ke Pusat Cluster 1	Jarak ke Pusat Cluster 2	Jarak ke Pusat Cluster 3	Terdekat	C 1	C 2	C 3
1	2.302	0.618	1.849	0.6187		1	
2	1.442	2.152	0.624	0.6247			1
3	1.442	2.152	0.624	0.6247			1
4	0.737	2.123	1.179	0.7374	1		
5	1.074	1.175	1.380	1.0740	1		
6	0.737	2.123	1.179	0.7374	1		
7	1.540	1.417	0.868	0.8683			1
8	1.327	2.265	0.488	0.4886			1
9	0.737	2.123	1.179	0.7374	1		
10	1.327	2.265	0.488	0.4886			1
Jumlah					47	10	33

3.7 Hasil Akhir Cluster

Hasil perhitungan dengan menggunakan fungsi SQRT pada *Ms.Excel*, pada iterasi ke empat didapatkan nilai yang sama dengan perhitungan pada iterasi ketiga. Berdasarkan nilai yang sama tersebut, maka proses perhitungan dihentikan cukup sampai dengan iterasi ketiga. Data hasil pengelompokan pada iterasi ketiga menghasilkan 3 *cluster* data dengan pengelompokan sebagai berikut:

- Cluster 1* (C1) berjumlah 47 siswa yang anggotanya data nomor urut ke: 4, 5, 6, 9, 12, 14, 15, 17, 18, 20, 21, 23, 26, 27, 28, 32, 33, 35, 36, 38, 39, 41, 45, 46, 47, 48, 49, 50, 51, 54, 55, 56, 58, 59, 63, 64, 66, 68, 73, 76, 77, 80, 81, 83, 84, 85, 90.
- Cluster 2* (C2) berjumlah 10 siswa yang anggotanya data nomor urut ke : 1, 19, 31, 34, 43, 62, 71, 72, 82, 88.
- Cluster 3* (C3) berjumlah 33 siswa yang anggotanya data nomor urut ke : 2, 3, 7, 8, 10, 11, 13, 16, 22, 24, 25, 29, 30, 37, 40, 42, 44, 52, 53, 57, 60, 61, 65, 67, 69, 70, 74, 75, 78, 79, 86, 87, 89.

Pengujian yang dilakukan dengan menggunakan *Software RapidMiner Studio Versi 9.2* mendapatkan hasil klasterisasi yang sama dengan pengujian secara manual dengan *Ms. Excel*. Hasil pengujian dengan *Software RapidMiner* adalah:

- Cluster 0* (C0) berjumlah 33 siswa yang anggotanya data nomor urut ke: 2, 3, 7, 8, 10, 11, 13, 16, 22, 24, 25, 29, 30, 37, 40, 42, 44, 52, 53, 57, 60, 61, 65, 67, 69, 70, 74, 75, 78, 79, 86, 87, 89
- Cluster 1* (C1) berjumlah 47 siswa yang anggotanya data nomor urut ke: 4, 5, 6, 9, 12, 14, 15, 17, 18, 20, 21, 23, 26, 27, 28, 32, 33, 35, 36, 38, 39, 41, 45, 46, 47, 48, 49, 50, 51, 54, 55, 56, 58, 59, 63, 64, 66, 68, 73, 76, 77, 80, 81, 83, 84, 85, 90
- Cluster 2* (C2) berjumlah 10 siswa yang anggotanya data nomor urut ke: 1, 19, 31, 34, 43, 62, 71, 72, 82, 88

Berdasarkan hasil pengujian dengan menggunakan *Software RapidMiner* memiliki perbedaan penamaan dengan perhitungan menggunakan *Ms.Excel*.

Didapatkan hasil berupa *cluster 0*, *cluster 1* dan *cluster 2*. Sedangkan anggota masing-masing *cluster* sama dengan hasil pengujian dengan menggunakan *Ms.Excel*. Data *cluster 0* pada *RapidMiner* sama dengan data pada *cluster 3 Ms.Excel*. Sedangkan untuk *cluster 1* dan *cluster 2* tidak terdapat perbedaan.

4. Kesimpulan

Hasil klasterisasi data 90 siswa dikelompokkan menjadi 3 *cluster* dengan jumlah data siswa untuk *cluster 1* dengan bobot sikap dan hasil belajar tinggi berjumlah 47 siswa, *cluster 2* dengan bobot sikap dan hasil belajar sedang berjumlah 10 siswa dan *cluster 3* dengan bobot sikap dan hasil belajar rendah berjumlah 33 siswa. Data hasil klasterisasi yang didapatkan akan dibagi secara merata kedalam 3 kelompok belajar siswa untuk mendapatkan komposisi kelas yang seimbang. Pembagian dilakukan oleh guru agar masing-masing kelas memiliki siswa dengan kemampuan dan sikap belajar yang merata sehingga pada akhirnya diharapkan nilai rata-rata setiap kelas meningkat. Penelitian ini dapat dilanjutkan untuk melakukan pembagian siswa kedalam kelompok belajar secara sistematis.

Daftar Rujukan

- Hanifah, E. N., & Wulandari, T. (2018). Penggunaan Metode Card Sort Untuk Meningkatkan Keaktifan Siswa dalam Pembelajaran IPS Kelas VIII E SMP Negeri 1 Majalengka. *Jipsindo*, 5(1), 61. DOI: <https://doi.org/10.21831/jipsindo.v5i1.20184> .
- Setiaji, G. G., Khoirudin, K., & Vydia, V. (2019). Komparasi Metode Clustering K-Means dan Fuzzy C-Means Untuk Memprediksi Ketepatan Waktu Lulus. *Jurnal Pengembangan Rekayasa dan Teknologi*, 15(1), 38. DOI: <https://doi.org/10.26623/jprt.v15i1.1488> .
- Siregar, A. M. (2018). Penerapan Algoritma K-Means Untuk Pengelompokan Daerah Rawan Bencana di Indonesia. *Internal (Information System Journal)*, 1(2), 1–10. DOI: <https://doi.org/10.32627/internal.v1i2.42> .
- Virgo, I., Defit, S., & Yunus, Y. (2020). Klasterisasi Tingkat Kehadiran Dosen Menggunakan Algoritma K-Means Clustering (Studi Kasus Institut Agama Islam Batusangkar). *Jurnal Sistem Informasi dan Teknologi*, 2(1), 24–29. DOI: <https://doi.org/10.37034/jsisfotek.v2i1.22> .
- Asroni, A., Fitri, H., & Prasetyo, E. (2018). Penerapan Metode Clustering dengan Algoritma K-Means pada Pengelompokan Data Calon Mahasiswa Baru di Universitas Muhammadiyah Yogyakarta (Studi Kasus: Fakultas Kedokteran dan Ilmu Kesehatan, dan Fakultas Ilmu Sosial dan Ilmu Politik). *Semesta Teknika*, 21(1). DOI: <https://doi.org/10.18196/st.211211> .
- Zulhendra., Nurcahyo, G. W., & Santony, J. (2018). Algoritma K-Means Clustering Untuk Analisa Data Service Berdasarkan Pengaduan Pelanggan. *UNES Journal of Information System*, 3(1), 1-7. DOI: <https://doi.org/10.31227/osf.io/5byfp> .
- Siburian, T., Safii, M., & Parlina, I. (2019). Penerapan Algoritma K-Means Clustering untuk Pengelompokan Harga Eceran Beras di Pasar Tradisional Berdasarkan Wilayah Kota. *Prosiding Seminar Nasional Riset Information Science (Senaris)*, 1, 927. DOI: <https://doi.org/10.30645/senaris.v1i0.101> .

- [8] Amelio, A., & Tagarelli, A. (2019). Data Mining: Clustering. *Encyclopedia of Bioinformatics and Computational Biology*, 1, 437–448. DOI: <https://doi.org/10.1016/b978-0-12-809633-8.20489-5> .
- [9] Suryadi, S. (2018). Penerapan Metode Clustering K-Means Untuk Pengelompokan Kelulusan Mahasiswa Berbasis Kompetensi. *Jurnal Informatika*, 6(1), 52–72. DOI: <https://doi.org/10.36987/informatika.v6i1.738> .
- [10] Vhallah, I., Sumijan, S., & Santony, J. (2018). Pengelompokan Mahasiswa Potensial Drop Out Menggunakan Metode Clustering K-Means. *Jurnal Resti (Rekayasa Sistem dan Teknologi Informasi)*, 2(2), 572–577. DOI: <https://doi.org/10.29207/resti.v2i2.308> .
- [11] Putra, R. R., & Wadisman, C. (2018). Implementasi Data Mining Pemilihan Pelanggan Potensial Menggunakan Algoritma K Means. *Intecom: Journal of Information Technology and Computer Science*, 1(1), 72–77. DOI: <https://doi.org/10.31539/intecom.v1i1.141> .
- [12] Hasanah, M., Defit, S., & Nurcahyo, G. W. (2019). Implementasi Algoritma K-Means untuk Klasterisasi Peserta Olimpiade Sains Nasional Tingkat SMA. *Jurnal Sistim Informasi dan Teknologi*, 1(3), 29–34. DOI: <https://doi.org/10.35134/jsisfotek.v1i3.7> .
- [13] Pradnyana, G. A., & Permana, A. A. J. (2018). Sistem Pembagian Kelas Kuliah Mahasiswa dengan Metode K-Means dan K-Nearest Neighbors Untuk Meningkatkan Kualitas Pembelajaran. *Juti: Jurnal Ilmiah Teknologi Informasi*, 16(1), 59. DOI: <https://doi.org/10.12962/j24068535.v16i1.a696> .
- [14] Mardalius, M. (2017). Implementasi Algoritma K-Means Clustering Untuk Menentukan Kelas Kelompok Bimbingan Belajar Tambahan (Studi Kasus : Siswa Sma Negeri 1 Ranah Pesisir). *Proceding SEMILOKA ROYAL*. DOI: <https://doi.org/10.31219/osf.io/6mec3> .
- [15] Pandiangan, H. (2019). Penerapan Data Mining dalam Clustering Produksi Daging Sapi di Indonesia Menggunakan Algoritma K-Means. *Journal of Computer Networks, Architecture and High Performance Computing*, 1(2), 37–44. DOI: <https://doi.org/10.47709/cnape.v1i2.239> .
- [16] Omolewa, O. T., Oladele, A. T., Adeyinka, A. A., & Oluwaseun, O. R. (2019). Prediction of Student's Academic Performance using k-Means Clustering and Multiple Linear Regressions. *Journal of Engineering and Applied Sciences*, 14(22), 8254–8260. DOI: <https://doi.org/10.36478/jeasci.2019.8254.8260> .
- [17] Amirulloh, I. (2019). Pemetaan Kelompok Kerja Siswa dengan Metode Clustering K-Means dan Algoritma Greedy. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, 1(2). DOI: <https://doi.org/10.36499/jinrpl.v1i2.2953> .
- [18] Pasina, I., Bayram, G., Labib, W., Abdelhadi, A., & Nurunnabi, M. (2019). Clustering Students Into Groups According to Their Learning Style. *MethodsX*, 6, 2189–2197. DOI: <https://dx.doi.org/10.1016%2Fj.mex.2019.09.026> .