



KLASTERISASI PESERTA OLIMPIADE LOMBA KOMPETENSI SISWA (LKS) MENGGUNAKAN ALGORITMA K-MEANS

Parsaroan Sitohang^{1*}

¹ Sistem Informasi, Universitas Putra Indonesia "YPTK"

^{1*} parsaroan.sitohang22@gmail.com

Article History:

Received Jan 23rd, 2025

Revised Jan 24th, 2025

Accepted Feb 2nd, 2025

Abstrak

Algoritma K-Means adalah metode pengelompokan yang mampu mengorganisasi kumpulan data besar menjadi informasi terstruktur berdasarkan kriteria tertentu. Penelitian ini bertujuan untuk membantu Sekolah Menengah Kejuruan Negeri (SMKN) 1 Bangkinang dalam memilih siswa terbaik yang akan mewakili sekolah pada kompetisi Lomba Kompetensi Siswa (LKS). Data yang digunakan meliputi nilai rapor siswa kelas XI dari program Teknik Komputer dan Jaringan (TKJ) tahun ajaran 2017–2018, dengan fokus pada mata pelajaran produktif yang relevan dengan materi kompetisi. Dengan menggunakan aplikasi RapidMiner, siswa dikelompokkan menjadi tiga kategori: sangat kompeten, kompeten, dan kurang kompeten. Analisis menunjukkan bahwa dari 57 siswa, terbentuk tiga kluster: 27 siswa sangat kompeten (Kluster 1), 19 siswa kompeten (Kluster 2), dan 11 siswa kurang kompeten (Kluster 3). Metode ini memungkinkan sekolah untuk lebih efektif menentukan kriteria dalam memilih perwakilan siswa untuk kompetisi LKS, sehingga meningkatkan kinerja sekolah dari tingkat kabupaten hingga nasional. Penelitian ini menunjukkan bahwa penerapan algoritma K-Means menyediakan solusi berbasis data untuk pengambilan keputusan yang lebih akurat dan sistematis dalam sektor pendidikan.

Kata Kunci : Clustering, K-Means, Lomba Kompetensi Siswa (LKS), Pengambilan Keputusan, RapidMiner.

Abstract

The K-Means algorithm is a clustering method capable of organizing large datasets into structured information based on specific criteria. This study aims to assist Sekolah Menengah Kejuruan Negeri (SMKN) 1 Bangkinang in selecting the best students to represent the school in the Lomba Kompetensi Siswa (LKS) competition. The data used includes the report card scores of 11th-grade students from the Computer and Network Engineering (Teknik Komputer dan Jaringan - TKJ) program for the 2017–2018 academic year, focusing on productive subjects relevant to the competition material. Using the RapidMiner application, students were grouped into three categories: highly competent, competent, and less competent. The analysis revealed that out of 57 students, three clusters were formed: 27 students in the highly competent category (Cluster 1), 19 students in the competent category (Cluster 2), and 11 students in the less competent category (Cluster 3). This method enables the school to establish more effective criteria for selecting student representatives for the LKS competition, thereby improving the school's performance from the district to the national level. The findings indicate that the implementation of the K-Means algorithm provides a data-driven solution for more accurate and systematic decision-making in the education sector.

Keyword : Clustering, Decision Making, K-Means, RapidMiner, Student Competency Competition (LKS).



PENDAHULUAN

Kemajuan teknologi informasi sudah semakin berkembang pesat dalam segala bidang kehidupan. Banyak sekali data yang dihasilkan oleh teknologi informasi yang canggih, mulai dari bidang ekonomi, industri, dan teknologi serta berbagai bidang kehidupan lainnya (Yunita, 2018). Sekolah Menengah Kejuruan Negeri (SMKN) 1 Bangkinang adalah salah satu SMK terbaik yang ada di kabupaten kampar propinsi Riau. SMK ini memiliki 9 jurusan dengan jumlah siswa lebih dari 1.000 siswa. Salah satu jurusan yang terfavorit adalah jurusan Teknik Komputer dan Jaringan (TKJ). Untuk meningkatkan kemampuan kompetensi siswa siswi SMK dibidangnya masing – masing, pemerintah melakukan lomba antar SMK yang ada di seluruh Indonesia. Setiap tahun pemerintah melaksanakan pesta akbar SMK yang di sebut dengan olimpiade Lomba Kompetensi Siswa (LKS), mulai dari tingkat kabupaten, propinsi, nasional hingga internasional. SMKN 1 Bangkinang adalah salah satu SMK yang menjadi peserta yang mewakili kabupaten kampar untuk bertanding di ajang LKS tingkat propinsi. Sekolah tersebut selalu mengutus wakilnya untuk berlomba di level propinsi. Namun TKJ selalu mendapatkan hasil yang belum memuaskan setiap tahunnya. Lomba Kompetensi Siswa (LKS) sendiri sudah dilaksanakan sejak tahun 1992 di kabupaten Malang Jawa Timur untuk pertama kalinya. Kegiatan LKS ini juga dilakukan sebagai seleksi oleh pemerintah terhadap SMK untuk meningkatkan mutu pendidikan SMK dan untuk mendapatkan siswa siswi SMK terbaik dibidang jurusannya masing - masing. Sekolah – sekolah yang akan ikut serta dalam lomba tersebut harus dipertandingkan terlebih dahulu, mulai dari tingkat kabupaten, propinsi hingga menjadi wakil tingkat nasional. Dari tingkat sekolah, siswa di pilih secara langsung oleh guru yang dianggap mampu untuk dikirim sebagai peserta lomba mewakili sekolah. Penelitian ini membantu mencari kriteria siswa yang tepat untuk mengikuti lomba LKS tersebut sehingga prestasi SMKN1 bangkinang lebih baik di masa yang akan datang. Untuk itu metode yang digunakan adalah metode Algoritma K-Means Clustering dimana peneliti mengelompokkan siswa menjadi beberapa kelompok. Siswa dibagi menjadi 3 kelompok yakni, siswa sangat kompeten, siswa kompeten dan siswa kurang kompeten. Data yang digunakan adalah data siswa TKJ kelas XI tamatan 2017 - 2018. Acuan variabel data yang akan digunakan dalam penyelesaian adalah nilai rapor siswa semester 3 dan semester 4. Mata pelajaran yang akan di kluster yakni mata pelajaran produktif atau mata pelajaran kejuruan yang berhubungan langsung dengan materi yang diperlombakan yang tertera di rapor siswa tersebut.

Algoritma K-means merupakan algoritma klasterisasi yang mengelompokkan data berdasarkan titik pusat kluster (centroid) terdekat dengan data. Algoritma Pengelompokan K-Means secara efektif mengelompokkan kompetensi TIK pada siswa, membantu dalam fasilitasi pendidikan dan identifikasi kelompok bimbingan tingkat lanjut (Faisal et al., 2022). Pembentukan cluster merupakan salah satu teknik yang digunakan dalam mengekstrak pola kecenderungan suatu data (Nasari & Darma, 2015:5). Teknik pengelompokan K-means dalam DM menganalisis data evaluasi pendidikan secara efektif, memberikan akurasi 95,6 persen dan memandu para pengambil keputusan dalam mengekstraksi, mengubah, menganalisis, dan membuat model (Liu, 2022). Tujuan dari K-means adalah pengelompokkan data dengan memaksimalkan kemiripan data dalam satu kluster dan meminimalkan kemiripan data antar kluster. Penggunaan centroid awal dalam algoritma K-means dapat meningkatkan efisiensi dan akurasi dalam menghitung jarak dan membuat kluster (Singh & Raheja, 2014). Dengan memilih titik awal yang lebih baik dan penyempurnaan kluster pasca-pemrosesan (N Sujatha et al., 2022). Ukuran kemiripan yang digunakan dalam kluster adalah fungsi jarak. Sehingga pemaksimalan kemiripan data didapatkan berdasarkan jarak terpendek antara data terhadap titik centroid (Asroni & Adrian, 2016). Pengelompokan hierarki aglomeratif merupakan metode pengelompokan hierarki dengan pendekatan bawah atas (bottom up). Proses pengelompokan dimulai dari masing-masing data sebagai satu buah kelompok, kemudian secara rekursif mencari kelompok terdekat sebagai pasangan untuk bergabung sebagai satu kelompok yang besar. Proses tersebut diulang terus sehingga tampak bergerak ke atas membentuk jenjang (hierarki). Cara ini membutuhkan parameter kedekatan kelompok. Pada metode single linked (jarak terdekat), kedekatan di antara dua kelompok ditentukan dari jarak terdekat (terkecil) diantara pasangan diantara dua data dari dua kelompok yang berbeda (satu dakeompok dari kelompok pertama dan satu lagi dari kelompok yang lain, atau disebut juga kemiripan termaksimal (Zahrotun, 2015). Pengertian clustering adalah pengelompokan sejumlah data atau objek ke dalam kluster (group) agar setiap kluster tersebut berisi data yang semirip mungkin dan berbeda dengan data / objek dalam cluster lainnya (Maukar et al., 2022). Penelitian ini bertujuan untuk membantu guru mencari kriteria siswa yang tepat untuk mengikuti Lomba Kompetensi Siswa dan menjadi wakil dari sekolah tersebut, menerapkan metode Algoritma K-Means dalam mengelaster siswa berdasarkan nilai rapor dari mata pelajaran produktif atau kejuruan, menentukan kelompok peserta didik berdasarkan siswa sangat kompeten, kompeten dan kurang kompeten, dan menerapkan metode K-Means dalam mengelompokkan siswa dengan menggunakan aplikasi RapidMiner. Tujuan penelitian diberi batasan permasalahan yang meliputi Algoritma Clustering yang digunakan adalah K-Means terhadap data rapor siswa TKJ kelas XI semester III dan IV di SMKN 1 Bangkinang. Nilai rapor yang dijadikan data adalah nilai mata pelajaran kejuruan yang ada di jurusan TKJ SMKN 1 Bangkinang, dan Jumlah mata pelajaran kejuruan yang akan di cluster adalah sebanyak 3 mata pelajaran kejuruan.

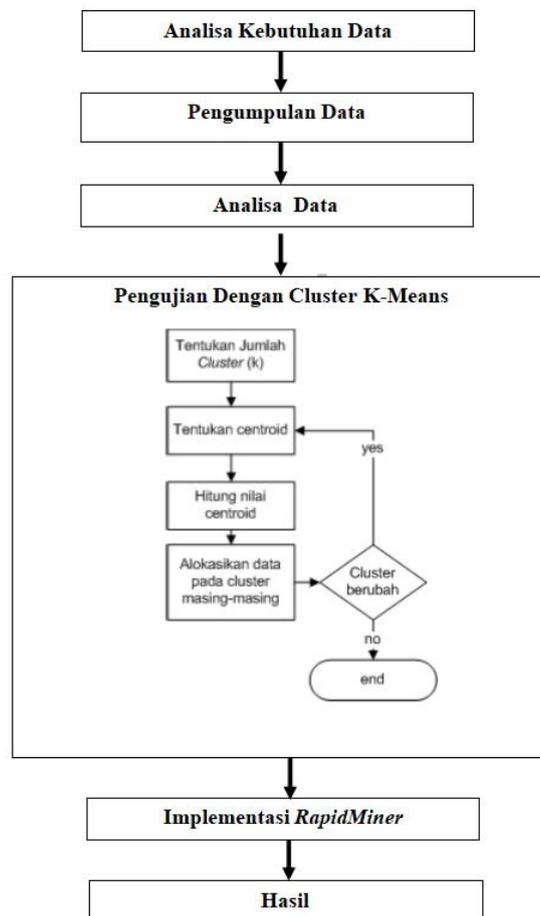
METODE

Kerangka Penelitian

Kerangka kerja penelitian adalah tahapan-tahapan yang di lakukan oleh penulis dalam menyelesaikan penelitian (Setiawan & Muhaqiqin, 2021). Tahapan-tahapan ini nantinya dilaksanakan secara beraturan oleh peneliti dalam awal penelitian sampai hasil yang akan diperoleh. Maka berikutnya dalam proses pembuatan kerangka kerja penelitian ini bisa menggunakan metode Clustering Algoritma K-Means pada Gambar 1. Berdasarkan kerangka kerja penelitian pada Gambar 1, maka masing-masing langkahnya dapat diuraikan tahapan tahapannya sebagai berikut:

a. Analisa kebutuhan data

Permasalahan yang di temukan kemudian akan dianalisa. Langkah dalam proses analisa masalah adalah langkah dalam mempelajari dimana akan di tentukan garis batasan masalahnya. Dengan menganalisa permasalahan dimana tadinya telah di tentukan tersebut, maka diharapkan masalah tersebut bisa dimengerti dengan seksama. Namun pada proses ini penulis menentukan data dalam bentuk leger nilai rapor siswa yang akan dikumpulkan yang



berkesinambungan dengan penelitian ini.

Gambar 1. Kerangka Kerja Penelitian

b. Pengumpulan Data

Pengumpulan data yang dilakukan oleh penulis adalah dengan cara melakukan pengumpulan data leger nilai rapor yang diminta kepada bagian kurikulum langsung. Pengumpulan data yang akan dilaksanakan di dalam penelitian kali ini dengan cara mengumpulkan nilai - nilai rapor siswa melalui leger nilai yang ada pada data bagian kurikulum SMKN1 Bangkinang yang meliputi semua kelas jurusan TKJ.

c. Analisa Data

Pada saat melakukan tahap penganalisaan data di mana data yang tadinya telah di dapat pada tahap proses pengumpulan data, data disusun dan dikelompokkan kedalam rangkaian tabel yang mudah dimengerti oleh semua orang. Kemudian data data yang digunakan dalam penelitian ini dipastikan data yang akurat, jelas, dan sesuai dengan kebutuhan penelitian. Pada penelitian ini data yang digunakan adalah nilai rapor siswa TKJ tamatan 2017 – 2018 mulai dari nilai rapor semester III Hingga semester IV. dalam penelitian ini data akan dikelompokkan menjadi 3 kelompok



yakni siswa sangat kompeten, siswa kompeten dan siswa kurang kompeten. Dan setelah dilakukan pengelompokan data data tersebut maka data tadi kemudian akan dihitung menggunakan metode Algoritma K-Means.

d. Pengujian dengan Cluster K-Means

Tahapan selanjutnya untuk menyelesaikan penelitian dilakukanlah pengelompokan data nilai siswa dengan menggunakan teknik K-Means , dengan langkah-langkah berikut ini:

1. Pilih jumlah cluster K
2. Inisialisasi K pusat cluster ini dapat sebenarnya bisa di lakukan dengan berbagai cara. Akan tetapi cara yang paling sering dilakukan adalah dengan cara memilih secara sembarangan atau secara acak. Pusat-pusat cluster diberi nilai awal dengan angka-angka yang random
3. Alokasi seluruh data/objek ke cluster terdekat. Dua objek ditentukan oleh jarak dari kedua objek tersebut sama halnya juga dengan cluster dengan suatu data juga ditentukan oleh jarak antara data tersebut dengan pusat cluster. Dalam tahap ini dapat dihitung jarak dari tiap data ke tiap data pusat cluster.untuk memasukan ke kluster yang mana data dari satu cluster tertentu akan menentukan data tersebut akan masuk ke cluster yang mana. Dalam menghitung jarak data data yang ada tertuju kesetiap pusat cluster kita dapat memanfaatkan rumus yang ada yakni rumus teori jarak Euclidean, rumus tersebut seperti (1)

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

Dimana:

$D_{(i,j)}$ = Jarak data ke i ke pusat cluster j

X_{ki} = Data ke i pada atribut data ke k

X_{kj} = Titik pusat ke j pada atribut ke k

4. Hitung kembali pusat cluster dengan keanggotaan cluster yang ada pada saat ini. Pusat cluster yaitu rata-rata dari semua data/ objek dalam cluster tertentu. Jika ingin menggunakan alternatif lain yaitu dapat menggunakan median dari cluster tersebut. Jadi rata-rata (mean) bukan satu-satunya ukuran yang dapat digunakan.
5. Tiap objek memakai pusat cluster yang baru apabila dilakukan perhitungan kembali. Dan apabila dari pusat cluster tidak lagi terjadi perubahan maka bisa dianggap proses clustering disebut dengan selesai.

e. Implementasi RapidMiner

Tahapan berikutnya yang akan dilakukan dalam penelitian ini adalah melakukan implementasi. Pada tahap implementasi ini penulis akan menghitung data nilai rapor siswa yang telah diperoleh menggunakan aplikasi RapidMiner dengan perhitungan algoritma K-Means

f. Hasil

Pada tahapan hasil penulis akan mengetahui hasil perhitungan Algoritma K-Means menggunakan excel dan aplikasi RapidMiner. Dari hasil penelitian ini akan didapatkan sebuah hasil pengelompokan dari siswa tersebut sesuai dengan cluster mereka masing-masing, dan penulis dapat merekomendasikan hasilnya kepada sekolah sebagai pertimbangan pengambilan keputusan.

HASIL DAN PEMBAHASAN

Data yang dipergunakan adalah data rapor siswa tamatan 2017 – 2018, format data yang yang digunakan penulis dalam penelitian ini adalah nilai dalam format excel. Data yang didapat diolah kembali menjadi data yang dapat dijadikan sebagai bahan uji coba dalam penelitian ini. Dalam data mentah atau data nilai rapor siswa awal, terlalu banyak mata pelajaran yang didapat sedangkan dalam penelitian ini hanya menggunakan mata pelajaran kejuruan lebih tepatnya hanya menggunakan 3 mata pelajaran kejuruan. Leger nilai rapor yang digunakan adalah nilai kejuruan TKJ semester 3 dan 4. Nilai semester 3 dijumlahkan dengan nilai semester 4 lalu dibagi 2 maka didapatkanlah nilai rata – rata dari semester 3 dan 4. Nilai kedua lokal yakni TKJ1 dan TKJ2 digabung secara langsung karna untuk perlombaan LKS tersebut yang akan diberangkatkan adalah siswa TKJ kelas 2 maka penulis menggunakan data langsung untuk 2 kelas, selengkapny tabel nilai dapat dilihat di tabel 1.

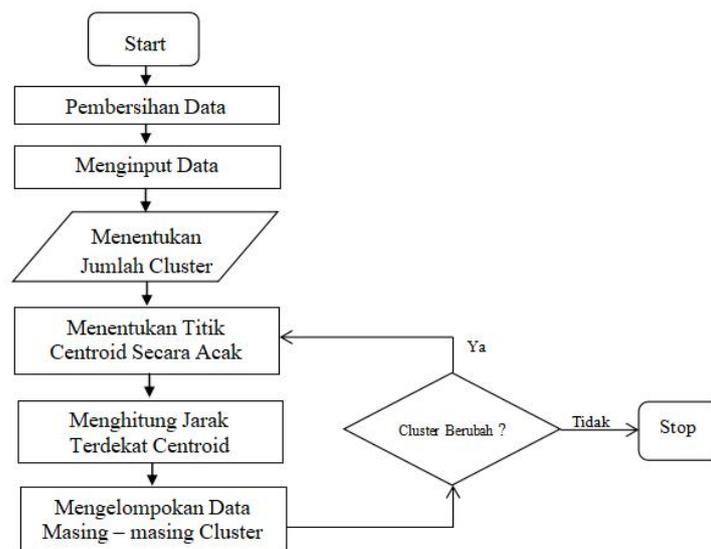
Tabel 1 Daftar Nilai Produktif TKJ Semester 3

No	Nama Siswa	Mata Pelajaran
----	------------	----------------



		PD	SK	KT	KD	SOJ	AS	RBJ
1	Adek Nur Apriliani	80	80	75	78	77	80	75
2	Alif Adrian Utama	88	86	95	85	86	80	95
3	Ani Riani	80	80	75	78	75	78	75
...
...
56	Wahyudi Pratama	84	82	95	82	91	82	95
57	Wulan Hapsari	84	80	95	80	88	84	95

Berdasarkan algorithma dari proses K-Means dapat digambarkan melalui flowchart pada Gambar 2.



Gambar 2 Flowchart Proses K-Means

Pembersihan data atau cleaning data merupakan suatu proses pengurangan data awal atau menghilangkan noise, seperti membuang duplikasi data, memeriksa data yang tidak konsisten, dan memperbaiki kesalahan terhadap data. Dalam hal ini kenapa penulis memilih 3 daftar pelajaran sebagai bahan data untuk perhitungan K-means dikarenakan 3 mata pelajaran inilah yang sangat relevan dan berhubungan langsung dengan materi yang diperlombakan dalam olimpiade LKS tersebut. 3 mata pelajaran tersebut yakni Sistem Operasi Jaringan (SOJ), Adminstrasi Server (AS) dan terakhir Rancang Bangun Jaringan (RBJ) data dapat dilihat pada tabel

Tabel 2. Nilai 3 Mata Pelajaran Untuk Perhitungan K-Means

No.	Nama Siswa	Mata Pelajaran		
		SOJ	AS	RBJ
1	Adek Nur Apriliani	78	80	75
2	Alif Adrian Utama	85	83	90
3	Ani Riani	77	79	75
...
56	Wahyudi Pratama	91	85	94
57	Wulan Hapsari	88	82	93

Menentukan Jumlah Cluster

Dalam metode K-Means kita harus menentukan jumlah cluster yang akan menjadi acuan dalam hasil pengambilan keputusan. Dalam penelitian ini cluster dibagi menjadi 3 cluster. Cluster ini menentukan nantinya siswa mana yang sangat kompeten, siswa mana yang kompeten dan yang terakhir siswa mana yang kurang kompeten

Menentukan Titik Centroid Secara Acak

Menentukan centroid awal dalam perhitungan ini dilakukan dengan cara acak, peneliti mengambil langkah yakni mengambil nilai tertinggi, nilai terendah dan nilai menengah siswa yang akan diuji coba. Jarak dari setiap objek



ke masing-masing centroid menggunakan rumus korelasi antar tiga objek yaitu Euclidean Distance (D) Untuk menentukan An diambil dari Nilai mata pelajaran SOJ (X), nilai mata pelajaran Administrasi Server AS (Y), dan nilai mata pelajaran Rancang Bangun Jaringan RBJ (Z) maka dari itu didapatkanlah yaitu:

- a. Semua data akan di kelompokkan ke dalam tiga cluster
- b. Center points dari ketiga cluster yang ditentukan secara random adalah :
 Pusat cluster 1 (C1) : di ambil dari siswa nomor 56 (91, 85, 94)
 Pusat cluster 2 (C2) : di ambil dari siswa nomor 19 (88, 85, 96)
 Pusat cluster 3 (C3) : di ambil dari siswa nomor 36 (86, 80, 92)

Untuk mencari kriteria siswa yang akan mewakili sekolah dalam mengikuti LKS maka dibutuhkan atribut atribut dalam bentuk 3 mata pelajaran kejuruan. Percobaan akan dilakukan dengan menggunakan parameter-parameter sebagai berikut :

- Jumlah data : 57
 - Jumlah cluster : 3 (sangat kompeten, kompeten dan kurang kompeten)
 - Jumlah atribut : 4 (Nama siswa, SOJ, AS dan RBJ)
- Adapun sampel data nilai siswa tampak pada tabel 4.5 dibawah ini :

Untuk menghitung pusat cluster 1 tahapan ini akan dihitung dari jarak setiap data kemasing masing centroid menggunakan rumus Euclidean Distance (D) iterasi pertama. Menghitung jarak masing masing data ke titik pusat cluster (C1).

- a. Pada stage (D1), yaitu atas nama siswa Adek nur apriliani atau dengan kode T1, memiliki jarak titik pusat cluster (C1) adalah : $D1 = \sqrt{(78-91)^2 + (80-85)^2 + (75-94)^2} = 23,55844$
- b. Pada stage (D2), yaitu atas nama siswa Alif adrian utama atau dengan kode T2, memiliki jarak titik pusat cluster (C1) adalah : $D2 = \sqrt{(85-91)^2 + (83-85)^2 + (90-94)^2} = 7,483315$

Untuk Dn seterusnya akan di tampilkan pada Tabel 3.

Tabel 3. Pengelompokan data dalam cluster dengan jarak terdekat.

No	Jarak ke C1	Jarak ke C2	Jarak ke C3	Jarak terdekat C1	Jarak terdekat C2	Jarak terdekat C3
1	23,5584	23,7908	18,7883			1
2	7,48331	7	3,74166			1
3	24,3516	24,454	19,2614			1
...
56	0	3,60555	7,34847	1		
57	4,3589	4,24264	3			1
JUMLAH				3	10	44

Setelah kita mendapatkan hasil pengelompokkan Kemudian lakukan iterasi kembali, sebelumnya kita tentukan terlebih dahulu posisi centroid baru dengan cara menghitung rata-rata dari data-data yang berada pada centroid yang sama. Berikut caranya kita menentukan centroid baru dari hasil iterasi pertama:

$$C1(X1) = \frac{(90+91+91)}{3}$$

$$C1(X1) = 90,67$$

$$C1(Y1) = \frac{(82+82+85)}{3}$$

$$C1(Y1) = 83$$

$$C1(Z1) = \frac{(94+93+94)}{3}$$

$$C1(Z1) = 93,67$$



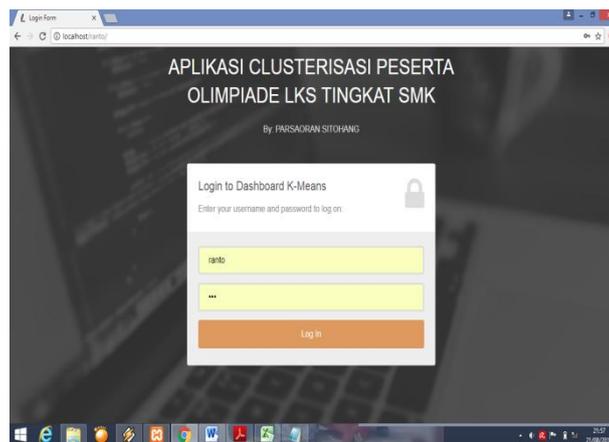
Setelah proses perhitungan di atas, maka akan diperoleh *centroid* baru yang dapat di lihat pada tabel 4. Setelah mendapatkan hasil pengelompokkan dari iterasi keenam dan telah mendapatkan nilai yang sama dengan hasil dari iterasi ke tujuh maka pencarian dihentikan dan tidak memerlukan *centroid* baru lagi, maka iterasi cukup sampai dengan iterasi ke enam dan dihentikan.

Tabel 4 Nilai *Centroid* Baru Untuk Iterasi II

CENTROID	X	Y	Z
C1	90,67	83	93,67
C2	84,9	85,8	95,1
C3	82,48	80,75	87,98

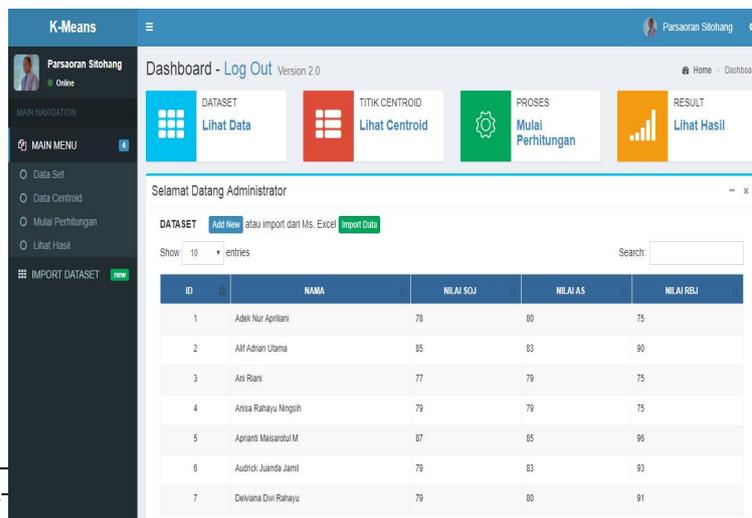
Implementasi

Untuk mengimplementasikan dan menguji rancangan Data Mining dengan menggunakan Algoritma K-Means diperlukan sebuah alat bantu yaitu software Data Mining yang disebut dengan RapidMiner. RapidMiner merupakan software open source yang berguna sebagai solusi untuk melakukan analisis terhadap Data Mining, text mining dan analisis prediksi. RapidMiner memiliki berbagai teknik deskriptif dan prediksi dalam memberikan wawasan kepada pengguna sehingga dapat membuat keputusan paling baik. Proses implementasi dan analisa data dilakukan dengan menggunakan program aplikasi RapidMiner, dengan melakukan pengujian pada data yang telah didapatkan dari SMKN 1 Bangkinang dalam bentuk leger nilai siswa TKJ tahun 2018 dengan paramater pelajaran kejuruan sebanyak 7 mata pelajaran yang dicleaning menjadi 3 mata pelajaran kejuruan dan dengan jumlah anak 57 siswa yang terdiri dari siswa TKJ1 dan siswa TKJ2.



Gambar 3 Halaman Login Sistem

Form login berfungsi memberikan akses kepada user atau admin untuk dapat masuk kedalam sistem aplikasi dengan cara memasukkan username dan password yang telah dibuat dalam database seperti gambar 3. Dashboard atau tampilan halaman muka merupakan tampilan awal pada form utama. Dalam halam ini terdapat beberapa menu yang dapat digunakan oleh user. Tampilan halam muka ini dapat dilihat pada gambar 4.





Gambar 4. Form Dashboard

KESIMPULAN

Hasil penelitian dapat membantu dalam menentukan siswa yang akan menjadi wakil sekolah dalam mengikuti olimpiade LKS dengan cara memperlombakan kembali siswa yang tergolong dalam cluster 1 yang berjumlah 27 siswa, dimana dalam cluster 1 (Siswa Sangat Kompeten). Dari hasil perhitungan dengan menggunakan metode K-Means Clustering maka didapatkanlah hasil dari pengelompokan siswa tersebut menjadi 3 cluster. Dalam penelitian ini tujuan utamanya adalah mengelompokkan 57 jumlah siswa menjadi 3 kelompok, dan setelah melalui proses perhitungan maka didapatkanlah hasil sesuai dengan cluster masing – masing yakni C1 berjumlah 27 siswa dengan kategori sangat kompeten, C2 berjumlah 19 siswa dengan kategori kompeten dan terakhir C3 berjumlah 11 siswa dengan kategori kurang kompeten.

DAFTAR PUSTAKA

- Astroni, A., & Adrian, R. (2016). Penerapan Metode K-Means Untuk Clustering Mahasiswa Berdasarkan Nilai Akademik Dengan Weka Interface Studi Kasus Pada Jurusan Teknik Informatika UMM Magelang. *Semesta Teknika*, 18(1), 76–82. <https://doi.org/10.18196/st.v18i1.708>
- Faisal, M., Nurdin, N., Fajriana, F., & Fitri, Z. (2022). Information and Communication Technology Competencies Clustering For Students For Vocational High School Students Using K-Means Clustering Algorithm. *International Journal of Engineering, Science and Information Technology*, 2(3), 111–120. <https://doi.org/10.52088/ijesty.v2i3.318>
- Liu, R. (2022). Data Analysis of Educational Evaluation Using K-Means Clustering Method. *Computational Intelligence and Neuroscience*, 2022. <https://doi.org/10.1155/2022/3762431>
- Maukar, A. L., Marisa, F., & Widodo, A. A. (2022). Analisis Data Penerimaan Mahasiswa Baru Berbasis K-Means. *JIKO (Jurnal Informatika Dan Komputer)*, 6(2), 142. <https://doi.org/10.26798/jiko.v6i2.558>
- N Sujatha, Latha Narayanan Valli, A Prema, SK Rathiha, & V Raja. (2022). Initial centroid selection for K- means clustering algorithm using the statistical method. *International Journal of Science and Research Archive*, 7(2), 474–478. <https://doi.org/10.30574/ijrsra.2022.7.2.0309>
- Nasari, F., & Darma, S. (2015). *Seminar Nasional Teknologi Informasi dan Multimedia 2015 PENERAPAN K-MEANS CLUSTERING PADA DATA PENERIMAAN MAHASISWA BARU (STUDI KASUS : UNIVERSITAS POTENSI UTAMA)*. 6–8.
- Setiawan, R. P., & Muhaqiqin, M. (2021). Sistem Informasi Manajemen Presensi Siswa Berbasis Mobile Studi Kasus SMAN 1 Sungkai Utara Lampung Utara. ... *Teknologi Dan Sistem Informasi*, 2(3), 119–124. <https://jim.teknokrat.ac.id/index.php/sisteminformasi/article/view/898%0Ahttps://jim.teknokrat.ac.id/index.php/sisteminformasi/article/viewFile/898/394>
- Singh, M., & Raheja, S. (2014). *Enhancing the Exactness of K-Means Clustering Algorithm by Centroids*. 5, 294–300.
- Yunita, F. (2018). Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru. *Sistemasi*, 7(3), 238. <https://doi.org/10.32520/stmsi.v7i3.388>
- Zahrotun, L. (2015). Analisis Pengelompokan Jumlah Penumpang Bus Trans Jogja Menggunakan Metode Clustering K-Means Dan Agglomerative Hierarchical Clustering (Ahc). *Jurnal Informatika*, 9(1), 1039–1047. <https://doi.org/10.26555/jifo.v9i1.a2045>