

Analisis Clustering Untuk Menentukan Siswa Berprestasi di SMK Swasta TI Panca Dharma Stungkit Menggunakan Metode K-Means

¹Dicky Apriandi, ²Rahayu Mayang Sari, ³Muhammad Irfan Sarif
Sistem Komputer, Universitas Pembangunan Panca Budi Medan

dickyapriandi123@gmail.com, rahayu@dosen.pancabudi.ac.id, irfanberbagi@gmail.com

ABSTRAK

Pendidikan memiliki peran yang amat penting dalam masyarakat karena pendidikan merupakan upaya untuk memperbaiki taraf hidup manusia. Peningkatan mutu pendidikan merupakan prioritas kebijakan umum pembangunan di Indonesia yang harus terus diperhatikan. Faktor yang memiliki pengaruh besar dalam pencapaian tersebut adalah guru yang memberikan pengajaran efektif dan berkualitas. SMK TI Panca Dharma Stungkit yang beralamat di Jln. Kata Jadi Stungkit, Kecamatan Wampu, Kabupaten Langkat, Sumut. Penilaian prestasi siswa yaitu mata pelajaran baik teori dan praktek, penilaian ekstrakurikuler kehadiran dan ketidakhadiran siswa dan pada saat mengikuti ekstrakurikuler, yang dikategorikan dalam nilai pengetahuan, nilai keterampilan dan penilaian sikap. metode untuk mengolahnya menjadi sebuah informasi dan pengetahuan yang dapat digunakan sebagai bahan pertimbangan pendidik dalam proses pengambilan kebijakan dan keputusan yaitu dengan metode K-Means.

Kata Kunci: Clustering, Siswa, Berprestasi, K-Means

PENDAHULUAN

Latar Belakang

Dewasa ini pengolahan data elektronika telah menjadi kebutuhan yang sangat utama. Perkembangan pesat dalam teknologi informasi yang menjadikan semua informasi dapat disimpan dalam jaringan komputer telah membuat munculnya sistem basis data yang sangat besar. Dalam hitungan detik, data-data dalam berbagai basis data akan senantiasa terbaru, baik dikarenakan adanya update maupun penambahan data baru (Sibuea & Safta, 2017).

Hal ini dijelaskan dalam UU No.20/2003 mengenai pendidikan yang menyatakan bahwa pendidikan merupakan usaha sadar dan terencana untuk menimbulkan suasana belajar serta proses pembelajaran supaya peserta didik dapat secara aktif dalam pengembangan potensi diri (Sovia et al., 2020). Peningkatan mutu pendidikan merupakan prioritas kebijakan umum pembangunan di Indonesia yang harus terus diperhatikan. Faktor yang memiliki pengaruh besar dalam pencapaian tersebut adalah guru yang memberikan pengajaran efektif dan berkualitas. Perkembangan kualitas dan keaktifan siswa dalam pembelajaran merupakan unsur dasar sebagai bentuk keberhasilan dalam proses pembelajaran yang tentu saja tidak semua siswa memiliki tingkat kecepatan dalam memahami suatu materi (Dewi et al., 2022).

Tujuan Penelitian

Adapun tujuan penelitian yang akan dicapai dari latar belakang tersebut adalah :

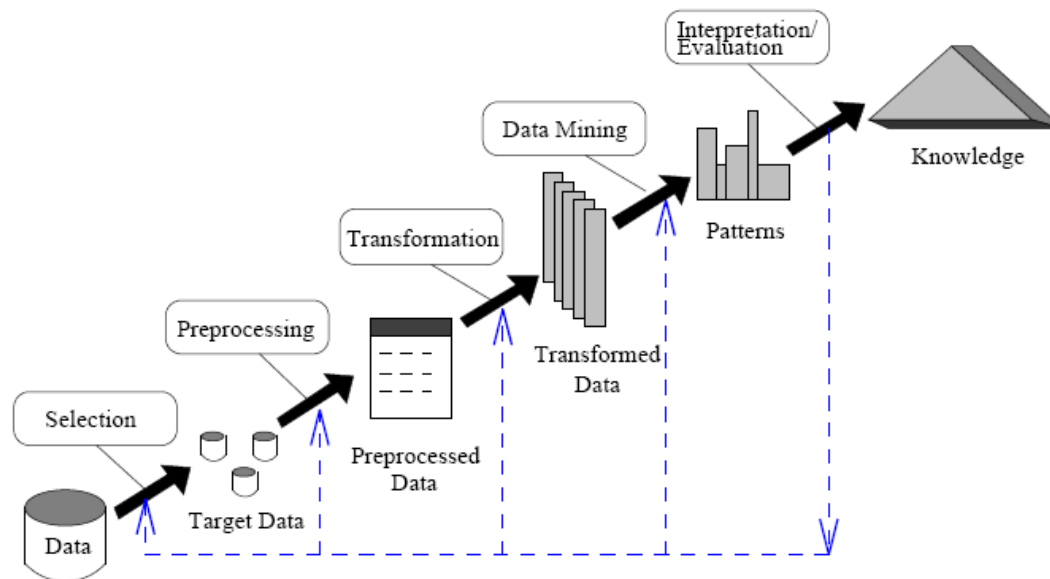
1. Untuk mengetahui siswa berprestasi di SMK TI Panca Dharma Stungkit.
2. Untuk mengimplementasikan metode K-Means dengan software Tanagra.

TINJAUAN PUSTAKA

Knowledge Discovery in Database

Ilmu pengetahuan diketahui sebagai basis data, dan dikenal dalam bahasa asing yaitu Knowledge Discovery in Database (KDD), hal ini merujuk pada cara yang dilakukan untuk metode untuk menemukan pengetahuan yang tersimpan pada database. Dalam basis data, terdapat tabel yang memiliki relasi satu sama lain. Proses KDD melibatkan tahapan-tahapan seperti pemrosesan data dan analisis untuk menghasilkan pengetahuan yang berharga. Pengetahuan yang dihasilkan atau diperoleh melalui proses KDD dapat digunakan sebagai dasar untuk memutuskan keputusan penting atas masalah yang dihadapi (Lowensky & Elisa, 2023)

Proses Knowledge Discovery in Databases (KDD) secara garis besar dapat dijelaskan sebagai berikut (Sari et al., 2023a):



Gambar 1. Tahapan Knowledge Discovery in Databases (KDD)

1. Data Selection

Pemilihan (seleksi) data dari sekumpulan data operasional perlu dilakukan sebelum tahap penggalian informasi dalam knowledge data discovery dimulai. Data hasil seleksi yang akan digunakan untuk proses Data Mining, disimpan dalam suatu berkas, terpisah dari basis data operasional.

2. Pre-processing atau cleaning

Sebelum proses Data Mining dapat dilaksanakan, perlu dilakukan proses cleaning pada data yang menjadi fokus knowledge data discovery. Proses cleaning mencakup antara lain membuang duplikasi data, memeriksa data yang inkosisten, dan memperbaiki kesalahan pada data, seperti kesalahan cetak (tipografi). Juga dilakukan proses enrichment, yaitu proses memperkaya data yang sudah ada dengan data atau informasi lain yang relevan dan diperlukan KDD, seperti data atau informasi eksternal.

3. Transformation

Coding adalah proses transformasi pada data yang telah dipilih, sehingga data tersebut sesuai untuk proses Data Mining. Proses coding dalam KDD merupakan proses kreatif dan sangat tergantung pada jenis atau pola informasi yang akan dicari dalam basis data.

4. Data Mining

Data Mining adalah proses mencari pola atau informasi menarik dalam data terpilih dengan menggunakan teknik atau metode tertentu. Teknik, metode, atau algoritma dalam Data Mining sangat bervariasi. Pemilihan metode atau algoritma yang tepat sangat bergantung pada tujuan dan proses KDD secara keseluruhan.

5. Interpretation atau Evaluation

Pola informasi yang dihasilkan dari proses Data Mining perlu ditampilkan dalam bentuk yang mudah dimengerti oleh pihak yang berkepentingan. Tahap ini merupakan bagian dari proses KDD yang disebut interpretation. Tahap ini mencakup pemeriksaan apakah pola atau informasi yang ditemukan bertentangan dengan fakta atau hipotesis yang ada sebelumnya.

Data Mining

Data Mining adalah suatu istilah yang digunakan untuk menguraikan penemuan pengetahuan dalam database. Data Mining adalah proses yang menggunakan teknik variabel, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar. Data Mining adalah suatu proses menemukan hubungan yang berarti, pola, dan kecenderungan dengan memeriksa dalam sekumpulan besar data yang tersimpan dalam penyimpanan dengan menggunakan teknik pengenalan pola seperti teknik 1088 variabel dan matematika (Sari et al., 2023b).

Data mining merupakan salah satu teknik untuk menggali atau “menambang” pengetahuan dari sekumpulan besar data. Data mining merupakan analisis dari peninjauan kumpulan data untuk menemukan hubungan yang tidak diduga dan meringkas data dengan cara yang berbeda dengan sebelumnya yang dapat dipahami dan bermanfaat bagi pemilik data. Data mining adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan dan machine learning untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai database besar

Data mining biasanya mengolah data dari database dengan ukuran yang besar. Dari data tersebut dilakukan pencarian pola atau trend sesuai dengan tujuan dari penerapan data mining tersebut. Hasil dari pengolahan data mining tersebut selanjutnya dapat digunakan untuk pengambilan keputusan maupun analisis yang dibutuhkan.

Terdapat beberapa alasan mengapa ilmu data mining dibutuhkan saat ini diantaranya terdapat sejumlah besar data di suatu perusahaan atau organisasi yang hanya tersimpan di dalam database tanpa dianalisis lebih lanjut untuk digunakan untuk perkembangan perusahaan atau organisasi tersebut. Selain itu dengan perkembangan internet yang sangat pesat, memberikan dampak positif dengan kemudahan akses data dengan berbagai perangkat hardware dan software yang memiliki daya komputasi dan kapasitas yang luar biasa.

Meskipun algoritma data mining biasanya diterapkan untuk ukuran data yang besar, beberapa algoritma bisa juga diterapkan untuk ukuran data yang relatif kecil. Kumpulan data yang digunakan dalam data mining sederhana dalam struktur dimana baris menjelaskan kasus-kasus individu (juga disebut sebagai pengamatan atau contoh) dan kolom menggambarkan atribut atau variabel dari kasus-kasus. Pilihan algoritma yang akan digunakan tergantung pada jenis data (yaitu, nominal, ordinal, rasio atau interval) (Widaningsih, 2019).

Clustering

Clustering pada suatu data adalah suatu tahapan untuk menggolongkan himpunan data yang atribut kelasnya belum dideskripsikan, secara konsep clustering adalah untuk memaksimalkan dan meminimalkan kemiripan intra antar kelas. sebagai contoh, ada suatu himpunan obyek, proses pertama dapat di klasterisasi menjadi beberapa himpunan kelas selanjutnya menjadi sebuah himpunan beraturan sehingga dapat diturunkan berdasarkan kelompok klasifikasi tertentu. Cluster juga dapat diartikan sebagai kelompok. Maka analisa clustering pada dasarnya akan menghasilkan sejumlah cluster (kelompok).

Tujuan dari pada clustering data adalah untuk meminimalisasikan fungsi tujuan yang ditetapkan didalam proses clustering, dan umumnya selalu meminimalisasi variasi suatu cluster dan memaksimalkan variasi antar cluster (Muliono & Sembiring, 2019).

K-Means

K-Means adalah salah satu teknik clustering pada Data Mining proses pemodelan tanpa supervisi dan metode pengelompokan data secara partisi. Data yang dikelompokkan metode K-Means menjadi beberapa kelompok dan setiap kelompok memiliki karakteristik yang mirip atau sama dengan lainnya tetapi dengan kelompok lainnya memiliki karakteristik berbeda. Dengan tujuan meminimalisasi perbedaan setiap data didalam satu cluster serta memaksimalkan perbedaan dengan cluster yang lain.

Istilah-istilah didalam algoritma k-means clustering (Muliono & Sembiring, 2019):

1. Cluster: adalah kelompok atau grup.
2. Centroid: adalah titik pusat untuk menentukan euclidian distance.
3. Iterasi: adalah pengulangan proses, berhenti ketika hasil iterasi telah konvergen.

Dasar algoritma K-Means adalah sebagai berikut (Sari et al., 2023a):

1. Inisialisasi : tentukan nilai K sebagai jumlah *cluster* yang diinginkan dan metrik ketidakmiripan (jarak) yang diinginkan. Jika perlu, tetapkan ambang batas perubahan fungsi objektif dan ambang batas perubahan posisi *centroid*.
2. Pilih K data dari set data X sebagai centroid.
3. Alokasikan semua data ke centroid terdekat dengan jarak metrik jarak yang sudah ditetapkan .

$$D = \sqrt{(m_1x - c_1x)^2 + (m_1y - c_1y)^2}$$

4. Hitung kembali centroid C berdasarkan data yang mengikuti *cluster* masing-masing.

$$C_k = \left(\frac{1}{C_k}\right) \sum d_i$$

5. Ulangi langkah 3 dan 4 hingga kondisi konvergen tercapai, yaitu :
 - a. Perubahan fungsi objektif sudah di bawah ambang batas yang diinginkan.
 - b. Tidak ada data yang berpindah *cluster*.
 - c. Perubahan posisi centroid sudah di bawah ambang batas yang ditetapkan.

K-Means digunakan dalam mempartisi cluster, yang berarti bahwa setiap bagian data harus dimasukkan ke dalam clusternya sendiri, dan setiap bagian data yang digunakan untuk cluster tertentu pada satu titik dalam proses dapat dipindahkan ke cluster lain pada fase berikutnya. Algoritma K-Means terkenal karena kemudahan penggunaan dan kemampuannya untuk dengan mudah mengidentifikasi sejumlah besar data dan outlier (Ikhwan & Aslami, 2020).

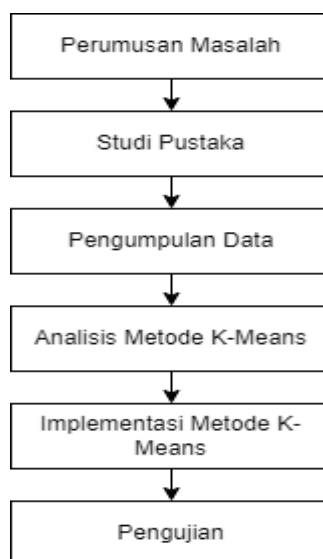
Tanagra

Aplikasi Tanagra adalah perangkat lunak bebas untuk tujuan akademik dan penelitian. Penelitian ini melibatkan beberapa metode pada data mining dimulai dari analisis eksplorasi data, pembelajaran statistik, pembelajaran mesin, dan basis data. Tanagra merupakan salah satu software data mining yang didalamnya disediakan beberapa metode data mining mulai dari mengeksplorasi analisis data, pembelajaran statistik, pembelajaran mesin, dan data mining kebanyakan.

Tanagra merupakan suatu software berbasis open source dimana semua orang dapat mengakses source code dan menambahkan algoritma mereka sendiri, sejauh dia setuju dan menyesuaikan dengan lisensi pendistribusian software. Tanagra adalah salah satu software data mining yang di dalamnya terdapat beberapa metode data mining, Tanagra dapat digunakan oleh semua orang karena bersifat open source (Adrianto et al., 2020).

METODE PENELITIAN

Tahapan Penelitian



Gambar 2 Tahapan Penelitian

Metode Pengumpulan Data

Beberapa teknik pengumpulan data yang harus dilakukan untuk bisa menyelesaikan penelitian dan memperoleh keperluan penelitian. Metode pengumpulan data yang dilakukan peneliti sebagai berikut.

1. Observasi

Observasi ini digunakan agar peneliti lebih mudah dalam melakukan pengamatan. Observasi ini melihat keadaan lapangan dan perencanaan analisis system yang akan diimplementasikan menggunakan software Tanagra pada SMK TI Panca Dharma Stungkit.

2. Wawancara

Pelaksanaan wawancara dilakukan secara semi terstruktur. Wawancara semi terstruktur ini masih bisa pertanyaan lain diluar daftar yang telah ditentukan agar lebih menjangkau data yang diperlukan melalui pertanyaan-pertanyaan yang telah disiapkan. Wawancara ini berupa pengumpulan informasi mengenai daftar absensi siswa, nilai tugas, nilai sikap, nilai UTS, nilai UAS serta nilai ekstrakurikuler pada SMK TI Panca Dharma Stungkit dan hal lain yang berkaitan dengan analisis yang akan dibuat.

Analisis Metode K-Means

1. Menentukan Kelompok Data

Jarak tiap objek ke masing-masing centroid menggunakan rumus kolerasi antar dua objek yaitu Euclidean Distance.

Tabel 1 Data Nilai Siswa

NO	NAMA	NILAI	EKTRAKURIKULER
1	Aji Rahma Dani	81	80
2	Andrian	79	80
3	Antoni Praniko	76	80
4	Ariya Amri Nata	80	80
5	Bayu Afrizal	77	80

NO	NAMA	NILAI	EKTRAKURIKULER
6	Benny Irwansah	76	80
7	Christ Steven Simamora	79	80
8	Dani Alfa Rizky	76	80
9	David Gun	84	89
10	Doni Surya Darma	85	89
11	Giffari	77	80
12	Hengky Febriansyah	76	80
13	Irfansyah	75	80
14	Lucky Andrian	77	80
15	M. Alvin Ramaddan	71	70
16	Marchel Amandika	76	77
17	Mhd. Tri Rizky Farehan	76	77
18	Muhammad Atan Tira	86	89
19	Muhammad Mahadir Ardyansyah	85	89
20	Nikolas Mhs Putra Sembiring	45	70
21	Pramudia Maulana	76	75
22	Rama Kurniawan	75	75
23	Rehan Fikri Rudyansyah	85	89
24	Wahyu Saputra	74	75

2. Proses Clustering

Proses algoritma K-means dapat dilakukan dengan tahapan sebagai berikut :

1. Tentukan nilai centroid secara manual dengan menggunakan rumus :

$$C_k = \left(\frac{1}{nk} \right) \sum d$$

C_k = dari kelompok data yang sudah di proses oleh Tanagra.

2. Langkah selanjutnya adalah melakukan perhitungan untuk menentukan jarak setiap data dan centroid yang telah di bentuk dengan menggunakan rumus Euclidean Distance :

$$D_{11} = \sqrt{(m_1x - c_1x)^2 + (m_1y - c_1y)^2}$$

HASIL DAN PEMBAHASAN

Perhitungan Iterasi 1

$$D_{1.1} = \sqrt{(m_1x - c_1x)^2 + (m_1y - c_1y)^2}$$

$$D_{1.1} = \sqrt{(81 - 81)^2 + (80 - 4)^2}$$

$$D_{1.1} = 0$$

$$D_{1.2} = \sqrt{(m_1x - c_2x)^2 + (m_1y - c_2y)^2}$$

$$D_{1.2} = \sqrt{(81 - 76)^2 + (80 - 8)^2}$$

$$D_{1.2} = 6,4$$

Tabel 2 Hasil Iterasi 1

Aji Rahma Dani
Andrian
Ariya Amri Nata
Bayu Afrizal
Christ Steven Simamora
David Gun
Doni Surya Darma
Giffari
Lucky Andrian
Muhammad Atan Tira
Muhammad Mahadir Ardyansyah
Rehan Fikri Rudyansyah

Perhitungan Iterasi 2

$$D_{1.1} = \sqrt{(m_1x - c_1x)^2 + (m_1y - c_1y)^2}$$

$$D_{1.1} = \sqrt{(81 - 81)^2 + (80 - 4)^2}$$

$$D_{1.1} = 0$$

$$D_{1.2} = \sqrt{(m_1x - c_2x)^2 + (m_1y - c_2y)^2}$$

$$D_{1.2} = \sqrt{(81 - 73)^2 + (80 - 8)^2}$$

$$D_{1.2} = 8,2$$

Tabel 3 Hasil Iterasi

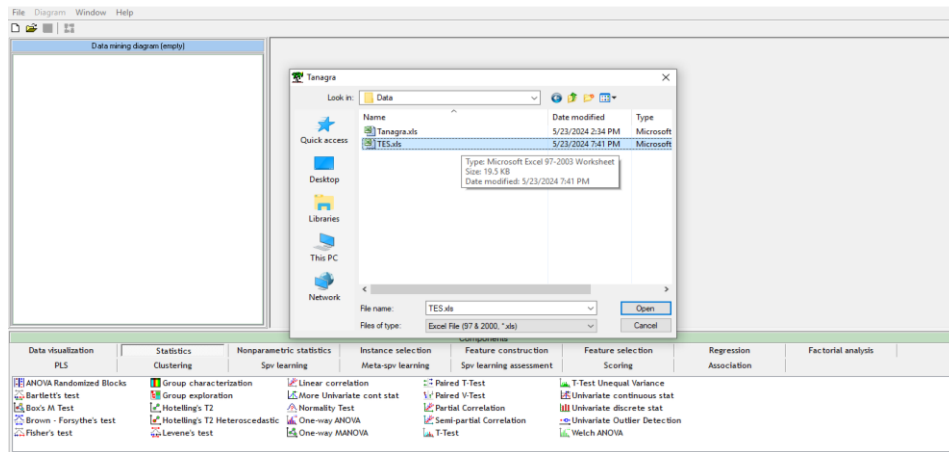
Antoni Praniko
Benny Irwansah
Dani Alfa Rizky
Hengky Febriansyah
Irfansyah
M. Alvin Ramaddan
Marchel Amandika
Mhd. Tri Rizky Farehan
Nikolas Mhs Putra Sembiring
Pramudia Maulana
Rama Kurniawan
Wahyu Saputra

Pengujian Sistem

Pengujian system Analisis Clustering Untuk Menentukan Siswa Berprestasi di SMK TI Panca Dharma Stungkit Menggunakan Metode K-Means menggunakan software Tanagra 1.4.50 dapat dilihat sebagai berikut :

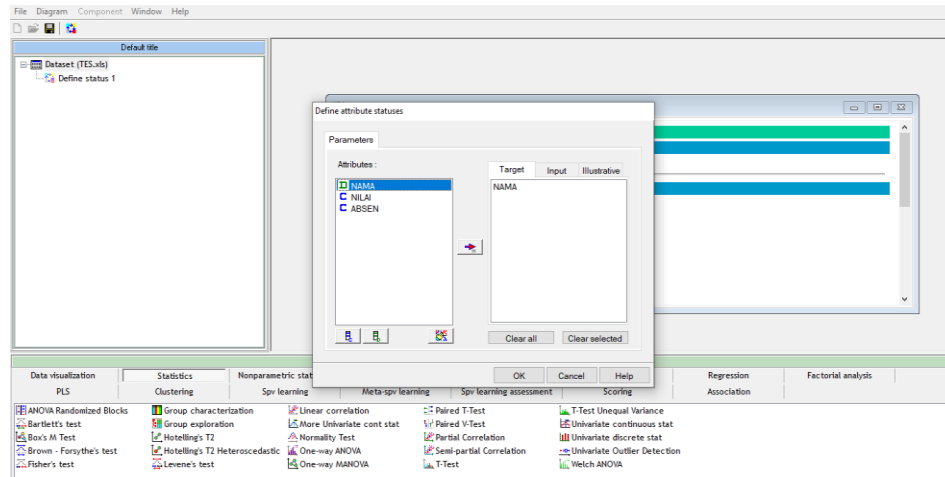
1. Input Dataset

Masukkan data mentah yang sudah disimpan di Microsoft excel adapun tampilannya sebagai berikut :

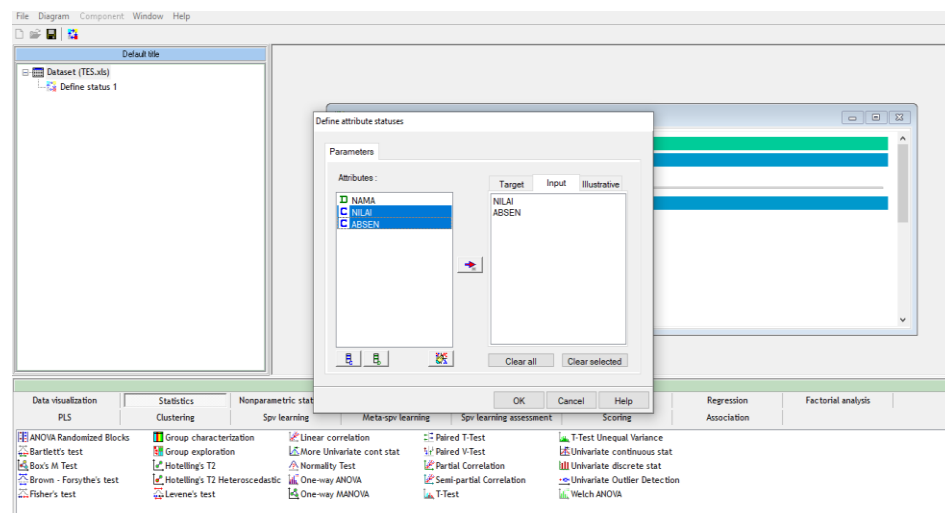


Gambar 3 Tampilan Dataset

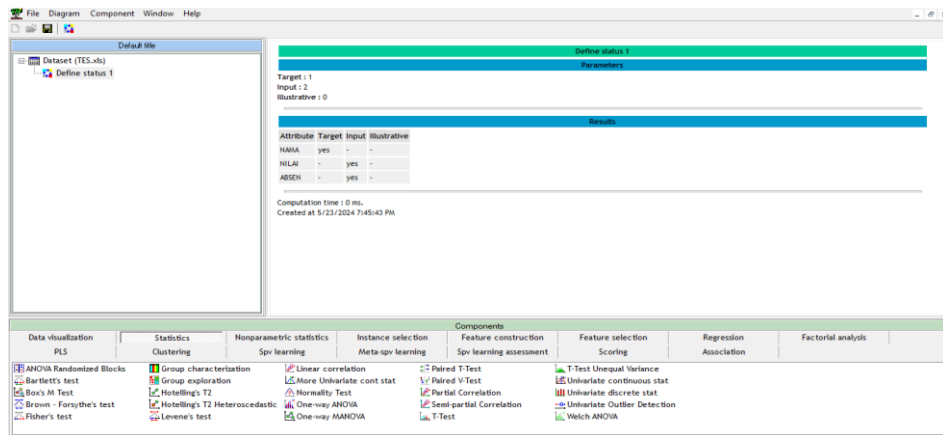
2. Define Status



Gambar 4 Tampilan Define Status Target

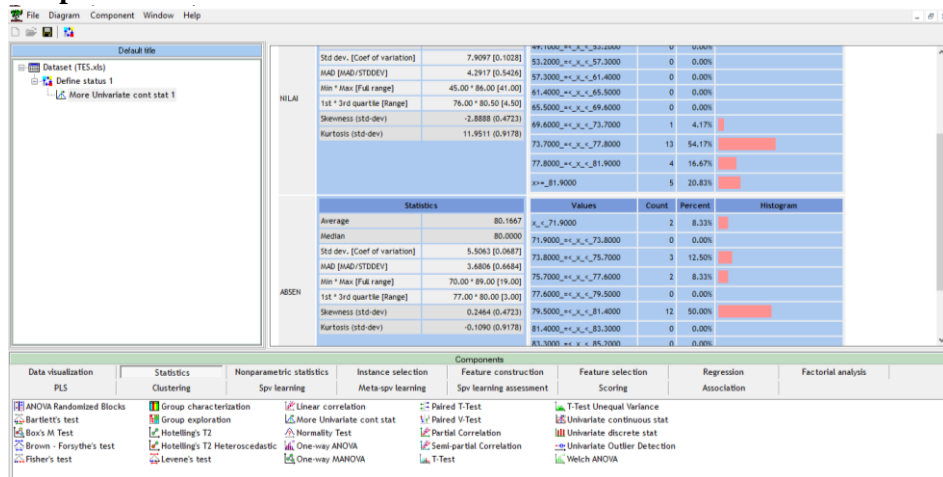


Gambar 5 Tampilan Define Status Input



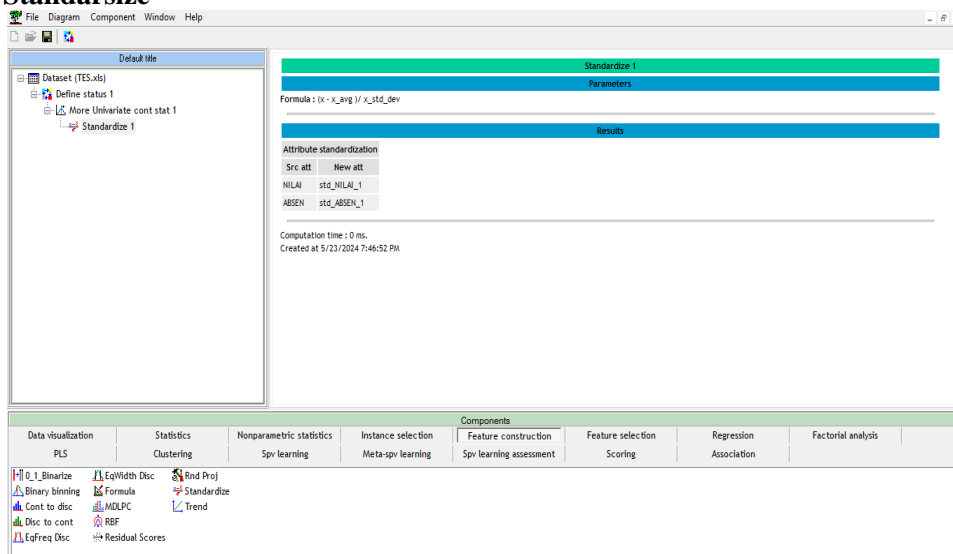
Gambar 6 Tampilan Hasil Define Status 1

3. Komponen More Variant Cont Stat 1



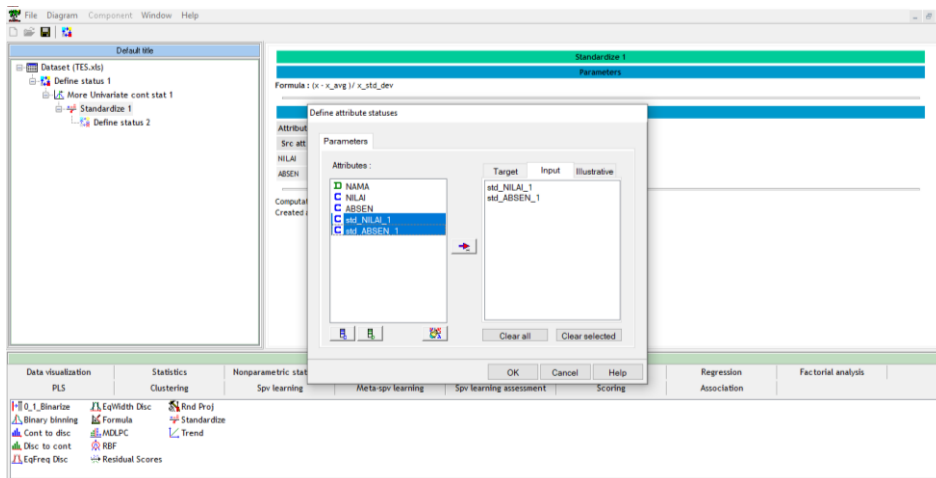
Gambar 7 Tampilan More Variant Cont Stat 1

4. Standarsize

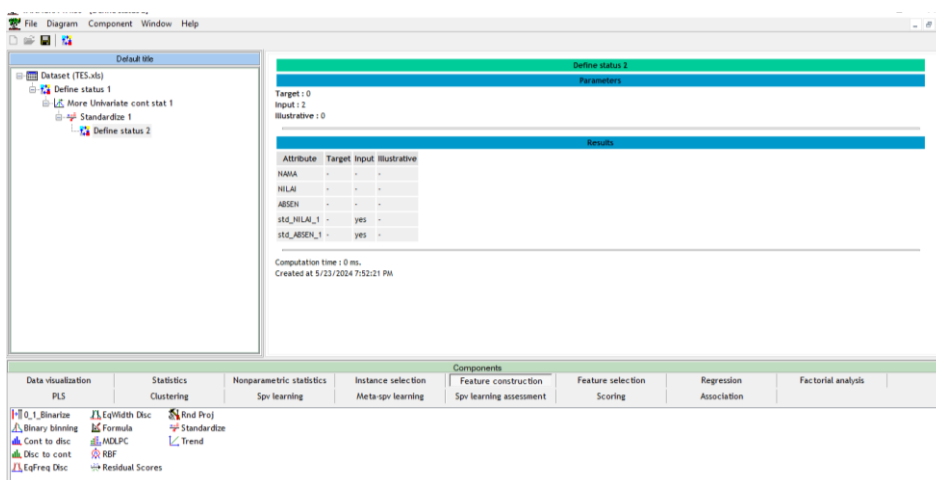


Gambar 8 Tampilan Standarsize

5. Define Status 2



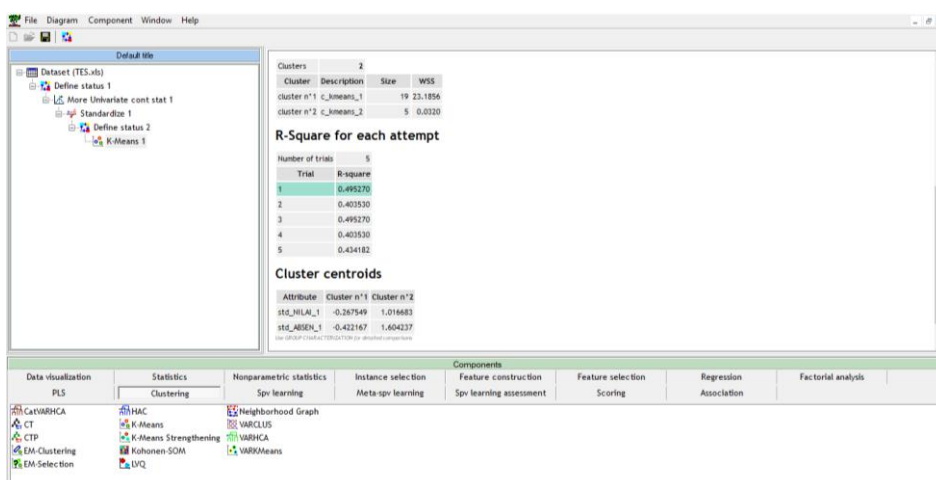
Gambar 9 Tampilan Input Define Status 2



Gambar 10 Tampilan Hasil Define Status 2

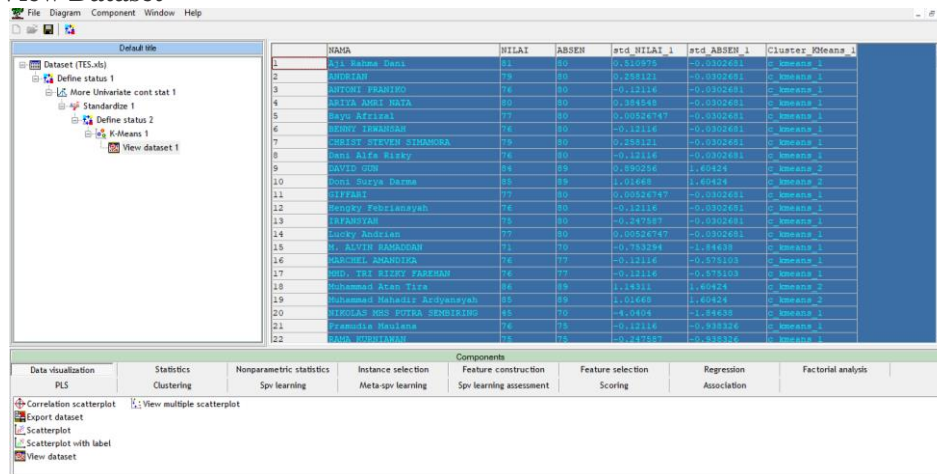
6. Clustering K-Means

K-Means merupakan algoritma clustering yang akan mengelompokkan data-data dan menentukan jumlah cluster yang akan dibentuk sebagai berikut :



Gambar 11 Tampilan Clustering K-Means

7. View Dataset

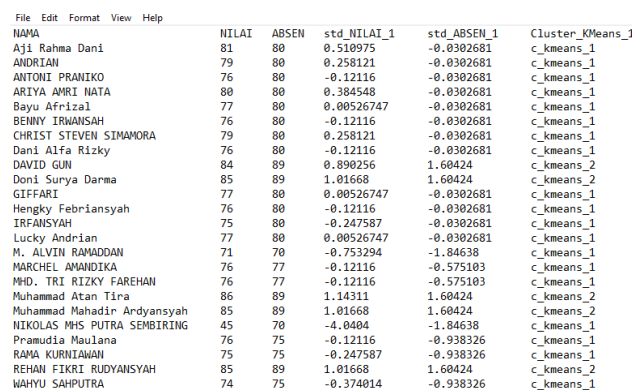


NAMA	NILAI	ABSEN	std_NILAI_1	std_ABSEN_1	Cluster_KMeans_1
Aji Rahma Dani	81	80	0.518975	-0.0302681	c_kmeans_1
ANDRIAN	79	80	0.258121	-0.0302681	c_kmeans_1
ANTONI PRANIKO	76	80	-0.12116	-0.0302681	c_kmeans_1
ARIYA AMRI NATA	80	80	0.384548	-0.0302681	c_kmeans_1
Bayu Afrizal	77	80	0.00526747	-0.0302681	c_kmeans_1
BENNY IRWANSAH	76	80	-0.12116	-0.0302681	c_kmeans_1
CHRIST STEVEN SIMAMORA	79	80	0.258121	-0.0302681	c_kmeans_1
Dani Alfa Rizky	76	80	-0.12116	-0.0302681	c_kmeans_1
DAVID GUN	84	89	0.890256	1.60424	c_kmeans_2
Doni Surya Darma	85	89	1.01668	1.60424	c_kmeans_2
GIFFARI	77	80	0.00526747	-0.0302681	c_kmeans_1
Hengky Febriansyah	76	80	-0.12116	-0.0302681	c_kmeans_1
IRFANSYAH	75	80	-0.247587	-0.0302681	c_kmeans_1
Lucky Andrian	77	80	0.00526747	-0.0302681	c_kmeans_1
M. ALVIN RAMADAN	71	70	-0.753294	-1.84638	c_kmeans_1
MARCHEL AMANDIKA	76	77	-0.12116	-0.575103	c_kmeans_1
MHD. TRI RIZKY FAREHAN	76	77	-0.12116	-0.575103	c_kmeans_1
Muhammad Atan Tira	86	89	1.14311	1.60424	c_kmeans_2
Muhammad Mahadir Ardyansyah	85	89	1.01668	1.60424	c_kmeans_2
NIKOLAS MHS PUTRA SEMBIRING	45	70	-4.0404	-1.84638	c_kmeans_1
Pramudia Maulana	76	75	-0.12116	-0.938326	c_kmeans_1
RAPA KURNIAWAN	75	75	-0.247587	-0.938326	c_kmeans_1
REHAN FIKRI RUDYANSYAH	85	89	1.01668	1.60424	c_kmeans_2
WAHYU SAHPUTRA	74	75	-0.374014	-0.938326	c_kmeans_1

Gambar 12 View Dataset

8. Exsprot Dataset

Merupakan hasil dari data-data yang sudah dianalisis dengan metode K-Means menggunakan software Tanagra sebagai berikut :



NAMA	NILAI	ABSEN	std_NILAI_1	std_ABSEN_1	Cluster_KMeans_1
Aji Rahma Dani	81	80	0.518975	-0.0302681	c_kmeans_1
ANDRIAN	79	80	0.258121	-0.0302681	c_kmeans_1
ANTONI PRANIKO	76	80	-0.12116	-0.0302681	c_kmeans_1
ARIYA AMRI NATA	80	80	0.384548	-0.0302681	c_kmeans_1
Bayu Afrizal	77	80	0.00526747	-0.0302681	c_kmeans_1
BENNY IRWANSAH	76	80	-0.12116	-0.0302681	c_kmeans_1
CHRIST STEVEN SIMAMORA	79	80	0.258121	-0.0302681	c_kmeans_1
Dani Alfa Rizky	76	80	-0.12116	-0.0302681	c_kmeans_1
DAVID GUN	84	89	0.890256	1.60424	c_kmeans_2
Doni Surya Darma	85	89	1.01668	1.60424	c_kmeans_2
GIFFARI	77	80	0.00526747	-0.0302681	c_kmeans_1
Hengky Febriansyah	76	80	-0.12116	-0.0302681	c_kmeans_1
IRFANSYAH	75	80	-0.247587	-0.0302681	c_kmeans_1
Lucky Andrian	77	80	0.00526747	-0.0302681	c_kmeans_1
M. ALVIN RAMADAN	71	70	-0.753294	-1.84638	c_kmeans_1
MARCHEL AMANDIKA	76	77	-0.12116	-0.575103	c_kmeans_1
MHD. TRI RIZKY FAREHAN	76	77	-0.12116	-0.575103	c_kmeans_1
Muhammad Atan Tira	86	89	1.14311	1.60424	c_kmeans_2
Muhammad Mahadir Ardyansyah	85	89	1.01668	1.60424	c_kmeans_2
NIKOLAS MHS PUTRA SEMBIRING	45	70	-4.0404	-1.84638	c_kmeans_1
Pramudia Maulana	76	75	-0.12116	-0.938326	c_kmeans_1
RAPA KURNIAWAN	75	75	-0.247587	-0.938326	c_kmeans_1
REHAN FIKRI RUDYANSYAH	85	89	1.01668	1.60424	c_kmeans_2
WAHYU SAHPUTRA	74	75	-0.374014	-0.938326	c_kmeans_1

Gambar 13 Tampilan Exsprot Dataset

KESIMPULAN

Hasil penelitian Analisis Clustering untuk Menentukan Siswa Berprestasi di SMK Swasta TI Panca Dharma Stungkit Menggunakan Metode K-Means yaitu

1. Dengan menggunakan metode K-Means dapat mengetahui pengelompokkan siswa/i yang berprestasi dan tidak berprestasi.
2. Dengan 24 sampel data siswa/i dapat membentuk 2 cluster, yaitu berprestasi dan tidak berprestasi. Terdapat 12 orang siswa/i masuk ke dalam cluster berprestasi dan 12 orang siswa/i masuk ke dalam cluster tidak berprestasi.

REFERENSI

- Adrianto, S., Khasanah, N., & Wahyuni, D. (2020). Implementasi Data Mining pada Penjualan Kartu Perdana Internet di Purnama Ponsel Menggunakan Metode Algoritma Apriori. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 5(2), 81–95. <https://doi.org/10.14421/jiska.2020.52-03>
- Ani Arnomo, S. (2022). Penerapan Data Mining Metode Apriori Terhadap Data Penjualan Pada Pt Tasindo Total Inproducts. *Jurnal Comasie*, 03.

- Aritonang, J., & Saragih, S. P. (2017). Implementasi Data Mining pada Penjualan Menggunakan Algoritma Apriori di PT. Selatan Indobatam Mandiri. *Comasie*, 2(2), 31–39.
- Dewi, F. P., Aryni, P. S., & Umaidah, Y. (2022). Implementasi Algoritma K-Means Clustering Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran. *JISKA (Jurnal Informatika Sunan Kalijaga)*, 7(2), 111–121. <https://doi.org/10.14421/jiska.2022.7.2.111-121>
- Erwansyah, K. (2019). *Implementasi Data Mining Untuk Menganalisa Hubungan Data Penjualan Produk Bahan Kimia Terhadap Persediaan Stok Barang Menggunakan Algoritma FP (Frequent Pattern) Growth Pada PT . Grand Multi Chemicals*. 2(2), 30–40.
- Ikhwan, A., & Aslami, N. (2020). Implementasi Data Mining untuk Manajemen Bantuan Sosial Menggunakan Algoritma K-Means. *Jurnal Teknologi Informasi*, 4(2), 208–217. <https://doi.org/10.36294/jurti.v4i2.2103>
- Lowensky, E., & Elisa, E. (2023). Analisa Pola Pembelian Konsumen Menggunakan Data Mining Dengan Algoritma Apriori (Studi Kasus: Edukits Batam Centre). *Computer and Science Industrial Engineering (COMASIE)*, 9(2). <https://doi.org/10.33884/comasiejournal.v9i2.7608>
- Muliono, R., & Sembiring, Z. (2019). Data Mining Clustering Menggunakan Algoritma K-Means Untuk Klasterisasi Tingkat Tridarma Pengajaran Dosen. *CESS (Journal of Computer Engineering, System and Science)*, 4(2), 2502–2714.
- Nabila, Z., Isnain, A. R., Permata, & Abidin, Z. (2021). Analisis Data Mining Untuk Clustering Kasus Covid-19 Di Provinsi Lampung Dengan Algoritma K-Means. *Jurnal Teknologi Dan Sistem Informasi (JTISI)*, 2(2), 100. <http://jim.teknokrat.ac.id/index.php/JTISI>
- Reisyer, K. E., & Harman, R. (2023). Implementasi Algoritma Apriori Dalam Pemodelan Pola Pembelian Produk Pada Edukits Btc. *Computer and Science Industrial Engineering (COMASIE)*, 9(3). <https://doi.org/10.33884/comasiejournal.v9i3.7656>
- Sari, R. M., Wahyuni, S., & Rizka, A. (2023a). Implementasi Algoritma K-Means Untuk Guru Berprestasi. *Jurnal Minfo Polgan*, 12(1), 531–538. <https://doi.org/10.33395/jmp.v12i1.12453>
- Sari, R. M., Wahyuni, S., & Rizka, A. (2023b). Pengklusteran Guru Berprestasi Menggunakan Algoritma K-Means. *Escaf*, 1087–1094. <https://semnas.univbinainsan.ac.id/index.php/escaf/article/view/424%0Ahttps://semnas.univbinainsan.ac.id/index.php/escaf/article/download/424/243>
- Sibuea, M. L., & Safta, A. (2017). Pemetaan Siswa Berprestasi Menggunakan Metode K-Means Clustering. *Jurteks*, 4(1), 85–92. <https://doi.org/10.33330/jurteks.v4i1.28>
- Siregar, J. A. S., & Handoko, K. (2021). Jurnal Comasie Jurnal Comasie. *Jurnal Comasie*, 6(2), 40–51. [http://ejournal.upbatam.ac.id/index.php/comasiejournal%0AJurnal Comasie ISSN \(Online\) 2715-6265%0APERANCANGAN](http://ejournal.upbatam.ac.id/index.php/comasiejournal%0AJurnal%20Comasie%20ISSN%202715-6265%0APERANCANGAN)
- Sovia, R., Mandala, E. P. W., & Mardhiah, S. (2020). Algoritma K-Means dalam Pemilihan Siswa Berprestasi dan Metode SAW untuk Prediksi Penerima Beasiswa Berprestasi. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 6(2), 181. <https://doi.org/10.26418/jp.v6i2.37759>
- Suhartini, S., & Yuliani, R. (2021). Penerapan Data Mining untuk Mengcluster Data Penduduk Miskin Menggunakan Algoritma K-Means di Dusun Bagik Endep Sukamulia Timur. *Infotek: Jurnal Informatika Dan Teknologi*, 4(1), 39–50.

<https://doi.org/10.29408/jit.v4i1.2986>

Widaningsih, S. (2019). Perbandingan Metode Data Mining Untuk Prediksi Nilai Dan Waktu Kelulusan Mahasiswa Prodi Teknik Informatika Dengan Algoritma C4,5, Naïve Bayes, Knn Dan Svm. *Jurnal Tekno Insentif*, 13(1), 16–25.
<https://doi.org/10.36787/jti.v13i1.78>