



IMPLEMENTASI ALGORITMA *K-MEANS CLUSTERING* PENDETEKSIAN DINI PERFORMA SISWA PADA PEMBELAJARAN BAHASA INDONESIA

Implementation of the k-means algorithm for early detection of student performance in Indonesian language learning.

Alsiningssi Loda Wahap¹, Arini Aha Pekuwali², Reynaldi Thimotius Abineno³

Program Studi Teknik Informatika, Fakultas Sains Dan Teknologi, Universitas Kristen Wira Wacana Sumba

Jl.R.Suprato No.35,Prailiu,Kec.Kota Waingapu,Kabupaten Sumba Timur,Nusa Tenggara Timur

Email : arini.pekuwali@unkriswina.ac.id¹, alsiningsiwahap@gmail.com²,
reynaldi@unkriswina.ac.id³

Corresponding author: alsiningsiwahap@gmail.com

ABSTRACT

Indonesian language subjects have a crucial role in the development of students in Indonesia. The general reasons for selecting Indonesian language subjects in this research involve aspects that focus on students' abilities based on the results of knowledge and skills scores. In Indonesian language subjects, with a focus on case studies at SMP Negeri 4 Mauluru. semester 1 and semester 2. Minimum in cluster 0 will get a score of 74 for all semesters. The maximum in cluster 0 is 74 for all semesters. The average value in cluster 0 is 77.69 for semester 1, has 74.48 for semester 2, and 78.00 for semester 2. The standard deviation of values in cluster 0 is 1.551 for semester 1. The data has a value of 1.712 for semester 2, and 2.472 for semester P1, K1, P2 and K2 show certain aggregate values in cluster 0 for each semester. Cluster 1 has higher knowledge and skills scores compared to Cluster 0, with an average score in the range of 60. In classes IXA and IXB at SMP Negeri 4 Mauluru, the school has difficulty determining the grades of students who are included in the class, so it is used data mining techniques using the K-Means Algorithm method which is grouped into several clusters.

Keywords: *k-means, clustering, iterasi, bahasa indonesia.*

ABSTRAK

Bahasa Indonesia sangat penting untuk perkembangan siswa di Indonesia. Penelitian ini memilih mata pelajaran Bahasa Indonesia karena fokusnya pada kemampuan siswa berdasarkan nilai pengetahuan dan keterampilan. Pada mata pelajaran Bahasa Indonesia di SMP Negeri 4 Mauluru. semester 1 dan semester 2. Minimum dalam cluster 0 akan mendapatkan nilai 74 untuk semua semester. Maksimum dalam cluster 0 adalah 74 untuk semua semester. Untuk nilai rata-rata dalam cluster 0 adalah 77.69 untuk semester 1, memiliki 74.48 untuk semester 2, dan 78.00 untuk semester 2. Standar deviasi nilai dalam cluster 0 1.551 untuk semester 1. Memiliki data nilai 1.712 untuk semester 2, dan 2.472 untuk semester P1, K1, P2 dan K2 ini menunjukkan nilai-nilai aggregate tertentu dalam cluster 0 untuk masing-masing semester. Cluster 1 memiliki nilai pengetahuan dan keterampilan yang lebih tinggi dibandingkan dengan Cluster 0, dengan nilai rata-rata di kisaran 60. Pada kelas IXA dan IXB di SMP Negeri 4 Mauluru pihak sekolah kesulitan untuk menentukan nilai siswa-siswi yang dimasukkan ke kelas, maka digunakan teknik dari data mining dengan metode Algoritma K-Means yang dikelompokkan ke beberapa cluster.

Kata kunci : K-Means, clustering, iterative,, bahasa indonesia

PENDAHULUAN

Pendidikan adalah bidang aspek penting dalam pembentukan masa depan anak-anak. Pendidikan adalah keseluruhan pengalaman belajar yang dialami setiap orang sepanjang hidupnya, yang tidak memiliki batas waktu dan berlangsung sepanjang usia, anak-anak, remaja. Untuk memaksimalkan kualitas pendidikan, sangat penting untuk lingkungan diciptakan oleh manusia seperti sekolah mengenali hal-hal yang dapat menyebabkan kegagalan siswa pada mata pelajaran tertentu.

Bahasa Indonesia peran dalam perkembangan siswa di Indonesia, dan kegagalan siswa dalam mata pelajaran ini dapat memiliki konsekuensi jangka panjang. ini menggunakan algoritma clustering K-Means untuk menemukan siswa SMP 4 Mauliru yang berisiko gagal sejak dini dalam bahasa Indonesia. . (Bernissa, 2020).

Pendidikan sangatlah penting dimana informasi yang dihasilkan salah satu bentuk inovasi dalam bidang pendidikan. Untuk membentuk kelompok belajar siswa dalam kelas sesuai dengan kriteria siswa. Sebagai upaya meningkatkan mutu pendidikan melalui layanan pembelajaran dapat dijadikan sebagai acuan bahan pertimbangan bagi pihak sekolah untuk melakukan langkah secara terstruktur untuk rangka meningkatkan kualitas dan Siswa yang mampu menulis dan membaca dengan baik memiliki kemampuan untuk memahami dan menginterpretasikan informasi secara efektif. Kemampuan literasi juga merupakan dasar bagi keberhasilan dalam bidang studi lainnya, karena kemampuan ini mendukung pemahaman mereka tentang materi di berbagai disiplin ilmu dikalkulasi sudah melewati cakupan tingkat akurasi nilai tertinggi dari standar batasan kriteria ketuntasan minimum (KKM) (Sulistiyawati & Supriyanto, 2020)

Literasi pendidikan Bahasa Indonesia berkelanjutan terkait dengan kemampuan bahasa indonesia. Data internal sekolah menunjukkan variasi yang cukup signifikan dalam literasi siswa tentang pelajaran bahasa indonesia tujuan perkembangan mereka. Sementara beberapa siswa menunjukkan kinerja yang luar biasa, yang lain menghadapi kesulitan dalam memahami dan menguasai materi. Sehingga intervensi yang tepat dapat digunakan untuk meningkatkan literasi secara keseluruhan, analisis yang mendalam diperlukan untuk menentukan kelompok siswa berdasarkan kinerja mereka. (Agustin, 2015).

MATERI DAN METODE

Education Data Mining

Analisis data yang sangat besar digunakan untuk mengetahui hubungan antar data dan disajikan dalam format yang mudah dipahami. Sebuah pola dibuat sebagai hasil dari pemrosesan data. Seperti yang disebutkan di atas, tujuan pemrosesan data adalah untuk memperoleh pengetahuan dari data yang sudah ada. Knowledge Discovery in Database Process (KDD) adalah salah satu metode data mining. Prosesnya meliputi pengumpulan data, pemrosesan data untuk memilih yang paling penting, integrasi data, data mining, pembuatan model, dan pengujian. (Bernissa, 2020).

Algoritma K- Means

K-Means adalah teknik data mining yang ada yang menggunakan sebagian dari populasi sebagai titik pusat untuk cluster awal. Dengan menghitung jarak antara setiap objek dan titik pusatnya, dapat menetapkan kelompok dalam satu cluster. Algoritma clustering K-Means dapat membagi data berdasarkan jarak antar kelompok. Algoritma ini menggunakan fungsi pengukuran data dengan karakteristik unik. Untuk menghitung jarak, fungsi geometri digunakan. Data kemudian dimasukkan ke dalam kelompok dengan jarak. Istilah "anomali" digunakan dalam

clustering untuk menggambarkan data yang telah diuji pada clustering tetapi tidak menemukan kelas yang sesuai. (Agustin, 2015)

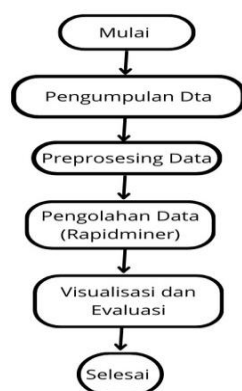
Davies Bouldin Index

Setelah proses klasifikasi selesai, Davis Bouldin Index (DBI) digunakan untuk menentukan jumlah optimal. Tujuan pendekatan DBI ini adalah untuk mengurangi tumpang tindih antara objek dalam satu kelas dan antara satu kelas dengan kelas lainnya. Nilai DBI yang diperoleh (non-negatif;0) berkorelasi positif dengan kualitas klasifikasi yang dihasilkan dari penerapan mode kelompok. Pada akhirnya, sebuah cluster akan memiliki skema clustering terbaik dengan DBI minimum. menegaskan bahwa Davies Bouldin Index digunakan untuk menilai kualitas setiap cluster dalam pengelompokan data. Indeks mempertimbangkan seberapa baik setiap objek dikelompokkan ke dalam cluster dengan menggabungkan dua ide, cohesion (kelompokan) dan soparation (permasalahan). Cohesion mengukur seberapa dekat setiap data dengan satu sama lain Data kemudian dimasukkan ke dalam kelompok dengan jarak terdekat. (Ranjawali et al., 2023).

Rapidminer

Data tentang penerimaan siswa bahasa baru. Rapidminer secara otomatis melakukan analisis statistika setelah import data berhasil. dengan menggunakan algoritma dan prinsip data mining. Dengan menggabungkan metode statistika, kecerdasan buatan, dan database, Rapidminer dapat mengekstrak pola dari dataset yang sangat besar. Dengan menggunakan operator-operator, Rapidminer memudahkan penggunaannya melakukan perhitungan data yang sangat besar. Operator-operator ini memungkinkan data ditambah. Setelah data dihubungkan ke operator, kita hanya perlu menghubungkannya ke node hasil untuk melihat hasilnya. Selain itu, hasil Rapidminer dapat ditampilkan secara visual melalui grafik. Siswa baru yang berhasil menjadi Rapidminer.

Berikut merupakan kerangka berpikir yang akan menggambarkan tahapan-tahapan yang akan dilakukan.



Gambar 1. Alur Penelitian

Pada tahap pengumpulan data adalah mencakup informasi dari pihak sekolah yang berkaitan dengan data-data yang diperlukan seperti data siswa, nilai siswa pada semester 1 tahun ajaran 2023/2024 menggunakan metode kualitatif, yang akan turun langsung dilapangan untuk mendapatkan data-data tersebut.

Pada tahapan ini akan dilakukan normalisasi data dan penelitian atribut yang akan digunakan pada penelitian ini proses normalisasi dilakukan untuk memastikan setiap variabel memiliki

rentang nilai yang sama satu dengan yang lain. Seperti jenis kelamin 0,1 Atribut yang semester 1 dan 2 nilai bahasa indonesia pada nilai pengetahuan dan keterampilan dan jenis kelamin pada SMP Negeri 4 Mauliru (5 variabel) Keterampilan semester 1, pengetahuan semester 2, keterampilan 1 dan pengetahuan semester 2 dan sama dengan ujian sekolah. Terdapat 5 variabel yang dibutuhkan dalam penelitian.

Pada tahapan ini, praprosesing selesai dilakukan pengelompokkan dataset yang akan. Dataset akan digunakan sebagai berikut.

Pada tahapan ini, dataset yang dihasilkan akan digunakan untuk melakukan clustering dengan menggunakan Rapidminer sehingga menghasilkan data pengelompokkan atau yang akan di analisis.

Pada tahapan ini, hasil tersebut akan di analisis dengan tujuan untuk mencari pola, data, hubungan antara variabel, karakteristik pada setiap cluster hingga pengetahuan/ wawasan yang akan digunakan untuk mengukur tingkat literasi bahasa indonesia pada penelitian ini hasil menggunakan tabel perbandingan nilai atau agregat value untuk mengukur nilai mean max, dan nilai rata-rata dari setiap variabel pada masing-masing cluster. Setelah dianalisis hasil analisis akan di visualisasikan sehingga mempermudah proses interpretasi hasil.

Pada tahapan ini akan dilakukan evaluasi metode clustering menggunakan metode Davis Bolding Indeks untuk mengukur seberapa baik akurasi dan model yang digunakan berdasarkan.

HASIL DAN PEMBAHASAN

Pengumpulan Data

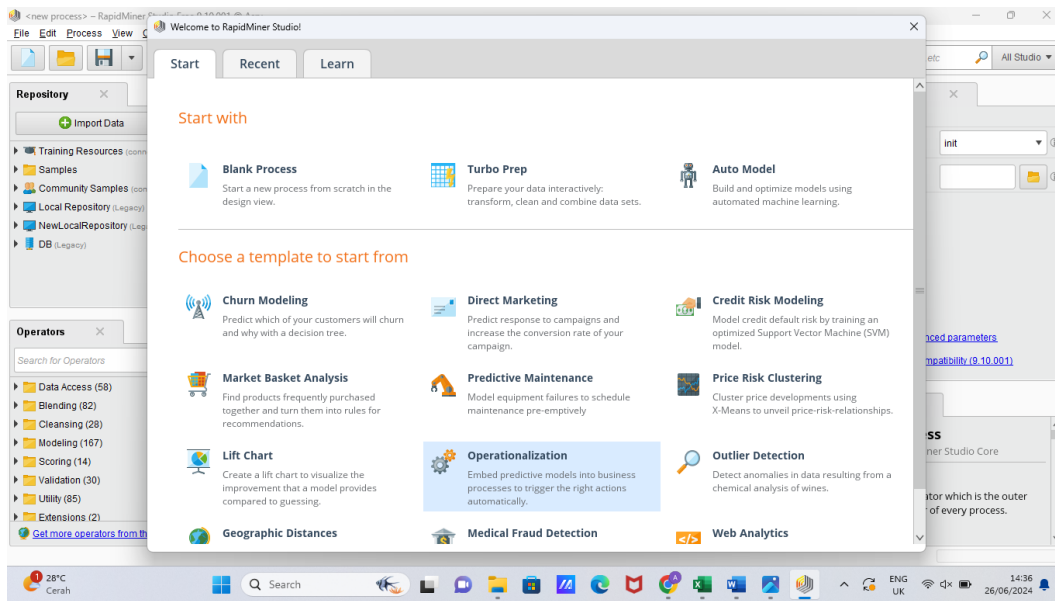
Data yang dibutuhkan untuk penelitian adalah data siswa. Untuk data siswa SMP menggunakan data siswa tahun masuk ajaran 2021/2020 selama 1 dan 2 semester dengan jumlah data sebanyak 60. Data siswa ini terdiri dari data dasar dan data akademik, Dimana kedua jenis yaitu data pengetahuan dan keterampilan ini akan menjadi fitur untuk proses pengklasifian kinerja akademik data.

Proseprocessing Data

Dataset dibersihkan selama tahap preprocessing data. Dataset nilai siswa dari *Exsel* diubah menjadi CSV yang terstruktur dengan kolom-kolom yang jelas. Setiap kolom menunjukkan nilai performa siswa yang berfungsi sebagai semester dan jenis keterampilan. Misalnya, "Pengetahuan dan Keterampilan Semester 1" menunjukkan nilai pengetahuan dan "Keterampilan 1" menunjukkan nilai keterampilan. Dengan asumsi bahwa jumlah dataset sebanyak n data dan jumlah inisialisasi centroid $k = 3$ sesuai dengan penelitian, algoritma *K-Means* digunakan. Data penelitian terdiri dari 60 data siswa yang diterapkan kedalam algoritma *K-Means*. Eksperimen dilakukan dengan parameter yang diberikan di bawah ini. Selain itu, data nominal diubah menjadi angka dalam beberapa langkah agar dapat diproses.

Tampilan Utama *RapidMiner*

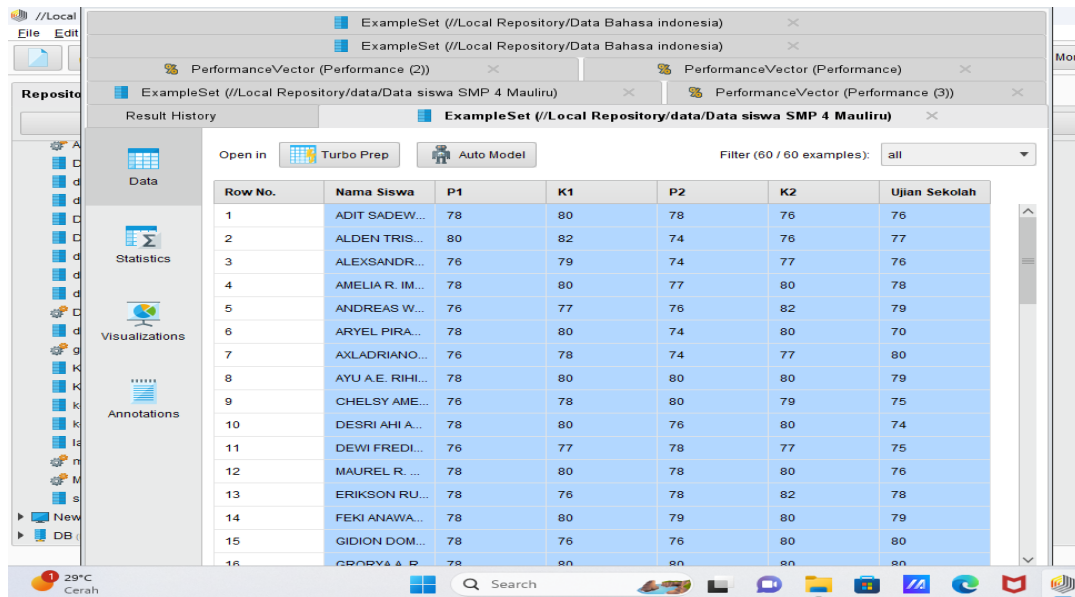
Menu awal *rapidminer* memiliki tampilan awal yang mencakup berbagai komponen dan fitur. Salah satunya adalah "New Process", yang merupakan langkah pertama dalam menggunakan aplikasi *rapidminer*.



Gambar 2. Tampilan Awal Aplikasi *Rapidminer*

Tampilan *Select The Cell To Import*

Sistem memberikan instruksi tentang bagaimana cara untuk menimport dalam data baru yang akan dilakukan dan diolah selanjutnya, dengan format data saat ini dalam bentuk *Exel*. Pada tahapan ini akan terlihat dalam gambar diberikut ini.

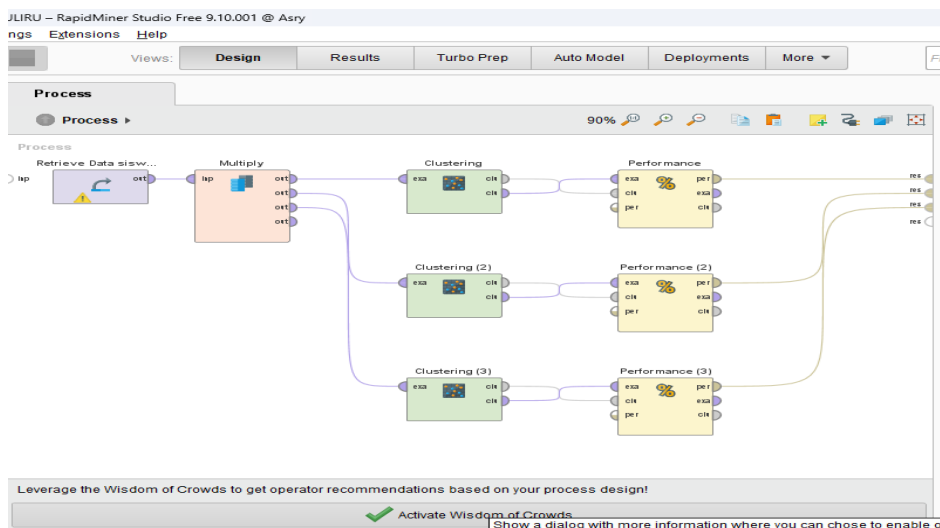


Gambar 3. Tampilan *Select The Cells To Import*

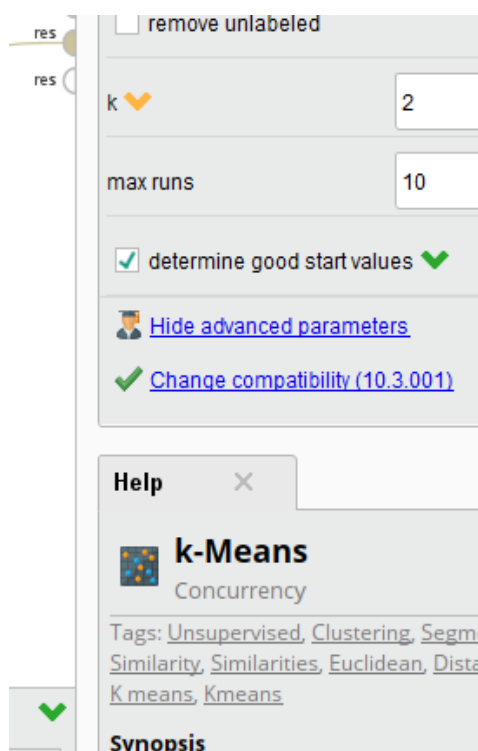
Tampilan *Clustering Rapidminer*

Algoritma *K-Means* akan diterapkan dalam proses yang telah dibentuk untuk akan dilakukan klusterisasi atau pebgelompokkan. Untuk operator klusterisasi yang digunakan adalah *K-Means*.

Setelah itu akan pembuatan model selesai, proses akan dijalankan untuk mendapatkan akan menghasilkan klasterisasi. Gambar 4.3 akan menunjukkan model proses yang menggunakan algoritma *K-Means*.



Gambar 4. Tampilan *Clustering Rapidminer*



Gambar 5. Parameter Jumlah *Cluster*

Berdasarkan tampilan gambar diatas 4.4, diperlukan untuk memasukkan nilai K yang diinginkan. Pada iterasi yang dilakukan sebanyak 10 kali, nilai K yang digunakan adalah k=2, 3, dan 4. Tujuan dari iterasi ini adalah untuk menemukan *cluster* dengan nilai DBI yang mendekati 0 atau yang paling rendah, sehingga dianggap sebagai cluster terbaik. Setelah iterasi selesai, proses

klasterisasi dilakukan untuk mendapatkan hasil dari algoritma *K-Means* yang diterapkan menggunakan *RapidMiner*. Berdasarkan literasi tersebut, ditemukan bahwa nilai K terbaik adalah k=3, dan hasil klasterisasi diperoleh dengan menggunakan algoritma *K-Means*.

Interpretasi Hasil

Dengan menggunakan analisis *cluster k-means*, kelompok-kelompok akan dibuat berdasarkan variabel-variabel yang terkait dengan data siswa, yaitu nilai-nilai pengetahuan dan keterampilan. Data siswa SMP Negeri 4 Mauluru dapat diidentifikasi dengan bantuan analisis kelompok. atau masalah dengan data siswa. Dengan mengelompokkan data siswa berdasarkan kelompok, ada kemungkinan bahwa data siswa akan diproses dengan cara yang tidak seimbang. Data dikelompokkan menjadi dua *cluster*. Cluster 0 dan Cluster 1 setelah proses klasterisasi. Setelah klasterisasi selesai, hasil *cluster* masing-masing siswa akan ditampilkan dalam bentuk tabel bawah ini.

Tabel 1 Anggota Hasil Cluster

No	Nama Siswa	Cluster	JK	P1	K1	P2	K2	US
1	A***	Cluster_0	0	78	80	78	76	76
2	AME***	Cluster_0	1	80	82	74	76	77
3	ALEXSAN***	Cluster_0	1	76	79	74	77	76
4	AME***	Cluster_0	0	78	80	77	80	78
5	ANDR***	Cluster_0	0	76	77	76	82	79
59	AS***	Cluster_1	1	78.0	78.0	75.0	78.0	87.0
60	AST***	Cluster_1	1	78.0	80.0	75.0	80.0	80.0

Tabel 1 menunjukkan hasil perhitungan yang digunakan aplikasi *Rapidminer*, yang akan dibagi menjadi 3 *cluster*. Cluster 0 dan 1 terdiri dari 60 cluster dengan faktor penyebab dataset. Dalam *cluster* ini antara cluster 0,1 dan 2 sehingga dapat nilai. Sedangkan ,cluster 0 terdiri 58 *cluster*.

Aggreget Value Clustering

Setelah melakukan proses pengelompokkan data yang di (*clustering*) berdasarkan tabel di atas, selanjutnya akan ditentukan nilai minimum (*max*), dan rata-rata (mean) dari karakteristik dari masing-masing *cluster*. Berikut adalah untuk mengidentifikasi *Aggregate Value*, dapat dilihat pada tabel 2

Tabel 2. Aggregate Value 0

Aggregate Value	P1	K1	P2	K2	Ujian Sekolah
Min	74	76	74	74	70
Max	82	84	82	82	82
Rata-rata	77,69	79,48	78,00	78,39	77,76
Standar deviasi	1,551	1,712	2,472	1,965	2,264

Tabel 3. Aggregate Value 1

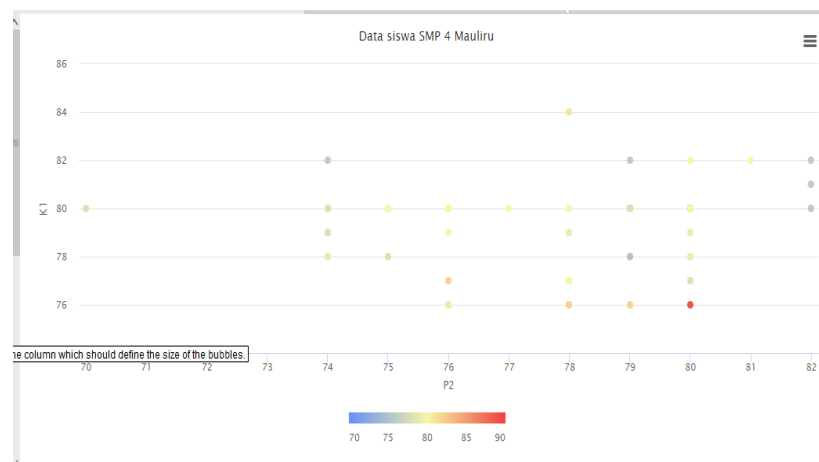
<i>Aggregate Value</i>	P1	K1	P2	K2	Ujian Sekolah
Min	78	76	78	80	80
Max	82	80	80	89	89
rata-rata	79,60	77,60	79,40	84,00	82,60
Standar deviasi	1,673	2,190	0,894	3,807	3,974

Tabel 4. *Aggregate Value 2*

<i>Aggregate Value</i>	P1	K1	P2	K2	Ujian Sekolah
Min	78	80	70	78	97
Max	78	80	70	78	98
rata-rata	78.00	80.00	70.00	78.00	97.00
Standar deviasi	1,673	2,190	0,894	3,807	3,974

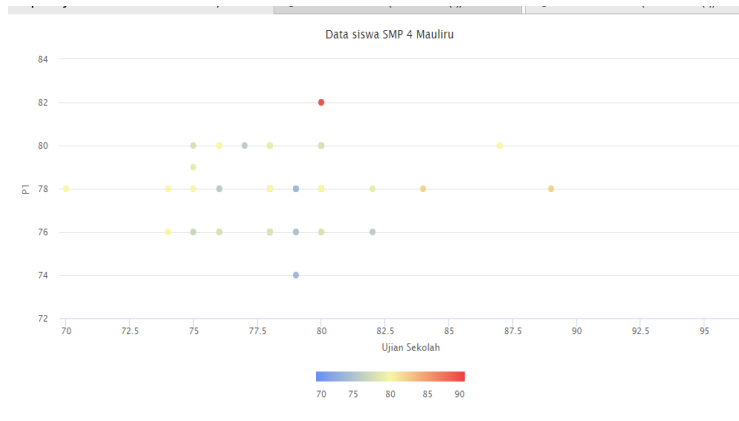
Visualisasi Hasil *Clustering*

Hasil pengujian data menggunakan perangkat lunak *Rapidminer* menghasilkan beberapa *output*. Dalam gambar tersebut visualisasi, terdapat garis yang menunjukkan jumlah kelompok *cluster*. Visualisasi tersebut menampilkan sumbu x keterampilan, sumbu y. pengetahuan.



Gambar 6. *Scatter Plot* Pengatahuan dan Keterampilan

Gambar 6 adalah gambaran *scatter plot* dari pengetahuan dan keterampilan yang berbagai variable x dan keterampilan sebagai variabel sementara sumbu y sebagai variabel, sehingga dapat dilihat bahwa kasus dataset diatas keterampilan berada pada *cluster* 1 yaitu berdasarkan pengetahuan dan keterampilan.

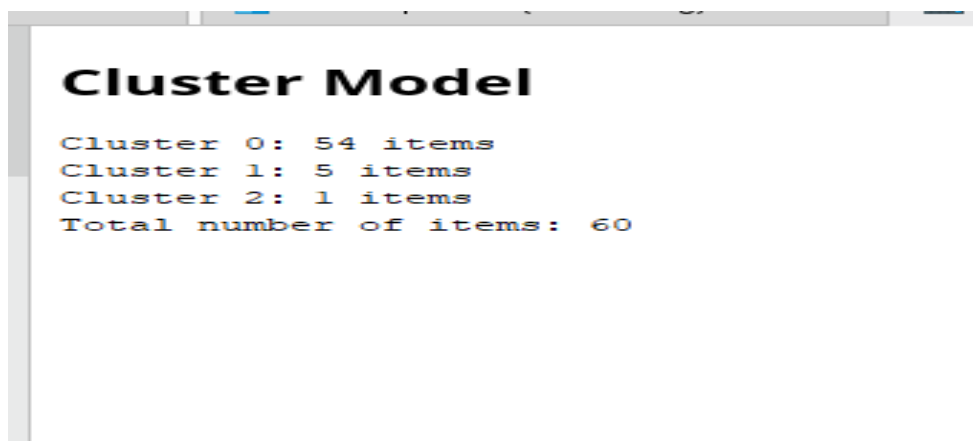


Gambar 7 Scatter Plot Pengetahuan dan Ujian Sekolah

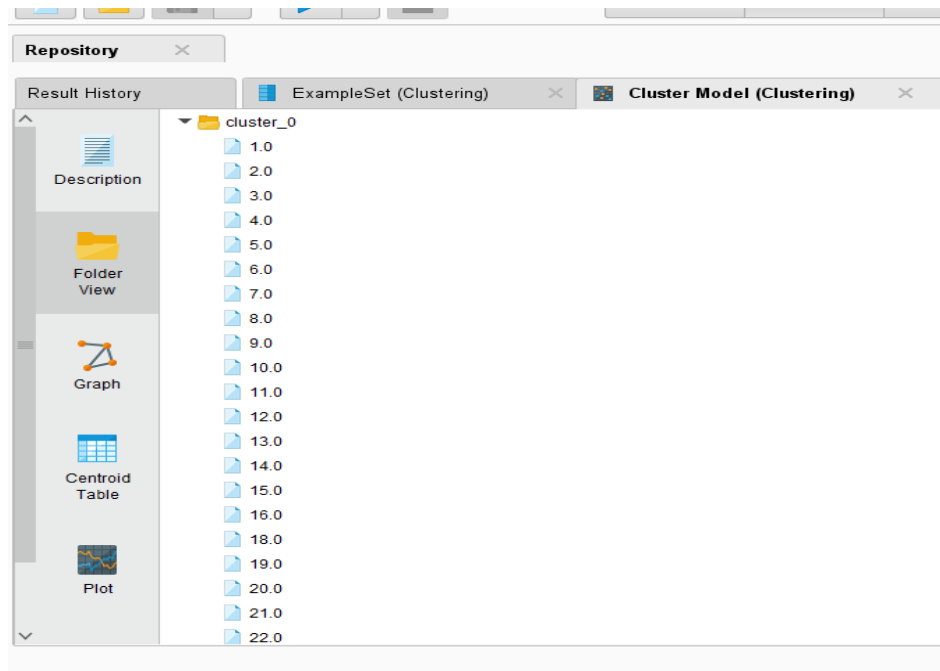
Gambar 7. ditampilkan diatas menunjukkan hasil akhir dari proses *clustering* data menggunakan aplikasi *Rapidminer* yang digunakan untuk melakukan pengelompokan data cluster. Berdasarkan hasil pengujian, dapat disimpulkan bahwa terdapat 60 data dalam cluster. *Cluster 0* ada 54 data yang ditandai dengan warna merah, sementara cluster 1 memiliki 5 data yang akan digunakan *Bar Plot Scatter Plot*, yang serupa dengan dengan contoh yang terlihat pada gambar diatas.

Cluster Model

Dalam model claster terdapat beberapa tampilan yang menampilkan hasil *cluster*, salah satunya tampilan yang menunjukkan hasil penglompokan berdasarkan *cluster* dan jumlah anggota dalam memiliki setiap cluster. Claster 0 memiliki 54 data, sedangkan *cluster 1* memiliki 5 data. Dan *cluster 2* memiliki 1 data. Tampilan tersebut dapat dilihat gambar dibawah ini.



Gambar 8. Cluster Model



Gambar 9. Folder View

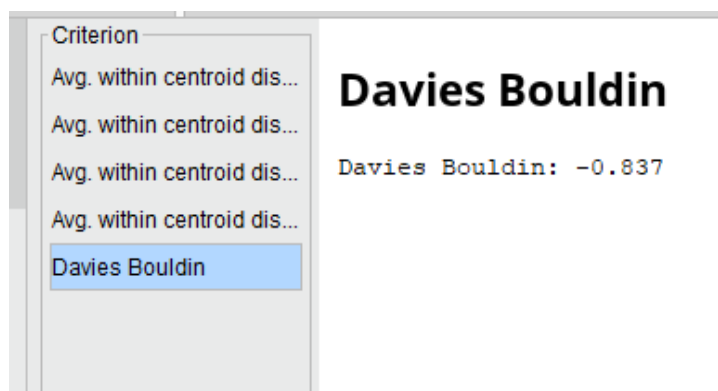
Dari gambar 9 diatas yang menunjukkan data secara komprehensif dari bagian-bagian *cluster*, dengan ini setiap *cluster* anggota dari dua *cluster* menampilkan cluster 0. Setelah menerapkan peggelompokan kasus menggunakan *Rapidminer*.

Evaluasi Hasil *Clustering*

Berberdasarkan tabel 3 dibawah ini, dengan meggunakan metode *k-means clustering* dan didukung oleh cluster perhitungan *Davies-Bouldin Index* (DBI), dilakukan 3 iterasi percobaan untuk menentukan cluster yang paling optimal berdasrkan nilai DBI.

Tabel 4. Nilai *Davies -Bouldin Index* (DBI)

K	Avg Within Centroid Distance	DBI
3	13.206	0.837



Gambar 10. Nilai *Davies Bouldin Index DBI* Ke-3

Gambar 10 akan memberikan informasi tentang nilai *Davies Bouldin Index DBI* dan *performa* klasterisasi informasi ini akan memperoleh pengujian dengan menggunakan perhitungan *Davies - Bouldin Index (DBI)* menggunakan *Rapidminer*. Hasil yang akan dapat dari nilai *Davies -Bouldin Index DBI* ditemukan *K-3 cluster*, dengan nilai DBI sebesar 0.838 yang dapat 0. Ini menunjukkan bahwa. Semakin kecil maka nilai *Davies Bouldin Index (DBI)* yang berhasikan ,maka hasil cluster pengompokan data, dan nilai yang lebih baik dan lebih satu sama lain. Dengan demikian, DBI adalah indicator penting dalam mengevaluasi keberhasilan metode klasterisasi *K-Means* dalam proses pengelompokan data.

KESIMPULAN

Berdasarkan penelitian yang dilakukan di SMP Negeri 4 Mauliru menggunakan metode *clustering K-Means*, dapat disimpulkan bahwa penggunaan data mining efektif dalam mengelompokkan siswa berdasarkan nilai pengetahuan, keterampilan, dan ujian sekolah. Hasil clustering menunjukkan bahwa siswa dapat dibagi ke dalam tiga kelompok berdasarkan nilai yang serupa. Ini membantu sekolah menemukan kebutuhan dan potensi masing-masing siswa. Selain itu, hasil analisis *Davies-Bouldin Index (DBI)* menunjukkan kekuatan cluster dengan nilai *Davies-Bouldin Index DBI* terbaik 0.837 dan dua cluster. Dalam penelitian ini, clustering *K-Means* berhasil memberikan informasi yang mendalam tentang pola prestasi siswa. Sekolah dapat menggunakan informasi ini untuk lebih memahami kinerja akademik siswa dan membuat strategi pembelajaran yang lebih sesuai dan efektif.

SARAN

Untuk memperluas proyek tugas akhir ini, beberapa hal harus dilakukan. Untuk memulai, dapat mempertimbangkan untuk meningkatkan kinerja klasterisasi dengan mengubah atau menggabungkan algoritma *k-means* dengan algoritma lain. Karena metode *Hasil k-means* cenderung *lebih baik dengan* lebih banyak data, dapat disarankan untuk menambah jumlah data yang digunakan. Selanjutnya, penelitian dapat dilakukan secara lebih rinci dan melihat data dari semua sekolah di Sumba Timur, terutama dari SMP Negeri 4 Mauliru. Untuk mengoptimalkan perkembangan, kegagalan awal dapat diidentifikasi untuk siswa di semua mata pelajaran.

DAFTAR PUSTAKA

- Agustin, F. E. M. (2015). IMPLEMENTASI ALGORITMA K-MEANS UNTUK MENENTUKAN KELOMPOK PENGAYAAN MATERI MATA PELAJARAN UJIAN NASIONAL (STUDI KASUS: SMP NEGERI 101 JAKARTA). *JURNAL TEKNIK INFORMATIKA*, 8(1), 73–78. <https://doi.org/10.15408/jti.v8i1.1938>
- Alhapizi, M. R., Nasir, M., & Effendy, I. (2020). *Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Untuk Menentukan Strategi Promosi Mahasiswa Baru Universitas Bina Darma Palembang*. 1(1), 1–14.
- Alifia, A. N., Setiawan, A. F., & Rudhistiar, D. (2024). PENERAPAN ALGORITMA K-MEANS CLUSTERING DALAM PERINGATAN DINI RESIKO KEGAGALAN SISWA PADA MATA PELAJARAN BAHASA INDONESIA (STUDI KASUS : SMP NEGERI 2 MALANG). 8(2), 1174–1181.
- Bernissa, A. B. (2020). Implementasi Algoritma K-Means Untuk Menentukan Kelas Unggulan Pada SMPN 1 Bojong. *POTENSI (EProsiding Sistem Informasi)*, 1(1), 128–135. <http://eprosiding.ars.ac.id/index.php/psi%0Ahttp://eprosiding.ars.ac.id/index.php/psi/article/view/192>
- Nanang, A., Teknik, F., Majalengka, U., Printing, O., Fotocopy, R., Komputer, S., & Sartika, J. D. (n.d.). *PENERAPAN DATA MINING MENGGUNAKAN ALGORITMA K-MEANS UNTUK*. 93–102.
- Pangestu, A. (2021). *ALGORITMA K-MEANS PENGELOMPOKAN PELANGGAN DASARKAN KUBIKASI AIR TERJUAL MELALUI WEKA*.
- Ranjawali, R., Talakua, A. C., & Abineno, R. T. (2023). Clustering Stunting Pada Balita Dengan Metode K- Means Di Puskesmas Kanatang. *SATI: Sustainable Agricultural Technology Innovation*, 80–92. <https://ojs.unkriswina.ac.id/index.php/semnas-FST/article/view/587/324>
- Wijaya, Y. A., Studi, P., Informatika, T., Studi, P., Informasi, S., Studi, P., & Informatika, M. 2022 K-MEANS DI SEKOLAH MENENGAH WAHIDIN DI CIREBON. 6(2), 552–559.11(3), 67–71.
- Sulistiyawati, A., & Supriyanto, E. (2020). *Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan*. 15(2), 25–36.