



ANALISIS DATA MINING MENGGUNAKAN METODE CLUSTERING TERHADAP PRESTASI SISWA I'DADIYAH SUKOREJO

Abdur Rohman Nurut Toyyibin^{a*}, Zaehol Fatah^b

^a Saintek / Sistem Informasi, ab34rohman@gmail.com, Universitas Ibrahimy Sukorejo, Situbondo Jawa Timur

^b Saintek / Sistem Informasi, Zaeholfatah@gmail.com, Universitas Ibrahimy Sukorejo, Situbondo Jawa Timur

* Korespondensi

ABSTRACT

This study analyzes the performance patterns of students at Madrasah I'dadiyah Sukorejo using data mining methods, specifically clustering. The analyzed factors include exam scores and participation in extracurricular activities, as both are considered to significantly influence academic performance. Exam scores reflect mastery of subjects, while extracurricular activities often positively impact students' social skills and learning motivation.[1] The K-Means algorithm was selected to classify students into three main groups: high-performing, average-performing, and low-performing students. The clustering results are expected to provide strategic guidance for the school to improve the quality of education. Low-performing students can receive additional guidance or motivational training, while average-performing students can be encouraged to participate more actively in extracurricular activities to enhance interpersonal skills. Understanding these performance patterns helps the school design more effective programs to maximize students' academic potential based on their needs. This study also opens opportunities for further exploration of other factors affecting academic performance, such as family conditions and the home learning environment. Thus, this approach becomes an essential step in creating a more inclusive and high-quality education system.

Keywords: Data Mining, Clustering K-Means, Student Achievement, Education, Madrasah I'dadiyah.

Abstrak

Penelitian ini menganalisis pola prestasi siswa Madrasah I'dadiyah Sukorejo menggunakan metode data mining, khususnya clustering. Faktor yang dianalisis meliputi nilai ujian dan partisipasi dalam kegiatan ekstrakurikuler, karena keduanya dianggap memiliki pengaruh signifikan terhadap prestasi akademik. Nilai ujian mencerminkan penguasaan materi pelajaran, sedangkan kegiatan ekstrakurikuler seringkali berdampak positif pada keterampilan sosial dan motivasi belajar siswa.[1] Algoritma K-Means dipilih untuk membagi data siswa ke dalam tiga kelompok utama, yaitu siswa berprestasi tinggi, sedang, dan rendah. Hasil clustering ini diharapkan memberikan panduan strategis bagi madrasah dalam meningkatkan kualitas pendidikan. Siswa dengan prestasi rendah dapat diberikan bimbingan tambahan atau pelatihan motivasi, sementara siswa dengan prestasi sedang diarahkan untuk lebih aktif dalam kegiatan ekstrakurikuler guna meningkatkan keterampilan interpersonal. Pemahaman terhadap pola prestasi ini membantu madrasah merancang program yang lebih efektif untuk memaksimalkan potensi akademik siswa sesuai kebutuhan. Penelitian ini juga membuka peluang eksplorasi lebih lanjut terhadap faktor lain yang memengaruhi prestasi, seperti kondisi keluarga dan lingkungan belajar di rumah. Dengan demikian, langkah ini menjadi bagian penting dalam menciptakan pendidikan yang lebih inklusif dan berkualitas.

Kata Kunci: Data Mining, Clustering K-Means, Prestasi Siswa, Pendidikan, Madrasah I'dadiyah..

1. PENDAHULUAN

Prestasi akademik siswa merupakan salah satu indikator penting dalam mengevaluasi kualitas pendidikan di lembaga pendidikan. Di Madrasah I'dadiyah Sukorejo, analisis terhadap faktor-faktor yang mempengaruhi prestasi siswa masih menghadapi tantangan dalam hal efisiensi yang memerlukan waktu dan tenaga yang cukup besar. Salah satu cara untuk mempercepat proses analisis ini adalah dengan menggunakan metode data mining, khususnya clustering, untuk mengelompokkan siswa berdasarkan prestasinya.[2] Dalam penelitian ini, pengelompokan dilakukan berdasarkan nilai siswa sebagai indikator utama prestasi akademik, tanpa memperhitungkan faktor kehadiran. Pendekatan ini diharapkan dapat membantu pihak sekolah memahami karakteristik masing-masing kelompok siswa berdasarkan nilai akademik mereka dan memungkinkan pengambilan keputusan yang lebih tepat terkait program pembinaan atau dukungan tambahan yang dibutuhkan oleh siswa dalam setiap kelompok.

Dengan penerapan metode clustering ini, pihak madrasah diharapkan dapat mengidentifikasi kelompok siswa yang memiliki pola prestasi serupa. Misalnya, siswa dengan nilai akademik tinggi, sedang, dan rendah dapat dikelompokkan ke dalam tiga kategori yang berbeda.[3] Hasil dari clustering ini akan memberikan gambaran mengenai distribusi prestasi siswa, sehingga madrasah dapat merancang strategi yang lebih terarah, seperti program pengayaan untuk siswa dengan nilai tinggi atau program pendampingan khusus untuk siswa yang membutuhkan perhatian lebih.

Selain itu, pendekatan ini diharapkan dapat memberikan insight baru mengenai faktor-faktor yang mempengaruhi perbedaan prestasi di kalangan siswa, serta dapat menjadi acuan bagi pengembangan metode pembelajaran yang lebih efektif. Penelitian ini tidak hanya memberikan manfaat bagi pihak sekolah dalam memahami pola prestasi siswa secara lebih mendalam, tetapi juga berpotensi menjadi referensi bagi penelitian lanjutan di bidang pendidikan. Dengan demikian, penerapan data mining dalam pengelompokan prestasi siswa dapat menjadi salah satu solusi inovatif dalam meningkatkan kualitas pendidikan di Madrasah I'dadiyah Sukorejo.

2. METODOLOGI PENELITIAN

2.1. Jenis Penelitian

Penelitian ini merupakan penelitian kuantitatif dengan pendekatan eksploratif. Penelitian kuantitatif ini bertujuan untuk menggali pola dan karakteristik prestasi siswa melalui analisis data secara sistematis dan objektif menggunakan metode data mining, khususnya clustering. Pendekatan eksploratif digunakan untuk mengeksplorasi dan mengidentifikasi pola-pola prestasi yang belum diketahui sebelumnya berdasarkan data yang tersedia.

2.1.1. Populasi dan Sampel Penelitian

Populasi dalam penelitian ini adalah seluruh siswa di Madrasah I'dadiyah Sukorejo. Sampel penelitian diambil menggunakan teknik purposive sampling, di mana siswa yang memiliki data lengkap mengenai nilai akademik dan partisipasi ekstrakurikuler akan dipilih sebagai sampel. Jumlah sampel disesuaikan dengan ketersediaan data yang valid untuk memastikan hasil analisis yang akurat.[4]

2.1.2. Teknik Pengumpulan Data

Data yang digunakan dalam penelitian ini terdiri dari data nilai akademik dan data partisipasi ekstrakurikuler siswa. Teknik pengumpulan data dilakukan melalui:

a. Dokumentasi

Mengumpulkan data nilai ujian semester siswa sebagai representasi dari prestasi akademik dan data partisipasi ekstrakurikuler yang tercatat di madrasah.

b. Observasi

Melakukan pengamatan terhadap keterlibatan siswa dalam kegiatan ekstrakurikuler untuk memperoleh informasi tambahan tentang tingkat partisipasi mereka.

c. Teknik Analisis Data

Untuk menganalisis data prestasi siswa 'Idadiyah Sukorejo, dilakukan pendekatan melalui metode clustering. Teknik ini bertujuan untuk mengelompokkan siswa berdasarkan pola kesamaan nilai mereka dalam berbagai mata pelajaran dan aspek lainnya yang tercantum dalam dataset.[5]

Berikut adalah langkah-langkah yang diterapkan dalam analisis ini:

1) Preprocessing Data

- a) Data yang diperoleh dari dataset divalidasi untuk memastikan tidak ada nilai yang hilang (missing values) atau outlier yang signifikan.
- b) Normalisasi dilakukan pada nilai-nilai untuk menyamakan skala setiap variabel, sehingga perbedaan skala tidak memengaruhi hasil clustering.

2) Implementasi Algoritma K-Means Clustering

Setelah data siap, langkah berikutnya adalah mengelompokkan data menggunakan algoritma K-Means Clustering. Prosesnya meliputi:

a) Menentukan Jumlah Kluster (K):

Berdasarkan kajian teoretis dan tujuan penelitian, jumlah kluster yang digunakan adalah tiga (3) kluster, yaitu:

- i. Siswa berprestasi tinggi
- ii. Siswa berprestasi sedang
- iii. Siswa berprestasi rendah

b) Pengelompokan Data

K-Means akan membagi data siswa ke dalam tiga kluster berdasarkan jarak antara data siswa dengan centroid dari setiap kluster.

Tabel 1. 1 Clustering Data Nilai Tertinggi

NAMA SISWA	Average
ABDUR ROHMAN	89.85
Ahmad AHSINUR ROFIQ	83.23.00
AHMAD AKMALUDIN	86.00.00
AHMAD ROJIEL GHUFRON	87.69
AHMED YASSER LUMA	81.92
DIMAS HABIBUR ROHMAN	85.00.00
LALU HAMZAH FANZURI HABIBI	84.31.00

Tabel 1. 2 Clustering Data Nilai Sedang

NAMA SISWA	Average
A.MARZUQI	75.31.00
Ahmad SYAUQI JAELANI	79.85
AMANDA RISKY SULTAN	78.46.00
AS'AD AL WI SHIHAB	80.92
AMHA NURUDDIN PUTRA MULYA	78.54.00
LALU MUHAMMAD TOBARI MUKHTAR	80.31.00
M.FAZLUR RAHMAN ASSAUQI	76.38.00

Tabel 1. 3 Clustering Data Nilai Rendah

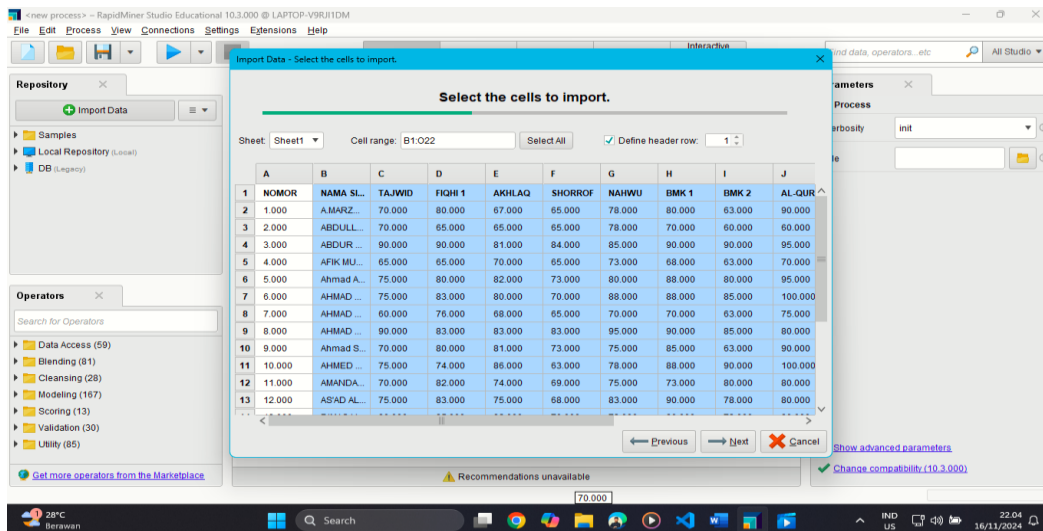
NAMA SISWA	Average
ABDULLAH KHAIRI LADUNI	72.08.00
AFIK MUZAKKI	71.46.00
AHMAD BAKRI	74.46.00
IKHSAN MAULANA	65.15.00
LALU HAIKAL AZMI	72.15.00
LUKMAN HAKIM	71.69
M.MISBAH HUDDIN	70.92

c) Evaluasi Kluster

Menggunakan metode evaluasi, seperti Elbow Method atau Silhouette Score, untuk menilai kualitas kluster yang terbentuk dan memastikan bahwa kluster tersebut telah mengelompokkan siswa secara optimal.[6]

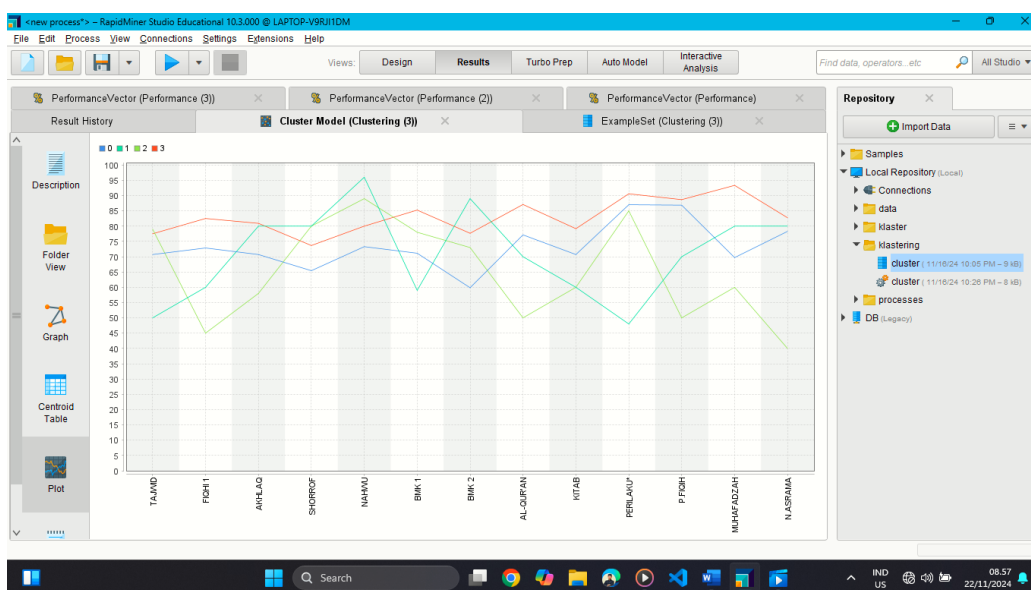
3) Interpretasi Hasil Clustering

Hasil dari proses clustering akan dianalisis dan diinterpretasikan untuk memahami karakteristik masing-masing kelompok. Karakteristik ini akan memberikan gambaran mengenai tingkat prestasi siswa dan memungkinkan pihak madrasah untuk merancang intervensi yang tepat untuk setiap kelompok.



Gambar 1. 1 Data Clustering

Pada Gambar 1. 1 ini menunjukkan antarmuka RapidMiner Studio yang sedang dalam proses mengimpor data dari file spreadsheet. Data mencakup kolom seperti NOMOR, NAMA SL, serta nilai mata pelajaran seperti TAJWID, Fiqih 1, AKHLAQ, SHOROF, NAHWU, BMK 1, BMK 2, dan AL-QUR. Rentang data yang dipilih adalah B1 hingga O22, dengan baris pertama sebagai header. Di sebelah kiri terdapat panel Repository untuk penyimpanan data dan Operators untuk berbagai fungsi analisis. Proses ini merupakan langkah awal untuk mempersiapkan data sebelum analisis lebih lanjut, seperti clustering atau klasifikasi.[7]



Gambar 1. 2 Hasil Data Clustering

Pada Gambar 1. 2 ini menunjukkan hasil analisis clustering menggunakan RapidMiner Studio. Grafik memvisualisasikan tiga cluster (0, 1, dan 2) yang dibedakan dengan garis berwarna (merah, biru, dan hijau).

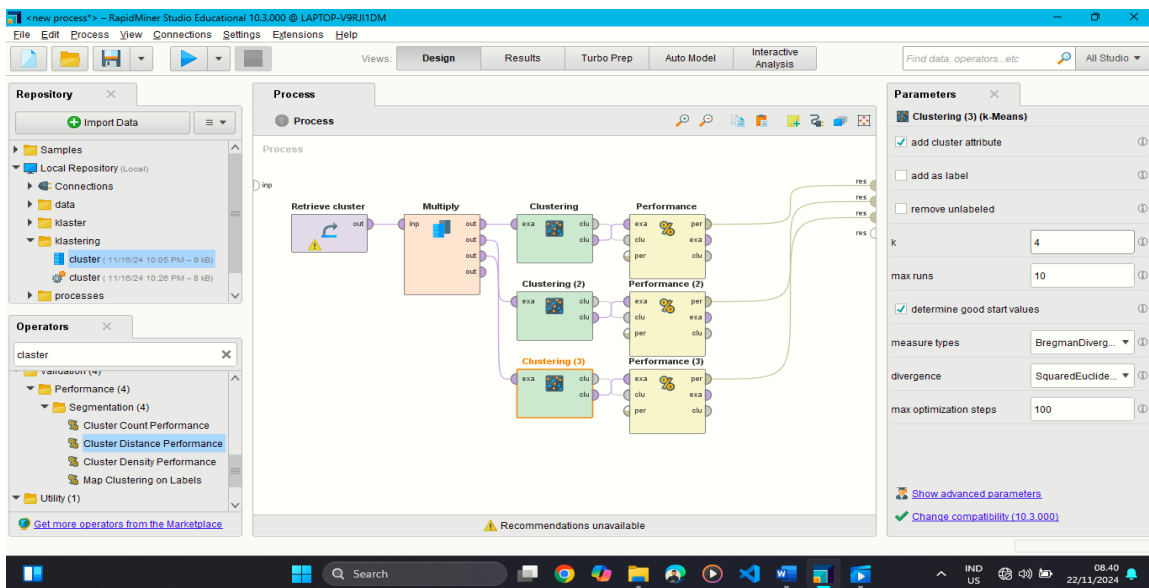
Sumbu horizontal mewakili atribut seperti TAJWID, FIQIH, AKHLAQ, dan lainnya, sedangkan sumbu vertikal menunjukkan nilai rata-rata untuk masing-masing atribut pada setiap cluster. Grafik ini membantu memahami perbedaan pola atau karakteristik antar cluster berdasarkan nilai atribut yang dianalisis. Panel di kanan menunjukkan folder dan proses data yang digunakan. Visualisasi ini berguna untuk mengidentifikasi pola dan membantu pengambilan keputusan.[8]

3. HASIL DAN PEMBAHASAN

Pada bagian ini dijelaskan hasil-hasil penelitian dan sekaligus diberikan pembahasan yang komprehensif. Hasil dapat disajikan dalam bentuk gambar, grafik, tabel dan lain-lain yang memudahkan pembaca. Pembahasan dapat dilakukan dalam beberapa sub-bab.

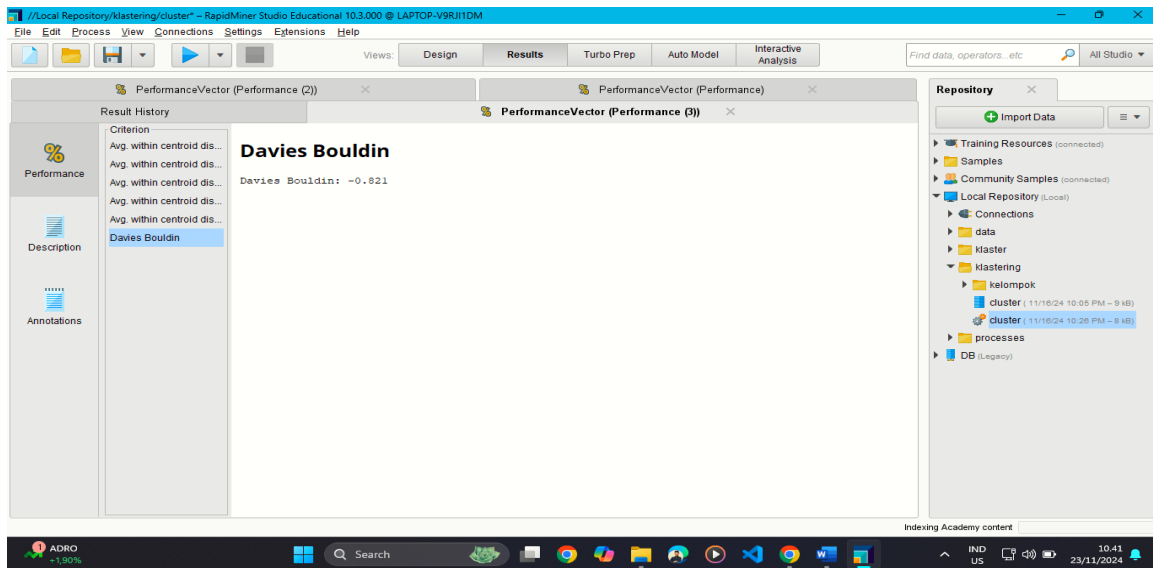
3.1 Kluster 1 (Prestasi Tinggi)

Kelompok siswa yang memiliki nilai akademik tinggi serta partisipasi yang baik dalam kegiatan ekstrakurikuler. Siswa dalam kelompok ini menunjukkan kemampuan akademik yang kuat dan keterlibatan aktif dalam kegiatan di luar kelas, yang sering kali dikaitkan dengan keterampilan sosial dan kepemimpinan yang baik. Berikut clustering data siswa, seperti gambar di bawah ini.



Gambar 1. 3 Proses Cluster Dengan (K) 4

Pada Gambar 1. 3 ini menunjukkan proses clustering menggunakan RapidMiner Studio. Data yang telah diproses sebelumnya diambil dengan operator Retrieve Cluster, lalu dimodifikasi menggunakan operator Multiply. Tiga operator Clustering menerapkan metode K-Means untuk membagi data menjadi beberapa cluster dengan parameter berbeda. Setelah itu, operator Performance digunakan untuk mengevaluasi hasil clustering. Di panel kanan, parameter untuk K-Means diatur, termasuk jumlah cluster (k), iterasi maksimal (max runs), dan pengukuran jarak antar cluster. Proses ini bertujuan untuk melakukan clustering dan mengevaluasi hasilnya.[9]

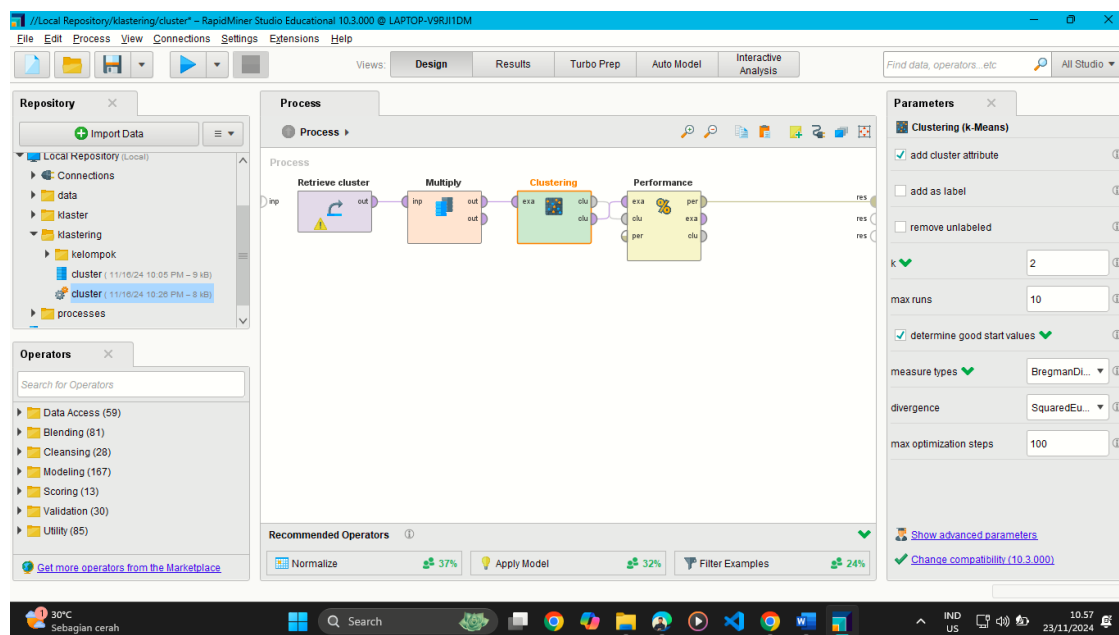


Gambar 1. 4 Nilai Tertinggi (Prestasi Tinggi)

Pada Gambar 1. 4 tersebut menampilkan antarmuka dari perangkat lunak RapidMiner Studio, yang digunakan untuk melakukan analisis data, khususnya evaluasi hasil klusterisasi. Pada panel tengah, ditampilkan hasil evaluasi model klusterisasi menggunakan metrik Davies-Bouldin Index (DBI) dengan nilai sebesar -0.821. Nilai ini menunjukkan tingkat validitas kluster yang dihasilkan, di mana secara umum, DBI yang lebih kecil (mendekati nol) menandakan klusterisasi yang lebih optimal. Gambar ini menggambarkan proses analisis data yang sedang berlangsung, khususnya tahap evaluasi untuk menentukan kualitas klusterisasi dataset.[10]

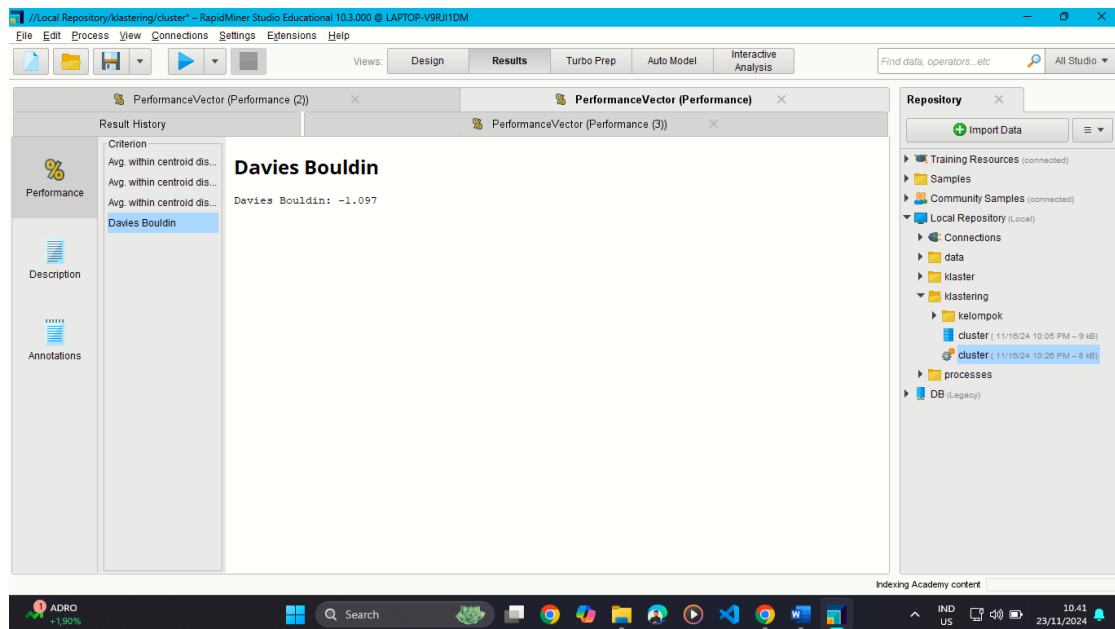
3.2 Kluster 2 (Prestasi Sedang)

Kelompok siswa dengan nilai akademik yang cukup baik, namun partisipasinya dalam kegiatan ekstrakurikuler tidak setinggi kelompok pertama. Siswa dalam kelompok ini umumnya menunjukkan pemahaman yang cukup terhadap materi pembelajaran, tetapi mungkin memerlukan lebih banyak motivasi atau dorongan untuk berpartisipasi dalam kegiatan di luar akademik. Berikut clustering data siswa, seperti gambar di bawah ini.



Gambar 1. 5 Proses Clustering Dengan (K)2

Pada Gambar 1. 5 ini menunjukkan antarmuka aplikasi RapidMiner Studio untuk analisis data menggunakan algoritma K-Means. Panel kiri menampilkan Repository berisi data dan proses, sementara panel tengah menunjukkan alur analisis yang terdiri dari beberapa operator: Retrieve Cluster untuk mengambil data, Multiply untuk menggandakan data, Clustering untuk pengelompokan menggunakan K-Means, dan Performance untuk evaluasi hasil. Panel kanan berisi konfigurasi parameter seperti jumlah kluster (k) sebanyak 2, iterasi maksimum 10, dan langkah optimasi maksimal 100. Gambar ini menggambarkan proses analisis clustering yang menghasilkan pengelompokan data berdasarkan karakteristik tertentu.[11]

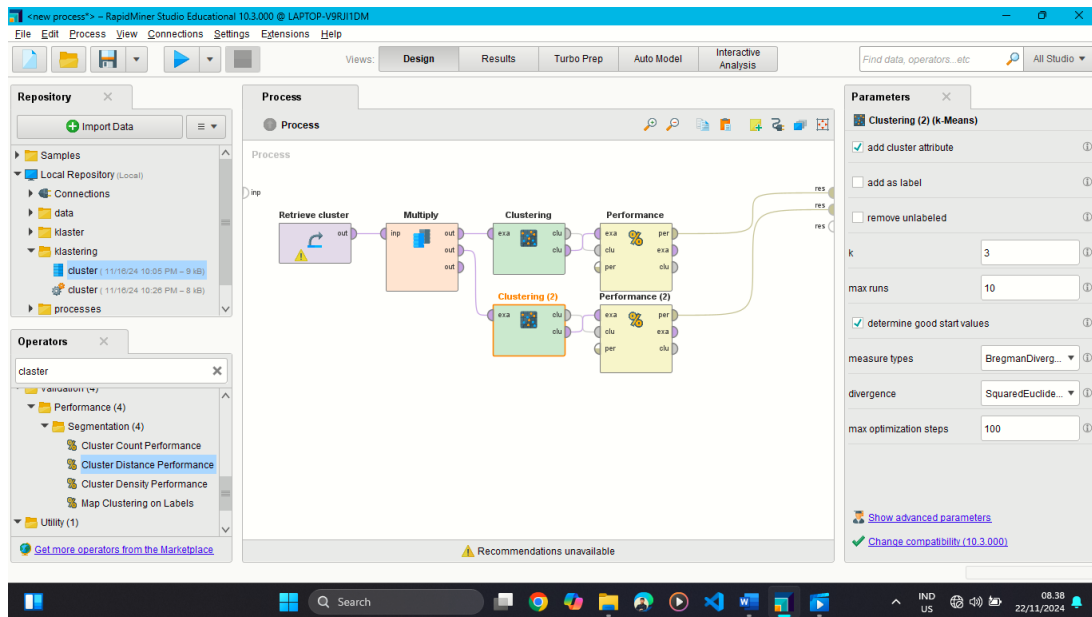


Gambar 1. 6 Nilai Sedang (Prestasi Sedang)

Pada Gambar 1. 6 ini menampilkan hasil evaluasi klusterisasi yang dilakukan menggunakan RapidMiner Studio. Pada hasil yang ditunjukkan, metrik yang digunakan adalah Davies-Bouldin Index (DBI), dengan nilai sebesar -1.097. Biasanya, DBI digunakan untuk mengukur kualitas kluster, di mana nilai yang lebih kecil menunjukkan kluster yang lebih baik. Namun, nilai negatif seperti yang ditampilkan ini memerlukan analisis lebih lanjut, karena secara umum DBI bernilai positif. Dan nilai tersebut yang di tampilkan adalah nilai yang tergolong sedang dari nilai-nilai yang lain.[12]

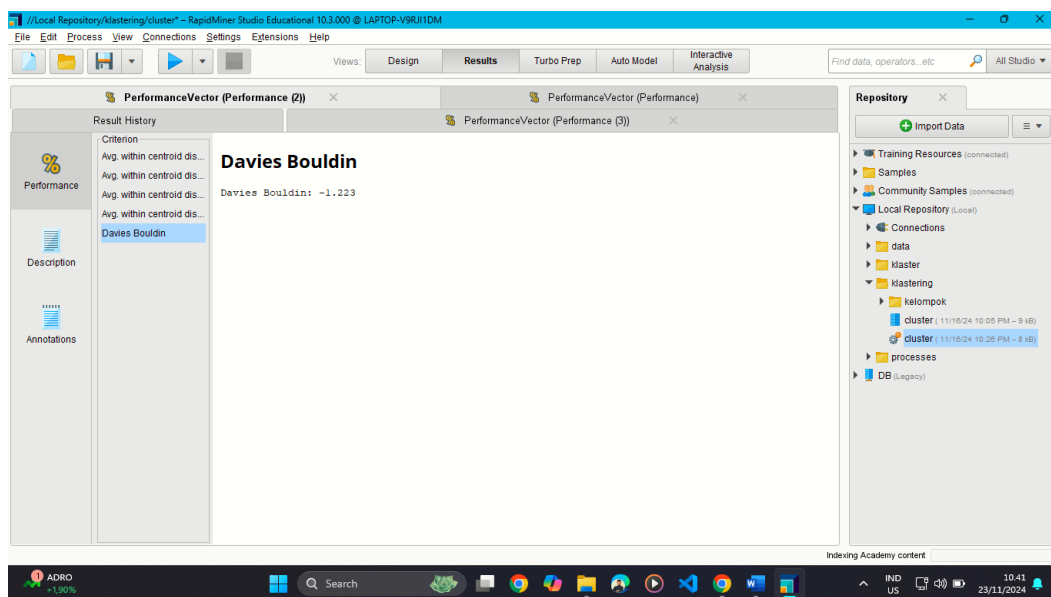
3.3 Kluster 3 (Prestasi Rendah)

Kelompok ini mencakup siswa yang memiliki nilai akademik rendah dan tingkat partisipasi dalam kegiatan ekstrakurikuler yang minim. Siswa dalam kluster ini kemungkinan besar membutuhkan perhatian tambahan, baik dalam aspek akademik maupun dalam motivasi untuk berpartisipasi lebih aktif dalam kegiatan sekolah. Berikut clustering data siswa, seperti gambar di bawah ini.



Gambar 1. 7 Proses cluster dengan (K) 3

Pada Gambar 1. 7 ini menunjukkan proses analisis clustering menggunakan algoritma K-Means di RapidMiner Studio. Data diambil melalui Retrieve Cluster dan digandakan dengan Multiply untuk digunakan pada dua proses Clustering. Kedua proses ini mengelompokkan data ke dalam tiga kluster, dengan parameter seperti jumlah kluster (k) sebesar 3 dan jarak Squared Euclidean. Hasil clustering dievaluasi menggunakan dua operator Performance untuk membandingkan kualitas pengelompokan. Tujuan analisis ini adalah mengevaluasi efektivitas parameter yang digunakan.[13]



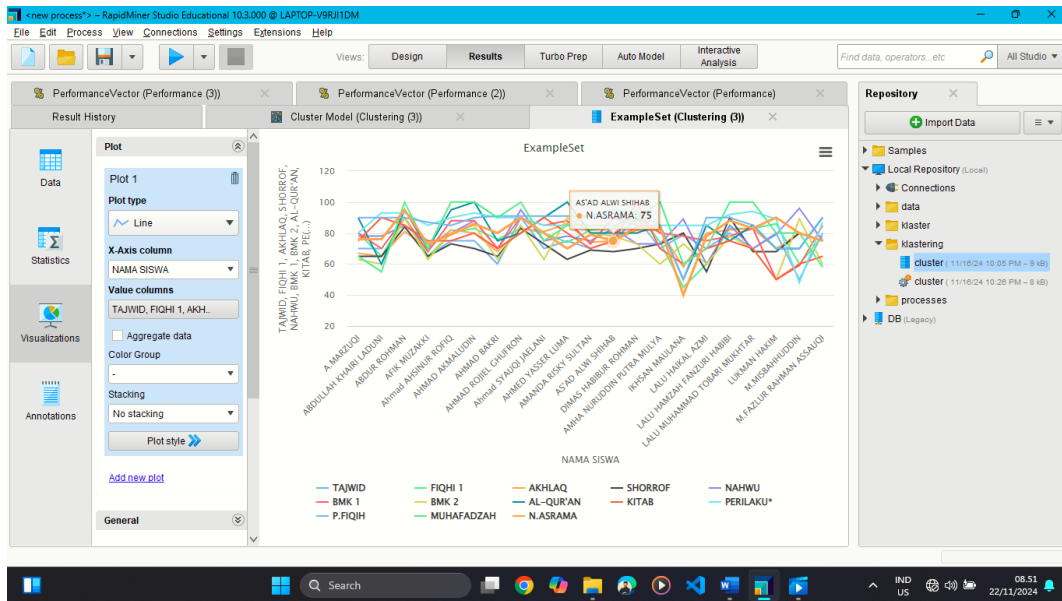
Gambar 1. 8 Nilai Terendah (Prestasi Rendah)

Pada Gambar 1. 8 ini menunjukkan hasil evaluasi model klusterisasi pada perangkat lunak RapidMiner Studio menggunakan metrik Davies-Bouldin Index (DBI). Hasil yang ditampilkan memiliki nilai DBI sebesar -1.223. Secara umum, nilai DBI yang lebih rendah menunjukkan kualitas kluster yang baik, dan nilai DBI yang lebih tinggi menunjukkan kualitas kluster yang kurang baik, karena mengindikasikan bahwa kluster lebih terpisah dan lebih kompak.[14]

3.4 Implikasi Hasil Penelitian bagi Madrasah I'dadiyah Sukorejo

Berdasarkan hasil clustering yang divisualisasikan dalam grafik, terdapat pola yang dapat digunakan untuk meningkatkan kualitas pendidikan di Madrasah I'dadiyah Sukorejo. Grafik ini menunjukkan variasi nilai

siswa dalam berbagai aspek, termasuk mata pelajaran inti seperti Tajwid, Fiqih, Akhlaq, hingga parameter perilaku dan partisipasi dalam kegiatan asrama.



Gambar 1. 9 Grafik Dari Hasil Clustering

Pada Gambar 1. 9 ini adalah tangkapan layar dari perangkat lunak RapidMiner Studio yang menunjukkan analisis data kinerja siswa menggunakan grafik line chart. Sumbu X menampilkan nama siswa, sementara sumbu Y menunjukkan nilai mereka dalam berbagai kategori, seperti TAJWID, FIQIH 1, AKHLAQ, dan lainnya. Garis-garis berwarna membandingkan performa antar siswa. Grafik ini interaktif, dengan fitur yang menyoroti data spesifik, seperti nilai "ASAD ALWI SHIRAB" sebesar 75 di kategori tertentu. Data yang digunakan tersimpan di repositori lokal dengan nama dataset "cluster," dan analisis ini menggunakan metode clustering untuk mengelompokkan siswa berdasarkan performa mereka.[15]

4. KESIMPULAN DAN SARAN

Penelitian ini berhasil mengelompokkan siswa ke dalam tiga kategori utama, yaitu siswa berprestasi tinggi, sedang, dan rendah. Algoritma K-Means yang digunakan terbukti mampu menganalisis data dengan baik dan menghasilkan kelompok yang memiliki karakteristik homogen, sehingga memberikan wawasan yang jelas tentang distribusi dan faktor-faktor yang memengaruhi prestasi siswa.

Hasil penelitian ini memberikan implikasi strategis bagi pihak Madrasah I'dadiyah Sukorejo dalam meningkatkan kualitas pendidikan. Siswa yang tergolong berprestasi rendah disarankan untuk diberikan bimbingan tambahan dan motivasi belajar, sedangkan siswa berprestasi sedang diarahkan untuk lebih aktif dalam kegiatan ekstrakurikuler guna meningkatkan keterampilan interpersonal mereka. Dengan adanya pemahaman yang lebih mendalam tentang pola prestasi siswa, pihak madrasah dapat menyusun program pengajaran yang lebih terarah dan efektif sesuai dengan kebutuhan masing-masing kelompok siswa. Penelitian ini juga menegaskan manfaat penerapan data mining dalam dunia pendidikan, khususnya dalam membantu perencanaan strategis untuk meningkatkan kualitas pembelajaran secara keseluruhan.

Ucapan Terima Kasih

Dengan penuh rasa syukur, kami mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan dan kontribusi dalam penyelesaian penelitian ini. Ucapan terima kasih yang sebesar-besarnya kami sampaikan kepada Madrasah I'dadiyah Sukorejo yang telah menyediakan data dan fasilitas untuk kelancaran penelitian ini.

Kami juga menghaturkan terima kasih kepada pembimbing dan pengajar yang telah memberikan arahan, masukan, dan motivasi dalam setiap tahapan penelitian. Tidak lupa, kepada keluarga dan rekan-rekan yang telah memberikan doa, semangat, serta dukungan moral yang tiada henti.

Semoga hasil penelitian ini dapat memberikan manfaat nyata bagi dunia pendidikan, khususnya dalam upaya meningkatkan kualitas pembelajaran di Madrasah I'dadiyah Sukorejo. Akhir kata, kami menyadari

bahwa penelitian ini masih memiliki kekurangan, sehingga kritik dan saran yang membangun sangat kami harapkan untuk penyempurnaan di masa mendatang. Terima kasih.

DAFTAR PUSTAKA

- [1] B. H. Khan, J. R. Corbeil, and M. E. Corbeil, *Responsible Analytics and Data Mining in Education: Global Perspectives on Quality, Support, and Decision Making*. Taylor & Francis, 2018. [Online]. Available: <https://books.google.co.id/books?id=xWN8DwAAQBAJ>
- [2] K. Anam and Z. Fata, “Korelasi Perilaku Dengan Tingkat Prestasi Siswa I ’ dadiyah Menggunakan Metode Apriori,” vol. 2, pp. 266–272, 2024.
- [3] E. Irwansyah and M. Faisal, *Advanced Clustering: Teori dan Aplikasi*. DeePublish, 2015. [Online]. Available: <https://books.google.co.id/books?id=8y80BgAAQBAJ>
- [4] L. M. Given, *The Sage Encyclopedia of Qualitative Research Methods: A-L ; Vol. 2, M-Z Index*. in A Sage Reference Publication. SAGE Publications, 2008. [Online]. Available: <https://books.google.co.id/books?id=IFrb6IPLISEC>
- [5] G. K. GUPTA, *INTRODUCTION TO DATA MINING WITH CASE STUDIES*. PHI Learning, 2014. [Online]. Available: <https://books.google.co.id/books?id=fzB9BAAAQBAJ>
- [6] K. Sud, P. Erdogmus, and S. Kadry, *Introduction to Data Science and Machine Learning*. IntechOpen, 2020. [Online]. Available: <https://books.google.co.id/books?id=THb-DwAAQBAJ>
- [7] S. Ramjan and J. Sunkpho, *Principles and Theories of Data Mining With RapidMiner*. in Advances in Computer and Electrical Engineering. IGI Global, 2023. [Online]. Available: <https://books.google.co.id/books?id=ISzCEAAAQBAJ>
- [8] M. Hofmann and R. Klinkenberg, *RapidMiner: Data Mining Use Cases and Business Analytics Applications*. in Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. CRC Press, 2016. [Online]. Available: https://books.google.co.id/books?id=Y_wYCwAAQBAJ
- [9] R. F. Putra *et al.*, *DATA MINING : Algoritma dan Penerapannya*. PT. Sonpedia Publishing Indonesia, 2023. [Online]. Available: <https://books.google.co.id/books?id=zLHGEEAAAQBAJ>
- [10] W. L. Martinez, A. R. Martinez, and J. Solka, *Exploratory Data Analysis with MATLAB*. in Chapman & Hall/CRC Computer Science & Data Analysis. CRC Press, 2017. [Online]. Available: <https://books.google.co.id/books?id=8RIwDwAAQBAJ>
- [11] J. Wu, *Advances in K-means Clustering: A Data Mining Thinking*. in Springer Theses. Springer Berlin Heidelberg, 2012. [Online]. Available: https://books.google.co.id/books?id=pI2_F8SqWcQC
- [12] G. Gan, C. Ma, and J. Wu, *Data Clustering: Theory, Algorithms, and Applications, Second Edition*. in Mathematics in Industry. Society for Industrial and Applied Mathematics, 2020. [Online]. Available: <https://books.google.co.id/books?id=r4wIEAAAQBAJ>
- [13] E. A. Novia, W. I. Rahayu, and C. Prianto, *SISTEM PERBANDINGAN ALGORITMA K-MEANS DAN NAÏVE BAYES UNTUK MEMPREDIKSI PRIORITAS PEMBAYARAN TAGIHAN RUMAH SAKIT BERDASARKAN TINGKAT KEPENTINGAN*. Kreatif. [Online]. Available: <https://books.google.co.id/books?id=MND9DwAAQBAJ>
- [14] S. Wierzchoń and M. Kłopotek, *Modern Algorithms of Cluster Analysis*. in Studies in Big Data. Springer International Publishing, 2017. [Online]. Available: <https://books.google.co.id/books?id=LeJEDwAAQBAJ>
- [15] Y. Ardilla *et al.*, *DATA MINING DAN APLIKASINYA*. Penerbit Widina, 2021. [Online]. Available: <https://books.google.co.id/books?id=53FXEAAAQBAJ>