

Implementasi Rapidminer Untuk Menentukan Siswa Unggulan Menggunakan Metode K-Means

Nabella Rosyefa Wahyudi¹, Yunianita Rahmawati², Suprianto³

^{1,2,3}Fakultas Sains dan Teknologi, Teknik Informatika, Universitas Muhammadiyah, Sidoarjo, Indonesia,

E-mail: ¹nabella2407@gmail.com, ²yunianita@umsida.ac.id, ³suprianto@umsida.ac.id

(* : corresponding author)

Abstrak

Penelitian ini bertujuan untuk menerapkan Metode K-Means dan RapidMiner dalam mengidentifikasi siswa unggulan berdasarkan nilai mata pelajaran Matematika, IPA, dan IPS di SMP Negeri 11 Sampit. Dengan dataset yang terdiri dari 24 atribut dan 31 data siswa, Metode K-Means berhasil mengelompokkan siswa berdasarkan kemiripan nilai yang siswa tersebut peroleh. Hasil evaluasi model menunjukkan bahwa nilai *cluster* terbaik berjumlah 2 *cluster*. Penggunaan Metode K-Means dalam penelitian ini merupakan solusi untuk kesulitan dalam menentukan siswa unggulan di SMP Negeri 11 Sampit karena kemampuan siswa cenderung seimbang di setiap semester. Penelitian ini memberikan wawasan bahwa Metode K-Means efektif dalam menentukan siswa unggulan dan dapat menjadi alat yang berguna dalam evaluasi Pendidikan.

Kata kunci: RapidMiner, K-Means, Siswa, Akademik

Abstract

This research aims to apply the K-Means Method and RapidMiner in identifying superior students based on the scores of Mathematics, Science, and Social Studies subjects at SMP Negeri 11 Sampit. With a dataset consisting of 24 attributes and 31 student data, the K-Means Method successfully clusters students based on the similarity of the grades the students obtained. The model evaluation results show that the best cluster value is 2 clusters. The use of the K-Means Method in this study is a solution to the difficulty in determining superior students at SMP Negeri 11 Sampit because students' abilities tend to be balanced in each semester. This research provides insight that the K-Means Method is effective in determining superior students and can be a useful tool in educational evaluation.

Keywords: RapidMiner, K-Means, Student, Academic

1. PENDAHULUAN

Di era globalisasi saat ini, cara belajar siswa sangat bergantung pada kehadiran media teknologi informasi. Hal ini mencerminkan karakteristik siswa abad ke-21 yang memegang kendali, lebih menyukai beragam pilihan pembelajaran, dan menikmati kolaborasi teknis [1]. Pentingnya media yang dihasilkan teknologi dapat menjadi sumber peningkatan standar mutu pendidikan dan sekolah. Pendidikan memegang peranan penting dalam kehidupan manusia, menciptakan kondisi bagi individu untuk mengekspresikan bakat, inovasi sebagai salah satu faktor penunjang mutu pendidikan di Indonesia. Setiap sekolah berupaya untuk meningkatkan mutu sekolahnya melalui pengelompokan siswa unggulan yang didasari kemampuan siswa dan kriteria sekolah.

Pengelompokan siswa unggulan, secara umum ditentukan melalui nilai siswa pada setiap mata pelajaran di kelas. Objek penelitian ini menggunakan data nilai siswa kelas 7 SMP Negeri 11 Sampit. Berdasarkan hasil observasi yang dilakukan bahwa SMP Negeri 11 Sampit mengalami kesulitan untuk menentukan perwakilan siswa lomba seperti lomba OSN (Olimpiade Sains Nasional) karena kemampuan siswa SMP Negeri 11 Sampit di setiap generasi cenderung seimbang. Selain itu penempatan siswa unggulan dapat berubah menyesuaikan nilai yang diperoleh siswa setiap semester. Salah satu solusi yang dapat dilakukan pada SMP Negeri 11 Sampit yaitu dengan menerapkan metode data mining *clustering*. Salah satu jenis teknik dalam data mining *clustering* adalah Metode K-Means. Metode ini berfungsi untuk membagi data ke dalam *cluster-cluster* sehingga yang serupa akan dikelompokkan ke dalam *cluster* yang sama, sementara data yang berbeda akan dikelompokkan dalam *cluster* yang serupa dengan data tersebut [2].

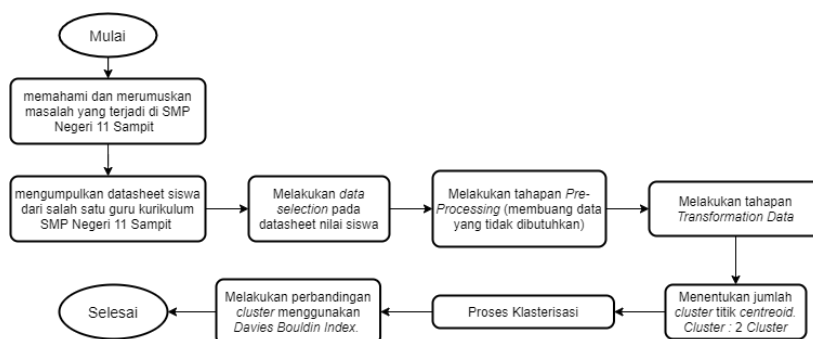
Penelitian ini akan melakukan pengelompokan data siswa dengan menerapkan Metode K-Means yang memiliki tingkat keakuratan tinggi, hal ini dibuktikan dalam paper [3] yang berjudul ‘Analisis Algoritma K-Means dan K-Medoids Untuk *Clustering* Data Kinerja Karyawan Pada Perusahaan Perumahan Nasional’. Selain itu, penerapan Algoritma K-Means lebih mudah dan sederhana serta sering digunakan pada proses data mining [4],[5]. Algoritma K-Means merupakan salah satu jenis algoritma *clustering* dalam mengelompokkan data dengan menentukan titik pusat *cluster* terdekat [6]. Penerapan metode ini termasuk metode berulang yang sederhana karena metode ini cenderung *flexible* dan tidak memakan waktu lama [7]. Algoritma K-Means juga dipahami meminimalkan banyak kesalahan [8].

Penelitian terdahulu juga telah menjelaskan tentang Algoritma K-Means seperti paper [9] yang berjudul ‘Implementasi Algoritma K-Means *Clustering* dalam Penentuan Siswa Kelas Unggulan’, penelitian ini bertujuan untuk mengelompokkan siswa pada kelas yang lebih tinggi dengan mengimplementasikan Metode K-Means *Clustering*. Kemudian, penelitian yang telah dijelaskan oleh Reno Supardi dalam paper [10], penelitian ini bertujuan untuk membantu pelaporan penjualan di Toko Edelwis. Selanjutnya, penelitian yang telah dijelaskan oleh Putra Primanda dalam paper [11], penelitian ini bertujuan untuk membantu mengklastering kelas unggulan dengan menerapkan Metode K-Means. Selanjutnya paper [12], penelitian ini bertujuan memudahkan proses klasterisasi untuk memantau perkembangan kemakmuran masyarakat di Daerah Aceh. Serta penelitian yang telah di jelaskan oleh Suraya dalam paper [13], penelitian ini bertujuan untuk membantu mengklastering IPK (Indeks Prestasi Mahasiswa) dengan menerapkan Algoritma K-Means. Penelitian sebelumnya yang bertujuan untuk mengelompokkan siswa yang unggul dengan mengimplementasikan algoritma K-Means telah dilakukan di berbagai lembaga pendidikan, namun penelitian saya hanya terfokus pada SMP Negeri 11 Sampit, dimana pengimplementasian Algoritma K-Means menggunakan RapidMiner belum dilakukan sebelumnya disana. Penelitian ini mengisi kesenjangan dalam literatur dengan memberikan lebih banyak informasi tentang penentuan siswa yang unggul di suatu institusi tertentu.

Berdasarkan pemaparan di atas, kajian ini bertujuan yaitu untuk mempermudah SMP 11 SAMPIT dalam mengelompokkan siswa unggulan sebagai bahan evaluasi meningkatkan mutu sekolah serta membuktikan tingkat keakuratan performa *cluster Davies Bouldin Index* (DBI). Penelitian ini menggunakan *tools* RapidMiner untuk mempermudah perhitungan pada data mining. Untuk mencapai tujuan tersebut maka dibuatlah solusi yaitu implementasi Rapidminer untuk menentukan siswa unggulann menggunakan metode k-means.

2. METODE PENELITIAN

Peneliti melakukan serangkaian proses di mana setiap proses berpotensi membantu penelitian ini menjadi lebih mudah. Diagram alur penelitian dari tahapan proses ditunjukkan oleh Gambar 1:



Gambar 1. Alur Penelitian

Pada Gambar 1 menampilkan langkah mengidentifikasi, merumuskan dan memahami permasalahan yang akan diteliti, salah satunya yaitu dengan melakukan studi literatur, selanjutnya mengumpulkan data berupa data nilai siswa kelas 7. Data yang terkumpul sebanyak 31 siswa.

Kemudian, melakukan analisis hasil dengan menerapkan Metode K-Means pada setiap data hingga menemukan hasil akhir.

2.1. Metode Pengambilan data

Metode pengumpulan data kuantitatif meliputi dokumentatif, yaitu mengumpulkan data nilai keseluruhan siswa kelas VII yang akan diteliti.

2.2. Data Mining

Data mining merupakan proses yang menerapkan matematika, berbagai teknik statistik, kecerdasan buatan yang bertujuan untuk mengidentifikasi dan mengekstraksi data/ informasi berharga dari sekumpulan data yang lebih besar [14]. Data mining memiliki beberapa teknik seperti *clustering*, analisis regresi, *text mining* dan lain lain.

2.3. Davies Bouldin Indeks (DBI)

Evaluasi DBI (*Davies Bouldin Index*) perlu dilakukan untuk mencari *cluster* terbaik dari metode *clustering*. Nilai *performance* proses *clustering* akan semakin baik apabila nilai tersebut semakin kecil [15]. DBI berperan cukup penting dalam penelitian ini karena dengan DBI, peneliti dapat mengetahui kevalidan tingkat peforma K-Means dalam menentukan K (*cluster*).

2.4. Metode K-Means

Metode K-Means merupakan metode *clustering* yang mengelompokkan data yang mirip satu sama lain ke dalam *cluster* yang sama. Data yang memiliki kemiripan atau sama dikelompokkan ke dalam satu *cluster* dan yang berbeda akan dikelompokkan ke *cluster* yang lain. [16]. Berikut adalah gambaran umum proses *clustering* Algoritma K-Means :

- a. Menentukan jumlah K (*cluster*).
- b. Memilih titik pusat *cluster* (*centroid*) secara acak. Titik *centroid* ditentukan acak sesuai dengan jumlah *cluster*.
- c. Menghitung jarak setiap *cluster* dengan menggunakan rumus *Euclidean Distance*. Rumus ini menghitung selisih data dengan titik *centroid* lalu menjumlahkan tiap data tersebut. Rumus *Euclidean Distance* ditunjukkan oleh Gambar 2.

$$\text{dist}(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$$

Gambar 2. Rumus *Euclidean Distance*

Keterangan Gambar 2 :

- $\sum_{i=1}^n$, $i=1$ bermakna bahwa perhitungan penjumlahan dimulai dari atribut pertama hingga n, n merupakan jumlah total dari atribut data yang akan dihitung.
 - $(x_i - y_i)^2$, x_i merupakan data awal dari atribut pertama dan nilai dari y_i merupakan hasil rata rata dari semua data *cluster* pada iterasi ke-2 dan seterusnya.
- d. Memperbarui titik pusat *cluster* (*centroid*).
Rumus mencari titik pusat *cluster* ditunjukkan pada Gambar 3.

$$C_j = \frac{1}{n_j} \sum_{i=1}^{n_j} x_i$$

Gambar 3. Rumus memperbarui *centroid*

Keterangan Gambar 3 :

- n_j , merepresentasikan jumlah data dari j (*cluster*) yang sudah dialokasikan sebelumnya.

- $\sum_{i=1}^{n_j} x_i$, $i=1$ bermakna bahwa perhitungan penjumlahan dimulai dari atribut pertama hingga n_j , n_j merupakan jumlah total dari atribut data *cluster* yang akan dihitung sesuai dengan posisi atribut ke i .
- e. Ulangi langkah-langkah 2 dan 3 hingga tidak ada perubahan yang signifikan dalam proses pengelompokan data.

2.5. Datasheet (Data Nilai Siswa)

Row No.	Nama	cluster	Agm_PA2	Agm_KA2	PPkn_PA2	PPkn_KA2	Bind_PA2	Bind_KA2	Mat_PA2	Mat_KA2	IPA_PA2	IPA_KA2	IPS_PA2
1	ADITYA RAMA...	cluster_0	87	86	78	78	78	79	76	71	71	80	80
2	Akbar Risky S...	cluster_0	87	86	78	78	79	80	75	73	71	85	81
3	Alam Setiago	cluster_0	92	94	78	78	81	83	78	78	73	85	85
4	Amelsa Putri ...	cluster_0	92	94	79	79	84	82	76	78	73	85	85
5	APRIL SIDIK	cluster_0	88	94	78	78	80	81	79	78	73	85	86
6	Aulia Aini Nur...	cluster_0	90	94	78	78	82	80	80	77	71	85	84
7	Ayu Sufiah	cluster_0	89	94	78	78	85	80	78	83	75	85	85
8	Beri Abdi Pa...	cluster_0	86	86	78	78	78	79	72	71	53	80	80
9	Caesha Athi ...	cluster_0	92	94	80	80	85	85	79	78	77	85	87
10	Dendy Putra ...	cluster_0	89	94	78	78	81	80	76	73	71	85	83
11	DIDA YULIAN...	cluster_0	92	94	80	80	87	84	79	78	73	85	85
12	DIMAS PUTRA	cluster_0	90	94	78	78	81	80	73	71	70	80	84
13	Gesang Mukti...	cluster_0	92	94	80	80	81	85	81	78	72	85	86
14	IKA ARIANIS ...	cluster_0	89	94	78	78	85	82	81	78	73	85	84
15	Ilyas Abdullh	cluster_0	88	94	80	80	80	81	75	73	78	85	86
16	M. Jordi Apry...	cluster_0	87	94	78	78	83	83	78	70	71	85	83
17	M. Naufal Tufail	cluster_0	90	94	78	78	80	80	78	81	72	85	86
18	M. RAQVA HA	cluster_0	92	94	80	80	82	83	82	78	87	86	88

Gambar 4. Datasheet (Data Nilai Siswa)

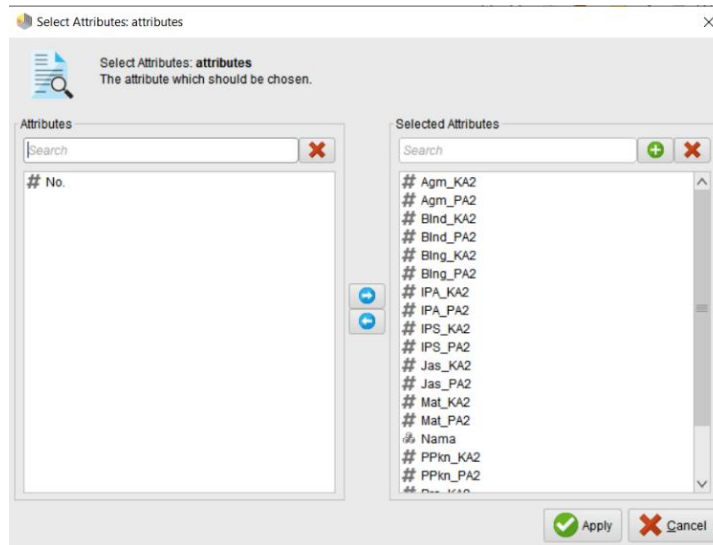
Datasheet nilai siswa kelas 7 yang akan dikelola ditampilkan pada Gambar 4. *Dataset* yang diolah memiliki 24 *attribute* sama seperti Nama, Agm_PA, Agm_KA, PPkn_PA, PPkn_KA, Bind_PA, Bind_KA, Mat_PA, Mat_KA, IPA_PA, IPA_KA, IPS_PA, IPS_KA, Bing_PA, Bing_KA, Seni_PA, Seni_KA, Jas_PA, Jas_KA, Pra_PA, Pra_KA. *Dataset* memiliki jumlah data siswa sebanyak 31 data. *Dataset* ini diperoleh dari guru kurikulum kelas 7 SMP Negeri 11 Sampit.

3. HASIL DAN PEMBAHASAN

Proses perhitungan penerapan Metode K-Means untuk menentukan siswa unggulan ini dilakukan menggunakan *tools* RapidMiner. Proses perhitungan ini hanya berfokus pada mata pelajaran yang diujikan OSN (Olimpiade Sains Nasional) saja seperti Matematika, IPA dan IPS. Berikut hasil dan pembahasan yang sudah dilakukan oleh peneliti dengan menggunakan *tools* yang sudah dijelaskan sebelumnya.

3.1 Data Selection

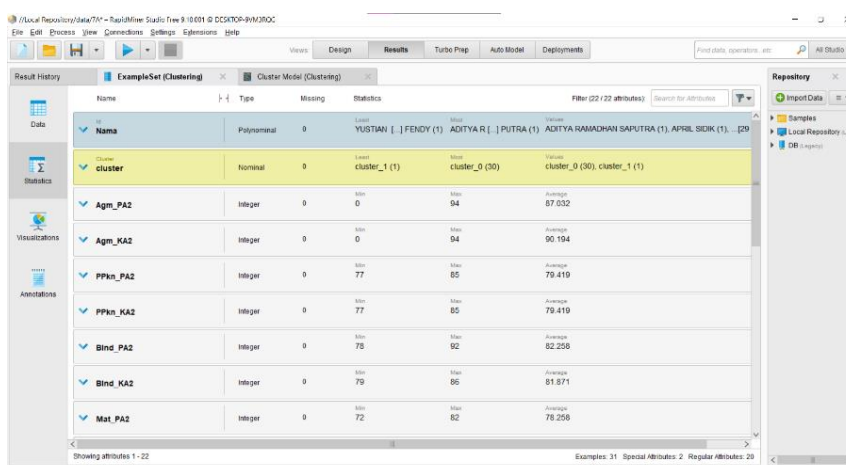
Setelah semua data terkumpul, tahapan yang perlu dilakukan selanjutnya yaitu menyeleksi *attribute* yang akan digunakan dalam pengimplementasian Algoritma K-Means. Tahapan seleksi data dalam rapidminer dapat menggunakan operator *select attribute*. Fitur ini berfungsi untuk memilah keseluruhan atribut data yang dibutuhkan oleh operator perhitungan. Proses tahapan seleksi atribut dari *datasheet* nilai siswa kelas 7 SMP 11 Sampit ditunjukkan oleh Gambar 5. Proses seleksi data kelas 7 menggunakan operator *select attribute* pada RapidMiner. Data tersebut memiliki 21 *attribute* yaitu Nama, Agm_PA, Agm_KA, PPkn_PA, PPkn_KA, Bind_PA, Bind_KA, Mat_PA, Mat_KA, IPA_PA, IPA_KA, IPS_PA, IPS_KA, Bing_PA, Bing_KA, Seni_PA, Seni_KA, Jas_PA, Jas_KA, Pra_PA, Pra_KA. Atribut nama diubah menjadi tipe data id untuk mempermudah proses pengolahan data. Setelah diubah, atribut tersebut akan dihitung sebagai label objek yang akan diteliti.



Gambar 5. Proses seleksi atribut data kelas 7

3.2 Pre-Processing

Setelah proses seleksi data nilai selesai, langkah berikutnya adalah melakukan *pre-processing* data pada hasil dari tahapan sebelumnya. *Pre-Processing* bertujuan untuk membersihkan data yang tidak perlu digunakan/ data kosong/ data ganda dan sejenisnya. Proses tahapan *Pre-Processing* dari datasheet nilai siswa kelas 7 SMP 1 Sampit ditunjukkan oleh Gambar 6.



Gambar 6. Proses *Pre-Processing* data kelas 7

Pada Gambar 6 ditampilkan stastitika hasil seleksi data kelas 7 setelah dilakukan *running* di *software* Rapid Minner. Data kelas 7 memiliki 2 atribut spesial dan 31 *example* di mana atribut Nama masuk ke dalam salah satu atribut spesial karena memiliki tipe *polynomial* yang tidak bisa diolah dengan perhitungan Algoritma *K-Means*.

3.3 Tranformation Data

Setelah tahapan *pre-processing* selesai, langkah berikutnya adalah melakukan *transformation* data pada hasil dari tahapan sebelumnya. *Transformation* data merupakan proses mengubah / memanipulasi data mentah menjadi lebih sesuai untuk dianalisis. Proses tahapan *transfarmation data* dari datasheet nilai siswa kelas 7 SMP 11 Sampit ditunjukkan pada Gambar 7.

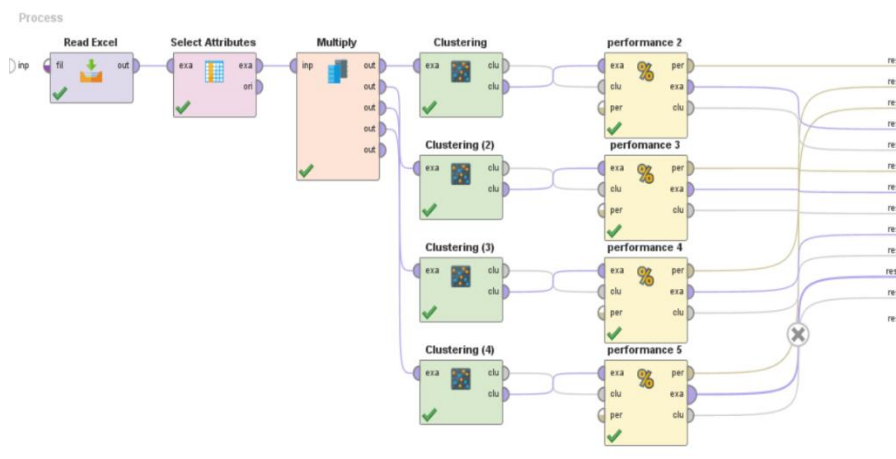
Row No.	Nama	cluster	Agri_PA2	Agri_KA2	PPhis_PA2	PPhis_KA2	Biol_PA2	Biol_KA2	Mat_PA2	Mat_KA2	IPA_PA2
1	ADITTA RAMA...	cluster_0	87	86	78	78	78	79	76	71	71
2	Adzar Rizki S...	cluster_0	87	86	78	78	79	80	75	73	71
3	Alan Setiapa	cluster_0	92	94	78	78	81	83	78	78	73
4	Amelia Putri	cluster_0	92	94	79	79	84	82	76	78	73
5	APRIL SIDIK	cluster_0	88	94	78	78	80	81	79	78	73
6	Aulia Aini Nur...	cluster_0	90	94	78	78	82	80	80	77	71
7	Ayu Sulhan	cluster_0	89	94	78	78	85	83	78	83	75
8	Beti Adid Pa...	cluster_0	88	86	78	78	79	72	71	71	63
9	Ceselia APR...	cluster_0	92	94	80	80	85	85	79	78	77
10	Denny Putra...	cluster_0	89	94	78	78	81	80	76	73	71
11	DIDA YULIANI	cluster_0	92	94	80	80	87	84	79	78	73
12	DMAS PUTRA	cluster_0	90	94	78	78	81	80	73	71	70
13	Gasang Mus...	cluster_0	92	94	80	80	81	85	81	78	72
14	IKA ARNAND...	cluster_0	89	94	78	78	85	82	81	78	73
15	Ryza Abdillah	cluster_0	88	94	80	80	80	81	75	73	78
16	M. Jora Agni...	cluster_0	87	94	78	78	83	83	78	70	71
17	M. Naulie Tatal	cluster_0	90	94	78	78	80	80	78	81	72
18	M. RAHYA HA...	cluster_0	85	84	80	80	82	81	82	78	87

Gambar 7. Proses Transformation Data Data kelas 7

Pada tahapan Transformation data, hasil data yang didapat sudah berupa integer/ angka sehingga pada tahapan ini tidak ada perlu diubah karena memenuhi syarat penerapan Algoritma K-Means.

3.4 Data Mining

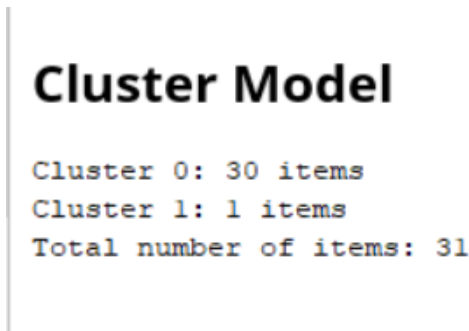
Setelah melalui tahapan Transformation data. Langkah berikutnya yaitu melakukan tahapan data mining pada hasil dari tahapan sebelumnya. Proses tahapan Data Mining dari datasheet nilai siswa kelas 7 SMP 11 Sampit ditampilkan pada Gambar 8, proses clustering dilakukan menggunakan tools RapidMiner. Proses tersebut menggunakan 2 - 4 cluster berbeda yang bermaksud untuk membandingkan nilai cluster lalu mencari nilai yang terbaik dari setiap cluster. Sebelum proses cluster dilakukan, data akan dibaca menggunakan operator read excel lalu selanjutnya dilakukan proses seleksi atribut menggunakan operator select attribute.



Gambar 8. Proses Clustering pada data kelas 7

3.5 Analisis Cluster

Hasil cluster dari proses penerapan Metode K-Means untuk menentukan siswa unggulan menggunakan Rapidminer dari datasheet nilai siswa kelas 7 yang sudah dilampirkan sebelumnya ditunjukkan pada Gambar 9. Hasil pemodelan jumlah cluster yang terbentuk dari perhitungan Metode K-Means di kelas 7, dimana cluster 0 mendapatkan 30 data dan cluster 1 mendapatkan 1 data. Cluster 0 merupakan representasi dari siswa unggul dan cluster 1 merupakan representasi dari siswa bukan unggulan dari seluruh mata pelajaran. Tampilan hasil dari keseluruhan data yang sudah melalui tahapan Clustering menggunakan Metode K-Means ditunjukkan pada Gambar 10.



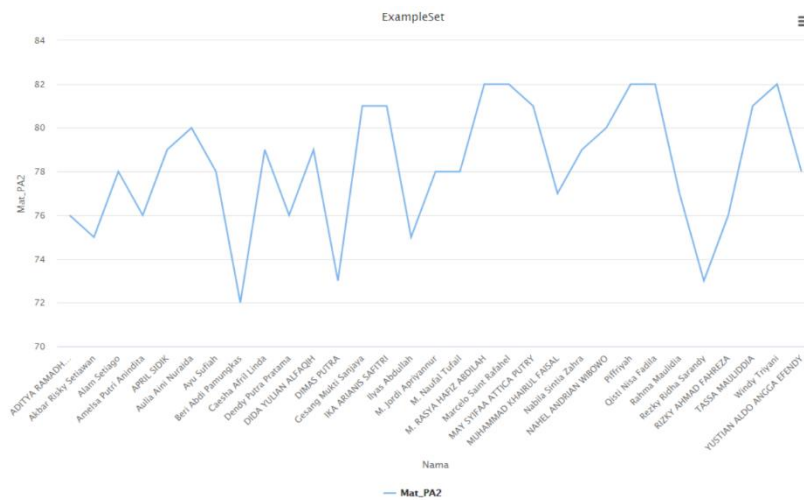
Gambar 9. Hasil Cluster Kelas 7

Row No.	Nama	cluster	Agm_PA2	Agm_KA2	PPkn_PA2	PPkn_KA2	Blind_PA2	Blind_KA2	Mat_PA2	Mat_KA2	IPA_PA2
1	ADITYA RAMA...	cluster_0	87	86	78	78	78	79	76	71	71
2	Akbar Risky S...	cluster_0	87	86	78	78	79	80	75	73	71
3	Alam Seliago	cluster_0	92	94	78	78	81	83	78	78	73
4	Amelsa Putri ...	cluster_0	92	94	79	79	84	82	76	78	73
5	APRIL SIDIK	cluster_0	88	94	78	78	80	81	79	78	73
6	Aulia Aini Nur...	cluster_0	90	94	78	78	82	80	80	77	71
7	Ayu Sufiah	cluster_0	89	94	78	78	85	80	78	83	75
8	Beri Abdi Pa...	cluster_0	86	86	78	78	78	79	72	71	53
9	Caesha Afril ...	cluster_0	92	94	80	80	85	85	79	78	77
10	Dendy Putra ...	cluster_0	89	94	78	78	81	80	76	73	71
11	DIDA YULIAN...	cluster_0	92	94	80	80	87	84	79	78	73
12	DIMAS PUTRA	cluster_0	90	94	78	78	81	80	73	71	70
13	Gesang Mukl...	cluster_0	92	94	80	80	81	85	81	78	72
14	IKA ARIANIS ...	cluster_0	89	94	78	78	85	82	81	78	73
15	Ilyas Abdull...	cluster_0	88	94	80	80	80	81	75	73	78
16	M. Inrdi Anrv	cluster_0	87	94	78	78	83	83	78	70	71

Gambar 10. Data Hasil Cluster Kelas 7

Data yang ditampilkan pada Gambar 10 dapat digambarkan juga dengan grafik untuk diketahui perankingan data di setiap mata pelajarannya Berikut tampilan grafik perankingan data kelas 7 dapat dilihat pada Gambar 10.

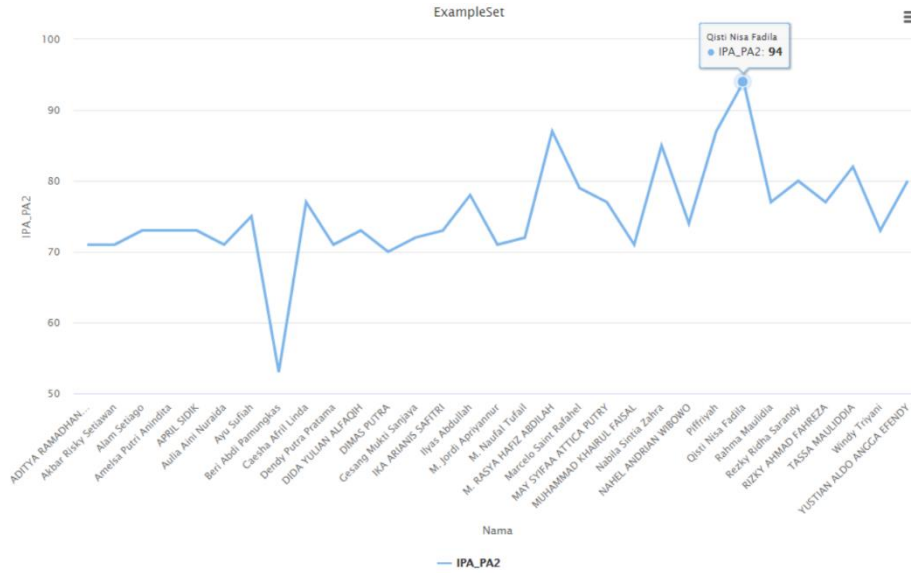
3.5.1 Mata Pelajaran Matematika



Gambar 11. Grafik Ranking Nilai Mata Pelajaran Matematika Kelas 7

Pada Gambar 11 diketahui secara visual dari grafik bahwa siswa dengan nama M. Rasya Hafiz Abdillah, Marcelo Saint Rahafel, Piffriyah, Qisti Nisa Fadilla dan Windy Triyani menjadi siswa paling unggul pada mata pelajaran Matematika dibandingkan temannya yang lain.

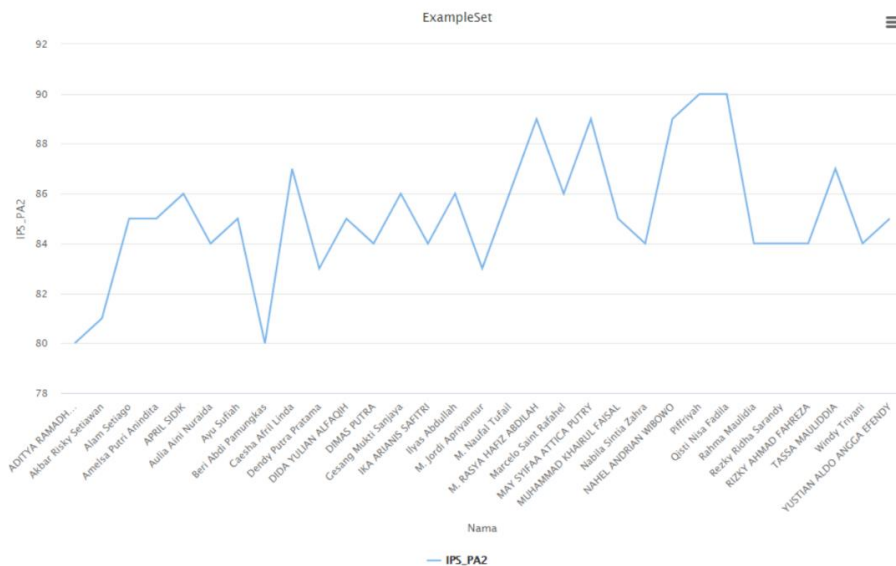
3.5.2 Mata Pelajaran IPA



Gambar 12. Grafik Ranking Nilai Mata Pelajaran IPA Kelas 7

Pada Gambar 12 diketahui secara visual dari grafik bahwa siswi dengan nama Qisti Nisa Fadilla menjadi siswa paling unggul pada mata pelajaran IPS dibandingkan temannya yang lain.

3.5.3 Mata Pelajaran IPS

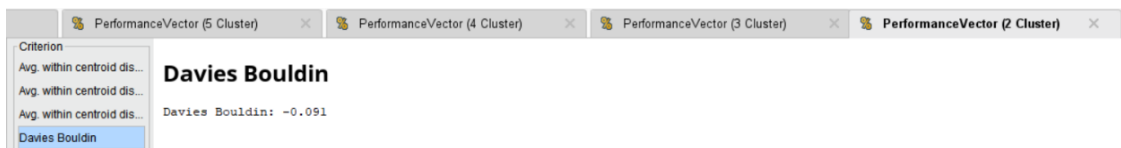


Gambar 13. Grafik Ranking Nilai Mata Pelajaran IPS Kelas 7

Pada Gambar 13 diketahui secara visual dari grafik bahwa siswi dengan nama Qisti Nisa Fadilla dan Piffriyah menjadi siswa paling unggul pada mata pelajaran IPS dibandingkan teman yang lain.

3.6 Evaluasi Model

Setelah proses seleksi data nilai selesai, langkah berikutnya adalah melakukan evaluasi model pada hasil dari tahapan sebelumnya. Setiap pengujian menggunakan algoritma K-Means menghasilkan *cluster* yang berbeda-beda tergantung pada nilai K yang digunakan. Dalam empat percobaan yang dilakukan, nilai K yang berbeda menghasilkan hasil yang berbeda pula. Berikut dibawah ini hasil nilai pengujian *performance* data kelas 7 yang paling baik yang ditampilkan pada Gambar 14 :



Gambar 14. Hasil Nilai pengujian *Performance* data Kelas 7

Mengacu tabel hasil pengujian performa K-2 hingga K-5 (K = Jumlah Cluster) pada Tabel 1, dijelaskan bahwa *cluster* yang cocok untuk diterapkan dalam proses perhitungan sebelumnya yaitu sejumlah 2 kluster dengan perolehan nilai performa sebesar -0,091 karena merupakan nilai terkecil dari nilai performa yang lain.

Tabel 1. Hasil Pengujian *Performance* data kelas 7

Nilai K	<i>Davies Bouldin Index (DBI)</i>
2	-0.091
3	-0,866
4	-0,681
5	-0,906

4. KESIMPULAN DAN SARAN

Berdasarkan hasil penelitian yang telah dilakukan oleh peneliti, dapat ditarik kesimpulan bahwa penerapan Metode K-Means menggunakan *tools* RapidMiner bekerja dengan baik dan mendapatkan nilai *cluster* yang akurat. Dari hasil perhitungan menggunakan *tools* RapidMiner mendapatkan beberapa nama siswa/i kelas 7 yang unggul di mata pelajaran Matematika, IPA dan IPS. Siswa yang bernama M. Rasya Hafiz Abdillah, Marcelo Saint Rahafel, Piffriyah, Qisti Nisa Fadilla dan Windy Triyani unggul di mata pelajaran Matematika, Siswi yang bernama Qisti Nisa Fadilla unggul di mata pelajaran IPA dan Siswi yang bernama Qisti Nisa Fadilla dan Piffriyah unggul di mata pelajaran IPS. Pada penelitian selanjutnya diharapkan penelitian ini dapat dikembangkan serta direalisasikan dalam bentuk aplikasi/website sehingga mempermudah pengguna untuk menerapkan Metode K-Means selain menggunakan *tools* RapidMiner.

DAFTAR PUSTAKA

- [1] S. D. Laksana, "Pentingnya Pendidikan Karakter dalam Menghadapi Teknologi Pendidikan Abad 21," *Jurnal Teknologi Pembelajaran*, vol. 1, no. 01, pp. 14–22, 2021.
- [2] F. P. Dewi, P. S. Aryni, and Y. Umaidah, "Implementasi Algoritma K-Means *Clustering* Seleksi Siswa Berprestasi Berdasarkan Keaktifan dalam Proses Pembelajaran," *JISKA (Jurnal Informatika Sunan Kalijaga)*, vol. 7, no. 2, pp. 111–121, 2022.
- [3] A. U. Fitriyadi, A. Kurniawati, "Analisis Algoritma K-Means dan K-Medoids untuk *Clustering* Data Kinerja Karyawan pada Perusahaan Perumahan Nasional," *KILAT*, vol. 10, no. 1, pp. 157–168, 2021.
- [4] Y. P. Sari, A. Primajaya, and A. S. Y. Irawan, "Implementasi Algoritma K-Means untuk *Clustering* Penyebaran Tuberkulosis di Kabupaten Karawang," *Jurnal INOVTEK Polbeng - Seri Informatika*, vol. 5, no. 2, pp. 229-239, 2020.

- [5] J. Hutagalung, "Pemetaan Siswa Kelas Unggulan Menggunakan Algoritma K-Means Clustering," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 9, no. 1, pp. 606–620, 2022.
- [6] D. Darmansah and N. W. Wardani, "Analisis Pesebaran Penularan Virus Corona di Provinsi Jawa Tengah Menggunakan Metode K-Means Clustering," *JATISI (Jurnal Teknik Informatika dan Sistem Informasi)*, vol. 8, no. 1, pp. 105–117, 2021.
- [7] H. Havaluddin, et al., "Implementasi Metode K-Means untuk Pengelompokan Rekomendasi Tugas Akhir," *Informatika Mulawarman Jurnal Ilmiah Ilmu Komputer*, vol. 16, no. 1, pp. 13-18, 2021.
- [8] S. Hajar, et al, "Penerapan K-Means Clustering pada ekspor minyak kelapa sawit menurut negara tujuan," *Seminar Nasional Teknologi Komputer & Sains (SAINTEKS) 2020*, pp. 314–318, 2020.
- [9] A. Sulistiyawati and E. Supriyanto, "Implementasi Algoritma K-means Clustering dalam Penentuan Siswa Kelas Unggulan," *Jurnal Tekno Kompak*, vol. 15, no. 2, p. 25, 2021.
- [10] R. Supardi and I. Kanedi, "Implementasi Metode Algoritma K-Means Clustering pada Toko Eidelweis," *Jurnal Teknologi Informasi*, vol. 4, no. 2, pp. 270–277, 2020.
- [11] R. P. Primanda, A. Alwi, and D. Mustikasari, "Data Mining Seleksi Siswa Berprestasi untuk Menentukan Kelas Unggulan Menggunakan Metode K-Means Clustering (Studi Kasus di MTS Darul Fikri)," *KOMPUTEK*, vol. 5, no. 1, pp. 88-100, 2021.
- [12] M. I. Zuhendra, R. Hidayat, H. Hendrawaty, "Penerapan Data Mining Untuk Klasterisasi Tingkat Kemiskinan Berdasarkan Data Terpadu Kesejahteraan Sosial (DTKS)," *SKANIKA (Sistem Komputer dan Teknik Informatika)*, vol. 7, pp. 32–41, 2024.
- [13] D. A. Suraya, Muhammad Sholeh, "Penerapan Metode Clustering Dengan Algoritma K-means Pada Pengelompokan Indeks Prestasi Akademik Mahasiswa," *SKANIKA (Sistem Komputer dan Teknik Informatika)*, vol. 6, no. 6, pp. 51–60, 2023.
- [14] D. P. Utomo and M. Mesran, "Analisis Komparasi Metode Klasifikasi Data Mining dan Reduksi Atribut Pada Data Set Penyakit Jantung," *Jurnal Media Informatika Budidarma*, vol. 4, no. 2, pp. 437-444, 2020.
- [15] I. T. Umagapi, et al., "Uji Kinerja K-Means Clustering Menggunakan Davies-Bouldin Index Pada Pengelompokan Data Prestasi Siswa," *Prosiding Seminar Nasional Sistem Informasi dan Teknologi (SISFOTEK) 7*, pp. 303–308, 2023.
- [16] F. Yunita, "Penerapan Data Mining Menggunakan Algoritma K-Means Clustering Pada Penerimaan Mahasiswa Baru," *Jurnal SISTEMASI*, vol. 7, no. 3, pp. 238-249, 2018.