



# Clustering Data Calon Siswa Baru Menggunakan Metode K-Means di Pusat Pengembangan Anak Fajar Baru Cengkareng

Kiki Setiawan <sup>1</sup>, Yulia Yanti Ayu Saputry <sup>2\*</sup>

<sup>1,2\*</sup> Program Studi Teknik Informatika, Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika, Kota Jakarta Timur, Daerah Khusus Ibu kota Jakarta, Indonesia.

## article info

### Article history:

Received 25 July 2023

Received in revised form

29 September 2023

Accepted 20 November 2023

Available online January 2024

### DOI:

<https://doi.org/10.35870/jti.k.v8i1.1426>

### Keywords:

K-Means; Clustering;  
Prospective New Student  
Data.

### Kata Kunci:

K-Means; Clustering; Data  
Calon Siswa Baru.

## abstract

Clustering is the process of partitioning a set of data objects into subsets known as clusters. K-means is an unsupervised learning algorithm, K-Means also has a function to group data into data clusters. The K-Means algorithm method was chosen because it has a fairly high accuracy of object size, so this algorithm is relatively more scalable and more efficient for processing large numbers of objects. In the world of education, in general, every new school year there will be something called registration of new prospective students, at the Fajar Baru Child Development Center, many prospective students are accepted from 3 years to 5 years old, therefore the authors hope that by using clustering data can easily group data so that it can make it easier to find the necessary data. By using the K-means algorithm method and using the RapidMiner application, it found 80% efficient results in grouping data.

## abstract

Clustering merupakan proses partisi satu set objek data ke dalam himpunan bagian yang disebut sebagai cluster. K-means merupakan salah satu algoritma yang bersifat unsupervised learning, K-Means juga memiliki fungsi untuk mengelompokkan data kedalam data cluster. Metode algoritma K-Means ini dipilih karena memiliki ketelitian yang cukup tinggi terhadap ukuran objek, sehingga algoritma ini relatif lebih terukur dan lebih efisien untuk pengolahan objek dalam jumlah yang besar. Dalam dunia pendidikan pada umumnya setiap tahun ajaran baru akan ada yang namanya pendaftaran calon siswa/siswi baru, pada Pusat Pengembangan Anak Fajar Baru banyak menerima calon siswa/siswi yang berusia mulai dari 3 tahun sampai 5 tahun, oleh karena itu penulis berharap dengan menggunakan clustering data dapat dengan mudah mengelompokkan data sehingga dapat memudahkan dalam menemukan data-data yang diperlukan. Dengan menggunakan metode algoritma K-means serta menggunakan aplikasi RapidMiner menemukan hasil yang efisien sebesar 80% dalam mengelompokkan data.

\*Corresponding Author. Email: [ayu.saputry.257@gmail.com](mailto:ayu.saputry.257@gmail.com) <sup>2\*</sup>.

© E-ISSN: 2580-1643.

Copyright © 2024 by the authors of this article. Published by Lembaga Otonom Lembaga Informasi dan Riset Indonesia (KITA INFO dan Riset). This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License. 

## 1. Latar Belakang

Pada era digital saat ini, data menjadi komoditas yang sangat berharga. Setiap organisasi, termasuk lembaga pendidikan, memiliki akses ke berbagai jenis data yang dihasilkan setiap hari. Namun, mengelola dan menganalisis data tersebut menjadi tantangan tersendiri. Dalam pendidikan, khususnya pada proses penerimaan calon siswa baru, mengelompokkan data calon siswa berdasarkan kriteria tertentu menjadi hal yang krusial. Inilah yang menjadikan teknik clustering, khususnya metode K-Means, menjadi sangat relevan dalam upaya pengelolaan data di Pusat Pengembangan Anak Fajar Baru Cengkareng.

Clustering merupakan salah satu teknik dalam analisis data yang bertujuan untuk mengelompokkan objek data ke dalam kelompok yang homogen berdasarkan kesamaan karakteristik. Dalam clustering, objek data yang serupa dikelompokkan bersama sedangkan objek yang berbeda dikelompokkan secara terpisah. Secara umum, proses clustering dapat dibagi menjadi dua jenis utama: clustering hierarkis dan clustering non-hierarkis. Clustering hierarkis menghasilkan struktur hierarkis berbentuk pohon yang menunjukkan hubungan antara cluster, sementara clustering non-hierarkis, seperti K-Means, mengelompokkan data menjadi sejumlah cluster yang telah ditentukan sebelumnya tanpa adanya struktur hierarkis. Clustering telah menjadi alat yang penting dalam berbagai bidang, termasuk ilmu komputer, statistik, dan ilmu sosial. Dalam pendidikan, clustering digunakan untuk berbagai tujuan, mulai dari pengelompokkan siswa berdasarkan kemampuan akademik hingga identifikasi pola perilaku siswa untuk peningkatan pembelajaran. Dengan menggunakan clustering, lembaga pendidikan dapat mengoptimalkan pengelolaan data siswa dan merancang strategi pendidikan yang lebih efektif.

Salah satu algoritma clustering yang paling populer dan banyak digunakan adalah algoritma K-Means. K-Means adalah metode clustering non-hierarkis yang bertujuan untuk membagi himpunan data menjadi  $k$  kelompok, di mana setiap data hanya tergabung dalam satu kelompok. Algoritma ini bekerja dengan cara menginisialisasi pusat-pusat cluster secara acak,

mengelompokkan data berdasarkan pusat-pusat tersebut, dan menghitung ulang pusat-pusat cluster berdasarkan rata-rata data dalam setiap cluster. Proses ini diulang hingga tidak ada perubahan signifikan dalam pusat-pusat cluster atau hingga jumlah iterasi maksimum tercapai. Keunggulan utama dari algoritma K-Means adalah kemampuannya dalam menangani data dalam jumlah besar dengan efisien. Algoritma ini juga relatif mudah dimengerti dan diimplementasikan, membuatnya menjadi pilihan yang populer di kalangan praktisi data mining dan analisis data. Namun, perlu diingat bahwa hasil clustering yang dihasilkan oleh K-Means sangat tergantung pada inisialisasi pusat-pusat cluster awal, sehingga dapat menghasilkan solusi yang berbeda-beda untuk inisialisasi yang berbeda.

Dalam Pusat Pengembangan Anak Fajar Baru Cengkareng, penggunaan algoritma K-Means untuk clustering data calon siswa baru memiliki banyak potensi manfaat. Dengan menggunakan teknik clustering, Pusat Pengembangan Anak dapat mengelompokkan calon siswa berdasarkan berbagai kriteria, seperti usia, latar belakang pendidikan, atau tingkat kemampuan. Hal ini dapat membantu dalam proses seleksi siswa baru, penempatan siswa dalam kelas yang sesuai dengan tingkat kemampuan mereka, serta perencanaan program pendidikan yang lebih terarah dan efektif. Selain itu, penggunaan algoritma K-Means juga dapat membantu dalam identifikasi pola-pola dalam data calon siswa yang mungkin tidak terdeteksi secara manual. Misalnya, dengan menggunakan teknik clustering, Pusat Pengembangan Anak dapat mengidentifikasi pola-pola korelasi antara karakteristik siswa dengan tingkat keberhasilan mereka dalam program pendidikan. Informasi ini dapat digunakan untuk merancang strategi pendidikan yang lebih personal dan adaptif sesuai dengan kebutuhan dan karakteristik individu siswa. Dengan demikian, penelitian ini bertujuan untuk menginvestigasi potensi penggunaan algoritma K-Means dalam clustering data calon siswa baru di Pusat Pengembangan Anak Fajar Baru Cengkareng. Melalui penelitian ini, diharapkan dapat ditemukan pola-pola yang bermanfaat dalam pengelolaan data siswa dan perancangan program pendidikan yang lebih efektif.

Penelitian terdahulu telah menunjukkan berbagai

aplikasi algoritma K-Means dalam berbagai kasus. Misalnya, Kamila *et al.* (2019) membandingkan algoritma K-Means dan K-Medoids untuk pengelompokan data transaksi bongkar muat di Provinsi Riau, dengan fokus pada mengurangi sensitivitas partisi yang dihasilkan terhadap nilai-nilai ekstrim dalam dataset [1]. Sementara itu, Parlina *et al.* (2018) memanfaatkan algoritma K-Means untuk menentukan pegawai yang layak mengikuti assessment center untuk clustering program SDP, menyoroti penggunaan pengelompokan untuk mengelompokkan dokumen atau benda yang tidak tersusun rapi [2]. Selain itu, Nishom (2019) melakukan perbandingan akurasi Euclidean Distance, Minkowski Distance, dan Manhattan Distance pada algoritma K-Means Clustering berbasis Chi-Square, dengan mengimplementasikan pengukuran jarak menggunakan metode Manhattan dan Minkowski [3]. Setyaningtyas *et al.* (2022) memberikan tinjauan pustaka sistematis tentang penerapan data mining teknik clustering algoritma K-Means, menekankan pentingnya metode tinjauan pustaka sistematis dalam mengidentifikasi dan mengevaluasi hasil penelitian [4]. Penelitian terkait juga telah dilakukan, seperti yang dilakukan oleh Satria dan Anggrawan (2021) yang mengembangkan aplikasi K-Means berbasis web untuk klasifikasi kelas unggulan, menunjukkan kesinambungan dalam penggunaan metode klasifikasi dengan algoritma K-Means [5]. Selain itu, Wahid *et al.* (2023) melakukan pengelompokan keberhasilan produksi peternak ayam broiler di Riau berdasarkan Index Performance menggunakan K-Means, menyoroti praktik pembudidayaan ayam broiler saat ini [6].

Sudi Suryadi (2018) menerapkan metode clustering K-Means untuk mengelompokkan kelulusan mahasiswa berdasarkan kompetensi, memberikan wawasan dalam pengelolaan data akademik [6]. Zulfa Nabila *et al.* (2021) menganalisis data Covid-19 di Provinsi Lampung menggunakan algoritma K-Means, memberikan kontribusi dalam manajemen data pandemi di tingkat regional [7]. Hani Prastiwi *et al.* (2022) mengimplementasikan K-Means Clustering dalam manajemen persediaan stok barang di mini market, memperkuat pemahaman dalam manajemen persediaan [8]. Sabrina Aulia Rahmah (2020) menerapkan K-Means Clustering untuk menganalisis

pola penjualan pestisida, memberikan wawasan penting dalam pengambilan keputusan di bidang penjualan [9]. Cornelia Selvi Dinta Sembiring *et al.* (2022) menggunakan K-Means untuk menentukan judul skripsi dan jurnal penelitian, memberikan kontribusi dalam pengelolaan data akademik [10].

Algoritma K-Means menawarkan kemampuan untuk mengelompokkan data berdasarkan kriteria tertentu, memungkinkan analisis yang lebih terarah dan efektif. Dengan aplikasi yang meluas, algoritma ini menemukan tempatnya dalam berbagai bidang, mulai dari pendidikan, kesehatan, hingga manajemen persediaan dan analisis penjualan. Diharapkan pengembangan lebih lanjut dan penerapan algoritma K-Means akan terus memberikan kontribusi yang signifikan dalam pengelolaan data dan pengambilan keputusan di berbagai bidang. Tujuan utama dari penelitian ini adalah untuk mengidentifikasi pola-pola dalam data calon siswa baru di Pusat Pengembangan Anak Fajar Baru Cengkareng menggunakan algoritma K-Means. Spesifiknya, penelitian ini akan mencoba mengelompokkan calon siswa berdasarkan beberapa kriteria tertentu, seperti usia, latar belakang pendidikan, dan tingkat kemampuan, untuk membantu dalam proses seleksi dan penempatan siswa baru. Penelitian ini juga akan menganalisis hasil clustering untuk mengidentifikasi pola-pola yang dapat digunakan untuk perancangan program pendidikan yang lebih efektif. Ruang lingkup penelitian ini akan terbatas pada penggunaan data calon siswa baru yang tersedia di Pusat Pengembangan Anak Fajar Baru Cengkareng. Data yang akan digunakan meliputi informasi tentang usia, latar belakang pendidikan, dan hasil tes kemampuan siswa. Penelitian ini tidak akan membahas proses pengumpulan data.

## 2. Metode Penelitian

### *Waktu dan Tempat Penelitian*

Penelitian ini dilakukan mulai bulan 1 April 2023 sampai 15 Mei 2023 yang bertempat di Pusat Pengembangan Anak Fajar Baru Cengkareng.

### Metode Pengumpulan Data

Metode pengumpulan data yang digunakan untuk mendapatkan data dari suatu informasi, maka metode yang digunakan dalam proses pengumpulan data sebagai berikut:

#### 1) Metode Observasi

Metode pengumpulan data yang akan dilakukan adalah melihat serta mempelajari permasalahan yang ada dilapangan yang erat kaitannya dengan objek yang diteliti.

#### 2) Metode Studi Pustaka

Metode yang dilakukan sebagai bahan pembelajaran dengan cara mencari bahan yang mendukung dalam pendefinisian masalah melalui buku-buku serta jurnal terdahulu. Pada penelitian ini penulis menggunakan metode penelitian yaitu Algoritma K-Means Clustering. Agar mencapai hasil tersebut, penulis akan melakukan pengujian dengan menggunakan data calon siswa baru pada Pusat Pengembangan Anak Fajar Baru Cengkareng. Sumber data dalam penelitian ini langsung diambil dari Pusat Pengembangan Anak Fajar Baru Fajar Baru Cengkareng dimana data tersebut merupakan data sekunder yang terdiri atas data calon siswa baru dengan banyak data 19 orang dari jumlah data seluruhnya. Data primer merupakan data yang dapat diperoleh dari para calon siswa secara langsung ke bagian objek yang sedang diteliti kemudian data yang telah diperoleh akan diolah dalam RapidMiner.

### Penerapan Metodologi

#### 1) Teknik pengumpulan data

Teknik pengumpulan data merupakan suatu cara yang dapat dilakukan oleh peneliti agar bisa memperoleh data-data yang diperlukan. Dalam penelitian ini memiliki beberapa teknik yang dapat menjadi acuan sebagai berikut:

##### a) Wawancara

Wawancara yang telah dilakukan pada yang menjadi penanggung jawab Pusat Pengembangan Anak Fajar Baru Cengkareng mengenai penerimaan calon siswa baru yang bertujuan untuk agar dapat mendalami permasalahan secara keseluruhan sehingga penelitian ini

memiliki konsep untuk memecahkan masalah.

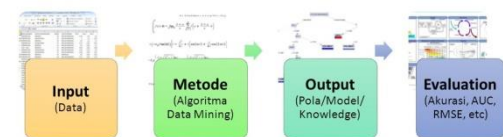
#### b) Survey

Survey penelitian dilakukan dengan meminta ijin kepada penanggung jawab atau koordinator yang bertanggung jawab atas semua data penerimaan calon siswa baru, serta data yang telah diperoleh akan dijadikan sebagai dataset dalam rapidminer.

#### c) Telaah dokumen

Telaah dokumen adalah pengumpulan data dengan cara mengumpulkan dan mempelajari dokumen-dokumen yang telah didapatkan, dijadikan sebagai data set untuk kelengkapan data set.

### Tahapan Utama Proses Data Mining



Gambar 1. Tahapan utama

#### 2) Proses Modeling

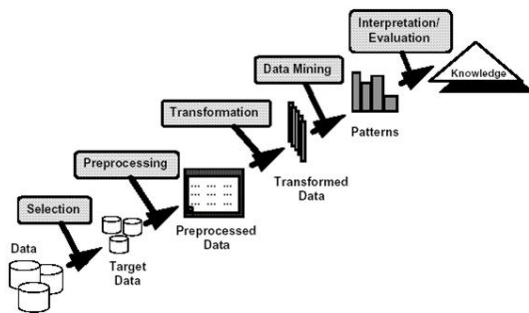
Memasuki proses penambahan data dengan menerapkan predictive modeling yang terdiri dari dua teknik, yakni classification dan value prediction. Adapula database segmentation, merupakan partisi database menjadi sejumlah bagian atau segmen, hingga cluster atau perekaman yang sama. Digunakan pula link analysis untuk menghubungkan record satu dengan lainnya.

#### 3) Evaluasi Data

Menggunakan teknik *deviation detection*, yakni identifikasi *outlier* dengan mengekspresikan deviasi dari ekspektasi yang ada. Kemudian dilakukan *nearest neighbour* yang dipakai untuk memprediksi pengelompokan, teknik tertua dalam data mining. Dan terakhir *clustering*, mengklasifikasikan data berdasarkan kriteria masing-masing.

#### 4) Penyajian Data

Digunakan model prediktif seperti *decision tree*, digambarkan layaknya seperti pohon karena setiap node di dalam struktur pohon mewakili sebuah pertanyaan yang digunakan untuk menggolongkan data bisa didapat dari jurnal data mining.



Gambar 2. Tahapan Utama.

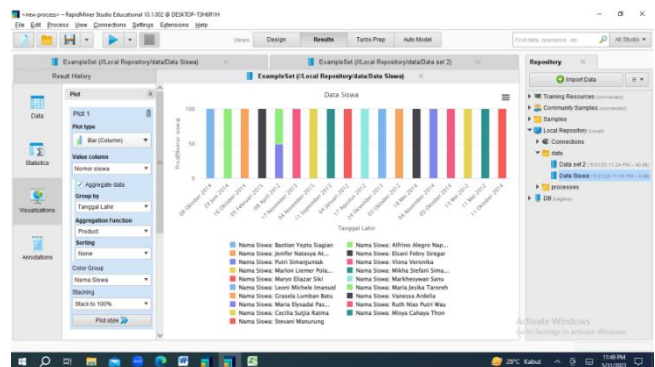
#### Rancangan Pengujian

Rencana pengujian ini bertujuan untuk mengevaluasi dampak data calon siswa baru terhadap kontribusi laporan bagi staf dan membandingkan hasil pengelompokan data menggunakan algoritma K-Means dengan pengelompokan manual menggunakan RapidMiner. Pengelompokan manual dilakukan dengan mengelompokkan data berdasarkan nama, usia, dan tanggal lahir untuk mempercepat proses kerja staf. Metode ini memudahkan staf dalam mengelompokkan data dengan melihat warna dan persentase, sehingga mempercepat proses pengolahan data. Contoh data siswa lama atau alumni yang telah menyelesaikan pendidikan di Pusat Pengembangan Anak Fajar Baru Cengkareng juga dapat diintegrasikan, memudahkan pengumpulan data alumni. Tampilan grafis data memungkinkan informasi yang dibutuhkan dapat ditemukan dengan cepat dan memberikan ringkasan informasi, menghemat waktu dalam menentukan fokus analisis.

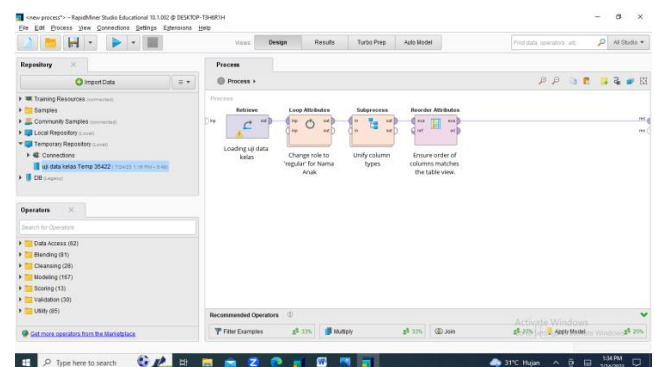
### 3. Hasil dan Pembahasan

Pengujian komponen internal laptop merupakan tahap penting dalam evaluasi kinerja perangkat keras. Langkah pertama yang dilakukan adalah ekstraksi data menggunakan Microsoft Excel untuk

menyediakan data yang akan dianalisis. Setelah data berhasil diekstraksi, langkah selanjutnya adalah mengkonversikannya ke dalam aplikasi RapidMiner, sebuah platform analisis data yang dapat melakukan berbagai macam analisis dengan berbagai algoritma, termasuk algoritma K-Means untuk pengelompokan data. Penggunaan RapidMiner membantu memudahkan proses analisis data dengan pengelompokan yang lebih akurat. Setelah data berhasil diimpor ke dalam RapidMiner, proses pengelompokan data dilakukan dengan menggunakan algoritma K-Means. Hasil dari pengelompokan ini kemudian divisualisasikan untuk memahami pola data yang ada. Visualisasi clustering yang dilakukan melibatkan berbagai kriteria, seperti tanggal lahir, jenis kelamin, kelas, dan umur anak, yang membantu memperjelas pola data yang ada dan memudahkan analisis lebih lanjut terhadap setiap kelompok data.



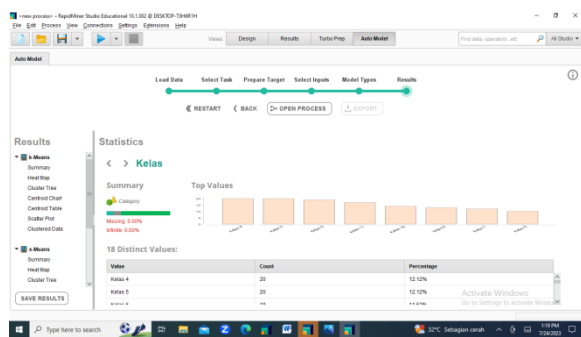
Gambar 3. Clustering berdasarkan tanggal lahir



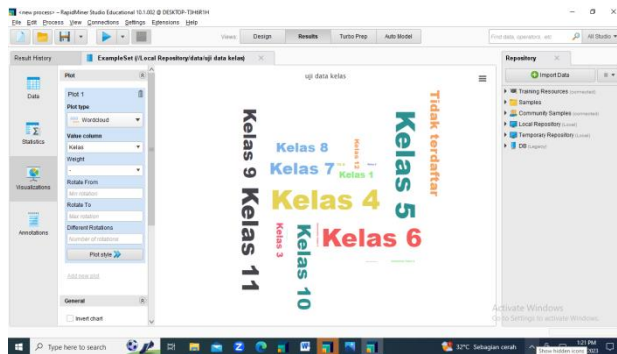
Gambar 4. Proses dalam menghubungkan data

Data yang telah di-clusterkan dapat digunakan sesuai kebutuhan, sedangkan statistik dalam pengolahan data memberikan pemahaman mendalam terhadap pola data yang ada.





Gambar 5. Statistics dalam clustering data anak



Gambar 6. Visualisasi data kelas anak

Visualisasi clustering berdasarkan tanggal lahir membantu untuk melihat bagaimana data terdistribusi secara temporal. Hal ini dapat memberikan wawasan yang berharga terkait dengan tren atau pola yang mungkin terjadi seiring berjalannya waktu. Misalnya, apakah ada korelasi antara bulan kelahiran dan kecenderungan dalam perilaku atau performa anak-anak. Selanjutnya, visualisasi berdasarkan jenis kelamin dapat memberikan pemahaman tentang perbedaan dalam distribusi data antara laki-laki dan perempuan. Analisis ini dapat memberikan wawasan yang berguna terkait dengan preferensi, kebiasaan, atau karakteristik khusus dari masing-masing jenis kelamin yang mungkin mempengaruhi keputusan atau kebijakan yang akan diambil. Selain itu, visualisasi berdasarkan kelas dapat membantu untuk memahami bagaimana distribusi data di antara berbagai kelompok kelas. Ini dapat membantu dalam mengidentifikasi kelompok kelas yang mungkin membutuhkan perhatian khusus atau mendapatkan manfaat dari pendekatan atau intervensi tertentu.

Visualisasi berdasarkan umur anak dapat memberikan pemahaman tentang bagaimana distribusi data berubah seiring bertambahnya usia anak-anak. Analisis ini dapat membantu dalam

memahami perkembangan atau perubahan yang mungkin terjadi seiring waktu dan memberikan wawasan tentang bagaimana pendekatan atau strategi yang tepat dapat disesuaikan dengan kebutuhan atau karakteristik dari berbagai kelompok usia. Selain visualisasi, statistik dalam pengolahan data juga memberikan pemahaman mendalam tentang pola data yang ada. Analisis statistik ini dapat mencakup berbagai metrik, seperti rata-rata, median, simpangan baku, dan lain-lain, yang membantu dalam mengidentifikasi pola, tren, atau anomali dalam data. Ini dapat memberikan informasi yang berharga dalam pengambilan keputusan, perencanaan, atau evaluasi program atau kebijakan.

Dengan menggunakan aplikasi RapidMiner, proses analisis data menjadi lebih efisien dan efektif. Pengguna dapat dengan mudah mengimpor data, menerapkan berbagai analisis, dan memvisualisasikan hasilnya dalam format yang informatif dan mudah dipahami. Hal ini memungkinkan pengguna untuk mendapatkan wawasan yang lebih dalam tentang data mereka dan membuat keputusan yang lebih baik berdasarkan informasi yang diberikan oleh data. Langkah-langkah menggunakan aplikasi RapidMiner relatif sederhana dan mudah dipahami. Setelah data diimpor, pengguna dapat memilih operator-operator yang sesuai untuk melakukan analisis data yang diinginkan. Selanjutnya, pengguna dapat melihat hasil analisis dalam bentuk visualisasi yang informatif dan memudahkan interpretasi. Dengan demikian, aplikasi RapidMiner menjadi alat yang sangat berguna dalam analisis data dan pengambilan keputusan. Pengujian komponen internal laptop menggunakan RapidMiner merupakan langkah yang efektif dalam menganalisis dan mengelompokkan data. Visualisasi dan statistik yang dihasilkan membantu pengguna untuk memahami pola data yang ada dan membuat keputusan yang lebih baik berdasarkan informasi yang diberikan oleh data. Diharapkan penggunaan aplikasi ini dapat terus dikembangkan dan diterapkan dalam berbagai bidang untuk meningkatkan efisiensi dan efektivitas pengelolaan data.

#### *Analisis dan Interpretasi*

Dari hasil pengujian, terlihat bahwa penggunaan aplikasi RapidMiner dalam pengelompokkan data memberikan hasil yang memuaskan. Visualisasi

clustering memungkinkan pengguna untuk dengan cepat dan mudah memahami pola-pola yang ada dalam data. Selain itu, statistik yang dihasilkan juga memberikan wawasan yang mendalam tentang distribusi data dan karakteristik kelompok data yang terbentuk. Dalam pengelolaan data anak, pengelompokan berdasarkan kriteria seperti jenis kelamin, kelas, dan umur anak memungkinkan para pengelola data untuk lebih memahami karakteristik setiap kelompok anak. Hal ini dapat membantu dalam pengambilan keputusan yang lebih tepat, misalnya dalam penyusunan program pendidikan atau kegiatan ekstrakurikuler yang sesuai dengan karakteristik anak-anak dalam setiap kelompok. Selanjutnya, penggunaan aplikasi RapidMiner tidak hanya memudahkan dalam proses pengelompokan data, tetapi juga membantu dalam analisis lebih lanjut. Dengan menggunakan operator-operator yang tersedia, pengguna dapat melakukan analisis statistik lebih lanjut, membangun model prediktif, atau mengidentifikasi pola-pola yang menarik dalam data.

#### 4. Kesimpulan dan Saran

Berdasarkan pembahasan dapat disimpulkan bahwa teknologi memang sangat penting bagi kehidupan sehari-hari, banyak hal yang masih menjadi kekurangan dalam pengolahan data ini. Penulis berharap agar semua program bisa berjalan dengan baik dan bermanfaat bagi masyarakat luas. Dalam penulisan jurnal ini banyak hal yang dialami kesan baik yang dapat penulis ambil yaitu dapat mengurangi tingkat kekeliruan serta kesalahan dalam mengelompokkan data yang lebih efisien, tepat serta cepat. Kekurangan yang dialami yaitu pada saat salah menempatkan tipe pada kolom akan tidak terbaca oleh sistem aplikasi yang ada didalam RapidMiner. Saran yang penulis sampaikan adalah dengan cara memilih tipe kolom yang sesuai dengan yang dibutuhkan sehingga tidak menciptakan eror pada sistem yang sudah ada.

Pengujian yang telah dilakukan yaitu seperti mengelompokkan data berdasarkan nama, kelas, jenis kelamin, dan usia karena ini yang sangat dibutuhkan dalam administrasi saat ini pada Pusat Pengembangan Anak Fajar Baru Cengkareng dan

dapat diambil kesimpulan seperti :

- 1) Metode K-Means sangat ideal serta lebih efektif dalam mengelompokkan data sehingga memudahkan menemukan kelompok data secara cepat dan tepat.
- 2) Aplikasi RapidMiner juga sangat mudah digunakan untuk mengelompokkan data atau menclustering data.

Penulis sangat berharap dengan adanya metode pengelompokkan data yang lebih efisien ini dapat mengurangi tingkat kesalahan pada saat mengelompokkan data-data yang ada di lapangan dan penulis berharap bisa lebih mudah untuk menggunakan aplikasi ini untuk kebutuhan di Pusat Pengembangan Anak Fajar Baru Cengkareng menggunakan metode K-Means untuk clustering data-data yang lain untuk kedepannya, karena aplikasi yang digunakan tidak memakan banyak penyimpanan sehingga bisa di instal dengan spesifikasi PC atau laptop yang standar. Untuk hasil dengan menggunakan metode K-Means dalam mengclustering data sebesar 80 % untuk pengelompokkan datanya.

#### 5. Daftar Pustaka

- [1] Nasir, J., 2020. Penerapan Data Mining Clustering Dalam Mengelompokkan Buku Dengan Metode K-Means. *Jurnal Simetris*, 11(2), pp.1-13.
- [2] Purnamaningsih, C., 2013. Pemanfaatan Metode K-Means Clustering dalam Penentuan Penjurusan Siswa SMA. UPT Perpustakaan Universitas Sebelas Maret.
- [3] Damanik, Y.F.S.Y., Sumarno, S., Gunawan, I., Hartama, D. and Kirana, I.O., 2021. Penerapan Data Mining Untuk Pengelompokan Penyebaran Covid-19 Di Sumatera Utara Menggunakan Algoritma K-Means. *Jurnal Ilmu Komputer Dan Informatika*, 1(2). DOI: <https://doi.org/10.54082/jiki.13>.

- [4] Damanik, N. and Sigiro, M., 2021. Penerapan data mining menggunakan algoritma K-Means clustering pada penerimaan mahasiswa baru sebagai metode promosi. *J. Tek. Inform. Komput. Univers*, 4(2), p.158.
- [5] Rohmah, A., Sembiring, F. and Erfina, A., 2021, September. Implementasi Algoritma K-Means Clustering Analysis Untuk Menentukan Hambatan Pembelajaran Daring (Studi Kasus: Smk Yaspim Gegerbitung). In *Prosiding Seminar Nasional Sistem Informasi dan Manajemen Informatika Universitas Nusa Putra* (Vol. 1, No. 01, pp. 290-298).
- [6] Febrivani, E. and Winanjaya, R., 2021. Penerapan Data Mining Asosiasi Pada Persediaan Obat. *Jurnal Ilmu Komputer Dan Sistem Informasi (JIKOMSI)*, 4(1), pp.23-35. DOI: <https://doi.org/10.9767/jikomsi.v4i1.141>.
- [7] Suryadi, S., 2018. Penerapan Metode Clustering K-Means Untuk Pengelompokan Kelulusan Mahasiswa Berbasis Kompetensi. *INFORMATIKA*, 6(1), pp.52-72. DOI: <https://doi.org/10.36987/informatika.v6i1.738>.
- [8] Nabila, Z., Isnain, A.R., Permata, P. and Abidin, Z., 2021. Analisis data mining untuk clustering kasus covid-19 di Provinsi Lampung dengan algoritma k-means. *Jurnal Teknologi Dan Sistem Informasi*, 2(2), pp.100-108. DOI: <https://doi.org/10.33365/jtsi.v2i2.868>.
- [9] Prastiwi, H., Pricilia, J. and Rasywir, E., 2022. Implementasi Data Mining Untuk Menentukan Persediaan Stok Barang Di Mini Market Menggunakan Metode K-Means Clustering. *Jurnal Informatika Dan Rekayasa Komputer (JAKAKOM)*, 2(1), pp.141-148. DOI: <https://doi.org/10.33998/jakakom.2022.2.134>.
- [10] Aulia, S., 2020. Klasterisasi Pola Penjualan Pestisida Menggunakan Metode K-Means Clustering (Studi Kasus Di Toko Juanda Tani Kecamatan Hutabayu Raja). *Djtechno: Jurnal Teknologi Informasi*, 1(1), pp.1-5. DOI: <https://doi.org/10.46576/djtechno.v1i1.964>.
- [11] Sembiring, C.S.D.B., Hanum, L. and Tamba, S.P., 2022. Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Menentukan Judul Skripsi Dan Jurnal Penelitian (Studi Kasus Ftik Unpri). *Jurnal Sistem Informasi dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, 5(2), pp.80-85. DOI: <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2393>.
- [12] Handoko, S., Fauziah, F. and Handayani, E.T.E., 2020. Implementasi Data Mining Untuk Menentukan Tingkat Penjualan Paket Data Telkomsel Menggunakan Metode K-Means Clustering. *Jurnal Ilmiah Teknologi dan Rekayasa*, 25(1), pp.76-88. DOI: <http://dx.doi.org/10.35760/tr.2020.v25i1.2677>.
- [13] Asmana, A., Wijaya, Y.A. and Martanto, M., 2022. Clustering Data Calon Siswa Baru Menggunakan Metode K-Means Di Sekolah Menengah Kejuruan Wahidin Kota Cirebon. *JATI (Jurnal Mahasiswa Teknik Informatika)*, 6(2), pp.552-559. DOI: <https://doi.org/10.36040/jati.v6i2.5236>.
- [14] Ningrat, D.R., Di Asih, I.M. and Wuryandari, T., 2016. Analisis cluster dengan algoritma K-Means dan Fuzzy C-Means clustering untuk pengelompokan data obligasi korporasi. *Jurnal Gaussian*, 5(4), pp.641-650. DOI: <https://doi.org/10.14710/j.gauss.5.4.641-650>.
- [15] Priyatman, H., Sajid, F. and Haldivany, D., 2019. Klasterisasi Menggunakan Algoritma K-Means Clustering untuk Memprediksi Waktu Kelulusan Mahasiswa. *Jurnal Edukasi Dan Penelitian Informatika (JEPIN)*, 5(1), p.62.



- [16] Intel., 2023. Meningkatkan Intelijen Bisnis dengan Analisis Dalam Memori. URL: <https://www.intel.co.id/content/www/id/id/artificial-intelligence/in-memory-analytics.html>. Diakses Tanggal 18 Juni 2023.
- [17] Yulianti, Y., Utami, D.Y., Hikmah, N. and Hasan, F.N., 2019. Penerapan Data Mining Menggunakan Algoritma K-Means Untuk Mengetahui Minat Customer Di Toko Hijab. *Jurnal Pilar Nusa Mandiri*, 15(2), pp.241-246. DOI: <https://doi.org/10.33480/pilar.v15i2.650>.
- [18] Chiang, H.D., Xu, T.S., Lv, X.L. and Dong, N., 2022. Hierarchical Trust-Tech-Enhanced K-Means Methods and Their Applications to Power Grids. *IEEE Open Access Journal of Power and Energy*, 9, pp.560-572. DOI: <https://doi.org/10.1109/OAJPE.2022.3230385>.
- [19] Purnama, J.J. and Rahayu, S., 2022. Klasifikasi konsumsi energi industri baja menggunakan teknik data mining. *Jurnal Teknoinfo*, 16(2), pp.395-407. DOI: <https://doi.org/10.33365/jti.v16i2.1984>.
- [20] Tamba, S.P., 2022. Penerapan Data Mining Algoritma Apriori Dalam Menentukan Stok Bahan Baku Pada Restoran Nelayan Menggunakan Metode Association Rule. *Jurnal Sistem Informasi dan Ilmu Komputer Prima (JUSIKOM PRIMA)*, 5(2), pp.97-102. DOI: <https://doi.org/10.34012/jurnalsisteminformasidanilmukomputer.v5i2.2407>.