

OPTIMASI ANALISIS *CLUSTERING* UNTUK AKTIVITAS DAN RESPON PENGGUNA MEDIA SOSIAL DENGAN *K-MEANS*

Mela Nur Afrilia¹, Nining Rahaningsih², Raditya Danar Dana³, Nisa Dienwati Nuris⁴

^{1,4} Teknik Informatika, ² Komputerisasi Akuntansi, ³ Manajemen Informatika, STMIK IKMI Cirebon
Jln Perjuangan No 10 B Karyamulya Kesambi Kota Cirebon Jawa Barat Indonesia
melaafria77@gmail.com

ABSTRAK

Media sosial sebagai *platform* utama interaksi dan aktivitas *online* memiliki peran sentral dalam membentuk pola perilaku pengguna. Analisis *clustering* aktivitas dan respons pengguna di media sosial menjadi penting untuk pengembangan strategi pemasaran dan manajemen *platform* yang efektif. Dalam lingkungan yang semakin kompleks ini, diperlukan pendekatan analisis yang sistematis. Penelitian ini bertujuan mendalami perilaku pengguna media sosial melalui algoritma *K-Means* dan *Davies Bouldin Index* (DBI) untuk mengoptimalkan jumlah *cluster*, sehingga memperkaya pemahaman pola aktivitas dan respons pengguna. Dengan menggunakan *RapidMiner*, penelitian berhasil mengidentifikasi tiga *cluster* optimal, dengan *Cluster 1* menonjol sebagai yang paling dominan, diisi oleh pengguna yang menunjukkan tingkat aktivitas dan respons lebih tinggi dibandingkan kelompok lainnya. Temuan ini dapat memberikan panduan berharga bagi pengembangan strategi media sosial yang lebih tepat sasaran.

Kata kunci : *Clustering, Aktivitas dan Respon, Media Sosial, Algoritma K-Means*

1. PENDAHULUAN

Dalam era perkembangan teknologi digital, penggunaan media sosial tumbuh pesat dan memengaruhi interaksi sosial manusia. Fenomena ini telah mengubah cara komunikasi, pertukaran informasi, dan pembentukan hubungan sosial dilakukan [1]. Di era akses internet yang mudah, pengguna media sosial dapat dengan cepat berinteraksi dan mengikuti perkembangan terkini. Meski demikian, terdapat risiko pengguna terjebak dalam kegiatan tanpa tujuan, menghabiskan waktu berharga di platform tersebut. Penelitian ini menggunakan analisis *cluster* dengan algoritma *K-Means* untuk menggali variabilitas perilaku pengguna media sosial. Dengan memahami pola aktivitas dan respon pengguna, penelitian ini bertujuan memberikan wawasan tentang cara pengguna media sosial berinteraksi dengan platform tersebut. Sorotan khusus juga diberikan pada respon emosional pengguna terhadap interaksi di media sosial.

Penggunaan media sosial, fenomena mendominasi kehidupan sehari-hari, memberikan peluang interaksi dan berbagi informasi, namun dapat menimbulkan permasalahan emosional akibat penggunaan berlebihan. Fitur menariknya memfasilitasi komunikasi jarak jauh, tetapi dampak negatifnya melibatkan potensi kecanduan. Tantangan di bidang informatika terletak pada identifikasi pola perilaku pengguna media sosial, termasuk intensitas penggunaan dan dampak emosional. Kategorisasi pola perilaku menjadi permasalahan utama, di mana pengguna cenderung menghabiskan waktu tanpa tujuan yang jelas, menyebabkan pemborosan waktu berharga. Respon emosional pengguna menciptakan kompleksitas yang perlu diidentifikasi dan dianalisis, menyoroti tantangan dalam memahami variasi aktivitas dan respon di media sosial [2].

Penelitian sebelumnya yang berjudul “Hubungan Tingkat Kepercayaan Diri Dengan Intensitas Penggunaan Media Sosial *Whatsapp* Pada Mahasiswa Kedokteran 2018” [3] menunjukkan bahwa pengguna yang menghabiskan lebih banyak waktu di media sosial cenderung mengalami peningkatan tingkat kepercayaan diri, terutama melalui perbandingan dengan orang lain di platform tersebut.

Namun, masih diperlukan analisis lebih mendalam untuk memahami variasi perilaku dan respon emosional pengguna. Beberapa aspek yang perlu dipertimbangkan meliputi rata-rata waktu penggunaan dan kecenderungan mengakses media sosial tanpa tujuan yang jelas [1]. Selain itu, penelitian terkait kualitas tidur menunjukkan bahwa pengguna media sosial mungkin mengalami masalah tidur. Namun, penelitian ini belum mempertimbangkan berbagai atribut aktivitas dan respon pengguna yang dapat mempengaruhi pola tidur, sehingga diperlukan penelitian lebih lanjut untuk menggali hubungan yang lebih dalam.

Di samping itu, [4] melibatkan analisis sentimen masyarakat terhadap *K-Popers* di Twitter dengan menggunakan metode *clustering K-Means*. Hasil penelitian ini memberikan gambaran yang signifikan untuk memecahkan masalah dalam pengelompokan data, khususnya dalam mengidentifikasi kecenderungan opini masyarakat terhadap *K-Pop*, apakah termasuk dalam sentimen positif, negatif, atau netral. Penelitian ini memberikan kontribusi penting dalam memahami persepsi dan pandangan masyarakat terhadap fenomena *K-Pop* di media sosial.

Penelitian ini bertujuan mendalam memahami pola perilaku pengguna media sosial dengan menganalisis variasi dalam aktivitas dan respon emosional pengguna. Melalui identifikasi dan pengelompokan pengguna ke dalam *cluster-cluster*

berdasarkan karakteristik ini, penelitian ini bertujuan memberikan pemahaman yang lebih mendalam tentang bagaimana interaksi individu dengan media sosial dapat mempengaruhi perilaku dan kesejahteraan emosional mereka. Lebih lanjut, penelitian ini memiliki tujuan untuk memberikan dasar yang dapat membantu individu dalam mengelola penggunaan media sosial mereka dengan lebih bijaksana dan sehat. Signifikannya, penelitian ini mengisi kesenjangan pengetahuan dalam literatur mengenai perilaku pengguna media sosial melalui pendekatan analisis *cluster*. Hasil penelitian ini tidak hanya akan memberikan kontribusi teoritis pada bidang informatika, khususnya dalam studi perilaku manusia di era digital, tetapi juga dapat memberikan wawasan praktis kepada para ahli psikologi, pendidik, dan praktisi kesehatan mental. Dengan demikian, penelitian ini memiliki potensi besar untuk memberikan manfaat yang signifikan bagi individu, masyarakat, dan penelitian lanjutan di bidang psikologi komunikasi dan ilmu komputer.

Dalam mengatasi kompleksitas perilaku pengguna media sosial, penelitian ini akan menerapkan pendekatan analisis *cluster* dengan menggunakan algoritma *K-Means*. Algoritma *K-Means*, dikenal karena kinerjanya yang baik akan digunakan untuk mengelompokkan aktivitas dan respon pengguna media sosial [5]. Pendekatan ini bertujuan mengklasifikasikan pengguna ke dalam *cluster-cluster* homogen berdasarkan atribut aktivitas dan respon mereka. Tahapan awal penelitian melibatkan pencarian data terkait atribut aktivitas dan respons pengguna dalam konteks media sosial.

Setelah data terkumpul, peneliti kemudian menganalisis dan mengelompokkan pola perilaku serupa di antara pengguna media sosial. Dengan demikian, pendekatan ini diharapkan memberikan pemahaman lebih dalam tentang variasi perilaku di *platform* media sosial. Melalui pengelompokan yang dihasilkan, diharapkan penelitian ini dapat mengidentifikasi kelompok-kelompok homogen, membuka jalan bagi interpretasi yang lebih baik mengenai interaksi dan respons pengguna di media sosial. Langkah-langkah ini menggambarkan komitmen penelitian untuk mengartikan pola perilaku pengguna media sosial dengan pendekatan analisis yang terpercaya dan relevan.

Dengan menganalisis pola aktivitas dan respons pengguna media sosial melalui algoritma *K-Means*, penelitian ini memberikan wawasan mendalam tentang perilaku di dunia digital. Implikasinya mencakup potensi peningkatan kesadaran digital dan pengelolaan waktu yang lebih bijak bagi pengguna media sosial. Pemahaman yang lebih baik juga membantu mengenali tanda-tanda kegelisahan atau masalah tidur, memungkinkan pengguna untuk mengurangi akses tanpa tujuan. Analisis respons emosional mendukung perancangan intervensi yang tepat, membantu individu mengatasi tantangan di dunia digital. Temuan ini berpotensi memengaruhi

kebijakan media sosial, mendorong pengembangan teknologi berfokus pada kesejahteraan pengguna, dan meluaskan pemahaman akademis tentang interaksi manusia dengan teknologi digital.

Penggunaan *K-Means* dalam penelitian ini dipilih untuk mengelompokkan pengguna media sosial berdasarkan pola aktivitas dan respon, memungkinkan identifikasi kelompok serupa. *K-Means* efektif untuk analisis perilaku dengan memisahkan data menjadi kelompok homogen. Dengan *K-Means*, penelitian mendapatkan wawasan mendalam tentang perilaku pengguna dan memahami respons serupa. Kelebihan *K-Means* dalam menangani data besar membantu analisis aktivitas kompleks. Penggunaan *Davies Bouldin Index* (DBI) dengan *K-Means* membantu mengoptimalkan jumlah *cluster* secara otomatis, memudahkan analisis data. Oleh karena itu, *K-Means* dipilih untuk mencapai tujuan penelitian dalam mengungkap pola perilaku pengguna media sosial dengan pendekatan clustering yang sistematis.

2. TINJAUAN PUSTAKA

2.1. Media Sosial

Media sosial adalah bentuk media online yang mencakup *blog*, jejaring sosial, wiki, dan dunia virtual, memungkinkan partisipasi, berbagi, dan penciptaan konten pengguna. Andreas Kaplan dan Michael Haenlein mendefinisikannya sebagai kelompok aplikasi internet di atas ideologi dan teknologi Web 2.0, memfasilitasi pertukaran konten pengguna. Jejaring sosial, seperti Facebook dan Twitter, memungkinkan pembuatan halaman web pribadi dan berbagi informasi. Media sosial menggunakan internet, mengundang partisipasi terbuka dan umpan balik. Dengan perkembangan teknologi internet dan perangkat seluler, media sosial tumbuh pesat, mengubah peran media massa, dan mempercepat penyebaran informasi global. Ciri khasnya melibatkan pengiriman pesan ke banyak orang tanpa *gatekeeper*, penyebaran cepat, dan pengguna memiliki kendali atas waktu interaksi, mendorong partisipasi aktif dan berbagi informasi secara terbuka [6].

2.2. Machine Learning

Machine learning, cabang Ilmu Komputer terkait kecerdasan buatan, berfokus pada pengembangan sistem yang belajar dari data. Data merupakan aspek krusial dalam algoritma *machine learning*. Terbagi menjadi data *training* dan data *testing*, data *training* digunakan untuk mengembangkan model, sedangkan data *testing* digunakan untuk evaluasi performansi algoritma terlatih pada data baru. Dengan memanfaatkan data, *machine learning* meningkatkan kemampuannya memahami pola, membuat prediksi, dan belajar dari pengalaman [7].

2.3. Data Mining

Data mining merupakan proses kompleks yang menggunakan statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstrak

informasi berharga dari *database* besar. Sebagaimana dinyatakan oleh [8], *data mining* bukan hanya mengenai pengumpulan data, tetapi juga melibatkan analisis dan prediksi informasi yang ingin ditampilkan. Informasi yang terkumpul diolah dalam *database* untuk mendukung pengambilan keputusan. [7] merinci tahap-tahap *data mining*, termasuk pembersihan data, integrasi data dari berbagai sumber, pemilihan data yang relevan, transformasi data, proses *mining*, evaluasi pola, dan penyajian pengetahuan. Tahap akhir melibatkan formulasi keputusan berdasarkan hasil analisis yang diperoleh.

2.4. Algoritma K-Means

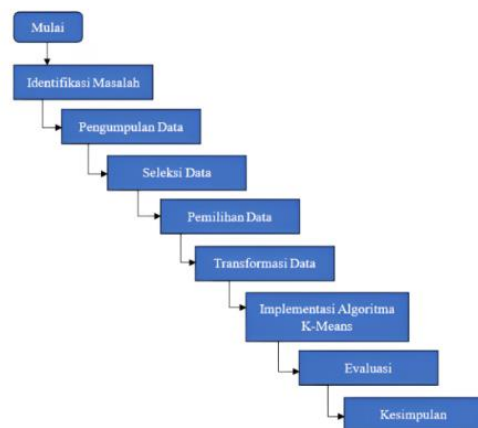
Algoritma *K-Means*, sebagaimana dijelaskan dalam teks, merupakan suatu algoritma teknik *clustering* yang dimulai dengan pemilihan secara acak terhadap *cluster* yang diinginkan untuk membentuk kelompok dari data yang akan di-*cluster*. *K-Means* dikenal sebagai salah satu algoritma paling populer dalam *clustering* karena sifatnya yang sederhana dan efisien, seperti yang diakui oleh penelitian [9] yang menempatkannya sebagai salah satu dari 10 algoritma *data mining* teratas. Algoritma ini efektif dalam mengelompokkan data ke dalam *cluster-cluster* yang saling berdekatan, memfasilitasi analisis pola dan struktur data dengan cara yang mudah dimengerti dan diterapkan [10].

2.5. Clustering

Clustering adalah metode efektif untuk mengorganisir sejumlah besar dokumen teks yang tidak berurutan menjadi kluster yang bermakna dan koheren. Tujuannya adalah menciptakan struktur yang memungkinkan navigasi dan penelusuran informasi yang intuitif serta informatif. Proses ini membantu dalam mengidentifikasi keterkaitan dokumen, mempermudah pemahaman konten terkait, dan meningkatkan efisiensi akses terhadap informasi yang tersebar luas [11].

3. METODE PENELITIAN

3.1. Metode Penelitian



Gambar 1. Tahapan metode penelitian

Gambar 1 menunjukkan tahapan-tahapan dari proses metode KDD (*Knowledge Discovery in Databases*) yang diterapkan dalam penelitian ini.

Tabel 1. Tahapan Penelitian

No	Tahapan	Aktivitas	Deskripsi Aktivitas
1.	Identifikasi Masalah	<i>Research Problem</i>	Menentukan latar belakang masalah, rumusan masalah, parameter, dan solusi masalah. Serta mencari metode yang cocok untuk digunakan dalam penelitian ini.
2.	Metode Penelitian	<i>Source Data</i>	Mengumpulkan dengan cara mencari data dan juga literatur-literatur yang menyangkut pembahasan pada penelitian ini
		<i>Data Selection</i>	Menseleksi data yang sudah terkumpul untuk digunakan sebagai proses <i>data mining</i>
		<i>Data Preprocessing</i>	Membuang data yang sudah tidak terpakai, memeriksa data dan memperbaiki data jika terjadi kesalahan
		<i>Data Transformation</i>	Pada fase ini dilakukan proses transformasi bentuk data yang belum jelas ke dalam bentuk data yang valid
		<i>Data Mining</i>	Dilakukan analisis data dengan menggunakan Algoritma <i>K-Means cluster</i>
3.	Kesimpulan	Evaluasi	Mencari tingkat performa terbaik
		Pengetahuan dan Informasi	Fase terakhir yaitu dilakukan proses penarikan kesimpulan dari hasil yang telah didapatkan

Tabel 1 di atas menjelaskan secara rinci aktivitas pada setiap tahapan penelitian yang dilakukan oleh peneliti.

3.2. Sumber Data

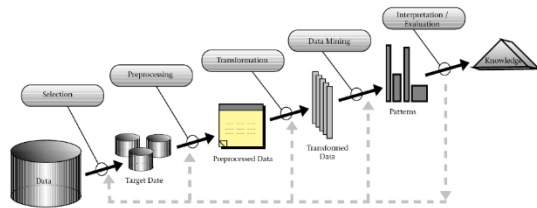
Data yang digunakan dalam penelitian ini berasal dari data publik di *Kaggle*. Sumber data *Kaggle* menyajikan informasi yang beragam terkait perilaku pengguna media sosial, mencakup atribut aktivitas pengguna dan respons pengguna.

3.3. Teknik Pengumpulan Data

Dalam pelaksanaan penelitian ini, pengumpulan data dilakukan dengan mengakses dataset yang tersedia di *Kaggle* melalui situs web <https://www.kaggle.com/>. Dataset yang digunakan mencakup berbagai variabel yang memberikan informasi terkait aktivitas pengguna media sosial, seperti waktu penggunaan dan respons pengguna. Proses pengumpulan data melibatkan pencarian

dataset, pemeriksaan kualitas data, dan pengunduhan dataset untuk analisis lebih lanjut.

3.4. Teknik Analisis Data



Gambar 2. Proses KDD (Knowledge Discovery in Database)

Gambar 2 menggambarkan teknik analisis dalam penelitian ini menggunakan metode *Knowledge Discovery in Database* (KDD) dengan tahapan sebagai berikut:

- Seleksi Data (Data Selection):** Dalam penelitian ini, seleksi data melibatkan pemilihan atribut yang relevan untuk analisis aktivitas dan respon pengguna media sosial. Contohnya, pemilihan *platform* media sosial, aktivitas pengguna, dan respon pengguna sebagai atribut yang menjadi fokus penelitian.
- Pembersihan Data (Data Cleaning):** Pada tahap pembersihan data, dilakukan eliminasi data yang tidak konsisten atau tidak relevan. Misalnya, data-data yang memiliki kesalahan atau format yang tidak sesuai dihapus dari *dataset*, sehingga *dataset* menjadi bersih dan siap untuk proses selanjutnya.
- Transformasi Data (Data Transformation):** Proses transformasi data melibatkan konversi atribut nominal menjadi bentuk numerik dan perubahan nama atribut. Contoh konkretnya adalah menggunakan operator *Rename* pada *RapidMiner* untuk mengganti nama atribut yang panjang menjadi lebih singkat dan lebih mudah dipahami.
- Data Mining:** Tahap *data mining* pada penelitian ini mencakup penggunaan algoritma *K-Means Clustering* untuk mengelompokkan pengguna ke dalam *cluster* berdasarkan aktivitas dan respon mereka. Ini termasuk proses mengidentifikasi nilai *cluster* terbaik menggunakan DBI.
- Evaluasi:** Dalam konteks penelitian, evaluasi dilakukan secara berkelanjutan untuk menilai validitas data mining. Misalnya, menggunakan metrik evaluasi seperti *Davies-Bouldin Index* untuk menentukan seberapa baik model *clustering* dapat mengelompokkan data aktivitas dan respon pengguna media sosial.

4. HASIL DAN PEMBAHASAN

4.1. Hasil Penelitian

Hasil dari proses *data mining* ini adalah analisis aktivitas dan respon pengguna media sosial menggunakan metode *Knowledge Discovery in Database Process* (KDD) untuk mengelompokkan

mereka ke dalam beberapa *cluster*. Proses ini dilakukan dengan menggunakan *RapidMiner Studio* versi 10.3. Tahapan pertama melibatkan pengambilan data dari berbagai *platform* media sosial, mencakup aktivitas dan respon pengguna, dengan total 481 *record* data. Setelah itu, dilakukan seleksi dan transformasi atribut untuk mempersiapkan data. Proses *data mining* menggunakan metode *clustering K-Means*, dengan pengujian model *clustering* dari 2 hingga 10 *cluster*.

4.2. Data

Pengujian data penelitian dengan dataset aktivitas dan respon pengguna media sosial pada tahun 2022 dari Kaggle (<https://www.kaggle.com/>) berjumlah 481 records. Proses import data dilakukan dengan menambahkan operator *Read Excel*, memilih lokasi file, *sheet* yang berisi sumber data, dan menentukan tipe data dan atribut.

4.3. Transformation Data

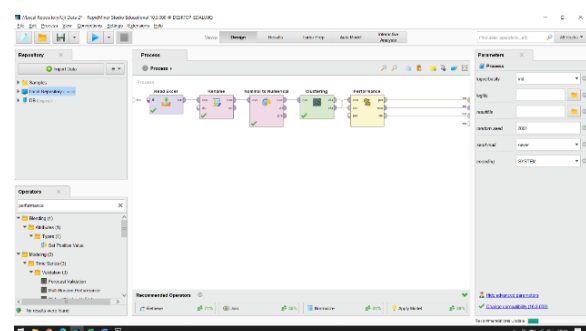
✓ Purposeless use of Social Media	Integer	0	1	5	Frequency	3.553
✓ Readinesses if Social Media n...	Integer	0	1	5	Frequency	2.508
✓ Difficulty in concentrating	Integer	0	1	5	Frequency	3.245
✓ Comparison of self to peers	Integer	0	1	5	Frequency	2.832
✓ Validation sought from Social ...	Integer	0	1	5	Frequency	2.465
✓ Sleep Issues	Integer	0	1	5	Frequency	3.202

Gambar 3. Hasil operator *Rename*

Gambar 6 menampilkan hasil dari proses *Rename* melibatkan pengubahan nama atribut dan konversi tipe data nominal menjadi numerik, memudahkan analisis. Operator *Rename* dan *Nominal to Numerical* digunakan untuk memudahkan analisis data.

4.4. Data Mining

Tambahkan operator *K-Means* untuk *clustering* dan operator *Cluster Distance Performance* untuk evaluasi dengan nilai DBI. Proses ini menghasilkan visualisasi *cluster*, *plot clustering*, dan *performance vector*.

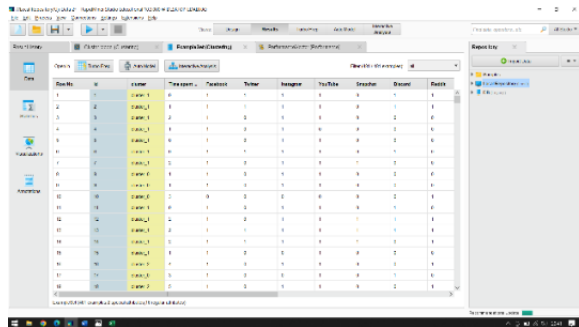


Gambar 4. Tampilan proses *running data*

Gambar 10 merupakan visualisasi proses *running data* yang kemudian menghasilkan *output* yaitu:

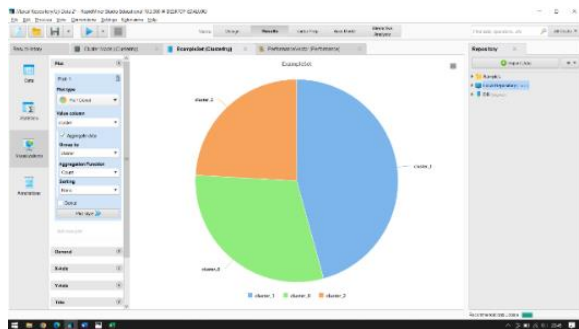
4.4.1. ExampleSet

Pada *ExampleSet* dapat dilihat beberapa tampilan hasil *cluster*, yaitu pada *Data View*. *Data View* merupakan tampilan hasil *cluster* data secara keseluruhan sesuai dengan data yang telah diinputkan.



Gambar 5. Tampilan hasil *cluster* pada *data view*

Gambar 11 memberikan visualisasi tampilan yang menggambarkan hasil dari proses *clustering* pada *data view*. Dengan menggunakan visualisasi ini, dapat lebih jelas memahami struktur dan pola dari pengelompokan data ke dalam *cluster-cluster* yang dihasilkan melalui analisis.



Gambar 6. *Pie chart*

Chart pada Gambar 12 merupakan tampilan grafik hasil pengelompokan atau *cluster dataset* dengan 3 *cluster*.

4.4.2. Cluster Model (Clustering)

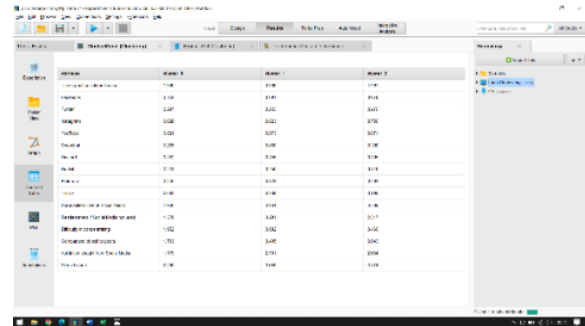
Pada *cluster model (clustering)* dapat dilihat berbagai tampilan hasil *cluster*, yaitu *text view* yang merupakan hasil pengelompokan berdasarkan *cluster* dan jumlah anggotanya.

Cluster Model

```
Cluster 0: 145 items
Cluster 1: 220 items
Cluster 2: 116 items
Total number of items: 481
```

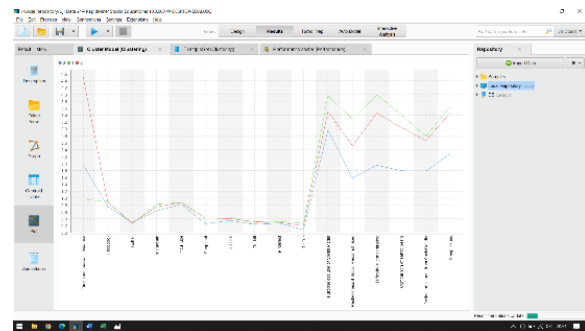
Gambar 7. *Cluster model*

Pada gambar diatas, terdapat 3 *cluster* yang dibentuk yaitu *cluster 0* memiliki 145 *items*, *cluster 1* memiliki 220 *items*, dan *cluster 2* memiliki 116 *items*.



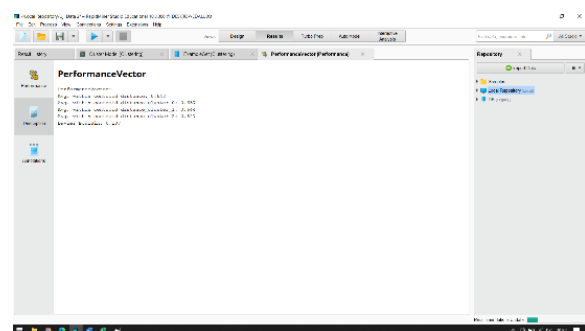
Gambar 8. Nilai *centroid*

Gambar 14 menampilkan visualisasi nilai *centroid* hasil proses *clustering*, menunjukkan lokasi pusat atau representasi titik tengah dari setiap *cluster*. Informasi ini memberikan wawasan tambahan mengenai karakteristik dan distribusi data dalam setiap *cluster*, berguna untuk interpretasi lebih lanjut terkait perilaku aktivitas dan respon pengguna media sosial.



Gambar 9. Tampilan *plot clustering*

Gambar 15 menampilkan *plot clustering* yang menggambarkan secara grafis sebaran dan keterkaitan data dalam *cluster*. Visualisasi ini memberikan pandangan intuitif terhadap pengelompokan aktivitas dan respon pengguna media sosial.



Gambar 10. Tampilan *performance vector*

Gambar 16 memvisualisasikan evaluasi kinerja model *clustering* pada data aktivitas dan respon pengguna media sosial. *Performance vector* ini menggunakan metrik evaluasi, seperti akurasi, presisi, *recall*, untuk memberikan gambaran keseluruhan tentang kehandalan model dalam mengelompokkan data ke dalam *cluster* yang sesuai.

4.5. Evaluation

Model *clustering* dibuat dengan jumlah *cluster* 2–10, diuji dengan *Cluster Distance Performance*, dan hasilnya direpresentasikan dalam Tabel 2. Terdapat tiga *cluster* dengan rata-rata jarak *centroid* 0,648, dan DBI sebesar 0,107, menunjukkan keoptimalan pengelompokan data aktivitas dan respon pengguna media sosial.

Tabel 2. Nilai DBI tiap *cluster*

Cluster	Nilai DBI
2	0,127
3	0,107
4	0,115
5	0,113
6	0,123

Cluster	Nilai DBI
7	0,121
8	0,115
9	0,114
10	0,118

Tabel 2 menunjukkan hasil pengujian aplikasi *RapidMiner* dengan tiga *cluster* yang telah ditentukan. Rata-rata jarak *centroid* adalah 0,648, dengan nilai rinci 0,657 untuk *cluster* 0, 0,644 untuk *cluster* 1, dan 0,647 untuk *cluster* 2. *Davies-Bouldin Index* mencapai nilai 0,107. Informasi ini menggambarkan distribusi dan kompaknya pengelompokan data dalam setiap *cluster*, serta indikasi kesesuaian model *clustering* terhadap *dataset* aktivitas dan respon pengguna media sosial.

Tabel 3. Hasil Analisis

No	Cluster 0	Cluster 1	Cluster 2
1.	Terdapat 145 item	Terdapat 220 item	Terdapat 116 item
2.	Terdiri dari pengguna media sosial dengan penggunaan waktu akses media sosial : 1 - 2 jam setiap hari sebanyak 21 pengguna 2 – 3 jam setiap hari sebanyak 35 pengguna 3 - 4 jam setiap hari sebanyak 36 pengguna 4 - 5 jam setiap hari sebanyak 5 pengguna < 1 jam setiap hari sebanyak 28 pengguna > 5 jam setiap hari sebanyak 20 pengguna	Terdiri dari pengguna media sosial dengan penggunaan waktu akses media sosial : 2 - 3 jam setiap hari sebanyak 66 pengguna 3 - 4 jam setiap hari sebanyak 57 pengguna < 1 jam setiap hari sebanyak 1 pengguna > 5 jam setiap hari sebanyak 96 pengguna	Terdiri dari pengguna media sosial dengan penggunaan waktu akses media sosial : 1 - 2 jam setiap hari sebanyak 49 pengguna 4 - 5 jam setiap hari sebanyak 62 pengguna < 1 jam setiap hari sebanyak 5 pengguna
3.	Terdiri dari pengguna media sosial yang mengakses tanpa tujuan : Bernilai 1 sebanyak 18 pengguna Bernilai 2 sebanyak 33 pengguna Bernilai 3 sebanyak 46 pengguna Bernilai 4 sebanyak 32 pengguna Bernilai 5 sebanyak 16 pengguna	Terdiri dari pengguna media sosial yang mengakses tanpa tujuan : Bernilai 1 sebanyak 1 pengguna Bernilai 2 sebanyak 14 pengguna Bernilai 3 sebanyak 46 pengguna Bernilai 4 sebanyak 90 pengguna Bernilai 5 sebanyak 69 pengguna	Terdiri dari pengguna media sosial yang mengakses tanpa tujuan : Bernilai 1 sebanyak 1 pengguna Bernilai 2 sebanyak 18 pengguna Bernilai 3 sebanyak 37 pengguna Bernilai 4 sebanyak 41 pengguna Bernilai 5 sebanyak 19 pengguna
4.	Terdiri dari pengguna media sosial yang merasa gelisah ketika tidak menggunakan media sosial : Bernilai 1 sebanyak 85 pengguna Bernilai 2 sebanyak 38 pengguna Bernilai 3 sebanyak 20 pengguna Bernilai 4 sebanyak 2 pengguna	Terdiri dari pengguna media sosial yang merasa gelisah ketika tidak menggunakan media sosial : Bernilai 1 sebanyak 10 pengguna Bernilai 2 sebanyak 45 pengguna Bernilai 3 sebanyak 73 pengguna Bernilai 4 sebanyak 55 pengguna Bernilai 5 sebanyak 37 pengguna	Terdiri dari pengguna media sosial yang merasa gelisah ketika tidak menggunakan media sosial : Bernilai 1 sebanyak 22 pengguna Bernilai 2 sebanyak 41 pengguna Bernilai 3 sebanyak 31 pengguna Bernilai 4 sebanyak 15 pengguna Bernilai 5 sebanyak 7 pengguna
5.	Terdiri dari pengguna media sosial yang merasa kesulitan berkonsentrasi : Bernilai 1 sebanyak 59 pengguna Bernilai 2 sebanyak 46 pengguna Bernilai 3 sebanyak 31 pengguna Bernilai 4 sebanyak 6 pengguna Bernilai 5 sebanyak 3 pengguna	Terdiri dari pengguna media sosial yang merasa kesulitan berkonsentrasi : Bernilai 1 sebanyak 4 pengguna Bernilai 2 sebanyak 13 pengguna Bernilai 3 sebanyak 45 pengguna Bernilai 4 sebanyak 80 pengguna Bernilai 5 sebanyak 78 pengguna	Terdiri dari pengguna media sosial yang merasa kesulitan berkonsentrasi : Bernilai 1 sebanyak 4 pengguna Bernilai 2 sebanyak 22 pengguna Bernilai 3 sebanyak 33 pengguna Bernilai 4 sebanyak 30 pengguna Bernilai 5 sebanyak 27 pengguna
6.	Terdiri dari pengguna media sosial yang membandingkan diri dengan orang lain di media sosial : Bernilai 1 sebanyak 76 pengguna Bernilai 2 sebanyak 39 pengguna Bernilai 3 sebanyak 19 pengguna Bernilai 4 sebanyak 6 pengguna	Terdiri dari pengguna media sosial yang membandingkan diri dengan orang lain di media sosial : Bernilai 1 sebanyak 18 pengguna Bernilai 2 sebanyak 45 pengguna Bernilai 3 sebanyak 49 pengguna Bernilai 4 sebanyak 46 pengguna	Terdiri dari pengguna media sosial yang membandingkan diri dengan orang lain di media sosial : Bernilai 1 sebanyak 16 pengguna Bernilai 2 sebanyak 24 pengguna Bernilai 3 sebanyak 33 pengguna Bernilai 4 sebanyak 25 pengguna

No	Cluster 0	Cluster 1	Cluster 2
	Bernilai 5 sebanyak 5 pengguna	Bernilai 5 sebanyak 62 pengguna	Bernilai 5 sebanyak 18 pengguna
7.	Terdiri dari pengguna media sosial yang mencari validasi dari media sosial : Bernilai 1 sebanyak 79 pengguna Bernilai 2 sebanyak 31 pengguna Bernilai 3 sebanyak 27 pengguna Bernilai 4 sebanyak 4 pengguna Bernilai 5 sebanyak 4 pengguna	Terdiri dari pengguna media sosial yang mencari validasi dari media sosial : Bernilai 1 sebanyak 47 pengguna Bernilai 2 sebanyak 43 pengguna Bernilai 3 sebanyak 57 pengguna Bernilai 4 sebanyak 55 pengguna Bernilai 5 sebanyak 18 pengguna	Terdiri dari pengguna media sosial yang mencari validasi dari media sosial : Bernilai 1 sebanyak 24 pengguna Bernilai 2 sebanyak 25 pengguna Bernilai 3 sebanyak 38 pengguna Bernilai 4 sebanyak 24 pengguna Bernilai 5 sebanyak 5 pengguna
8.	Terdiri dari pengguna media sosial yang memiliki masalah kesulitan tidur : Bernilai 1 sebanyak 51 pengguna Bernilai 2 sebanyak 41 pengguna Bernilai 3 sebanyak 22 pengguna Bernilai 4 sebanyak 22 pengguna Bernilai 5 sebanyak 9 pengguna	Terdiri dari pengguna media sosial yang memiliki masalah kesulitan tidur : Bernilai 1 sebanyak 21 pengguna Bernilai 2 sebanyak 30 pengguna Bernilai 3 sebanyak 31 pengguna Bernilai 4 sebanyak 59 pengguna Bernilai 5 sebanyak 79 pengguna	Terdiri dari pengguna media sosial yang memiliki masalah kesulitan tidur : Bernilai 1 sebanyak 14 pengguna Bernilai 2 sebanyak 20 pengguna Bernilai 3 sebanyak 16 pengguna Bernilai 4 sebanyak 29 pengguna Bernilai 5 sebanyak 37 pengguna
9.	Terdiri dari beberapa pengguna <i>platform-platform</i> media sosial: Facebook = 109 pengguna Twitter = 43 pengguna Instagram = 91 pengguna YouTube = 119 pengguna Snapchat = 37 pengguna Discord = 51 pengguna Reddit = 34 pengguna Pinterest = 40 pengguna TikTok = 12 pengguna	Terdiri dari beberapa pengguna <i>platform-platform</i> media sosial: Facebook = 196 pengguna Twitter = 56 pengguna Instagram = 181 pengguna YouTube = 192 pengguna Snapchat = 99 pengguna Discord = 100 pengguna Reddit = 55 pengguna Pinterest = 71 pengguna TikTok = 59 pengguna	Terdiri dari beberapa pengguna <i>platform-platform</i> media sosial: Facebook = 102 pengguna Twitter = 32 pengguna Instagram = 87 pengguna YouTube = 101 pengguna Snapchat = 45 pengguna Discord = 47 pengguna Reddit = 37 pengguna Pinterest = 34 pengguna TikTok = 23 pengguna

Tabel 3 menunjukkan bahwa *Cluster 1* (220 anggota) adalah pengguna media sosial yang aktif, menghabiskan waktu akses lebih dari 2-3 jam sehari, sering mengakses tanpa tujuan, dan mencari validasi dari media sosial. *Cluster 0* (145 anggota) cenderung mengakses sekitar 1-2 jam sehari, dengan tingkat perbandingan diri dan kesulitan berkonsentrasi yang sedang. Sementara *Cluster 2* (116 anggota) memiliki waktu akses yang bervariasi, termasuk dalam rentang 1-2 jam dan 4-5 jam sehari, dengan tingkat perbandingan diri dan pencarian validasi yang lebih rendah. Preferensi *platform* media sosial juga bervariasi di setiap *cluster*, memberikan wawasan tentang pola perilaku dan preferensi pengguna.

5. KESIMPULAN DAN SARAN

Dalam penelitian ini, penerapan algoritma *K-Means* menghasilkan tiga *cluster* dengan karakteristik berbeda. *Cluster 1* menonjol sebagai yang terbesar, didominasi oleh pengguna aktif dengan waktu akses lebih dari 2-3 jam sehari, sering mengakses tanpa tujuan, dan mencari validasi dari media sosial. *Cluster 0* cenderung mengakses sekitar 1-2 jam sehari, dengan tingkat perbandingan diri dan kesulitan berkonsentrasi yang sedang. Sementara *Cluster 2* memiliki waktu akses yang bervariasi, termasuk dalam rentang 1-2 jam dan 4-5 jam sehari, dengan tingkat perbandingan diri dan pencarian validasi yang lebih rendah. Preferensi *platform* media sosial juga bervariasi di setiap *cluster*, memberikan wawasan tentang pola perilaku dan preferensi pengguna. Evaluasi menggunakan *Davies-Bouldin Index* menunjukkan keefektifan pengelompokan, dan hasil penelitian memberikan

landasan untuk pemahaman yang lebih baik tentang perilaku pengguna media sosial. Saran untuk penelitian selanjutnya mencakup pengembangan tujuan penelitian yang lebih fokus, eksplorasi algoritma pengelompokan lainnya, dan pemahaman lebih mendalam tentang struktur data aktivitas dan respon pengguna.

DAFTAR PUSTAKA

- [1] K. Woran, R. M. Kundre, and F. A. Pondaag, "Analisis Hubungan Penggunaan Media Sosial Dengan Kualitas Tidur Pada Remaja," *J. Keperawatan*, vol. 8, no. 2, p. 1, 2021, doi: 10.35790/jkp.v8i2.32092.
- [2] N. N. Zahara and K. Wisnusakti, "Hubungan Intensitas Penggunaan Media Sosial Terhadap Kestabilan Emosi Pada Anak Usia Remaja," *Open Access Jakarta J. Heal. Sci.*, vol. 1, no. 3, pp. 85–93, 2022, doi: 10.53801/oajjhs.v1i3.10.
- [3] A. S. A. N. Annisa, I. Yuliadi, and D. Nugroho, "Hubungan Tingkat Kepercayaan Diri Dengan Intensitas Penggunaan Media Sosial Whatsapp Pada Mahasiswa Kedokteran 2018," *Wacana*, vol. 12, no. 1, pp. 86–109, 2020, doi: 10.13057/wacana.v12i1.170.
- [4] M. A. Z. Larasati, N. A. S. Winarsih, M. S. Rohman, and G. W. Saraswati, "Penerapan Metode K-Means Clustering Dalam Menganalisis Sentimen Masyarakat Terhadap K-Popers Pada Twitter," *Progresif J. Ilm. Komput.*, vol. 18, no. 2, p. 201, 2022, doi: 10.35889/progresif.v18i2.877.
- [5] S. Anwar, T. Suprapti, G. Dwilestari, and I. Ali,

- “PENGELOMPOKKAN HASIL BELAJAR SISWA DENGAN METODE CLUSTERING K-MEANS,” *J. Sist. Inf. dan Teknol. Informasi*, vol. 4, no. 2, pp. 60–72, 2022.
- [6] A. G. Gani, “Pengaruh Media Sosial Terhadap Perkembangan Anak Remaja,” *J. Mitra Manaj.*, vol. 7, no. 2, pp. 32–42, 2020.
- [7] Oon Wira Yuda, Darmawan Tuti, Lim Sheih Yee, and Susanti, “Penerapan Penerapan Data Mining Untuk Klasifikasi Kelulusan Mahasiswa Tepat Waktu Menggunakan Metode Random Forest,” *SATIN - Sains dan Teknol. Inf.*, vol. 8, no. 2, pp. 122–131, 2022, doi: 10.33372/stn.v8i2.885.
- [8] R. H. Sukarna and Y. Ansori, “Implementasi Data Mining Menggunakan Metode Naive Bayes Dengan Feature Selection Untuk Prediksi Kelulusan Mahasiswa Tepat Waktu,” *J. Ilm. Sains dan Teknol.*, vol. 6, no. 1, pp. 50–61, 2022, doi: 10.47080/saintek.v6i1.1467.
- [9] A. R. M. Rochcham, “Implementasi Algoritma K-Means Untuk Clustering Kepuasan Mahasiswa Terhadap Pelayanan Akademik,” *J. Fak. Tek. Univ. Pandanaran*, vol. 6, no. 2, pp. 42–45, 2020.
- [10] C. Kamila, “Systematic Literature Review: Penggunaan Algoritma K-Means Untuk Clustering di Indonesia dalam Bidang Pendidikan,” *Intech*, vol. 2, no. 1, pp. 19–24, 2021, doi: 10.54895/intech.v2i1.866.
- [11] I. Widaningrum, D. Mustikasari, R. Arifin, S. L. Tsaqila, and D. Fatmawati, “Algoritma Term Frequency-Inverse Document Frequency (TF-IDF) dan K-Means Clustering Untuk Menentukan Kategori Dokumen,” *Pros. Semin. Nas. Sist. Inf. dan Teknol.*, pp. 145–149, 2022.