

SEGMENTASI PENDENGAR MUSIK BERDASARKAN KEBIASAAN DAN DAMPAKNYA TERHADAP KESEHATAN MENTAL MENGGUNAKAN *K-MEANS* DAN *RANDOM FOREST*

Joang Ipmawati*

Universitas Nahdlatul Ulama Yogyakarta, Yogyakarta, DIY, Indonesia

Email*: joang@unu-jogja.ac.id

(*) *Corresponding Author*

ABSTRAK

Musik memiliki peran penting dalam kehidupan manusia, tidak hanya sebagai bentuk hiburan, tetapi juga sebagai alat untuk mengelola emosi dan kesehatan mental. Penelitian ini bertujuan untuk menganalisis hubungan antara kebiasaan mendengarkan musik dengan kondisi psikologis individu, serta mengidentifikasi faktor utama yang mempengaruhi efek musik terhadap kesehatan mental. Metode yang digunakan dalam penelitian ini adalah *K-Means Clustering* untuk mengelompokkan pendengar musik berdasarkan kebiasaan dan kondisi mental mereka, serta *Random Forest* untuk menentukan faktor dominan yang mempengaruhi pengalaman musik seseorang. Hasil penelitian menunjukkan bahwa pendengar musik dapat dikategorikan ke dalam empat klaster utama, dengan kecemasan (*Anxiety*) dan *insomnia* sebagai faktor paling dominan dalam menentukan efek musik. *Silhouette Score* sebesar 0.24 menunjukkan bahwa *clustering* memiliki pemisahan yang cukup baik, meskipun masih dapat ditingkatkan. Evaluasi model *Random Forest* menunjukkan akurasi tertinggi sebesar 86% pada Klaster 3, sementara Klaster 0 memiliki akurasi terendah sebesar 55%, mengindikasikan perbedaan pola dalam tiap kelompok. Kesimpulan dari penelitian ini adalah bahwa musik memiliki dampak yang berbeda pada setiap kelompok pendengar, tergantung pada faktor psikologis mereka. Temuan ini dapat digunakan sebagai dasar untuk pengembangan sistem rekomendasi musik berbasis kesehatan mental, serta dalam pengembangan terapi musik untuk individu dengan tingkat kecemasan atau gangguan tidur yang tinggi.

Kata Kunci: *Clustering*, *K-Means*, kesehatan mental, musik, *Random Forest*

ABSTRACT

Music plays a crucial role in human life, not only as a form of entertainment but also as a tool for managing emotions and mental health. This study aims to analyze the relationship between music listening habits and individuals' psychological conditions, as well as to identify the key factors influencing the effects of music on mental health. The methods used in this research include K-Means Clustering to group music listeners based on their habits and mental states, and Random Forest to determine the dominant factors affecting their music experience. The findings reveal that music listeners can be categorized into four main clusters, with anxiety and insomnia emerging as the most dominant factors influencing music effects. The Silhouette Score of 0.24 indicates a moderately effective clustering performance, though there is room for improvement. The evaluation of the Random Forest model shows the highest accuracy at 86% in Cluster 3, while Cluster 0 records the lowest accuracy at 55%, suggesting variations in pattern recognition across clusters. This study concludes that music impacts each group of

listeners differently, depending on their psychological characteristics. These findings can serve as the foundation for developing personalized music recommendation systems based on mental health conditions, as well as advancing music therapy for individuals with high levels of anxiety or sleep disorders.

Keywords: Clustering, K-Means, mental health, music, Random Forest

1. PENDAHULUAN

Musik telah menjadi bagian penting dalam kehidupan sehari-hari dan memiliki peran yang signifikan dalam memengaruhi kondisi psikologis seseorang [1]. Sebagai bentuk ekspresi dan terapi, musik dapat membantu mengurangi kecemasan, meningkatkan suasana hati, hingga mengatasi gangguan tidur [2]. Namun, efek musik terhadap kesehatan mental tidak bersifat universal, melainkan dipengaruhi oleh berbagai faktor, seperti frekuensi dan durasi mendengarkan, genre yang dipilih, serta kondisi psikologis individu. Oleh karena itu, memahami segmentasi pendengar musik berdasarkan kebiasaan dan dampaknya terhadap kesehatan mental menjadi hal yang penting untuk diteliti [3] [4].

Dengan meningkatnya jumlah data terkait kebiasaan mendengarkan musik, teknik *machine learning* dapat digunakan untuk menganalisis pola-pola yang tersembunyi dalam preferensi pendengar. Model pembelajaran mesin dapat memprediksi apakah lagu tertentu memiliki efek terapeutik pada orang tertentu, membantu terapis musik dan pasien dalam memilih musik yang tepat untuk perawatan [5]. Algoritma *K-Means* merupakan *unsupervised machine learning* yang dianggap salah satu algoritma mesin pembelajaran populer di kalangan peneliti [6]. Algoritma ini digunakan untuk mengelompokkan musik berdasarkan karakteristiknya, sehingga dapat mengenali pola kesamaan antara berbagai genre. Dengan mengekstrak fitur-fitur seperti tempo, nada, dan elemen musik lainnya, metode ini membantu dalam mengklasifikasikan musik yang memiliki kemiripan ke dalam satu kategori yang sama [7]. Hakim dkk berhasil mengelompokkan musik menggunakan algoritma ini menjadi lima kluster [8]. Dalam penelitian ini, metode *K-Means Clustering* diterapkan untuk mengelompokkan pendengar musik berdasarkan kebiasaan mereka, seperti durasi mendengarkan musik per hari, preferensi genre, serta kondisi mental seperti kecemasan, depresi, dan *insomnia*. *K-Means* memungkinkan identifikasi kelompok - kelompok yang memiliki karakteristik serupa, sehingga dapat memberikan pemahaman lebih mendalam tentang perbedaan dampak musik pada berbagai tipe pendengar [9].

Selain segmentasi, penting untuk memahami faktor utama yang mempengaruhi dampak musik terhadap kesehatan mental. Untuk itu, penelitian ini menggunakan *Random Forest* sebagai metode prediksi dan analisis faktor dominan. *Random Forest* lebih efektif dalam klasifikasi data penurunan stress menggunakan musik dibandingkan SupportVector Musik [10]. *Random forest* membantu mengidentifikasi variabel mana yang paling berkontribusi terhadap efek musik pada setiap kelompok pendengar. Penelitian Fadil dengan kombinasi kedua metode ini, penelitian ini tidak hanya dapat mengelompokkan pendengar berdasarkan pola konsumsi musik mereka, tetapi juga memberikan wawasan tentang hubungan antara kebiasaan mendengarkan musik dan kondisi psikologis mereka.

Tujuan utama dari penelitian ini adalah untuk menganalisis segmentasi pendengar musik berdasarkan kebiasaan mereka menggunakan *K-Means Clustering* serta mengidentifikasi faktor dominan yang mempengaruhi dampak musik terhadap kesehatan mental menggunakan *Random Forest*. Hasil dari penelitian ini diharapkan dapat menjadi

dasar dalam pengembangan sistem rekomendasi musik berbasis kesehatan mental serta memberikan wawasan bagi terapi musik yang lebih efektif, khususnya bagi individu dengan tingkat kecemasan atau gangguan tidur yang tinggi

2. METODE

Pemahaman Dan Eksplorasi Dataset

Tahap ini dilakukan untuk memahami struktur dataset, mengidentifikasi variabel yang relevan, dan mendeteksi adanya nilai kosong (*missing values*) atau nilai yang tidak wajar (*outliers*) [11]. *Dataset* yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari www.kaggle.com. Dataset terdiri dari 736 entri dengan 33 variabel yang mana dipilih 8 variabel yang relevan dengan informasinya terlihat pada tabel 1.

Tabel 1. Informasi *dataset*

| No | Atribut | | |
|----|----------------------|--------------------|--|
| | Nama | Jenis | Deskripsi |
| 1 | <i>Hours per day</i> | Numerik | Durasi mendengarkan musik dalam jam per hari |
| 2 | <i>Anxiety</i> | Numerik | Tingkat Kecemasan |
| 3 | <i>Depression</i> | Numerik | Tingkat Depresi |
| 4 | <i>Insomnia</i> | Numerik | Tingkat <i>Insomnia</i> |
| 5 | <i>OCD</i> | Numerik | Tingkat <i>Obsessive-Compulsive Disorder</i> |
| 6 | <i>Fav genre</i> | <i>Categorical</i> | <i>Genre</i> musik favorit |
| 7 | <i>Music effects</i> | <i>Categorical</i> | Efek musik terhadap pendengar |
| 8 | <i>While working</i> | <i>Categorical</i> | Kebiasaan mendengarkan musik saat bekerja |

Preprocessing Data

Tahap *Preprocessing Data* dilakukan untuk memastikan bahwa data dalam kondisi optimal sebelum dianalisis lebih lanjut. Langkah pertama adalah menghapus data kosong serta menangani outlier guna menghindari distorsi dalam hasil analisis. Selanjutnya, variabel kategorikal seperti *genre* musik, efek musik, dan kebiasaan mendengarkan musik saat bekerja dikonversi menjadi format numerik melalui proses *encoding*, sehingga dapat diproses oleh algoritma *machine learning*. Terakhir, dilakukan normalisasi data untuk menyamakan skala setiap variabel, sehingga mempermudah proses *Clustering* (pengelompokan data) agar hasilnya lebih akurat dan bermakna [12].

Implementasi Algoritma *K-Means*

K-Means clustering merupakan metode untuk mengelompokkan data non-hierarki yang mempartisi data ke dalam bentuk dua atau lebih kelompok, sehingga data berkarakteristik sama dimasukkan ke dalam satu kelompok yang sama dan data yang berkarakteristik berbeda dikelompokkan ke dalam kelompok yang lain [13].

Pada penelitian ini, metode *Clustering (K-Means)* digunakan untuk mengelompokkan pendengar musik berdasarkan kebiasaan mendengarkan dan kondisi kesehatan mental. Proses ini diawali dengan menentukan jumlah kluster optimal menggunakan *Elbow Method*, yang membantu mengidentifikasi titik terbaik di mana penambahan jumlah kluster tidak lagi memberikan peningkatan signifikan dalam pemisahan data.

Setelah jumlah kluster optimal ditentukan, dilakukan proses pengelompokan data berdasarkan beberapa variabel utama, seperti durasi mendengarkan musik, tingkat kecemasan, depresi, *insomnia*, gangguan *obsesif-kompulsif* (OCD), genre musik favorit, efek musik terhadap pendengar, serta kebiasaan mendengarkan musik saat bekerja. Dengan menggunakan pendekatan ini, individu yang memiliki pola mendengarkan musik dan kondisi mental yang serupa dapat dikelompokkan dalam satu kategori.

Hasil dari proses ini memungkinkan pemahaman yang lebih mendalam tentang bagaimana musik dapat memengaruhi kesehatan mental, serta memberikan wawasan yang dapat digunakan dalam pengembangan strategi intervensi berbasis musik atau sistem rekomendasi musik yang disesuaikan dengan kondisi emosional pendengar [14].

Implementasi Prediksi Faktor Dominan (*Random Forest*)

Dalam penelitian ini, metode *Random Forest* diterapkan untuk menganalisis faktor utama yang mempengaruhi efek musik dalam setiap kluster yang telah terbentuk sebelumnya melalui proses *clustering*. *Random Forest* digunakan sebagai teknik prediksi dan analisis fitur penting, di mana model ini dapat mengidentifikasi variabel yang paling berpengaruh terhadap respons individu terhadap musik.

Langkah pertama dalam implementasi ini adalah melatih model *Random Forest* menggunakan data yang telah dikelompokkan, dengan tujuan mengetahui faktor mana yang paling menentukan efek musik pada setiap kluster. Variabel-variabel seperti durasi mendengarkan musik, tingkat kecemasan, depresi, *insomnia*, dan kebiasaan mendengarkan musik dianalisis untuk melihat pengaruhnya terhadap efek yang dirasakan pendengar.

Setelah model dilatih, dilakukan evaluasi model guna memastikan keakuratan prediksi dan relevansi faktor yang ditemukan. Evaluasi ini mencakup pengukuran akurasi serta analisis *feature importance*, yang menunjukkan kontribusi masing-masing variabel terhadap efek musik dalam setiap kelompok pendengar. Dengan cara ini, penelitian dapat mengidentifikasi faktor-faktor dominan yang menentukan bagaimana musik mempengaruhi individu, sehingga dapat digunakan dalam pengembangan terapi musik yang lebih efektif atau sistem rekomendasi musik yang berbasis pada kondisi psikologis pendengar [15].

Evaluasi Model

Untuk memastikan keakuratan dan efektivitas metode yang digunakan dalam penelitian ini, dilakukan evaluasi model melalui dua pendekatan utama

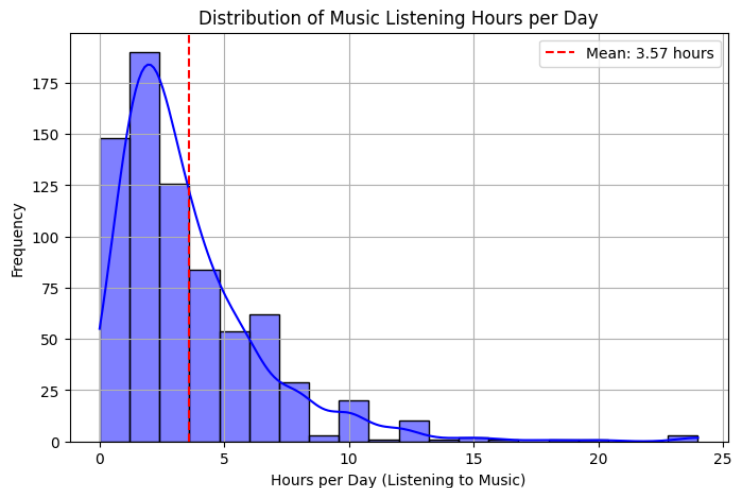
- a. Evaluasi *Clustering* dengan *Silhouette Score*, digunakan untuk mengukur kualitas pengelompokan. Skor ini menilai seberapa baik setiap data dalam kluster sesuai dengan kelompoknya, dibandingkan dengan kluster lain. Jika skor mendekati 1 maka kluster yang terbentuk memiliki pemisahan yang baik dan valid. Sebaliknya jika mendekati 0 maka titik data berada di perbatasan antar kluster, sehingga kurang jelas pengelompokannya [16].
- b. Evaluasi Model Prediksi dengan *Random Forest* menggunakan Akurasi. *Random Forest* digunakan untuk melihat seberapa baik model dalam memprediksi [17]. Evaluasi dilakukan dengan membandingkan hasil prediksi model dengan data sebenarnya, sehingga dapat diketahui apakah faktor utama yang teridentifikasi dalam *clustering* memiliki korelasi yang kuat dengan efek musik yang dirasakan oleh pendengar. Semakin tinggi akurasi model, semakin relevan variabel yang digunakan dalam prediksi efek musik.

Hasil dari evaluasi ini akan digunakan untuk menilai apakah model *clustering* dan prediksi faktor dominan sudah cukup baik, atau perlu dilakukan penyesuaian parameter dan optimasi lebih lanjut guna meningkatkan keakuratan analisis.

3. HASIL DAN PEMBAHASAN

Eksplorasi *Dataset*

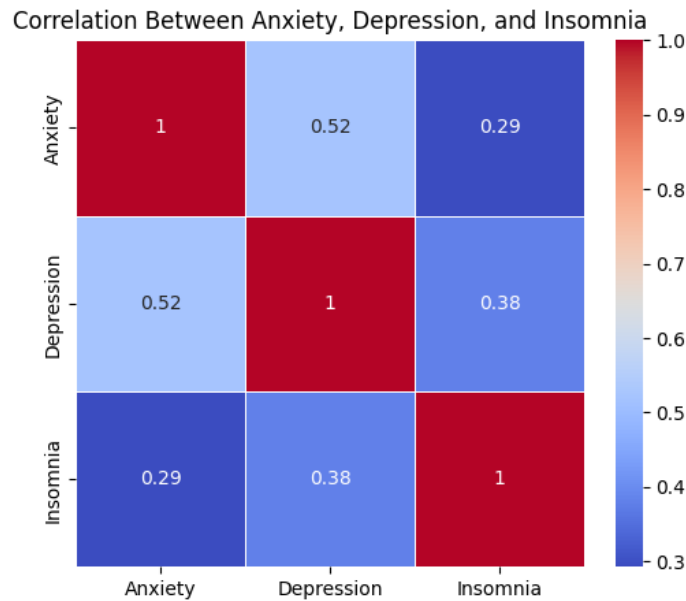
Dalam penelitian ini, analisis diawali dengan eksplorasi data untuk memahami distribusi variabel utama dalam dataset. Salah satu aspek yang dianalisis adalah durasi mendengarkan musik per hari, yang mencerminkan kebiasaan pendengar dalam mengonsumsi musik yang hasilnya dapat dilihat dalam gambar 1.



Gambar 1. Distribusi durasi mendengarkan *music* per hari

Gambar1 di atas menunjukkan rata-rata durasi mendengarkan musik adalah 3.57 jam per hari yang ditunjukkan oleh garis merah putus-putus. Distribusi data ini menunjukkan pola *skewed right* (condong ke kanan), yang berarti sebagian besar pendengar menghabiskan waktu mendengarkan musik kurang dari 5 jam per hari, sementara beberapa individu mendengarkan musik dalam durasi yang jauh lebih lama, mencapai lebih dari 15 jam dalam sehari.

Korelasi antara kecemasan, depresi, dan *insomnia* menunjukkan hubungan positif/negatif antara variabel.

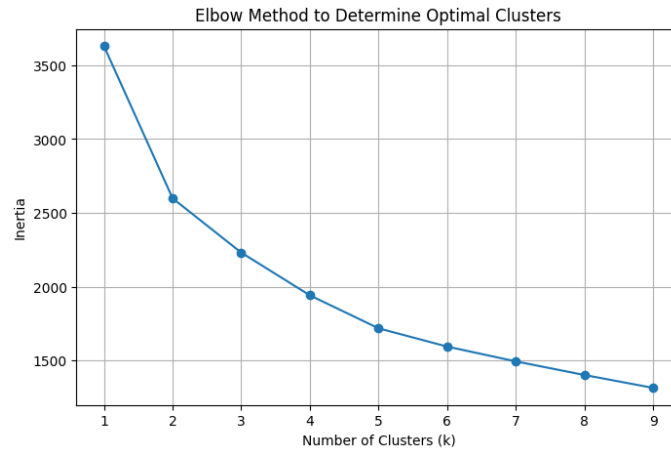


Gambar 2. Korelasi antara kecemasan, depresi, dan *insomnia*

Gambar 2 adalah *heatmap* korelasi antara tiga variabel psikologis utama kecemasan (*Anxiety*), depresi (*Depression*), dan *insomnia*. Nilai korelasi berkisar antara -1 hingga 1, di mana 1 menunjukkan korelasi sempurna positif, 0 berarti tidak ada korelasi, dan -1 menunjukkan korelasi sempurna negatif. Korelasi antara kecemasan dan depresi (0.52) menunjukkan hubungan yang cukup kuat. Hal ini mengindikasikan bahwa individu dengan tingkat kecemasan yang lebih tinggi cenderung memiliki tingkat depresi yang lebih tinggi pula. Korelasi antara depresi dan *insomnia* (0.38) juga cukup signifikan, yang berarti bahwa individu dengan depresi cenderung mengalami gangguan tidur lebih sering. Sedangkan korelasi antara kecemasan dan *insomnia* (0.29) lebih rendah dibandingkan yang lain, tetapi tetap menunjukkan adanya hubungan. Ini menunjukkan bahwa kecemasan dapat mempengaruhi pola tidur, meskipun tidak sekuat hubungannya dengan depresi.

Implementasi Algoritma *K-Means*

Penentuan jumlah kluster optimal dalam penelitian ini dilakukan dengan menggunakan metode Elbow. Metode ini melibatkan pengujian nilai k dalam rentang 1 hingga 5, di mana untuk setiap nilai k dihitung nilai *within-cluster sum of squares* (WCSS). Hasil perhitungan ditampilkan dalam bentuk grafik dapat dilihat pada Gambar 3, yang menunjukkan bahwa nilai WCSS mengalami penurunan signifikan dari $k=1$ hingga $k=3$, kemudian cenderung melandai setelahnya. Titik siku (*elbow point*) teridentifikasi pada $k=4$, yang diinterpretasikan sebagai jumlah kluster optimal. Nilai tersebut selanjutnya digunakan dalam proses klusterisasi untuk mengelompokkan pendengar musik berdasarkan pola kebiasaan serta kondisi kesehatan mental mereka.

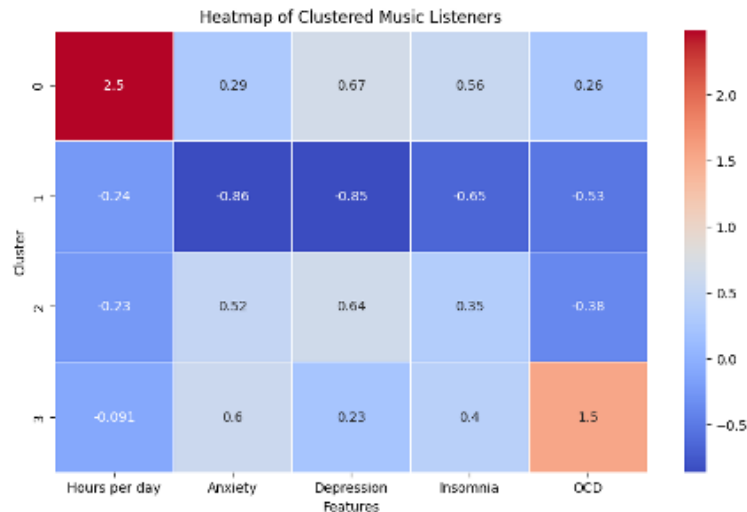
Gambar 3. Hasil perhitungan metode *elbow*

Proses *clustering* menggunakan algoritma *K-Means* menghasilkan pembagian data ke dalam empat kluster yang ditentukan berdasarkan hasil evaluasi menggunakan *Elbow Method*. Hasil *clustering* dianalisis untuk memahami karakteristik setiap kluster berdasarkan atribut utama yang hasilnya dapat dilihat pada tabel 2.

Tabel 2. Hasil *clustering*

| <i>Hours per day</i> | <i>Anxiety</i> | <i>Depression</i> | <i>Insomnia</i> | <i>OCD</i> | <i>Fav genre</i> | <i>Music effects</i> | <i>While working</i> | <i>Cluster</i> |
|----------------------|----------------|-------------------|-----------------|------------|------------------|----------------------|----------------------|----------------|
| 0.138 | 0.417 | 0.728 | 2.0334 | -0.226 | 15 | 1 | 0 | 2 |
| -0.359 | 1.135 | 0.728 | -0.239 | 0.125 | 6 | 0 | 1 | 2 |
| 0.138 | 0.4170 | -0.929 | 0.410 | 2.238 | 12 | 0 | 1 | 3 |
| 0.469 | 0.776 | 1.061 | 1.059 | 1.534 | 6 | 0 | 1 | 3 |
| -0.193 | -0.661 | 1.060 | 0.734 | -0.930 | 15 | 0 | 1 | 2 |
| ... | ... | ... | ... | ... | ... | ... | ... | ... |
| 0.800 | -1.379 | -0.929 | -0.563 | -0.226 | 13 | 0 | 1 | 1 |
| 0.469 | -1.379 | -0.598 | -0.563 | -0.578 | 0 | 0 | 1 | 1 |
| -0.524 | -1.379 | -0.929 | -0.563 | 0.829 | 5 | 0 | 1 | 1 |

Visualisasi hasil *clustering* menggunakan *heatmap* dapat dilihat pada gambar 4 dianalisis untuk memahami karakteristik setiap kluster berdasarkan atribut utama seperti durasi mendengarkan musik, kecemasan, depresi, *insomnia*, dan OCD.



Gambar 4. Heatmap clustering pendengar musik

Pada gambar 4, warna yang lebih terang menunjukkan nilai yang lebih tinggi dalam kluster tersebut, sementara warna lebih gelap menunjukkan nilai yang lebih rendah atau negatif.

Analisis kluster yang terbentuk mempunyai karakteristik tiap kluster sebagai berikut

1. Kluster 0 memiliki durasi mendengarkan musik tertinggi (2.5 jam/hari) dan tingkat kecemasan, depresi, serta *insomnia* yang sedang.
2. Kluster 1 ditandai dengan kecemasan (-0.86) dan depresi (-0.85) yang lebih rendah, serta tingkat *insomnia* yang negatif, menunjukkan pendengar dengan kondisi mental yang lebih stabil.
3. Kluster 2 menunjukkan kecemasan (0.57) dan depresi (0.64) yang lebih tinggi, dengan *insomnia* juga cukup tinggi.
4. Kluster 3 memiliki nilai OCD tertinggi (1.5) dan kecemasan yang cukup tinggi (0.6), tetapi durasi mendengarkan musik relatif lebih rendah.

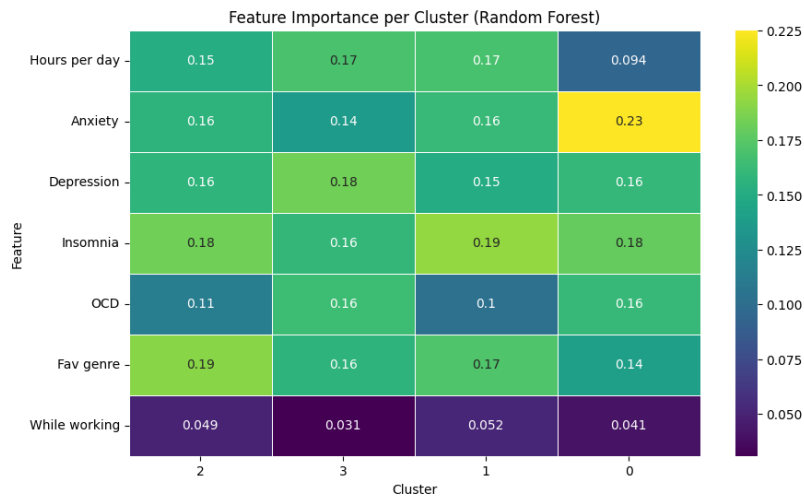
Evaluasi Clustering dengan Silhouette Score

Berdasarkan perhitungan *silhouette score* dilakukan di *google colab* menggunakan kode *python*, model mendapatkan nilai 0.24, yang mengindikasikan bahwa kualitas pengelompokan dalam *K-Means Clustering* masih moderate. Dengan nilai 0.24 *clustering* sudah cukup membentuk kelompok yang berbeda, tetapi masih ada beberapa data yang berdekatan dengan kluster lain. Hal ini dapat diatasi dengan penyesuaian jumlah kluster atau penggunaan metode *clustering* lain yang lebih kompleks untuk meningkatkan pemisahan antar kelompok.

Analisis Faktor Dominan (Random Forest)

Untuk memahami faktor utama yang memengaruhi efek musik terhadap kesehatan mental, penelitian ini menggunakan metode *Random Forest* untuk menganalisis *feature importance* dalam masing-masing kluster. *Random Forest* menghitung kontribusi setiap fitur berdasarkan rata-rata penurunan impuritas (*impurity decrease*) pada setiap pemisahan (*split*) pohon keputusan dalam model. Nilai *feature importance* yang lebih tinggi menunjukkan bahwa fitur tersebut berperan lebih besar dalam proses prediksi. Visualisasi hasil analisis ditampilkan pada Gambar 5 dalam bentuk *heatmap*, yang

menggambarkan tingkat kepentingan relatif dari masing-masing fitur terhadap tiap klaster yang terbentuk.



Gambar 5. Heatmap feature importance

Dari gambar 5 hasil *Random Forest* yang mengidentifikasi faktor utama yang mempengaruhi efek musik dalam setiap klaster yaitu

1. *Anxiety* (Kecemasan) memiliki pengaruh terbesar dalam Klaster 0 (0.23), menunjukkan bahwa kecemasan adalah faktor dominan yang mempengaruhi efek musik dalam kelompok ini terutama dalam klaster 0
2. *Insomnia* dan Depresi memiliki peran penting dalam semua klaster dengan nilai yang cukup konsisten (0.16 - 0.19), menunjukkan bahwa gangguan tidur dan tingkat depresi memainkan peran besar dalam pengalaman musik pendengar.
3. *Hours per day* (Durasi mendengarkan musik) memiliki pengaruh yang lebih rendah dalam Klaster 0 (0.094) dibandingkan dengan klaster lain, menunjukkan bahwa dalam kelompok ini, durasi mendengarkan musik bukan faktor utama yang menentukan dampak musik terhadap individu.
4. *While working* memiliki pengaruh paling rendah dalam semua klaster, menunjukkan bahwa kebiasaan mendengarkan musik saat bekerja tidak terlalu menentukan bagaimana musik mempengaruhi kondisi mental individu.

Evaluasi Akurasi Model *Random Forest*

Untuk mengukur kinerja model *Random Forest*, dilakukan evaluasi akurasi guna memastikan bahwa model dapat secara efektif memprediksi faktor utama yang mempengaruhi efek musik dalam setiap klaster. Hasil dari evaluasi dilihat pada tabel 3.

Tabel 3. Hasil akurasi *Random Forest*

| Klaster | Akurasi |
|---------|---------|
| 0 | 0,55 |
| 1 | 0,65 |
| 2 | 0,75 |
| 3 | 0,86 |

Pada tabel 3 terlihat akurasi lebih tinggi pada Klaster 2 dan 3, yang menunjukkan bahwa pola dalam klaster ini lebih jelas dan dapat dikenali oleh model *Random Forest*. Akurasi lebih rendah pada Klaster 0 dan 1, yang mungkin disebabkan oleh data yang lebih bervariasi atau faktor yang belum sepenuhnya terwakili dalam model.

4. KESIMPULAN DAN SARAN

Penelitian ini menunjukkan bahwa pendengar musik dapat dikelompokkan ke dalam empat klaster utama berdasarkan kebiasaan dan kondisi psikologis mereka. Kecemasan dan *insomnia* terbukti menjadi faktor dominan yang mempengaruhi efek musik pada individu. *Silhouette Score* sebesar 0.24 menunjukkan bahwa *clustering* cukup baik, sementara evaluasi akurasi *Random Forest* tertinggi pada Klaster 3 (86%) dan terendah pada Klaster 0 (55%), menunjukkan perbedaan pola dalam tiap kelompok. Hasil ini mengungkap bahwa musik memiliki dampak yang bervariasi tergantung pada karakteristik pendengarnya.

Untuk meningkatkan penelitian ini, disarankan mencoba metode *clustering* lain seperti DBSCAN untuk meningkatkan akurasi segmentasi. Peningkatan model dengan teknik *boosting* juga dapat membantu memperbaiki prediksi. Penelitian selanjutnya dapat menggunakan dataset yang lebih luas dan bahkan mengembangkan sistem rekomendasi musik berbasis kesehatan mental agar lebih aplikatif dalam dunia nyata.

5. DAFTAR PUSTAKA

- [1] I. Syaefudin, "Pengaruh Psikologis Musik Gambus Di Kampung Arab Surabaya," *Repertoar Journal*, vol. 3, no. 2, pp. 195–217, Feb. 2023, doi: 10.26740/rj.v3n2.p195-217.
- [2] Lussy Putri Khadijah, "Efektivitas Terapi Musik Untuk Menurunkan Tingkat Stres Dan Kecemasan," *Detector: Jurnal Inovasi Riset Ilmu Kesehatan*, vol. 1, no. 3, pp. 91–98, Jul. 2023, doi: 10.55606/detector.v1i3.2101.
- [3] A. N. Najla, "Dampak Mendengarkan Musik Terhadap Kondisi Psikologis Remaja," *Jurnal Edukasi*, vol. 1, no. 1, Dec. 2020, doi: 10.31234/osf.io/fwujn.
- [4] M. Malikhah, "Dinamika Pengaruh Musik pada Kesejahteraan Psikologis Peserta Didik: Analisis Literatur tentang Respons Neurologis dan Emosional," *Journal of Education Research*, vol. 5, no. 4, pp. 5109–5118, Oct. 2024, doi: 10.37985/jer.v5i4.1751.
- [5] H. A. Modran, T. Chamunorwa, D. Ursuțiu, C. Samoilă, and H. Hedeșiu, "Using Deep Learning to Recognize Therapeutic Effects of Music Based on Emotions," *Sensors*, vol. 23, no. 2, p. 986, Jan. 2023, doi: 10.3390/s23020986.
- [6] M. Ahmed, R. Seraj, and S. M. S. Islam, "The k-means Algorithm: A Comprehensive Survey and Performance Evaluation," *Electronics (Basel)*, vol. 9, no. 8, p. 1295, Aug. 2020, doi: 10.3390/electronics9081295.
- [7] D. B. Rarasati, "A Grouping of Song-Lyric Themes Using K-Means Clustering," *JISA(Jurnal Informatika dan Sains)*, vol. 3, no. 2, Dec. 2020, doi: 10.31326/jisa.v3i2.658.
- [8] B. Hakim, F. J. Kaunang, C. Susanto, J. Salim, and R. Indradjaja, "Implementasi Machine Learning Dalam Pengelompokan Musik Menggunakan Algoritma K-Means Clustering," *IDEALIS : InDonEsiA journal Information System*, vol. 8, no. 1, pp. 74–83, Jan. 2025, doi: 10.36080/idealisis.v8i1.3357.

- [9] Y. Xiao, Z. He, H. Liao, K. Huang, J. Zhao, and Y. Wu, "Music genre influence and artist similarity based on data analysis," *J Phys Conf Ser*, vol. 1903, no. 1, p. 012007, Apr. 2021, doi: 10.1088/1742-6596/1903/1/012007.
- [10] D. Muriyatmoko, A. Musthafa, M. Fa-Idzaa, U. Darussalam Gontor, and J. Timur Indonesia, "Seminar Nasional Amikom Surakarta (Semnasa) 2024 Perbandingan Metode Support Vector Machine Dan Random Forest Dalam Menganalisis Pengaruh Musik Terhadap Penurunan Tingkat Stress Mahasiswi Semester 7 Saat Skripsi (Studi Kasus : Universitas Darussalam Gontor)," *Seminar Nasional Amikom Surakarta (Semnasa) 2024*, vol. 2, no. 1, pp. 128–135, Nov. 2024, Accessed: Mar. 23, 2025. [Online]. Available: <https://ojs.amikomsolo.ac.id/index.php/semnasa/issue/view/2>
- [11] Nauval Satriani Siregar, "EDA & Data Preprocessing in R," https://rpubs.com/Satria_Siregar/743806.
- [12] J. Ipmawati and I. Unggara, "Analisis Status Gizi Anak Menggunakan Metode Klastering pada Dataset Anthropometri," *bit-Tech*, vol. 7, no. 2, pp. 494–504, Dec. 2024, doi: 10.32877/bt.v7i2.1869.
- [13] N. W. Utami and A. A. I. I. Paramitha, "Penerapan Data Mining Untuk Mengetahui Pola Pemilihan Program Studi Di Stmik Primakara Menggunakan Algoritma K-Means Clustering," *Jurnal Teknologi Informasi dan Komputer*, vol. 7, no. 4, Oct. 2021, doi: 10.36002/jutik.v7i4.1540.
- [14] M. Daffa Rachman and A. Voutama, "Implementasi Algoritma K-Means Dalam Sistem Rekomendasi Musik Menggunakan Python," *JATI (Jurnal Mahasiswa Teknik Informatika)*, vol. 8, no. 3, pp. 3857–3862, Jun. 2024, doi: 10.36040/jati.v8i3.9635.
- [15] K. Abdi, A. Warjaya, I. Muthmainnah, and P. H. Pahutar, "Penerapan Algoritma Random Forest dalam Prediksi Kelayakan Air Minum," *Jurnal Ilmu Komputer dan Informatika*, vol. 3, no. 2, pp. 81–88, Jan. 2024, doi: 10.54082/jiki.81.
- [16] E. Virantika, K. Kusnawi, and J. Ipmawati, "Evaluasi Hasil Pengujian Tingkat Clusterisasi Penerapan Metode K-Means Dalam Menentukan Tingkat Penyebaran Covid-19 di Indonesia," *JURNAL MEDIA INFORMATIKA BUDIDARMA*, vol. 6, no. 3, p. 1657, Jul. 2022, doi: 10.30865/mib.v6i3.4325.
- [17] X. Zhou, P. Lu, Z. Zheng, D. Tolliver, and A. Keramati, "Accident Prediction Accuracy Assessment for Highway-Rail Grade Crossings Using Random Forest Algorithm Compared with Decision Tree," *Reliab Eng Syst Saf*, vol. 200, p. 106931, Aug. 2020, doi: 10.1016/j.ress.2020.106931.