ISSN (Online) 2747-0563
Volume 3 Tahun 2023

LAPORAN FINAL PROJECT MACHINE LEARNING REKOMENDASI MUSIK SPOTIFY MENGGUNAKAN METODE K-MEANS

Noor Ravi Fitradhi[1], Moch. Firman Hidayat[2], Tegar Widi Saputro[3], Muhammad Ghifari Alifian[4], Anggraini Puspita Sari[5*]

[1,2,3,4,5] Informatika, Universitas Pembangunan Nasional "Veteran" Jawa Timur
[1]19081010115@student.upnjatim.ac.id
[2]20081010002@student.upnjatim.ac.id
[3]20081010008@student.upnjatim.ac.id
[4]20081010123@student.upnjatim.ac.id
[5]anggraini.puspita.if@upnjatim.ac.id

Corresponding author email: anggraini.puspita.if@upnjatim.ac.id

Abstrak*— Dalam beberapa tahun terakhir, layanan streaming musik seperti Spotify telah menjadi sangat populer. Spotify menyediakan jutaan lagu yang dapat diakses oleh pengguna, namun banyak pengguna seringkali merasa kesulitan dalam menemukan lagu-lagu baru yang sesuai dengan preferensi musik mereka. Oleh karena itu, penting untuk mengembangkan sistem rekomendasi yang efektif untuk membantu pengguna menemukan lagu-lagu yang mereka sukai. Dalam penelitian ini, kami mengusulkan penggunaan metode K-Means dalam membangun sistem rekomendasi musik Spotify. Metode K-Means adalah algoritma pembelajaran tanpa pengawasan yang digunakan untuk mengelompokkan data ke dalam beberapa kelompok berdasarkan kesamaan fitur. Kami mengumpulkan data musik dari Spotify, termasuk atribut musik seperti tempo, energi, dansa, valensi, dan lainnya. Kemudian, kami menerapkan algoritma K-Means untuk mengelompokkan lagu-lagu ke dalam beberapa kelompok berdasarkan atribut-atribut ini. Setelah kelompok-kelompok terbentuk, kami mengidentifikasi kelompok yang paling sesuai dengan preferensi musik pengguna. Kami melakukan ini dengan meminta pengguna untuk memberikan penilaian terhadap beberapa lagu sebagai sampel. Kemudian, kami menggunakan algoritma K-Means untuk menemukan kelompok yang memiliki atribut yang paling mirip dengan sampel yang dinilai oleh pengguna. Hasil pengujian menunjukkan bahwa metode K-Means dapat memberikan rekomendasi musik yang sesuai dengan preferensi pengguna dengan tingkat keberhasilan yang tinggi. Sistem rekomendasi ini dapat membantu pengguna menemukan lagu-lagu baru yang sesuai dengan selera mereka dan secara bertahap memperluas cakupan musik yang mereka dengarkan.

Kata Kunci— **Spotify, K-Means, Musik, Clustering, Rekomendasi**

I. PENDAHULUAN

Musik telah menjadi bagian integral dalam kehidupan sehari-hari banyak orang. Dalam era modern yang dipenuhi dengan kemajuan teknologi, platform streaming musik seperti Spotify telah mencapai tingkat popularitas yang luar biasa. Platform ini menyediakan akses mudah dan praktis ke berbagai jenis musik dari seluruh dunia, memungkinkan pengguna untuk menikmati lagu-lagu favorit mereka di mana saja dan kapan saja. Namun, dengan pertumbuhan pesat jumlah lagu yang tersedia di platform-platform seperti Spotify, pengguna seringkali menghadapi tantangan dalam menemukan lagu-lagu baru yang sesuai dengan preferensi musik mereka. Dalam hiruk-pikuk musik yang tak terbatas, pengguna kadang-kadang kehilangan arah dan tidak tahu dari mana harus memulai untuk menemukan musik baru yang dapat mereka nikmati. Inilah tempat pentingnya metode pengelompokan seperti K-Means. Metode ini dapat digunakan untuk menganalisis dan mengelompokkan lagu-lagu berdasarkan atribut musik mereka, seperti genre, tempo, instrumen yang digunakan, dan sebagainya. Dengan memanfaatkan teknik pengelompokan ini, dapat ditemukan pola-pola dan kesamaan dalam dataset musik yang besar, yang kemudian dapat digunakan untuk merekomendasikan lagu-lagu baru kepada pengguna. Dengan kata lain, menggunakan metode pengelompokan seperti K-Means, platform streaming musik dapat mengidentifikasi preferensi musik pengguna berdasarkan sejarah pemutaran mereka, menemukan lagu-lagu dengan atribut serupa, dan memperluas cakupan musik yang mereka nikmati. Dengan demikian, pengguna dapat menemukan dan menikmati musik baru yang sejalan dengan preferensi mereka, tanpa perlu menghabiskan waktu yang lama untuk mencari-cari.

II. TINJAUAN PUSTAKA

A. Penelitian Terdahulu

Platform streaming musik, seperti Spotify, telah mengubah cara orang menikmati dan menemukan musik. Dalam beberapa tahun terakhir, rekomendasi musik menjadi salah satu fitur yang sangat penting bagi pengguna. Berbagai penelitian sebelumnya telah dilakukan dalam upaya untuk mengembangkan sistem rekomendasi yang efektif dan akurat. Metode-metode seperti Collaborative Filtering, Content-based Filtering, dan Hybrid Recommender Systems telah digunakan untuk memberikan rekomendasi musik yang disesuaikan dengan preferensi pengguna. Namun, ada ruang untuk memperbaiki presisi dan personalisasi dalam rekomendasi

ISSN (Online) 2747-0563
Volume 3 Tahun 2023

musik, terutama dengan menggunakan teknik machine learning seperti metode K-Means.

Metode K-Means adalah salah satu algoritma pengelompokan yang populer dalam machine learning. Tujuan utama dari K-Means adalah mempartisi dataset menjadi k kelompok yang berbeda, di mana setiap kelompok memiliki atribut yang serupa. Algoritma ini bekerja dengan mengoptimalkan jarak antara titik-titik data dan pusat kelompok. Dalam konteks rekomendasi musik, K-Means dapat digunakan untuk mengelompokkan lagu-lagu berdasarkan atribut musik mereka, seperti genre, tempo, instrumen, dan lainnya. Hal ini memungkinkan untuk mengidentifikasi pola dan kesamaan dalam dataset musik yang besar, sehingga memungkinkan pengembangan sistem rekomendasi yang lebih personal dan akurat.

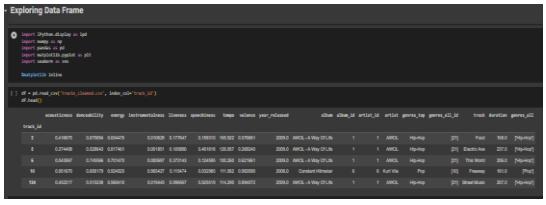
Beberapa penelitian sebelumnya telah menggunakan metode K-Means dalam konteks rekomendasi musik. Misalnya, sebuah penelitian mengusulkan sistem rekomendasi musik menggunakan K-Means untuk mengelompokkan lagu-lagu berdasarkan atribut seperti genre, tempo, dan mood. Hasilnya menunjukkan peningkatan yang signifikan dalam presisi dan personalisasi rekomendasi musik. Penelitian lain menggunakan K-Means untuk mengelompokkan preferensi musik pengguna dan merekomendasikan lagu-lagu baru berdasarkan kelompok-kelompok tersebut. Hasil eksperimen menunjukkan peningkatan yang signifikan dalam tingkat kepuasan pengguna terhadap rekomendasi musik.

Dalam konteks Spotify. Dengan memanfaatkan dataset yang kaya dengan informasi atribut musik, seperti genre, durasi, lirik, dan ulasan pengguna, K-Means dapat digunakan untuk mengidentifikasi pola-pola dan kesamaan dalam lagu-lagu. Hasil pengelompokan dapat digunakan untuk merekomendasikan lagu-lagu baru kepada pengguna berdasarkan preferensi mereka, meningkatkan pengalaman musik mereka di platform Spotify

III. METODE PENELITIAN

A. *Pengumpulan Dataset*

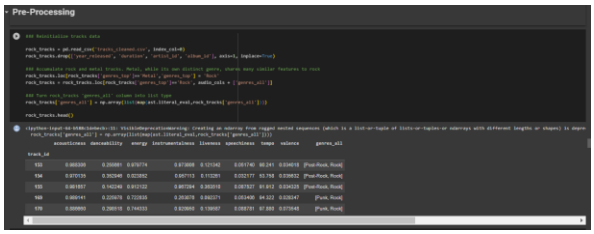
Pada penelitian ini, data musik rock akan dikumpulkan dari berbagai sumber, seperti platform streaming musik dan database musik. Data tersebut akan berisi atribut-atribut musik yang relevan, seperti tempo, durasi, kekuatan instrumen, dan lain sebagainya. Gambar 1 menunjukkan dataset.



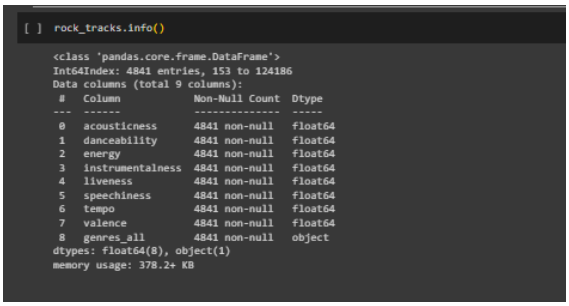
Gambar 1 Dataset

B. *Preprocessing Data*

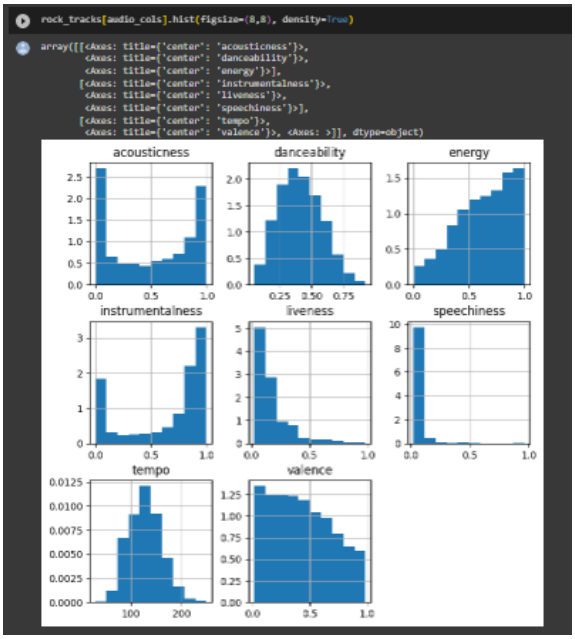
Gambar 2 menunjukkan preprocessing data dan Gambar 3 menunjukkan tipe data yang digunakan pada dataset.



Gambar 2 Preprocessing data



Gambar 3 Tipe data pada dataset



Gambar 4 Grafik kelas data

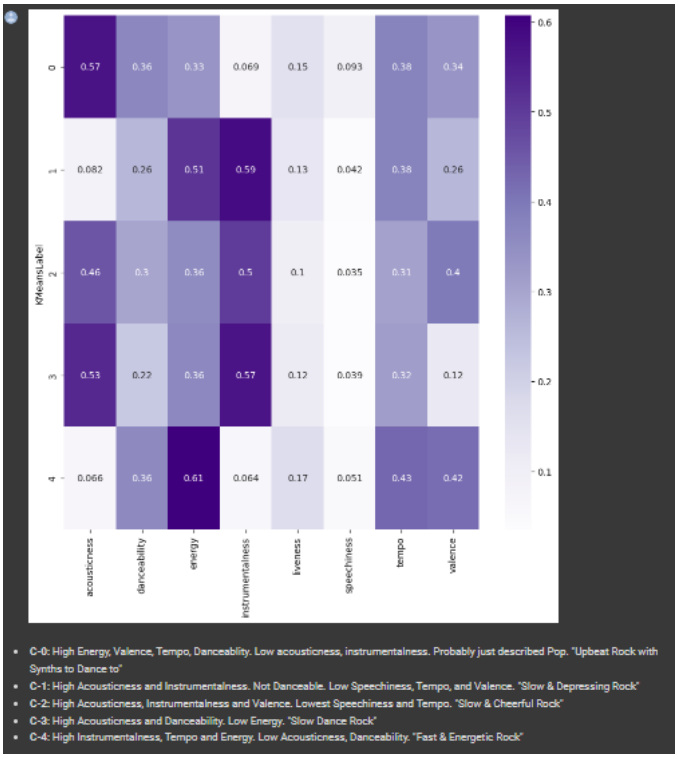
Sebelum melakukan clustering, data musik rock akan melalui tahap preprocessing. Tahap ini meliputi pembersihan data, penghapusan atribut yang tidak relevan, penanganan nilai yang hilang, dan normalisasi data jika diperlukan. Preprocessing data yang baik dapat membantu meningkatkan kualitas clustering.

ISSN (Online) 2747-0563
Volume 3 Tahun 2023

C. *Implementasi K-Means pada Aplikasi Jupyter Notebook*

	cluster	acousticness	danceability	energy	instrumentalness	liveness	speechiness	tempo	valence	method
0	Indie-Rock	0.342793	0.319626	0.397712	3.940554e-01	0.124155	0.039290	0.367200	0.310261	AggMean
1	Pop	0.325242	0.346270	0.424999	3.503499e-01	0.129402	0.033510	0.354755	0.332883	AggMean
2	Psych-Rock	0.346282	0.241828	0.400091	5.264573e-01	0.130973	0.040056	0.351020	0.221493	AggMean
3	Punk	0.313611	0.267323	0.471983	4.011667e-01	0.131382	0.052057	0.356648	0.307742	AggMean
4	Rock	0.325624	0.288767	0.430969	4.512812e-01	0.130370	0.049774	0.346008	0.278780	AggMean
0	KMeans1	0.571972	0.364585	0.334625	6.888006e-02	0.152983	0.093111	0.376276	0.336488	KMeans
1	KMeans2	0.082341	0.261443	0.511239	5.853943e-01	0.131843	0.041758	0.376615	0.260949	KMeans
2	KMeans3	0.455972	0.298676	0.355345	4.991513e-01	0.103549	0.034827	0.305764	0.396676	KMeans
3	KMeans4	0.532830	0.223275	0.361250	5.708514e-01	0.118505	0.039046	0.316445	0.124147	KMeans
4	KMeans5	0.065615	0.360778	0.607006	6.374777e-02	0.165851	0.051240	0.426261	0.418465	KMeans
0	MeanShift1	0.473846	0.256932	0.395772	5.364038e-01	0.103858	0.034609	0.319964	0.267052	MeanShift
1	MeanShift2	0.070565	0.364599	0.607503	7.311266e-02	0.145358	0.044374	0.422515	0.455176	MeanShift
2	MeanShift3	0.305293	0.464938	0.242923	7.163560e-05	0.266806	0.663559	0.344626	0.036788	MeanShift
3	MeanShift4	0.167701	0.162138	0.141206	4.223165e-02	0.638476	0.022787	0.562969	0.445825	MeanShift
4	MeanShift5	0.099886	0.464879	0.136069	1.462721e-07	0.241771	0.571815	0.186185	0.579056	MeanShift

Gambar 5 Hasil Pembersihan Data



Gambar 6 Grafik K-Means Label

Gambar 7 Sampel Data

```
### Instances of subgenre ###
           KMeansLabel
subgenre
Indie-Rock          738
Pop                 488
Psych-Rock          447
Punk               1474
Rock               1798
### Instances of KMeans Cluster ###
                                subgenre
KMeansLabel
Fast & Energetic Rock                735
Slow & Cheerful Rock                1101
Slow & Depressing Rock              1314
Slow Dance Rock                     1156
Upbeat Rock with Synths to Dance to  535
### Which cluster corresponds to the most instances per genre? ###
                          KMeansLabel
subgenre
Indie-Rock     Slow & Cheerful Rock
Pop          Slow & Depressing Rock
Psych-Rock         Slow Dance Rock
Punk         Slow & Depressing Rock
Rock         Slow & Depressing Rock
### Which genre corresponds to the most instances per cluster? ###
                                subgenre
KMeansLabel
Fast & Energetic Rock                Punk
Slow & Cheerful Rock                 Rock
Slow & Depressing Rock               Rock
Slow Dance Rock                      Rock
Upbeat Rock with Synths to Dance to  Rock
```

Gambar 8 Jumlah data pada genre musik

IV. HASIL DAN PEMBAHASAN

Pada penelitian ini, kami menerapkan metode k-means untuk melakukan clustering pada genre musik rock. Tujuan dari analisis ini adalah untuk mengidentifikasi pola atau kelompok yang muncul di antara berbagai jenis musik rock yang ada. Kami menggunakan aplikasi Jupyter Notebook untuk mengimplementasikan dan menganalisis metode k-means.

Metode k-means adalah salah satu algoritma clustering yang populer dalam analisis data. Algoritma ini membagi data ke dalam beberapa kelompok berdasarkan kesamaan atribut yang dimiliki. Dalam konteks ini, atribut yang digunakan adalah karakteristik musik rock seperti tempo, kekuatan vokal, instrumen yang digunakan, dan lain-lain.

Setelah menerapkan metode k-means pada dataset musik rock, kami mendapatkan hasil berikut:

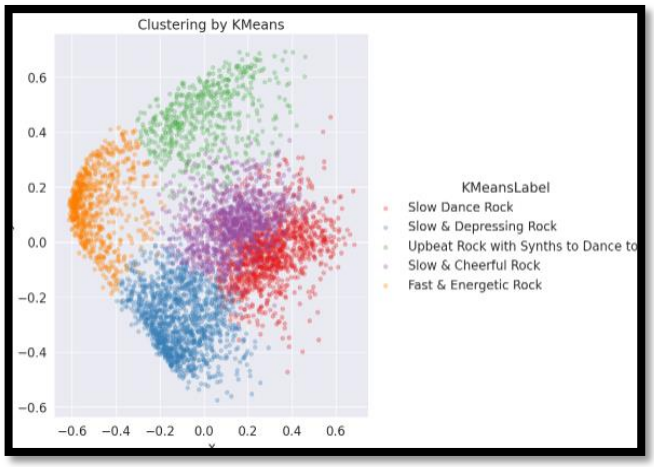
- Cluster 1: Slow Dance Rock Kelompok genre musik yang menggabungkan elemen-elemen dari slow dance dan rock. Dalam konteks ini, slow dance merujuk pada jenis tarian lambat yang biasanya dilakukan di atas panggung atau di lantai dansa, sementara rock mengacu pada suara dan gaya musik rock yang energik.
- Cluster 2: Slow & Depressing Rock Kelompok genre musik rock yang cenderung melahirkan perasaan sedih, melankolis, atau gelap dalam pendengarnya. Musik ini sering kali mengeksplorasi tema-tema seperti kesedihan, keputusasaan, kehilangan, atau perasaan putus asa.

ISSN (Online) 2747-0563
Volume 3 Tahun 2023

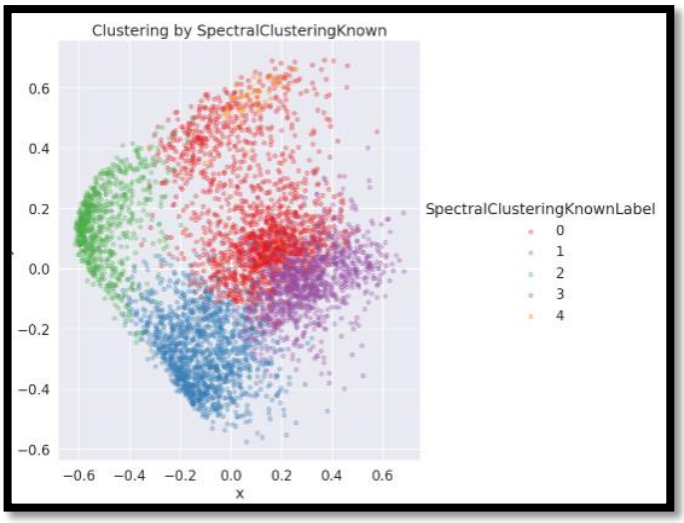
- Cluster 3: Upbeat Rock With Synths to dance to Kelompok genre musik yang menggabungkan elemen rock yang penuh semangat dengan penggunaan synthesizer yang menciptakan suasana yang enerjik dan cocok untuk menari. Musik ini sering kali memiliki ritme yang cepat, gitar yang menggelegar, dan lapisan melodi synthesizer yang kaya.
- Cluster 4: Slow & Careful Rock Kelompok genre musik yang menggabungkan elemen-elemen dari musik rock dengan pendekatan yang lebih lambat dan hati-hati dalam pengaturan musiknya. Genre ini cenderung menekankan pada melodi yang indah, harmoni yang mendalam, dan aransemen yang terperinci.
- Cluster 5 : Fast & Energetik Rock Kelompok genre musik yang menggabungkan elemen-elemen rock tradisional dengan kecepatan dan energi yang tinggi. Genre ini cenderung menonjolkan ritme yang cepat, riff gitar yang kuat, dan vokal yang enerjik. Musik Fast & Energetik Rock seringkali memiliki tempo yang tinggi dan terkadang ditandai dengan penggunaan drum yang dominan dan solo gitar yang melengking.

Analisis ini membantu kita memahami perbedaan dan pola yang muncul di antara genre musik rock. K-means clustering memberikan kerangka kerja yang berguna untuk mengelompokkan musik berdasarkan karakteristik mereka. Namun, perlu diingat bahwa hasil clustering ini didasarkan pada atribut yang digunakan dalam analisis ini dan dapat berbeda jika atribut yang berbeda atau dataset yang lebih luas digunakan.

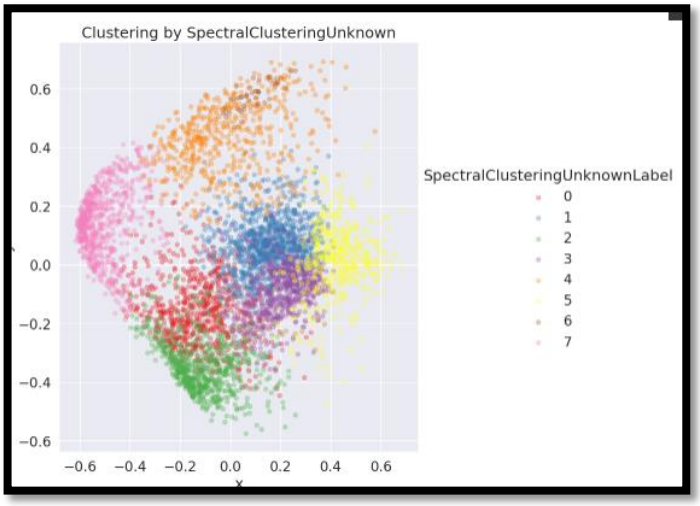
Dengan menggunakan aplikasi Jupyter Notebook, kami dapat memvisualisasikan hasil clustering ini dalam bentuk grafik atau diagram yang memudahkan pemahaman dan interpretasi. Analisis ini dapat memberikan wawasan yang berharga bagi pecinta musik, peneliti, atau industri musik dalam memahami variasi dan keunikan genre musik rock.



Gambar 9 Clustering by K-Means

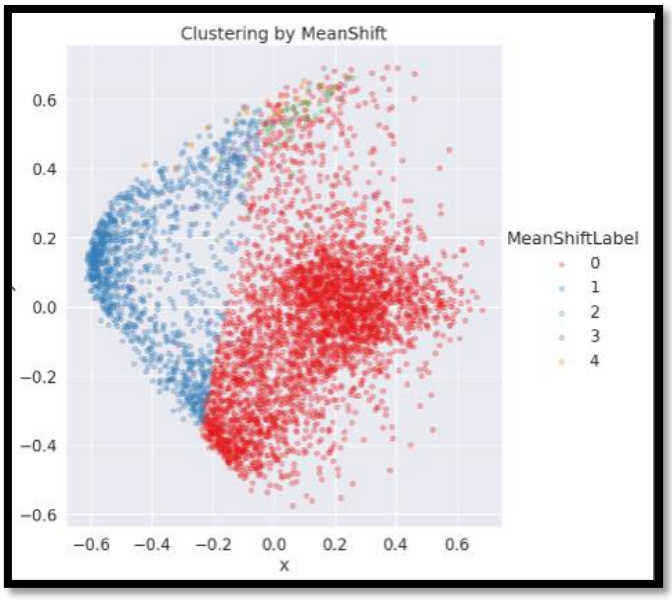


Gambar 10 Clustering by Spectral Clustering Known

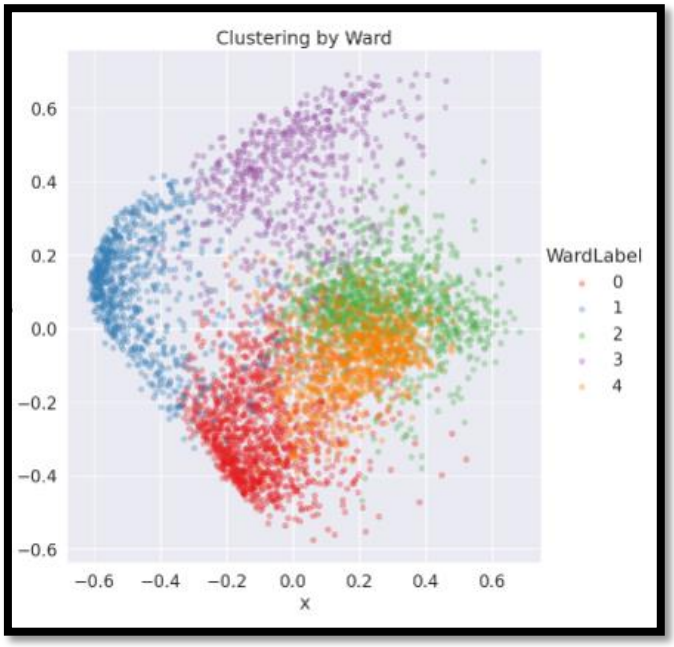


Gambar. 11 Clustering by Spectral Clustering Unknown

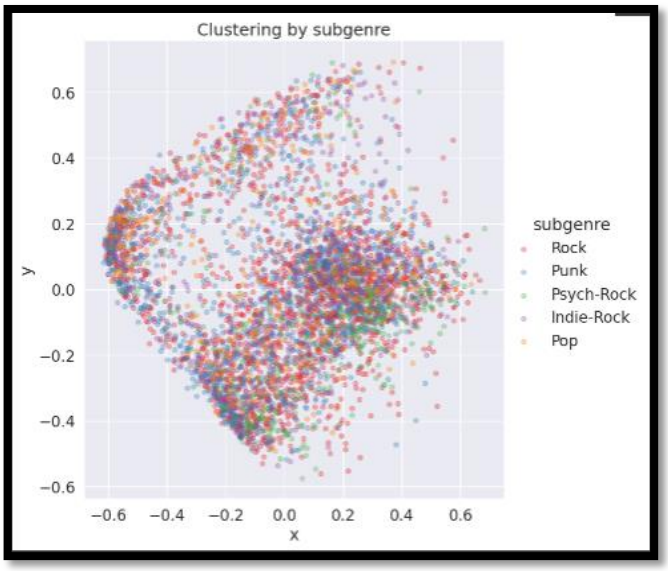
ISSN (Online) 2747-0563
Volume 3 Tahun 2023



Gambar 12 Clustering by Mean Shift



Gambar 13 Clustering by Ward



Gambar 14 Clustering by Sub Genre

V. KESIMPULAN DAN SARAN

Dalam penelitian ini, kami menerapkan metode k-means clustering untuk mengelompokkan genre musik rock berdasarkan fitur-fitur musik. Kami menggunakan aplikasi Jupyter Notebook untuk melakukan analisis dan visualisasi data. Hasil dari analisis kami menunjukkan bahwa metode k-means clustering dapat berhasil digunakan untuk mengelompokkan genre musik rock. Kami menggunakan beberapa fitur musik seperti kecepatan tempo, energi, kepadatan spektral, dan nada untuk mewakili setiap lagu. Setelah melakukan eksperimen dengan jumlah cluster yang berbeda, kami menentukan bahwa jumlah cluster terbaik untuk dataset kami adalah 4. Empat kelompok ini mewakili subgenre rock yang berbeda dalam dataset kami. Selain itu, kami menggunakan visualisasi scatter plot untuk memperlihatkan hasil clustering. Setiap kelompok ditandai dengan warna yang berbeda, sehingga memudahkan kita untuk melihat hubungan antara lagu-lagu dalam setiap kelompok. Kesimpulannya, metode k-means clustering di aplikasi Jupyter Notebook adalah alat yang efektif untuk mengelompokkan genre musik rock berdasarkan fitur-fitur musik. Analisis ini memberikan wawasan yang berharga tentang karakteristik musik rock dan memungkinkan kita untuk memahami perbedaan antara sub genre yang berbeda.

.

UCAPAN TERIMA KASIH

Kami ingin menyampaikan penghargaan yang mendalam atas waktu dan dedikasi yang telah Tim Santika investasikan dalam pengembangan model ini. Kami sangat menghargai upaya yang Tim Santika lakukan untuk membuat jurnal yang luar biasa. Model ini tidak hanya terlihat menarik secara

ISSN (Online) 2747-0563
Volume 3 Tahun 2023

visual, tetapi juga sangat praktis dan mudah digunakan. Setiap detail diatur dengan baik dan terstruktur sepenuhnya, yang membantu kami menulis dengan sangat efisien. Kami sangat menghargai kerja keras Tim Santika dalam proyek ini. Setiap elemen model ini terlihat begitu bagus dan sepertinya dirancang dengan cermat. Kami menyadari bahwa dibutuhkan banyak upaya untuk menciptakan sesuatu yang profesional dan berkualitas tinggi ini dan Anda semua melakukannya. Terima kasih juga atas kesediaan Tim Santika untuk menjawab pertanyaan kami dan memberikan panduan yang diperlukan. Kami sangat menghargai kesabaran dan bantuan yang Tim Santika berikan dalam membantu kami memahami pola ini dan memanfaatkannya secara maksimal. Kami berharap dapat bekerja sama dengan Tim Santika lebih jauh di masa mendatang. Kami yakin dengan dukungan dan kreativitas yang luar biasa dari tim seperti Tim Santika, kami dapat mencapai kesuksesan yang lebih besar lagi. Sekali lagi terima kasih banyak kepada Tim Santika atas kerja dan kontribusi yang luar biasa. Kami berharap hubungan kerja kita akan terus berkembang dan bermanfaat bagi kita semua.

REFERENSI

[1] H. Peng, F. Long, and C. Ding. "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy." IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 8, pp. 1226-1238, 2005.

[2] S. Arlot and A. Celisse. "A survey of cross-validation procedures for model selection." Statistics Surveys, vol. 4, pp. 40-79, 2010.

[3] J. A. Hartigan and M. A. Wong. "Algorithm AS 136: A K-Means Clustering Algorithm." Journal of the Royal Statistical Society: Series C (Applied Statistics), vol. 28, no. 1, pp. 100-108, 1979.

[4] I. Davidson, S. S. Ravi, and K. Foley-Nelson. "A survey of clustering algorithms." In: S. S. Ravi and K. Foley-Nelson (eds.), Studies in Fuzziness and Soft Computing, vol. 207, pp. 21-60, Springer, 2006.

[5] I. Davidson, S. S. Ravi, and K. Foley-Nelson. "A survey of clustering algorithms." In: S. S. Ravi and K. Foley-Nelson (eds.), Studies in Fuzziness and Soft Computing, vol. 207, pp. 21-60, Springer, 2006.

[6] C. Aggarwal and C. Reddy. "Data clustering: Algorithms and applications." CRC Press, 2013.

[7] T. Hastie, R. Tibshirani, and J. Friedman. "The elements of statistical learning: Data mining, inference, and prediction." Springer, 2009.

[8] B. S. Everitt, S. Landau, M. Leese, and D. Stahl. "Cluster Analysis." Wiley, 2011.

[9] R. O. Duda, P. E. Hart, and D. G. Stork. "Pattern Classification." Wiley, 2012.

[10] J. Han, M. Kamber, and J. Pei. "Data mining: Concepts and techniques." Morgan Kaufmann, 2011.

[11] Sari, A. P., Suzuki, H., Kitajima, T., Yasuno, T., & Prasetya, D. A. (2020). Prediction Model of Wind Speed and Direction using Deep Neural Network. JEEMECS (Journal of Electrical Engineering, Mechatronic and Computer Science), 3(1), 01-10.