

Analisis Pengelompokan Atribut Musik menggunakan Algoritma K-means++ untuk Top Hits Spotify dari tahun 2000 hingga 2019

Proposal Tugas Akhir

Diajukan Untuk Memenuhi
Persyaratan Guna Meraih Gelar Sarjana
Informatika Universitas Muhammadiyah Malang



Alviatul Nazila
202010370311293

Bidang Minat

Data Science

**PROGRAM STUDI INFORMATIKA
FAKULTAS TEKNIK
UNIVERSITAS MUHAMMADIYAH MALANG
2024**

LEMBAR PERSETUJUAN

Analisis Pengelompokan Atribut Musik menggunakan Algoritma K-means++ untuk Top Hits Spotify dari tahun 2000 hingga 2019

TUGAS AKHIR

**Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang**



Menyetujui,
Malang, 31 Mei 2024

Dosen Pembimbing 1



Ir. Wildan Suharso S.Kom., M.Kom

NIP. 10817030596PNS.

Dosen Pembimbing 2



Didih Rizki Chandranegara S.kom.,

M.Kom

NIP. 180302101992PNS.

LEMBAR PENGESAHAN

**Analisis Pengelompokan Atribut Musik menggunakan Algoritma
K-means++ untuk Top Hits Spotify dari tahun 2000 hingga 2019**

TUGAS AKHIR

Sebagai Persyaratan Guna Meraih Gelar Sarjana Strata 1
Informatika Universitas Muhammadiyah Malang

Disusun Oleh :

Alviatul Nazila

202010370311293

Tugas Akhir ini telah diuji dan dinyatakan lulus melalui sidang majelis penguji
pada tanggal 31 Mei 2024

Menyetujui,

Dosen Penguji 1



Setio Basuki MT., Ph.D.

NIP. 10809070477PNS.

Dosen Penguji 2



Christian Sri Kusuma Aditya

S.Kom., M.Kom

NIP. 180327021991PNS.

Mengetahui,

Ketua Jurusan Informatika



Ir. Galih Wasis Wicaksono S.kom. M.Cs.

NIP. 10814100541PNS.

LEMBAR PERNYATAAN

Yang bertanda tangan dibawah ini :

NAMA : Alviatul Nazila

NIM : 202010370311293

FAK./JUR. : Informatika

Dengan ini saya menyatakan bahwa Tugas Akhir dengan judul **“Analisis Pengelompokan Atribut Musik menggunakan Algoritma K-means++ untuk Top Hits Spotify dari tahun 2000 hingga 2019”** beserta seluruh isinya adalah karya saya sendiri dan bukan merupakan karya tulis orang lain, baik sebagian maupun seluruhnya, kecuali dalam bentuk kutipan yang telah disebutkan sumbernya.

Demikian surat pernyataan ini saya buat dengan sebenar-benarnya. Apabila kemudian ditemukan adanya pelanggaran terhadap etika keilmuan dalam karya saya ini, atau ada klaim dari pihak lain terhadap keaslian karya saya ini maka saya siap menanggung segala bentuk resiko/sanksi yang berlaku.

Mengetahui,
Dosen Pembimbing



Ir. Wildan Suharso S.Kom., M.Kom

Malang, 31 Mei 2024
Yang Membuat Pernyataan



Alviatul Nazila
alviatul Nazila

ABSTRAK

Penelitian ini bertujuan untuk menganalisis pengelompokan atribut musik dari lagu-lagu yang menjadi hits di Spotify selama periode 2000 hingga 2019 menggunakan algoritma K-means++. Dataset yang digunakan dalam penelitian ini diperoleh dari Kaggle, yang berjudul "Top Hits Spotify from 2000-2019" dan mencakup atribut musik seperti durasi, popularitas, danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, liveness, valence, dan tempo. Hasil penelitian ini menunjukkan bahwa terdapat tiga kelompok klaster yang berbeda berdasarkan atribut musik yang dimilikinya. Selain itu, penelitian ini juga mengungkapkan bahwa atribut musik seperti energy, loudness, dan valence yang tinggi dapat mempengaruhi kesuksesan sebuah lagu. Ini mengindikasikan bahwa karakteristik musik seperti tingkat energi, tingkat keras suara, tingkat kebahagiaan, dapat berperan penting dalam menarik minat audiens dan meningkatkan popularitas sebuah lagu.

Kata Kunci: *K-means++*, *Spotify*, *atribut musik*, *clustering*, *analisis data*, *algoritma*, *popularitas lagu*.

ABSTRACT

This study aims to analyze the clustering of musical attributes of songs that became hits on Spotify during the period from 2000 to 2019 using the K-means++ algorithm. The dataset used in this research was obtained from Kaggle, titled "Top Hits Spotify from 2000-2019," and includes musical attributes such as duration, popularity, danceability, energy, key, loudness, mode, speechiness, acousticness, instrumentalness, liveness, valence, and tempo. The results of this study show that there are three distinct clusters based on the musical attributes. Additionally, this study reveals that musical attributes such as high energy, loudness, and valence can influence a song's success. This indicates that musical characteristics like energy levels, loudness, and happiness can play a significant role in attracting audience interest and increasing a song's popularity.

Keywords: *K-means++*, *Spotify*, *musical attributes*, *clustering*, *data analysis*, *algorithm*, *song popularity*.

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan proposal tugas akhir yang berjudul "Analisis Pengelompokan Atribut Musik menggunakan Algoritma K-means++ untuk Top Hits Spotify dari tahun 2000 hingga 2019" ini dengan baik dan lancar. Proposal ini disusun sebagai salah satu syarat untuk menyelesaikan program studi Strata 1 (S1) di Program Studi Informatika, Fakultas Teknik, Universitas Muhammadiyah Malang. Dalam penyusunan proposal ini, penulis banyak mendapatkan bantuan, bimbingan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Bapak Ir. Wildan Suharso S.Kom., M.Kom dan Didih Rizki Chandranegara S.kom., M.Kom selaku dosen pembimbing yang telah memberikan bimbingan, arahan, dan dukungan dalam penyusunan proposal ini.
2. Ibu saya, yang selalu memberikan doa, dukungan, dan motivasi yang tiada henti.
3. Rekan-rekan yang telah memberikan bantuan dan semangat selama proses penyusunan proposal ini.
4. Semua pihak yang tidak dapat disebutkan satu per satu yang telah membantu dan mendukung penulis dalam menyelesaikan proposal ini.

Penulis menyadari bahwa proposal ini masih jauh dari sempurna, baik dari segi penyusunan maupun isi. Oleh karena itu, penulis mengharapkan kritik dan saran yang membangun dari pembaca demi perbaikan dan penyempurnaan proposal ini di masa mendatang. Akhir kata, penulis berharap semoga proposal ini dapat memberikan manfaat dan kontribusi positif bagi perkembangan ilmu pengetahuan, khususnya dalam bidang analisis data dan pengelompokan atribut musik.

Malang, 09 Juni 2024

Alviatul Nazila

DAFTAR ISI

ABSTRAK	2
KATA PENGANTAR	4
DAFTAR ISI	5
DAFTAR GAMBAR	7
DAFTAR TABEL	8
BAB I PENDAHULUAN	9
1.1 Latar Belakang.....	9
1.2 Rumusan Masalah	12
1.3 Tujuan Penelitian.....	13
1.4 Batasan Masalah.....	13
BAB II TINJAUAN PUSTAKA	14
2.1 Penelitian terdahulu.....	14
2.2 Spotify	15
2.3 Atribut Musik	16
2.4 Unsupervised Learning.....	18
2.5 Clustering	18
2.6 K-means++ Clustering	19
2.7 Metode elbow	21
BAB III METODOLOGI PENELITIAN	23
3.1 Persiapan Data	23
3.1.1 Pengumpulan Data.....	23
3.1.2 Penyimpanan Data.....	24
3.2 Data Cleaning	24
3.2.1 Penghapusan data duplikat	24
3.2.2 Mengatasi data yang hilang	25
3.2.3 Seleksi Fitur.....	26
3.3 Normalisasi Data	27
3.4 K-means++ Clustering	28
3.4.1 Pemilihan jumlah kluster	28
3.4.2 Implementasi K-means++	30

BAB IV HASIL DAN PEMBAHASAN	32
4.1 Persiapan data	32
4.2 Data Cleaning	33
4.2.1 Penghapusan data duplikat	33
4.2.2 Mengatasi data yang hilang	34
4.2.3 Seleksi fitur.....	35
4.3 Normalisasi Data	37
4.4 K-means++ Clustering	38
4.4.1 Pemilihan jumlah kluster	38
4.4.2 Implementasi K-means++	41
BAB V KESIMPULAN	44
DAFTAR PUSTAKA.....	45



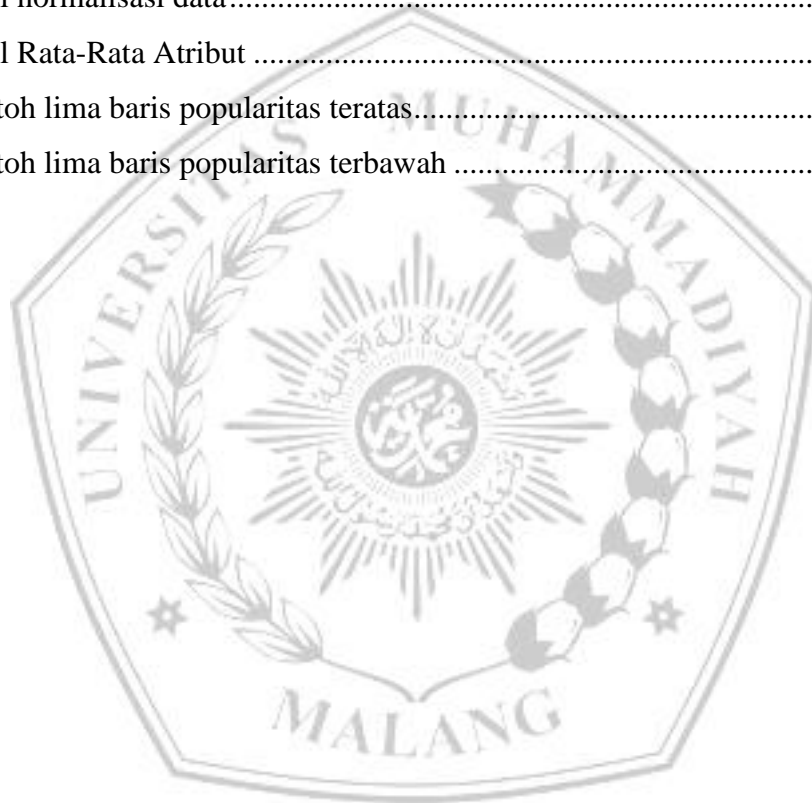
DAFTAR GAMBAR

Gambar 3.1 Tahapan Clustering	23
Gambar 3.2 Contoh data CSV	24
Gambar 3.3 Contoh Hapus data Duplikat	25
Gambar 3.4 Contoh Penghapusan Baris	26
Gambar 3.5 Contoh Metode Filter	27
Gambar 3.6 Contoh Normalisasi Data	28
Gambar 3.7 Contoh Hasil Metode Elbow	29
Gambar 3.8 Contoh Implementasi K-means++	31
Gambar 4.1 Visualisasi Matriks Korelasi	36
Gambar 4.2 Hasil Metode Elbow	39



DAFTAR TABEL

Tabel 2.1 Penelitian Terdahulu	14
Tabel 2.2 Atribut Musik.....	17
Tabel 1 Tabel 2.3 Perbedaan K-means dan K-means++	20
Tabel 4.1 Contoh lima baris teratas pada dataset.....	32
Tabel 4.2 Contoh hasil data duplikat pada dataset.....	33
Tabel 4.3 Contoh hasil data hilang pada datase	34
Tabel 4.4 Hasil seleksi fitur	37
Tabel 4.5 Rentang nilai pada dataset	37
Tabel 4.6 Hasil normalisasi data.....	38
Tabel 4.7 Hasil Rata-Rata Atribut	41
Tabel 4.8 Contoh lima baris popularitas teratas.....	42
Tabel 4.9 Contoh lima baris popularitas terbawah	42



DAFTAR PUSTAKA

- [1] Rokach L, Maimon O. Clustering methods.
- [2] Wicht LA, Waldfogel J, Waldfogel S. Playlisting Favorites: Is Spotify Gender-Biased? Joint Research Centre (Seville site); 2018 Dec.
- [3] Aguiar L, Waldfogel J, Waldfogel S. Playlisting favorites: Is Spotify gender-biased?. JRC Digital Economy Working Paper; 2018.
- [4] Bahmani B, Moseley B, Vattani A, Kumar R, Vassilvitskii S. Scalable k-means++. arXiv preprint arXiv:1203.6402. 2012 Mar 29.
- [5] Al-Beitawi Z, Salehan M, Zhang S. Cluster analysis of musical attributes for top trending songs.
- [6] Privandhani NA. Clustering Pop Songs Based On Spotify Data Using K-Means And K-Medoids Algorithm. Jurnal Mantik. 2022 Jul 5;6(2):1542-50.
- [7] Resmadi I. Jurnalisme musik dan selingkar wilayahnya. Kepustakaan Populer Gramedia; 2021 Dec 15.
- [8] Suharso W, Arifianto S, Wibowo H, Chandranegara DR, Syaifuddin S. Music Features Pada Bidang Ilmu Komputer Menggunakan Modularity Clustering. JUST IT: Jurnal Sistem Informasi, Teknologi Informasi dan Komputer. 2022 Dec 15;13(1).
- [9] Netti SY, Irwansyah I. Spotify: Aplikasi Music Streaming untuk Generasi Milenial. Jurnal Komunikasi. 2018 Jul 31;10(1):1-6.
- [10] Wibowo H, Suharso W, Azhar Y, Wicaksono GW, Minarno AE, Harmanto D. Music information retrieval based on active frequency. Makara Journal of Technology. 2021;25(2):5.
- [11] Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. Journal of computational and applied mathematics. 1987 Nov 1;20:53-65.
- [12] FALAH I FM. PENERAPAN METODE CLUSTERING UNTUK PENGELOMPOKAN MAHASISWA POTENSIAL DROP OUT MENGGUNAKAN.
- [13] Arthur D, Vassilvitskii S. K-means++ the advantages of careful seeding. In Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms 2007 Jan 7 (pp. 1027-1035).
- [14] Datasans. Machine Learning Cheatsheet : Unsupervised Model (Clustering). Instagram Datasans. <https://www.instagram.com/datasans.book/>. Diakses pada 19 September 2023.
- [15] Sinaga KP, Yang MS. Unsupervised K-means clustering algorithm. IEEE access. 2020 Apr 20;8:80716-27.

- [16]Netti SY, Irwansyah I. Spotify: Aplikasi Music Streaming untuk Generasi Milenial. *Jurnal Komunikasi*. 2018 Jul 31;10(1):1-6.
- [17]LIONO MS. STRATEGI LEMBAGA PENYIARAN RADIO MEMPERTAHANKAN PENDENGAR DI TENGAH BERKEMBANGNYA STREAMING MUSIK BERBAYAR SPOTIFY (Studi Pada Lembaga Penyiaran Radio D! Radio Lampung)(Skripsi) Oleh: MEDA SANDENA LIONO JURUSAN ILMU KOMUNIKASI FAKULTAS ILMU SOSIAL DAN ILMU POLITIK UNIVERSITAS LAMPUNG BANDAR LAMPUNG 2022.
- [18]Nuraeni F, Kurniadi D, Dermawan GF. Pemetaan Karakteristik Mahasiswa Penerima Kartu Indonesia Pintar Kuliah (KIP-K) menggunakan Algoritma K-Means++. *Jurnal Sisfokom (Sistem Informasi Dan Komputer)*. 2023 Jan 25;11(3):437-43.
- [19]Sinaga, K.P. and Yang, M.S., 2020. Unsupervised K-means clustering algorithm. *IEEE access*, 8, pp.80716-80727.
- [20]Windarto, A.P., Hasan Siregar, M.N., Suharso, W., Fachri, B., Supriyatna, A., Carolina, I., Efendi, Y. and Toresa, D., 2019, August. Analysis of the K-Means Algorithm on Clean Water Customers Based on the Province. In *Journal of Physics: Conference Series* (Vol. 1255, No. 1, p. 012001). IOP Publishing.
- [21]Dewi, D.A.I.C. and Pramita, D.A.K., 2019. Analisis Perbandingan Metode Elbow dan Silhouette pada Algoritma Clustering K-Medoids dalam Pengelompokan Produksi Kerajinan Bali. *Matrix: Jurnal Manajemen Teknologi dan Informatika*, 9(3), pp.102-109.
- [22]Nuriska D, Irawan B, Bahtiar A, Dikananda AR. KLASTERISASI DATA LAGU TERPOPULER SPOTIFY 2023 BERDASARKAN SUASANA HATI MENGGUNAKAN ALGORITMA K-MEANS. *JATI (Jurnal Mahasiswa Teknik Informatika)*. 2023;7(6):3843-50.
- [23]Sari FI, Sazali H. ANALISIS PENGGUNAAN PODCAST PADA APLIKASI SPOTIFY SEBAGAI MEDIA PEMBELAJARAN DAN INFORMASI BAGI MAHASISWA ILMU PERPUSTAKAAN UIN SUMATERA UTARA MEDAN. *Kabillah (Journal of Social Community)*. 2023 Jun 27;8(1):139-49.
- [24]Navisa S, Hakim L, Nabilah A. Komparasi Algoritma Klasifikasi Genre Musik pada Spotify Menggunakan CRISP-DM. *Jurnal Sistem Cerdas*. 2021 Aug 31;4(2):114-25.
- [25]Mayuna, T. and Witanti, A., 2023. Identifikasi Genre Musik pada Platform Spotify dengan Menggunakan Metode K-Nearest NeighboR (KNN). *Jurnal Syntax Admiration*, 4(12), pp.2277-2285.

- [26]Kaggle. (n.d.). In Wikipedia. Retrieved February 28, 2024, from <https://en.wikipedia.org/wiki/Kaggle>
- [27]Nugroho N, Adhinata FD. Penggunaan Metode K-Means dan K-Means++ Sebagai Clustering Data Covid-19 di Pulau Jawa. *Teknika*. 2022 Oct 20;11(3):170-9.
- [28]Fikriah, F.K., Hayati, N. and No, J.K., Feature Selection Dengan Komparasi Algoritma Untuk Prediksi Telemarketing Bank. *Jurnal Ilmu Komputer*.
- [29] Campello, R.J., Kröger, P., Sander, J. and Zimek, A., 2020. Density-based clustering. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(2), p.e1343.





UNIVERSITAS
MUHAMMADIYAH
MALANG



FAKULTAS TEKNIK

INFORMATIKA

informatika.umm.ac.id | informatika@umm.ac.id

FORM CEK PLAGIARISME LAPORAN TUGAS AKHIR

Nama Mahasiswa : Alviatul Nazila

NIM : 202010370311293

Judul TA : Analisis Pengelompokan Atribut Musik menggunakan Algoritma K-means++ untuk Top Hits Spotify dari tahun 2000 hingga 2019

Hasil Cek Plagiarisme dengan Turnitin

No.	Komponen Pengecekan	Nilai Maksimal Plagiarisme (%)	Hasil Cek Plagiarisme (%) *
1.	Bab 1 – Pendahuluan	10 %	3%
2.	Bab 2 – Daftar Pustaka	25 %	4%
3.	Bab 3 – Analisis dan Perancangan	25 %	2%
4.	Bab 4 – Implementasi dan Pengujian	15 %	2%
5.	Bab 5 – Kesimpulan dan Saran	5 %	0%
6.	Makalah Tugas Akhir	20%	0%

*) Hasil cek plagiarisme diisi oleh pemeriksa (staf TU)

*) Maksimal 5 kali (4 Kali sebelum ujian, 1 kali sesudah ujian)

Mengetahui,

Pemeriksa (Staff TU)



Kampus I
Jl Dendung 1 Malang Jawa Timur
P +62 341 551 253 (Hunting)
F +62 341 460 435

Kampus II
Jl Dendungan Sutami No 188 Malang, Jawa Timur
P +62 341 551 149 (Hunting)
F +62 341 582 060

Kampus III
Jl Raya Tlogomas No 246 Malang, Jawa Timur
P +62 341 464 318 (Hunting)
F +62 341 460 435
E webmaster@umm.ac.id