

Klasifikasi Genre Musik Menggunakan Support Vector Machine Berdasarkan Spectral Features

I Gusti Agung Ngurah Diputra Wiraguna^{a1}, Luh Arida Ayu Rahning Putri^{a2}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Udayana
Jalan Raya Kampus Udayana, Bukit Jimbaran, Kuta Selatan, Badung, Bali Indonesia
¹ngurahdiputra075@student.unud.ac.id
rahningputri@unud.ac.id

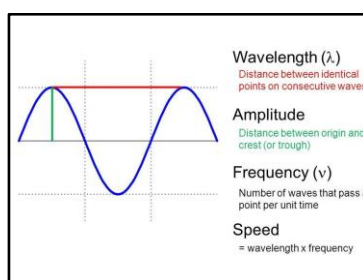
Abstract

This research focuses on music genre classification based on spectral features and Support Vector Machine (SVM). Features such as Spectral Centroid, Spectral Rolloff, Spectral Flux, and Spectral Bandwidth are extracted from MP3 music audio. The dataset comprising 4 music genres is utilized for training and testing the system. The extracted spectral features are fed into the SVM classifier to predict the genre of test samples. Python and machine learning are both used in developing the system while the experimental results demonstrate the effectiveness of SVM in accurately classifying music genres based on the current extracted features. The proposed approach contributes to automated music genre classification systems, facilitating music organization, recommendation, and retrieval. This research promotes advancements in music information retrieval and enhances user experience in music-related applications.

Keywords: Music Feature Extraction, MP3, Music, Spectral Features, SVM

1. Pendahuluan

Musik merupakan salah satu bentuk seni yang melibatkan indra pendengaran manusia. Konsep musik sama halnya dengan gelombang bunyi dimana terjadi perubahan tekanan udara dan gelombang akan terus bertransmisi hingga mencapai telinga manusia. Gelombang bunyi tergolong gelombang longitudinal sehingga sudut antara arah getaran dan arah rambatan gelombangnya adalah nol. Selama perambatannya gelombang longitudinal akan membentuk pola rapatan dan regangan yang menjadi pemklasifikasian satu gelombang longitudinal. Gelombang bunyi juga memiliki komponen gelombang pada umumnya seperti amplitudo, frekuensi, periode, fase, panjang gelombang, dan cepat rambat gelombang [11].



Gambar 1. Gelombang Bunyi dan Komponenya

Perbedaan antara musik dengan gelombang bunyi pada umumnya terletak pada karakteristik gelombang tersebut. Dalam musik dikenal beberapa istilah seperti melodi, ritme, harmoni, dan lainnya yang memiliki karakteristik berbeda dengan gelombang bunyi pada umumnya sehingga terkesan enak didengar. Karakteristik setiap gelombang bunyi berbeda beda bergantung pada frekuensi gelombang tersebut. Dalam mendengarkan musik sering disadari bahwa musik keluaran sekarang terdengar berbeda dengan musik beberapa tahun lalu. Bahkan bisa saja

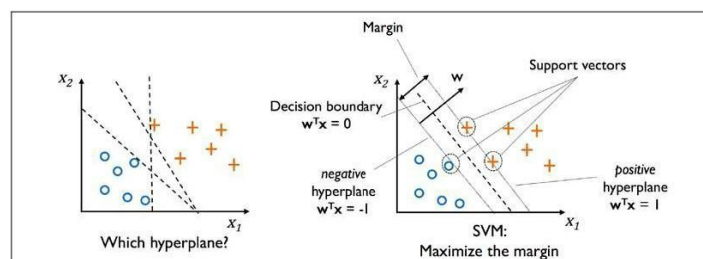
musik - musik yang dirilis pada tahun yang sama terdengar berbeda. Perbedaan dari musik tersebut terletak dari gaya pembawaan musik. Perbedaan gaya musik biasa disebut sebagai genre musik. Setiap genre musik berbeda - beda sesuai dengan musisinya. Contoh musik klasik yang memiliki gaya musik yang tenang dan volume suara tergolong rendah. Sedangkan musik rock memiliki volume suara yang keras dan tempo yang lebih cepat. Semakin berkembangnya teknologi, musik tidak harus didengarkan secara langsung dari sumber suara namun juga dapat didengarkan jarak jauh. Musik dapat disimpan dalam sebuah file dengan berbagai jenis format seperti WAV dan MP3 yang masing - masing memiliki kelebihan tersendiri. Terdapat banyak platform musik yang memungkinkan kita untuk mendengarkan musik secara gratis bahkan mengunduh file musik tersebut untuk disimpan dan didengarkan kapan saja kita mau. Sebuah platform musik digital biasanya memberi pengguna fitur untuk mencari jenis musik yang mereka inginkan. Selain itu juga terdapat fitur rekomendasi musik yang menampilkan beberapa musik yang sekiranya mungkin dinikmati oleh pengguna. Sistem rekomendasi membagi musik ke dalam beberapa kategori salah satunya genre. Platform tersebut mampu memberi rekomendasi dengan genre yang sesuai karena sistem mencatat riwayat pengguna seperti musik apa yang dicari dan musik apa yang didengarkan. Setelah mengakses musik pada riwayat pengguna sistem pada platform tersebut akan mengklasifikasikan musik berdasarkan metadata musik sistem. Sebuah sistem rekomendasi berbasis kecerdasan buatan dapat dijabarkan berdasarkan sumber pengetahuannya. Beberapa sumber tersebut meliputi social, individual, dan content. Social berupa sumber secara umum, individual berupa sumber yang lebih spesifik ke tiap perorangan, dan content merupakan sumber yang berdasarkan musik itu sendiri mulai dari fitur yang sederhana hingga yang cukup rumit [6]. Dalam mengklasifikasikan musik berdasarkan genre maupun aspek lainnya. Perlu dilakukan ekstraksi fitur dari musik tersebut. Fitur sama dengan karakteristik yang dibahas sebelumnya dan membuat setiap musik/suara memiliki perbedaan masing - masing. Salah satu ekstraksi fitur pada musik adalah fitur frequency domain. Ekstraksi fitur dari sebuah musik penting dilakukan untuk mengubah nilai musik tersebut kedalam bentuk numerik karena komputer hanya dapat memproses input bentuk numerik. Selain itu ekstraksi fitur juga berfungsi sebagai ruang fitur dalam metode klasifikasi yang akan menjadi pembeda setiap kelas.

1.1 SVM (Support Vector Machine)

Salah satu metode klasifikasi yang dapat digunakan adalah Support Vector Machine. Support Vector Machine merupakan sebuah metode klasifikasi yang merupakan penjabaran dari PLA (Perceptron Learning Algorithm). Jika pada PLA kita dapat mengklasifikasikan data dalam bentuk linear regresi, maka SVM akan memaksimalkan margin antara kelas data. Margin pada SVM merujuk pada jarak antara hyperplane pembagi dengan setiap sampel data latih pada tiap kelas data dengan jarak terdekat ke hyperplane [10]. Margin dapat dibedakan menjadi positif dan negatif berdasarkan letak kelas pada hyperplane yang dapat dirumuskan sebagai berikut [10].

$$w\theta + wt x_{\text{positive}} = 1 \tag{1}$$

$$w\theta + wt x_{\text{negative}} = -1 \tag{2}$$



Gambar 2. Hyperplane pada 2 Kelas dan Hyperplane Optimal yang Diperoleh [10]

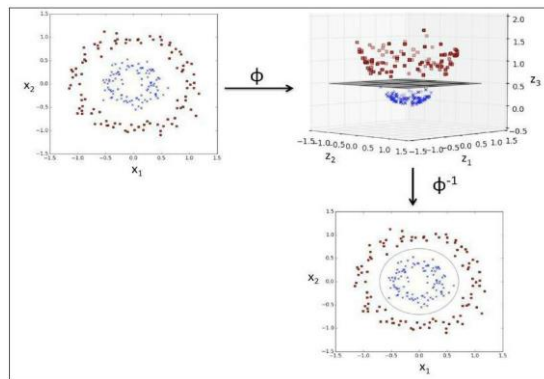
Support Vector Machine dibagi menjadi beberapa bentuk berupa jenis kernel yang digunakan. Kernel tersebut antara lain linear, RBF (Radial Basis Function), Sigmoid, dan Polinomial.

Penggunaan kernel pada SVM berfungsi dalam klasifikasi data dengan bentuk non-linier. Data non-linier yang dimaksud berupa kelas data yang memiliki banyak dimensi nilai sehingga sulit direpresentasikan secara linier. Penggunaan kernel dalam menangani kelas non-linear yaitu membuat kombinasi non-linear dari fitur asal untuk diproyeksikan ke dalam dimensi yang lebih tinggi melalui mapping (ϕ) [10]. Sebagai contoh kelas data berdimensi 2 akan dibentuk menjadi dimensi 3 yang nantinya dapat dipisahkan secara linear dengan rumus berikut [10].

$$\phi(x_1, x_2) = (z_1, z_2, z_3) = (x_1, x_2, x_1^2 + x_2^2) \quad (3)$$

Dimana:

- a. x_1, x_2 = fitur awal pada dimensi 2
- b. z_1, z_2, z_3 = fitur hasil konversi ke bentuk dimensi 3

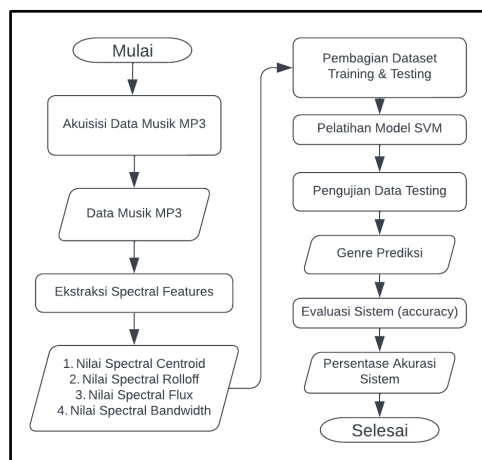


Gambar 3. SVM dengan Fitur Kelas 2 dan 3 Dimensi [10]

Setiap file musik tentu memiliki nilai fitur yang berbeda - beda. Dengan dilakukannya klasifikasi sebuah musik dapat dikategorikan ke dalam label tertentu dan dapat dijadikan sebagai acuan dalam mengklasifikasikan musik lainnya. Dengan dilakukannya penelitian ini diharapkan mampu mengklasifikasikan musik berdasarkan genre sekaligus menguji metode SVM dalam mengklasifikasikan genre musik sesuai data yang diberikan.

2. Metode Penelitian

Terdapat beberapa langkah dalam sistem mulai dari persiapan data hingga proses klasifikasi dan evaluasi.



Gambar 4. Flowchart Sistem

2.1 Akuisisi Data

Data yang digunakan berupa data sekunder yang diperoleh dari website www3.projects.science.uu.nl yang merujuk pada penelitian “**Studying emotion induced by music through a crowdsourcing game**” [3]. Data musik berformat MP3 yang merupakan data musik terkompresi. Data disimpan pada folder setiap genre berbeda yang dijadikan satu directory berupa folder “source”.

Tabel 1. Jumlah dan Pembagian Data Musik

Genre	Jumlah
Classical	100
Electronics	100
Pop	100
Rock	100

Audio selanjutnya akan melalui Python modules librosa berupa “librosa.load()”. Beberapa variabel seperti y dan sr akan muncul dimana y merupakan bentuk denotasi sinyal audio (data musik) berbentuk array numpy dan sr merupakan sampling rate yang melambangkan frekuensi dari y [8]. Penggunaan modul ini berfungsi untuk memperoleh nilai amplitudo (y) dan sampling rate (sr) dari setiap file musik yang sudah diakuisisi.

2.2 Ekstraksi Fitur Musik

Ekstraksi fitur yang dilakukan pada data musik berfokus pada fitur frekuensi (spectral features). Proses ekstraksi menggunakan bantuan modules Python yaitu librosa dan numpy. Beberapa fitur yang diekstrak antara lain spectral centroid, spectral rolloff, spectral flux, dan spectral bandwidth. Karena setiap data musik memiliki durasi yang berbeda maka nilai dari spectral features dapat memiliki jumlah yang berbeda saat proses ekstraksi. Sehingga digunakan fungsi “mean. ()” untuk mencari rata - rata dari setiap nilai spectral features tersebut.

Tabel 2. Contoh Hasil Ekstraksi Spectral Features

File	Label	Centroid	Rolloff	Flux	Bandwidth
1.mp3	Classical	1308.785	2067.708	0.953	1428.417
2.mp3	Classical	667.434	865.712	0.861	906.712
3.mp3	Electronic	2385.439	4882.837	1.194	2528.466
4.mp3	Electronic	1242.056	2518.873	1.408	1622.783
5.mp3	Pop	2516.364	5698.397	1.342	2766.53
6.mp3	Pop	2590.617	5800.017	1.273	2737.064
7.mp3	Rock	1566.794	3348.488	1.507	2086.703
8.mp3	Rock	1875.557	4144.43	1.16	2153.481

2.3 Pembagian Dataset Training dan Data Testing

Data musik pada folder setiap genre kemudian akan dipilih secara acak untuk dijadikan data uji dan data latih. Pembagian data dibantu oleh Python modules berupa “sklearn.model_selection” dengan ukuran data tes sebesar 0,2. Hal ini berarti sebanyak 20% dari data akan dialokasikan sebagai data uji dan 80% sisanya sebagai data latih. Data uji dipilih secara acak dari setiap genre [9].

2.4 Pelatihan Model SVM

Pelatihan model SVM menggunakan Python modules “sklearn.svm” untuk memanggil modul “SVC”. Secara default jenis SVM yang digunakan adalah kernel RBF. Pada dasarnya rumus perhitungan kernel RBF berupa [12].

$$K(x, x') = \exp(-\gamma \|x - x'\|^2) \quad (4)$$

Nilai γ secara default dapat didefinisikan:

$$\gamma = \frac{1}{n \text{ features} * \sigma^2} \quad (5)$$

Dimana:

- $\|x - x'\|^2$ menyatakan pangkat dari jarak Euclidean data antar fitur.
- γ menyatakan skalar pengaruh setiap data kelas dalam mencapai decision boundary model.

Setelah pemanggilan model SVM perlu digunakan fungsi “fit. ()” untuk memasukkan data latih yang sudah dibagi sebelumnya ke dalam model SVM. terdapat sebanyak 320 sampe musik sebagai data latih yang terpilih secara acak.

2.5 Pengujian

Pengujian sistem dilakukan dengan menguji data uji dengan model SVM yang sudah dilatih. Sebanyak 80 sampel musik acak dijadikan data uji saat pembagian data latih dengan data uji. Pengujian data testing menggunakan fungsi “predictions ()” pada SVC. Selain pengujian menggunakan data uji juga dilakukan evaluasi berupa akurasi dari sistem dalam menguji data uji. Hasil dari perhitungan akurasi nantinya akan dikonversikan kedalam bentuk persen.

$$\text{accuracy} = \frac{\sum \text{TP}}{\sum \text{Prediction}} \times 100\% \quad (6)$$

Dimana:

- TP (True Positive) merupakan berapa banyak prediksi genre yang sesuai dengan genre asalnya.
- Prediction merupakan jumlah total pengujian.

3. Hasil dan Diskusi

Hasil Pengujian berupa hasil data uji dan evaluasi sistem berupa perhitungan akurasi.

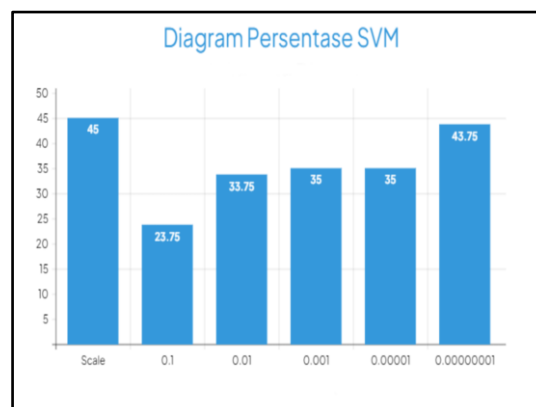
Tabel 3. Perbandingan Hasil Prediksi Data Uji

File (MP3)	Label	Prediksi	File (MP3)	Label	Prediksi
1.mp3	pop	classical	45.mp3	rock	pop
10.mp3	pop	classical	46.mp3	classical	classical
100.mp3	classical	pop	47.mp3	classical	rock
11.mp3	pop	rock	48.mp3	classical	pop
12.mp3	classical	classical	49.mp3	classical	pop
13.mp3	classical	rock	5.mp3	rock	rock
14.mp3	rock	rock	50.mp3	electronic	classical
15.mp3	classical	classical	51.mp3	electronic	electronic
16.mp3	pop	rock	52.mp3	classical	pop
17.mp3	electronic	pop	53.mp3	electronic	pop
18.mp3	classical	classical	54.mp3	classical	pop
19.mp3	rock	rock	55.mp3	rock	rock
2.mp3	classical	classical	56.mp3	rock	pop
20.mp3	classical	pop	57.mp3	classical	pop
21.mp3	electronic	classical	58.mp3	rock	rock
22.mp3	classical	classical	59.mp3	classical	rock
23.mp3	pop	pop	6.mp3	pop	pop
24.mp3	rock	rock	60.mp3	electronic	classical
25.mp3	pop	rock	61.mp3	classical	classical
26.mp3	rock	rock	62.mp3	electronic	pop
27.mp3	classical	classical	63.mp3	electronic	rock
28.mp3	classical	classical	64.mp3	pop	rock
29.mp3	rock	rock	65.mp3	pop	pop

3.mp3	pop	pop	66.mp3	classical	classical
30.mp3	classical	classical	67.mp3	electronic	classical
31.mp3	rock	rock	68.mp3	classical	pop
32.mp3	electronic	rock	69.mp3	rock	rock
33.mp3	pop	rock	7.mp3	pop	classical
34.mp3	pop	pop	70.mp3	rock	electronic
35.mp3	electronic	rock	71.mp3	classical	classical
36.mp3	rock	rock	72.mp3	rock	rock
37.mp3	electronic	rock	73.mp3	electronic	classical
38.mp3	pop	pop	74.mp3	classical	classical
39.mp3	classical	classical	75.mp3	classical	classical
4.mp3	pop	rock	76.mp3	rock	rock
40.mp3	electronic	rock	77.mp3	pop	rock
41.mp3	electronic	pop	78.mp3	pop	classical
42.mp3	rock	electronic	79.mp3	rock	rock
43.mp3	electronic	classical	8.mp3	electronic	pop
44.mp3	electronic	classical	80.mp3	pop	classical

3.1. Hasil Akurasi

Perhitungan akurasi dapat dilakukan dengan menghitung seberapa banyak data uji yang sesuai dengan label sebenarnya. Karena model SVM menggunakan kernel RBF (kernel default) maka pengujian juga dilakukan dengan mengganti nilai gamma. Dengan catatan nilai gamma $\neq 0$.



Gambar 5. Persentase Akurasi Model SVM dengan Gamma Berbeda

Berdasarkan hasil pengujian akurasi pada model SVM dengan perhitungan akurasi dan nilai gamma yang divariasikan dalam pengujian diperoleh bahwa nilai γ default memiliki akurasi terbaik sebesar 45%. Selain itu dapat dilihat bahwa semakin kecil nilai γ maka semakin bertambah pula akurasi dari pengujian sistem.

4. Kesimpulan

Berdasarkan hasil pengujian diperoleh bahwa akurasi model SVM dalam mengklasifikasikan genre musik berada pada 45% akurat. Ada beberapa faktor yang mempengaruhi performa dari SVM dalam mengklasifikasikan data seperti nilai parameter gamma, jenis kernel, jenis ekstraksi fitur, hingga jumlah dan variasi data khususnya musik dimana setiap data musik memiliki durasi dan fitur yang berbeda meskipun terletak pada kategori/label yang sama.

Daftar Pustaka

- [1] Jaya, E., Santosa, 2016. Klasifikasi Genre Musik Menggunakan Metode Support Vector Machine. *Jurnal Ilmiah Fasilkom* 81–82.
- [2] Lu, G., 2001. Indexing and retrieval of audio: A survey. *Multimedia Tools and Applications* 15, 269–290.
- [3] Aljanaki, A., Wiering, F., Veltkamp, R.C., 2016. Studying emotion induced by music through a crowdsourcing game. *Information Processing and Management* 52, 115–128.
- [4] As Sarofi, M.A., Irhamah, I., Mukarromah, A., 2020. Identifikasi Genre Musik dengan Menggunakan Metode Random Forest. *Jurnal Sains dan Seni ITS* 9.
- [5] Putri, L.A.A.R., 2017. Seleksi Fitur Dalam Klasifikasi Genre Musik. *Jurnal Ilmiah Ilmu Komputer Udayana* 10, 19–26.
- [6] Burke, R., Felfernig, A., Göker, M.H., 2011. Recommended systems: An overview. *AI Magazine* 32, 13–18. doi:10.1609/aimag.v32i3.2361
- [7] https://musicinformationretrieval.com/spectral_features.html
- [8] McFee, B., Raffel, C., Liang, D., Ellis, D., McVicar, M., Battenberg, E., & Nieto, O. (2015). librosa: Audio and Music Signal Analysis in Python. In *Proceedings of the 14th Python in Science Conference* (pp. 18–24). SciPy.
- [9] Scikit-learn Contributors. (2023). `sklearn.model_selection.train_test_split`. Tersedia: <https://scikit-learn.org>
- [10] S. Raschka and V. Mirjalili, *Machine Learning and Deep Learning with Python, scikit-learn, and Tensorflow 2*, 3rd ed. Mumbai: Packt, 2019.
- [11] R. Zakwandi, *Bahan Ajar Gelombang Bunyi*. Bandung, 2017.
- [12] J. Doe, "SVM Classifier and RBF Kernel: How to Make Better Models in Python," *Towards Data Science*. [Online]. Tersedia: <https://towardsdatascience.com>