

Klasifikasi *Mood* Musik Menggunakan *Support Vector Machine* Berdasarkan Fitur Spektral

I Gusti Agung Ngurah Diputra Wiraguna^{a1}, Luh Arida Ayu Rahning Putri^{a2}, I Ketut Gede Suhartana^{a3},
Gst. Ayu Vida Mastrika Giri^{a4}

^aProgram Studi Informatika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas
Udayana, Jimbaran, Kuta Selatan, Badung, Bali, Indonesia

¹gungwah1515@gmail.com

²rahningputri@unud.ac.id

³ikg.suhartana@unud.ac.id

⁴vida@unud.ac.id

Abstrak

Klasifikasi musik merupakan salah satu topik MIR (Music Information Retrieval) yang berkembang pesat. Berbagai aspek musik telah menjadi topik penelitian seputar klasifikasi musik mulai dari *genre*, instrumen, hingga *mood*. Salah satu teori taksonomi *mood* musik yang cukup terkenal adalah *circumplex model* yang memetakan *mood* musik berdasarkan dimensi *arousal* dan *valence*. *Arousal* untuk tingkat intensitas dari *mood* tersebut. Sedangkan *valence* untuk kondisi emosional positif atau negatifnya suatu *mood*. Penelitian mengklasifikasikan 4 kelas kuadran *mood* musik berdasarkan *circumplex model* menggunakan model pembelajaran mesin SVM (*Support Vector Machine*). Penelitian menggunakan jenis fitur spektral yang merupakan distribusi fitur sinyal musik pada domain frekuensi untuk mengklasifikasikan 4 kelas *mood* kuadran musik. Klasifikasi dilakukan melalui beberapa pendekatan berbeda seperti mengklasifikasikan 4 kelas secara langsung serta kombinasi model *arousal* dan *valence* terpisah. Hal ini dilakukan untuk menilai performa fitur spektral dalam klasifikasi *mood* musik berdasarkan 4 kelas kuadran dan melihat apakah ada kecenderungan fitur spektral lebih cocok dalam klasifikasi *arousal* atau *valence*. Penelitian juga mengujicobakan beberapa tahap lain seperti *parameter tuning* SVM dan mengaplikasikan seleksi fitur hingga reduksi dimensi fitur. Hasil penelitian diperoleh model SVM dengan akurasi terbaik untuk pendekatan klasifikasi 4 kelas yang memperoleh akurasi sebesar 66,67%. Hasil terbaik diperoleh setelah mengaplikasikan seleksi fitur dan reduksi dimensi fitur. Selain itu hasil menunjukkan bahwa fitur spektral cenderung mampu mengklasifikasikan *arousal* lebih baik dibandingkan *valence* dalam konteks *circumplex model* untuk *mood* musik.

Kata kunci : *Mood Musik, Circumplex Model, Fitur Spektral, Klasifikasi, Support Vector Machine*

1. Pendahuluan

Beberapa artikel menyimpulkan musik memiliki pengaruh terhadap *mood* pendengarnya [1], [2]. Pengaruh musik terhadap *mood* pendengar dapat terjadi pada siapa saja tanpa memandang usia atau jenis kelamin [1]. Beberapa aspek musik yang dinilai mampu mempengaruhi *mood* dapat berupa topik musik tersebut, lirik, hingga bagaimana musik tersebut dimainkan. Berdasarkan aspek - aspek tersebut musik dapat dikategorikan berdasarkan kategori tertentu seperti jenis instrumen, *genre*, hingga *mood*. *Mood* sebagai salah satu bentuk pengkategorian musik cenderung bersifat subjektif karena terkadang sebuah musik menggambarkan kondisi atau keadaan sesuai dengan keinginan musisinya. Secara psikologis, penelitian terdahulu telah mengemukakan skema pengkategorian *mood* dalam kelas tertentu. Salah satu yang cukup terkenal adalah *circumplex model* yang mengemukakan *mood* berdasarkan dimensi *arousal* dan *valence* [3], yang kemudian dijadikan landasan dalam penelitian seputar *mood* termasuk musik.

Dalam penelitian seputar musik banyak dijumpai jenis fitur musik salah satunya yaitu fitur spektral. Fitur spektral merupakan persebaran fitur sinyal suara dalam representasi domain frekuensi. Selain domain frekuensi terdapat domain lain seperti domain waktu dan *cepstral*. Namun fitur spektral dipilih karena merupakan fitur dasar sinyal suara dan representasi domain frekuensi dapat diekstraksi banyak jenis fitur spektral. Fitur spektral mencerminkan nilai karakteristik frekuensi sebuah sinyal yang

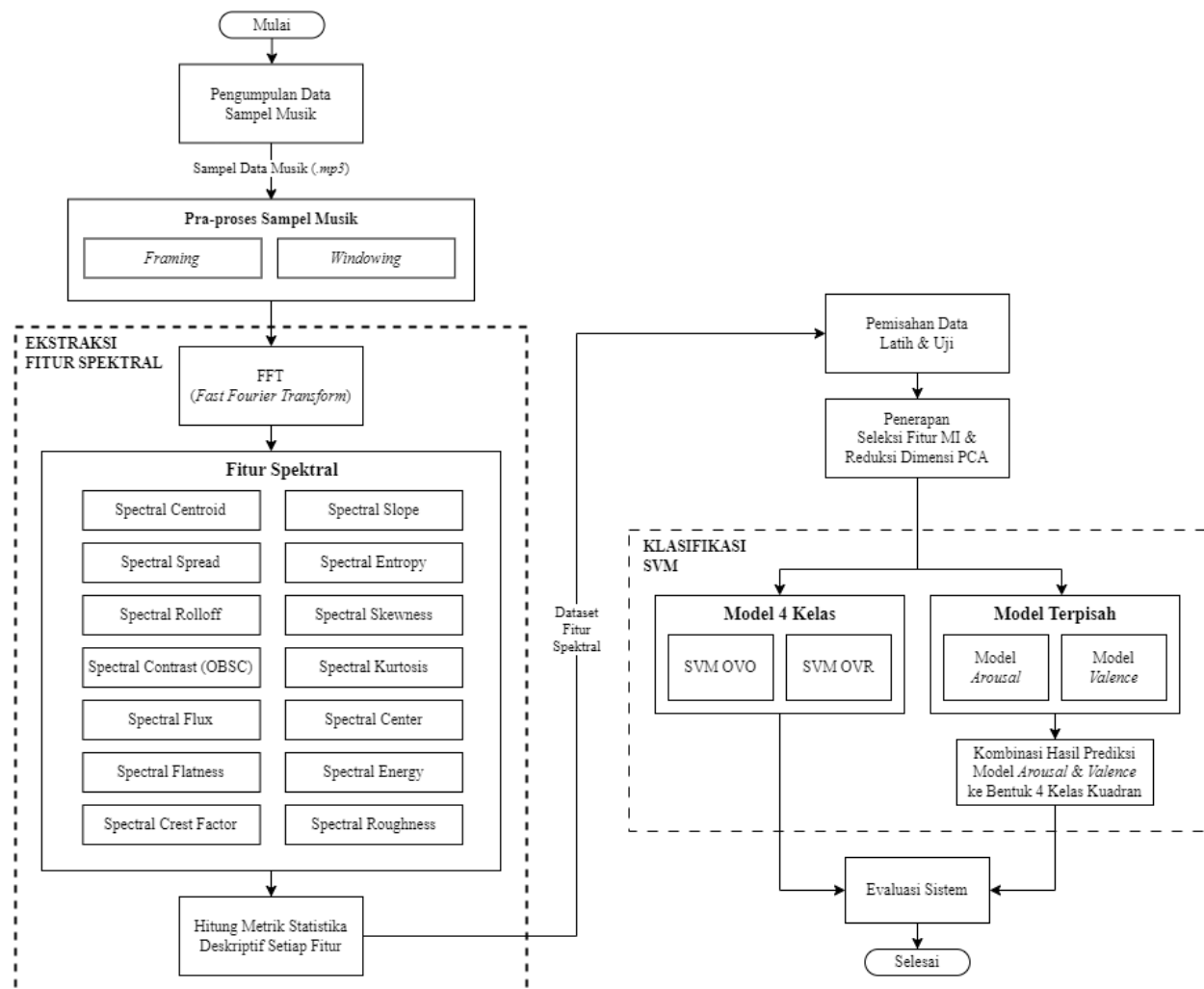
merepresentasikan harmoni, nada, hingga derau secara rasional [4]. Hal ini membuat fitur spektral lebih unggul dibandingkan fitur domain waktu dan pada beberapa kasus, fitur spektral juga mengungguli fitur domain *cepstral* [5], [6].

Menggunakan dasar teori pengkategorian *mood* circumplex model dan fitur spektral dirancang sebuah model klasifikasi *mood* musik menggunakan SVM. Beberapa penelitian sebelumnya telah mengaplikasikan SVM pada ranah musik baik dalam klasifikasi *mood*, hingga aspek lain seperti *genre* dan memperoleh performa yang baik [7], [8]. Salah satu alasan utama dipilihnya SVM adalah kemampuan klasifikasi *multiclass* meskipun pada dasarnya SVM adalah model klasifikasi biner. Hal ini sesuai digunakan pada kasus klasifikasi *mood* 4 kuadran yang tersusun kembali oleh aspek *arousal* dan *valence*.

Penelitian ini bertujuan untuk menguji performa SVM dalam melakukan Klasifikasi musik dengan kelas *mood* berupa 4 kuadran *circumplex model* dengan beberapa pendekatan berbeda. Pengujian dilakukan melalui konfigurasi kelas dataset pada klasifikasi. Hasil penelitian diharapkan mampu menjadi acuan terkait pengkategorian *mood* dan performa fitur spektral pada ranah musik, serta menjadi motivasi untuk penelitian berikutnya baik seputar musik maupun pembelajaran mesin.

2. Metode Penelitian

Klasifikasi *mood* musik menggunakan SVM berdasarkan fitur spektral terdiri dari beberapa tahapan utama meliputi pengumpulan data sampel musik, ekstraksi fitur spektral, klasifikasi dengan model SVM, serta pengujian dan evaluasi model. Alur penelitian dapat dilihat pada gambar 1.



Gambar 1. Alur Penelitian

Secara umum penelitian dimulai dari mengumpulkan data sampel musik 4 kuadran *circumplex model*. Data sampel musik akan melalui tahap pra-proses melalui *framing* dan *windowing* sebelum diaplikasikan FFT (*Fast Fourier Transform*) sebagai bentuk representasi sinyal pada domain frekuensi untuk kemudian diekstraksi nilai ciri atau karakteristik sinyal berupa fitur spektral. Berbagai jenis fitur spektral diekstraksi dan dijadikan sebagai dataset fitur dalam melatih dan menguji model pembelajaran mesin SVM melalui beberapa pendekatan model Klasifikasi SVM berbeda.

2.1. Pengumpulan Data

Dalam menunjang perancangan model klasifikasi *mood* musik berdasarkan 4 kuadran *circumplex model*, diperlukan dataset berupa sampel musik yang relevan. Dataset tergolong jenis sekunder karena berasal dari sumber *open-source* yaitu [Kaggle.com](https://www.kaggle.com) berupa kumpulan sampel musik berformat “.mp3” yang terbagi pada 4 kelas berbeda sesuai kuadran *circumplex model*. Selain sampel musik, dataset juga terdiri dari beberapa file “.csv” yang berisikan metadata seperti anotasi kelas, judul lagu, dan musisinya [9].

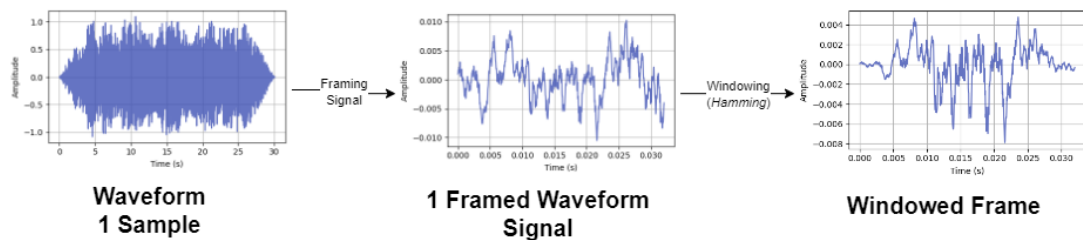
Hasil pengumpulan data memperoleh dataset sampel musik dengan jumlah seimbang untuk masing-masing 4 kelas kuadran dalam bentuk *folder*. Satu kelas berisi 225 sampel “.mp3” sehingga total sampel sebanyak 900 sampel musik. sampel musik berdurasi 30 detik dan bervariasi, ada yang berisi nyanyian lirik dan ada yang tidak. Setiap sampel musik akan melalui tahap pemrosesan sinyal dan ekstraksi fitur sebelum digunakan dalam melatih dan menguji model klasifikasi SVM.

2.2. Ekstraksi Fitur

Berdasarkan gambar 1, tahapan ekstraksi fitur dimulai dari pra-proses data sampel musik, penerapan FFT (*Fast Fourier Transform*), lalu hasil FFT digunakan untuk ekstraksi fitur spektral. Hasil fitur spektral yang masih berbentuk banyak *frame* kecil untuk sebuah sampel musik akan diaplikasikan metrik statistika deskriptif untuk menghitung nilai fitur spektral untuk keseluruhan 1 sampel musik.

a. Pra-proses Sinyal

Tahap pra-proses berfungsi mempersiapkan sampel sinyal sebelum ekstraksi fitur spektral. Pra-proses umumnya dilakukan untuk menyeragamkan data baik dari segi bentuk, tipe, hingga format disimpannya data. Tahap pra-proses yang dilakukan adalah *framing* dan *windowing*.



Gambar 2. Pra-proses Sampel Sinyal Musik

Tahap *framing* yang mengubah sinyal menjadi beberapa bagian kecil berfungsi untuk mencari karakteristik fitur sinyal musik secara detail. Mengingat sinyal terutama dalam kasus ini berupa musik merupakan gelombang non-stasioner sehingga frekuensi suara musik bisa berbeda pada durasi tertentu. Untuk tahap *windowing* menggunakan jenis “*Hamming*” yang terlampir dalam persamaan (1).

$$w(n) = 0,54 - 0,46 \cos = 2\pi \left(\frac{n}{N} \right) \quad (1) [10]$$

Fungsi *windowing* membuat transisi antar *frame* sinyal tidak terpotong secara langsung. *Windowing* mampu mengantisipasi *spectral leakage* yang terjadi jika sinyal hanya dibagi melalui *framing* saja. Setiap *frame* yang telah diaplikasikan fungsi *windowing* akan ditumpuk sebagian (*overlapping frames*) untuk mengurangi kehilangan informasi dan menjaga kontinuitas antar *frame* sesuai dengan sampel musik aslinya [10].

b. Ekstraksi Fitur Spektral

Setiap *frame* sinyal akan diaplikasikan FFT untuk merubahnya ke representasi domain frekuensi yang mengandung spektrum dan bin frekuensi penyusun sinyal. Spektrum dan bin frekuensi digunakan dalam ekstraksi berbagai jenis fitur spektral. Adapun fitur spektral yang digunakan terlampir pada tabel 1. Fitur spektral diekstraksi untuk setiap *frame* yang menggambarkan fitur representasi sinyal pada domain frekuensi [11].

Tabel 1. Fitur Spektral [11], [12], [13]

Fitur	Keterangan
<i>Spectral Centroid</i>	Pusat massa spektrum penyusun sinyal.
<i>Spectral Roll-off</i>	Frekuensi dengan spektrum dibawah persentasi tertentu.
<i>Spectral Spread</i>	Persebaran frekuensi di sekitar <i>spectral centroid</i> .
<i>Spectral Contrast</i>	Perbedaan frekuensi puncak dan lembah pada <i>range</i> frekuensi tertentu.
<i>Spectral Flux</i>	Perubahan spektrum antar <i>frame</i> .
<i>Spectral Flatness</i>	Tingkat kedataran frekuensi penyusun sinyal.
<i>Spectral Crest Factor</i>	Rasio spektrum puncak dengan rata – rata-nya.
<i>Spectral Slope</i>	Perbandingan spektrum setiap bertambahnya frekuensi.
<i>Spectral Entropy</i>	Ketidakteraturan distribusi spektrum sebuah sinyal.
<i>Spectral Skewness</i>	Tingkat simetri/asimetri distribusi spektrum.
<i>Spectral Kurtosis</i>	Ketajaman puncak distribusi spektrum sinyal.
<i>Spectral Energy</i>	Total energi spektrum dalam sebuah sinyal.
<i>Spectral Roughness</i>	Fluktuasi/cepatnya perubahan antar spektrum.

Untuk 1 sampel musik akan diekstraksi fitur spektral sebanyak jumlah pada tabel 1. Fitur spektral diekstraksi secara manual menggunakan deskripsi dan rumus sesuai artikel tertentu [11], [12], [13]. Setiap ekstraksi 1 jenis fitur akan menghasilkan vektor fitur karena tahap pra-proses yang membuat sebuah sampel terbagi kedalam banyak *frame*. Untuk menangani hal ini diterapkan metrik statistika deskriptif meliputi *mean*, *median*, *variance*, *minimum*, *maximum*, *range*, *standard deviation*, *interquartile*, dan *variance*. Metrik statistika diterapkan pada tiap jenis fitur spektral untuk setiap sampel dengan hasil berupa dataset fitur spektral. Dataset fitur kemudian akan disertakan dengan kolom nama file sampel serta kelas sampel untuk dipisahkan menjadi data latih dan uji sebagai bahan perancangan model klasifikasi SVM.

2.3. Klasifikasi SVM

Klasifikasi SVM dimulai dengan mengakses dataset fitur dan memisahkannya menjadi 80% data latih dan 20% data uji. Model SVM akan mencari *hyperplane* optimal yang mampu mengklasifikasikan data antar kelas. SVM sendiri pada dasarnya merupakan klasifikasi biner sehingga penerapan dalam kelas 4 kuadran *circumplex model* perlu dilakukan penyesuaian. Adapun persamaan dasar *hyperplane* sebagai berikut.

$$w \cdot x + b = 0 \quad (2) \text{ [14]}$$

Dari rumus dasar *hyperplane* pada persamaan (2), perhitungan *hyperplane* optimal untuk SVM dimodifikasi dalam bentuk fungsi objektif SVM yang penyelesaiannya diperoleh menggunakan *lagrange multiplier* (α). Bentuk persamaan fungsi objektif ini sering juga disebut fungsi *wolfe dual* [14]. Adapun rumus dari fungsi *wolfe dual* SVM sebagai berikut.

$$\begin{aligned}
 \text{maximize}_w W(a) &= \sum_{i=1}^m a_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m a_i a_j y_i y_j x_i \cdot x_j \\
 \text{subject to, } C &\geq a_i \geq 0, \text{ for } i = 1, \dots, m \\
 \sum_{i=1}^m a_i y_i &= 0
 \end{aligned} \quad (3) \text{ [14]}$$

Menerapkan persamaan (3) akan menghasilkan nilai α yang dapat digunakan untuk menghitung bobot dan bias model yang kemudian digunakan dalam prediksi data. Dalam membentuk model SVM, data latih digunakan untuk menentukan nilai bobot dan bias yang dinilai menghasilkan *hyperplane* optimal sesuai persamaan (2) dan (3). Kemudian nilai bobot dan bias digunakan untuk memprediksi data uji dan dibandingkan apakah hasil prediksi sesuai dengan kelas asal.

Dalam melatih model dikenal istilah parameter tuning dengan beberapa jenis parameter seperti nilai *cost* (C), *gamma* (γ), dan *kernel trick*. Kernel khususnya membuat SVM mampu melakukan klasifikasi data dengan dimensi tinggi [14]. Beberapa variasi nilai parameter diujicobakan secara *grid-search* dalam SVM untuk mencari kombinasi parameter dengan performa terbaik. Adapun beberapa jenis kernel yang digunakan sebagai berikut.

$$\text{kernel linear, } k(x_i, x_j) = (x_i \cdot x_j) \quad (4) [14]$$

$$\text{kernel polynomial, } k(x_i, x_j) = (x_i \cdot x_j + c)^d \quad (5) [14]$$

$$\text{kernel RBF, } k(x^{(i)}, x^{(j)}) = \exp(\gamma \|x^{(i)} - x^{(j)}\|^2) \quad (6) [14]$$

Untuk menangani klasifikasi kelas 4 kuadran, diterapkan pendekatan berupa konfigurasi jenis SVM berupa OVO, OVR, serta klasifikasi model *arousal* dan *valence* terpisah. Klasifikasi OVO dan OVR menyusun skema klasifikasi antar kelas kuadran dengan merubahnya menjadi klasifikasi biner. Skema OVO akan menyusun klasifikasi antar 1 kelas dengan 1 kelas lainnya hingga terbentuk seluruh pasangan yang mungkin dari seluruh 4 kelas. Sedangkan skema OVR langsung mengklasifikasikan 1 kelas dengan 3 kelas lainnya. Untuk pendekatan model *arousal* dan *valence* terpisah akan dibentuk 2 model klasifikasi dimana pada model *arousal* dibedakan menjadi kelas *arousal* tinggi (Q1, Q2) dan rendah (Q3, Q4). Sedangkan model *valence* dibedakan menjadi kelas *valence* positif (Q1, Q4) atau negatif (Q2, Q3). Pengujian model *arousal* dan *valence* terpisah menggunakan data uji 4 kelas kuadran melalui *mapping* hasil prediksi. Data uji dianggap diklasifikasikan benar hanya jika diprediksikan sesuai kelas asal pada kedua model *arousal* dan model *valence*.

2.4. Evaluasi Sistem

Pengujian performa sistem dilihat dari hasil prediksi data uji apakah sesuai kelas asalnya atau tidak. Hasil prediksi data uji dapat digambarkan dalam bentuk *confusion matrix* sebagai berikut.

Tabel 2. Confusion Matrix

Confusion Matrix		Kelas Asal	
		Positive	Negative
Kelas Prediksi	Positive	TP	FP
	Negative	FN	TN

Berdasarkan *confusion matrix* pada tabel 2 dapat disusun persamaan (7) untuk menghitung akurasi sistem. Pengukuran akurasi menghitung seberapa banyak data uji yang diklasifikasikan sesuai kelas asalnya berbanding dengan keseluruhan jumlah prediksi data uji. Salah satu alasan pemilihan metrik akurasi adalah seimbangnya jumlah sampel data musik tiap kelas.

$$\text{akurasi} = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (7) [8]$$

Adapun keterangan dari *confusion matrix* pada tabel 2 serta perhitungan akurasi pada persamaan (7) dimana TP (*True Positive*) merupakan data yang diprediksikan positif dan memang kelas asalnya positif. FP (*False Positive*) merupakan data yang diprediksikan positif namun kelas sebenarnya adalah negatif. FN (*False Negative*) merupakan banyak data yang diprediksikan negatif namun kelas sebenarnya adalah positif. TN (*True Negative*) merupakan data yang diprediksi negatif dan memang kelas asalnya negatif.

Pengukuran akurasi diterapkan pada seluruh pendekatan model klasifikasi SVM. Hasil akurasi selain menunjukkan performa sistem juga menunjukkan bagaimana pengaruh fitur spektral dalam mengklasifikasikan musik berdasarkan kategori 4 kuadran menggunakan SVM. Skenario pengujian meliputi pendekatan model SVM OVO, OVR, hingga kombinasi model *arousal* dan *valence* yang dilakukan dalam bentuk *grid-search* untuk menentukan kombinasi parameter yang menghasilkan akurasi terbaik. Parameter SVM yang diujicobakan terlampir dalam tabel 3.

Tabel 3. Parameter Model SVM

Kernel	C (Cost)	Gamma (γ)
RBF	1, 10, 100	0.001, 0.01, 0.1, 0.5, 1
Linear	1, 10, 100	0.001, 0.01, 0.1, 0.5, 1
Polynomial	1, 10, 100	0.001, 0.01, 0.1, 0.5, 1

Pengujian dilanjutkan dengan mengujicobakan seleksi fitur MI dan reduksi dimensi fitur PCA. Seleksi fitur MI akan menghitung nilai MI setiap fitur dalam dataset terhadap kelas data yang ada. Nilai fitur MI memiliki rentang 0 hingga 1, Dimana nilai MI tinggi mengindikasikan fitur tersebut relevan terhadap kelas data [15]. Sedangkan nilai MI kecil atau bahkan 0 berarti fitur tersebut independen dan tidak menggambarkan kelas data. Berbeda dengan seleksi fitur MI, reduksi dimensi fitur PCA akan membentuk dataset fitur baru dengan dimensi yang lebih ringkas namun mempertahankan relevansi informasi fitur data baru dengan fitur data pada dimensi asalnya [16]. Sehingga PCA cocok digunakan pada dataset berdimensi tinggi yang menyebabkan beban komputasi yang besar.

Penerapan seleksi fitur MI dan reduksi dimensi PCA bertujuan mengidentifikasi fitur spektral yang dinilai relevan dalam Klasifikasi *mood* 4 kuadran. Dengan mereduksi dimensi dataset fitur membuat pelatihan model lebih mudah dan meringankan beban komputasional. Selain itu menerapkan seleksi fitur MI dan reduksi dimensi fitur PCA diharapkan mampu meningkatkan performa model SVM yaitu peningkatan akurasi model.

3. Hasil dan Pembahasan

Hasil implementasi sistem menunjukkan beberapa perolehan akurasi model SVM untuk setiap pendekatan berbeda. Perolehan akurasi terbaik setiap pendekatan ditampilkan dengan *confusion matrix* untuk melihat karakteristik hasil klasifikasi. Pengujian dilakukan melalui *grid-search* parameter SVM yang mungkin terbentuk pada tabel 2, Pengujian model dilakukan menggunakan dataset fitur yang lengkap hingga yang telah diterapkan seleksi fitur MI dan reduksi dimensi fitur PCA.

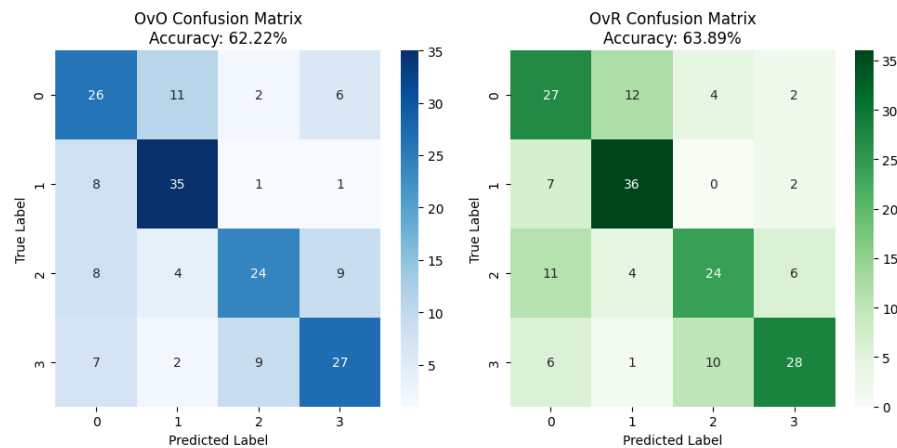
3.1. Pendekatan Model Klasifikasi

Terdiri dari pendekatan klasifikasi 4 kelas (*multi-class*) secara langsung dan klasifikasi model terpisah. Perolehan akurasi setiap pendekatan model SVM sebagai berikut.

a. Klasifikasi 4 Kelas

Klasifikasi dilakukan dengan menerapkan konfigurasi OVO dan OVR untuk menangani klasifikasi *multi-class*. Pendekatan klasifikasi OVO dan OVR diterapkan karena kasus klasifikasi yaitu 4 *mood* kuadran *circumplex model* sedangkan model SVM pada dasarnya adalah model klasifikasi biner.

Model klasifikasi OVO akan menyusun skema 4 kelas menjadi sebanyak pasangan antar 2 kelas yang mungkin terbentuk. Prediksi akan dilakukan secara *voting* untuk memilih hasil prediksi yang lebih banyak terpilih. Berbeda dengan model OVO, model klasifikasi OVR akan menyusun skema klasifikasi biner antar 2 kelas dimana kelas 1 melawan sisa kelas lainnya secara langsung. Hasil pengujian model OVO dan OVR ditunjukkan dalam bentuk *confusion matrix* berikut.

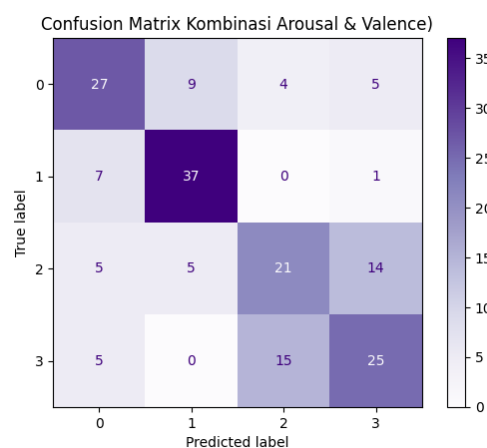


Gambar 3. *Confusion Matrix* Model OVO dan OVR

Gambar 3 memperlihatkan *confusion matrix* model SVM OVO dan OVR yang memperoleh akurasi terbaik. Model OVO memperoleh akurasi terbaik sebesar 62,22% dengan kombinasi parameter berupa nilai C senilai 10, *gamma* 0,01, dan jenis kernel RBF. Sedangkan model OVR memperoleh akurasi terbaik sebesar 63,89% dengan parameter C sebesar 1, *gamma* 0,01, dan jenis kernel RBF. Meskipun OVR memperoleh akurasi lebih baik, *confusion matrix* OVR menunjukkan banyak misklasifikasi menjadi Q1. Sedangkan model OVO persebaran hasil prediksi lebih sesuai dengan teori *circumplex model* yang menunjukkan model lebih mampu dalam klasifikasi *arousal*.

b. Klasifikasi Model *Arousal* dan *Valence* Terpisah

Klasifikasi model terpisah menyusun skema model *arousal* dan ,model *valence* dengan menyesuaikan kelas pada dataset fitur. Kedua model diuji dengan data uji *mood* 4 kelas kuadran.



Gambar 4. *Confusion Matrix* Kombinasi Model *Arousal* dan *Valence*

Gambar 4 menunjukkan *confusion matrix* kombinasi model *arousal* dan *valence* dengan perolehan akurasi terbaik sebesar 61,11%. Akurasi tersebut diperoleh melalui konfigurasi model *arousal* berparameter C senilai 10, *gamma* 0,001, dan jenis kernel RBF. Sedangkan model *valence* berkonfigurasi C senilai 100, *gamma* 0,01, dan jenis kernel RBF. Hasil kembali menunjukkan model SVM dengan fitur spektral lebih akurat dalam klasifikasi aspek *arousal*.

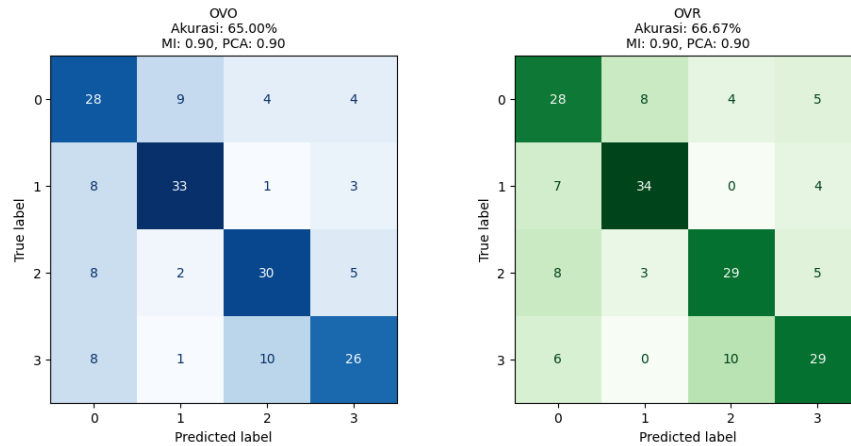
3.2. Penerapan Seleksi Fitur MI & Reduksi Dimensi Fitur PCA

Kombinasi seleksi fitur dan reduksi dimensi fitur diterapkan setelah memperoleh akurasi terbaik dari setiap pendekatan klasifikasi. Seleksi fitur yang digunakan adalah MI (*Mutual Information*) dan reduksi dimensi PCA. MI dan PCA diterapkan berpasangan untuk mengidentifikasi

kombinasi parameter MI dan PCA yang mampu meningkatkan akurasi model SVM setiap pendekatan sebelumnya. Parameter MI dan PCA yang digunakan senilai 70%, 80%, dan 90%.

a. Klasifikasi 4 Kelas

Penerapan seleksi fitur MI dan reduksi dimensi fitur PCA terbukti meningkatkan akurasi model. Karakteristik klasifikasi dan peningkatan akurasi ditunjukkan melalui *confusion matrix* model OVO dan OVR setelah seleksi fitur MI dan reduksi dimensi PCA.

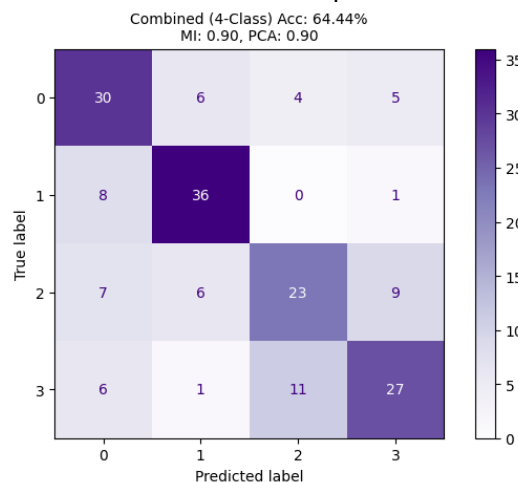


Gambar 5. *Confusion Matrix* Model OVO dan OVR (Setelah MI dan PCA)

Gambar 5 memperlihatkan peningkatan akurasi dari model OVO dan OVR setelah diterapkan seleksi fitur MI dan reduksi dimensi PCA. Akurasi terbaik masih diperoleh oleh model OVR yang meningkat menjadi 66,67%. Akurasi tersebut diperoleh oleh konfigurasi parameter C senilai 10, *gamma* 0,01, dan jenis kernel RBF. Model OVO memperoleh akurasi 65% dengan konfigurasi parameter yang sama. Terlihat bahwa penggunaan 90% fitur MI dengan jumlah *principal component* hasil PCA dengan karakteristik 90% sama – sama memperoleh akurasi terbaik di kedua model.

b. Klasifikasi Arousal dan Valence Terpisah

MI dan PCA juga diterapkan pada model *arousal* dan *valence* terpisah untuk melihat apakah terjadi peningkatan akurasi serta karakteristik hasil prediksi kedua model.



Gambar 6. *Confusion Matrix* Model Kombinasi (Setelah MI dan PCA)

Akurasi meningkat menjadi 64,44% yang diperlihatkan pada gambar 6. Akurasi terbaik diperoleh pada kombinasi model *arousal* berparameter C senilai 10, *gamma* 0,001, dan kernel RBF. Untuk model *valence* memiliki konfigurasi nilai C senilai 10, *gamma* 0,01, dan kernel RBF. Akurasi terbaik diperoleh dari penerapan seleksi fitur MI dan reduksi dimensi fitur PCA berparameter 90%. Secara keseluruhan hasil sebelum dan sesudah penerapan MI dan PCA memiliki karakteristik yang sama pada *confusion matrix* dimana model lebih mampu mengklasifikasikan aspek *arousal*.

Hasil penelitian menunjukkan adanya pengaruh parameter tuning hingga penerapan seleksi fitur MI dan reduksi dimensi fitur PCA pada akurasi model SVM. Pendekatan model SVM berbeda menunjukkan karakteristik hasil yang cukup berbeda meskipun akurasi model tidak berbeda jauh. Pada gambar 3 model OVR memiliki akurasi yang lebih baik namun *confusion matrix* model OVO memperlihatkan hasil klasifikasi yang lebih seimbang dan sesuai dengan taksonomi *circumplex model*. Melalui pendekatan kombinasi model *arousal* dan *valence* terpisah terbukti bahwa model SVM dengan fitur spektral lebih mampu mengklasifikasikan aspek *arousal* dibandingkan *valence*.

Penerapan seleksi fitur MI dan reduksi dimensi fitur PCA pada dataset fitur mampu meningkatkan akurasi setiap pendekatan model SVM. Berdasarkan gambar 5 dan 6, karakteristik hasil Klasifikasi model SVM setelah diterapkan seleksi fitur MI dan reduksi dimensi fitur PCA masih menunjukkan kecenderungan model yang lebih akurat dalam klasifikasi *arousal*. Hasil penerapan MI dan PCA juga menunjukkan bahwa model memiliki performa baik saat parameter MI dan PCA bernilai tinggi, seperti 90%. Melihat perolehan akurasi setiap pendekatan model SVM, ditampilkan beberapa fitur dengan perolehan nilai MI tertinggi.

Tabel 4. Fitur Spektral dengan Nilai MI Tertinggi

Fitur & Metrik Statistika	Nilai MI
<i>roughness_median</i>	0,365027
<i>flux_median</i>	0,340344
<i>contrast_band3_median</i>	0,336179
<i>roughness_iqr</i>	0,335787
<i>contrast_band3_mean</i>	0,331809

Terlihat dari tabel 4 fitur spektral dan metrik statistika yang diekstraksi seperti *spectral roughness*, *spectral flux*, dan *spectral contrast* pada *subband* frekuensi rendah memperoleh nilai MI tertinggi di kisaran 0,3. Seleksi fitur MI memperoleh nilai MI seluruh fitur yang kemudian dapat diurutkan dari tertinggi ke terendah. Sebagai tambahan, perolehan nilai MI terendah menyentuh nilai 0 pada beberapa fitur yang mayoritas terdiri dari *spectral contrast* pada *subband* frekuensi tinggi. Dimana nilai MI 0 mengartikan fitur tidak relevan sama sekali.

Peningkatan akurasi di seluruh pendekatan model SVM sesuai dengan tujuan penerapan seleksi fitur MI yaitu memilih fitur yang relevan terhadap kelas klasifikasi berdasarkan nilai MI tiap fitur. Sedangkan metode PCA mereduksi fitur yang kompleks dengan membentuk *principal component* dengan ukuran lebih kecil namun tetap mempertahankan informasi fitur sesuai persentase PCA yang digunakan

Berdasarkan hasil penelitian hampir keseluruhan pendekatan model klasifikasi menunjukkan karakteristik performa yang lebih akurat dalam klasifikasi *arousal* dibandingkan *valence*. Beberapa teori menjelaskan terkait hasil penelitian. Fitur spektral yang merupakan representasi domain frekuensi sebuah sinyal lebih mampu dalam klasifikasi *arousal* sesuai dengan karakteristik domain sinyal tersebut. Spektrum beserta bin frekuensi bervariasi antar *frame* sinyal lebih menunjukkan gairah atau intensitas sebuah sinyal [17]. Dalam ranah musik aspek *arousal* bisa dilihat dari cara sebuah musik dimainkan dan seberapa cepat/lambatnya. Untuk *valence* yang merupakan kondisi emosional positif atau negatif, terlihat fitur spektral kesulitan dalam melakukan klasifikasi. Hal ini menyesuaikan dengan karakteristik domain frekuensi yang terdiri dari spektrum dan bin frekuensi penyusun sinyal, secara garis besar melambangkan energi atau intensitas sebuah sinyal musik [18], [19]. Berdasarkan hasil penelitian dan penjelasan tersebut dapat dikatakan bahwa fitur spektral memang lebih cocok dalam mengenali aspek *arousal* dalam *mood* 4 kuadran *circumplex model* dibandingkan aspek *valence*.

4. Kesimpulan

Berdasarkan hasil penelitian dapat disimpulkan bahwa model SVM memperoleh akurasi terbaik sebesar 63,89% dan meningkat menjadi 66,67% setelah diterapkan seleksi fitur MI dan reduksi dimensi PCA. Hasil bentuk *confusion matrix* keseluruhan model klasifikasi menunjukkan karakteristik klasifikasi *arousal* lebih baik dibandingkan *valence*. Hasil tersebut mengimplikasikan fitur spektral lebih mampu mengklasifikasikan *arousal* atau tingkat gairah dan intensitas sebuah kategori *mood*. Berbeda dengan

valence yang merupakan kondisi emosional positif atau negatif yang cukup sulit digambarkan dalam sebuah sinyal musik, fitur spektral dapat mengidentifikasi dan mengenali aspek *arousal* lebih mudah.

Referensi

- [1] A. Mamdouh and M. Ahmed, "The Impact of Music on Mood: Descriptive Observations of Listening Experiences and Their Effect on Mental Health," 2024. [Online]. Available: <https://www.researchgate.net/publication/379145461>
- [2] A. Tewari, "The Impact of Music on Mood Regulation," *The International Journal of Indian Psychology*, vol. 12, no. 2, p. 5, Jun. 2024, doi: 10.25215/1202.390.
- [3] J. A. Russell, "A circumplex model of affect," *J Pers Soc Psychol*, vol. 39, no. 6, pp. 1161–1178, Dec. 1980, doi: 10.1037/h0077714.
- [4] C. Constantinescu and R. Brad, "An Overview on Sound Features in Time and Frequency Domain," *International Journal of Advanced Statistics and IT&C for Economics and Life Sciences*, vol. 13, no. 1, pp. 45–58, Dec. 2023, doi: 10.2478/ijasitels-2023-0006.
- [5] F. Alías, J. C. Socoró, and X. Sevilano, "A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds," 2016, *Balkan Society of Geometers*. doi: 10.3390/app6050143.
- [6] C. Prabha, S. Kaur, M. Malik, M. Uddin, and D. Nandan, "A Cutting-Edge Feature Extraction Approach for Speaker Recognition Leveraging Optimized Variance Spectral Flux and Daubechies Wavelet," *Traitement du Signal*, vol. 40, no. 6, pp. 2845–2852, Dec. 2023, doi: 10.18280/ts.400645.
- [7] C. A. Qurniaty and Kusnawi, "Ekspresi Emosi Berdasarkan Suara Menggunakan Algoritma Multi Layer Perceptron dan Support Vector Machine," *Indonesian Journal of Computer Science Attribution*, vol. 12, no. 6, pp. 2023–4014, 2023.
- [8] B. R. Ismanto, T. M. Kusuma, and D. Anggraini, "Indonesian Music Classification on Folk and Dangdut Genre Based on Rolloff Spectral Feature Using Support Vector Machine (SVM) Algorithm," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 1, p. 11, Jan. 2021, doi: 10.22146/ijccs.54646.
- [9] R. Panda, R. Malheiro, and R. P. Paiva, "Novel Audio Features for Music Emotion Recognition," *IEEE Trans Affect Comput*, vol. 11, no. 4, pp. 614–626, Oct. 2020, doi: 10.1109/TAFFC.2018.2820691.
- [10] O. Kamil, "Frame Blocking and Windowing Speech Signal," 2018, [Online]. Available: <https://www.researchgate.net/publication/331635757>
- [11] G. Sharma, K. Umapathy, and S. Krishnan, "Trends in audio signal feature extraction methods," *Applied Acoustics*, vol. 158, Jan. 2020, doi: 10.1016/j.apacoust.2019.107020.
- [12] J. Tao, D.-N. Jiang, L. Lu, H.-J. Zhang, J.-H. Tao, and L.-H. Cai, "Music type classification by spectral contrast feature MUSIC TYPE CLASSIFICATION BY SPECTRAL CONTRAST FEATURE 1," 2002. [Online]. Available: <https://www.researchgate.net/publication/313484983>
- [13] G. Peeters, "A Large Set Of Audio Features For Sound Description (Similarity and Classification) In The CUIDADO Project," 2003.
- [14] A. Kowalczyk, *Support Vector Machine Succinctly*. Syncfusion, 2017. [Online]. Available: www.syncfusion.com.
- [15] X. Cheng, "Signals and Systems Ins," 2024. [Online]. Available: <https://soapubs.com/index.php/ICSS>
- [16] M. Greenacre, P. J. F. Groenen, T. Hastie, A. I. D'Enza, A. Markos, and E. Tuzhilina, "Principal component analysis," *Nature Reviews Methods Primers*, vol. 2, no. 1, Dec. 2022, doi: 10.1038/s43586-022-00184-w.
- [17] Y. H. Yang, Y. C. Lin, Y. F. Su, and H. H. Chen, "A regression approach to music emotion recognition," *IEEE Trans Audio Speech Lang Process*, vol. 16, no. 2, pp. 448–457, Feb. 2008, doi: 10.1109/TASL.2007.911513.
- [18] U. Oberst, "The fast Fourier transform," *SIAM J Control Optim*, vol. 46, no. 2, pp. 496–540, 2007, doi: 10.1137/060658242.
- [19] D. T. Kusuma, "Fast Fourier Transform (FFT) Dalam Transformasi Sinyal Frekuensi Suara Sebagai Upaya Perolehan Average Energy (AE) Musik," *PETIR*, vol. 14, no. 1, pp. 28–35, Oct. 2020, doi: 10.33322/petir.v14i1.1022.