

# Penggolongan Musik Terhadap Suasana Hati Menggunakan Metode K-Means

I Gede Harsemadi<sup>1)</sup>, I Made Sudarma<sup>2)</sup>

Program Studi Sistem Informasi STMIK STIKOM Bali <sup>1)</sup>  
Program Studi Magister Teknik Elektro Universitas Udayana <sup>2)</sup>  
e-mail : harsemadi@stikom-bali.ac.id<sup>1)</sup>, debronxs@yahoo.co.id<sup>2)</sup>

## Abstrak

*Music Information Retrieval* merupakan bagian dari bidang sistem temu kembali informasi (*Information Retrieval*) yang merupakan bidang keilmuan yang berhubungan dengan pencarian informasi dalam dokumen, dan hubungan metadata berbagai dokumen dalam basis data yang saling terhubung. Data yang digunakan dalam penelitian ini adalah musik. Karena setiap musik yang tercipta memiliki energi emosi tersendiri yang terpancar, serta memiliki kaitan yang kuat dengan psikologi manusia, hal ini menandakan bahwa musik berkaitan dengan emosi dan *mood* (suasana hati) tertentu pada diri manusia. Dalam penelitian ini akan mengelompokkan data musik berdasarkan 4 kategori *mood* model Thayer, dengan metode *data mining* menggunakan algoritma *clustering K-Means*. Keunggulan algoritma ini melakukan proses pemodelan tanpa supervisi (*unsupervised learning*) dengan meminimalkan variasi antar data yang ada di dalam suatu *cluster* dan memaksimalkan variasi dengan data yang ada di *cluster* lainnya. Penelitian ini diharapkan dapat menghasilkan sebuah rekayasa sistem untuk mengelompokkan data musik menggunakan algoritma *K-Means*.

**Kata kunci:** musik, mood, pengelompokkan, clustering, k-means.

## 1. Pendahuluan

*Music Information Retrieval* merupakan bagian dari bidang dari sistem temu kembali informasi (*Information Retrieval*) yang merupakan bidang keilmuan yang berhubungan dengan pencarian informasi dalam dokumen, dan hubungan metadata berbagai dokumen dalam basis data yang saling terhubung. Musik hingga kini menjadi bidang penelitian dalam penerapan teknik temu kembali informasi yang menarik untuk diteliti. Ranah penelitian dalam bidang musik termasuk dalam klasifikasi, *clustering* data musik, serta analisa musik untuk representasi pengetahuan di dalamnya.

Musik memiliki kaitan yang kuat dengan psikologi manusia, hal ini menandakan bahwa musik berkaitan dengan emosi dan *mood* (suasana hati) tertentu pada diri manusia. Musik memiliki banyak kebaikan bagi tubuh dan jiwa manusia, musik dapat menenangkan bahkan membangkitkan semangat seseorang yang mendengarkan musik. Para ahli psikologi menyatakan bahwa terdapat bagian pada otak manusia yang merasakan musik dekat dengan bagian otak yang berhubungan dengan ekspresi emosional manusia [1]. Dalam penelitian yang dilakukan oleh Song, dkk [2] menyatakan bahwa setiap musik yang telah tercipta memiliki energi emosi tersendiri yang terpancar, maka dari itu mulai banyak penelitian yang telah dilakukan pada pengenalan emosi musik. Dari sudut pandang psikologi musik, elemen yang paling penting dalam penggolongan musik itu sendiri adalah pada emosi apa yang tersirat dari musik tersebut

Thayer melakukan penelitian dengan judul *The biopsychology of mood and arousal*, muncul pendekatan lain yang lebih sederhana terhadap kategorisasi pada suasana hati dalam musik. Thayer mengajukan model 2 dimensi yang memetakan suasana hati dalam musik, pendekatan 2 dimensi ini mengangkat teori yang menyatakan bahwa emosi dan suasana hati disebabkan oleh dua faktor, yaitu *stress* (senang dan cemas) dan *energy* (santai dan energetik). Selanjutnya dari model 2 dimensi ini dibagi menjadi 4 *cluster* yaitu *contentment/kepuasan*, *exuberance/gembira*, *depression/depresi*, dan *anxious/cemas*; kalut, [3].

Menentukan emosi atau suasana hati apa yang terkandung sebuah lagu merupakan tantangan dalam penelitian ini. Untuk mengetahui lebih efektif mengenai suasana hati yang muncul dalam sebuah lagu, dapat diketahui pada saat mendengarkan lagu tersebut pada bagian *intro* dan *refrain* lagu [4]. Bagian *intro* merupakan bagian awal yang dapat menentukan atmosfer lagu, *intro* juga secara intuitif memiliki kata-kata dan nada-nada yang intensif untuk memberikan informasi mengenai suasana hati. Bagian *refrain* merupakan bagian dengan kata-kata atau nada-nada yang paling sering diulang-ulang dalam sebuah lagu. Secara umum bagian *refrain* adalah bagian dengan pengulangan kata dan nada yang paling

---

sering dan hal ini yang menjadi bagian *refrain* yang paling menentukan *mood* apa yang terkandung dalam lagu tersebut.

Dalam hal menentukan jenis *mood* yang terkandung dalam lagu diperlukan metode khusus yang menggabungkan komputasi numerik dengan penambahan data berupa fitur-fitur unik dalam sebuah lagu, hal ini dikenal sebagai *Music Information Retrieval* (MIR). MIR adalah salah satu bagian dalam data mining dimana informasi yang akan digali dari sumber data berupa musik. Banyak penelitian yang telah dilakukan mengenai MIR khususnya pada klasifikasi emosi dan *mood* dalam bidang musik, psikologi, pemrosesan sinyal, *machine learning* maupun kombinasi dari beberapa model penelitian tersebut. MIR menggunakan berbagai macam metode *data mining* untuk pengelompokan termasuk di dalamnya klasifikasi dan *clustering* data seperti K-NN, C4.5, decision tree, Support Vector Model (SVM), Artificial Neural Network, Self Organization Map, K-Means dan lain sebagainya.

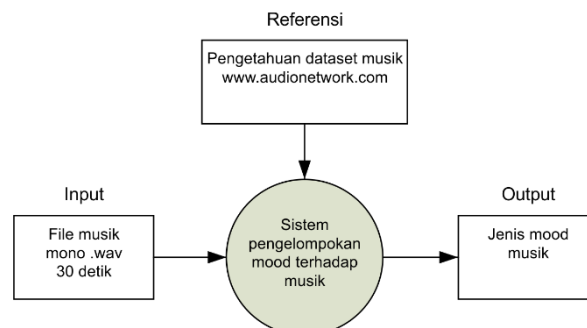
Dalam proses *data mining*, sebelum berbagai jenis lagu tersebut dikelompokkan berdasarkan *mood*, maka harus melewati tahap awal pengolahan data (*preprocessing*). Tahapan *preprocessing* ini dikenal dengan *feature selection* [1]. Tahapan selanjutnya dari *feature selection* adalah proses *feature extraction* atau ekstraksi ciri dari sebuah file lagu untuk menemukan ciri yang untuk masuk ke tahapan pengolahan data berikutnya, perlu dilakukan pengolahan sinyal dengan menggunakan transformasi fourier yang berfungsi mengubah fungsi atau sinyal dalam domain waktu ke domain frekuensi. Tahapan *feature extraction* terdiri dari tiga proses, yaitu pengambilan sampel lagu (bagian *refrain*), penerapan *fast fourier transform* dan *spectral analysis* untuk mendapatkan nilai *spectral feature* yang menjadi atribut dasar untuk dilakukannya klasifikasi *mood*.

Untuk proses *clustering* dalam penelitian ini menggunakan algoritma K-Means, dimana algoritma *data mining* ini melakukan proses pemodelan tanpa supervisi (*unsupervised learning*) dan merupakan salah satu algoritma yang melakukan pengelompokan data dengan sistem partisi. Metode *K-Means* berusaha mengelompokkan data yang ada ke dalam beberapa kelompok, dimana data dalam satu kelompok mempunyai karakteristik yang sama satu sama lainnya dan mempunyai karakteristik yang berbeda dengan data yang ada di dalam kelompok yang lain. Dengan kata lain, algoritma ini berusaha untuk meminimalkan variasi antar data yang ada di dalam suatu *cluster* dan memaksimalkan variasi dengan data yang ada di *cluster* lainnya.

Berdasarkan pemaparan penjelasan latar belakang tersebut maka penelitian ini dilakukan untuk dapat membangun sebuah rekayasa sistem yang dapat mengelompokkan musik ke dalam empat parameter *mood* atau suasana hati yang terdapat dalam musik. Sistem ini akan dibangun berdasarkan pengelompokkan parameter *mood* yang diperoleh dengan menggunakan algoritma K-Means. Sistem ini nantinya diharapkan dapat membantu penggunaannya untuk menemukan kelompok kategori *mood* apa yang terkandung dalam sebuah file musik, serta memperoleh nilai *silhouette coefficient* untuk mengukur tingkat kualitas dan kekuatan pada *cluster mood*. Perlu diperhatikan juga bahwa file musik yang digunakan ini hanya instrumental musik saja tanpa mengandung lirik vokal. Koleksi musik yang digunakan dalam proses pelatihan dan pengujian ini diperoleh dari situs [www.audionetwork.com](http://www.audionetwork.com).

## 2. Metode Penelitian

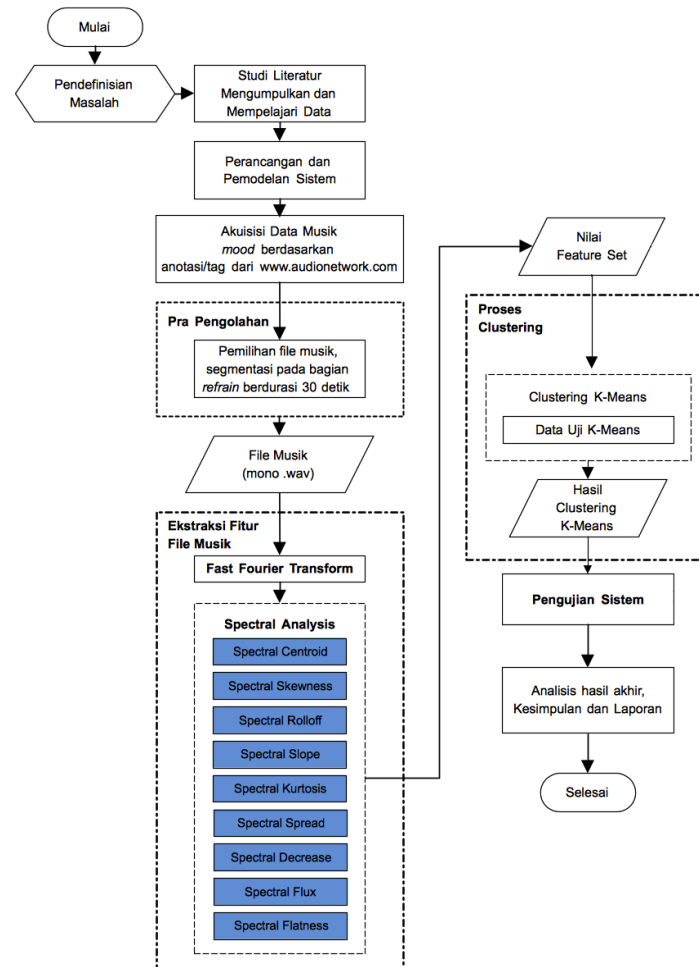
Sistem pengelompokan musik terhadap suasana hati dengan menggunakan algoritma K-Means ini memiliki 3 komponen utama yaitu komponen *input*, sistem pengelompokan/klasifikasi musik terhadap *mood*, dan komponen *output* yaitu label jenis *mood* musik yang dapat dilihat pada Gambar 1.



Gambar 1. Blok diagram sistem pengelompokan suasana hati terhadap musik

Gambar 1 dapat dijelaskan dimana sistem akan menerima masukan (*input data*) berupa file musik format mono .wav berdurasi rata-rata 30 detik, yang selanjutnya sistem melakukan proses pengelompokan musik terhadap *mood* dengan menggunakan K-Means *clustering*. Kemudian sistem menghasilkan output berupa label jenis *mood* model Thayer yaitu 1) *contentment*/ketenangan; relaksasi,

2) *exuberance*/bersemangat; riuh; gembira, 3) *depression*/depresi; sedih, dan 4) *anxious*/cemas; kalut; kacau. Selanjutnya alur sistem yang lebih detail dari pembuatan rekayasa sistem ini memiliki beberapa tahapan yang dapat dilihat pada Gambar 2.



Gambar 2. Alur Analisis Penelitian

Pendefinisian masalah melingkupi penjabaran latar belakang dan masalah yang akan ditangani oleh sistem. Studi literatur/kepuustakaan merupakan tahapan awal yang penting dalam memulai penelitian, karena tahap ini berhubungan dengan pengumpulan teori dan merangkum penelitian-penelitian sebelumnya yang akan menjadi dasar penelitian ini. Tahapan akuisisi data pada penelitian ini menggunakan *dataset* suasana hati dari situs [www.audionetwork.com](http://www.audionetwork.com). Audio Network adalah sebuah perusahaan musik yang menggunakan suasana hati untuk pengelompokan musik, perusahaan ini menggunakan 56 anotasi/tag sosial untuk pelabelan suasana hati, dan label ini telah ditentukan sebelumnya oleh para spesialis bidang musik.

Tahapan pra-pengolahan file musik diawali dengan sebuah file musik yang diedit menggunakan aplikasi Adobe Audition yang mengubah tampilan file menjadi spektrum audio, selanjutnya mengambil potongan klip musik yang dilakukan secara manual dengan mendengarkan sebuah file musik secara berulang-ulang kemudian mendapatkan posisi *refrain* dari musik tersebut [4]. Hasil klip file musik tersebut disimpan dengan format *.wav* dengan *mono audio channel*.

Tahapan ekstraksi fitur file musik yang telah diinput (*mono \*.wav*) akan diproses untuk mendapatkan ciri-ciri khusus yang disebut dengan ekstraksi fitur. Proses ekstraksi dimulai dengan mengubah sinyal *file* musik yang diinput menjadi domain frekuensi menggunakan metode *Fast Fourier Transform* (FFT) [5]. Selanjutnya menerapkan 9 jenis *spectral analysis* [6] untuk menghasilkan seperangkat nilai atau *feature set* yang terdiri dari 9 atribut data untuk setiap file lagu. Dan nilai-nilai tersebut yang selanjutnya akan dimasukkan ke proses analisis data, yaitu sebagai nilai masukkan pada pembentukan *dataset* data uji untuk tahapan *clustering* K-Means selanjutnya.

Tahapan *clustering* K-Means ini diawali dengan penentuan jumlah *cluster* (*k*) untuk

mengelompokkan data uji sebagai data input yang akan dikelompokkan pada sejumlah  $k$  cluster. Sebagai contoh apabila terdapat 400 baris data dalam *dataset* dan ada 4 *cluster mood* yang perlu dibentuk, maka  $k$  adalah 4 *cluster*. Algoritma ini selanjutnya mengambil 4 *record* data secara acak dari *dataset* sebagai pembentuk *cluster* awal. Setiap *cluster* yang terbentuk akan dihitung rata-ratanya (*means*), dimana rata-rata dari suatu *cluster* adalah rata-rata dari semua *record* yang terdapat pada *cluster* tersebut. Selanjutnya data dialokasikan ke *cluster* terdekat, dalam hal ini menggunakan perhitungan *euclidean distance*. Selanjutnya menghitung kembali nilai *centroid* dari rata-rata setiap *cluster* yang ada, dan alokasikan kembali data ke *cluster* terdekat. Hal ini dilakukan berulang kali sampai terbentuk *cluster* yang stabil dan prosedur algoritma *K-Means* selesai. *Cluster* yang stabil terbentuk ketika iterasi atau pengulangan perhitungan *K-Means* tidak membuat *cluster* baru dan tidak ada data yang berpindah [7].

Seperti yang diilustrasikan pada Gambar 2.5, sistem akan menghasilkan 4 *cluster* data dimana masing-masing *cluster* tersebut belum memiliki label. Penentuan label *mood* ini dilakukan secara manual dengan melihat kemiripan karakteristik nilai *feature set* yang ada pada *dataset* hasil *clustering*. Selanjutnya berdasarkan kemiripan ini barulah dapat ditentukan *mood* apakah tergolong ke kelas *mood 1. Contentment/ ketenangan;relaksasi, 2. Exuberance/ bersemangat; riuh; gembira, 3. Depression/ depresi; sedih, dan 4. Anxious/ cemas; kalut; kacau*. Selain itu pula sistem menghasilkan lama waktu yang diperlukan oleh sistem dalam proses klasifikasi data uji.

### 3. Hasil dan Pembahasan

Antarmuka sistem dari proses ekstraksi fitur *mood* musik telah dikembangkan. Secara umum dalam sebuah file audio yang digunakan dalam pembentukan *dataset* ini memiliki 10 nilai atribut (9 dari *spectral analysis* dan 1 sebagai kode kategori *mood* musik). Berikut pada Tabel 1 adalah *dataset* fitur audio untuk masing-masing kategori *mood* musik, namun data yang ditampilkan hanya 20 baris data untuk keempat jenis kategori *mood*.

Tabel 1 Kategori *feature set mood* musik

No.	Nama File	Nilai Spectral Analysis									
		Centroid	Skewness	Roll Off	Slope	Kurtosis	Spread	Decrease	Flux	Flatness	Ktgr
1	C-01	4073.2	29.808	10374	-1.27E-07	953.41	2988	0.14463	0.00014336	0.11385	1
2	C-02	1538.4	28.692	3400.6	-6.08E-07	915.05	954.66	0.096774	0.00012711	0.047206	1
3	C-03	2602.9	29.415	11493	-1.14E-07	923.1	2783.2	0.091045	0.00012063	0.13927	1
4	C-04	1184.6	22.86	10073	-3.28E-07	620.58	1868.4	0.081259	0.00011634	0.088511	1
5	C-05	10745	30.017	18639	-1.79E-11	978.98	6357.2	0.11864	0.00014202	0.84807	1
6	E-01	3098.3	16.702	9416.2	-1.77E-06	414.92	3059.9	0.057931	0.00012605	0.078833	2
7	E-02	4485.5	27.575	13118	-5.02E-07	866.8	5183.4	0.070906	0.00015421	0.11747	2
8	E-03	1408.6	25.932	10256	-8.40E-07	754.85	2396	0.070485	0.00013687	0.20239	2
9	E-04	3153.7	17.838	9222.5	-8.48E-07	402.74	2718.2	0.059337	0.00014705	0.12391	2
10	E-05	6898.7	28.616	14388	-2.17E-07	917	4891.8	0.080382	0.0002008	0.13711	2
11	D-01	1316.2	27.958	8060.2	-5.25E-08	878.66	1626.2	0.11	0.00014245	0.073075	3
12	D-02	942	21.903	4186.2	-9.60E-07	597.5	870.07	0.16725	0.00012038	0.041157	3
13	D-03	858.98	29.39	2098.5	-7.00E-08	932.66	572.71	0.047736	0.00015712	0.04961	3
14	D-04	3249.5	26.621	11848	-4.63E-07	793.65	3949.4	0.12695	0.00017632	0.01428	3
15	D-05	1367.1	28.07	13236	-6.37E-07	856.3	3325.9	0.080783	0.00018148	0.010534	3
16	A-01	1904.8	26.857	6758.1	-6.43E-07	815.93	1703.3	0.063516	0.0001079	0.082729	4
17	A-02	3205.4	18.779	11375	-1.38E-06	407.25	4118.9	0.064084	0.00012322	0.09827	4
18	A-03	2119.9	24.457	13592	-1.08E-06	715.16	3273	0.097417	0.00012796	0.24331	4
19	A-04	10969	28.344	18628	-4.39E-12	870.9	6361.5	0.083638	0.00013916	0.86102	4
20	A-05	1606.7	27.646	11256	-5.17E-08	828.9	2887.8	0.10072	0.00013088	0.10209	4

#### 3.1 Antarmuka Sistem Pengelompokan Mood Musik Menggunakan K-Means

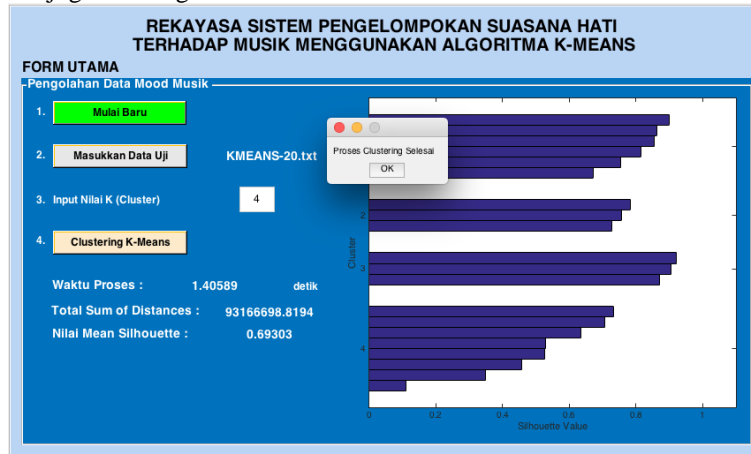
Antarmuka tahapan *clustering* atau pengelompokkan musik terhadap *mood* menggunakan algoritma *K-Means* ini dilakukan dengan mengelompokkan data uji yang telah dibuat menjadi beberapa file *dataset* uji. Data *input* untuk proses *clustering* ini dibentuk dari sekumpulan file berjumlah 40-400 *file* musik yang telah memiliki nilai *spectral feature* yang diperoleh pada tahapan pra-pengolahan *mood* musik. Data uji ini diambil dari musik yang ada pada situs [www.audionetwork.com](http://www.audionetwork.com), terbagi menjadi 10 *file dataset* uji, dimana masing-masing file data uji ini berformat \*.txt serta memiliki 40, 80, 120, 160, 200, 240, 280, 320, 360 dan 400 baris data yang berisikan nilai atribut berupa 9 nilai *spectral feature*.

Antarmuka *clustering K-Means* ini terdiri dari 2 panel utama, yang pertama untuk menginputkan data uji serta melakukan proses *clustering* data. Panel kedua adalah panel untuk memvisualisasikan hasil *clustering* data uji dalam bentuk histogram atau diagram batang. Terdapat tombol “Mulai Baru” yang

digunakan untuk memulai baru proses *clustering*. Kemudian dilanjutkan dengan memasukkan data uji yang dalam format data \*.txt, sistem akan membuka *window* untuk memasukkan *file* data uji.

Setelah data uji dimasukkan ke dalam sistem, selanjutnya akan ditentukan berapa jumlah *k cluster* yang akan terbentuk dari data uji ke dalam kelompok-kelompok. Dapat dilihat pada Gambar 3, dalam tahapan ini nilai *k* yang digunakan adalah 4, yang mewakili jumlah kategori *mood* musik.

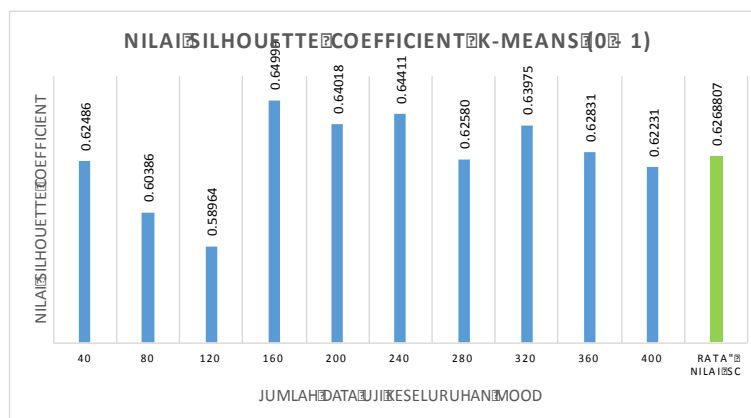
Tahapan selanjutnya adalah memulai proses *clustering* yaitu dengan menekan tombol “Clustering K-Means”, proses ini dimulai ketika data uji dan nilai *k* sudah diinputkan dengan benar. Sistem akan secara otomatis melakukan klasifikasi menggunakan fungsi *K-Means* dan menghasilkan pengelompokan data yang divisualisasikan dalam bentuk histogram atau diagram batang, waktu yang diperlukan untuk proses klasifikasi, nilai total jarak *sum of distances* untuk keseluruhan *cluster* dan nilai *silhouette coefficient* juga dihitung.



Gambar 3. Antarmuka hasil *clustering* data uji

### 3.2 Hasil Pengujian Menggunakan K-Means

Untuk mengetahui tingkat kualitas dan kekuatan *cluster* dapat diketahui dengan pengujian *silhouette coefficient*. Pengujian ini juga dimaksudkan untuk mengetahui seberapa baik suatu objek ditempatkan dalam suatu *cluster*. Kisaran nilai *silhouette coefficient* ini berada pada -1, 0 dan +1. Hasil *clustering* dikatakan baik apabila nilai *silhouette coefficient* bernilai positif dan mendekati +1, berarti objek tersebut telah berada pada *cluster* yang tepat. Sedangkan apabila nilai *silhouette coefficient* = 0, berarti objek tersebut berada diantara dua *cluster*, sehingga objek tersebut tidak jelas harus dimasukkan ke dalam *cluster* A atau B. Namun, jika nilai *silhouette coefficient* = -1 artinya terdapat struktur *cluster* yang *overlapping*, sehingga objek tersebut lebih tepat dimasukkan ke dalam *cluster* lainnya [8]. Nilai rata-rata *silhouette coefficient* dari tiap objek dalam suatu *cluster* adalah suatu ukuran yang menunjukkan seberapa ketat/fit data dikelompokkan dalam *cluster* tersebut.



Gambar 4. Nilai *silhouette coefficient* untuk setiap jumlah data uji

Berikut ini akan disajikan nilai rata-rata *silhouette coefficient* dari data uji oleh sistem menggunakan algoritma K-Means berdasarkan jumlah data uji dimana jumlah *cluster* yang dibentuk berdasarkan nilai *k=4*. Gambar 4 menunjukkan rata-rata nilai *silhouette coefficient* diperoleh berdasarkan

---

data uji yang digunakan berurut-urut adalah 40 hingga 400 data uji. Dapat diketahui bahwa nilai *silhouette coefficient* terendah berada saat pemrosesan *clustering* pada data uji 120 file yaitu 0,58964, dan tertinggi pada data uji 160 file yaitu 0,64999. Rata-rata nilai *silhouette coefficient* untuk *clustering* keseluruhan data uji yaitu 0,6268807. Hal ini dapat diartikan bahwa secara keseluruhan proses *clustering* menghasilkan *medium structure* (karena nilai SC berada pada,  $0.5 < SC \leq 0.7$ ).

#### 4. Simpulan

Telah dibangun sistem pengelompokan musik terhadap suasana hati dengan menerapkan metode data mining menggunakan algoritma K-Means, dimana pada *cluster* yang ditentukan berdasarkan penggolongan 4 jenis suasana hati model Thayer (*contentment/* menenangkan;bahagia, *exuberance/*bersemangat;riuh, *depression/* sedih;depresi, dan *anxious/*panik;kalut;cemas). Hasilnya adalah sistem ini mampu melakukan *clustering* dengan nilai rata-rata *silhouette coefficient* yang dihasilkan dari 40 – 400 data uji menunjukkan nilai 0,6268807 yang tergolong dalam *medium structure* atau tingkat kekompakan data pada suatu *cluster* yang tergolong baik.

Untuk pengembangan penelitian selanjutnya dapat dilakukan dengan penerapan model algoritma *clustering* lainnya seperti K-Medoid serta membandingkan performa kecepatan *clustering* dan hasil *silhouette coefficient*.

#### Daftar Pustaka

- [1] Samira Pouyanfar, Hossein Sameti. (2014). *Music Emotion Recognition Using Two Level Classification*. International Conference on Intelligent System (ICIS).
- [2] Song, Y.et al. (2012). Evaluation of Musical Features for Emotion Classification. *Proceedings of the 13th International Society for Music Information Retrieval Conference*. Porto, Portugal.
- [3] Thayer. (1989). *The biopsychology of mood and arousal*. Oxford University Press.
- [4] Seungwon, Oh., Minsoo Hahn, Jinsul Kim. (2013). *Music Mood Classification Using Intro and Refrain Parts of Lyrics* (ICISA).
- [5] Reonaldo Y. S. (2014). *Simulasi Sistem Pengacak Sinyal dengan Metode FFT (Fast Fourier Transform)*. E-Journal Teknok Elektro dan Komputer ISSN 2301-8402.
- [6] Lerch, Alexander. 2012. *An introduction to Audio Content Analysis – Applications in Signal Processing and Music Informatics*. IEEE Press.
- [7] Tahta, Alfina., Budi Santosa, Ali Ridho Barakbah. 2012. *Analisa Perbandingan Metode Heirarchical Clustering, K-Means dan Gabungan Keduanya dalam Cluster Data*. Jurnal Teknik ITS, Vol. 1, ISSN: 2301-9271
- [8] Anggara, M., Herry Sujiani, Helfi Nasution. 2016. *Pemilihan Distance Measure pada K-Means Clustering Untuk Pengelompokkan Member di Alvaro Fitness*. Jurnal Sistem dan Teknologi Informasi (JUSTIN) Vol. 1, No.1.