

KLASIFIKASI ALIRAN MUSIK BERDASAR CIRI SINYAL AUDIO DENGAN METODE ARTIFICIAL NEURAL NETWORK DAN RANDOM FOREST

Muhamad Aldiansyah

Program Studi Fisika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Jakarta
Jln. Rawamangun Muka, Jakarta Timur 13220

Email: muhamad.aldiansyah@ymail.com

Abstrak

Pengklasifikasian aliran musik secara komputasi diperlukan untuk keperluan skala industri dan merupakan bidang kajian *Music Information Retrieval* yang masih berkembang. Dalam mengklasifikasi musik secara komputasi diperlukan ciri-ciri yang paling bisa membedakan aliran musik satu dengan yang lain. Atas dasar itu, penelitian ini dilakukan untuk melihat kekuatan dari pengklasifikasian aliran musik berdasar ciri spektral dan temporal pada musik. Berdasar ciri tersebut, tiga aliran musik: metal, pop, dan rock telah diklasifikasi dengan metode *Artificial Neural Network* menghasilkan tingkat akurasi 80% dan metode *Random Forest* menghasilkan tingkat akurasi 70%. Hasil ini menunjukkan bahwa ciri spektral dan temporal merupakan ciri yang terbukti dapat mengklasifikasi aliran musik, dan metode klasifikasi yang berbeda berpengaruh pada hasil tingkat akurasi klasifikasi.

Kata-kata kunci: klasifikasi, musik, ciri audio, jaringan syaraf tiruan, random forest

1. Pendahuluan

Pengklasifikasian aliran musik secara komputasi merupakan salah satu bidang kajian *music information retrieval* (MIR). Manusia pada umumnya dapat menentukan aliran musik hanya dengan mendengar beberapa detik iringan musik pada sebuah lagu. Meski begitu, tak jarang subjektivitas manusia dalam menentukan aliran musik yang samar antara dua aliran musik memungkinkan rasa kurang yakin muncul. Sebuah standar pengklasifikasian aliran musik diperlukan untuk kebutuhan pribadi atau dalam skala industri.

Dari penelitian [1] dapat diketahui bahwa manusia pun dapat membuat komputer mengklasifikasi aliran musik melalui ciri-ciri yang terdapat pada sinyal audio. Metode yang digunakan dalam penelitian [1] adalah *Gaussian Mixture Model* dan *k-Nearest Neighbor*. Ciri yang dipakai dalam penelitian [1] adalah ciri spektral, ciri temporal, dan ciri ritme pada musik.

Dalam penelitian ini, pengklasifikasian aliran musik dilakukan dengan menggunakan *database* sampel musik GTZAN. *Database* merupakan database yang digunakan dalam penelitian [1]. *Database* ini berisi 10 aliran musik berbeda. Setiap aliran musik berisi 100 set lagu dengan durasi masing-masing 30 detik. Dari 10 aliran musik pada *database* tersebut, dalam penelitian ini hanya digunakan 3 aliran musik, yaitu: *rock*, *pop*, dan *metal*. Ciri yang digunakan hanya 3 ciri spektral, yaitu *spectral centroid*, *spectral roll-off*, dan *spectral flux*; dan 2 ciri temporal, yaitu: *zero-crossing rate* dan *enegy entropy*. Metode klasifikasi yang digunakan adalah *Artificial Neural Network* (ANN) dan *Random Forest*. Dari pegklasifikasian ini dapat dilihat apakah ciri tersebut cukup berpengaruh

dalam mengklasifikasi musik dan melihat metode klasifikasi yang lebih baik di antara ANN dan Random Forest untuk mengklasifikasi aliran musik dalam penelitian ini.

2. Metode Penelitian

Musik adalah rangkaian suara dari berbagai instrumen pengolah nada dan ritme, sehingga dapat ditinjau bahwa musik mengandung berbagai sinyal audio yang tergabung menjadi satu rangkaian sinyal audio. Sinyal audio musik dapat dibaca oleh komputer digital secara diskrit. Sinyal audio diskrit ini mengandung informasi dan dapat diambil melalui proses ekstraksi.

Di dalam *database* sampel musik GTZAN ada 100 lagu dari setiap aliran musik. Dari 3 aliran musik, yaitu: *rock*, *pop*, dan *metal*, 300 lagu digunakan dalam penelitian ini. Sebanyak 9 *mid-time feature* dari setiap lagu dihitung. Ciri yang diekstrak dalam penelitian ini adalah 9 *mid-time feature* dari sinyal audio musik yang dihitung secara statistik, di antaranya adalah: *mean-spectral centroid*, *mean-spectral roll-off*, *mean-spectral flux*, *mean-zero-crossing rate*, *std-spectral centroid*, *std-spectral roll-off*, *std-spectral flux*, *std-zero-crossing rate*, dan *min-energy entropy*. *Mean* adalah nilai rata-rata, *std* adalah nilai standar deviasi, dan *min* adalah nilai terendah.

Penghitungan secara statistik tersebut didapat dengan memotong durasi setiap lagu dari 30 detik - 1 potong menjadi 2 detik - 15 potong. Setiap potong inilah yang disebut sebagai *mid-time feature*. Kemudian dari setiap 2 detik ini dapat dipotong lagi dengan durasi 0.02 detik untuk menghitung *short-time feature*-nya. *Short-time feature* adalah 5 ciri berikut:

spectral centroid, spectral roll-off, spectral flux, zero-crossing rate, dan energy entropy. Sehingga alur pengekstrakan ciri adalah menghitung *short-time feature*, menghitung *mid-time feature*, kemudian dari 2 detik -15 potong *mid-time feature* ini dihitung rata-ratanya supaya untuk satu lagu memiliki 9 ciri. Dari sini didapat matriks 300 lagu x 9 ciri.

Berdasar pada [2, 3], masing-masing ciri dapat dihitung dengan persamaan berikut.

1. Spectral Centroid

$$C = \frac{\sum_{n=1}^N f(n) x(n)}{\sum_{n=1}^N x(n)} \quad (1)$$

Keterangan:

C = *spectral centroid*

f(n) = sampel frekuensi ke-n

x(n) = koefisien transformasi fourier diskrit ke-n

N = jumlah koefisien transformasi fourier diskrit

2. Spectral Roll-off

$$\sum_{n=1}^m x(n) = 0.85 \sum_{n=1}^N x(n) \quad (2)$$

Keterangan:

x(n) = koefisien transformasi fourier diskrit ke-n (sebelah kanan)

N = jumlah koefisien transformasi fourier diskrit

m = jumlah penghitungan batas roll-off

3. Spectral Flux

$$F = \sum_{n=1}^N (EN_i - EN_{i-1}) \quad (2)$$

$$EN_i(k) = \frac{x(n)}{\sum_{l=1}^N x(l)} \quad (3)$$

Keterangan:

F = *spectral flux*

EN = *power spectrum*

x(n) = koefisien transformasi fourier diskrit ke-n

x(l) = koefisien transformasi fourier diskrit ke-l

N = jumlah koefisien transformasi fourier diskrit

4. Zero-Crossing Rate

$$Z = \frac{1}{2N} \sum_{n=1}^N |sgn[x(n)] - sgn[x(n-1)]| \quad (4)$$

$$sgn[x(n)] = \begin{cases} 1, & x(n) \geq 0 \\ -1, & x(n) < 0 \end{cases} \quad (5)$$

Keterangan:

Z = *zero-crossing rate*

sgn[x(n)] = fungsi tanda

x(n) = amplitudo sinyal diskrit ke-n

N = jumlah amplitudo sinyal diskrit (panjang potongan *short time*)

5. Energy Entropy

$$H = - \sum_{j=1}^K e_j \cdot \log_2(e_j) \quad (6)$$

$$e_j = \frac{E_{subtime j}}{E_{short-time}} \quad (7)$$

$$E_{short-time} = \frac{1}{N} \sum_{n=1}^N |x(n)|^2 \quad (8)$$

Keterangan:

H = *energy entropy*

E_{short-time} = *energy*

E_{sub-time} = *sub-energy (short-energy dibagi menjadi beberapa sub-energy)*

e_j = perbandingan *short-energy* dan *sub-energy*

L = jumlah

x(n) = amplitudo sinyal diskrit ke-n

N = jumlah amplitudo sinyal diskrit (panjang potongan *short-time*)

K = jumlah perbandingan energi short dan *sub-energy*

Ciri temporal seperti *zero-crossing rate* dan *energy entropy* dapat langsung dihitung dari sinyal audio domain waktu. Sedangkan untuk *spectral centroid, spectral roll-off, dan spectral flux* perlu dilakukan transformasi fourier cepat untuk mengubah sinyal dari domain waktu ke domain frekuensi. transformasi fourier cepat untuk sinyal diskrit adalah sebagai berikut,

$$X_k = \sum_{n=1}^N x_n e^{-i2\pi k \frac{n}{N}} \quad k = 1, \dots, N \quad (9)$$

Keterangan:

X_k = Koefisien transformasi fourier diskrit ke-k

x_n = amplitudo sinyal diskrit ke-n

N = jumlah amplitudo sinyal diskrit (panjang potongan *short-time*)

k = sampel frekuensi

Setelah dihitung, sebanyak 300 lagu masing-masing memiliki 9 ciri. Dari data tersebut dibagi menjadi dua bagian: data latih dan data uji. Data latih adalah data yang digunakan untuk membuat model klasifikator aliran musik, sedangkan data uji adalah data yang digunakan untuk menghitung nilai akurasi model klasifikator yang dibuat. Sehingga data dibagi dari matriks 300 x 9 menjadi matriks 240 x 9 untuk data latih dan matriks 60 x 9 untuk data uji.

Metode klasifikasi yang digunakan adalah *Artificial Neural Network* dan *Random Forest*. Algoritma ANN yang digunakan adalah Algoritma Backpropagation Levenberg-Marquardt dan algoritma *Random Forest* yang digunakan adalah Algoritma Breiman. Proses komputasi ANN dan *Random Forest* dapat dilakukan dengan menggunakan MATLAB.

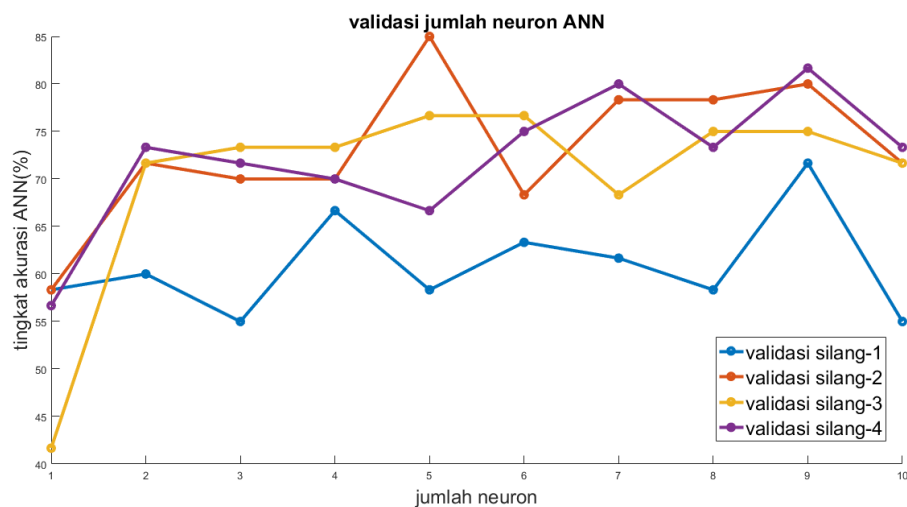
Dalam proses pelatihan membuat model klasifikator, dilakukan *model selection* dengan validasi silang seperti yang ditunjukkan pada buku [4] sebanyak 4 kali. Pembagian data untuk validasi silang ini dilakukan dengan cara membagi matriks 240×9 menjadi matriks 180×9 dan matriks 60×9 dengan 4 pembagian yang berbeda. Matriks 180×9 digunakan sebagai data latih, sedangkan matriks 60×9 dilakukan sebagai data uji.

Validasi yang pertama ini bertujuan untuk menentukan parameter di setiap metode klasifikasi. Untuk metode ANN, dilakukan validasi silang untuk memilih jumlah *neuron* pada sebuah hidden layer yang memiliki tingkat akurasi rata-rata paling tinggi dalam pelatihannya, sedangkan untuk *Random Forest* dilakukan validasi silang untuk memilih jumlah *tree* yang memiliki tingkat akurasi rata-rata paling tinggi dalam pelatihannya. Tingkat akurasi dihitung melalui *confusion matrix*-nya.

Setelah jumlah *neuron* dan *tree* didapatkan sebagai parameter membuat model klasifikasi melalui tingkat akurasi validasi tertinggi, langkah selanjutnya adalah menguji tingkat akurasi model yang dibangun berdasar data latih. Pengujian dilakukan pada data uji. Validasi yang kedua pun dilakukan untuk mendapat tingkat akurasi yang diusahakan berlaku *general*. Validasi silang dilakukan sebanyak 5 kali. Matriks 240×9 dapat dilatih menjadi model klasifikator untuk mengklasifikasi data uji matriks 60×9 . Tingkat akurasi model ini dapat divalidasi dengan pembagian data yang berbeda sehingga tingkat akurasi diupayakan merupakan tingkat akurasi yang *general*.

3. Hasil dan Pembahasan

Metode klasifikasi yang pertama digunakan adalah metode ANN. Sebelum pengujian, 4 validasi silang dilakukan dengan pembagian data yang berbeda untuk data latih dan data uji. Hasil tingkat akurasi beragam seperti gambar di bawah.



Gambar 1. Validasi jumlah neuron pada metode ANN

Berdasar hasil di atas, jumlah neuron yang akan digunakan dipilih dengan melihat tingkat akurasi rata-rata tertinggi, yaitu 9 neuron dengan tingkat akurasi rata-rata 77.08 %.

Setelah validasi untuk *model selection* seperti yang ditunjukkan gambar 1 telah dilakukan. Pengujian pun dilakukan dengan menggunakan data latih dan data uji. Kemudian untuk mendapatkan tingkat akurasi yang

berlaku *general*, maka validasi untuk tingkat akurasi hasil uji dilakukan dengan validasi silang sebanyak 5 kali, membagi data latih dan data uji dengan pembagian data yang berbeda. Hasil dari validasi hasil uji adalah sebagai berikut.

Tabel 1. Confusion Matrix Validasi Hasil Uji 1 ANN

	Metal	Pop	Rock
Metal	18	1	1
Pop	0	12	8
Rock	7	3	10

Tabel 2. Confusion Matrix Validasi Hasil Uji 2 ANN

	Metal	Pop	Rock
Metal	17	0	3
Pop	0	17	3
Rock	2	1	17

Tabel 3. Confusion Matrix Validasi Hasil Uji 3 ANN

	Metal	Pop	Rock
Metal	19	0	1
Pop	0	19	1
Rock	0	4	16

Tabel 4. Confusion Matrix Validasi Hasil Uji 4 ANN

	Metal	Pop	Rock
Metal	18	0	2
Pop	0	18	2
Rock	5	3	12

Tabel 5. Confusion Matrix Validasi Hasil Uji 5 ANN

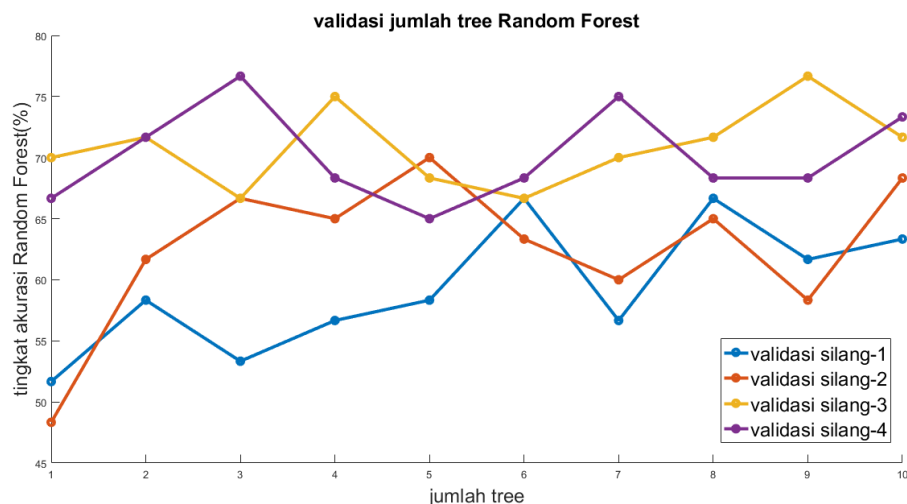
	Metal	Pop	Rock
Metal	16	1	3
Pop	0	19	1
Rock	7	1	12

Tabel 6. Tingkat Akurasi Validasi Hasil Uji ANN

Validasi Hasil Uji	Tingkat Akurasi (%)
1	66.67
2	85.00
3	90.00
4	80.00
5	78.33

Dengan menjumlahkan 5 tingkat akurasi dibagi dengan 5 maka diperoleh rata-rata tingkat akurasi hasil uji keseluruhan sebesar 80 %. Nilai ini merupakan nilai akurasi pengklasifikasi aliran musik berdasar bentuk permukaan musik dengan metode ANN.

Metode pengklasifikasi yang kedua adalah metode Random Forest. Sebelum pengujian, 4 validasi silang pun dilakukan dengan pembagian data yang berbeda untuk data latih dan data uji. Hasil tingkat akurasi beragam seperti gambar di bawah.



Gambar 2.. Validasi jumlah tree pada metode Random Forest

Berdasar hasil di atas, jumlah tree yang akan digunakan dipilih dengan melihat tingkat akurasi rata-rata tertinggi, yaitu 10 tree dengan tingkat akurasi rata-rata 69.17 %.

Setelah validasi untuk *model selection* seperti yang ditunjukkan gambar 1 telah dilakukan. Pengujian pun dilakukan dengan menggunakan data latih dan data uji. Kemudian untuk mendapatkan tingkat akurasi yang

berlaku *general*, maka validasi untuk tingkat akurasi hasil uji dilakukan dengan validasi silang sebanyak 5 kali, membagi data latih dan data uji dengan pembagian data yang berbeda. Hasil dari validasi hasil uji adalah sebagai berikut.

Tabel 7. Confusion Matrix Validasi Hasil Uji 1 Random Forest

	Metal	Pop	Rock
Metal	6	1	7
Pop	8	18	0
Rock	6	1	13

Tabel 8. Confusion Matrix Validasi Hasil Uji 2 Random Forest

	Metal	Pop	Rock
Metal	15	5	0
Pop	4	13	4
Rock	1	2	16

Tabel 9. Confusion Matrix Validasi Hasil Uji 3 Random Forest

	Metal	Pop	Rock
Metal	19	3	0
Pop	1	10	1
Rock	0	7	19

Tabel 10. Confusion Matrix Validasi Hasil Uji 4 Random Forest

	Metal	Pop	Rock
Metal	11	7	4
Pop	4	13	1
Rock	5	0	15

Tabel 11. Confusion Matrix Validasi Hasil Uji 5 Random Forest

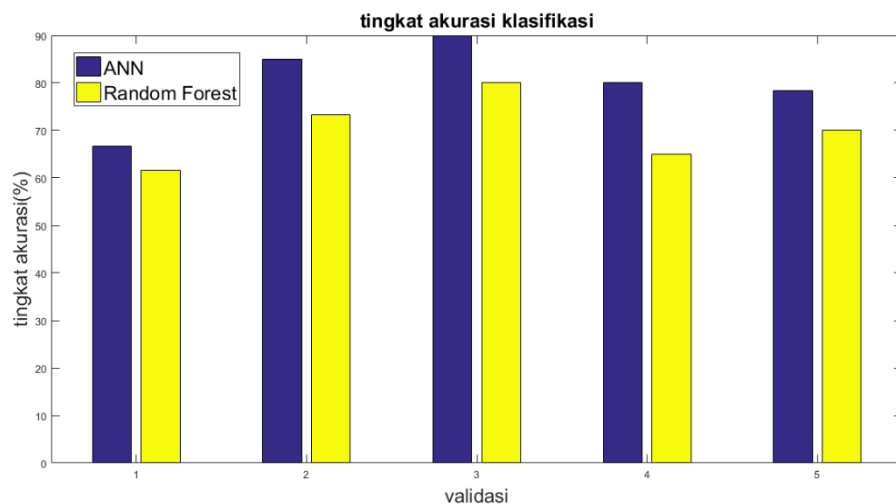
	Metal	Pop	Rock
Metal	13	8	1
Pop	7	10	0
Rock	0	2	19

Tabel 12. Tingkat Akurasi Validasi Hasil Uji Random Forest

Validasi Hasil Uji	Tingkat Akurasi (%)
1	61.67
2	73.33
3	80.00
4	65.00
5	70.00

Dengan menjumlahkan 5 tingkat akurasi dibagi dengan 5 maka diperoleh rata-rata tingkat akurasi hasil uji keseluruhan sebesar 70 %. Nilai ini merupakan nilai akurasi pengklasifikasi aliran musik berdasar bentuk permukaan musik dengan metode Random Forest.

Dari hasil tingkat akurasi yang telah didapat maka dapat dilihat hasil tingkat akurasi keseluruhan pada gambar di bawah ini.



Gambar 2.. Validasi jumlah tree pada metode Random Forest

Berdasar hasil di atas, tingkat akurasi rata-rata untuk klasifikasi aliran musik dengan metode ANN adalah 80.00 %, sedangkan dengan metode Random Forest adalah 70.00%.

4. Simpulan

Dapat dilihat bahwa klasifikasi aliran musik dapat dilakukan berdasar ciri spektral dan temporal, yaitu 9 ciri seperti: *mean-spectral centroid*, *mean-spectral roll-off*, *mean-spectral flux*, *mean-zero-crossing rate*, *std-spectral centroid*, *std-spectral*

roll-off, *std-spectral flux*, *std-zero-crossing rate*, dan *min-energy entropy* menghasilkan tingkat akurasi 80.00% dengan metode ANN dan 70.00% dengan metode Random Forest.

Sebelum pengklasifikasian dilakukan dengan data uji, perlu dilakukan *model selection* untuk *tuning* parameter yang digunakan dalam metode klasifikasi dengan cara validasi silang seperti jumlah *neuron* pada ANN atau jumlah *tree* pada *Random Tree*. Karena seperti yang dapat dilihat dari Gambar 1 dan Gambar 2, beberapa tingkat akurasi didapatkan sangat rendah untuk 4 kali validasi, atau sangat tinggi di beberapa pembagian data validasi namun sangat rendah di beberapa pembagian data validasi lainnya. Hasil yang diinginkan adalah tingkat akurasi yang tinggi secara umum dan didapat dalam penelitian kali ini jumlah neuron adalah 9 dengan rata-rata tingkat akurasi validasi 77.08 % dan jumlah tree adalah 10 dengan rata-rata tingkat akurasi validasi 80%.

Ucapan Terimakasih

Terimakasih kepada Pak Bambang Heru Iswanto yang telah membimbing dalam penelitian ini.

Daftar Acuan

Jurnal

- [1] G. Tzanetakis, and P. Cook, Musical genre classification of audio signal, IEEE Transactions on Speech and Audio Processing vol. 10 no. 5, IEEE, Jul 2002, pp. 293- 302.
- [2] G. Peeters, A Large Set of Audio Features for Sound Description, IRCAM, (2003), p. 11-13.

Buku

- [3] T. Giannakopoulos, Introduction to Audio Analysis: Matlab Approach, San Diego, Elsevier. (2014), p. 59-97.
- [4] T. Hastie, J. H. Friedman, R. Tibshirani, The Element of Statistical Learning: Data Mining, Inference, and Prediction: 2nd edition, Springer, (2009), p. 241 - 242