

PENERAPAN JARINGAN SYARAF TIRUAN DENGAN RADIAL BASIS FUNCTION UNTUK PENGENALAN GENRE MUSIK

Tangguh Gradhianta[♣], Dr. Yusuf Fuad, M. App. Sc.[♣]

[♣]Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Surabaya
60231

email : t.gradhianta@gmail.com

ABSTRAK

Kecerdasan buatan dapat diaplikasikan dalam banyak bidang dalam kehidupan. Penerapan kecerdasan buatan diantaranya dapat dicapai dengan pendekatan jaringan syaraf tiruan (JST). Salah satu contoh metode jaringan syaraf tiruan yang dikenal adalah metode *radial basis function* (RBF). Jaringan syaraf tiruan *radial basis function* (JST RBF) dikenal sebagai salah satu jaringan syaraf yang memiliki tiga lapis bersifat *feedforward* yang dapat memecahkan masalah klasifikasi atau pengenalan pola. Dalam penelitian ini JST RBF digunakan untuk mengklasifikasi musik ke dalam *genre* (jenis) musik berdasarkan kedekatannya dengan target. Sebagai kebutuhan, jenis musik yang dipakai pada penelitian ini adalah campursari, keroncong, *pop*, dan *rock* dengan 3 macam durasi yaitu 2 detik, 5 detik, dan 10 detik pada setiap musik. Sedangkan banyak neuron yang dipakai dalam lapisan tersembunyi sebanyak 56 neuron. Bahan masukan (*input*) yang digunakan dalam JST RBF ini berformat **.mp3* yang diunduh dari internet yang selanjutnya dikonversi ke dalam format **.wav* dan diekstraksi dengan menggunakan *mel-frequency cepstrum coefficients* (MFCC). Teknik ini mengekstraksi fitur suara yang terdapat pada data musik. Koefisien yang digunakan dalam penelitian ini sebanyak 7 koefisien untuk setiap data musik. Dari hasil simulasi program menunjukkan bahwa JST RBF dapat mengklasifikasi musik dengan akurasi paling tinggi pada data uji berdurasi 10 detik sebesar 75%.

Kata kunci : *Genre*, jaringan syaraf tiruan, kecerdasan buatan, *mel-frequency cepstrum coefficients*, musik, *radial basis function*.

I. PENDAHULUAN

Kecerdasan buatan merupakan salah satu bagian pemrograman komputer yang membuat mesin agar dapat melakukan pekerjaan sebaik manusia [6]. Kecerdasan buatan dapat dicapai dengan pendekatan jaringan syaraf tiruan (JST).

JST bermanfaat semisal untuk pengklasifikasian dan peramalan. Pada JST terdapat dua jenis pembelajaran yaitu metode tanpa pengawasan (*unsupervised learning method*) dan metode dengan pengawasan (*supervised learning method*). Salah satu contoh metode JST diantaranya *radial basis function* (RBF) dimana JST RBF memiliki dua tahap pembelajaran yaitu metode tanpa pengawasan yang menghasilkan bobot latih dan tahap kedua yaitu metode dengan pengawasan yang membuahkan solusi linier [1].

Salah satu contoh aplikasi JST RBF adalah seperti dalam Turnbull dan Elkan (2005) dengan artikel berjudul "*Fast Recognition of Musical Genres Using RBF Networks*". Dalam salah satu percobaan JST RBF tersebut digunakan 800 data latih dan 200 data uji dengan 90 neuron pada lapis tersembunyi yang menghasilkan akurasi pengenalan jenis musik sebesar 71,5%. Pada percobaan ini, salah satu cara untuk menginisialisasi parameter fungsi basis adalah dengan menggunakan algoritma *kmeans*. *Input* yang digunakan adalah hasil ekstraksi dengan algoritma *mel-frequency cepstrum coefficients* (MFCC) sehingga menghasilkan vektor citra audio.

Ciri khas suatu musik dapat diketahui dari *refrain*. *Refrain* atau *chorus* adalah bagian musik yang paling representatif dan menonjol secara tematis dari sebuah lagu. Bagian ini sering diulangi dan merupakan kesan dari sebuah lagu [4]. Oleh karena itu penulis mengupayakan pengenalan jenis musik dengan pendekatan JST RBF menggunakan data musik berupa lagu dimana durasi data diambil dari awal *refrain*.

II. DASAR TEORI

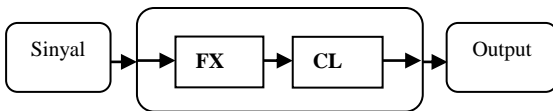
2.1. Musik dan Jenis Musik

Musik adalah kombinasi suara yang terorganisasi dalam suatu cara dengan ritme, melodi, dan harmoni sehingga dapat dinikmati oleh pendengar. Musik dikembangkan, dimanipulasi dan dirancang ulang oleh manusia. Pada kenyataannya

musik merupakan fenomena alam yang berakibat dari prinsip-prinsip matematika dan fisika [2]. Jenis (*genre*) musik adalah kelas suatu musik dengan kumpulan sifat umum yang dibuat melalui persepsi manusia dengan cara membedakan satu musik dengan yang lain. Sebuah jenis musik ditandai oleh struktur instrumentasi dan ritme musik [5].

2.2. Pengenalan Jenis Musik

Seperti klasifikasi sinyal, pengenalan musik adalah mengenai pengklasifikasian musik ke dalam beberapa kategori yang telah ditentukan dengan tujuan mengembangkan sistem dengan *input* berupa sinyal musik dalam bentuk digital dan *output* berupa informasi mengenai keterkaitan *input* dengan kelas jenis musik yang beragam. Berikut disajikan skema pengenalan jenis musik



Gambar 1. Skema Pengenalan jenis musik

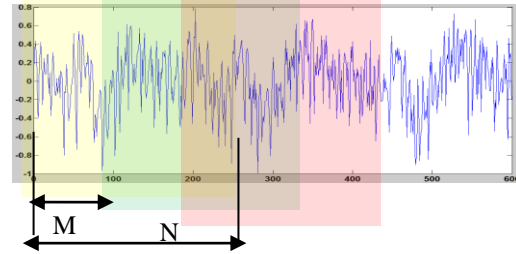
Desain *black box* pada skema diatas yang menghasilkan *output* dari sinyal *input* merupakan subyek penelitian mengenai pengenalan jenis musik. Pada skema diatas, pengenalan musik adalah permasalahan klasifikasi yang berisi dua langkah dasar yang harus dilakukan yaitu ekstraksi fitur (*feature extraction*/FX) dan klasifikasi (*classification*/CL). Tujuan langkah peng ekstraksian fitur adalah untuk memperoleh informasi penting dari data *input*. Langkah klasifikasi bertujuan untuk menemukan kombinasi nilai fitur yang sesuai dengan kategori yang telah tersedia [5].

2.3. Ekstraksi Fitur Suara MFCC

Untuk membuat suatu perbandingan yang selanjutnya dilakukan suatu pengklasifikasian, data musik yang diperoleh harus ditransformasikan supaya informasi yang terkandung dapat diakses. Proses ini dinamakan dengan ekstraksi fitur. Metode peng ekstraksian yang paling populer adalah *mel-frequency cepstrum coefficients* (MFCC) yang didasarkan pada informasi spektrum suara untuk menangkap bagian spektrum yang relevan [5]. Algoritma Peng ekstraksian fitur dengan menggunakan MFCC melalui tahapan-tahapan sebagai berikut [3]:

Langkah 1: *Frame blocking*. Untuk memudahkan dan mempercepat proses analisis suara, dilakukan

pembagian sinyal menjadi beberapa partisi yang selanjutnya disebut dengan *frame*. Berikut adalah contoh representasi *frame blocking*



Gambar 2. Ilustrasi *frame blocking* dengan $M = 100$ dan $N = 256$

Langkah 2: *Windowing*. Untuk setiap *frame*, dilakukan *windowing* dengan menggunakan persamaan

$$y(n) = x(n) w(n), \quad 0 \leq n \leq N - 1 \quad (1)$$

dengan

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

dimana $n = 0, 1, \dots, N - 1$, $x(n)$ sinyal *input*, $y(n)$ sinyal hasil *windowing*, dan $w(n)$ adalah fungsi *hamming window* [7].

Langkah 3: *Fast Fourier transform*. Pada tiap *frame*, dihitung FFT hasil *windowing* dengan persamaan DFT yang telah terdekomposisi menjadi FFT yang diterapkan oleh Cooley dan Tukey pada 1965 [7] yaitu

$$X(m) = X_0(m) + W^m X_1(m) \quad (3)$$

dengan

$$X_0(m) = \sum_{r=0}^{\frac{N}{2}-1} x(2r) W^{2rm} \quad \text{dan}$$

$$X_1(m) = \sum_{r=0}^{\frac{N}{2}-1} x(2r+1) W^{(2r+1)m}$$

dimana $x(2r)$ adalah kelompok sinyal berindeks genap dan $x(2r+1)$ adalah kelompok sinyal berindeks ganjil. Dalam hal ini, banyak titik sampel harus memenuhi bilangan perpangkatan 2 (*radix-2*).

Langkah 4: *Mel-frequency wrapping*. Proses ini dilakukan dengan melibatkan persamaan

$$X_i = \log_{10}(\sum_{k=0}^{N-1} |X(k)| H_i(k)) \quad (4)$$

dengan $|X(k)|$ adalah magnitudo sinyal hasil FFT, $H(k)$ adalah nilai filter frekuensi, dan $i = 1, 2, \dots, M$; M adalah banyak filter [3].

Langkah 5: *Cepstrum*. Langkah selanjutnya menentukan *cepstrum* atau koefisien spektral.

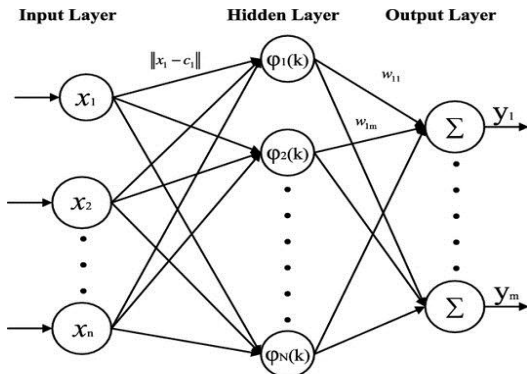
Hasil dari proses ini dinamakan *mel-frequency cepstrum coefficients* (MFCC). Persamaan yang digunakan adalah

$$C_j = \sum_{i=1}^M X_i \cos\left(\frac{\pi j(i-0,5)}{2M}\right) \quad (5)$$

dengan $j = 1, 2, 3, \dots, K$. K adalah banyaknya koefisien MFCC yang diinginkan dan M adalah banyak filter yang digunakan. Dalam hal ini C_j disebut koefisien ke- j MFCC [3].

2.4. Radial Basis Function (RBF)

Jaringan syaraf tiruan metode fungsi basis radial merupakan JST yang memiliki 2 tahap pelatihan. Pada tahap awal, parameter fungsi basis ditentukan secara cepat dengan menggunakan *unsupervised method* yang hanya memerlukan data *input* saja. Tahap kedua pelatihan ini adalah membawa hasil dari unit tersembunyi ke unit *output* secara linier. Fungsi basis radial merupakan fungsi yang bergantung pada jarak antara data dengan suatu pusat data. Fungsi basis radial yang digunakan umumnya nonlinier yaitu $\phi(\|\bar{x} - \bar{c}_i\|)$ dengan $\bar{x} \in \mathbb{R}^n$ dan $\|\cdot\|$ adalah *norm* Euclidean dan himpunan vektor $c = \{\bar{c}_i \in \mathbb{R}^n; i = 1, 2, 3, \dots, m\}$ adalah pusat data dari sampel data [1].



Gambar 3. Ilustrasi Proses JST RBF

Pada gambar diatas, proses dari lapisan *input* menuju lapisan tersembunyi menggunakan metode pembelajaran tidak terawasi (*unsupervised learning*) dan proses yang terjadi dari lapisan tersembunyi menuju lapisan *output* menggunakan metode pembelajaran terawasi (*supervised learning*). Algoritma JST RBF adalah:

1. proses pelatihan :

Langkah 1: menentukan pusat data dan standar deviasi dari data latih.

Langkah 2: menghitung *output* tiap fungsi basis.

Fungsi basis yang digunakan adalah fungsi Gaussian yaitu

$$\phi_{ij}(\|\mathbf{x} - \mathbf{c}_i\|) = \exp\left(-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2\sigma_j^2}\right) \quad (6)$$

dengan \mathbf{x} adalah *input*, \mathbf{c}_i pusat data ke- i , $i = 1, 2, \dots, n$ dan σ_j^2 adalah standar deviasi ke- j , $j = 1, 2, \dots, n$ dengan n adalah banyak pusat data

Langkah 3: menghitung bobot pelatihan dengan menggunakan persamaan

$$\bar{w} = \Phi^+ \bar{t} \quad (7)$$

dengan \bar{w} adalah matriks bobot pelatihan, Φ^+ adalah pseudo invers matriks Gaussian Φ , dan \bar{t} adalah vektor target.

Langkah 4: menghitung *output* JST RBF dengan menggunakan persamaan

$$\bar{y} = \Phi \bar{w} \quad (8)$$

dimana \bar{y} adalah vektor *output*.

Langkah 5: menghitung galat (*error*) antara *output* hasil pelatihan dengan target. Persamaannya yaitu

$$error = t_k - y_k \quad (9)$$

2. Proses Pengujian:

Langkah 1: menentukan pusat data dan standar deviasi dari data uji.

Langkah 2: menghitung *output* tiap fungsi basis. Fungsi basis yang digunakan adalah fungsi Gaussian yaitu

$$\phi_{ij}(\|\mathbf{x} - \mathbf{c}_i\|) = \exp\left(\frac{-\|\mathbf{x} - \mathbf{c}_i\|^2}{2\sigma_j^2}\right) \quad (6)$$

dengan \mathbf{x} adalah *input*, \mathbf{c}_i pusat data ke- i , $i = 1, 2, \dots, n$ dan σ_j^2 adalah standar deviasi ke- j , $j = 1, 2, \dots, n$ dengan n adalah banyak pusat data

Langkah 3: menghitung *output* JST RBF untuk data latih dengan persamaan (8) dimana bobot yang digunakan adalah hasil pelatihan yaitu \bar{w} .

Langkah 4: menghitung galat (*error*) antara *output* hasil pengujian dengan target menggunakan persamaan (9).

III. KESIMPULAN

Dari hasil simulasi yang dilakukan, percobaan ini telah berhasil menerapkan metode MFCC sebagai pengekstraksi ciri data audio yang

digunakan sebagai bahan masukan JST RBF seperti yang telah diterapkan oleh [8]. Pada simulasi ini telah menghasilkan akurasi pengenalan jenis musik sebesar: 54,17% pada data musik berdurasi 2 detik, 70,83% pada data musik berdurasi 5 detik, dan 75% pada data musik berdurasi 10 detik. Dari hasil penelitian ini dapat disimpulkan bahwa akurasi pengenalan jenis musik paling tinggi terdapat pada data musik berdurasi 10 detik dengan rata-rata akurasi sebesar 75%. Hasil ini lebih baik daripada hasil penelitian oleh [8] yang menghasilkan akurasi sebesar 71,5%.

Tabel 1. Prosentase hasil klasifikasi musik dalam %

	2 detik	5 detik	10 detik
Percobaan 1	54,17%	70,83%	75%
Percobaan 2	54,17%	70,83%	75%
Percobaan 3	54,17%	70,83%	75%
	54,17%	70,83%	75%

Hasil klasifikasi yang diambil adalah pada hasil proses pengujian. Dari tabel tersebut terlihat bahwa hasil klasifikasi yang paling baik terdapat pada implementasi algoritma jaringan syaraf tiruan RBF dengan menggunakan data musik berdurasi 10 detik dengan presentase akurasi sebesar 75%.

IV. SARAN

Saran dari penulis yang dapat diberikan untuk penelitian selanjutnya yang berkaitan adalah dapat diupayakannya suatu variasi dalam penelitian tersebut baik dari segi metode pengekstraksian, banyaknya jenis musik, banyak data musik, banyak durasi, dan banyak neuron yang digunakan pada penelitian yang terkait baik dengan menggunakan spesifikasi komputer yang sama maupun berbeda.

DAFTAR PUSTAKA

- [1] Bishop, Christian M. (1995) *Neural Networks for Pattern Recognition*. Universitas Aston: Birmingham, Britania Raya.
- [2] David Jr, Jack H. (1995) *The Mathematics of Music*. Paper Math 1513.5097 [Online]. <http://jackhdavid.thehouseof david.com> [Diakses: 11 April 2013].
- [3] Do, Minh (1994) *DSP Mini-Project: An Automatic Speaker Recognition System*

[Online]. <http://www.ifp.uiuc.edu/> [Diakses: 20 November 2012].

- [4] Goto, Masataka (2006) *A Chorus Section Method for Musical Audio Signals and its Application to a Music Learning Station*. IEEE Transactions on Audio, Speech, and Language Processing vol. 14 no. 5 [Online]. <http://citeseerx.ist.psu.edu> [Diakses: 14 April 2103].
- [5] Kosina, Karin (2002) *Music Genre Recognition*. Skripsi Sarjana. Politeknik Media Teknologi dan Desain Hagenberg: Hagenberg, Jerman [Online]. www.music.mcgill.ca [Diakses: 18 Desember 2011].
- [6] McCarthy, John (2007) *What is Artificial Intelligence?*. Artikel. Universitas Stanford: Amerika Serikat [Online]. <http://www.formal.stanford.edu> [Diakses: 11 April 2013].
- [7] Rao, R. K. dan Kim, D. N dkk. (2010) *Fast Fourier Transform: Algorithms and Applications*. Springer: Amerika Serikat.
- [8] Turnbull, Douglas dan Elkan, Charles (2005) *Fast Recognition of Musical Genres Using RBF Networks*. IEEE Transactions of Knowledge and Data Engineering Vol. 17 No. 4 [Online]. web.cs.swarthmore.edu/~turnbull [10 Januari 2011].