

# KLASIFIKASI JENIS MUSIK BERDASARKAN FILE AUDIO MENGUNAKAN JARINGAN SYARAF TIRUAN LEARNING VECTOR QUANTIZATION

Rocky Yefrenes Dillak<sup>1)</sup>, Dwi Murdaningsih Pangestuty<sup>2)</sup>, Martini Ganantowe Bintiri<sup>3)</sup>

<sup>1)</sup>Jurusan Teknik Informatika AMIKOM Yogyakarta

Jl. Ring Road Utara Condong Catur Sleman-Yogyakarta

<sup>2)</sup>Jurusan Ilmu Komputer Universitas Gadjah Mada Yogyakarta

Bulaksumur, Yogyakarta 55281, Telp (0274) 588688, Fax. (0274) 565223

<sup>3)</sup>Jurusan Teknik Sipil Universitas Sintuwu Maroso

Jl. P. Timor No. 1 Poso Sulawesi Tengah 94615

e-mail : [rocky\\_dillak@yahoo.com](mailto:rocky_dillak@yahoo.com), [dwi\\_186@yahoo.com](mailto:dwi_186@yahoo.com), [gana75mart@yahoo.com](mailto:gana75mart@yahoo.com)

## Abstrak

Klasifikasi merupakan proses identifikasi kelompok dari suatu obyek berdasarkan kesamaan fitur tertentu, dimana setiap kelompok telah terbentuk melalui suatu proses tertentu. Keunikan setiap kelompok direpresentasikan dengan suatu vector fitur. Penelitian ini bertujuan mengembangkan suatu metode yang dapat digunakan untuk melakukan klasifikasi terhadap jenis musik berdasarkan file audio dengan format wav menggunakan algoritma jaringan syaraf tiruan learning vector quantization (LVQ). Pengklasifikasian file audio ke dalam suatu kelompok atau kelas, memerlukan ciri atau fitur dari file audio tersebut. Metode ekstrak fitur yang digunakan untuk memperoleh ciri atau fitur dari file yang dimaksud adalah Sort Time Energy (STE) dan Zero Crossing Rate (ZC) yang diturunkan dalam domain waktu (*time domain*) yang merupakan salah satu komponen data audio. Berdasarkan hasil dari penelitian ini menunjukkan bahwa pendekatan yang diusulkan mampu melakukan klasifikasi terhadap jenis musik berdasarkan file audio berformat wav dengan akurasi sebesar 80%.

**Kata Kunci :** klasifikasi, jenis musik, file audio, format wav, jaringan syaraf tiruan, learning vector quantization

## 1. PENDAHULUAN

Pengenalan pola adalah suatu cabang dalam *artificial intelligence* (AI) dimana komputer dapat dibuat untuk belajar (*learning*) atas suatu pengalaman (pola) dan mampu menyimpan hasil proses belajar tersebut menjadi suatu *knowledge*. Knowledge tersebut dapat digunakan untuk mengenali pola data yang baru yang belum pernah diajarkan. Pola data yang diajarkan kepada komputer dalam konsep pengenalan pola merupakan pasangan kumpulan ciri (*vector* fitur) serta solusi atas sebuah permasalahan (*target*). Pemanfaatan pengenalan pola dalam kehidupan sehari – hari sudah banyak dilakukan misalnya untuk permasalahan klasifikasi, deteksi, identifikasi atau pengelompokan. Tujuan dari *paper* ini adalah mengkaji pemanfaatan pengenalan pola untuk melakukan klasifikasi terhadap music berdasarkan data audio menggunakan jaringan syaraf tiruan *learning vector quantization* (LVQ). Klasifikasi merupakan pengelompokan fitur ke dalam kelas yang sesuai. Dimana vektor fitur pelatihan tersedia dan telah diketahui kelas-kelasnya, kemudian vektor fitur pelatihan tersebut dimanfaatkan untuk merancang pemilah maka pengenalan pola ini disebut terbimbing (*supervised*). Agar dapat mengklasifikasi objek ke dalam suatu kelompok atau kelas, maka perlu diketahui terlebih dulu ciri atau fitur dari objek tersebut. Untuk kasus klasifikasi lagu berdasarkan jenis musik ini, metode ekstraksi fitur yang dapat digunakan adalah *Sort Time Energy* (STE) dan *Zero Crossing Rate* (ZC) yang diturunkan dalam domain waktu (*time domain*) yang merupakan salah satu komponen data *audio*.

Proses klasifikasi dapat dilakukan dengan menggunakan *Artificial Neural Network* (ANN) atau Jaringan Syaraf Tiruan metode LVQ untuk melatih atau *training* sekaligus untuk menguji atau *testing* data. LVQ merupakan salah satu model pelatihan dengan supervisi (*supervised learning*) yang arsitektur jaringannya berlayer tunggal (*single layer*). Jaringan terdiri dari beberapa unit masukan (*input vector*), dan memiliki beberapa unit keluaran (*output vector*). LVQ melakukan pembelajaran pada lapisan kompetitif yang terawasi. Lapisan kompetitif akan secara otomatis belajar untuk mengklasifikasikan vektor-vektor *input*. Kelas-kelas yang didapat sebagai hasil dari lapisan kompetitif ini hanya tergantung pada jarak antara vektor-vektor *input* dengan vektor bobot dari kelas-kelas. Jika vektor *input* mendekati sama dengan vektor bobot suatu kelas maka lapisan kompetitif akan mengklasifikasikan vektor *input* tersebut ke dalam kelas yang sama.

## 2. TINJAUAN PUSTAKA

Sejumlah penelitian mengenai *pattern recognition* berbasis *machine learning* menggunakan data audio sebagai data input pada sistem telah banyak dilakukan. Algoritma biasanya diterapkan untuk identifikasi suara, *clustering* data audio, klasifikasi suara dan musik, *retrieval* musik dan audio berdasarkan similaritas, segmentasi audio, dan lain sebagainya.

Dewi, et al (2010) mengembangkan suatu sistem klasifikasi lagu anak-anak berdasarkan parameter mood. Penelitian ini menggunakan fitur *rhythm pattern* dari musik, sedangkan untuk penggunaan parameter mood berdasarkan pada model energi-stres Robert Thayer. Hasil klasifikasi diperoleh dari penggunaan dua buah metode, yaitu menggunakan *K-Nearest Neighbor* dan *Self-Organizing Map*, kemudian dibandingkan dengan mood yang diperoleh dari ahli psikologi anak. Sistem ini mampu memperoleh akurasi sebesar 66.67% untuk klasifikasi musik menggunakan penggabungan K-NN dan SOM dengan data sebanyak 30 lagu. Sedangkan untuk klasifikasi dengan jumlah data sebanyak 120 lagu, akurasi mencapai 73.33% untuk metode K-NN dan 86.67% untuk metode SOM.

Mayer, et al (2010) mengembangkan klasifikasi *genre* musik dengan mengkombinasikan fitur yang diekstrak dari domain audio musik dengan konten tekstual dari liriknya. Pada penelitian ini klasifikasi dilakukan dengan menggunakan beberapa paradigma *machine learning* yang berbeda. Hasil penelitian menunjukkan bahwa kombinasi fitur yang berasal dari dua *subspace* (audio dan lirik) ini mampu mencapai hasil yang lebih baik bila dibandingkan dengan hanya menggunakan fitur audio tunggal.

Setiawan (2009) melakukan penelitian tentang klasifikasi suara berdasarkan *gender* menggunakan algoritma *K-Means*. Metode ekstraksi fitur sinyal suara yang digunakan pada penelitian ini berbasis pada domain waktu dan domain frekuensi. Dari hasil penelitian menunjukkan bahwa sistem mampu mengklasifikasi pola data dengan baik.

Dan untuk penggunaan JST LVQ yang juga merupakan salah satu metode berbasis *machine learning* telah banyak digunakan. Metode ini biasanya diterapkan untuk masalah klasifikasi dan peramalan.

Charami, et al (2007) melakukan evaluasi dengan membandingkan performa dua metode klasifikasi dalam hal *retrieval* informasi musik. Dari hasil evaluasi peneliti tidak bisa mencapai kesimpulan yang valid.

Li, et al (2011) melakukan penelitian tentang peramalan terjadinya lidah api matahari dengan mengkombinasikan salah satu metode *unsupervised clustering* (algoritma *K-Means*) dan LVQ. Algoritma *K-Means* digunakan untuk mengatasi masalah ketidakseimbangan dalam *dataset*. Dari hasil penelitian diperoleh akurasi sebesar 74.58%.

LVQ merupakan salah satu algoritma dalam JST untuk melakukan pembelajaran secara terawasi. Pada algoritma ini setiap kelas memiliki vektor bobot, dimana banyaknya komponen vektor bobot tersebut sesuai dengan banyaknya komponen vektor *input*. Penentuan kelas atau *output* suatu vektor *input* adalah dengan melakukan pengukuran jarak antara vektor *input* tersebut terhadap vektor bobot dari setiap kelas, selanjutnya kelas dari vektor bobot yang memiliki jarak yang terkecil yang menjadi hasil *output*nya. Sedangkan untuk mengupdate bobot dalam proses pembelajaran dapat dihitung dengan menggunakan persamaan (1) atau (2).

Jika  $T = C_j$  maka :

$$W_j = W_j + \alpha(X_i - W_j) \quad (1)$$

Jika  $T \neq C_j$  maka :

$$W_j = W_j - \alpha(X_i - W_j) \quad (2)$$

dimana,

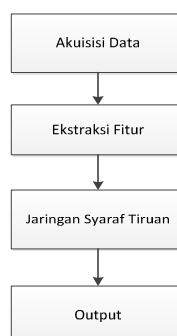
$W_j$  : Bobot ke-j suatu kelas

T : Target

$C_j$  : Hasil *output*

### 3. METODE PENELITIAN

Tahapan dalam penelitian ditunjukkan pada Gambar 1.



Gambar 1. Arsitektur Sistem

### 3.1. Akuisisi Data

Data pada penelitian ini berupa data audio dalam format wav yang diperoleh dari beberapa lagu dengan jenis musik yang berbeda. Sebagai sampel, untuk setiap lagu dilakukan pemotongan dengan panjang 10 detik dan hanya diambil pada bagian refrainnya saja. Dan format wav yang digunakan bertipe 16 bit dengan frekuensi 8000 Hz, artinya terdapat 8000 sampel per detik.

### 3.2. Ekstraksi Fitur

Dalam penelitian ini metode ekstraksi fitur yang digunakan adalah *Sort Time Energy* (STE) dan *Zero Crossing Rate* (ZC) yang diturunkan dalam domain waktu yang merupakan salah satu komponen data audio.

*Sort Time Energy* → Menandakan kekerasan suara pada waktu yang pendek :

$$STE = \frac{1}{N} \sum_{n=1}^N X(n)^2 \quad (3)$$

N : Jumlah Sampel

X(n) : Nilai Sinyal dari sampel

*Zero Crossing Rate* → Sampel berurutan pada sebuah sinyal digital memiliki perbedaan tanda, ukuran dari noise sebuah sinyal pada fitur domain :

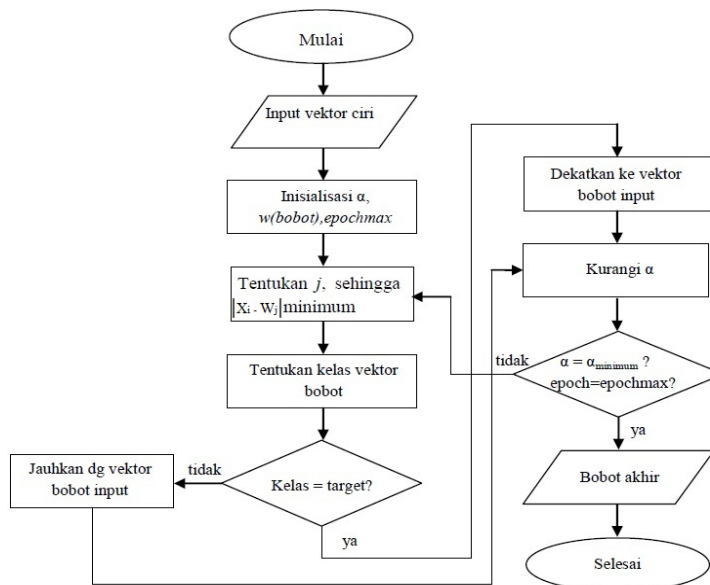
$$ZC = \frac{\sum_{n=1}^N |\text{sgn} x(n) - \text{sgn} x(n-1)|}{2N} \quad (4)$$

sgn x(n) : nilai dari x(n), bernilai 1 jika x(n) positif, -1 jika x(n) negatif

N : Jumlah Sampel

### 3.3. Pelatihan JST Learning Vector Quantization (LVQ)

Fitur yang digunakan sebagai *input* bagi pelatihan JST adalah *sort time energy* dan *zero crossing rate*, sehingga terdapat dua (2) neuron dalam *input layer* pada jaringan LVQ. Sedangkan target dari JST adalah "Pop", "R&B" atau "Rock" yang merepresentasikan kelas-kelas yang ada, sehingga terdapat tiga (3) neuron dalam *output layer* pada jaringan LVQ. Algoritma JST LVQ seperti ditunjukkan pada Gambar 2.



Gambar 2. Pelatihan JST LVQ

#### 4. HASIL DAN PEMBAHASAN

Data yang digunakan sebagai data uji adalah sebanyak 10 buah file wav seperti ditunjukkan pada Tabel 1

**Tabel 1.** Data pengujian

Nama file	Kelas	Hasil JST
1.wav	pop	jazz
2.wav	pop	pop
3.wav	pop	pop
4.wav	jazz	jazz
5.wav	jazz	jazz
6.wav	jazz	jazz
7.wav	jazz	jazz
8.wav	dangdut	dangdut
9.wav	dangdut	jazz
10.wav	dangdut	dangdut

Berdasarkan hasil pelatihan JST LVQ terhadap data latih yang ada maka diperoleh vektor bobot terbaik yang kemudian akan digunakan untuk melakukan pengujian dalam mengenali pola data *input*, sehingga dapat ditentukan kelas dari data uji yang ada.

Hasil pengujian yang diperoleh (Tabel 1) digunakan untuk menghitung akurasi dengan persamaan :

$$\text{Akurasi} = (\text{jumlah klasifikasi benar} / \text{jumlah data}) \times 100\% \quad (5)$$

Hasil pengujian menunjukkan bahwa sistem dapat mengenali pola data dengan baik dan mencapai akurasi sebesar 80%.

#### 5. KESIMPULAN

Metode penelitian yang dikembangkan dapat digunakan untuk melakukan klasifikasi jenis musik dari suatu lagu dengan menggunakan fitur data audio sebagai *input* sistem. Hasil pelatihan menggunakan Jaringan Syaraf Tiruan – *Learning Vector Quantization* (LVQ) memiliki kinerja yang cukup baik karena mampu mengenali pola data dengan akurasi mencapai 80 %.

#### DAFTAR PUSTAKA

- Dewi, K.C., Harjoko, A., 2010, Kid's Song Classification Based on Mood Parameters Using K-Nearest Neighbor Classification Method and Self Organizing Map, *The Second International Conference on Distributed Frameworks and Applications 2010*
- Mayer, R., Rauber, A., 2010, Building Ensembles of Audio and Lyrics Features to Improve Musical Genre Classification, *In Proceedings of the ACM 13th International Conference on Multimedia 2010*
- Setiawan, A., 2009, Analisis Klasifikasi Suara Berdasarkan Gender Dengan Format Wav Menggunakan Algoritma K-Means, *Sains dan Teknologi*, 2 (2). ISSN 1979-6870
- Charami, M., Halloush, R., Tsekeridou, S., 2007, Performance Evaluation of TreeQ and LVQ Classifiers for Music Information Retrieval, *International Federation for Information Processing, Volume 247, Artificial Intelligence and Innovations 2007*
- Li, R., HuaNing, W., YanMei, C., Xin, H., 2011, Solar flare forecasting using learning vector quantity and unsupervised clustering techniques, *Science China Press and Springer-Verlag Berlin Heidelberg 2011*