Indonesian's Dangdut Music Classification Based on Audio Features

Ferdinand Mahardhika
Computer Science Departement,
BINUS Graduate Program - Doctor of Computer Science.
Bina Nusantara University
Jakarta, Indonesia 11480
ferdinand.mahardhika001@binus.ac.id

Harco Leslie Hendric Spits Warnars
Computer Science Departement,
BINUS Graduate Program - Doctor of Computer Science.
Bina Nusantara University
Jakarta, Indonesia 11480
spits.hendric@binus.ac.id

Yaya Heryadi
Computer Science Departement,
BINUS Graduate Program – Doctor of Computer Science.
Bina Nusantara University
Jakarta, Indonesia 11480
yayaheryadi@binus.edu

Lukas
Cognitive Engineering Research Group (CERG),
Faculty of Engineering, Universitas Katolik Atma Jaya.
Jakarta, Indonesia 12930
lukas@atmajaya.ac.id

***Abstract*—The uniqueness of modern dangdut music today is in the beat and melody of music that is relatively faster like techno dance music, and the progressive melody arrangement from synthesizer keyboard. In this paper, we will extract audio features from music, and then used that features for classification based on machine learning methods. Support Vector Machine (SVM) was used in this study. In this paper's experiment we use 32 audio music files, consisting of 16 for dangdut music and 16 for techno dance music. The results of this testing get varying accuracy levels, between 80% to 90% for each music audio file that has been successfully classified.**

Keywords—dangdut music classification, machine learning, Support Vector Machine

I. Introduction

Very extraordinary music in Indonesia today, where music listeners in Indonesia vary widely, from a young age to old age. Various types of music exist in this country, from Rock, Pop, HipHop, RnB, Rap, Traditional, etc., are growing rapidly in Indonesia. Lately there is an extraordinary phenomenon, where the type of music "Dangdut" starts to get a very extraordinary response[1]. In the era of the 80s, this music was typical of "old style" music that was not so popular, and its listeners were dominated by the lower classes and more often heard in the villages. Dangdut music is basically wilderness music from the Arab region with a combination of "tabla" musical instruments from India, which then enter the country of Indonesia and experience alkulturation with local culture.

Time has changed, nowadays dangdut music is music that is quite popular [2] because it has a more modern "look", especially with the influence of western music. "Dangdut Koplo Remix" is so popular today, a modern form of dangdut music where a blend of dangdut music with "techno dance music" that originating from European countries. The uniqueness of modern dangdut music today is in the beat and melody of music that is relatively faster like techno dance music[3], and the progressive melody arrangement from synthesizer keyboard. Between dangdut music and techno dance music, sometimes there are similarities that are at first glance difficult to distinguish.

In this paper, we will extract audio features from music, and then used that features for classification based on machine learning methods. The goal in this study is to do music clustering between two class of genre music, which is dangdut music and techno dance music.

II. Related Work

There have been several studies conducted to classify music using audio features. Basically this is part of music information retrieval (MIR). In 2014, Konstantin Markov [4] uses several features to classify music, one of these features is: MFCC (Mel Frequency Cepstral Coefficients) features and TMBR (timbre features). Classification of 8 musical genres, including: Blues, Electronic, Rock, Classical, Folk, Jazz, Country and Pop. The algorithm used as a classifier is Support Vector Machine (SVM).

In 2016, Prafulla et al [5] used 5 features of audio that are used as a basis for classification. The features that they extract are: rhythm, timbre, pitch, tonality and dynamic features. This research was conducted on Indian music genre. The feature is extracted using MIR Toolbox in MATLAB. The algorithm that they used as a classifier is the Neural Network and SVM.

In 2016, Khonglah et al [6] they use audio features, such as zero crossing rate, spectral roll-off, spectral flux, spectral centroid and energy in their study. In 2017, Aisha et al [7] they clustered traditional Indonesian music using 11 features on the time domain and frequency domain. These features include energy, entropy energy, zero-crossing rate, spectral centroid, spectral entropy, spectral flux, spectral rolloff, mel-frequency cepstral coefficient, harmonic, chroma vector, and spectral zone. Meanwhile in the terms of success for doing the classification, there is few example show us that they success in their study using SVM [4],[5],[8],[9],[10] as classifier in audio classification.

978-1-5386-9422-0/18/$31.00 ©2018 IEEE

III. METHODOLOGY

In this study we used nine features from audio [7] to do music classification. The features include time-domain audio features, which consist of energy, entropy energy and zero crossing rate. And also frequency-domain audio features, which consist of spectral centroid, spectral entropy, spectral flux, spectral roll-off, harmonic, mel-frequency cepstral coefficient, and chroma vector.

Following is the equation of audio features:

1. Energy:

$$E(i) = \sum_{n=1}^{W_L} |x_i(n)|^2 \quad (1)$$

2. Entropy Energy:

$$H(i) = -\sum_{j=1}^{K} e_j \log_2(e_j) \quad (2)$$

3. Zero Crossing Rate:

$$Z(i) = \frac{1}{2W_L} \sum_{n=1}^{W_L} |\mathrm{sgn}[x_i(n)] - \mathrm{sgn}[x_i(n-1)]| \quad (3)$$

4. Spectral Centroid:

$$C_i = \frac{\sum_{k=1}^{Wf_L} kX_i(k)}{\sum_{k=1}^{Wf_L} X_i(k)} \quad (4)$$

5. Spectral Entropy:

$$H = -\sum_{f=0}^{L-1} nf.\log_2(nf) \quad (5)$$

6. Spectral Flux:

$$Fl_{(i,i-1)} = \sum_{k=1}^{Wf_L} (EN_i(k) - EN_{i-1}(k))^2 \quad (6)$$

7. SpectralRoll-Off:

$$\sum_{k=1}^{m} X_i(k) = C \sum_{k=1}^{Wf_L} X_i(k) \quad (7)$$

8. Harmonic:

$$r(i,k) = \frac{\sum_{j=m}^{m+n-1} s(j)s(j-k)}{\left(\sum_{j=m}^{m+n-1} s(j)^2 \times \sum_{j=m}^{m+n-1} s(j-k)^2\right)^{0.5}}$$

$$\text{where } H(i) = \max_{k=Q} r(i,k) \quad (8)$$

9. Chroma Vector:

$$v_k = \sum_{n \in S_k} \frac{Xi(n)}{N_k}, k \in 0 \quad (9)$$

IV. EXPERIMENT

This experiment is generally divided into four major steps, starting from data collection, pre processing data, testing data and experimental results. Each section we will explain one by one. The whole process flow in this experiment is illustrated in the Fig. 1.

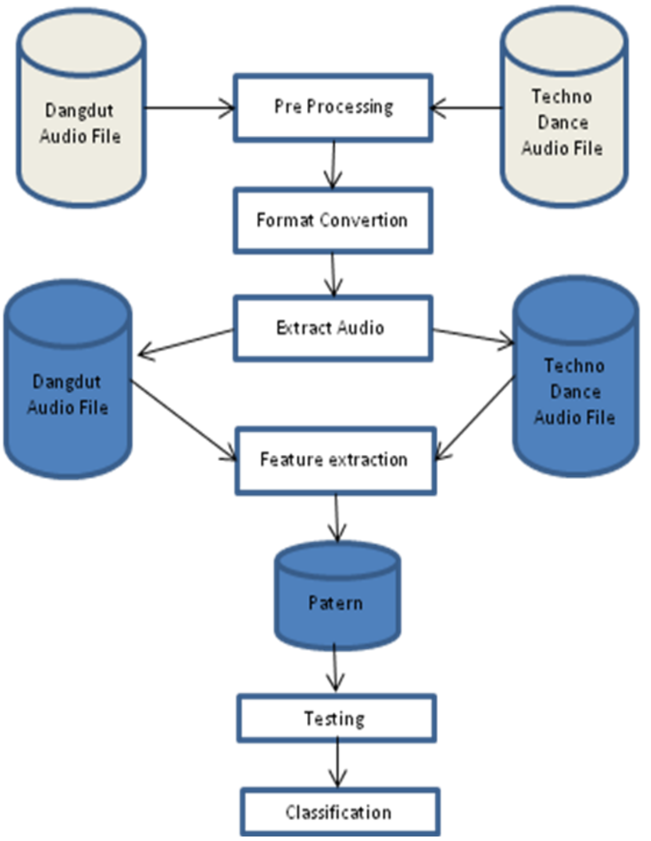

Fig. 1. Process flow in experiment

A. Data Collection

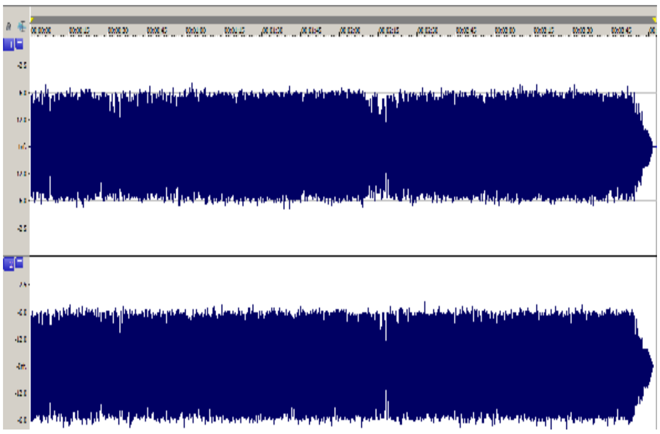

Fig. 2. Stereo audio waveform

We collect audio files for dangdut music and techno dance music using Internet search engine and Youtube[11]. From the results of this collection, 80 audio files for dangdut music and 80 audio files for techno dance music were obtained, so that the total audio files collected were 160 audio files. All audio files collected are in MP3 format with a frequency sampling of 44.1 kHz, stereo with the duration of each song about four minutes. Fig. 2 show stereo audio waveform. In table I, is a list of dangdut song title used in this study:

TABLE I. LIST OF DANGDUT SONG TITLE

No	Song Title (Artist)	No	Song Title (Artist)
01	5 Centi (Via Vallen)	30	Aku Mah Apa Tuh (Cita Citata)
02	Abang Ghoib (Barbie Felia)	31	Aku Tak Biasa (Ria Amelia)
03	Abang Goda (Duo Serigala)	32	Alamat Palsu (Ayu Ting Ting)
04	Aisah Jamilah (Sandrina)	33	Baru 6 Bulan (Connie Nurlita)
05	Aku Jijik (Sandrina)	34	Berdiri Bulu Romaku (Bebizy)
06	Cie Cie (Nella Kharisma)	35	Kau Tercipta Bukan Untukku (Nella Kharisma)
07	Cikini Gondangdia (Duo Anggrek)	36	Keong Racun (Lissa)
08	Cinta Basi (Regina)	37	Konco Mesra (Meggy Diaz)
09	Cinta Ganjil Genap (Susie Legit)	38	Kost Kostan (Duo Serigala)
10	Cinta Kurang Gizi (Via Vallen)	39	Kuper (Susi Ngapak)
11	Cinta Pertama (Hesty Klepek Klepek)	40	Lagi Syantik (Siti Badriah)
12	Cinta Putih (Duo Biduan)	41	Makan Hati (2Racun Youbi sister)
13	Cinta Satu Malam (Melinda)	42	Makan Hati ver 2 (Baby Shima)
14	Cinta Tak Terbatas Waktu (Ucie Sucita)	43	Mama Minta Pulsa (Siti Badriah)
15	Cinta Tulalit (Bebizy)	44	Melanggar Hukum (Siti Badriah)
16	CKC Cuma Kamu Cin (Camel Petir)	45	Pacar Baru (Yuni R)
17	Coba Coba (iMeyMey)	46	Pacar Satu Satunya (Ratu Idola)
18	Cowok Ayam Kampung (Ayu Maharany)	47	Perawan Atau Janda (Cita Citata)
19	Cuit Cuit Witwiw (Neng Oshin)	48	Pernikahan Dini (Cita Citata)
20	Depan Belakang (Salsiah)	49	Pura Pura Bahagia (Cita Citata)
21	Di Tikung Teman (Sandrina)	50	Pusing Pala Barbie (Lula Lula)
22	Dibalas Dusta (Ratu Idola)	51	Sakit Sakit Hatiku (Via Vallen)
23	Digenjot Cinta (Ucie Sucita)	52	Sakitnya Luar Dalam (Ratu Meta)
24	Duda Anak 2 (Nyimas Idola)	53	Sambalado (Ayu Ting Ting)
25	Duyeh (Neng Oshin)	54	Sebelas Duabelas (Nella Kharisma)
26	Enaknya Dikamu (Salsiah)	55	Selfie (Viola Arsa)
27	Gak Ada Waktu Beib (Ghea Youbi)	56	Sianida (Duo Serigala)
28	Galau Ting Ting (Ratu Idola)	57	Suamiku Kawin Lagi (Siti Badriah)
29	Gantung Aku Di Monas (Meggy Diaz)	58	Sudah Cukup Sudah (Zakia Gotik)
59	Goyang 2 Jari (Sandrina)	70	Sumpah Aku Nggak Sakit (Ratu Idola)
60	Goyang Dumang (Cita Citata)	71	Susah Move ON (Duo Anggrek)
61	Goyang Nasi padang (Duo Anggrek)	72	Tak Sanggup Lagi (Dinda Permata)
62	Goyang Pantura (Velline Ratu Begal)	73	Teganya (Susi Legit)
63	Goyang Pokemon (Varra Selvarra)	74	Tercyduk (Desy Ning Nong)
64	Hello Sayang (Baby Sexyola)	75	Terong Dicabein (Siti Badriah)
65	Hey Siapa Kamu (2Racun Youbi Sister)	76	Tua Tua Keladi (Hesty Klepek Klepek)
66	Jagung Bakar (Lynda Moy)	77	Tuh Kan (Ayurasi)
67	Jagung Rebus (Maya Jasika)	78	Undangan Mantan (Siti Badriah)
68	Jaran goyang (Nella Kharisma)	79	Ya Nasib (Susi Legit)
69	Kamu Pelakor (Ratu Idola)	80	Yank Haus (2TikTok)

B. Data Pre Processing

Before the data is used, initial preparations are made for existing data. Audio music data that has been collected, is converted to mono and with a sampling rate of 44.1 kHz then changed the format to WAV. The song duration is taken only 20 seconds from the beginning, and then audio normalization process is carried out to get the same sound hardness level to -16 dB. All this process is done using Praat software[12]. Fig. 3 show the wave form after alteration to 20s and from stereo to mono with format convertion to WAV.

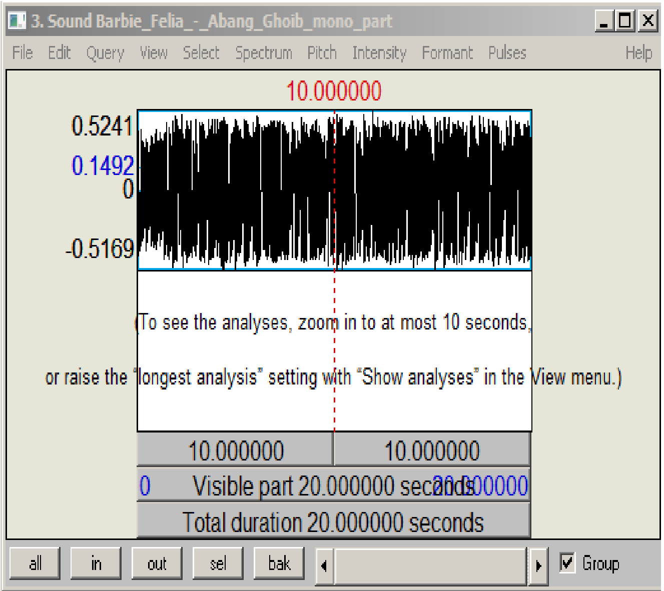

Fig. 3. Wave form after alteration

C. Training

For training and testing purposes, we use Python[13] based programming languages and specific python open library for audio extraction features. In this experiment pyAudioAnalysis [14] was used. The advantage of this open library is its ability to extract audio features and also as a classifier. All can be done automatically, using only changes to few parameters. The classifier in this experiment uses SVM[15]. For each type of music, 80% is taken for training and 20% for testing. A total of 64 audio dangdut music files are placed into a file folder named "DGD", and 64 techno dance music files are placed into a file folder called "TCH". See Fig. 4 dan Fig. 5 for more details.

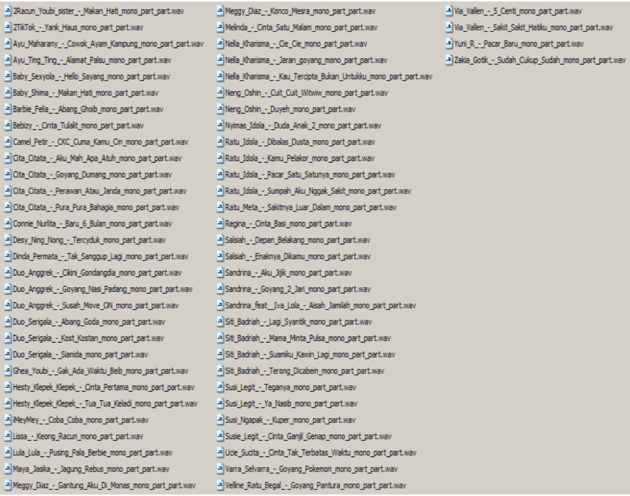

Fig. 4. List in file folder DGD

Training is done by setting the parameter "isSignificant = 0.8" in pyAudioAnalysis, this means that the threshold level for accuracy in this experiment is set on 80%. The training process runs 600 times iteration. Training result are shown in Fig. 6.

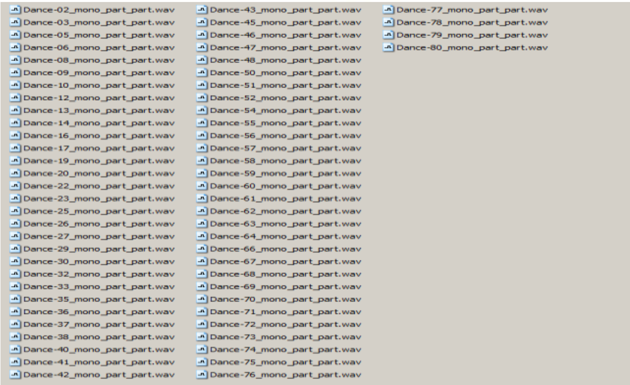

Fig. 5. List in file folder TCH

D. Testing and Result

At this stage an experiment will be conducted to find out whether training results can recognize a music file correctly or not. There are 32 audio music files, consisting of 16 for dangdut music and 16 for techno dance music. The results of this testing get varying accuracy levels, between 80% to 90% for each music audio file that has been successfully classified. Fig. 7, show example of testing result.

Table II shows the overall results in the classification between dangdut music and techno dance music.

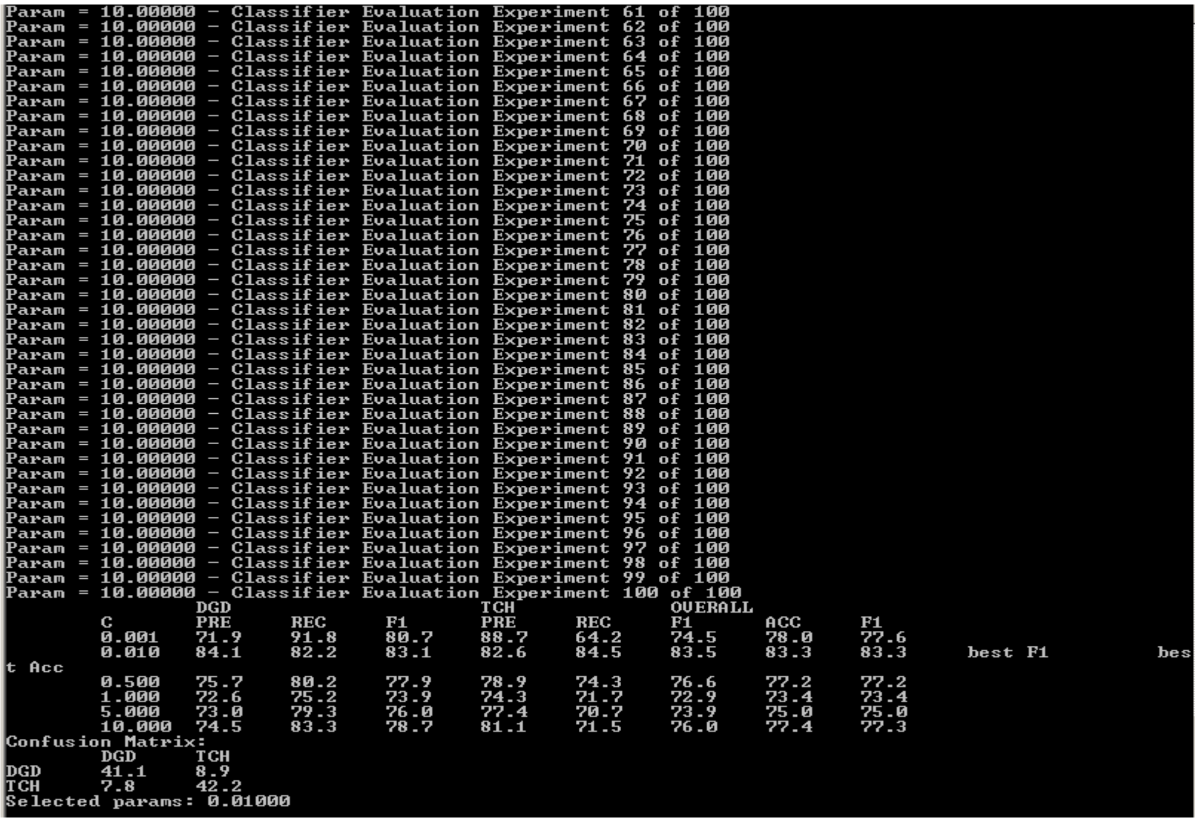

Fig.6. Training Result

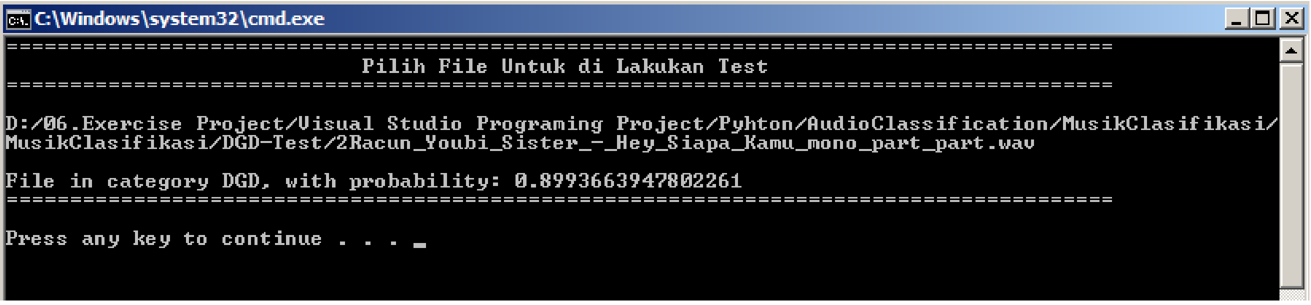

Fig.7. Testing Result

TABLE II. CLASSIFICATION RESULT

	DGD	TCH
Classification success	8	7
Unsuccessful	8	9
	16	16

Explanation for table II, from 16 audio dangdut music files, 8 audio music files were successfully classified and 8 audio music files were not successfully classified. Meanwhile, from 16 audio techno dance music files, 7 audio music files were successfully classified and 9 audio music files were not successfully classified.

V. Conclusion and Future Research

The results from table II are not very satisfying. This is understandable, given the relatively small amount of data for training, and then in the pre processing stage, the music duration is limited to 20 seconds at the beginning.

There are still wide open opportunities for research in this area. For example, with the use of more and varied data, and also not limited to two types of music. Also can be considered to examine similarities in different types of music by using the feature extraction that from audio music.

References

[1] Dangdut Jawa Koplo mengorbit tinggi, retrieved 28 August 2018 from https://lifestyle.sindonews.com/read/1264422/157/dangdut-jawa-koplo-mengorbit-tinggi-1512890757

[2] Dari dangdut koplo bisa tenar ini sejarahnya musik dengan lirik galau mengundang goyang, retrieved 28 August 2018 from http://www.tribunnews.com/seleb/2018/01/24/dari-dangdut-koplo-bisa-tenar-ini-sejarahnya-musik-dengan-lirik-galau-mengundang-goyang

[3] Top definition Electronic Music, retrieved 28 August 2018 from https://www.urbandictionary.com/define.php?term=Electronic%20Music

[4] Markov, Konstantin., Matsui, Tomoko. (2014). Music Genre and Emotion Recognition Using Gaussian Processes. *IEEE Access*, vol. 2, pp. 688 -698.

[5] Kalapatapu, Prafulla., Goli, Srihita., Arthum, Prasanna., Malapati, Aruna. (2016). A Study On Feature Selection And Classification Techniques Of Indian Music. *The 7th International Conference on Emerging Ubiquitous Systems and Pervasive Networks (EUSPN 2016)*, pp. 125-131.

[6] Khonglah, Banriskhem K & Prasanna, S.R. Mahadeva. (2016). Speech / music classification using speech-specific features. *Digital Signal Processing*, vol. 48, pp. 71-83.

[7] Jondya, Aisha Gemala., Iswanto, Bambang Heru. (2017). Indonesian's Traditional Music Clustering Based on Audio Features. *2nd International Conference on Computer Science and Computational Intelligence (ICCSCI)*, pp. 174-181, Bali, Indonesia, 13-14 October 2017.

[8] Hsiao, Shih-Wen., Chen, Shih-Kai., Lee, Chu-Hsuan. (2017). Methodology for stage lighting control based on music emotions. *Information Sciences*, vol. 412–413, pp. 14-35.

[9] Nanni, Loris., Costa, Yandre M.G., Lumini, Alessandra., Kim, Moo Young., Baek, Seung Ryul. (2016). Combining visual and acoustic features for music genre classification. *Expert Systems With Applications*, vol. 45, pp. 108 - 117.

[10] Lacerda, Everton B., Mello, Carlos A.B. (2017). Automatic classification of laryngeal mechanisms in singing based on the audio signal. *Procedia Computer Science*, vol. 112, pp. 2204 - 2212.

[11] https://www.youtube.com/results?search_query=dangdut+koplo

[12] Boersma, Paul & Weenink, David. (2018). Praat: doing phonetics by computer [Computer program]. Version 6.0.42, retrieved 15 August 2018 from http://www.praat.org/

[13] Stef Maruch and Aahz Maruch, *Python for DUMMIES*, 1th ed, John Wiley & Sons, 2006.

[14] Giannakopoulos, T. (2015). pyAudioAnalysis: An Open-Source Python Library for Audio Signal Analysis. *PLoS ONE* 10(12): e0144610. doi:10.1371/journal.pone.0144610.

[15] Yang, Yujun., Li, Jianping., Yang, Yimei. (2015). The research of the fast SVM classifier method. *12th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP)*. 121-124. doi:10.1109/ICCWAMTIP.2015.7493959.

[16] Heryadi, Y. and Warnars, H.L.H.S. 2017. Learning temporal representation of transaction amount for fraudulent transaction recognition using CNN, Stacked LSTM, and CNN-LSTM, CyberneticsCom 2017, Phuket, Thailand, 20-22 Nov 2017.

[17] Heryadi, Y., Miranda,E. and Warnars, H.L.H.S. 2017. Learning decision rules from incomplete biochemical risk factor indicators to predict cardiovascular risk level for adult patients. Cybernetics and Computational Intelligence (CyberneticsCom), 2017 IEEE International Conference on, Phuket, Thailand, 20-22 Nov 2017.

[18] Ricky, M.Y., Warnars, H.L.H.S., Budiharto, W. and Abbas, B.S. 2017. **Machine Learning Approach to Task Ranking.** Pervasive Systems, Algorithms and Networks & 2017 11th International Conference on Frontier of Computer Science and Technology & 2017 Third International Symposium of Creative Computing (ISPAN-FCST-ISCC), 2017 14th International Symposium on, Exeter, UK, 21-23 June 2017.

[19] Munir, B.S., Trisetyarso, A., Gaol, F.L., Soewito, B. and Warnars, H.L.H.S. 2017. **Response Time Assessment of Hilbert Huang Transform and Prony Analysis Using Phasor Measurement Data.** Advanced Science Letters, 23(4), 3722-3724, April 2017.