Music Genre Classification using Naïve Bayes Algorithm

Ardiansyah*, Boy Yuliadi**, Riad Sahara***
Department of Information System, Faculty of Computer Science
UniversitasMercuBuana

Abstract

*Classification of music genres that are processed consist of, rock, pop, dangdut and jazz. The number of each genre for training data is 20 data for rock music, 20 data for pop music, 20 data for jazz music, and 17 data for dangdut music. While for data testing there are 10 data for rock music, 8 data for pop music, 10 data for jazz music, and 8 data for dangdut music. So that the total digital music that will be used is 113 digital music. Data taken in the form of * .mp3 data is changed into * .wav data. The classification method used is Naive Bayes. And feature features that are used are short time energy, zero crossing rate, spectral centroid and spectral flux. The results obtained from the overall testing data are correctly classified as 15 data from 36 data so that the success percentage is 41.67%. To classify the highest success is the highest classification of pop music by 70.5%, followed by dangdut music by 62.5%, then rock music by 30%. For jazz music the success of the classification is very small, which is only 10%.*

Keywords : *Music genre, Classification, Naïve Bayes.*

I. INTRODUCTION

The word music comes from the Greek language, namely mousike. Music is the science or art of composing tones or sounds in sequences, combinations, and temporal relationships to produce compositions (sounds) that have unity and continuity according to the Indonesian dictionary (http://kbbi.web.id/musik).

Rapid technological developments also affect the development of music where music is now stored in digital form. Digital music is the reproduction of sound from digital signals that have been converted into analog signals, digital voice recording by means of encoding binary numbers resulting from changes in analog sound signals with the help of sampling frequencies. Forms of digital signal storage in computer-based media. Digital format can store large amounts of data [12]. Where digital music is very much on the internet, so a lot of research is done to classify music genres. Music genre is a grouping of music according to their similarities to each other [12].

Classifying music based on genre is expected to make it easier for music lovers to enjoy music in accordance with the desired genre. To get a good automatic classification it is necessary to choose the right method [2] [4].

In this research using the Naive Bayes method for the stage of digital music classification. There are several studies that have carried out the classification using the Naive Bayes method as carried out by Kour et al, Rathore, and Wongso et al. [6] [9] [11]. And for the initial process of classification, a feature extraction is needed in order to distinguish each genre of music. Feature extraction in question is short time energy, zero crossing rate, spectral centroid and spectral flux such as feature extraction in the study used by Rathore [9].

Given the rapid development of music in Indonesia, especially digital-based music, it allows one to have a lot of digital music libraries. The music you want to enjoy will be adjusted to the mood or the situation of the digital music lovers themselves, so that when the digital music library is very large it is likely that someone will find it difficult to choose the desired music genre at the right time.

Look at the problems that exist so that this study classifies the digital music genre using the Naive Bayes method and feature extraction of short time energy, zero crossing rate, spectral centroid and spectral flux for the initial stage before classification. With a combination of naive bayes and feature extraction used, it is expected to produce good classification accuracy.

This research is limited only to classify the genres of rock, pop, dangdut, and jazz music by using feature extraction of short time energy, zero crossing rate, spectral centroid and spectral flux, and the classification method of naive bayes.

II. REVIEWLITERATURE

A. Digital Music Genre

Digital music is the reproduction of sound from digital signals that have been converted into analog signals, digital sound recording by encoding binary numbers resulting from changes in analog sound signals with the help of sampling frequencies. Whereas the

music genre is a grouping of music according to its resemblance to one another. Music can also be grouped according to other criteria, such as geography. A genre can be defined by musical techniques, style, context, and music themes [12]. To classify music by genre has also been done by several other researchers [8]. Below is the music genre that will be calcified in this study [5]:

1. Pop

Pop music is music with a simple rhythm that is easily known and liked by many people (general). Pop music is not only sung or carried by bands, but there are also solo singers who join in this pop genre.
Characteristics: melodies are easy to apply with various lyric characters, are flexible and easily combined with other types, songs are easy to hum and easy to understand, harmony is not complicated, tempos vary.

2. Rock

Rock music is a popular genre of music that became known generally in the mid 50s. Its roots come from rhythm and blues, country music from the 40s and 50s and various other influences. Furthermore, rock music also took a style from various other music, including folk music (folk music), jazz and classical music.
The distinctive sound of rock music often revolves around electric guitars or acoustic guitars, and the use of back beats is very noticeable in rhythm sections with bass and drum guitars, and keyboards such as organs, pianos or since 70s, synthesizers. Besides guitars or keyboards, saxophone and blues-style harmonica are sometimes used as solo musical instruments. In its pure form, rock music "has three chords, consistent and striking bakcbeat and attractive melody".
In the late 60's and early 70's, musk rock developed into several types. Mixed with folk music (regional music in america) became folk rock, with blues being blues-rock and with jazz, being jazz-rock fusion. In the 70s, rock combined the effects of soul, funk, and latin music.

3. Jazz

Jazz music is a type of music that grew out of blues, ragtime and european music, especially band music. In jazz music, the usual instruments are guitar, trombone, piano, trumpet and saxophone. Tone of jazz music has a unique characteristic in the game, sometimes not everyone can enjoy jazz music.
Characteristics of jazz music: 1) vocals and lyrics tend to be considered as part of the sound of the instrument, 2) the harmony is complicated, has a broad tonality and frequent modulation, 3) rhythm and melody have an improvised tendency

4. Dangdut

Dangdut is one of the emerging music arts genres in indonesia. This form of music is rooted in malay music in the 1940s. In the evolution towards contemporary forms now comes the influence of elements of indian music (especially from the use of tabla) and arabic (on crooked and harmonized). Changes in the flow of indonesian politics in the late 1960s opened the entry of strong western music influences with the inclusion of the use of electric guitars and also forms of marketing. Since the 1970s dangdut can be said to have matured in its contemporary form. As popular music, dangdut is very open to the influence of other forms of music, ranging from keroncong, langgam, degung, gambus, rock, pop, even house music. The mention of the name "dangdut" is an onomatopoeia of the sound of the tabla game (in the world of dangdut called drums only) which is typical and dominated by dang and ndut sounds. This name is actually a cynical designation in an early 1970s magazine article for a form of malay music that was very popular among the working class community at that time.

B. Feature extraction

Before being processed into a classification algorithm, the music data must be found for the extraction characteristics to be processed [3] [7]. There are four types of feature extraction used in this study, namely short time energy, zero crossing rate, spectral centroid, and spectral flux with standard deviation statistical parameters. Research using variant parameters has been carried out by al shoshan [1]. Al shoshan et al said that the combination of feature extraction from zero crossing rate (zcr) and short time energy (ste) would produce a good estimate [1].
The description of the four feature extractions is explained below:

1. Short Time Energy (STE)

Indicates the shortness of sound

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

Information: STE = Sort time energy
N = Number of Samples
X (n) = The value of the signal from the sample

2. Zero Crossing Rate (ZCR)

This feature measures the rate of change of a signal in a domain where the successive sample of a digital signal has a size difference.

$$Z_n = \frac{1}{2N}\sum_{m=n-N+1}^{N} | sgn[x(m)] - sgn[x(m-1)] |$$

Information :

ZC = *Zero crossing rate*

sgn x (n) = the value of x (n), is 1 if x (n) is positive, -1 if x (n) is negativeN = JumlahSampel

3. *Spectral Centroid*

Balancing the spectrum size point from spectral association with spectral brightness. High centroid values indicate high frequency.

$$C = \frac{\sum_{n=1}^{N} Mt(n) \cdot n}{\sum_{n=1}^{N} Mt(n)}$$

Information:

C = *Spectral Centroid*

Mt = value of*Centroid*

4. *Spectral Flux*

Measuring changes in spectral peaks in the signal. This is obtained from the calculation of large spectrum changes from layer to layer.

$$F = \sum_{n=1}^{N} (Nt(n) - Nt_{-1}(n))^2$$

Information :

F = *Spectral Flux*

Nt = Value of change in spectrum

C. Classification

To classify the genre of digital music data this research uses the Naive Bayes algorithm. The Naive Bayes algorithm is a machine learning algorithm that utilizes probability methods. With Bayes Theorem theory is a theorem used in statistics to calculate opportunities for a hypothesis, Bayes Optimal Classifier calculates the probability of a class from each group of existing attributes, and determines which class is the most optimal [2]

Wongso uses the SVM classification to produce an accuracy rate of 88.33%. Mauludiya et al produced an accuracy of 93% using kernel polynomials [5]. Kour et al. Used Back Propagation Neural Network (BPNN) and Support vector machine (SVM) to classify their music genres by using feature extraction of MFCCs [6]. Whereas Zhou et al classifies emotion queries based on SVM [13]

III. RESEARCH METHODOLOGY

The work order for the research to be carried out can be described in the flowchart below:

Figure1. Research Method

Below is an explanation of the steps taken in the research method.

A. Data

Digital music data used for this research is file type * .mp3 with digital rock, pop, dangdut, and jazz types. The number of each genre for training data is 20 digital music and testing data as many as 10 digital music. So that the total digital music data for all genres is 120 data. Then the collected digital music data is converted into the * .wav file to be processed. Some studies have also changed the data file format into * .wav has been done by other researchers such as Edwards and Liu et al (Edwards 2006; Liu et al 1997). To make it easier and so that the data that is processed is not too large digital music data is then cut randomly, the length of the music used is 30 seconds, this has also been done in other studies such as those conducted by Rathore [2].

B. Feature extraction

Data that has been in the form of * .wav is processed into the MATLAB program to find feature extraction values for each music data consisting of the value of STE, ZCR, spectral centroid and spectral flux. The statistical parameters used for all four feature

extractions are standard deviations. The data is saved in the * .txt file format.

C. Classification

Using the Naïve Bayes function on MATLAB to begin classifying the data that has been searched for the extraction characteristics.

D. Evaluation

Evaluation techniques are used in a manual way where each type of song tested is known.

E. Research data

The total number of data collected is 113 music which is divided into 25 dangdut music, 30 rock music, 30 jazz music and 28 pop music. Then each type is divided into two groups where 17 dangdut music data for training data, 20 rock music data, 20 jazz music data, and 20 pop music data for training data while for testing data each 8 dangdut music data for training data, 10 rock music data, 10 jazz music data, and 8 pop music data.

To be processed in the matlab application, music data is converted into * .wav (mono, 16 bit, 6000Hz) files by using the audacity apalication.

IV. RESULTS AND DISCUSSION

A. Feature Extraction

Extraction Characteristics taken on each type of song, both dangdut and keroncong songs consist of four features, namely Centroid, flux, STE, and ZCR. Training data and testing data are taken and processed separately to facilitate the classification process.

From the collected data, feature extraction is sought from each data both training data and testing data using the MATLAB program. The program used is available to calculate the required feature extraction (MIRTOOLBOX) (Lartillot and Toiviainen, 2007).Here is the work area in matlab, which is shown in the image below.

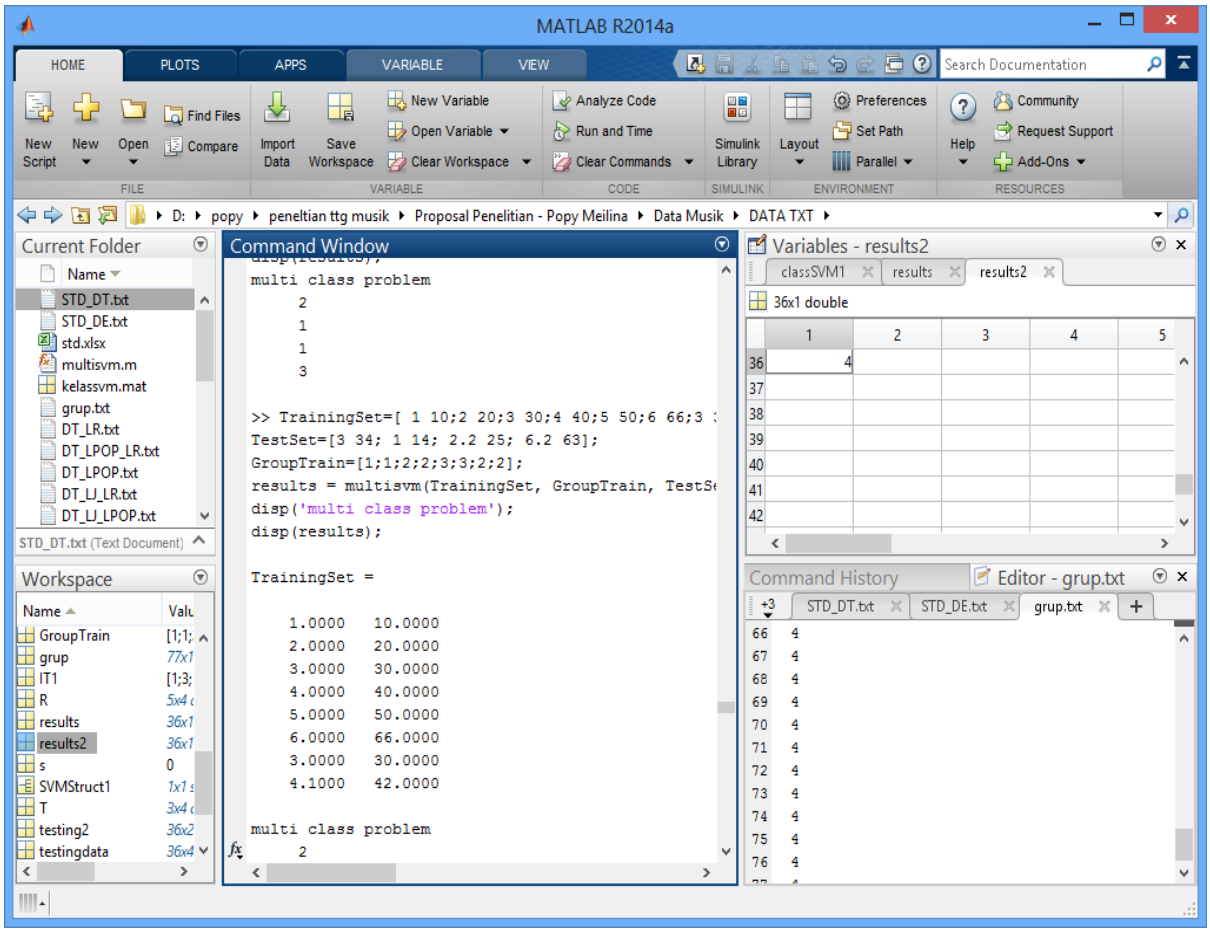

Figure2. Window Matlab

Below is shown feature extraction results for training data for dangdut songs.

Table 1. Characteristics of Dangdut Song Training Data

Centroid	Flux	STE	ZCR
37,75589	22,29433	0,021453	52,25027
86,20543	38,75021	0,103616	51,9616
58,53289	42,61633	0,029097	51,9616
67,61961	39,141	0,026049	51,9616
58,42801	13,86038	0,027783	52,25027
52,09621	22,156	0,035693	52,53895
43,27946	16,41737	0,028226	51,9616
57,13694	21,04631	0,023788	52,25027
47,24117	21,63703	0,038512	51,9616
73,95551	29,06938	0,021114	51,9616
65,3965	16,10937	0,039096	52,25027
82,70056	23,00693	0,0318	51,9616
62,18767	21,82888	0,019869	51,67292
74,42641	41,14114	0,043848	52,25027
100,5862	49,17387	0,0617	51,9616
50,61575	18,57183	0,032771	52,25027
67,2324	39,69604	0,049585	51,9616
79,2071	26,35192	0,127869	62,50661
93,8583	38,03233	0,02291	77,99562
66,7316	16,09445	0,068452	84,98834
53,0073	13,63289	0,053173	73,9763
80,0494	22,44046	0,117337	67,02361
48,9237	19,89586	0,054686	83,16868
80,8851	19,61568	0,054533	74,57203
65,7566	21,19969	0,114202	62,38594
84,0580	33,09778	0,134287	73,3504
60,4507	18,24119	0,105909	70,59043
48,9935	23,83064	0,107241	71,94027
68,5851	32,44899	0,125109	62,46889
67,2693	29,64271	0,11132	90,91316
65,8022	25,18907	0,1481	74,04419
86,4693	17,16614	0,101351	64,36919
61,8209	16,74221	0,11918	81,70572
95,9013	23,47348	0,111994	76,41204
57,1010	6,463295	0,094177	73,57663
87,0303	26,09613	0,137346	75,65793
67,1825	22,6468	0,095697	75,59009

Below is a feature extraction from 20 songs song training data.

Table 2. Extracts of Characteristics of POP Song Training Data

Centroid	Flux	STE	ZCR

Below is a feature extraction from the Jazz song training data of 20 songs.

Table 3. Extracts of Characteristics of Training Data Songs Jazz

Centroid	Flux	STE	ZCR

54,86481	26,82874	0,041836	93,69579	51,31667	33,51002	0,185508	54,06835
90,50009	19,45286	0,109409	74,02157	55,69606	24,53724	0,118621	86,18504
88,59724	21,89228	0,120981	80,78576	54,40701	25,5876	0,165091	71,88749
85,97048	29,77939	0,094883	87,15783	67,14586	20,39069	0,085131	73,36551
38,58186	19,5387	0,054408	63,17773	71,0394	44,15827	0,151571	55,02603
62,15613	13,22913	0,098342	81,80377	76,59948	41,6568	0,142936	73,52385
74,5683	22,29291	0,114127	74,17991	60,69182	20,75839	0,093559	65,07803
59,66935	20,70205	0,156697	64,13546	67,6101	11,1243	0,094958	55,38798
75,65187	21,37767	0,139516	65,50036	70,24769	14,84224	0,136622	53,97029
63,32895	28,58694	0,113182	80,86115	98,49528	36,93273	0,157957	63,46429
55,87575	21,54659	0,052783	88,91485	66,81753	33,27031	0,108797	77,36217
107,3836	29,32196	0,12501	53,9854	48,06873	15,55408	0,156881	76,05005
70,61725	15,2114	0,138058	80,15231	40,83353	15,58296	0,180638	60,74204
67,05919	18,08189	0,082275	67,92095	77,55645	21,12742	0,17421	73,48618
63,69315	28,97139	0,107434	302,6466	57,92812	34,2516	0,132211	64,61052
67,85026	22,62328	0,015129	200,3618	69,66506	24,49539	0,140008	55,14669
80,85742	19,72594	0,054533	74,57203	36,54545	22,90749	0,096327	26,34043
44,42422	13,9956	0,052488	60,61492	76,56107	31,00619	0,115597	54,85258
62,06137	37,16997	0,111125	71,31438	61,92323	16,22797	0,148942	60,82498
65,75988	21,19972	0,114202	62,38594	52,25087	38,55024	0,07482	27,51683

Below is a feature extraction from the training data of Rock songs as many as 20 songs.

Table 4. Extracts of Characteristics of Rock Song Training Data

Centroid	Flux	STE	ZCR

Below are 10 feature extraction of Testing data for rock songs.

Table 5. Extraction Value Characteristics of Rock Song Testing data

Centroid	Flux	STE	ZCR

63,86619	25,31839	0,194078	64,3843
38,75022	19,87472	0,181429	61,17943
39,45355	27,14284	0,166131	71,15604
75,68176	33,75777	0,043038	66,27704
48,01347	18,53927	0,133616	72,51339
71,17875	17,22902	0,121347	63,35118
56,92944	44,63081	0,127273	59,95025
53,95012	22,14747	0,139214	66,67672
57,0714	44,29743	0,118419	59,9578
72,80039	27,45611	0,150459	47,89234

Testing data for pop songs is shown in the table below.

Table 6. Extraction Value Characteristics of Pop Song Testing data

Centroid	Flux	STE	ZCR
85,68622	40,15112	0,146122	59,30925
67,13596	17,83954	0,092694	61,38299
77,3154	24,80331	0,113383	70,94487
78,34461	22,08608	0,124784	66,71444
67,05784	18,08188	0,082275	67,92095
72,46888	19,19087	0,165258	71,81205
59,11752	23,79067	0,143116	67,36294
79,12612	23,15739	0,114297	70,94487

Testing data for jazz songs is shown in the table below.

Table 7. Value of Extraction Characteristics of Jazz Song Testing data

Centroid	Flux	STE	ZCR
59,04277	21,49759	0,072813	62,71773
61,00053	40,36189	0,014342	88,8168
52,73359	40,61241	0,043533	63,69807
110,3567	26,09383	0,136809	64,67086
69,65207	14,61282	0,097507	80,71787
65,99527	21,95137	0,047693	80,77065
78,09049	35,93171	0,0508	62,45383
78,13003	35,93653	0,0508	62,45383
81,10214	31,4309	0,065105	77,94284
59,43071	16,12231	0,138411	68,85602

Testing data for jazz songs is shown in the table below.

Table 8. Extraction Value Characteristics of Dangdut Song Testing data

Centroid	Flux	STE	ZCR
61,21025	25,98564	0,043986	51,9616
68,97169	24,88077	0,04955	52,25027
52,48988	23,33542	0,026654	52,53895
81,27569	29,42863	0,024409	51,9616
73,55514	26,13268	0,078284	51,9616
78,69891	41,39643	0,024731	52,25027
56,88917	24,04786	0,024556	51,9616
66,30637	30,01247	0,022744	52,25027

B. Classification

Training data and testing data from songs are stored in the form of * .txt files to be processed in the matlab application.Before classifying the feature extraction data from training data must be entered into a matrix. The commands in MATLAB that can be used and exemplified below.

```
>>training=dlmread('DATA_TRAINING.txt');
>>testingdata=dlmread('DATA_TESTING.txt');
>>label=fitNaiveBayes(training, grup_predik)
>>grup =coba2.predict(testingdata);
```

>>grup=[1;4;4;2;1;2;4;4;4;1;2;1;2;2;2;2;1;2;4;2;4;3;2;
2;2;2;2;1;4;4;4;2;1;1;4;4]

Results from training and testing data that are processed against rock, pop, jazz and dangdut songs can be seen in the table below.

Table 10. Results of classification with naive bayes

Rock	POP	Jazz	Dangdut
1	2	4	4
4	1	2	4
4	2	4	4
2	2	3	2
1	2	2	1
2	2	2	1
4	1	2	4
4	2	2	4
4		2	
1		1	

Classification using naive bayes shows that there are three correctly classified rock music, there are six correctly classified pop music, there are 1 correctly classified jazz music and five correctly classified dangdut music.

C. Evaluation

From 113 music data that has been extracted, the characteristics are then classified with the Naive Bayes method. Where the data is divided into two groups of data namely training data and testing data. Based on the classification error of the testing data totaling 36 data, shown in table 11 below.

Table 11. Results data

		Results				Total
TestingData	Class	Rock	POP	Jazz	Dangdut	
	Rock	3	2	0	5	10
	POP	2	6	0	0	8
	Jazz	1	6	1	2	10
	Dangd ut	2	1	0	5	8
Total of Testing Data						36

Based on the results table above the overall testing data is classified correctly there are 15 data from 36 data so that the success percentage is 41.67%. To classify the highest success is the highest classification of pop music by 70.5%, followed by dangdut music by 62.5%, then rock music by 30%. For jazz music the success of the classification is very small, which is only 10%.

REFERENCES

[1] Al-Shoshan, Abdullah I. Speech and Music Classification and Separation: A Revie. J. King Saud Univ., Vol. 19, Eng. Sci. (1), pp. 95-133. Riyadh, Saudi. 2006.

[2] Dewi, Andi Imrah., Hidayat, Andi Nurul. Analisis Music Mining Information Retrieval Untuk Klasifikasi Jenis Music Bergenre Menggunakan Algoritma Naive Bayes. Jurnal Elektronik Sistem Informasi dan Komputer. Vol 1 No 2. Juli-Desember 2015. p. ISSN: 2777-888 **e.** ISSN**:** 2502-2148.

[3] Deshpande, Hrishikesh., Singh, Rohit., Nam, Unjung. Classification Of Music Signals In The Visual Domain. Proceedings of the COST G-6 Conference on Digital Audio Effects (DAFX-01). Limerick, Ireland. December 6-8,2001.

[4] Genre Musik. (accessed on July 25th, 2018)

[5] https://klinikmusik.wordpress.com/2014/10/19/genre-musik/http://dompet-inspirasi.blogspot.co.id/2013/11/macam-macam-genre-musik-dan-pengertianya.html

[6] Kour, Gursimran., Mehan, Neha. Music Genre Classification using MFCC, SVM and BPNN. International Journal of Computer Applications (0975 – 8887), Volume 112 No. 6, February 2015. www.ijcaonline.org.

[7] Lartillot, Olivier danToiviainen, Petri. A Matlab Toolbox For Musical Feature Extraction From Audio. Proc. of the 10th Int. Conference on Digital Audio Effects (DAFx-07). Bordeaux, France. September 10-15, 2007.

[8] Li, Tao.,Ogihara, Mitsunori., Li, Qi. A Comparative Study on Content-Based Music Genre Classification. Proceeding of SIGIR'03.Toronto, Canada. 28 Juli – 1 Agustus, 2003.

[9] Rathore, Archit. Music Genre Classification. (accessed on July 30, 2018) http://home.iitk.ac.in/~margaux/cs365/project/report.pdf

[10] Mauludiya, Rosyita Ayuning., Magdalena, Rita.,Ramatryana, I Nyoman Apraz. Simulation And Analysis Of Music Genre Classification Based On Fft And Support Vector Machine. Teknik Telekomunikasi, Fakultas Teknik Elektro, Universitas Telkom.

[11] Wongso, Rini., Santika, Diaz D. Automatic Music Genre Classificaation Using Dual Tree Complex Wavelet Transform and Support Vector Machine. Journal of Theoretical and Applied Information Technology, Volume 63 No 1, 10 Mei 2014.

[12] Wikipedia (accessed on July 30, 2018) https://id.wikipedia.org/wiki/Audio_digital https://id.wikipedia.org/wiki/Genre_musik

[13] Zhou, Lijuan., Lin, Hongfei., Liu, Wenfei. Enriching Music Information Retrieval Using Emotion Detection. SIGIR 2011 Workshop on Enriching Information Retrieval (ENIR 2011), Beijing, China. July 28, 2011.