

## Music Emotion Classification based on Lyrics-Audio using Corpus based Emotion

Fika Hastarita Rachman<sup>1</sup>, Riyanarto Sarno<sup>2</sup>, Chastine Fatichah<sup>3</sup>

<sup>1,2,3</sup>Department of Informatics, Institut Teknologi Sepuluh Nopember, Indonesia

<sup>1</sup>Department of Informatics, University of Trunojoyo Madura, Indonesia

---

### Article Info

#### Article history:

Received Jan 10, 2018

Revised Apr 2, 2018

Accepted Apr 11, 2018

#### Keyword:

Audio features

CBE

Corpus based emotion

Emotion

Lyric features

Music Classification

---

### ABSTRACT

Music has lyrics and audio. That's components can be a feature for music emotion classification. Lyric features were extracted from text data and audio features were extracted from audio signal data. In the classification of emotions, emotion corpus is required for lyrical feature extraction. Corpus Based Emotion (CBE) succeed to increase the value of F-Measure for emotion classification on text documents. The music document has an unstructured format compared with the article text document. So it requires good preprocessing and conversion process before classification process. We used MIREX Dataset for this research. Psycholinguistic and stylistic features were used as lyrics features. Psycholinguistic feature was a feature that related to the category of emotion. In this research, CBE used to support the extraction process of psycholinguistic feature. Stylistic features related with usage of unique words in the lyrics, e.g. 'ooh', 'ah', 'yeah', etc. Energy, temporal and spectrum features were extracted for audio features. The best test result for music emotion classification was the application of Random Forest methods for lyrics and audio features. The value of F-measure was 56.8%.

Copyright © 2018 Institute of Advanced Engineering and Science.  
All rights reserved.

---

### Corresponding Author:

Fika Hastarita Rachman,

Department of Informatics,

Institut Teknologi Sepuluh Nopember,

Jl. Teknik Kimia, Gedung Teknik Informatika, Kampus ITS Sukolilo Surabaya 60111, Indonesia.

Email: fika14@mhs.if.its.ac.id

---

## 1. INTRODUCTION

Music emotion represented in 2 models: categorical and dimensional models. Categorical and dimensional models are actually interrelated. Each model has advantages and disadvantages. Categorical model use human language for category label, so it's easy for user to understand it. The dimensional model is an emotional model that describes emotions in a dimensional vector space. The Corpus that used for the Categorical model reference is WordNet Affect of Emotion (WNA) [1]. WNA is a development of WordNet that have an emotional labels based on Ekman Emotion [2]. There are 6 categories of emotions: sadness, anger, joy, disgust, fear, and surprise. While Affective Norms for English Words (ANEW) is the dataset that is often used for research on dimensional model [3]. Each term in ANEW contained 3 dimension values. There are Valence, Arousal, and Dominance. Valence and Arousal are seen from each personal. Valence same as level of pleasure. In vector space, the range of valence value from negative into positif and arousal has value low until high. Different with Valence and Arousal, Dominance is relation between people and their environment. Although any research focus to combination 2 model [3], but more research using only one emotional model (i.e. [4]-[7]).

Music has lyrics and audio features that can be used for a reference in music classification. The lyrics of music are more dominant to Valence than Arousal. And audio of music more represent to Arousal dimension. The previous research that discusses the extraction of emotions from audio [8]-[10] have

conclusion that audio can be a feature in music emotion classification. Similarly with the previous research using lyrics extraction [7]-[10], it proves that lyrics are important too for feature in music emotion classification.

Two main features of music, audio and lyrics, was used in research [11]-[13]. But in the process, they didn't use a combination of dimensional and categorical models. The purpose of this research is to prove the linkage of categorical models and dimensional models for classification of emotions in music. We use the combined emotional corpus between the dimensional model and the category model for extraction of lyric feature. This is our contributed for this research.

Lyrics become an important part in the detection of emotions. One of the lyrics features is psycholinguistic feature. these features can be presented differently depending on model of emotion and type of corpus used. Vipin Kumar extract psycholinguistic features lyrics from Sentiwordnet [14]. Sentiwordnet is a corpus that has a positive-negative score [15]. Using sentiwordnet, the lyrics analyzed positive-negative of sentiment value not emotional value. Dimensional models can also affect the value of psycholinguistic features. With an ANEW emotion corpus, this feature can be worth the value of the valance and arousal dimensions [16]. Our research use CBE for extracting psycholinguistic feature of emotion from lyric. Corpus Based Emotion (CBE) applies the combined concept of categorical and dimensional datasets. Not only combining, but also expand corpus using similarity word and euclidean distance concepts. General Inquirer (GI) and Wordnet datasets are used to support the success of this research too.

Audio signal can be from speech or not. The speech feature is taken from the human voice without the instrument. Speech feature can also be classified into emotion [17]. But, for this music document, the audio features to be used are non-speech-shaped signals or wav signals. Audio can be extracted into Standard Audio and Melodic Audio features. Using the application of toolbox, extracted audio features can reach more than a hundred. The application of the relief algorithm and PCA (Principle Component Analysis) is used for reducing dimension and selection of audio features, so it is known which features are more important to use [4]. Van Loi Nguyen [18] divided audio features into two subsets of dimension: Arousal and Valence. The subsets is used for emotional classification with the dimensional model and convert it into categorical using Thayer's model. 9 kinds of spectral shape audio extraction results can also be used as a feature of music emotion classification [19]. Roughness feature in audio spectrum is can be meant as spectral flux. And continued implementation of emotional clasification can lead to application in the music recommendation system [20]. Our research will be tested using standard audio features obtained from the Toolbox: MIR Toolbox and Psysound.

In categorical music of emotion, there have been previous research using audio and lyric features [12], [13]. But the two of research did not use the emotional corpus in its feature extraction process. One uses the Jlyrics framework to obtain statistical features [12]. And others see the word sparcity that appears in the lyrics [13].

The difference of this research with previous research is on the lyrics and audio features. Lyrics features that extracted is a combination of psycholinguistic and stylistic features. While the audio feature used was taken from MIR Toolbox and Psysound. Previous research using audio and lyrics features with categorical model approach. This research will be combined lyrics and audio features with the approach of the two models of emotion, categorical and Dimensional.

## 2. RESEARCH METHOD

Emotion detection process to be performed in this research include multimodal features. The music features extracted from lyrics and audio components. Formal sentence structure is not owned by lyrics. Lyrics has a small of words with limited vocabulary. In the lyrics there is a phrase or ideom that makes it difficult to know the true meaning. It is a challenge to be able to express the emotion of music based on lyrics. There are several features that can be extracted: psycholinguistic and stylistic features of text. Psycholinguistic features are psychological of language features in the lyrics. This feature can be found with the help of emotional corpus: GI and CBE. Stylistic features of lyrics are interjection words (e.g., "ooh," "ah") and special punctuations (e.g., "!", " "?"). Audio features extracted using toolbox Psysound 3 and MIRToolbox. Figure 1 is a proposed model in this research. The Features include feature of energy and features of spectrum. We used 2 main feature, because energy and spectrum of audio always successful for detection emotion [16], [13]. Dynamic loudness represent from energy feature. Roughness and inharmonicity represent spectrum feature. The result of feature extraction are used in music emotion classification.

The dataset used for the extraction of psycholinguistic features in the lyrics are ANEW, WNA and GI.ANEW dataset is often used as a reference in emotional detection for dimensional models. While WNA is a reference dataset used in emotional detection for categorical model. In ANEW there is a big of word with emotion label. Six basic emotions of Ekman are used as its emotional labels. Data distribution after the merge

process between ANEW and WNA is from 1030 terms in ANEW and 1197 of WNA. There are 105 terms that have Valance-Arousal-Dominance value and emotional label (Figure 2). Corpus Based Emotional (CBE) represents 2 model, categorical and dimensional models.

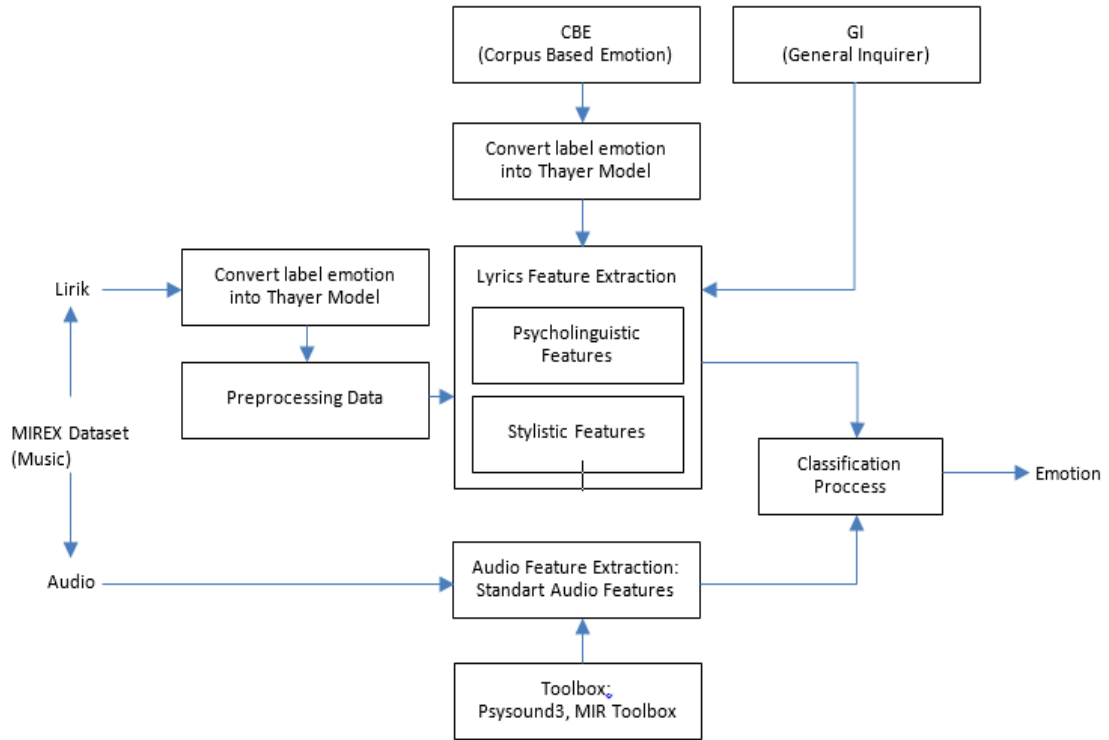


Figure 1. Proposed model

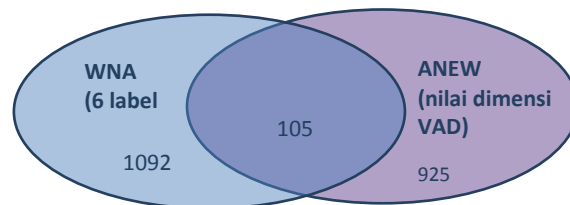


Figure 2. Data distribution in ANEW and WNA datasets

In this research, data result or prediction used categorical model only. But to obtain psycholinguistic features still refers to the CBE which represents categorical and dimensional models. CBE is a combined corpus between ANEW and WNA which has procedures of automatic incomplete data and CBE expand [21]. Automated data procedures for incomplete data are used for data that has no label or dimension values. The concept of Merging WNA and ANEW, causing incomplete data. In [21], ISEAR Dataset used for expand CBE. Figure 3 show the CBE scheme [21].

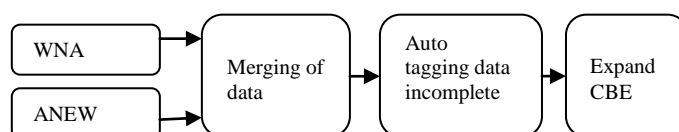


Figure 3. CBE scheme [21]

It can be seen in Figure 2, there are 2017 data is incomplete. Even for new data, it is certain that the term has no label or dimension value. Algorithm for Autotagging of incomplete data was created to handle this case. This procedure is made with the concept of synonym of word (synset), relatedness measure and Euclidean Distance. The method of relatedness measure used Adapted LESK [22]. We have test the Pearson Correlation value of Adapted LESK and Euclidean Distance, the value is -0.92. It means there is a relationship between Adapted LESK and Euclidean Distance. The T-test value between Adapted LESK with Euclidean Distance in terms labeled 'joy' is 6,6043, hence there can be correlation between that variable. CBE expand algorithm is also added to expand the corpus. The data testing is data that not contained in the previous CBE. Thus, CBE is corpus of emotional term that has a VAD dimension value and an emotion label.

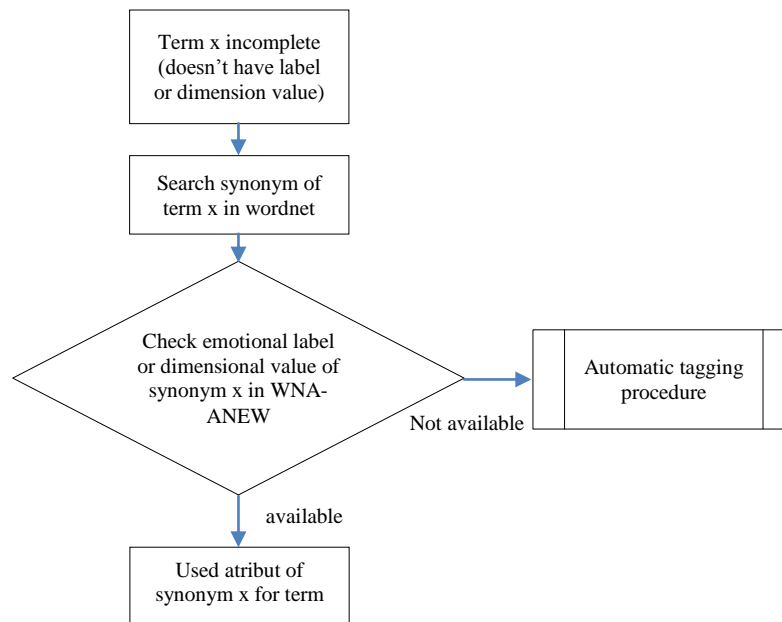


Figure 4. Filtering term using synset in WordNet

Before the autagging of incomplete data procedure is executed, it is necessary to check the synonyms of term in WordNet (Figure 4). If a synonym of term is found in the previous CBE, automatically the term has a same label value. Automatic tagging procedure handling for cases that do not have labels but have VAD values and want instead.

In previous CBE [20], POS Tagging and word filtering in the Emotion category on GI has not been used. It cause the automatic tagging algorithm has not produced the optimal output. The determination of the emotional cluster center before the Automatic Tagging algorithm is still based on the researcher's assumptions, so there are still not maximal results. In this research, we improvement CBE with used POS Tagging, GI filtering and cluster center determination. As well as the concept of K-Nearest neighbors (KNN), this research uses the close nodes to the model. The difference is K- nearest is used for classification [23], while this research is used to look for the score of Valance-Arousal-Dominance (VAD). Cluster center determination is based on VAD average value in every emotion label. With that improvement, CBE is expected to do better and produce more accurate output result.

Step of autotagging procedures of incomplete data are:

- 1) Define the center of cluster for each label of emotion 'Joy', 'Sad', 'Anger', 'Disgust', 'Fear', 'Surprise'. The center of the cluster is taken from the closest term to the average term data of each cluster. The center of cluster certainly has VAD value and emotion label.

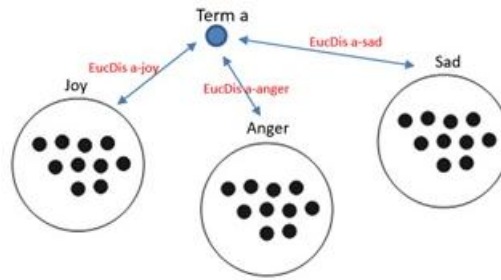


Figure 5. Illustration position of terms according to VAD value

- 2) For term a that has no emotion label:
  - a. If ‘term a’ does not have VAD, check the Adapted LESK value between term x with each center of cluster. Adapted LESK value is the value of the proximity between the two terms. The highest value describe the closest term. The emotion label between two terms is considered same. The label of term a equals with the emotion label on its term cluster.
  - b. If ‘term a’ has VAD value, euclidean distance value becomes reference to find the nearest term. euclidean distance value becomes reference to find the nearest term. Seen in Figure 5, 'EucDis a-sad' is the Euclidean distance between term a with the center of the sad cluster. And 'EucDis a-joy' is the Euclidean distance between term a with the center of the joy cluster. Formula (1) is a formula to find the value of euclidean distance between ‘term a’ and term center of cluster (‘term pc’). The smallest euclidean value represents the proximity between term a and the term cluster. So the term center of cluster labeled is considered equal to term a.

$$EucDis(pc; a) = \sqrt{(Vpc - Va)^2 + (Apc - Aa)^2 + (Dpc - Da)^2} \tag{1}$$

We used *Music* Information Retrieval Evaluation eXchange (MIREX) dataset for Music document. MIREX [24] is a music dataset for Mood Classification Task in International Society for *Music* Information Retrieval (ISMIR) conferences. This model classifies emotions into five distinct groups or clusters (Table 1), each cluster comprising five to seven related emotions (adjectives). There are 903 data in 30-seconds of audio. Its divided into 5 mood clusters. Each cluster has balances number of data (170, 164, 215, 191, 163 excerpts). Of 903 audio data, 764 has audio and lyrics. But because there is a process of converting data into Thayer model, then the test data used is 456 data.

Table 1. Five cluster in MIREX Dataset

Clusters	Mood adjectives
Cluster 1	Passionate, rousing, confident, boisterous, rowdy
Cluster 2	Rollicking, cheerful, fun, sweet, amiable/good natured
Cluster 3	Literate, poignant, wistful, bittersweet, auntumnal, brooding
Cluster 4	Humorous, silly, campy, quirky, whimsical, witty, wry
Cluster 5	Aggressive, fiery, tense/anxious, intense, volatile, visceral

Emotional label of MIREX derived from Russel Model, so it uses 5 emotion cluster labels. CBE have 6 emotional label derived from six Ekman basic emotion. In order to cooperate to support this research, it takes conversion process to get the uniform data for emotional label. This research uses Thayers model for the uniformity of label data, because it has 4 class emotions with clear limits on dimension spaces [25]. From the data conversion process generated 4 class emotions, namely: class 1, class2, class 3, and class 4.

Figure 6 is the mapping of MIREX label to Thayer model. Of the image is clearly visible that ‘Cluster 5’ on the MIREX will be converted to ‘Class 2’ on Thayer, ‘Cluster 2’ will be converted to ‘Class 1’ on Thayer, and ‘Cluster 3’ will be converted to ‘Class 3’ Thayer. In dimension spaces, ‘Cluster 1’ and ‘Cluster 4’ of MIREX are located in the slice area between 2 emotional classes of Thayer. ‘Cluster 1’ MIREX is located in the area of ‘Class 1’ and ‘Class 2’. Whereas ‘Cluster 4’ MIREX is located in the area of ‘Class 1’ and ‘Class 4’. In this research there is no handling of it, so only music data on MIREX with ‘Cluster 2’, ‘Cluster 3’, and ‘Cluster 5’ that used.

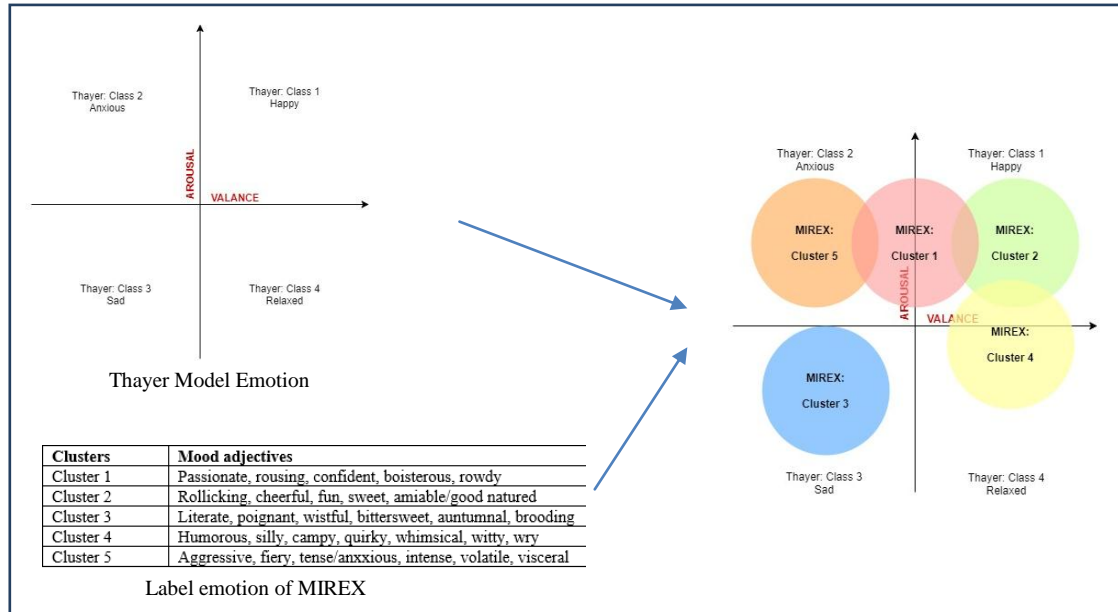


Figure 6. Conversion mapping label of emotion MIREX to Thayer model

The conversion mapping of the emotional label from CBE labels to Thayer labels is shown in Figure 7. In the Figure, it appears that the 'Happy' and 'Surprise' labels are converted to the 'Class 1' in Thayer class. 'Disgust', 'Fear', and 'Anger' labels converted on 'Class 2' Thayer. And 'Sadness' label are converted to 'Class 3' in Thayer.

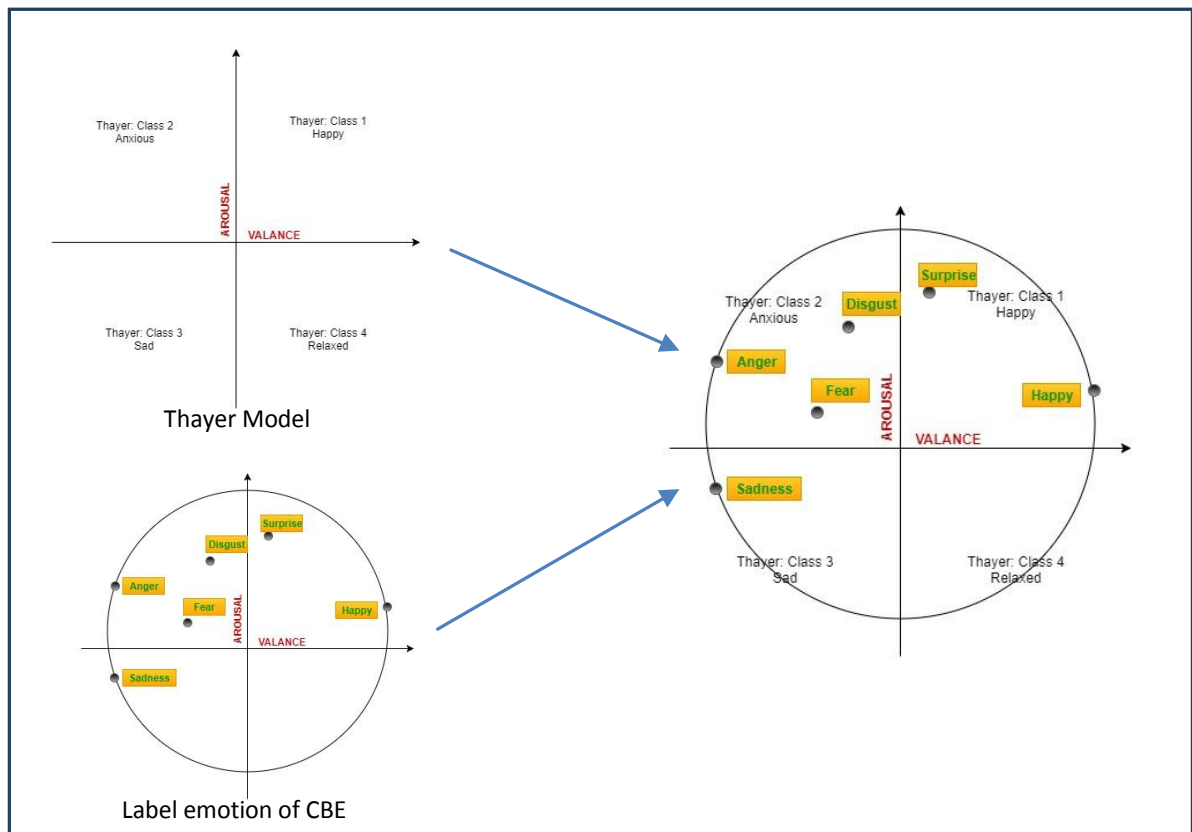


Figure 7. Conversion mapping label emotion of CBE to Thayer model

Figure 9 is the data flow process from the convert emotional label process, preprocessing until lyrics feature extraction. Before the lyrics feature extraction process, there is a preprocessing data. In the lyrics there are many informal words, such as 'n' which means 'and', 'll' which means 'will', 'em' which means 'them', and others. It is necessary for repair data to improve the word into a formal word structure form. Before the repair data process, there is a checking position of words using POS Tagging Stanford. This is important because there are parts of sentences that will not be processed (preposition, article, possessive pronoun, etc.). Parts of words that are not included in the repair process will be filtered from the data.

In accordance with [26], an audio feature that affects the music emotion recognition as shown in Table 2. In this study, audio features were extracted using the Pysound3 and MIR Toolbox. Features include feature of energy (dynamic loudness), feature of temporal (tempo) and features of spectrum (roughness and inharmonicity).

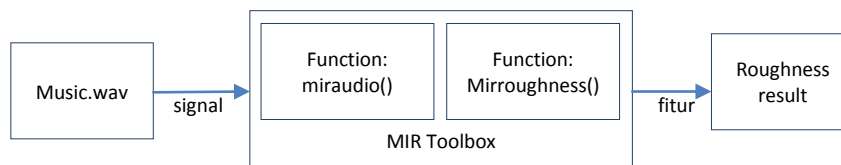


Figure 8. Flowprocess for extraction feature roughness

The process of feature extraction is shown in Figure 9. That process used mirtoolbox function. For roughness feature, the functions miraudio () and mirroughness () are used. The extraction flow feature for other features uses a same path like Figure 9 with the use of different functions on mirtoolbox or pysound 3.

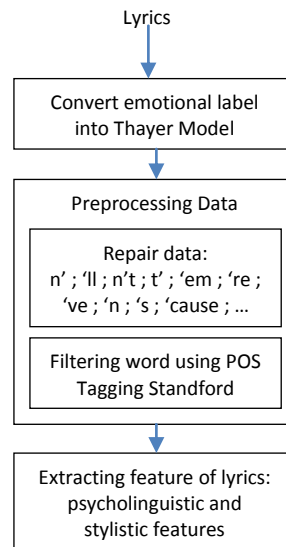


Figure 9. Process of lyrics

In the classification process, lyrics and audio features are tested using three classification methods: the Support Vector Machine (SVM) method, the Random Forest method and the Naive Bayes method.

Table 2. Audio Feature for Music Emotion Recognition [26]

Feature set	Extractor	Features
Energy	Psysound SDT	Dynamic loadness Audio power, total loudness, and specific loudness sensation coefficients
Rhythm	Marsyas MA toolbox, RP extractor MIR toolbox	Beat histogram Rhythm pattern, rhythm histogram, and tempo Rhythm strength, rhythm regularity, rhythm clarity, average onset frequency, and average tempo
Temporal Spectrum	SDT Marsyas, SDT  MA toolbox, Marsyas, SDT MATLAB	Zero-crossings, temporal centroid, and log attack time Spectral centroid, spectral rolloff, spectral flux, spectral flatness measures, and spectral crest factors Mel-frequency cepstral coefficients Spectral contrast, Daubechies wavelets coefficient histogram, tristimulus, even-harm, and odd-harm.
Harmony	MIR toolbox MIR toolbox  Marsyas PsySound	Roughness, irregularity, and inharmonicity Salient pitch, chromagram centroid, key clarity, musical mode, and harmonic change Pitch histogram Sawtooth waveform inspired pitch estimate

**2. RESULTS AND ANALYSIS**

The dataset used as test data is the MIREX-like mood dataset [24]. In MIREX, 764 data has lyrics and audio. But because of conversion label emotion process to Thayer model, we used 456 data. Thats data has a 'cluster 2', 'cluster 3', and 'cluster 5' emotion labels. The data will be used in the music emotion classification. There are 2 testing models: CBE accuracy testing for emotional classification based on psycholinguistic features, and emotion classification testing of music with various features. The first test was conducted with the aim of analyzing the best CBE case to be used for psycholinguistic feature extraction. The second test is done with the aim of finding the best feature that will be used for the classification of emotion music.

In the first test, there are 3 cases of CBE, namely: CBE1, CBE2, and CBE3. CBE1 is a merging ANEW and WNA dataset with no expansion process or automatic tagging procedure. CBE2 is CBE1 which has undergone automatic tagging process using Wordnet synonym concept. And CBE3 is the development of CBE2 which has undergone automatic tagging process using Euclidean Distance concept. For these tests, the three CBEs are used interchangeably for the extraction of psycholinguistic features in the lyrics. The result of its feature extraction is used for the classification of emotion music.

Figure 10 shows the deployment of CBE1 data in dimension space, where C1 is Valence, C2 is Arousal and C3 is Dominance. From Figure 10, it appears that the data is scattered well on the dimensional space. It making easier for the conversion process into the thayer model. The obstacle is the difference existence of dimensional. CBE have 3 dimension while Thayer model only have 2 dimension. For the time being, we adjusted the data with Thayer model using 2 dimension (Valence-Arousal).

CBE2 is formed with the help of Synset of Wordnet. Figure 11 shows the central position of the cluster in dimension space. The center of cluster is the initial step result of autotagging procedure of incomplete data. This center of cluster will be the center term to formation of CBE3.

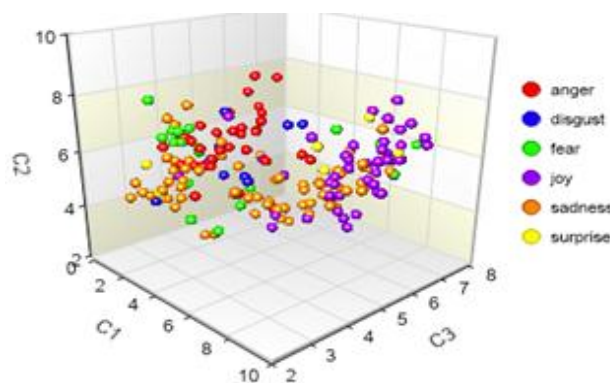


Figure 10. Spread of CBE1 data



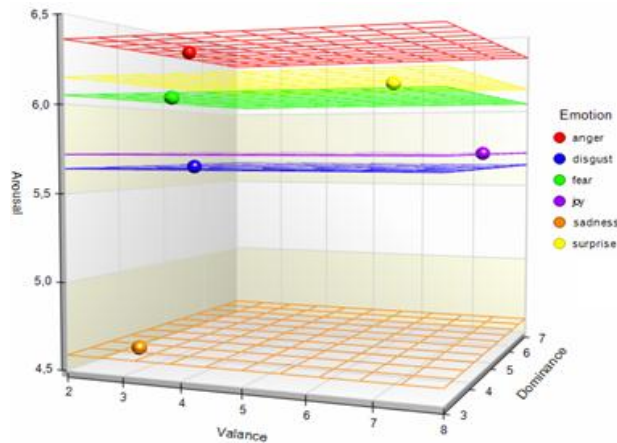


Figure 11. Cluster center position in dimension space

The formula of accuracy testing data terlihat pada formula (1). The formula is ratio between the sum of the true predicted with total document. And Table 3 shows the accuracy values obtained for different CBE uses in the extraction process of psycholinguistic features. It is seen that the use of CBE3 has a better percentage accuracy value of 0.37. So the psycholinguistic feature used for the classification process is a psycholinguistic feature using CBE3.

$$Accuracy = \frac{\text{sum of true predicted}}{\text{total document}} \times 100\% \tag{2}$$

Table 3. The Accuracy Values for different Case of CBE

Case	True prediction	Accuracy
CBE1	118	0,258
CBE2	137	0,3004
CBE3	169	0,37

The classification process was tested using SVM, Random Forest and Naive Bayes models. The classification process using the help of Weka tools with percentage split 66%.The F-Measure value of each method is shown in the Table 4. There are 4 test cases, each case using different features. The first case using the audio feature only. The audio feature are dynamic loudness, tempo, roughness and inharmonicity. It appears that the best results obtained by using Naive Bayes with value of 0,460. The second case use stylistic feature. The feature is only capable of bringing the result of 0,456 with Naive Bayes method. Unique to the third case, the psycholinguistic features are not affected by three classification methods. The accuracy results are equal for all. The value of accuracy is 0,354.

Table 4. F-Measure Value of Classification Methode

Feature	SMO	Random Forest	Naive Bayes
Audio	0,281	0,428	0,460
Stylistic	0,358	0,433	0,456
Psycholinguistic	0,354	0,354	0,354
Audio, Stylistic, psycholinguistic	0,437	0,568	0,456

The last case is used all of features in case one, case two and case tree. It is seen that the Random Forest method with the use of Audio, Stylistic and Psycholinguistic features has the best F-measure value. There is 0.568.

### 3. CONCLUSION

This research show that the use of CBE is able to support the process of classification emotion of music. With the best F-measure for Random Forest method of 56.8%. For further research, additional process

will be developed to improve the extraction performance of lyrics and audio features. So that can be obtained better accuracy value. Once analyzed, the likelihood of errors occurring in lyrical feature extraction is the absence of the concept of Word Sense Disambiguation (WSD) [27], Adverb-Adjective Component (AAC) or Negation word. And for the audio feature needs to do more combination of feature extraction, so it can be done testing the best audio feature for the emotional classification of this music.

## REFERENCES

- [1] C. Strapparava and A. Valitutti, "WordNet-Affect: an Affective Extension of WordNet", *LREC*, pp. 1083-1086, 2004.
- [2] P. Ekman, "An-Argument-For-Basic-Emotions.pdf", *Cogn. Emot.*, vol. 6, no. 3, pp. 169-200, 1992.
- [3] H. Corona and M.P.O. Mahony, "An Exploration of Mood Classification in the Million Songs Dataset", *12th Sound and Music Computing Conference*, 2015.
- [4] B. Rocha, R. Panda, and R.P. Paiva, "Music Emotion Recognition: The Importance of Melodic Features", in *International Workshop on Machine Learning and Music (MML)*, no. 2008, 2013.
- [5] Y. Hu, X. Chen, and D. Yang, "Lyric-Based Song Emotion Detection with Affective Lexicon and Fuzzy Clustering Method", *ISMIR 2009*, pp. 123-128, 2009.
- [6] J.S. Downie, "When Lyrics Outperform Audio for Music Mood Classification: A Feature Analysis", *ISMIR 2010*, pp. 619-624, 2010.
- [7] M. Kim and H. Kwon, "Lyrics-based Emotion Classification using Feature Selection by Partial Syntactic Analysis", 2011.
- [8] J.A. Ridoean, R. Sarno, D. Sunaryo, and D.R. Wijaya, "Music mood classification using audio power and audio harmonicity based on MPEG-7 audio features and Support Vector Machine", *2017 3rd Int. Conf. Sci. Inf. Technol.*, pp. 72-76, 2017.
- [9] L. Lu, D. Liu, and H. Zhang, "Automatic Mood Detection and Tracking of Music Audio Signals", *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 14, no. 1, pp. 5-18, Jan 2006.
- [10] A. Schindler and A. Rauber, "Capturing the Temporal Domain in Echonest Features for Improved Classification Effectiveness", *International Workshop on Adaptive Multimedia Retrieval*, pp. 1-15, 2015.
- [11] R. Malheiro, R. Panda, P. Gomes, and R.P. Paiva, "Music Emotion Recognition from Lyrics: A Comparative Study", in *International Workshop on Machine Learning and Music (MML)*, pp. 9-12, 2013.
- [12] R. Panda, R. Malheiro, B. Rocha, A. Oliveira, and R.P. Paiva, "Multi-Modal Music Emotion Recognition: A New Dataset, Methodology and Comparative Analysis", *10th International Symposium on Computer Music Multidisciplinary Research*, pp. 1-13, 2013.
- [13] F. Xue, Hao; Xue, Like; Su, "Multimodal Music Mood Classification by Fusion of Audio and Lyrics", in *21st International Conference, MultiMedia Modeling*, pp. 26-37, 2015.
- [14] V. Kumar, "Mood Classification of Lyrics using SentiWordNet", *Int. Conf. Comput. Commun. Informatics (ICCCI -2013)*, pp. 1-5, 2013.
- [15] A. Esuli, F. Sebastiani, and V.G. Moruzzi, "SENTIWORDNET: A Publicly Available Lexical Resource for Opinion Mining", *Proc. Lr. 2006*, pp. 417-422, 2006.
- [16] A. Jamdar, J. Abraham, K. Khanna, and R. Dubey, "Emotion Analysis of Songs Based on Lyrical and Audio Features", *Int. J. Artif. Intell. Appl.*, vol. 6, no. 3, pp. 35-50, 2015.
- [17] H.K. Palo and M.N. Mohanty, "Classification of Emotional Speech of Children Using Probabilistic Neural Network", *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 5, no. 2, pp. 311-317, 2015.
- [18] V.L. Nguyen, D. Kim, V.P. Ho, and Y. Lim, "A New Recognition Method for Visualizing Music Emotion", *Int. J. Electr. Comput. Eng.*, vol. 7, no. 3, pp. 1246-1254, 2017.
- [19] M. Sudarma and I.G. Harsemadi, "Design and Analysis System of KNN and ID3 Algorithm for Music Classification based on Mood Feature Extraction", *Int. J. Electr. Comput. Eng (IJECE)*, vol. 7, no. 1, pp. 486-495, 2017.
- [20] C. Science, A. Harjoko, B. Jimbaran, and S. Utara, "Music Recommendation System Based on Context Using Case-Based Reasoning and Self Organizing Map", *Indones. J. Electr. Eng. Comput. Sci.*, vol. 4, no. 2, pp. 459-464, 2016.
- [21] F.H. Rachman, R. Sarno, and C. Fatichah, "CBE: Corpus-Based of Emotion for Emotion Detection in Text Document", in *ICITACEE*, 2016, pp. 331-335.
- [22] S. Banerjee and T. Pedersen, "An Adapted Lesk Algorithm for Word Sense Disambiguation Using WordNet", in *Third International Conference on Computer Linguistics and Intelligent Text Processing*, pp. 136-145, 2002.
- [23] B.Y. Pratama and R. Sarno, "Personality Classification Based on Twitter Text Using Naive Bayes, KNN and SVM", in *2015 International Conference on Data and Software Engineering (ICoDSE)*, pp. 170174, 2015.
- [24] X. Hu, J.S. Downie, C. Laurier, M. Bay, and A.F. Ehmann, "The 2007MIREX Audio Mood Classification Task: Lesson Learned University of Illinois at Urbana-Champaign Music Technology Group, Universitat Pompeu Fabra claurier@iua.upf.edu", in *Proceedings of the International Symposium on Music Information Retrieval*, pp. 462-467, 2008.
- [25] R.E. Thayer, *The Biopsychology of Mood and Arousal*, New York: Oxford University Press, 1989.
- [26] Y. Yang and H.H. Chen, *Music Emotion Recognition*, CRC Press, 2011.
- [27] B.S. Rintyarna and R. Sarno, "Adapted Weighted Graph for Word Sense Disambiguation", *IcoICT*, 2016.

**BIOGRAPHIES OF AUTHORS**

**Fika Hastarita Rachman** was born in Denpasar Bali, in March 1983. She lecturer in Informatics Departement in University of Trunojoyo Madura. In 2014, she joined to study at the Ph.D. degree in Department of Informatics, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. She current research activities are focused on Text Information Retrieval and Text Mining.



**Riyanarto Sarno** is a currently a professor at the Informatics Departement, Institut Teknologi Sepuluh Nopember, Indonesia. His research interests include: Internet of Things, Bussiness Process Management, Process Aware Information Systems, Knowledge Engineering and Smart Grids. He is a contributing author of a number of refereed journal, Informatics books and Proceeding papers.



**Chastine Fatichah** is a currently lecturer at the Informatics Departement, Institut Teknologi Sepuluh Nopember, Indonesia. Research topic interests are Image processing, Soft Computing, and Machine learning. She is a contributing author of a number of refereed journal and proceeding papers.