Title	A Natural Language Search Engine for Music driven by Moods
Author(s)	Trung Thanh, Dang
Citation	
Issue Date	2009-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/8143
Rights	
Description	Supervisor:Kiyoaki Shirai, School of Information Science, Master

A Natural Language Search Engine for Music driven by Moods

Dang Trung Thanh (710046)

School of Information Science,
Japan Advanced Institute of Science and Technology

February 5, 2009

Keywords: music information retrieval, mood classification, text categorization, machine learning.

Nowadays music is playing a more and more important role in human's life, whereas digital catalogs rapidly become larger and more inconvenient to access. If we don't have a good method to explore music, a large amount of music will be fallen into oblivion. In this research, we aim at building a natural language search engine for music driven by moods that integrates two important approaches to explore effectively music collections: searching music by natural language queries and exploring related songs by mood. Especially, mood of songs is an important feature in our system. Human has a habit listening to a song that fit best his current emotion. A grasp of emotions in songs would be a great help for us to effectively discover music. Imagine that one day you are walking with your girl friend in a sakura park. You would be very happy. So I'm sure that both of you might like listening to happy songs. You can just turn on your iPod, input "happy" mood category and "sakura" query, search and enjoy happy songs together.

There are some previous music search engines before but almost of them can treat with only metadata. Using Knees's approach, we build a music search engine system that enables searching not only with metadata but also with arbitrary keywords and moods.

In order to search songs by mood, the mood of each song should be identified. There are two major approaches to solve mood detection problem. The first one, which is the main stream, is based on acoustical data of music. Almost current mood detection researches extract features from acoustical data to discover the mood of a song. The second one uses textual data of song like metadata and lyric. Lyrics contain almost meaning and mood of a song. However it is not concerned much in current researches. All textual-based researches apply semantic analysis techniques using a common knowledge database to extract emotion from lyrics. The supervised machine learning methods are not applied yet although it is more practical than unsupervised methods. The major reason is lacking a large amount of music dataset tagged with correct moods. In this research, we choose supervised learning as the approach for our mood classification system. There are three major problems to be solved: how to define mood categories, how to construct a training dataset for mood detection and how to decide the mood of a song.

The first problem is how to define categories of mood or emotion. There are two approaches to model emotion: categorical and dimensional. The categorical model often uses a list of basic adjective words as emotion categories. The dimensional model uses some dimensions, each one of which relates to some specific acoustical features. Our mood classification system is not based on acoustical features, so dimensional model is not suitable. Thus we chose categorical model. In this thesis, mood clusters in a famous contest about Audio Music Mood Classification (MIREX 07) are used as mood categories.

The second problem is how to construct a large amount of mood-tagged song database. We use a big blog site LiveJournal (www.livejournal.com) which has more than 9,000 users. Each blog entry is tagged with mood and music by users. Users can choose a mood tag from 132 predefined moods of LiveJournal or input freely their current mood. After mining from LiveJournal, we collected about 6.3 millions posts which contains both correct music and mood tags. Using only 50 most frequent mood keywords, we extracted about 665,000 songs of 6,800 artists/bands which

are tagged with moods. Then, rules to map from a LiveJournal mood keyword to our mood categories are manually designed. In LiveJournal, a song can be tagged with many mood keywords. Thus, how to choose the correct mood of a song is a problem. We propose four methods to decide the mood of a song and choose the best method by closed tests.

The third problem is how to classify mood categories for each song. We applied some state of the art text categorization methods and proposed new methods to classify moods of songs using lyric and metadata. Three classifiers are obtained from the training data: Support Vector Machine, Naive Bayes and Graph-based.

The first classifier is SVM. We utilize sentiment word and artist name as features. Weighting models are proposed to consider these features more important than others. An effective weighting model using entropy is also applied to improve the system. Experiments show that sentiment word and artist features are effective for mood classification. However, sentiment word features are not as good as we expected. The reason is that number of sentiment words in lyrics is very few, so this method can not capture moods of songs.

The second classifier is Naive Bayes model. In several tasks, it is reported that Naive Bayes classifier is comparable to SVM. This inspired us to continue investigating various lyrical features and metadata features using Naive Bayes for mood classification.

Each song has a specific structure in which each part plays a different role. A lyric is generally divided into several parts (Title, Introduction, Verse, Chorus, Bridge and Outro) in which the most important parts showing almost meaning and emotion of a song are chorus and title. We proposed a weighting model that puts more weights for word features occurring in chorus and title. The effectiveness of this weighting model is confirmed by the experiment.

Each artist tends to compose or sing a similar music. For example, Eric Clapton often sings sad songs but Bob Marley likes singing happy songs. Inspired by this phenomenon, we proposed a new model based on Naive

Bayes to consider artist feature. This model achieved the best performance comparing to the other models in this research.

The third method for mood classification is a graph-based method. In many tasks, the usual approach for classification problem does not exploit the available information about relationships between data items. Applying the Oh's graph-based approach, we proposed a new model using artist to create links among songs in the graph. Experiments showed that this approach is good for mood classification although only artist relationship is used to construct a graph. If other types of relationship are used, we believe that the result will be much better.

Here we describe best experiments with the three methods. A dataset consisting 5,580 data items are divided into 5 parts, one for testing and the remaining for training. The mood category 3 has highest proportion (54.12%) in this dataset. We consider this proportion as the baseline of our system. The experiments with this dataset show that the best method of NB and graph-based classifier outperforms the baseline with accuracy 57.44% and 57.00%, respectively. On the other hand, the best SVM's experiment is 52.73% which is lower than the baseline.

The accuracy of mood classification methods is not really high and good enough to apply for a real music search engine system. However, the accuracy of our method is comparable and a little better than that of acoustical methods. There are two main reasons for this low accuracy: mood is a subjective metadata; lyric is short and contains many metaphors which only human can understand. However, the experiments showed that artist, sentiment words, putting more weight for words in chorus and title parts are effective for mood classification. Graph-based method promises a good improvement if we have rich relationship information among songs.