Report template for the project in the course DD2380 at KTH

GROUP31

Akash Patel
BIRTHDATE1
MAIL1@kth.se

Christoph Kaiser
920416-T336
16/04/92
ckai@kth.se

Lisa Schmitz
920606-T203
06/06/92
lschmitz@kth.se

Nikolaos Tatarakis
BIRTHDATE4
MAIL4@kth.se

Abstract

This paper explores the potential accuary of the analysis of song lyrics. Different text analysers were tested for their ability to categorize lyrics as *negative* or *positive*. The focus lies on the comparison of different feature extraction methods and classifier. The identification of emotion in lyrics is a problem which has no satisfying solution yet.

NOTE

Related Works, Experimental Results, Discussions, Summary are sections that MUST be contained.

The section *Contributions* is a place to express any difference in contributions. The default assumption is that you all agree that all of you had an equal part to play in the project.

1 Introduction (1–2 pages)

Music has a great impact on people. Everyone knows the phenomen that a song can influence our mood, it can make us sad and it can make us happy. This amazing control over people's feelings is something which can be used for many different purposes. For example music provider like Spotify offer playlists labelled with a certain mood. But industries are not the only area of application. Researchers see a use for it in edutainment and even psychological therapy [2]. Unfortunately, the task of predicting the correct assoziated mood is not an easy one due to the comlexity of how emotion is transferred in songs. Obviously emotion is encoded both in the audio and the lyrics of a song [6]. This paper compares methods to identify emotion by analysing the text of song lyrics. In order to do this different variants of text analysers were tested. The modification was conducted by using different categories of emotions and classifiers.

1.1 Contribution

1.2 Outline

Since our work deals with different approaches of categorising and classifying song lyrics, previous work should be taken into count. The related work is therefore presented in Section 2. We based our text analyser in the results of this previous studies. Section 3 explains the method we used to realise and implement the analyser in detail. We used different variations of our text analyser, modifying the the categorsiation and the classifier. The results we were able to gather are described in Section 4. Moreover, problems that came up during the research are mentioned in this section. The results are sumarized in Section 5 and possible further research areas are given eventually.

2 Related work

Our work is mainly based the on paper of Youngmoo E. Kim et al. [2]. It gives an overview of recent approaches of emotion recognition in lyrics. Most of the presented approaches are content-based and are therefore relevant references for our work. Not only do they deal with different categorisations of mood but also treat variations of classifiers. This work provides a good inside into what has already been done and what worked well. Therefore it can be seen as the foundation we built our work on. We realised some of the presented methods and compared them to each other.

Different bag-of-words model classifiers are explained by Li et al. [3].

3 Method

Bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla

3.1 Implementation

The basic model we used for our text analyser is the bag-of-words model [4]. The functionality is illustrated in the figure ??. In this model a database with labeled data such as lyrics are stored. The next step would be to preprocess the lyrics and create a dicitonary with n-grams based on it. The database is divded into two parts: a training and a test set. Using this data, a classifier is trained with a training set and its accuracy can then be evaluated with the classification of the test set. In the following abstracts the individual parts of the bag-of-words model will be explained in detail.

Database The database was one of the most difficult parts of this research. Due to copyright issues it was not possible to access a already available database and we had to built one on our own. Some earlier researches use `Allmusic.com` as a basis for the database [2], since it provides emotional labels for songs. For this reason a program has been written to extract the lyrics with its song's emotional label.

Even though this prelabeling of the songs was helpful, with over ... emotional labels it were simply to many categories to use. Therefore we had to find supercategories for the existing labels. Downie et al. suggest in [1] to use five clusters of emotions. Whereas Yang and Lee only suggests a binary

distinction into positive and negative emotions [5]. Both variants were tested during our research.

Pre-processing The lyrics have been pre-processed the same way for each of the variants of the text analyser. Stop-words are filtered out using the Natural Language Toolkit and special character are deleted. In a next step, every word of the lyric is stemmed which means that it is transferred into its basic form. For example "wait" and "waiting" would be interpreted as the same word after the stemming has been done. The pre-processing is performed due to avoid misclassifications.

Feature extraction The feature extraction is condtucted for all lyrics. The procedure consists of transferring the lyrics into n-gram representation. We considered only 1-, 2- and 3-grams as reasonable parameters which is why only these have been tested in our study.

Classification The Natural Language Toolkit and the SciKit-Learn library provide the most commonly used text classifiers. In the following section the tested classifiers will be introduced.

- *Naive Bayes* Let $\mathcal{C} = (c_1, \ldots, c_n)$ be the categories and n the number of categories. Given a document $\mathcal{D}$ and its word list $\mathcal{W} = (w_1, \ldots, w_m)$ with m being the number of words of a certain document, the Naive Bayes classifier determines the category of a document as follows:

  $$c^*_{NB} = argmax_{c_j \in \mathcal{C}} P(c_j) \prod_{i=1}^{d} P(w_i|c_j)$$

  with $P(c_j)$ being the *a priori* probability of a class c_j and $P(w_i|c_j)$ being the conditional probability of the word w_i given class c_j.

  Naive Bayes will only work well for large enough datasets.

- *Multinomial Naive Bayes*
- *Bernoulli Naive Bayes*
- *Logistic regression*
- *Stochastic gradient descent*
- *Support vector machine*

- *Linear support vector machine*
- *Combination of multiple classifiers* We combined all introduced classifier into one classifier. It basically classifies documents according to how many classifiers resulted into a certain category. The category, to which most of the classification results belong to, gets selected as the resulting category of the song.

4 Experimental results

4.1 Experiemntal setup

Bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla

4.2 Experiment ...

Drawbacks We were aware of the difficulty of the task of designing a sufficient lyrics classifier. Especially in the last years a lot of research has been conducted in this area but there is still a lack of a reliable lyrics analyser. But during our research we were able to gain a better inside into the nature of this difficulty. Feelings and emotions are something very personal, and the perception of the emotion which is transferred by a certain songtext highly depends on the person who is reading the lyrics. It is not unusal that two different people might disagree on the general mood of a lyrical text. This is why it is of supreme importance to select a representative model of categories.

During our research we came to the conclusion that the use of only two categories - positive and negative - seem to have let to an overgeneralisation of emotions. Even the categories itself which were provided by `Allmusic.com` could not clearly be mapped to one of these two categories. To illustrate the difficulty we discovered, a research has been done. We used a questionnaire were we asked people to evaluate the categories from `Allmusic.com` as either negative or postive and to do the same with some randomly picked song lyrics. The results are shown in figure 1. It can be seen that there was a high difference of opinion across participants. This emphasizes the personal perception of emotion. Due to this, it cannot be expected to gain a higher accuracy by a classifier in emotion categorisation.

A second model of categorisation used in this research has been proposed by Downie et al. [1]. They used five cluster of emotion to categorise lyrics as shown in table 1.

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, autumnal, brooding
Cluster 4	humorous, silly, campy, quirky, whimsical, witty, wry
Cluster 5	aggressive, fiery, tense/anxious, intense, volatile, visceral

Table 1: Clusters of mood adjectives used in the MIREX Audio Mood Classification task [1].

Variations in categorisation

Variations in classifiers Bla bla bla bla bla Figure 1 bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla

Bla bla bla bla bla Table **??** bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla

5 Summary and Conclusions

Bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla bla

6 Contributions

We the members of project groupXX unanimously declare that we have all equally contributed toward the completion of this project. (PLEASE CHANGE THIS SUITABLY WITH DETAILS, IF IT IS NOT TRUE)

References

[1] XHJS Downie, Cyril Laurier, and MBAF Ehmann. The 2007 mirex audio mood classification task: Lessons learned. In *ISMIR 2008: Proceedings of the 9th International Conference of Music Information Retrieval*, page 462. Lulu. com, 2008.

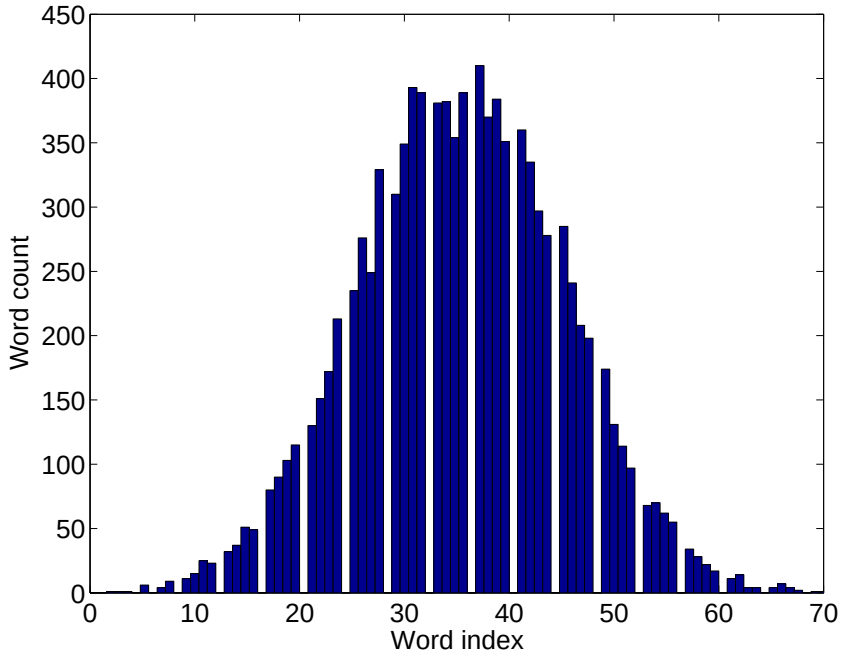

Figure 1: A description that makes browsing the paper easy and clearly describes what is in the picture. Make sure that the text in the figure is large enough to read and that the axes are labelled.

[2] Youngmoo E Kim, Erik M Schmidt, Raymond Migneco, Brandon G Morton, Patrick Richardson, Jeffrey Scott, Jacquelin A Speck, and Douglas Turnbull. Music emotion recognition: A state of the art review. In *Proc. ISMIR*, pages 255–266. Citeseer, 2010.

[3] Yong H Li and Anil K. Jain. Classification of text documents. *The Computer Journal*, 41(8):537–546, 1998.

[4] Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.

[5] Dan Yang and Won-Sook Lee. Music emotion identification from lyrics. In *Multimedia, 2009. ISM'09. 11th IEEE International Symposium on*, pages 624–629. IEEE, 2009.

[6] Yi-Hsuan Yang, Yu-Ching Lin, Ya-Fan Su, and Homer H Chen. A regression approach to music emotion recognition. *Audio, Speech, and Language Processing, IEEE Transactions on*, 16(2):448–457, 2008.