



Music Mood Classification Based On Lyrics

George Mathis Jr., Debzani Deb

Department of Computer Science

Winston-Salem State University, Winston-Salem, NC



WSSU

Purpose

- Recent studies on Music Information Retrieval (MIR) found that music mood is increasingly becoming an important access point to the music repositories and collections.
- A number of automatic mood-based classification methods have been explored which rely on different audio and instrumental features and some of them use song lyrics in their classifications.
- The studies provided contradictory results and are mostly based on small-scale datasets.
- In this study, we focus on the analysis of song lyrics for music mood classification using Apache Spark platform.
- Utilized Spark's MLlib library for the classification purpose in order to apply the classification algorithms on some sizable datasets in an efficient way.

Acknowledgement



NSF RIA Award #1600864

WSSU RIP Award



Input Datasets

- We utilized J.A. Russell's psychological model [1] to derive mood categories.
- Our training dataset contains 1000 songs from the popular Million Songs Dataset (MSD) [2] that are available to us with lyrics and with social-tag based mood categories ("happy" and "sad").
- We are also working on building a much bigger training dataset by finding a subset of songs in MSD for which lyrics are available and all 4 proposed mood categories are derivable

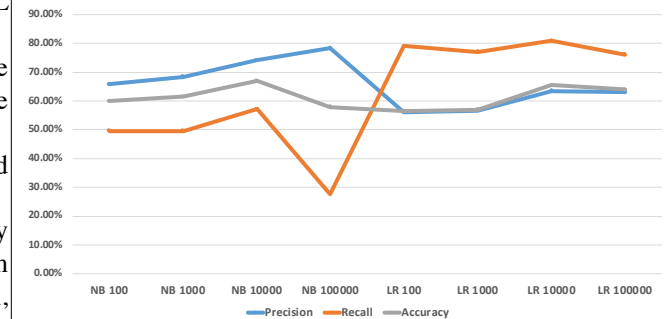
Methodology

- Spark has a library containing common machine learning (ML) functionality, called MLlib [3]. MLlib has many types of ML algorithms for classification.
- Preprocessing:** The dataset was obtained in CSV format so we developed a python program to parse the data and print out the lyrics for analyzation.
- Once preprocessed, the training dataset from the MSD contained 554 Sad songs and 446 Happy songs.
- Classification:** During the classification process, term frequency feature vectors are created from training dataset lyrics and then two classification algorithms, Naïve Bayes and logical regression, are utilized to train our models.
- Experiments were performed using separate validation test dataset containing 200 songs (105 Happy Songs and 95 Sad Songs) and the accuracy of the various classification algorithms are compared.
- Results are recorded as % precision, % recall, and % accuracy of correctly classified moods for unlabeled lyrical data.

Results

	Happy	Sad	Precision	Recall	Accuracy
NB 100	52	68	66%	50%	60%
NB 1000	52	71	68%	50%	62%
NB 10000	60	74	74%	57%	67%
NB 100000	29	87	78%	28%	58%
LR 100	83	30	56%	79%	57%
LR 1000	81	33	57%	77%	57%
LR 10000	85	46	63%	81%	66%
LR 100000	80	48	63%	76%	64%

%Precision, %Recall & %Accuracy



References

- J. A. Russell: "A Circumplex Model of Affect," Journal of Personality and Social Psychology, 39: 1161-1178, 1980.
- T. Bertin-Mahieux, et. al. "The million song dataset," ISMIR, 2011.
- H. Karau, A. Konwinski, P. Wendell, & M. Zaharia Learning Spark: Lightning-Fast Data Analysis, O'Rielly, 2015

WINSTON-SALEM STATE UNIVERSITY