# PREDICTING THE EMOTION OF PLAYLISTS USING TRACK LYRICS

**Sara Giammusso**[1] **Mario Guerriero**[1] **Pasquale Lisena**[1]
**Enrico Palumbo**[1,2] **Raphaël Troncy**[1]
[1] EURECOM, Sophia Antipolis , France
[2] ISMB, Turin, Italy
{firstname.lastname}@eurecom.fr

## ABSTRACT

One interesting approach to playlists generation is using emotional information extracted from songs to try to group them according to a certain mood. The goal of this study is to unravel the emotion patterns underlying the sequences of songs in a playlist using automatic approaches of Emotion Recognition on the lyrics.

## 1. INTRODUCTION

In the last few years, the online music streaming services have enabled users to create and share custom playlists, giving to Recommender Systems (RS) a crucial role in the playlists continuation task. Modern RSs rarely rely on emotions of songs, mainly because of the subjectivity and difficulties in getting this information.

The task of emotion recognition often involves the analysis of human expressions in multimodal forms such as texts, audio, or video. In this work, we are interested in using the textual modality where the task is closed to Sentiment Analysis [8], which is the computational treatment of opinions, sentiments and subjectivity of a natural language text. Moreover, it can be used to improve the way a RS gathers information about a playlist. Performing emotion recognition is quite challenging and most of the existing works focus on those data sources which ease this task by the presence of specific words and parts of text, as it happens for the hash-tags in tweets [7]. Other works use the pleasure (valence), arousal and dominance (PAD) labels extracted from the words comprising a songs lyric [5].

This paper presents a novel approach for classifying the main emotion of playlists within four different classes: *relaxed*, *happy*, *sad* and *angry*. The prediction relies on aggregating the emotion prediction computed at the song level through the analysis of their associated lyrics (we consider English lyrics). Unlike most of the previous techniques [4], our approach heavily relies on Machine Learning and Natural Language Processing both for feature extraction and classification.



## 2. EMOTION RECOGNITION IN SONG LYRICS

Specific studies have identified a set of stylometric, structural, orientation and vocabulary-based features, which can be extracted from lyrics and successfully used as input to classifiers to predict the emotion of a song [3]. A smaller set of these features has been selected in order to maximise the performance of the algorithm described later:

- **%Past_tense_verbs**: the percentage of the past tense verbs over the total number of verbs;
- **%Present_tense_verbs**: the percentage of the present tense verbs over the total number of verbs;
- **%Future_tense_verbs**: the percentage of the future tense verbs ("will" or "'ll" + base form) over the total number of verbs;
- **%ADJ**: percentage of adjectives over the total of words;
- **%PUNCT**: percentage of punctuation over the total number of words;
- **%Echoism**: percentage of echoism over the total number of words, where an echoism is either a sequence of two subsequent repeated words or the repetition of a vowel in a word;
- **%Duplicate_lines**: number of duplicated lines of the total number of lines in the lyrics;
- **isTitleInLyrics**: *true* if the lyrics contain the title string;
- **Sentiment_polarity**: sentiment polarity, between -1 (negative) and 1 (positive);
- **Subjectivity_degree**: degree of subjectivity of the text between 0 and 1.

We normalized these features by subtracting the mean and scaling to unit variance. For each song, we defined a feature vector by concatenating a 300-dimensioned word embedding vector of the lyrics and the 10 features described above. The word embeddings of the lyrics rely on a pre-trained language model coming from SpaCy [1] and based on GloVe [6]. The SpaCy word embedding vector of each song consists of an average of the word embedding vector of each token in the song.

We tested the effectiveness of this feature vector for the emotion prediction by using four different classifiers [2] : an artificial neural network (ANN), a support vector machine, a logistic regression and Xgboost [2]. We used as ground truth the perfectly balanced *MoodyLyrics4Q* dataset [1], which contains 2000 manually annotated songs with the

[1] https://spacy.io/
[2] The source code is available at https://goo.gl/7Z5wiC.

| Classifier | Accuracy |
|---|---|
| Neural Network (ANN) | 58.45% |
| Logistic Regression | 57.87% |
| Support Vector Machine | 58.04% |
| xgboost | 56.89% |

**Table 1**. Accuracy results on MoodyLyrics4Q

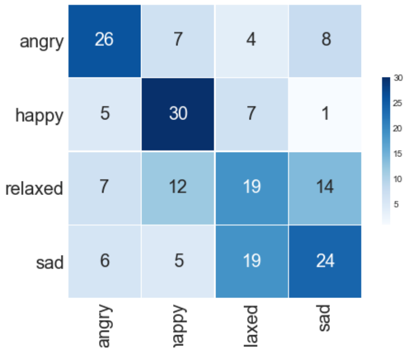


**Figure 1**. ANN Confusion Matrix

four different emotion labels *relaxed*, *happy*, *sad* and *angry*. The training has been performed on 90% of the entire dataset, while the randomly sampled remaining 10% has been used for testing. The lyrics have been collected from LyricsWikia[3]. Table 1 contains the accuracy computed with a 10-fold cross validation. ANN obtains the best performances.

A confusion between the *"sad"* and *"relaxed"* classes is a common pattern in all classifiers (see Figure 1). To investigate the reasons, we downloaded and read the lyrics of some songs and we discovered that discriminating between *"sad"* and *"relaxed"* emotions is hard, also for humans.

## 3. PLAYLIST CLASSIFICATION

For each song, the ANN produces in output a probabilistic distribution (rather then an hard assignment), in the form of vector `[sad%, angry%, happy%, relaxed%]`, where the sum of the percentages is equal to 100. Then, the emotion score $s_i^x$ for the emotion $i$ in the playlist $x$ is equal to the sum of all the individual song probabilities normalised by the number of songs $l$.

$$s_i^x = \sum_{song \in x} \frac{s_i^{song}}{l} \tag{1}$$

The dominant emotion within a playlist is chosen by selecting the one with the highest score. Computing $s_i^x$ for each emotion would give us an emotional vector, similar to the ones we have generated for individual lyrics.

## 4. EVALUATION AND DEMO

In order to face the absence of a dataset of annotated playlists, a perfectly balanced dataset has been generated by manually picking 40 playlists from Spotify, chosen among the ones representing an emotion in the playlist title, e.g. *"Sad songs"* or *"angry music"*. Within this environment, our algorithm reaches **80%** of accuracy.

For better appreciating the predictor, we developed a web app using this algorithm available at `http://data.doremus.org/emotion`. The apps requests access to a user Spotify account in order to reach the Spotify Developer API. The application receives as input from the user the Spotify URI of a playlist (some predefined playlists can also be selected). The lyrics are downloaded and classified, and then, the results for predicting the dominant emotion are aggregated.

## 5. CONCLUSIONS

The results we obtained show an certain emotion homogeneity in playlists which suggests that emotion recognition produces interesting information for improving Recommender Systems. Training the model is limited by the small dataset size. More complex and effective classification approaches would be possible with bigger datasets.

## 6. REFERENCES


[1] Erion Çano and Maurizio Morisio. Music mood dataset creation based on last.fm tags. In *International Conference on Artificial Intelligence and Applications (AIAP)*, 2017.

[2] Tianqi Chen and Carlos Guestrin. Xgboost: A scalable tree boosting system. In *22$^{nd}$ ACM International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016.

[3] Michael Fell and Caroline Sporleder. Lyrics-based analysis and classification of music. In *25$^{th}$ International Conference on Computational Linguistics (COLING)*, 2014.

[4] Youngmoo E. Kim et al. Music emotion recognition: A state of the art review. In *11$^{th}$ International Society for Music Information Retrieval Conference (ISMIR)*, 2010.

[5] Owen Craigie Meyers. A mood-based music classification and exploration system. Master's thesis, Massachusetts Institute of Technology, 2007.

[6] Jeffrey Pennington, Richard Socher, and Christopher D. Manning. Glove: Global vectors for word representation. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.

[7] Ashequl Qadir and Ellen Riloff. Learning emotion indicators from tweets: Hashtags, hashtag patterns, and phrases. In *Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 2014.

[8] Hassan Walaa, Medhat Ahmed and Korashy Hoda. Sentiment analysis algorithms and applications: A survey. *Ain Shams Engineering Journal*, 1994.


[3] `http://lyrics.wikia.com/`