Emotion Patterns in Music Playlists

Sara Giammusso[12] Mario Guerriero [12]

[1]MSc student in Data Science Department, EURECOM, Télécom ParisTech, France
[2]MSc student in Department of Control and Computer Engineering, Politecnico di Torino, Italy

Sixth Project meeting

Table of Contents

Table of Contents

Introduction

Previously On Sara&Mario Project...

We worked on feature engineering to see how we could improve our classifiers

Next steps:

- Make sure the POS tagger we are using is good
- Make some experiments with other datasets
- Keep working on feature engineering
- Start looking at playlists

Table of Contents

POS Tagger Analysis

- We are using spaCy's built-in POS tagger
- There is no authoritative paper regarding its internal details
- Our experiments suggest that it is good for our purpose [1]
  - Each line is treated as a separate sentence
  - The tagger is smart enough to recognize less common words and abbreviations (very common in songs)

[1] Please refer to the following Jupyter Notebook to know more about what we did on this topic: `https://github.com/sgiammy/emotion-patterns-in-music-playlists/blob/master/src/POS_tagger_verification.ipynb`

More on Feature Engineering

- We had built a big list of features in the past few weeks
- We concluded that we do not need all of them (of course)
- We got the best performances just by using a subset of 10 of them [2]
  - WORD_COUNT, ECHOISMS, SELFISH_DEGREE, DUPLICATE_LINES, IS_TITLE_IN_LYRICS, VERB_PRESENT, VERB_PAST, VERB_FUTURE, ADJ_FREQUENCIES and PUNCT_FREQUENCIES

[2] For more information, please refer to `https://github.com/sgiammy/emotion-patterns-in-music-playlists/blob/master/src/Advanced_Feature_Engineering.ipynb`

MoodyLyrics + EmoInt

- We are very limited by MoodyLyrics [3]
- EmoInt is made of tweets classified according to four emotions: anger, fear, joy and sadness
  - We have mapped anger to angry, joy to happy, sadness to sad and discarded fear
- We were able to increase our dataset by 2254 items
- No good results at the end

[3] https://github.com/sgiammy/emotion-patterns-in-music-playlists/blob/master/src/EmoInt.ipynb

Table of Contents

Two More Features

We wanted to expand our model with two features

- Sentiment/Polarity
- Subjectivity

To achieve this goal we used TextBlob[4]

- Built on top of NLTK and Pattern
- Makes use of Pattern's analyzer to generate polarity and subjectivity

At the end we will have 12 features (311 in reality, because the lyrics content vector is expanded)

[4] `http://textblob.readthedocs.io/en/dev/index.html`

Obtained Results: ANN

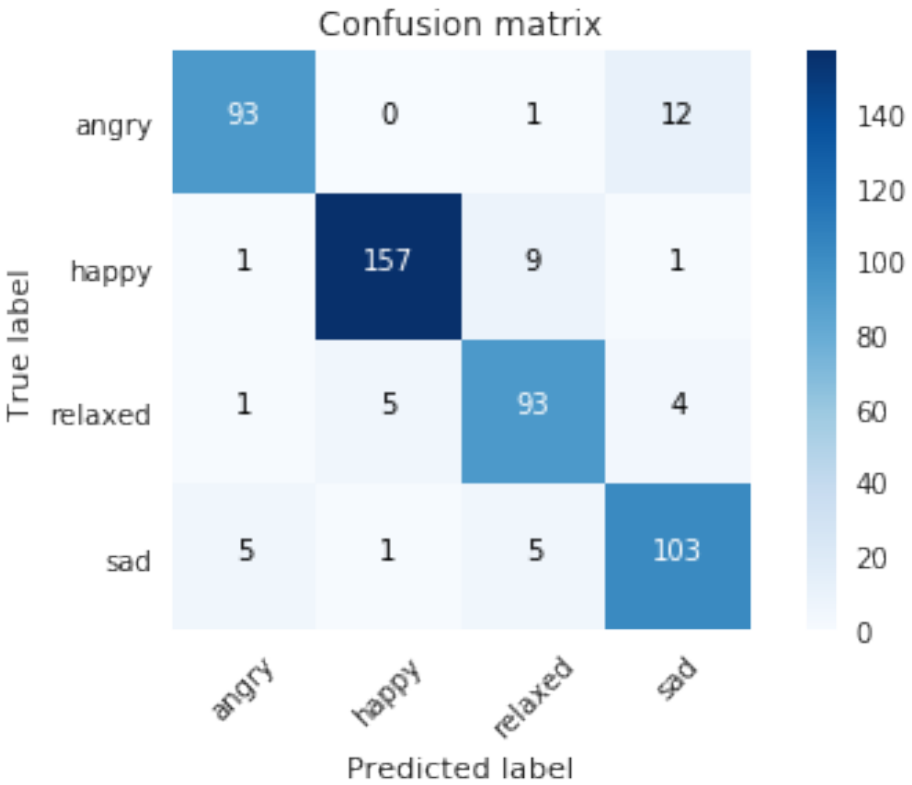

Figure 1: 90.84% accuracy on MoodyLyrics

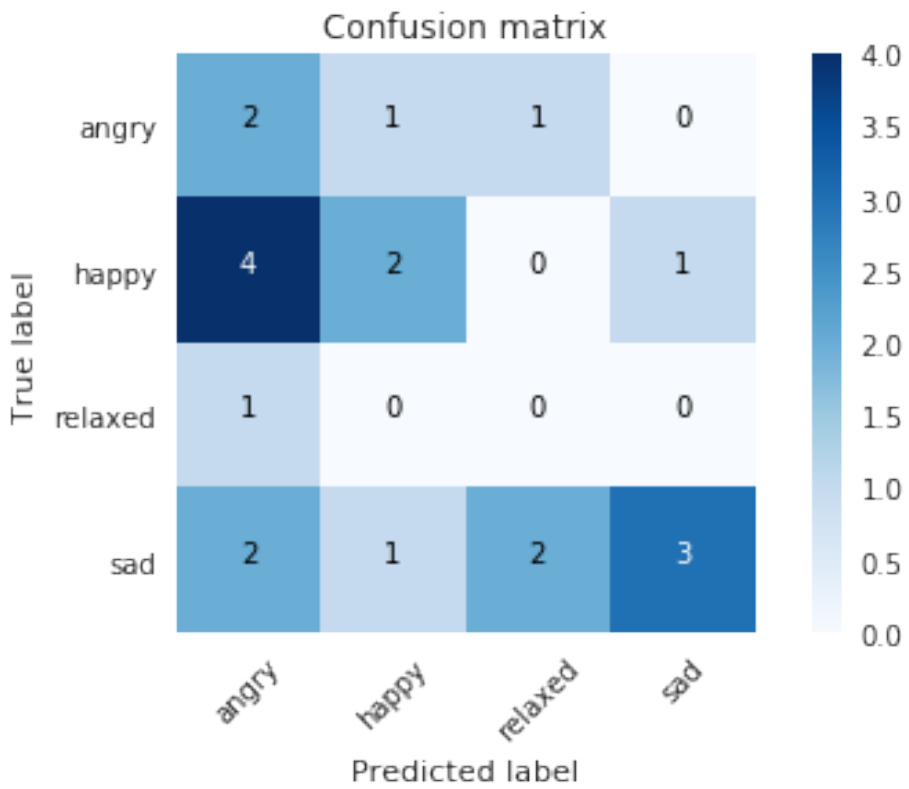

Figure 2: 50% accuracy on extra test

Obtained Results: Logistic Regression[5]

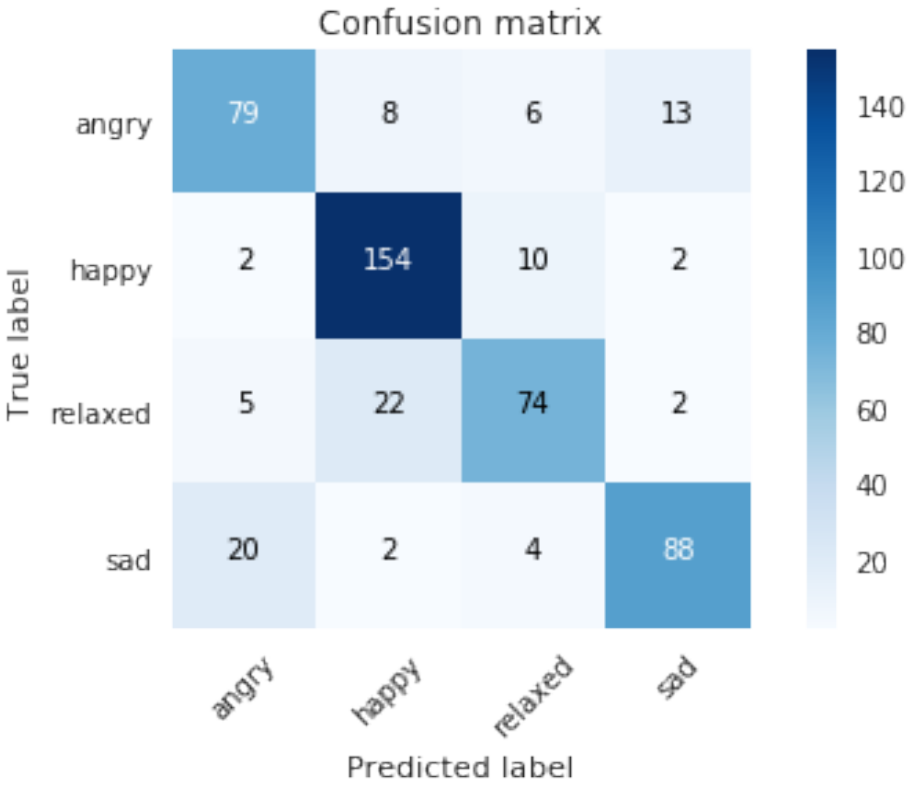

Figure 3: 80% accuracy on MoodyLyrics

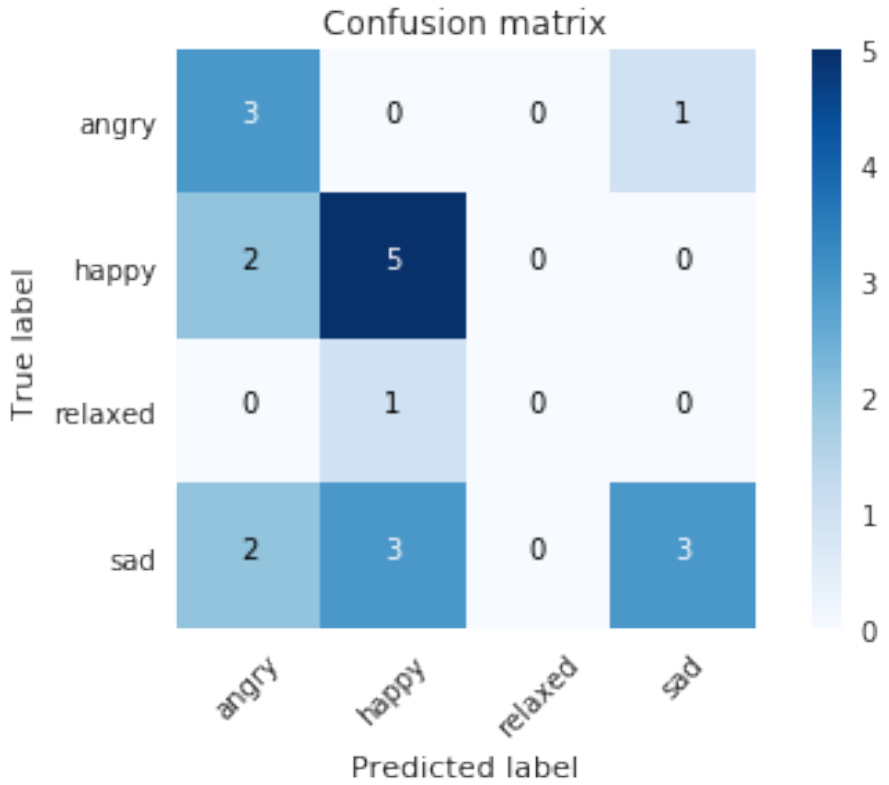

Figure 4: 55% accuracy on extra test

[5]We experimented with several probabilistic classifiers and Logistic Regression was the one giving the best results

Table of Contents

Which classification model?

We considered different approaches

- Give each song in the playlist an emotion label and choose using a majority rule
- Consider the text of all the songs in a playlist as belonging to a single song and classify it
- Compute an "emotion vector" for each song and average all of those vectors over the playlist [6]
  - This is the one which makes more sense to us
  - It produces 4 playlist features (one per emotion) you may want to use in your recommendation system

[6] Remember the sliders approach? We got inspired from this

Our Results[7]

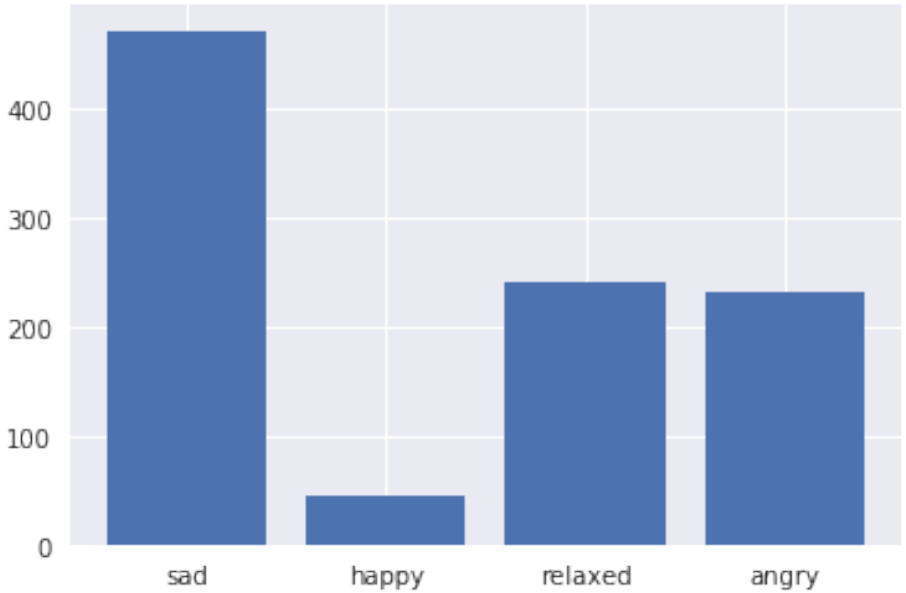

Figure 5: Playlists emotion distribution with ANN

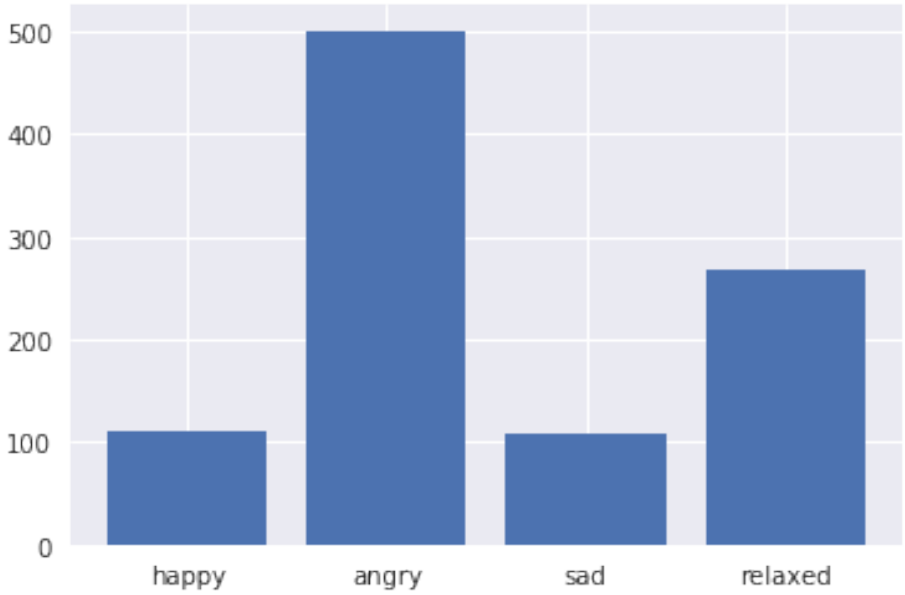

Figure 6: Playlists emotion distribution with Logistic Regression

[7] We will try to keep this notebook updated with all future statistics: `https://github.com/sgiammy/emotion-patterns-in-music-playlists/blob/master/src/Playlist_Classification_Stats.ipynb`

Table of Contents

What's next?

- We can't evaluate our results
- Could you please test them inside your recommendation system?
- Can we start writing our report? (if yes, any hint?)
- ...