Emotion detection in song lyrics

Emanuele Morales - 941935

Università degli Studi di Milano - Data Science and Economics

1 Introduction

The aim of this project is to create a model able to classify the emotions contained in the lyrics of the songs and to create a tool to suggest to users a playlist of songs containing the more appropriate tracks, based on the genre's preferences of the listeners and their emotional status in the moment they want to listen a song.

Nowadays, for music platforms like Spotify, Apple music, Google play, etc. it has became fundamental to improve the service quality in order to obtain a competitive advantage with respect to the competitors. One of the aspects that could increase the service level is the ability of the platform to recommend to users the right songs at the right moment.

Recommendation of songs based on the emotion communicated by the song is a good strategy, because emotions are an important part of human life and highly influence decision making and it is one of the drivers that make a user choose a song when he connects to his favourite platform.

Emotion detection can be performed on the audio features of the songs or from the analysis of the lyrics. In this project, it will be exploited the latter approach.

One of the more popular model applied to predict emotion in text is Support Vector Machine (SVM) that allows to obtain good performance if applied to web blog data [1] .

2 Research question and methodology

In order to detect emotions in lyrics, in this work it is proposed an approach based on *logistic regression model* that is trained on labeled tweets. This model tries to identify the presence of the four fundamental emotions (*anger, fear, joy, sadness*) into the lyrics of four different music's genres (*Rock, Pop, Hip Hop, Electronic*).

The four main emotions detected in the lyrics are then combined by exploiting the *Plutchik's wheel of emotions* to obtain further and more specific emotions (*frozeness, pride, envy, excitement, despair,bittersweetness*).

Then it is analyzed the dependency between genres and emotions by executing an analysis of the *contingency table* and the *Pearson's chi-square test.*

Finally, a recommendation system based on the favourite genres of the users and the *Jaccard similarity* between the feelings of the users and the ones contained in the lyrics is implemented to suggest a playlist of tracks to the final user.

3 Experimental results

3.1 Datasets

The training of the model has been performed on the *Wassa-2017* dataset [2]. The dataset is provided for four emotions: joy, sadness, fear, and anger. For example, the anger training dataset has tweets along with a real-valued score between 0 and 1 indicating the degree of anger felt by the speaker.

The lyrics dataset has been obtained by exploiting the *Genius Lyrics API* that allows to scrape lyrics from the Genius Lyrics website [3]. A dataset containing the 6 most popular songs of 20 artists for each genre considered (*Rock, Hip Hop, Pop, Electronic*) has been created, including the name of the artist, the title of the song, the genre and the lyrics, for a total of 480 lyrics scraped.

3.2 Pre-processing

Both the datasets have been subjected to a common part of pre-processing consisting in removing non-english words from the text, remove punctuation, remove double spaces and numbers, remove stopwords, and lemmatization..

Moreover:

- tweets have been cleaned by Twitter's features (retweet, hashtag, mentions, etc.).
- the intensity of the emotion, initially expressed as a continuous number between 0 and 1, has been categorized. In particular, the "noisy tweets", containing ambiguous levels of emotions (that correspond to an intensity between 0.35 and 0.65), have been removed in order to obtain a more precise model. The intensity has become a binary variable: 1 if the tweet contains an emotion with a intensity $>= 0.65$ and 0 if the tweet contain an emotion with intensity $<= 0.35$, obtaining a binary variable suitable for the logistic model.
- from the lyrics, the *chorus* has been extracted in order to perform a deeper analysis in the case in which the detection of the emotions by the model in the whole lyrics would result ambiguous. The chorus represents the more significant part of the song and it can be supposed it summarizes the emotion of the whole song.

3.3 Logistic Model for emotion detection

The logistic model is a discriminative model applied in order to categorize binary variables. It models the probability of one element to belong to a particular class given some features. In order to train the model, the text of the tweets of each emotion have been "vectorized" with a binary representation of the single words, checking the presence/absence of a word in a tweet. The result is a matrix for each emotion that will be given in input to the logistic regression model. For each emotion is created a model, trained on the correspondent emotion training set, that is able to identify the emotion in a new text given in input.

The validation of the model has been performed by applying a 5-fold cross-validation, obtaining the following results:

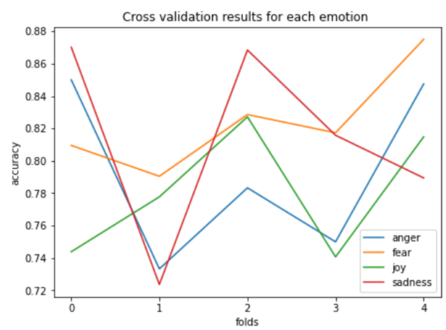

Fig. 1. Cross validation result for each of the 5 folds and emotions

	Logistic_regression	Avg_cross_validation
0	Anger_model	0.79
1	Fear_model	0.82
2	Joy_model	0.78
3	Sadness_model	0.81

Fig. 2. Average cross validation results

3.4 Application of logistic models on song lyrics

Each of the model fitted on the different emotion training set is applied one time on the whole lyrics of the songs and one time on the chorus only. The result is the detection of the emotions contained in each songs.

In order to obtain a wider spectrum of emotions and not the four principals emotions only, the *Plutchik's wheel of emotions* theory [4] is applied. Starting from the primary emotions and combining them it is possible to obtain a wider range of emotions.

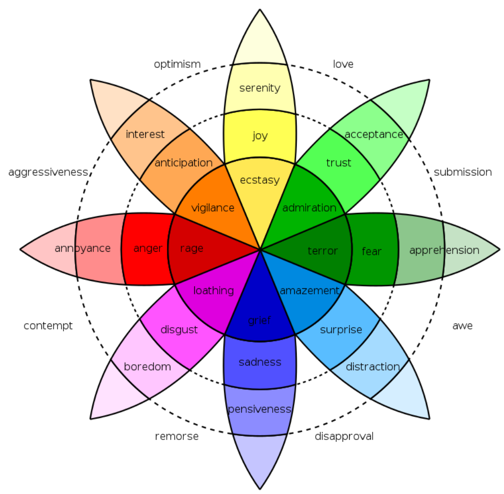

Fig. 3. Plutchik's wheel of emotion

It is possible to apply this theory for the lyrics for which the logistic model identified two of the principal emotions. In particular:

- anger + fear = *frozeness*
- anger + joy = *pride*
- anger + sadness = *envy*
- fear + joy = *excitement*
- fear + sadness = *despair*
- joy + sadness = *bittersweetness*

The two main difficulties verify in the case in which the model does not recognize any emotion in the song, and in the case in which the model recognizes more than two emotions.

In order to overcome these two issues the following approach is adopted:

– *more than two emotions detected*:
  the model could be "confused" by the large number of words contained in the lyrics. For this reason the model is applied on text of the chorus only, that is the most important part of the song that should summarize the sense of the song and so the emotions contained. This approach reduces the number of ambiguous emotion found previously.
– *no emotion detected*:
  a correlation between the genre and a particular emotion is obtained in order to assign to "unclassified" song the emotion related to the genre it belongs to.

3.5 Contingency table and chi-square test: analysis of dependencies

In order to discover if there is a particular correlation between the genre of the songs and the emotions, it is built a contingency table. This tool is a tabular mechanism with at least two rows and two columns used in statistics to present categorical data in terms of frequency counts.

emotion	anger	bittersweetness	despair	envy	excitement	fear	frozeness	joy	pride	sadness	All
genre											
Electronic	10	9	8	0	4	8	2	36	1	9	87
Hip-Hop	12	3	10	1	7	12	10	24	5	1	85
Pop	8	3	3	3	8	7	3	48	0	8	91
Rock	8	5	12	4	12	6	4	34	1	6	92
All	38	20	33	8	31	33	19	142	7	24	355

Fig. 4. Contingency table to represent the frequency by emotions and genre

To understand if there is an association between the row and column variables, the chi-square test is applied starting from the contingency table above. The null hypothesis H0 assumes that there is no association between the variables, while the alternative hypothesis is verified in case of association. This test is based on the divergence of the observed data from the values that would be expected under the null hypothesis of no association. The expected observation (in case of independence) for each entry of a contingency table and the statistic are calculated as:

$$E = \frac{f_{j\cdot} * f_{i\cdot}}{N} \quad \text{and} \quad \chi^2 = \sum \frac{(observed - expected)^2}{expected}$$

The result of the test on the contingency table gives the following result, that suggests a dependency between the two variables with a significance set to 90%.

```
p value is 0.06
Dependent (reject H0)
```

Fig. 5. Chi square test result

The alternative hypothesis does not specify the type of association, so in order to interpret the information obtained from the test, it is necessary an analysis of the distances between actual and expected values. Here it follows the distance matrix (observed - expected) represented on bar charts for each genre.

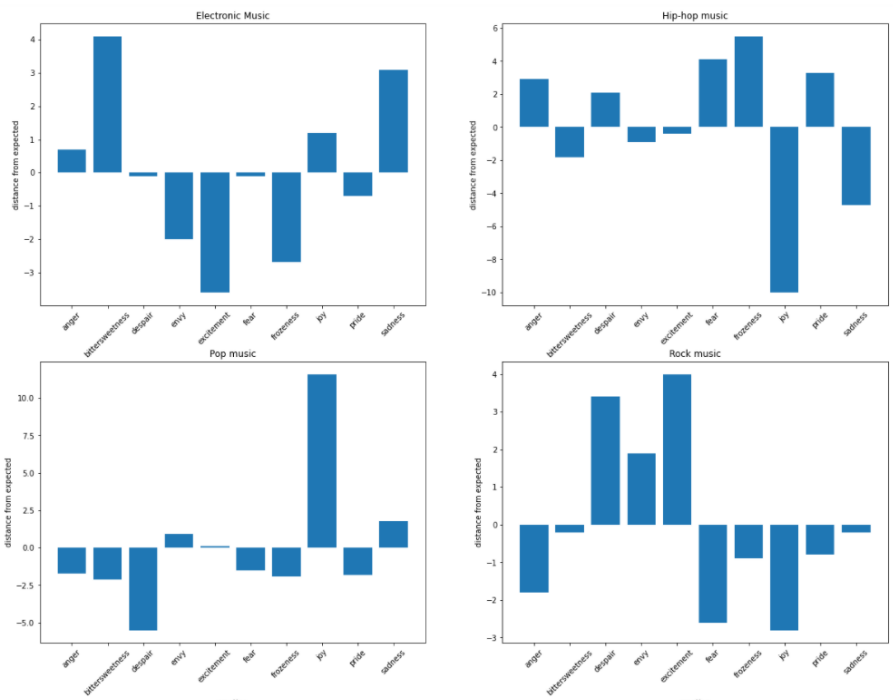

Fig. 6. Bar charts of the distance

From these plots it can be observed how some genres are related to some emotions. Particularly:

– Electronic lyrics are associated positively to bitter-sweetness and sadness and negatively to excitement and envy,
– Hip-Hop lyrics are associated positively to frozeness, fear, anger, pride and negatively to joy and sadness.
– Pop lyrics are associated positively to joy and negatively to despair.
– Rock lyrics are associated positively to excitement, envy and despair and negatively to fear joy and anger.

These information have been exploited to assign a emotion to the songs for which no emotion was found by the logistic model, by considering the most significant result for each genre, that is:

– Electronic → bittersweetness
– Hip Hop → frozeness
– Pop music → joy
– Rock music → excitement

3.6 Creation of an automatic playlist given user preferences about genres and mood

Once all the lyrics have been categorized, it can be created a playlist based on the preference of the user by asking her favourite genre and mood, returning the songs categorized with the correspondent labels. However this method is not the more appropriate. For example, if a user likes electronic music and select "envy" as emotion, the system would no return any results, as can be observed from the contingency table.

So it is better to create a recommendation system based on the similarity of the choices of the user and the features of the songs. The user will be asked a genre and to select her mood in the way it is presented in the following screens:

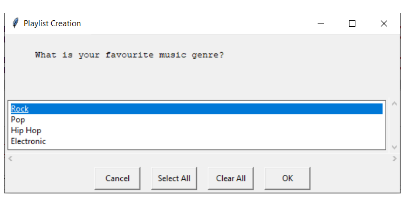

Fig. 7. User can select her favourite genre

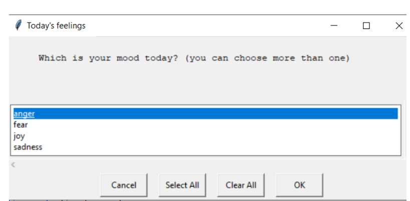

Fig. 8. User can select her fundamental emotions

The recommendation system implemented is based on the notion of *Jaccard similarity*. This measure is computed as:

$$Sim(A,B) = \frac{\mid A \cap B \mid}{\mid A \cup B \mid}$$

.

It is built a vector with the mood inputs of the user and it is calculated the jaccard similarity with the vector of emotions retrieved from the whole text of the songs in the dataset.

Two types of playlist are proposed:

- The first suggests the user with a playlist of ten songs of her favourite genre/genres she gives in input.
- The second suggests the user with a playlist of ten songs of other genres, different from the one she gives in input, with the idea of exploring new genres, but with high similarity with the mood of the user.
  In the example below, the user inserted rock and pop as favourite genres and angry and joy as mood.

artist	title	genre	jaccard_similarity
Queen	Don't Stop Me Now	Rock	1.000000
Iron Maiden	2 Minutes to Midnight	Rock	0.700000
Black Sabbath	War Pigs	Rock	0.700000
Fleetwood Mac	Rhiannon	Rock	0.500000
David Bowie	Space Oddity	Rock	0.500000
David Bowie	★ (Blackstar)	Rock	0.500000
David Bowie	Heroes	Rock	0.500000
David Bowie	Life on Mars?	Rock	0.500000
David Bowie	Starman	Rock	0.500000
Fleetwood Mac	Dreams	Rock	0.500000

Fig. 9. Playlist with the genre selected by user and mood similar to the one of user

artist	title	genre	jaccard_similarity
2Pac	Changes	Hip-Hop	1.000000
Marshmello	You & Me	Electronic	1.000000
Kendrick Lamar	Money Trees	Hip-Hop	1.000000
Lil Wayne	Love Me	Hip-Hop	1.000000
Justin Bieber	As Long As You Love Me	Pop	0.700000
Jennifer Lopez	All I Have	Pop	0.700000
Eminem	Godzilla	Hip-Hop	0.700000
2Pac	Dear Mama	Hip-Hop	0.700000
Ice-T	6 'N the Mornin'	Hip-Hop	0.700000
Aphex Twin	Milk Man	Electronic	0.700000

Fig. 10. Playlist of song similar to the mood of the user but genre different from the one given in input

References

1. Binali, H., Wu, C., Potdar, V. (2010, April). Computational approaches for emotion detection in text. In 4th IEEE International Conference on Digital Ecosystems and Technologies (pp. 172-177). IEEE
2. Author, F., Author, S.: Title of a proceedings paper. In: Editor, F., Editor, S. (eds.) CONFERENCE 2016, LNCS, vol. 9999, pp. 1–13. Springer, Heidelberg (2016). https://doi.org/10.10007/1234567890
3. Genius Lyrics website `https://genius.com/`.
4. Author, F., Author, S., Author, T.: Book title. 2nd edn. Publisher, plutchik's wheel of emotion, `https://en.wikipedia.org/wiki/Emotion_classification`