

Ranking Music Recommendations Using Hashtags

Aashik Ahamed A

Department Of Computer Science
St.Joseph's College Of Engineering
Chennai,India
ahamedajju.70@gmail.com

Dhanasekar K

Department Of Computer Science
St.Joseph's College Of Engineering
Chennai,India
mail2dhanasekar.kd@gmail.com

Dr.D.Rosy Salomi Victoria

Department Of Computer Science
St.Joseph's College Of Engineering
Chennai,India
drosysalomi@gmail.com

Abstract—The aim of the Project is to recommend music using hashtags. We recommend music using the emotions of the song. We analyze users emotion and their music choices. We extract information from hashtags by applying sentimental approach. The rawtext is preprocessed for mining and machine learning algorithm is applied. The data set is taken from Pandas and using NLTK we tokenize the words. We then combine sentimental computation and ranking of individual music track to recommend the music.

Keywords- Emotion of Music, Detection Of Sentimental Approach, Ranking, Recommendation, hashtags.

I. INTRODUCTION

In Recent Years, a number of people have been attracted to social medias like facebook, instagram, twitter. These social sites are used to express ones feeling and emotions .It brings out the personalities of an individual. People listen to music for different reasons Some of them include boredom, social cohesion, communication, emotional regulation, etc. We can determine users emotion using his musical preference . The emotional state can be considered as important information for Recommendation of music..

II. LITERATURE SURVEY

Many researches are trying to combine text mining and sentimental analysis. In sentimental analysis document level is widely used. Subjectivity and sentiment are properties of language. Phase Level is proposed by Theresa and it used to identify contextual polarity. Yi et al proposed polarity classification which is used to classify positive and negative sentiment. Document level was proposed by Turney. Wang et proposed supervised learning methods which has proven its worth in sentimental classification. It is based on large amount of labelled data which is time consuming and expensive. Semi Suprvised learning consists of small scale labelled data.

III. EXISTING SYSTEM

Existing system uses manual intervention and usage based suggestion. The recommendation of any service is based on the number of users who already requested similar services on the same demographic. The results were not always accurate, as they are not considering the individual users need on requirement. This systems assumes that the popular music played by majority of people were good for recommendation.

Drawbacks of Existing System

Current technology used in twitter for music recommendation system are usage based recommendation system which mostly recommends music based upon the majority of user's song request like a set of people requesting on services based upon their personal emotions and recommendations are usually done based upon music recommender system(MRS). The drawbacks associated with the current technology is that we can listen to songs based upon the number of people requesting the similar services and based upon their views the songs are recommended.

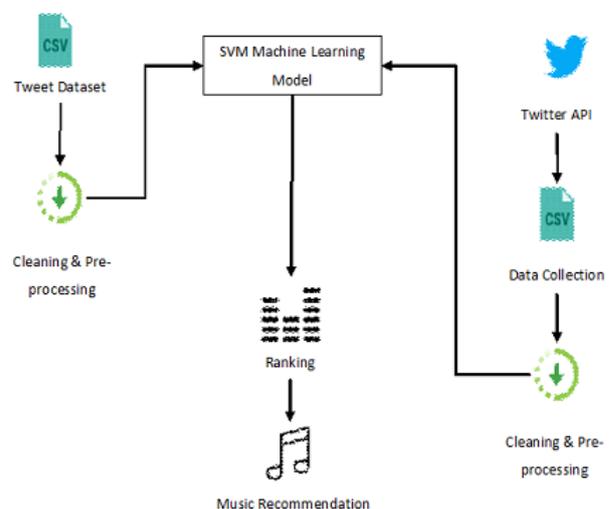
IV. PROPOSED SYSTEM

We propose a system where the contextual data of the user is considered for the recommendation. When a user tweets about his interested music, as hashtag, he also writes about his view. We try to analyze the sentiment of the content pertaining to the user. The sentiment value and the song the user is interested is considered together. The music is ranked based on the users view and their sentiment

Key Concerns

The Key concerns related to our proposed system involves just ranking and recommending music based upon the sentiment analysis made up by the users of the system as of like what type of mood and in what genre the music can be recommended to users of the system.

V. System architecture



VI. MODULES

- ✓ Data collection and pre-processing
- ✓ Tweet Sentiment
- ✓ Ranking and Recommendation

VII. MODULE EXPLANATION

A. Data collection and pre-processing



The objective of preprocessing is to clean the tweets.

1. Removing twitter handles (@user)

The tweets contain a lot of twitter handles which is used to acknowledge a tweet. We will remove all the handles as they do not convey much emotion about the state of the user.

2. Removing punctuations, numbers, and special characters

All these have to be removed in order the tweet gets cleaned and we find the emotion of the user.

3. Removing short words

We have to remove words which are very short like “hmm”, “oh”.

4. Tokenization

We will tokenize all the cleaned tweets in our dataset. Tokens are individual words and tokenization is a process of splitting a string of text into tokens.

5. Stemming

Stemming is a rule based process of stripping suffixes from a word. Example: player, plays, playing are different variations of play.

B. Pandas

Pandas is a library for python programming and is used for data manipulation. It offers data structures and operations for numerical tables. Its key data structure is called data frame and it is used to manipulate tables. Pandas is built on the Numby package.

Importing a dataset:

The function read_csv() is used to read a csv file. Lets import a dataset which consists of furnitures

```
>>>furniture=pd.read_csv("furniture.csv")
>>>furniture
```

Now a dataset consisting of Furniture is displayed.

C. Numpy

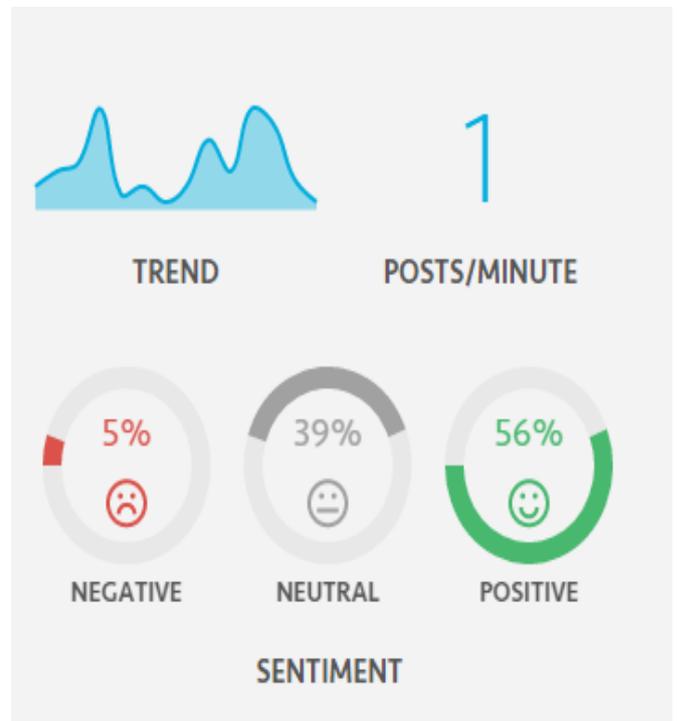
Numpy is used for adding multi dimensional arrays and matrices with a collection of mathematical functions.

D.NLTK

NLTK stands for Natural Language Toolkit. It is used for parsing, tokenization, semantic reasoning.

B. Sentimental Analysis

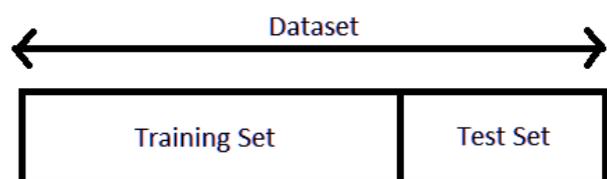
Sentimental Analysis can provide insights from social medias by detecting large volume of data. Sentimental analysis consists of 3 polarity classes. They are Positive, Negative and Neutral. The polarity of each tweet is determined by assigning a score from -1 to 1. We use Textblob, SentiWordNet to analyze the polarity .



C. Ranking and Recommendation:

Python Machine Learning:

We usually split data between 20-80% in training and testing stages. We split data into training data and testing data under supervised learning.



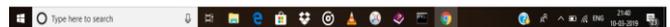
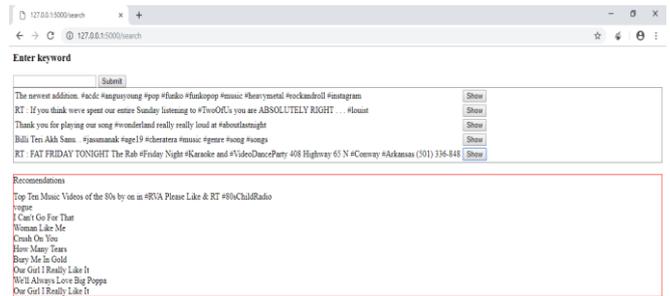
Python Flask:

Flask is a web framework of python. It is used for building complex ,database driven websites. Some names around python flask includes,

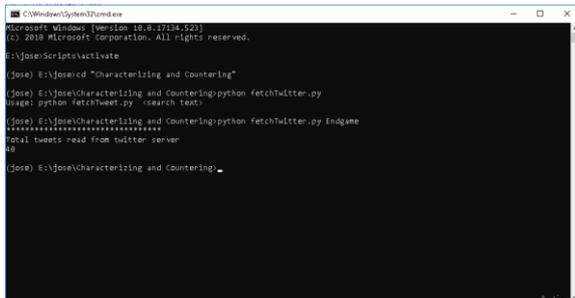
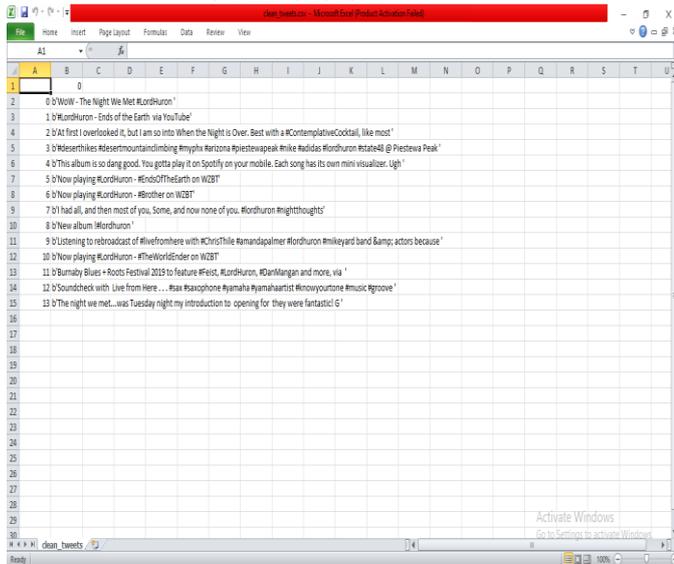
- **WSGI (Web Server Gateway Interface)**– A standard for Python web application development.
- **Werkzeug**– A WSGI toolkit implementing requests, response objects, and other utility functions.
- **Jinja 2**– A Python templating engine

The main purpose of ranking is listing the most suitable tracks for the user on top. The main building blocks for computing a ranking for a set of recommendation candidates are users, tracks and hashtags that are extracted from the graph. We implement different ranking like, user based ranking, track based ranking, twitter tags based ranking.

The Recommendation of music was carried out by combining sentimental computation and individual track.



SCREENSHOTS



VIII. CONCLUSION AND FUTURE WORK

The future concerns related to ranking and recommending music involves location based upon music recommendation by the developers. For people residing in particular place the recommendation of music are done.

IX. ACKNOWLEDGMENT

I hereby acknowledge that this work was carried out under the guidance of CSE department.

X. REFERENCES

- [1] Ms K . Nirmala Devi, MS. K. Mouthami, Dr. V. Murali Basakaran 'Sentimental analysis Based on textual reviews',2012.
- [2] Vapnik, V., 'The nature of statistical learning theory' pp.863-884,2000I .
- [3] Theresa Wilson, JanyceWiebe, Paul Hoffmann, 'Recognizing contextual polarity in phrase level sentimental analysis Vancouver,pp. 347-354 2004
- [4] Li, S., Wang., Zhou, G., Semi Sypervised Learning For Imbalance Sentimental Classification pp.1826-1831,2012.
- [5] Pang, B., &Lee, L., 'A sentimental education.Sentimental analysis usingsubjectivity summarization of minimum cuts'.