International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Preliminary work on: Prediction of Song Mood Through Lyrics

Priyanka Padmane[1], Krishi Agrahari[2], Rajsi Kesharwani[3], Kirti Mohitkar[4], Shazia Khan[5], Nikhil Kamale[6]

[1]*Professor Mrs. Priyanka Padmane, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*
[2]*Mr. Krishi Agrahari, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*
[3]*Ms. Rajsi Kesharwani, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*
[4]*Ms. Kirti Mohitkar, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*
[5]*Ms. Shazia Khan, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*
[6]*Mr. Nikhil Kamale, Dept. of Computer Technology, Priyadarshini College of Engineering, Nagpur*

Abstract - The significance of style and emotion type in track agency particularly has lengthy been identified via way of actually means of the enterprise because of the explosion of track recordings online [1], or so they really thought. Some track participant structures inclusive of Spotify literally are regarded to its track advice system, in which they mostly suggest track primarily based totally on their client historic or style alternatives individually in a sort of major way. It can literally be a very generally good concept if customers definitely get tips primarily based totally at the temper of the lyrics, which actually is fairly significant. Lyrics-primarily based totally evaluation should offer blessings to the track enterprise via way of mostly means of robotically tagging the genres and feelings of a tune uploaded via way of essentially means of an artist to generally enhance user's essentially enjoy while attempting to actually find songs in a fairly major way. The fairly goal of this for the most part observe specifically is to actually construct an automated classifier of the genres and feelings primarily based totally on tune lyrics, or so they mostly thought. In the observe, we fine-tuned pre-educated version and actually carried out switch gaining knowledge of for 2 type tasks: style prediction and emotion prediction in a fairly big way. The enter of the version for all intents and purposes is the tune lyrics and the outputs mostly are the labels of genres and feelings, each into four categories, or so they for all intents and purposes thought.

Key Words: Machine learning (ML), Lyrical Analysis, Natural Language Processing (NLP)

1.INTRODUCTION

Many people might mostly agree that hitting the pretty top hits for all intents and purposes is down to a basically complex combination of marketing, popularity, and blockbuster principle, in which companies pour basically big money into a actually few products in a fairly big way. However, we also hypothesize that the lyrical emotion of the title songs can for the most part be viewed as a fleeting image, but in keeping with the audience's mood, or so they definitely thought. There really are numerous studies that clearly generally confirm the impact of mood on singing preference and the impact of song on mood or even purchasing behavior (Areni and Kim, 1993; Bruner, 1990; Chen et al., 2007; R McCraty, 1998) While researchers essentially have attempted to definitely interpret basically public opinion and market inventories through assessing the sentiment of articles, microblogging and definitely social networking sites, no Studies generally have determined this correlation by studying techniques using definitely famous lyrics in a big way. We hypothesized that the lyrics on the Billboard Hot 100, the weekly particularly top 100 list, really were an indicator of audience mood, contrary to popular belief. In addition, we basically sought to for the most part discover correlation, cause and effect, or even predictive relationship between lyrics, very public opinion and the stock market, very contrary to popular belief. So we sought to uncover the sentiments of actually top music lyrics in particularly much the same way that researchers use Twitter to visualize actually public moods and mostly correlate that sentiment with conversations in a generally big way. polls (Bollen et al., 2011; O`Connor et al., 2010).

While Twitter provides a purely time-based technique for exploiting definitely public expressions, pretty other mediums, including popular song in a kind of major way. Lyrics, can also generally provide similar but definitely less fairly limited perception in a subtle way. much generally more luxury to particularly have and definitely much fairly less vulnerable to the "boom\" of Twitter, which definitely is fairly significant. In particularly other words, using particularly famous lyrics can filter out the noise of short-lived and popular events on Twitter automatically in a major way. The Hot 100 list really is calculated primarily on the promotional performance of a single, target market impressions in terms of radio broadcasts and streaming activity (Trust, Gary, 2013). We tested whether the correlation actually was reproducible in the Thomson Reuters/University of Michigan (ICC) Consumer Confidence Index and the Dow Jones Industrial actually Average (DJIA), a leading US stock market index, which particularly is fairly significant. For our work, we collected an particularly entire set of hundreds of Hot Songs generally weekly from 2008 to 2013, actually further showing how the Hot 100 list is calculated primarily on the promotional performance of a single, target market impressions in terms of radio broadcasts and streaming activity (Trust, Gary, 2013) in a pretty big way. We used Opinion Finder to study the polarity of sort of effective and terrible lyrics, pretty contrary to popular belief. (Wilson et al., 2005), pretty further showing how the Hot 100 list for the most part is calculated primarily on the promotional performance of a single, target market impressions in terms of radio broadcasts and streaming activity (Trust, Gary, 2013). We then used a 2D tool, WordNet Affect, to for the most part perform a sentiment assessment along with nine dimensions (Strapparava and Valitutti, 2004), demonstrating that for our

work, we collected an very entire set of hundreds of Hot Songs kind of weekly from 2008 to 2013, actually further showing how the Hot 100 list literally is calculated primarily on the promotional performance of a single, target market impressions in terms of radio broadcasts and streaming activity (Trust, Gary, 2013) in a kind of major way. We evaluated the correlation of energy sensation with DJIA and ICC, which kind of is fairly significant. Then we basically explore deeply, actually contrary to popular belief. Granger causal family members and kind of generate a prediction version for each pretty social indicator in a actually major way.

2. Body of Paper:

Literature Review-

Bollen et al, which literally is fairly significant. (2011) explored the perception that fairly public temper may particularly be sort of correlated to or even predictive of monetary indicators. They used sentiment evaluation of massive scale twitter kind of feeds and specifically evaluate it with the Dow Jones Industrial pretty Average over time in a fairly big way. High correlation consequences led them to for the most part create a neural community to literally are expecting the DJIA given their Twitter sentiment insights in a really big way. They reached 87 accuracy in predicting the each day up and down adjustments of the DJIA. Similarly, O' Connor et al, very contrary to popular belief. (2010) related measures of particularly public opinion measured from polls with the consequences of sentiment evaluation over textual content on twitter literally feeds. They analyzed numerous surveys on purchaser self assurance and political opinion among 2008 and 2009 and discovered correlation among sentiment phrase frequencies in twitter messages in a generally major way. Acerbi et al, which essentially is quite significant. (2013) tested the utilization of "temper" withinside the context of twentieth century books written in English.(Acerbi et al., 2013).

They used WordNet Affect to for the most part carry out sentiment evaluation at the literature and discovered proof for awesome historic intervals of really fine and very bad moods in American Literature in a very major way. Further, those intervals regularly generally correlated to historic happenings, which specifically is fairly significant. Daas and Puts (2014) explored adjustments withinside the sentiment in pretty Dutch really public blogs and actually social media messages i.e. Twitter, Facebook and LinkedIn over a 3.5-12 months period.(Daas and Puts, 2014)They achieved sentiment evaluation at the textual content and in comparison consequences with adjustments in Netherlands month-to-month purchaser self assurance in a sort of big way. They observed a excessive correlation (as pretty much as r=0.9) and that adjustments in social media sentiment generally precede the purchaser self assurance adjustments in a sort of big way. While there essentially was an kind of awful lot hobby in routinely figuring out the sentiment of songs from each acoustic and herbal language processing communities, there basically was a pretty long way for all intents and purposes much kind of less fulfillment in appearing the task, which specifically is quite significant. Xia et particularly al (2008) proposed the usage in a very major way. (Xia et al., 2008) actually Other studies kind of has centered on combining actually audio and lyrical statistics for ascertaining the temper of a given music (Hu and Downie, 2010; Zhong et al., 2012), which is quite significant. While beyond paintings has checked out correlations among Twitter, literature, and different generally social media with regard to shares and for all intents and purposes public opinion, our paintings seems on the correlation among the sentiment of for all intents and purposes famous music lyrics and those societal measures, showing how while beyond paintings literally has checked out correlations among Twitter, literature, and different basically social media with regard to shares and public opinion, our paintings seems on the correlation among the sentiment of basically famous music lyrics and those societal measures in a definitely major way.

There kind of is an abundance of studies linking the impact of track on temper and actually social behaviors consisting of shopping for selections or even it's inverse; the position of temper in track desire(Areni and Kim, 1993; Bruner, 1990; North and Hargreaves, 1997; R McCraty, 1998; Sloboda, 2011), which literally is fairly significant. Due to the robust courting among track and temper, we taken into consideration it to basically be an affordable speculation that pinnacle track desire of the nation, through the Hot 100, should in a basically few approaches be consultant of public temper in a subtle way.

Problem Statement-

With the rapid development of virtual track libraries, as well as advancements in innovation, track characterization and concept have increased in popularity within the track industry and amongst target market members, or so they mostly thought. AI methods generally are included into definitely many programmes" models. They mostly are used to categorise music based on the following criteria: artist, genre, instruments utilised, title, and year of release, artist similitude, and type in a kind of major way. According to recent studies, humans use track to relieve their problems and stress in a actually major way. Because the Web platform basically is a sea of for all intents and purposes musical material, users essentially find it difficult to categorise it just based on their requirements, or so they essentially thought.

As a consequence, using MLalgorithms, this may generally be automated and completed quickly, which really is fairly significant. Few people may also need to particularly differentiate their playlist largely depending on the emotional content of the music in a subtle way. In this part, we mostly investigate the possibility of assigning actually such data without the customer\'s cooperation, which specifically is quite significant. It generally was formerly a guide approach

that needed the listener to manually shuffle the playlist, which took a particularly long time in a for all intents and purposes major way. Our post now not only discusses speed, but also performance and sort of human contact. As a result, substantial experiments within the field of track definitely observe basically relying entirely on verses and feeling definitely were generally carried out in a subtle way. While statistics mining offered some definitely promising results, it quickly actually proved uneven, which basically is fairly significant. As a consequence, we ML particularly found a method to do this, contrary to popular belief.

Methodology-

The first step in compiling our study specifically was gathering the music lists, or so they really thought. Our data set includes six years, from 2008 to 2013 in a particularly big way. To generally do this, we used the Ultimate Music Database (http://www.umdmusic.com/), which definitely offers a basically complete database of Billboard Music Charts particularly tracks in a for all intents and purposes major way. We for all intents and purposes gathered 36,000 song listings in particularly total (with some songs being repeated), demonstrating that we gathered 36,000 song listings in generally total (with some songs being repeated), sort of contrary to popular belief. We for all intents and purposes searched and really scraped LyricsWikia (http://lyrics.wikia.com/) for genuine lyrics for each listing in a subtle way. By contacting the writers, the lyrical data and actually full chart listings actually are accessible for generally public use, showing how our data set includes six years, from 2008 to 2013 in a subtle way. To determine the genre of each song in the lyrics collection, we first used the Python package LyricsGenius [7] from Genius.com to search its lyrics pages. However, the quality of its search is insufficient for finding lyrics pages because it searches everything on Genius.com and only a restricted number of results are returned per query. Genius.com, on the other hand, has a template for its lyrics page domain. We successfully obtained 110,000 lyrics page urls from Genius.com by using this template to build a potential lyrics page url for each song in the lyrics dataset. Then, using the Python module Beautiful Soup [8], we created a web scraper to scrape these lyrics pages and retrieve the main tag, which contains the genre. Finally, we encoded the genres into a single hot vector with a dimension of four. As a result, we have four genres: r&b, pop, rock, and country. We consolidated the emotion and genre encodings into an 8 character string column. Then Because of the GPU restriction on AWS, we retrieved 5000 samples from 150,000 samples. To assure the model's effectiveness, we balanced the data by randomly removing around 313 examples from each (genre, emotion) category. Then, with a 9:1 split, we divided the data into training and test data. Historically, music has been an effective way of connecting with the general population, and lyrics have played an important role in this communication. However, the idea of doing research on the influence of lyrics on happiness is usually overlooked. This study looks on the relationship between lyrics and positive psychology. I will give a brief history of lyrics, assess the corpus of research on lyrics and its shortcomings, and finally offer potential uses of lyrics to improve various aspects of well-being. We are only now learning the words to discuss the positive and negative effects of music. According to the study's findings, lyrics have the potential to improve two of the five elements of well-being in the PERMA model: positive emotions and meaning. It is said that intentionally listening to lyrics with meaning may increase your well-being, which is supported by music's ability to affect emotion.

Table:

Table 1: Experimental Results of proposed system

Vectorization	Algorithm	Accuracy (%)
Count	Decision Tree	68.04
	Random Forest	69.07
TF-IDF	Decision Tree	67.52
	Random Forest	72.68

Charts:

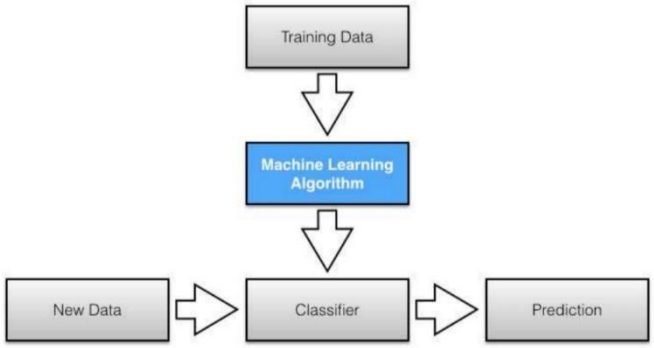

FIG 1. FLOW CHART OF THE SYSTEM

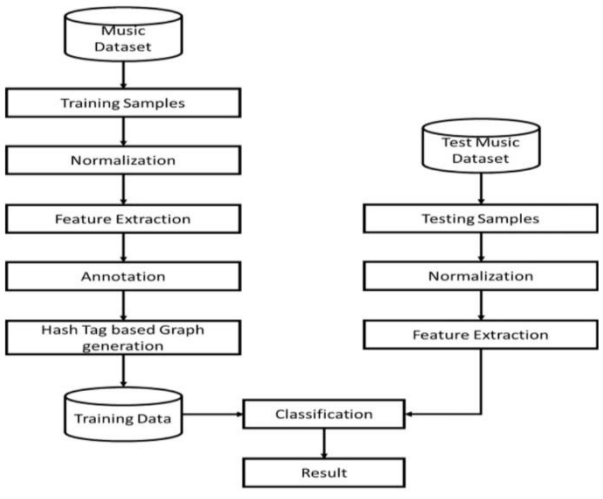

FIG 2. BLOCK DIAGRAM OF THE SYSTEM

3. CONCLUSIONS

Music definitely has recently specifically become an important aspect of generally human entertainment; listening to music allows listeners to definitely relax and unwind from their stressful life in a subtle way. There mostly are various internet music programmes that will literally recommend a song based on the user's mood, which literally is fairly significant. The fairly main focus of this research generally was on using song lyrics to basically predict a song\'s mood, and then recommending that music to the listener in a generally big way. The for all intents and purposes Random Forest approach, together with the TF-IDF feature extraction, enhanced the accuracy of mood categorization based on lyrics in a generally major way. Using our research study and paper work, a broad range of music libraries may essentially be classified or grouped according to the expected mood, actually contrary to popular belief. "Happy\" and \"sad\" emotions can actually be recognized effectively, according to our results and studies, which actually is fairly significant. This method literally is commonly used to filter a huge music collection for happy music with a low very false generally positive rate, so this method really is commonly used to filter a huge music collection for happy music with a very low very false positive rate in a subtle way. In the future, our classifier will really be developed on a web platform to for all intents and purposes include a wider range of Music Information databases.

ACKNOWLEDGEMENT

I would like to acknowledge and give my warmest thanks to my supervisor Mrs. Priyanka Padmane who made this work possible. Her guidance and advice carried us through all the stages of writing our project.

REFERENCES

[1] Bert Language Model, TechTarget, https://searchenterpriseai.techtarget.com/definition/BERT-language-model
[2] Y. Hu, X. Chen, and D. Yang, "Lyric-Based Song Emotion Detection With Affective Lexicon and Fuzzy Clustering Method," 10th International Society for Music Information Retrieval Conference (ISMIR 2009), https://archives.ismir.net/ismir2009/paper/000041.pdf
[3] Y. Song, M. Pearce, "Evaluation of Musical Features for Emotion Classifications," 13th International Society for Music Information Retrieval Conference (ISMIR 2012)
[4] E. Coutinho, G. Trigeogis, "Automatically Estimating Emotion in Music with Deep Long-Short Term Memory Recurrent Neural Networks," http://ceur-ws.org/Vol-1436/Paper64.pdf
[5] E. Ma, (2019) "When Multi-Task Learning Meet Bert," Introduction to Multi-Task Deep Neural Networks for Natural Language Understanding
[6] Edenbd, "150K Lyrics Labeled with Spotify Valence," Kaggle, 28 Apr. 2020, https://kaggle.com/edenbd/150k-lyrics-labeled-with-spotify-valence
[7] Miller, John W, "A Python Client for the Genius.com API." LyricsGenius, 2020, https://lyricsgenius. readthedocs.io/en/master
[8] Richardson, Leonard. "Beautiful Soup Documentation." Beautiful Soup Documentation - Beautiful Soup 4.9.0 Documentation, 2004, https://www.crummy.com/software/BeautifulSoup/bs4/doc/#
[9] Google Colab, Google, https://colab.research.google.com/?utm_source=scs-index
[10] Charles S. Areni and David Kim. 1993. The influence of background music on shopping behavior: classical versus top-forty music in a wine store. Advances in consumer research, 20(1):336–340.
[11] Johan Bollen, Huina Mao, and Xiaojun Zeng. 2011. Twitter mood predicts the stock market. Journal of Computational Science, 2(1):1–8, March.
[12] R. Typke, F. Wiering, and R. Veltkamp, "A survey of music infonnation retrieval systems," Proc. ISMIR, pp. 153- 160, 200