ORIGINAL ARTICLE

Emotional sentiment analysis of social media content for mental health safety

Ferdaous Benrouba[1] · Rachid Boudour[1]

Abstract
Text sentiment analysis is mostly used for the assessment of the author's mood depending on the context. The purpose of sentiment analysis is to discover the exactness of the underlying emotion in a given situation. It has been applied to various fields, including stock market predictions, social media data on product evaluations, psychology, the judiciary, forecasting, illness prediction, agriculture, and more. Many researchers have worked on these topics and generated important insights. These outcomes are useful in the field because they (outcomes) help people comprehend the general summary quickly. Additionally, sentiment analysis aids in limiting the harmful effects of some posts on various social media sites such as Facebook and Twitter. For these reasons and more, we are proposing an approach to filter the social media content that could be emotionally harmful to the user, through getting the social networks content; for that, we have used Twitter API to get the user posts, and then, we have used API natural understanding language API tool to extract and classify the emotions of the Twitter content into five basic emotional categories—Joy, sadness, anger, fear, disgust—into an array of emotions; after that, we have defined a perfect emotion array from over 450 words from the English language. The main purpose of this comprehensive research article is to examine the proposed solution that we have conducted to improve the quality of content displayed to users emotionally.

Keywords Sentiment analysis · Twitter API · Artificial intelligence · IBM API · Natural language processing

1 Introduction

Excessive use of social media has created imbalances in people's mental health and has become a major problem for this generation, scientists, and society (Boivin et al. 1995). This excessive use of social media is one of the most popular hobbies of today's generation. Social media can be in any form of website and application, from chat platforms like Facebook and Instagram, to online gaming and virtual worlds like YouTube and Reddit. These modern sites are expanding rapidly and serving as accessible gateways for communication and enjoyment for the young population. Virtual social media platforms such as Facebook, Twitter, and others have greatly enhanced the virtual environment over the last

decade by allowing users to share their thoughts, ideas, personal information, images and videos at an unprecedented rate (Chou et al. 2011). Moreover, as the number of daily users is increasing, social media users have quickly accepted online social communication as an integral element of daily life (Davila et al. 2012). In fact, since 2015, Facebook has counted an estimated 1 billion active members. This implies that one in seven individuals on the planet is using Facebook at least once in a single day. As a result, social media has a significant impact on several elements of modern digital life, including marketing, politics, education, health, and fundamental human connection. Even though the social media phenomenon is still relatively young, some empirical studies have analyzed the overall impact of frequent usage of social media on well-being and users' mental health. This lack of awareness is particularly concerning in the context of today's younger generation, as teenagers and young people spend a significant amount of time online socializing and so are at a higher risk of harmful consequences. Some studies have identified a link between social media use and negative consequences such as increased anxiety (Deters and Mehl

✉ Ferdaous Benrouba
   ferdaous.benrouba@univ-annaba.org

   Rachid Boudour
   rachid.boudour@univ-annaba.dz

1 Computer Science Department, Badji Mokhtar University,
   BO 12, 23000 Annaba, Annaba, Algeria

2012), stress (Dick 2013), depression (Drouin et al. 2012), loneliness (Davila et al. 2012), and eating disorders. This usage of social media by the younger generation generates major concerns about its negative consequences. For these reasons and more, we are proposing an approach to filter the social media content that could be emotionally harmful to the user, by getting the social media content; for that, we have used Twitter API to get the user posts, then we have used API natural understanding language API tool to extract and classify the emotions of the Twitter content into five basic emotional categories—Joy, sadness, anger, fear, disgust—into an array of emotions, and after that, we defined a perfect emotion array from 450 words from the English language. Then, we calculate the Euclidean distance (similarity distance) between the emotions array of each tweet content and the perfect emotions array; after that, we define a value called the threshold which will be 1.08; if the distance is equal to or less than the threshold, it is ok to display, otherwise display a warning message that the content could be emotionally harmful. This paper is organized as follows: Section 2 describes the related works and some previous research, Sect. 3 gives an overview of the proposed approach, and in Sect. 4, we discuss some results. Finally, Sect. 5 concludes the paper and outlines some future work. For all of these reasons, we are proposing an approach to filter the social media content that might be emotionally harmful to the user, To do this, we first obtained Twitter user posts by using Twitter API, then we have used IBM natural language understanding API to extract and classify the emotions of the Twitter content into five basic emotional categories—Joy, sadness, anger, fear, disgust—into an array of emotions, moreover, we defined a perfect emotion array from 450 words from the English language.

2 Literature review

In this section, we explore original research on the topic of emotion detection using artificial intelligence filters. In emotion detection, researchers have proposed many emotion models to understand emotions and the expressions behind emotions. Therefore, emotion models play an important role in text-based emotion detection on social media platforms. There have been several general approaches developed in the last 10 years regarding text-based emotion detection, we can list some of these attempts as follows: Boivin et al. (1995) explained the keyword-based emotion recognition approach. Each sentence is considered as a combination of content words or Emotional Function words (EFW). They can take three forms: emotional keywords, modifiers, and metaphors. Emotion keywords can take six emotion labels from Ekman's emotion model and assign specific weights. Chou

et al. (2011) presented an emotion classification system for tweets in the English language. Initially, text preprocessing has performed using word and sentence tokenization, stemming, lowercasing, and POS-tagging. Next, feature extraction is accomplished using n-gram features, lexicon features, and various semantic and syntactic features. Then, to solve the multi-class multi-label problem, a set of eleven binary classifiers was created for each possible emotion class, anger, anticipation, disgust, trust, fear, love, joy, pessimism, optimism, surprise, and sadness where each model gets the previous models' predictions as supplementary features to create a multi-label representation of the predictions. Davila et al. (2012), Deters and Mehl (2012), Dick (2013) and Drouin et al. (2012) studied deep Learning-based approach; deep learning is a variant of machine learning in artificial intelligence with networks capable of unsupervised learning from unstructured or unlabeled data. This approach allows neural networks to learn complex concepts by constructing them from simpler concepts. Initial preprocessing was performed on the data set. After that, the embedding layer is built, where tokens are represented in the form of numbers. Then, depending on the number of emotion labels, these feature vectors are input into one or more Deep Neural Network layers. Patterns are learned from data and used to predict the labels by using classification. Souri et al. (2018) proposed a deep learning system, for challenges with multi-label emotion identification problems, microblogging, preprocessing, they used the Ekphrasis tool. Word normalization, tokenization, spell correction, and segmentation are all performed by Ekphrasis. GloVe, a pre-trained word embedding algorithm, was used. The features from the embedding layer are provided to two Bidirectional-Gated recurrent unit layers. Then, the output of the first Bi-GRU layer is given to the first attention layer. The output combined from the first and second Bi-GRU and embedding layer is given to the second attention layer. Then, combined two attention layers are provided to a DNN with a sigmoid activation function for classification. The authors used 11 emotion categories to classify emotions: anger, anticipation, disgust, trust, fear, love, joy, pessimism, optimism, surprise, and sadness. Neural network model as a solution to filter negative emotion on social media platforms: In the neural models, unique words are utilized as the input in the parse trees which provide the synthetic data and semantic data. Hence, emotion composition is derived from the best. Recurrent neural networks and convolution neural networks are becoming more popular, and they do not require parse trees to split their features from the given sentences. Instead, recurrent neural networks and convolution neural networks (CNNs) use word embedding as inputs that already inscribes the semantic and synthetic data. Additionally, the architecture of the convolution

neural networks and recurrent neural networks helps in learning the connectivity between the words in a statement (Morente-Molinera et al. 2019). The recursive automatic coding network (RAN) in the semi-supervised model for the sentence-level, sentiment analysis provided a low-dimensional vector representation. Risch and Krestel (2018) proposed a new matrix-vector recursive neural network (MVRNN), each context is also associated with a tree representation of the matrix (Subramani et al. 2018). The structure of the tree is derived from the parser that is used externally. It describes a collaboration between the Recurrent Neural Network (RNN) and CNN architecture for sentiment classification from the short context that are privileged in the coarse-grained features generated by a CNN (Subramani et al. 2019). An approach based on the linguistics LSTM for effective sentiment prediction incorporates the sentiment lexicon as both a highly intensive context and negative context. In Abualigah et al. (2021), the LSTM includes these features during analyzing the sentiment to provide a useful view of the context. Authors have presented a traditional CNN/LSTM model which consists of two sections—local CNN and LSTM—to predict the attitude illustrated. Z Mo, E Sam - Journal of Student Academic Research, 2021, used biometric data to detect the emotional state of social media users as the first step toward early anxiety detection and intervention card against anxiety. Abualigah (2019) presented a personality classification model by analyzing user activities on Facebook using machine learning. The proposed method uses Facebook API to collect the data from 100 users. The proposed method also can recommend friends in Facebook groups. As many researchers reported on sentiment analysis (Ahmed and Abdelhamid 2022; Shaha and Amal 2022; Mountaga et al. 2021; LEUNG 2009), Ahmad et al. (2019) leveraged a sentiment analysis technique to address group decision-making using social media. The proposed method helps experts to generate preference relations which may be applied to group decision-making. Budiharto and Meiliana (2018) proposed a deep learning-based model to identify the aggregation of user-generated content in social media. The proposed method uses a recurrent neural network based on a bidirectional gated recurrent unit. They used a machine-translated to augment a dataset used for model training. Prabhu et al. (2019) proposed a domestic violence identification system using deep learning. The proposed system uses a dataset collected from Facebook. The proposed system uses a binary text classification approach that can detect if the content created by a user is recognized as critical or uncritical. Cury (2019) also presented a deep learning-based domestic violence identification system. Unlike their previous work (Drouin et al. 2012), this system can be applied for a multi-class post categorization including five categories, namely general,

empathy, awareness, personal story, and fundraising. The proposed system was evaluated using a dataset collected from Facebook users' generated content and achieved a high accuracy rate. Al Shehhi et al. (2019) used a deep learning model with a sentiment analysis technique to classify user-generated content on Twitter (tweets) as a binary classification (extremist tweets or non-extremist tweets). In their work, Subramani et al. (2019) used sentiment analysis techniques to predict the result of the Indonesian presidential election. The prediction of the proposed method was correct according to the result of the Indonesian presidential election. Pong-inwong and Songpan (2018) used sentiment analysis techniques to analyze tweets collected from Twitter users in the United Arab Emirates (UAE) to measure happiness. They used tweets in English and Arabic. They found that 7:00 am is the happiest hour and Friday is the happiest day. In Aloufi and Saddik (2018), the authors proposed a sentiment analysis system that can be applied to analyze posts in the teaching evaluation system. The proposed system collects student feedback and analyzes sentiment analysis phrases to obtain the classification of teaching attitude. Arias et al. (2013) proposed a sentiment analysis approach to analyze football fans' sentiments through tweets posted by them and their reaction to game events (i.e., penalty kicks, scoring goals, etc.). The proposed approach can conclude fan interaction through football game events. Jansen et al. (2009) proposed a toxic detection model based on a convolutional neural network (CNN), bidirectional gated recurrent units (GRU), and bidirectional long short-term memory (LSTM). They used the Wikipedia dataset to evaluate the proposed method and achieved an F1 score of 87.2% for predicting toxicity types and 82.2% for the toxic/non-toxic classification. In Abualigah et al. (2021), the authors presented a personal attack analysis method based on a combination of machine learning and crowdsourcing. In Abualigah et al. (2021), the authors proposed a deep learning-based approach to classify toxic comments. They used Kaggle data to test their approach. Sentic computing, a sentiment analysis paradigm that makes use of AI and Semantic Web approaches, has been presented by the authors in Ji et al. (2022) to improve the recognition and interpretation of sentiment in natural language processing. A chained-fusion approach is provided to jointly learn face dynamics and appearance in Ansari et al. (2022) authors' proposal. They have demonstrated how such sequential fusion might offer a probabilistic perspective on the model complementarity and correlation between two various data modalities for improved depression recognition. The authors of Ji et al. (2020) have suggested using relation networks to identify mental illnesses and suicidal tendencies along with associated risk markers. In Han et al. (2022), authors employed two sets of methods—hybrid

and ensemble—to early diagnosis depression. The results indicated that numerous feature combinations and careful feature selection can lead to higher performance, and Ji et al. (2020) provides a survey of the techniques currently utilized to identify depression and its shortcomings. A model that makes sense has been put forth by the authors of Han et al. (2022) for identifying depression on Twitter. It includes a unique encoder that combines feed-forward neural networks with hierarchical attention methods. Google and Jigsaw presented an API (Perspective API, https://www.perspectiveapi.com/ (accessed on 10 February 2021)) for classifying toxic comments. However, this API handles a simple binary text. Moreover, another study addressed the classification of toxic comments presented by Ringsquandl and Petkovic (2013). This study also used a simple binary text classification to test the proposed model.

3 Methodology

In this section, we will describe our proposed solution and examine what we have conducted to improve the quality of content displayed to the user emotionally through getting the content of the social networks; for this, we have used the Twitter API to get the user posts, analyzing and classifying the content into five basic emotional categories: Joy, sadness, anger, fear, disgust. For this, we have used IBM understanding language API tool, an online open source API for textual emotion recognition; after that, we output the dataset into an array of emotions; then, we have defined a perfect emotion array extracted from 450 words from English language; then, we calculate the Euclidean distance (similarity distance) between the emotions array of each tweet content and the perfect emotions array; after that, we define a value called the threshold which will be (1.08); if the distance is equal or less than the threshold, it is ok to display, otherwise display a warning message that the content could be emotionally harmful.

3.1 Data preprocessing

For the first step of building our application, we needed data from social media, so we used the Twitter API "tweets" to get all the data that we needed to proceed with the analysis; then, to analyze and classify content into five basic emotion categories: Joy anger disgust sadness and fear, we used IBM natural language understanding API (Komal and Joshi 2016).

3.2 IBM natural language understanding API tool

IBM API tool is used to analyze various features of text content at scale. Provide text, raw HTML, or a public URL and IBM Watson Natural Language Understanding will give you results for the features you request. The service cleans HTML content before analysis by default, so the results can ignore most advertisements and other unwanted content. With Natural Language Understanding, we can do a full emotional analysis which includes a set of text analytics features that you can use to extract meaning from unstructured data (Komal and Joshi 2016).

3.3 Sentiment analysis of tweet content

To perform a sentiment analysis on the dataset (tweets) we got from the Twitter API, we have used IBM natural language understanding Api to extract the emotions from about 450 words from the English language that people usually use to express positive emotions with every tweet presented as an emotion array $X[i, j]$ for further analysis. After getting data of every tweet is presented in form of an array of emotions $X[i, j]$ from the IBM natural language understanding Api, the content will be analyzed in the next step.

3.4 Emotion analysis of tweet content

For this step and in order to filter content, emotionally we have collected more than 450 words that people usually use to express positive emotions in English language. Then, we have analyzed them using the IBM API, the output that we got is going to be defined as the perfect emotions array, and its values are as next:

```
"emotion": {
    "sadness": 0.005441,
    "joy": 0.997533,
    "fear": 0.009288,
    "disgust": 0.001843,
    "anger": 0.002108
}
```

We defined a perfect emotion array as a matrix Y, where Y is a $m * n$ matrix, where each row represents a sample and each column represents the 5 emotions, respectively.

3.4.1 Evaluation of emotion

In order to classify the emotion array we got from our given tweet in the dataset, we will compute the Euclidean distance between every tweet in the dataset array $X[i, j]$ from the tweet data content and the perfect emotion array $Y[i, j]$

3.4.2 Euclidean distance

Euclidean distance is a measure of similarity and dissimilarity, the distance between two vectors X and Y is defined as shown in equation 1, in simple words, the Euclidean distance is the sum of squared differences between the similar elements of the two vectors X and Y, the reason behind taking the Euclidean distance is that it proved its effectiveness for text classification and clustering (see Cambria et al. 2010 for more details), and also, it is one of the most common distance metrics and acts as an inverse similarity function. To classify tweets as emotionally safe or harmful content, we set a threshold of (1.083274218344552), and we have proceeded with this value practically after the experimentation results process that we have made:

- First, we have collected more than 800 words that can be used to express negative emotions.
- Next, we analyzed the words, to get the following output

```
{"sadness": 0.268492,
"joy": 0.006753,

"fear": 0.296736,
"disgust": 0.012359,
"anger": 0.201871}
```

- The next point in the process is to calculate the Euclidean distance between the perfect emotion array and the output that we get from the previous step; this value is going to be defined as the threshold.

The value that we have obtained (1.08) expresses how much the contents are different, as we have previously mentioned the greater the distance, the greater the difference (check Eq. 1, E.D stands for Euclidean distance). Now if the distance is equal to or less than (1.08), then we assume that the tweet content is emotionally safe. Otherwise, we display a warning message that this content could be emotionally harmful.

$$E.D(x,y) = \sqrt{\sum_{i=1}^{d}(x_i - y_i)^2 + \sum_{j=1}^{d}(x_j - y_j)^2} \tag{1}$$

Figure 1 illustrates in detail the proposed approach.

 We built the application using Nodejs and Angular framework and made it available on Github, you can find the source code and installation instructions in Yuan et al. (2016).
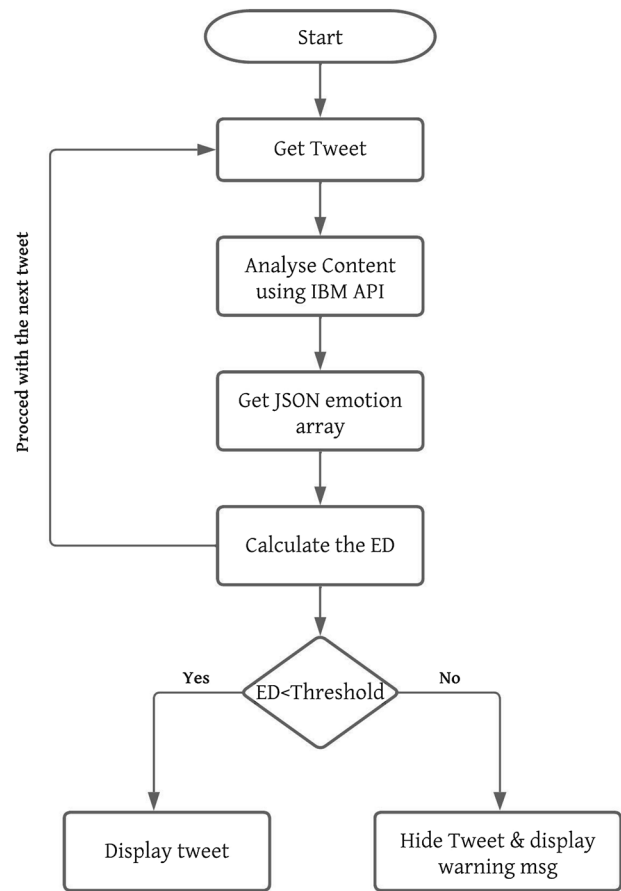

Fig. 1 Flowchart of the proposed approach

4 Results and discussion

Text emotion detection is one of the hardest tasks in natural language processing because even humans struggle to analyze sentiments accurately, here are some of the limitations that we faced when we were building the application:

- The language and Intentions: We faced a lot of problems concerning languages because the intention of the tweets sometimes varies on what real means.
- The media content: Text is not the only content in social media, there are also photographs and videos that can be emotionally harmful in the same way as text.

Furthermore, there are certain threats to the validity of our approach primarily in inherent differences between the compared tools. First, the method of calculating the Euclidean distance, which could not have been avoided, might have produced certain inconsistencies. Second, the nature of this tool type itself might have affected the results. IBM tool does not treat neutral as a special type, so we had to determine a threshold which was 1.08, the dataset combined

the emotions of anger and disgust into one emotional type (label). Neviarouskaya et al. had a similar problem with certain inconsistencies in the emotional model. Most IBM tool errors come from cases where additional context is required for the correct interpretation of textual emotion. This emphasized the need to include word sense disambiguation in our future research. Further research is needed to determine how much the difference between the data set and the second data set emotion (the difference sometimes is very small) affects IBM's final accuracy. The third kind of error is due to the vocabulary and style used in the language of fairy tales which is, arguably, significantly different from the contemporary use of language. These errors might be reduced by expanding our lexicon, but it is uncertain whether this would make the algorithm sufficiently powerful to take it out of the context of fairy tales and into the "wilderness" of the Web.

We have gathered a collection of tweets that we intuitively thought might hurt users emotions, some of which we included in Table 1(We have done the same process with content that we thought might be emotionally positive to certain users, results are included in table 2). The results demonstrate the efficiency and effectiveness of our approach in filtering any suspicious content that may cause emotional harm to end users and by consequence protect users mental health to a certain extent.

According to the results shown in the table below, we have noticed that the Euclidean distance calculated from all the tweets that we have performed an emotional analysis did not exceed the threshold that we have set earlier (1,08); therefore, the emotion classification of these tweets is positive and content is emotionally safe and by consequence, it is okay to be displayed. We have noticed as well that the joy value in all tweets emotion extraction is significantly higher than all other 4 emotions (sadness, anger, disgust, fear) which are considered to be negative emotions and this proves the positive results that we have obtained, in addition to that if we analyze the sentences individually we can see for sure that their content affirms something good and reflects happy and joyful emotions like for example in the first tweet it reflected the fact that PLA's participation in drills aims to deepen practical and friendly cooperation with the militaries of participating countries, which is good news, in the other hand the second tweet expressed the feeling of the first meeting with his lover, and for third tweet, it was a Congratulations message; as for the last tweet, it was a motivational message and helpful advice to feel better about yourself. To conclude this observation, we saw that if the joy value is significantly higher than the other emotions, the content will be automatically positive and safe according to the Euclidean distance.

5 Conclusion

Artificial intelligence (AI) has gotten a lot of press in recent years. The idea that computers would be able to imitate human cognitive processes to solve complicated computational/analytical issues is enticing, but it might be deceptive. It is easy for someone who is not familiar with AI technologies to think. For example, these systems can mysteriously and almost simply extract valuable insights from data. In this paper, we have presented a solution to create an emotion filter that detects based-text emotions extracted from tweets and monitors them based on their harmfulness. We presented an approach for textual emotion recognition positioned within the context of tweets from API Twitter. Our classification algorithm receives a text tweet as input and classifies it according to the five emotional types. The output

Table 1 Examples of tweets that might have negative impact on users

Tweet content	Emotion extraction	E.D	Classification
Europe facing its worst drought in 500 years - study https://t.co/jY0vqojAYC https://t.co/ZAXZslBekK	Sadness: 0.603305, joy: 0.061472, fear: 0.160444, disgust: 0.040136, anger: 0.05097	1.1226546857542616	Negative
Spectator sues Nick Kyrgios over '700 drinks' accusation during Wimbledon final https://t.co/0xPsmIj9vF	Sadness: 0.523502, joy: 0.060197, fear: 0.034915, disgust: 0.124649, anger: 0.257796	1.1081969079211509	Negative
From the Cold War to the "war on terror," the United States has used Somalia as a battleground for its geopolitical schemes with profoundly destructive consequences for Somalis.	Sadness: 0.516421, joy: 0.018284, fear: 0.327, disgust: 0.013176, anger: 0.044437	1.1501696659515066	Negative
Extreme 'whiplash' between drought and floods makes it harder to recover from climate disasters https://t.co/meX7XGh7YV	Sadness: 0.634944, joy: 0.074562, fear: 0.181029, disgust: 0.006534, anger: 0.055189	1.1315847573085278	Negative
RT @Seeker: Yelp will flag crisis pregnancy centers listings to avoid misleading abortion seekers https://t.co/Gs5ALAM5mc https://t.co/p85d...	Sadness: 0.162586, joy: 0.403488, fear: 0.29928, disgust: 0.097456, anger: 0.067443	1.2540924644562697	Negative

Table 2 Examples tweets that might have positive impact on users

Tweet content	Emotion extraction	E.D	Classification
PLA's participation in the Vostok-2022 drills aims to deepen practical and friendly cooperation with the militaries of participating countries, enhance strategic coordination and strengthen the ability to deal with various security threats: Defense Ministry spokesperson https://t.co/xhbwad1YEs	Sadness: 0.02296, joy: 0.932942, fear: 0.118979, disgust: 0.023001, anger: 0.024796	0.13218711900055975	Positive
Three years ago today @EmmaBostian and I met for the first time, Hoping we get to spend Christmas together this year in Sweden!!	Sadness: 0.2004453, joy: 0.863676, fear: 0.023482, disgust: 0.017436, anger: 0.009059	0.23757168240133333	Positive
Congratulations to my PhD adviser @geoffreyhinton for winning the Royal Medal! Totally unsurprising :)	"Sadness": 0.233232,"joy": 0.676328, "fear": 0.018988,"disgust": 0.019905,"anger": 0.019745	0.4075833218557894	Positive
Have the courage to walk away from the drama and people who create it around you. It's most often about them and not you. Surround yourself with those who support you & believe in you. Focus on the positive. Love those who treat you well, pray for the ones who don't. They need it.	Sadness: 0.134433, joy: 0.625341, fear: 0.081753, disgust: 0.018836, anger: 0.097448	0.41206246484664	Positive

emotional array can be used to determine the emotional type by calculating the Euclidean distance. The majority of recent studies employed supervised and unsupervised classifiers using machine learning. With regards to sentiment and emotion recognition from text, we take for example Chen et al. (2021) in their paper they look into sentence-level categorization, as well as review-level categorization, and their research and analysis make use of the Scikit-learn program which is a Python-based accessible software library which they use unsupervised techniques to measure similarity. On the other hand, IBM API analysis and similarity calculation using the Euclidian distance performed better than other machine learning methods and eventually outperforms the Scikit-learn program. Therefore, for our study, we decided to categorize tweets into their relevant sentiment and emotion classes using the score we got from calculating the Euclidean distance. Also, our contribution is based on the way IBM natural language understanding API libraries are created and how they are used in conjunction with heuristic rules. Our approach is illustrated through a software system, IBM API. Besides the affect recognition engine, IBM's natural language understanding API is, to the extent of our knowledge, the most reliable free open-source textual emotion recognition API published on the web. The evaluation study showed promising results in terms of high classification accuracy, highlighting the importance of text emotion recognition. Future efforts will be centered on refining the algorithm in the IBM natural language understanding API software for better accuracy.

Declarations

References

Abualigah LMQ (2019) Feature selection and enhanced Krill Herd algorithm for text document clustering. Springer, Berlin

Abualigah L, Gandomi AH, Elaziz MA, Hamad HA, Omari M, Alshinwan M, Khasawneh AM (2021) Advances in meta-heuristic optimization algorithms in big data text clustering. Electronics 10:101

Abualigah L, Diabat A, Mirjalili S, Abd Elaziz M, Gandomi AH (2021) The arithmetic optimization algorithm. Comput Methods Appl Mech Eng 376:113609

Abualigah L, Yousri D, Abd Elaziz M, Ewees AA, Al-qaness M, Gandomi AH (2021) Aquila optimizer: a novel meta-heuristic optimization algorithm. Comput Ind Eng 157:107250

Ahmad S, Asghar MZ, Alotaibi FM, Awan I (2019) Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. Hum Centric Comput Inf Sci 9:24

Ahmed CM, Abdelhamid D (2022) Sentiment analysis of Algerian dialect using machine learning and deep learning with Word2vec. Informatica 46:67–78. https://doi.org/10.31449/inf.v46i6.3340

Al Shehhi A, Thomas J, Welsch R, Grey I, Aung Z (2019) Arabia Felix 2.0: a cross-linguistic Twitter analysis of happiness patterns in the United Arab Emirates. J Big Data 6:33

Aloufi S, El Saddik A (2018) Sentiment identification in football-specific tweets. IEEE Access 6:78609–78621

Ansari L, Ji S, Chen Q, Cambria E (2022) Ensemble hybrid learning methods for automated depression detection. IEEE Trans Comput Soc Syst

Arias M, Arratia A, Xuriguera R (2013) Forecasting with Twitter data. ACM Trans Intell Syst Technol 5:8

Boivin M, Hymel S, Bukowski WM (1995) The roles of social withdrawal, peer rejection, and victimization by peers in predicting loneliness and depressed mood in childhood. Dev Psychopathol 7(04):765–785

Budiharto W, Meiliana M (2018) Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis. J Big Data 5:51

Cambria E, Hussain A, Durrani T, Havasi C, Eckl C, Munro J (2010) Sentic computing for patient centered applications. In: IEEE 10th international conference on signal processing proceedings. IEEE, pp 1279–1282

Chen Q, Chaturvedi I, Ji S, Cambria E (2021) Sequential fusion of facial appearance and dynamics for depression recognition. Pattern Recogn Lett 150:115–121

Chou KL, Liang K, Sareen J (2011) The association between social isolation and DSMIV mood, anxiety, and substance use disorders: wave 2 of the national epidemiologic survey on alcohol and related conditions. J Clin Psychiatry 72(11):1468–1476

Cury RM (2019) Oscillation of tweet sentiments in the election of João Doria Jr. for Mayor. J Big Data 6:42

Davila J, Hershenberg R, Feinstein BA, Gorman K, Bhatia V, Starr LR (2012) Frequency and quality of social networking among young adults: associations with depressive symptoms, rumination, and co-rumination. Psychol Pop Media Cult 1(2):72–86

Deters F, Mehl MR (2012) Does posting Facebook status updates increase or decrease loneliness? An online social networking experiment. Soc Psychol Personal Sci 1948550612469233

Dick J (2013) Why do social networks increase stress? The Huffington post 11/09/2013. Retrieved 9 October 2016 from http://www.huffingtonpost.in/entry/social-networks-and-stress_b_3534170

Drouin M, Kaiser DH, Miller DA (2012) Phantom vibrations among undergraduates: prevalence and associated psychological characteristics. Comput Hum Behav 28:1490–1496

Han S, Mao R, Cambria E (2022) Hierarchical attention network for explainable depression detection on Twitter aided by metaphor concept mappings. arXiv preprint arXiv:2209.07494

https://cloud.ibm.com/apidocs/naturallanguage-understanding

https://github.com/benrouba/Emotional-Sentiment-Analysis-of-Social-Media-content-for-Mental-Health-Safety

Jansen BJ, Zhang M, Sobel K, Chowdury A (2009) Twitter power: Tweets as electronic word of mouth. J Am Soc Inf Sci Technol 60:2169–2188

Ji S, Pan S, Li X, Cambria E, Long G, Huang Z (2020) Suicidal ideation detection: a review of machine learning methods and applications. IEEE Trans Comput Soc Syst. https://doi.org/10.1109/TCSS.2020.3021467

Ji S, Li X, Huang Z, Cambria E (2022) Suicidal ideation and mental disorder detection with attentive relation networks. Neural Comput Appl 34(13):10309–10319

Komal M, Joshi MS (2016) Effectiveness of different similarity measures for text classification and clustering. Int J Comput Sci Inf Technol 7:1715–1720

Leung CWK (2009) Sentiment analysis of product reviews. In: Encyclopedia of data warehousing and mining, 2nd edn. IGI Global, pp 1794–1799

Morente-Molinera JA, Kou G, Samuylov K, Urena R, Herrera-Viedma E (2019) Carrying out consensual Group Decision Making processes under social networks using sentiment analysis over comparative expressions. Knowl Based Syst 165:335–345

Mountaga D, Chayma F, Hatem H (2021) Bambara language dataset for sentiment analysis, conference paper ICLR

Pong-inwong C, Songpan W (2018) Sentiment analysis in teaching evaluations using sentiment phrase pattern matching (SPPM) based on association mining. Int J Mach Learn Cybern 10:2177–2186

Prabhu BA, Ashwini B, Khan TA, Das A (2019) Predicting election result with sentimental analysis using Twitter data for candidate selection. In: Innovations in computer science and engineering. Springer, Berlin, pp 49–55

Ringsquandl M, Petkovic D (2013) Analyzing political sentiment on Twitter. In: Proceedings of the 2013 AAAI spring symposium series, Stanford, pp 25–27

Risch J, Krestel R (2018) Aggression identification using deep learning and data augmentation. In: Proceedings of the first workshop on trolling, aggression and cyberbullying (TRAC-2018), Santa Fe, NM, USA, pp 150–158

Shaha A-O, Amal A-R (2022) A review and comparative analysis of sentiment analysis techniques. Informatica 46:33–44. https://doi.org/10.31449/inf.v46i6.3991

Souri A, Hosseinpour S, Rahmani AM (2018) Personality classification based on profiles of social network users and the five-factor model of personality. Hum Centric Comput Inf Sci 8:24

Subramani S, Wang H, Vu HQ, Li G (2018) Domestic violence crisis identification from Facebook posts based on deep learning. IEEE Access 6:54075–54085

Subramani S, Michalska S, Wang H, Du J, Zhang Y, Shakeel H (2019) Deep learning for multi-class identification from domestic violence online posts. IEEE Access 7:46210–46224

Yuan S, Huang H, Wu L (2016) Use of word clustering to improve emotion recognition from short text. J Comput Sci Eng 10(4):103–110

Ji, S., Pan, S., Li, X., Cambria, E., Long, G., & Huang, Z. (2020). Suicidal ideation detection: A review ofmachine learning methods and applications. IEEE Transactions on Computational Social Systems, 8(1), 214-226.

Han, S., Mao, R., & Cambria, E. (2022). Hierarchical attention network for explainable depressiondetection on Twitter aided by metaphor concept mappings. arXiv preprint arXiv:2209.07494.