Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4674 - 4680

Received 25 April 2021; Accepted 08 May 2021.

Emotion Analysis of Social Platform, Blogs and Audio Speeches of People

(1)Ankur Chakraborty, (2) Kishan Kumar, (3) P .Visalakshi

SRM INSTITUTE OF SCIENCE & TECHNOLOGY

(1)ap6800@srmist.edu.in , (2)kr3797@srmist.edu.in, (3) visalakp@srmist.edu.in

ABSTRACT

In this social world where whole world is just combined under the social platform, every people have his opinion on every topic whether it is politics, education, or even economics. Sentiment analysis is the term which will play a crucial role to compete in this world because everybody wants something new every time no one likes static you have to be on your knees every time and should analysis every time what people want from you .Like if you have electronics products market you have to keep keen eye on your prices of products, processing time ,camera mega pixel. If you are on politics you have to keep keen eye on growth of economy, removal of article 370, india - china war opinions of people are important to analysis the sentiment and change our plans according to it as we know once a poet said "if you change nothing nothing will change". In this paper we will predict the mood of people using machine learning approaches by extracting the data from social platforms.

Keywords: Machine learning, Natural language processing, Python, Mood prediction

INTRODUCTION

Social platform likes twitter, Instagram, facebook have a vast reach among people and provide variety of opinions what our paper does is to analyze to sentiment of people from their tweets, post gather the information and provide the result in a simple and structured manner about what people think about their product. Sentiment analyze is the way to form a structured data from unstructured data like tweets because it is very difficult for companies to read each and every tweet manually and predict mood of people. So by using machine learning approaches, and by natural language processing and computer linguistics methods we can differentiate data based on positive and negative words, identify stop words, comparison with a specific keyword for a product and using approaches like naïve bayes and neural networks.

Review is important part of sentiment analysis each and every person's view is important if you want to grow because each and every person is different in his way and will look the things differently in his own way. Opinions are of two types one is direct opinion another is comparison opinion.

For example : The processing of xiaomi is good.(DIRECT OR OBJECTIVE OPINION)

The processing of xiaomi is better than Lenovo and Micromax.(COMPARISON OPINION)

Like this opinion analysis is further classified as objective and subjective sentence.

Objectify: I bought a car few days ago.

Subject: The car is so much efficient.

For subjective sentence we can classify sentence as positive and negative .

Positive: The new Tata motors car is awesome.

Negative: New tata motors cars have a poor mileage.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4674 - 4680

Received 25 April 2021; Accepted 08 May 2021.

Each feature of product is classified on different aspects and based on various approaches mood or sentiment of people can be predict. In this paper we will extract the data from twitter so to remove unspecific data which are not related to issue as in twitter people tweets in a more genuine manner as respect to instagram or facebook.

In this paper we first extract data from twitter called a (Pre- processing) stage then we move to feature extraction using supervised and unsupervised machine learning approaches and the fourth section tells us about comparison between existing system and the new proposed

system and at the end we will deal with conclusion and future scope.

BACKGROUND RESEARCH WORK:

As we have stated above sentiment analysis focus to determine the sentiment and mood of people .In this paper we try to compact and make our model in such way that it can combine views of large amount of people that's why we are using social platform like twitter ,irrelevant text are less in twitter as compare to social platforms like Instagram and facebook so genuine opinion can be formed easily if we gathered data from twitter.

When we are studying about research papers we found that there are many approaches for sentiment classification like supervised and unsupervised algorithms ,support vector machine ,maximum entropy , natural language processing from word net dictionary we try to read and analyze different aspects of approaches.

The first paper we have read is done by Lina zoh she used to estimate review of films using expert system and semantics of sentence. Using nlp techniques and text classification is considered to classify the review. For semantic orientation unsupervised method is used because unsupervised methods require less methods to define explicitly . When she compares both approaches supervised learning approach 85% while unsupervised learning has 75% accuracy of movie reviews. We can conclude that supervised has more accuracy but it requires more time to train model while if we look unsupervised learning approach it is easy to use in real time application.

The next paper we read is written by bang he used machine learning techniques are more effective than human produced sentiment data. In this he uses naïve bayes, support vector machine also neural networks to see how they can improve accuracy of estimated data. He differentiated data based on positive and negative words characteristics based on unigram and bigram are used for classification.

After that we read our 3rd paper which was written by Zhu which produced data based on opinion polling free from textual reviews , the aspect related terms was learnt using bootstrapping method. It converts the multi aspect into a single aspect which was used for opinion polling .They held a survey on restaurant which have 76% accuracy. This method is easy to implement and can be applied to other domains like politics, movie reviews etc. .

In our 4th paper we found a paper written by jeonghee he used a mood predict analyzer which extract data from online documents , he used natural processing techniques. He find out all the aspect identification words then he estimate sentiment polarity of each identification word. To estimate and predict the sentiment it uses sentiment lexicon and extract pattern database also .This approach is also used in online product reviews for digital reviews.

In our next paper we read about mood prediction using linguistic knowledge which is a machine learning method gathered through synonymy graphs for effective classification .This approach shows the degree of relationship between documents have on their mood prediction. This is done by graph cut technique and also uses opinion words which were found through synonym graphs of word net. The proposed system also improves the prediction in classification task. It also improves accuracy up to 90 % accuracy with an added advantage of reduction in processing time , with minor difference in financial accuracy.

In our recent last paper which we read does sentiment analysis methods to classify web forum opinions in multiple languages they used stylistic and syntactic features to evaluate text and predict the sentiment in English and Arabic content. The entropy weighted generic algorithm used to improve classifier and achieve the true opinions of key features.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4674 - 4680

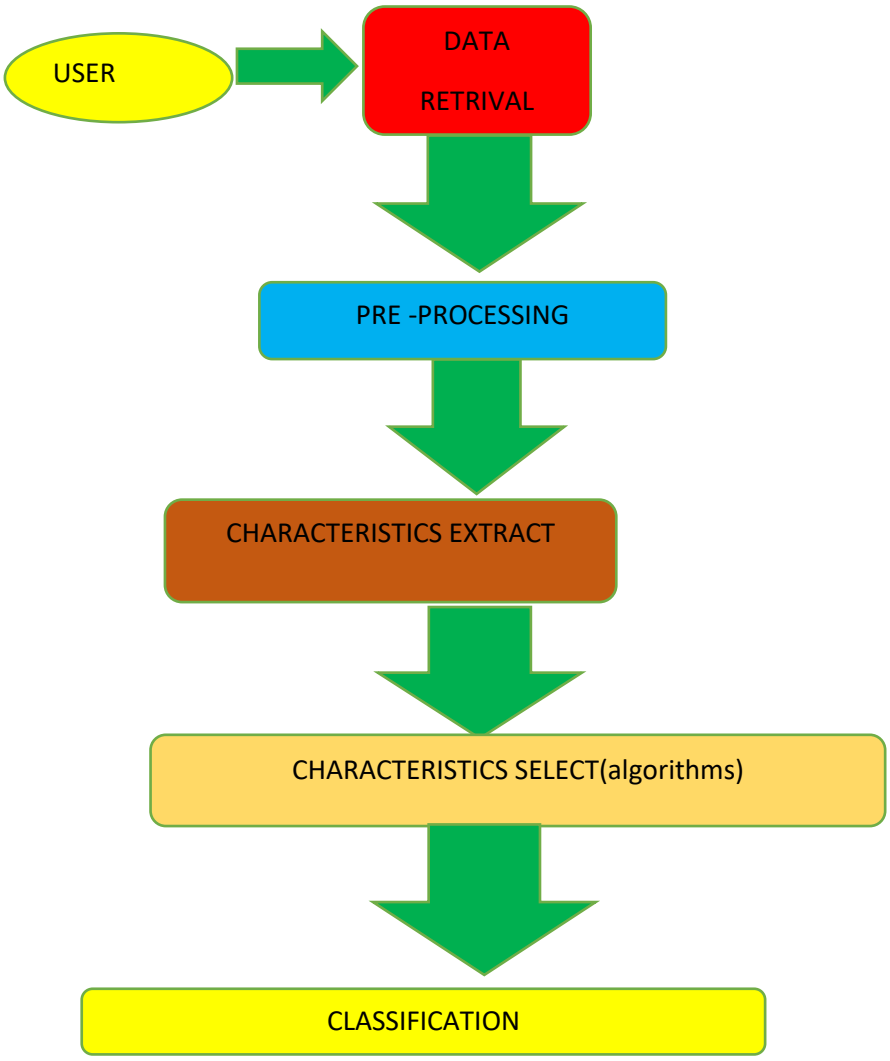
Received 25 April 2021; Accepted 08 May 2021.

EXISTING SYSTEM:

In the existing system most of approaches are domain specific . The existing system also doesn't take accountability of future impacts of results if you apply the output of sentiment analysis in your system or method . It also doesn't allow you to download data from application. If we say it in a simpler manner it works more like a static manner rather than dynamic data . In these approaches you can't use unsupervised algorithm like vector categorization , natural language processing methods , facial and speech recognition methods and therefore can't produce accurate sentiment data because they can't use large data set methods.

Proposed methodology: Our proposed methodology is to try and compact large data set to our system which we are trying to implement using these 5 steps: (1) Data retrieval (2) Pre-processing (3) Characteristics extraction (4) Characteristics selection (5)Classification

SYSTEM ARCHITECTURE DIAGRAM



DATA RETRIVAL: we will build a software which will be built using python because in python it is easy to code with respect to others the software we will take is tweepy , text blob , nlk etc.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4674 - 4680

Received 25 April 2021; Accepted 08 May 2021.

PRE-PROCESSING: in this stage we will filter data retrieval extract using software removal of non-essential things . There are many stages in pre-processing

(1) Conversion from uppercase to lower case

(2) using parts of speech method which we study in natural language processing we will categorize the text words once this stage is we will move to next step of pre-processing.

(3) Emojis categorization is another important thing in mood prediction certain words and symbols represent positive and negative words .

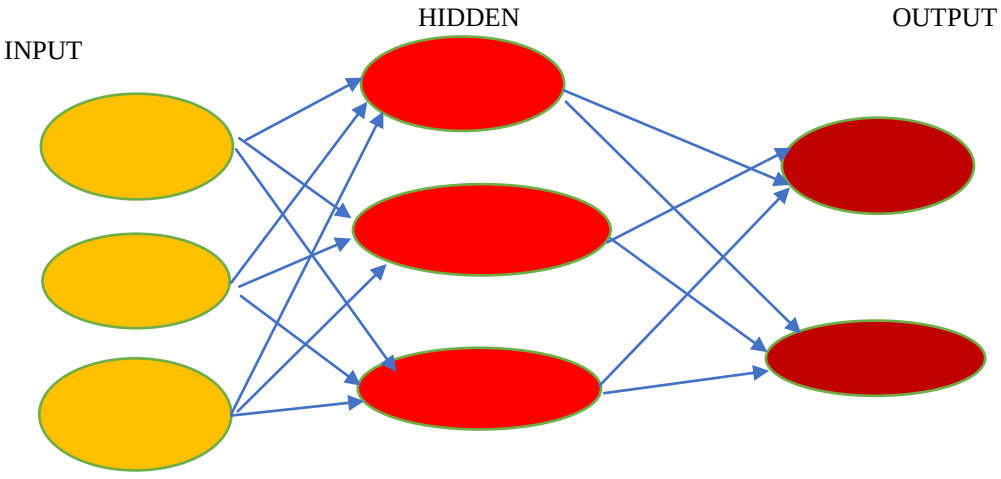
(4)Removal of stop words which are so common in text and it improves the efficiency of application stop words are “he, she, the, a” etc. the selection of words to be removed is done using listed words in nltk form.

Characteristics extraction: in this stage we try to focus to extract the text which will define uniqueness of an object .

(1) The first method is using parts of speech of tag to categorize the word based on adjective adverbs and nouns .
(2) Based on n-gram model there are 3 features like unigram , bigram , trigram model.

Possible algorithms : In this stage we will consider all possible algorithms that can be applied on our application .
(1)Naïve bayes: this classifier based on bayes probability theorem in this each feature is considered to be independent of another feature and calculate the posterior probability of each feature.
P(k/y) = (p(y/k)*p(k))/p(y)
P(k/y) is the posterior probability.
P(y/k) is the likelihood probability.

(3) Neural networks: this algorithm is one of most important way for mood prediction . it is also used as a alternative because of accuracy many are keen to use this method . in this we will train our system which will handle the correlation between the input variables. By activating the nodes in hidden layer it will perform back propogation.



There are 2 phases in the neural networks (1) the first one is training and the (2) one is testing . In the training phase they categorize the positive and negative words and assigned them weights. This phase focus on creating a bank of collection of words. Our system is trained with a labelled data to produce a meaningful doutput. The process of learning from a labelled data is called back propagation. Each node in this layer points to every node in the next layer. We fed the data in the input layer and from the edges it moves to the next layer until it reaches the output. Each edge has a weight and when input passes through edge it is multiplied with the weight associated with it. After that we evaluate the system performance then calculate whether the text is positive or negative.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4674 - 4680

Received 25 April 2021; Accepted 08 May 2021.

Support vector machine: another popular method which is considered in mood prediction widely is svm(support vector machine) .It is a supervised machine learning algorithm which can be used both to classify and estimate changes. In this approach each feature is pointed as a n point in n dimensional space value of each feature is pointed as a unique coordinate. It tries to find a hyper plane which separates the feature in two classes.

Problems faced :The main challenges faced are:

Detection of irrelevant content: Since we told in this modern world people are more expressive but at the same time they are less responsible than before anybody can tweet or post anything many of times they are irrelevant , so it is important to remove irrelevant post so to improve accuracy .

Domain-free : During the development of application domain specific limitation is one of the major problems in most of applications so it is important to remove this problem so that we have a wide reach among the users .

Remove limitation of analysis data : When we are looking on different mood prediction analysis application we see most of applications on the amount of data that can be analyzed or filter so we need to focus on increasing the amount of data that can be filter .

Removal of gramitical and ambiguity : Since in this mood prediction application we use many features of natural language processing techniques during classification and categorization of text so there are chances of ambiguity and removing implicitness and inference .

Applications of mood predictions:

(1)Buying of products : During the buying of products mood prediction tool is very useful so that you can analyze the things easily it gives you more than 100 of opinions in a single place you don't have to pay any third party or so called expert counsellor who doesn't give you 100% assurance and there are great chances of you can loose your money on them. In this approach you can take the responsibility on yourself and can easily things and take decisions.

(2) Communication with clients: In this application we are focusing on also to communicate with clients to be well informed about what information they want and also some expert views which users can want to know and also tell them about future impacts of these decisions.

(3)Review collection: In this application we an easily get reviews of 1000's of customer from the application which will help both customer and manufacture because they will also get information on what users really want and where there product lack and they will get information on improvement in product or services so that users can trust on their application and make their reach vast.

(4)Easy research : More than buying or selling of products mood prediction are also used in stock prediction which stock can be in profit in the upcoming days also they are used by political parties to predict what people think about their policies and what changes should they made in their policies so that they can increase their trust in public .

(5)Smart system: In this we will also try to implement smart recommendation system based on your previous search it will tell you web pages from where you will get information which can be useful for client so recommendation also plays a important role many times.

RESULTS AND OUTPUTS:

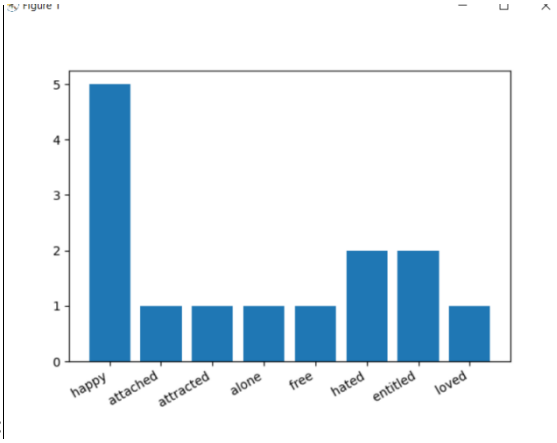
In this project we have given a new way of analyzing things from analyzing tweets , speeches , to voice .In this project we have successfully implemented all these three things by using natural language techniques to analyze sentiment of people.

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4674 - 4680

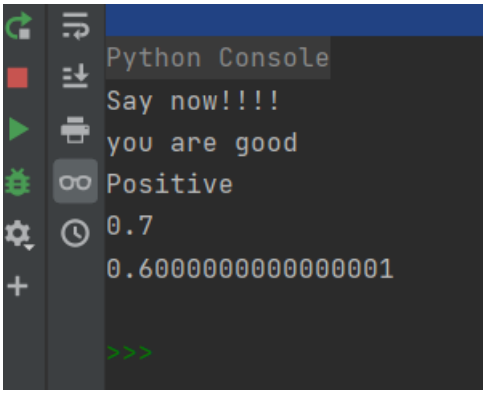
Received 25 April 2021; Accepted 08 May 2021.

OUTPUTS:

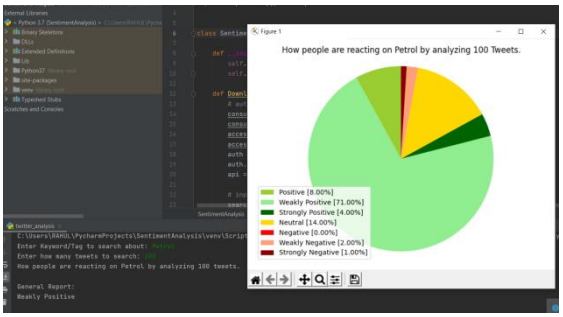
BLOG MOOD PREDICTION:



VOICE ANALYSIS:



TWEETS ANALYSIS:



CONCLUSION AND FUTURE SCOPE :

Since we have seen that mood prediction is the new term which is growing and will have a great impact on future upcoming so called “social world” where everyone is on social platform people are more expressive they are free to present their views and also to know what people think about them . It can play a crucial role in election prediction result sense the sentiment of public or in share or stock market where you analyze the ground sentiment and also got to know opinion of experts what they have views on trade also in this pandemic of covid world moves much towards the online mode so it is important that these sentiment or mood prediction

Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 4674 - 4680

Received 25 April 2021; Accepted 08 May 2021.

application will guide and train the new comers to sustain in this new era or competitive world which will also help in growing of economy of country which will focus on this and use it in a efficient manner .Lot of research is going on and improvements are made in short a lot has been done and a lot has to be done .There is a vast scope of improvement since this is such a broad topic where there is no limit of change a lot of things can be change.

REFERENCES :

(1)S.SIDDHARTH , (2)R.DARSINI (3) Dr. M. SUJITHRA "Sentiment Analysis on Twitter Data Using Machine Learning Algorithms in Python" Department of Computing Coimbatore Institute of Technology.

(2) Ali Hasan , Sana Moin , Ahmad Karim and Shahaboddin Shamshirband "Machine Learning-Based Sentiment Analysis for Twitter Accounts" Department of Computer Science, Air University, Multan Campus

(3) Torben Antrettera , Ivo Blohmb , Dietmar Grichnika , Joakim Wincent " Predicting new venture survival: A Twitter-based machine learning approach to measuring online legitimacy"

(4) Brinda Hegde, Nagashree H , Madhura Prakash "Sentiment analysis of Twitter data: A machine learning approach to analyse demonetization tweets" VTU & Information Science and Engineering BNMIT.

(5) Ankit Pradeep Patel, Ankit Vithalbhai Patel "Survey on Sentiment Analysis of Twitter Data using Machine Learning Approaches"

(6)" Sentiment Analysis A Literature Survey" Subhabrata Mukherjee Indian Institute of Technology, Bombay Department of Computer Science and Engineering.

(7) Saumya Chaturvedi AKTU, Lucknow, INDIA ,Vimal Mishra IERT, Allahabad, INDIA "Sentiment analysis using machine learning for business intelligence IEEE".