TEST Engineering & Management

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*

# Predicting User Behavior With Tweets Posted In Twitter

**RK. Ahmadh Rifai Kariapper**

South Eastern University of Sri Lanka. rk@seu.ac.lk

***Article Info***
***Volume 82***
***Page Number: 17702 - 17712***
***Publication Issue:***
***January-February 2020***

***Article History***
***Article Received:*** *18 October 2019*
***Revised:*** *14 November 2019*
***Accepted:*** *12 December 2019*
***Publication:*** *29 February 2020*

***Abstract***

Twitter is the most well-known microblogging platform around the globe. It has been one of the fastest-growing social network platforms and has a good position in the area of microblogging. More than 500 million registered users post 340 million tweets every day, sharing their thoughts and daily activities. Compared with typical microblogging platforms. Meanwhile, various researches have shown that people tend to express their emotions on Twitter. These tweets usually clearly express user behavior. The most popular emotional analysis task in Twitter is the automatic classification of tweets into sentiment categories such as positive, negative, and neutral. It does not give an idea about user emotion. (e.g., sad, surprise, love). Therefore, our goal is to design an emotion-based algorithm to analyze emotion. To develop an algorithm, we used a dictionary of words for each emotion (lexicon approach). However, we used the support vector machine and naïve Bayes classifiers as machine learning. We have shown that combining these methods provides 82%, 71% accuracy for support vector machine, and naive Bayesian classifiers in the emotional analysis.

***Keywords:*** *Bag of Words model, lexicon approach, machine learning, Twitter;*

## INTRODUCTION

The development of social network platforms [1]–[3] has given people a new way to produce and consume a great deal of information on the web [4], [5]. People used to acquire information from portal websites [6]–[9]. Many websites provide a long list of topics varying from politics to entertainment [10], [11]. These conventional online data sources are helpful yet less useful because they regularly contain redundant data. However, since the advent of online social network platforms, people like to obtain information from these platforms because of their quick and efficient features.

These platforms are available for people to select the source of information they are interested in. Besides, a large number of social networking sites, such as Twitter [12]–[14], Google+ [15], and Facebook [16], provide their users with data. Twitter is the world's most popular microblogging site [17]–[19]. It is also the fastest rising social network site and has a healthy place in microblogging. More than 500 million authorized users post 340 million tweets every day, share their opinions and everyday activities. Compared to the usual microblogging sites, Twitter messages are much shorter. You are only required to post 280 characters or less in one Twitter message. This makes it easier for people to get Twitter to a critical point because of the large amount of information available online. Depending on the needs of users, Twitter users will follow whatever people and sources of information they want. With all the advantages listed above, Twitter has become a powerful platform [20] with many details, ranging from breaking news worldwide to buying goods at home.



*Published by: The Mattingley Publishing Co., Inc.*

TEST Engineering & Management

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*

The information streams on Twitter have witnessed an unprecedented rise in the popularity of this social network in the last few years. Users have a vast amount of knowledge on various aspects. Not all the details, however, is useful to users, and each user has his or her interests and preferences. For consumers, there is an urgency to provide customized services. Nowadays, to support consumers, more and more personalized services are offered. To make their fast-paced lives more productive, people need this customized service. A vast amount of information is published each day on the Twitter website by users. In the last few years, several features and techniques have been researched with varying results for sentiment classifiers on the Twitter platform [21], [22], [31], [32], [23]–[30]. On Twitter, there are also several other research studies relating to emotion analysis.

**Research problem**

The message-level polarity classification is referred to as the task of automatically classifying tweets into sentiment categories. This problem has been successfully solved by representing tweets from a corpus of hand-annotated examples using feature vectors and training classification algorithms on them[33]. Not as much study has been carried on this identify the polarity of the text in terms of emotion. Classifying tweets in various forms of emotions is one step ahead of classifying text based on sentiment as emotions can be categorized as happy, sad, caring, angry, surprised, etc.

In particular, this technology can calculate the public mood of people in a society that can allow social scientists to understand the quality of life of the population. Measuring and controlling the living standards and quality of life of a community is essential to public policymaking. Quality of life can be calculated based on various facets of life, including physical, emotional, psychological, life satisfaction, and employment. Methods that measure living standards, however, fail to measure what people think and feel about their lives [34].

**Research Objective**

How to detect emotion from Twitter data?

Emotion detection in tweets is often a key topic in Twitter studies. In previous works, various approaches have been suggested. Much of this emotion study is linked to the preparation of the classifier for the classification of emotions. Instead of seeking a new approach to detecting emotions on Twitter, the study of emotions is based. The features of Twitter are short and casual. To see users' emotive responses from Twitter messages for our approaches of emotion detection should be carried out with the aim of high accuracy to identify the emotional states.

## LITERATURE REVIEW

Prabu Palanisamy et al. [35] have experimented with a lexicon-based approach to the study of sentiments on Twitter. It uses a lexicon that consists of words with the respective sentiment scores of each moment. The term is related to a single noun, phrase, or language. The lexicon is based on Serendio taxonomy and contains positive, negative, negation, stop words, and representations where sentiment is defined based on the presence or absence of words in the lexicon. This strategy would separate tweets as positive and negative. As a result of this analysis, it could be found that this approach established a 0.8004 F-score of the test data set used for this study.

Sushree Das et al. [36] completed a study to identify stork marketing forecasting using real-time sentimental data from Twitter. The data was taken from Twitter based on a specific organization. In order to do this analysis, this study focused on the classification model. From the results of this analysis, we could identify that model predicts high precision stock marketing with sentimental data.

PING-FENG PAI and CHIA-HSIN LIU [37] accomplished a study to predict vehicle sales with a Sentimental Analysis of Twitter data stock market values. This study contains a multivariate regression

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*

model with social media data, stock market value, and time-series data to calculate monthly vehicle sales. Also, they have employed the Least squares support vector regression models(LSSVR ) to assist with the multivariate model. To forecast monthly sales, they used hybrid data, sentimental tweets, and stock market values. They used the naïve base model, the exponential smoothing model, Seasonal autoregressive integrated moving average model, Autoregressive integrated moving average model, and backpropagation neural networks for time-series data. The result of this study provides more accurate vehicle sales predictions than any other system.

Bing Liu [34] demonstrates that the corpus-based approach can be made in two ways. The first method is recognizing opinion words and their polarities in the subject area corpus utilizing a given collection of opinion words. Another technique is building a new dictionary inside the specific subject area from another dictionary using a domain corpus. The discoveries suggest that regardless of whether opinion words are field dependent, it can happen that a similar word will have inverse orientation relying upon the context. The research done by Hazivassiloglou and McKeown [38] is famous in the literature about the corpus-based technique.

An unsupervised method for sentiment classification of movie reviews was applied by Rothels & Tibshirani [39]. Zagibalov & Carroll [40] proposed a way to classify the Chinese text as the fundamental idea of using positive words retrieved from the document. Zagibalov & Carroll enriched the list of positive seed words through iterative classification. Through the inspiration of their process, Rothels & Tibshirani also composed an initial seed set. The text of the document is classified into zones by separation of a piece of text by punctuation characters.

“An easy way of measuring the difference between positive and negative terms and marking it according to the sign of the difference.”Taboada and Thelwall have suggested more complex aggregation functions, including negation rules and opinion shifters[41].

A typical way is to construct the problem as a managed learning problem. The fundamental idea behind this approach is to train a function that can evaluate the feeling of an intangible text by means of a corpus of sentimentally marked documents. Pang, Lee and Vaithyanathan [42] were trained by film reviews of the Internet Movie Database (IMDb) in a bilateral (positive/negative) classification. They used features, including unigrams, bigrams and part of speech tags and algorithms for learning: including vector support machines (SVM), maximum entropy classification, and naive bayes.

## METHODOLOGY

This study is divided into three parts: collecting tweets, Data analysis, and Emotional analysis.

### 1. Collecting Tweet

Users need an app to communicate with the Twitter API to connect Twitter data on a programmatic ground. The Twitter API consists of two main divisions in general context, namely REST and streaming APIs, respectively. The REST APIs have programmatic access to Twitter-read and write data, such as the latest Post, the profile of the author and the followers, and more. The REST API recognizes Twitter applications, and OAuth users and answers are available in JSON format.

The streaming APIs constantly provide new responses in long-term HTTP connection to REST API requests. On the other hand, in order to get updates to the recent Tweets which match your search query, stay tuned for user profile updates, and etc. Streaming APIs provide low latency access for developers to Twitter’s global Tweet info stream.

### Authentication



*Published by: The Mattingley Publishing Co., Inc.*

January-February 2020
ISSN: 0193-4120 Page No. 17702 - 17712

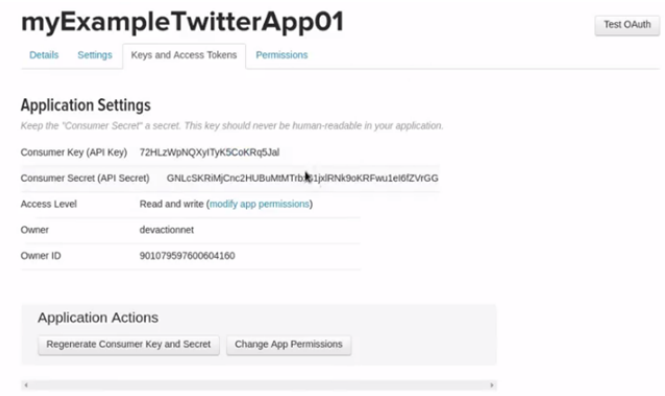


**Figure 1: Twitter App**

Twitter uses OAuth to provide authenticated access to the API. OAuth can be used for two main reasons: safe and standardized, respectively. Being secure means that accounts are secure because we do not need to share accounts with third-party applications. On the other hand, since it follows the standard, there are many examples and libraries to understand this topic. In general, there are two forms of authentication. The most common form of resource authentication in a Twitter OAuth 1.0a application is user authentication and application-only authentication. Signed requests identify the identity of an application in addition to the ID associated with the granted permission of the end-user whose application makes the API call represented by the user's access token. An application-only authentication is a form of authentication that an application performs its API request without a user context. Although API calls still have rate limits for each API method, the pool that each method retrieves belongs to the entire application, not per user limit. For API methods supporting this form of authentication, the document contains two rate limits per user (for application user authentication) and application by application (for application-specific authentication). Because some methods require a user context (for example, only a logged-in user can create a Tweet, a user context is required for that operation), and all API methods support application-specific authentication, There is none.

It is necessary to prepare a Consumer key and Consumer secret (API secret) in order to obtain a response from Twitter. With the creation of an app in the development area of Twitter, which is able to react with the consumer key and the consumer secret.

### Tweepy API

```
import tweepy
from tweepy import OAuthHandler

consumer_key = 'YOUR-CONSUMER-KEY'
consumer_secret = 'YOUR-CONSUMER-SECRET'
access_token = 'YOUR-ACCESS-TOKEN'
access_secret = 'YOUR-ACCESS-SECRET'

auth = OAuthHandler(consumer_key, consumer_secret)
auth.set_access_token(access_token, access_secret)

api = tweepy.API(auth)
```

**Figure 1: Tweepy API**

Tweepy makes it easier to use the Twitter streaming API by handling authentication, connection, creating and destroying the session, reading incoming messages, and partially routing messages.

## 2. Data Analysis

When tweet data has been obtained from Twitter, it must be prepared for review by deleting unnecessary pieces of data such as images, icons, digits, and links into that data.


**Figure 1: Twitte Data with Noise**

So, we had to clean it. In general, the steps required for text preprocessing depend on the requirements or applications of interest. Here, tweets are preprocessed to analyze emotions. Unlike other text documents, Twitter is a domain for flexibly expressing messages and comments freely. For Twitter status updates and tweets, several attributes are identified. The



Published by: The Mattingley Publishing Co., Inc.

TEST Engineering & Management

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*

maximum length of a Twitter message is 280 characters, including user comments, hashtags, URLs, etc. The frequency of elongated words, slang, acronyms, and emoticons is much higher than any other domain.

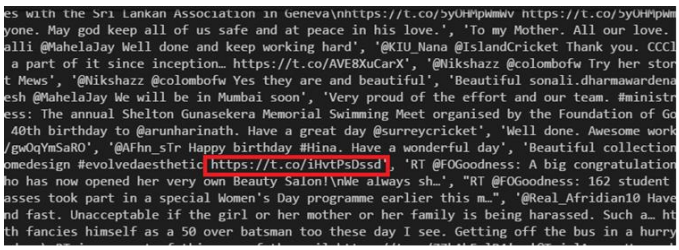


Figure 1: Remove Url

### Removing Links/URL

The tweet includes the direction to the URL. Embedded URLs are usually used to provide a source of detailed descriptions of the content that is listed in the text. However, in our approach, to analyze the feelings of the tweets, the description of the web content is not taken. Accordingly, such web direction is removed from the tweets by identifying such patterns.

### Removing emojis

Figure 1: Emoticons

We need to remove emojis and punctuations because they were not needed in our analysis. We removed them because when they were being extracted, they appeared in square boxes instead of proper emoticons.

### Converting all Tweets into Lower case

This allows for further transformation and classification without worrying about data inconsistencies. This function solved the case sensitivity problem by converting all alphabets with lowercase letters and making all data consistent.

### Split by Whitespace and Remove Punctuation

In general, text data is just a set of characters at an initial stage. Every process of text analysis requires words that are available in the dataset. This process is straightforward, as the text is already saved in a machine-readable format.

### Filter Stopped words

| Stopwords | | |
|---|---|---|
| a | it | these |
| about | its | they |
| again | itself | this |
| all | just | those |
| almost | kg | through |
| also | km | thus |
| although | made | to |
| always | mainly | upon |
| among | make | use |
| an | may | used |
| and | mg | using |
| another | might | various |
| any | ml | very |
| are | mm | was |
| as | most | we |
| at | mostly | were |

Figure 1: stopwords

Stop words are words that are generally considered useless. Since most search engines ignore these words, it is common to significantly increase the size of the index without improving accuracy or recall. The NLTK comes with a stop word corpus that contains a list of 128 English stop words.

### Convert words into the base form



*Published by: The Mattingley Publishing Co., Inc.*

TEST Engineering & Management

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*


**Figure 1: Cleaned Data**

There are various changes to the text operation. We need to deal with different forms of the same word and let the computer understand that these different words have the same basic form. For example, the word singing is displayed in many forms, such as songs, singers, songs, and singers. These are the same things. Humans can quickly identify these primary forms and derive contexts.

Extracting these primary forms is very useful. It can extract useful statistics to analyze the input text. Obstruction is a way to achieve this goal. The purpose of stemming is to reduce different forms of words into common necessary forms. This is a heuristic process that cuts off the end of a word and extracts the basic shape.

## 3. Emotional analysis

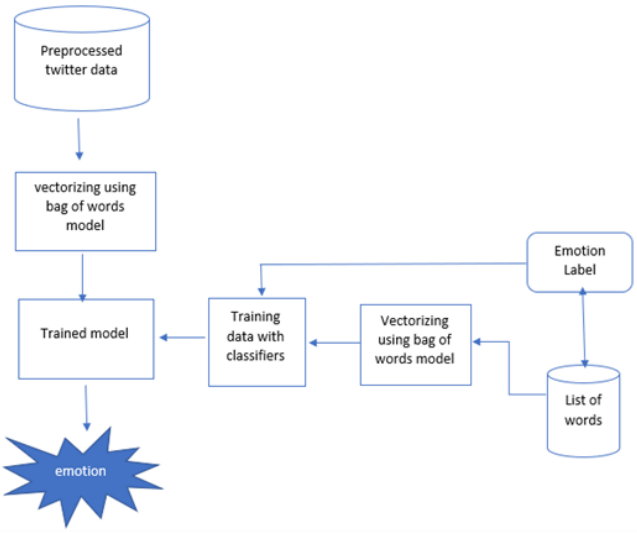


**Figure 1: Proposed Research Model**

The above diagram in Figure 8 shows a clear idea of the process we used to do the emotion analysis. Combination of the lexicon approach, machine learning, and a bag of words model. All the steps we follow for emotion analysis in the above diagram described below.

### Training data

six categories of emotion are used to analyses twitter (anger, fear, joy, love, sadness, surprise)as; the first step. Words are listed for each category. The below diagram is displayed that.

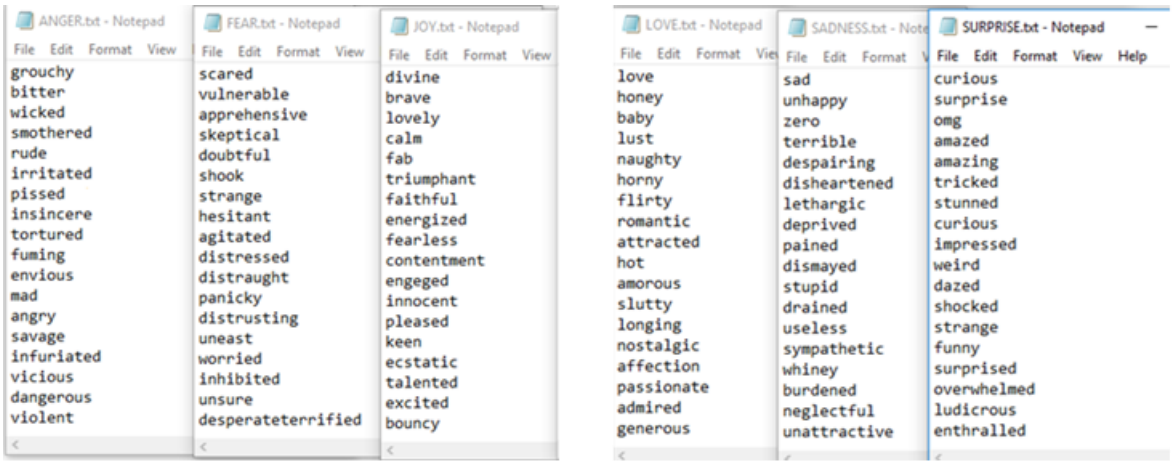


**Figure 1: Lexicon 1 & Lexicon 2**

Machine learning algorithms cannot work with the raw text directly. The text must be converted into



*Published by: The Mattingley Publishing Co., Inc.*

January-February 2020
ISSN: 0193-4120 Page No. 17702 - 17712

numbers. Specifically, vectors of numbers. Then we use the bag of words model to vectorizing the training data. Python code that displayed below.

```
count_vect = CountVectorizer()
train_vectors =count_vect.fit_transform(train_data)
```

**Figure 1: Python Code**

## Bag-of-Words Model approach

{'alleviation': 10, 'amusement': 13, 'animation': 15, 'bliss': 32, 'charm': 36, 'charming': 37, 'cheer': 38, 'cheering': 39, 'comfort': 41, 'delectation': 52, 'delight': 54, 'diversion': 69, 'ecstasy': 77, 'elation': 78, 'exultation': 82, 'exulting': 83, 'felicity': 88, 'festivity': 89, 'frolic': 96, 'fruition': 97, 'gaiety': 98, 'gem': 99, 'gladness': 102, 'glee': 103, 'good': 107, 'gratification': 109, 'hilarity': 117, 'humor': 122, 'indulgence': 129, 'jewel': 133, 'jubilance': 135, 'liveliness': 140, 'luxury': 150, 'merriment': 154, 'mirth': 156, 'pride': 186, 'proude': 189, 'prize': 187, 'rapture': 192, 'ravishment': 194, 'refreshment': 195, 'regalement': 196, 'rejoicing': 197, 'revelry': 201, 'satisfaction': 203, 'solace': 216, 'sport': 220,}

**Figure 1: Words to vector representation**

The bag-of-words model is one of the most straightforward language models used in NLP. It makes a unigram model of the text by keeping track of the number of occurrences of each word. This can be used for Text Classifiers. In this bag-of-words model, we only take individual words into account and give each word a specific subjectivity score. The bag-of-words model is a way of representing text data when modelling text with machine learning algorithms. The bag-of-words model is simple to understand and implement and has seen great success in problems such as language modelling and document classification.

```
[[0 0 0 ..., 1 0 0]
 [0 1 0 ..., 0 0 0]
 [0 0 0 ..., 0 0 0]
 [0 0 1 ..., 0 0 0]
 [1 0 0 ..., 0 1 0]
 [0 0 0 ..., 0 0 1]]
```

**Figure 1:binary representation of training data**

## Classification Algorithms

Grouping instances of datasets has long been the focus of machine learning. This problem is considered a classification problem if the category aiming to be instantiated known in advance. At the heart of the classification problem are tested datasets, training data sets, and classifiers. The training data set consists of instances where class information is available beforehand. This information can not be used with the reverse test data set. The classifier is an algorithm that models a training set to infer the class label of the instance in the test set. In the modelling process, the classifier makes use of the observable features of the instance.

There are several kinds of classification algorithms. Some differences come from feature space modelling in the reasoning process. Some classifiers use linear prediction functions, but other classifiers use nonlinear prediction functions. There are binary and multi-class classifiers. The binary classification task contains only two classes. Some of the multi-class classifiers are derived from binary classifiers, but some are inherently multi-class. Some classifiers generate trust values with possible class labels, but not others. In this paper, several kinds of classification algorithms are used. The Bayesian classifier is selected because it gives the probability that the instance is in the class. The SVM is chosen to utilize both linear and nonlinear models. As nearly a multi-class classifier, k-NN using instance similarity is chosen.

## Result and discussion

We used 100 Twitter profiles to verify the algorithm for this analysis. These are the few Twitter profiles we used for analysis, and the result got from using SVM and Naïve Bayes classifiers

**Table 1.** Result of Twitter profiles

| | SVM classifier | Naïve Bayes classifier | Emotion check By hand |
|---|---|---|---|



Published by: The Mattingley Publishing Co., Inc.

January-February 2020
ISSN: 0193-4120 Page No. 17702 - 17712

| | | | |
|---|---|---|---|
| Kumara Sangakkara | love | love | love |
| Bill gate | surprise | anger | surprise |
| Donald j trump | sadness | anger | sadness |
| Barack Obama | love | love | love |
| Britney spears | love | love | love |
| Emma Watson | love | love | sadness |
| Justin Bieber | love | joy | joy |
| Katy Perry | love | love | love |
| Lady Gaga | love | love | love |
| Pitbul | love | love | love |
| Rihanna | love | love | love |

| | | Predicted class | |
|---|---|---|---|
| Actual Class | | Class = Yes | Class = No |
| | Class = Yes | True Positive | False Negative |
| | Class = No | False Positive | True Negative |

**Figure 1: Evaluation Model**

There are some dissimilarities between actual and predicted. Accuracy is measured from the result we got. Performance is calculated as the following equations.

Accuracy = (TP+TN) /(TP+TN +FP+FN)

Precision = TP /(TP + FP)

Recall = TP /(TP + FN)

In which TP, FN, FP, and TN refer respectively to the number of true positive, the number of false negative, the number of false positive, the number of false negative.

**Table 2.** Classifier-based result

| | **SVM classifier** | **Naive Bayes classifier** |
|---|---|---|
| accuracy | 0.818181818182 | 0.727272727273 |
| precision | 0.767676767677 | 0.647727272727 |
| recall | 0.818181818182 | 0.727272727273 |

Classifying tweets into separate emotional classes is a challenging task. As mentioned, many tweets present more than one emotion. We strongly believe that this would allow us to have a more accurate description of the emotions in the tweet and solves the main issue that we encountered in this work. To accomplish the task, we used the bag-of-words model. It creates a vocabulary of all the unique words occurring in all the tweets. It helps to identify overall emotion in the Twitter profile. However, we train a classifier from a list of words in a particular domain area (sad, surprise and etc.). Ghiassi et al.[26] showed how a hybrid classifier might take advantage of multiple sentiment analysis approaches and how they perform in a Twitter dataset. This method can be used to classify any category you prefer. Improvement may have needed in this study for not adding emoticons to analyze the emotions. Sometimes Twitter users express their feelings through emoticons. It would be corrected in the future. In this research, we used a limited number of words list of each domain area. By increasing the word list, we can achieve higher accuracy. This research showed the performances of SVM and naïve Bayes classifiers. According to that, SVM has greater accuracy. Other classifiers did not test in this study. In the future, it must be tested to identify the best.

## CONCLUSION

Twitter is the most well-known microblogging platform in the world. It is also the fastest growing social network platform and has a good position in the area of microblogging. More than 500 million registered users post 340 million tweets every day, sharing their thoughts and daily activities. Besides, many studies show that people tend to express their emotions on Twitter.

The final results of the research were fairly positive. we have done the classify the six basic emotions for using SVM and naive Bayes classifiers.it gives 81%,72% accuracy respectively. The rates obtained as sensitivity and specificity are acceptable.



Published by: The Mattingley Publishing Co., Inc.

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*

Incorporating the information encoded in emotion lexicons, mainly domain-specific lexicons, can drastically improve the accuracy of the proposed emotional analysis. According to that result, SVM is the best classifier for emotional analysis for this approach.

## REFERENCES

1. R. Watermeyer, "Social Networking Sites," in Encyclopedia of Applied Ethics, 2012.
2. S. Media and W. Sites, "Social Media Web Sites," Media, 2010.
3. L. Rainie and A. Smith, "Social networking sites and politics," Pew Res. Center's Internet Am. Life Proj., 2012.
4. R. Gross, A. Acquisti, and H. J. Heinz, "Information revelation and privacy in online social networks," in WPES'05: Proceedings of the 2005 ACM Workshop on Privacy in the Electronic Society, 2005, doi: 10.1145/1102199.1102214.
5. B. Osatuyi, "Information sharing on social media sites," Comput. Human Behav., 2013, doi: 10.1016/j.chb.2013.07.001.
6. H. Van der Heijden, "Factors influencing the usage of websites: The case of a generic portal in The Netherlands," Inf. Manag., 2003, doi: 10.1016/S0378-7206(02)00079-4.
7. Y. Liu, X. Chen, and X. Wang, "Evaluating government portal websites in China," in PACIS 2010 - 14th Pacific Asia Conference on Information Systems, 2010.
8. B. Detlor, M. E. Hupfer, U. Ruhi, and L. Zhao, "Information quality and community municipal portal use," Gov. Inf. Q., 2013, doi: 10.1016/j.giq.2012.08.004.
9. L. Yuan, C. Xi, and W. Xiaoyi, "Evaluating the readiness of government portal websites in China to adopt contemporary public administration principles," Gov. Inf. Q., 2012, doi: 10.1016/j.giq.2011.12.009.
10. Z. Huang and M. Benyoucef, "Usability and credibility of e-government websites," Gov. Inf. Q., 2014, doi: 10.1016/j.giq.2014.07.002.
11. K. Vance, W. Howe, and R. P. Dellavalle, "Social Internet Sites as a Source of Public Health Information," Dermatologic Clinics. 2009, doi: 10.1016/j.det.2008.11.010.
12. P. Gragg and C. L. Sellers, "Twitter," Law Library Journal. 2010, doi: 10.4324/9781315658643-26.
13. Twitter Inc., "About Twitter," Twitter About. 2014.
14. "Twitter and society," Choice Rev. Online, 2014, doi: 10.5860/choice.52-0916.
15. S. Kairam, M. J. Brzozowski, D. Huffaker, and E. H. Chi, "Talking in circles: Selective sharing in Google+," in Conference on Human Factors in Computing Systems - Proceedings, 2012, doi: 10.1145/2207676.2208552.
16. U. Mathur and R. Mack, "Facebook," in Ethics in Marketing: International Cases and Perspectives, 2012.
17. A. Java, X. Song, T. Finin, and B. Tseng, "Why We Twitter: Understanding Microblogging," Proc. 9th WebKDD 1st SNA-KDD 2007 Work. Web Min. Soc. Netw. Anal., 2007.
18. C. Honeycutt and S. C. Herring, "Beyond microblogging: Conversation and collaboration via twitter," in Proceedings of the 42nd Annual Hawaii International Conference on System Sciences, HICSS, 2009, doi: 10.1109/HICSS.2009.89.
19. A. O. Larsson and H. Moe, "Studying political microblogging: Twitter users in the 2010 Swedish election campaign," New Media Soc., 2012, doi: 10.1177/1461444811422894.
20. P. Ghosh, G. Schwartz, and S. Narouze, "Twitter as a powerful tool for communication between pain physicians during COVID-19 pandemic," Regional Anesthesia and Pain Medicine. 2020, doi: 10.1136/rapm-2020-101530.



*Published by: The Mattingley Publishing Co., Inc.*

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*

21. R. K. Bakshi, N. Kaur, R. Kaur, and G. Kaur, "Opinion mining and sentiment analysis," in Proceedings of the 10th INDIACom; 2016 3rd International Conference on Computing for Sustainable Global Development, INDIACom 2016, 2016, doi: 10.1145/2915031.2915050.
22. A. Shelar and C. Y. Huang, "Sentiment analysis of twitter data," in Proceedings - 2018 International Conference on Computational Science and Computational Intelligence, CSCI 2018, 2018, doi: 10.1109/CSCI46756.2018.00252.
23. A. Go, R. Bhayani, and L. Huang, "Twitter Sentiment Classification using Distant Supervision," Processing, 2009.
24. A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010, 2010, doi: 10.17148/ijarcce.2016.51274.
25. M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment in Twitter events," J. Am. Soc. Inf. Sci. Technol., 2011, doi: 10.1002/asi.21462.
26. M. Ghiassi, J. Skinner, and D. Zimbra, "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network," Expert Syst. Appl., 2013, doi: 10.1016/j.eswa.2013.05.057.
27. A. Sarlan, C. Nadam, and S. Basri, "Twitter sentiment analysis," in Conference Proceedings - 6th International Conference on Information Technology and Multimedia at UNITEN: Cultivating Creativity and Enabling Technology Through the Internet of Things, ICIMU 2014, 2015, doi: 10.1109/ICIMU.2014.7066632.
28. E. Kouloumpis, T. Wilson, and J. Moore, "Twitter Sentiment Analysis: The Good the Bad and the OMG!," in Int'l AAAI Conference on Weblogs and Social Media (ICWSM), 2011.
29. E. Martínez-Cámara, M. T. Martín-Valdivia, L. A. Ureña-López, and A. R. Montejo-Ráez, "Sentiment analysis in Twitter," Nat. Lang. Eng., 2014, doi: 10.1017/S1351324912000332.
30. H. Saif, Y. He, and H. Alani, "Semantic sentiment analysis of twitter," in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2012, doi: 10.1007/978-3-642-35176-1-32.
31. V. A. and S. S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques," Int. J. Comput. Appl., 2016, doi: 10.5120/ijca2016908625.
32. A. Giachanou and F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods," ACM Computing Surveys. 2016, doi: 10.1145/2938640.
33. S. M. Mohammad, S. Kiritchenko, and X. Zhu, "NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets," in *SEM 2013 - 2nd Joint Conference on Lexical and Computational Semantics, 2013.
34. B. Liu, "Sentiment analysis and opinion mining," Synth. Lect. Hum. Lang. Technol., 2012, doi: 10.2200/S00416ED1V01Y201204HLT016.
35. P. Palanisamy, V. Yadav, and H. Elchuri, "Serendio: Simple and practical lexicon based approach to sentiment analysis," *SEM 2013 - 2nd Jt. Conf. Lex. Comput. Semant., vol. 2, no. SemEval, pp. 543–548, 2013.
36. S. Das, R. K. Behera, M. Kumar, and S. K. Rath, "Real-Time Sentiment Analysis of Twitter Streaming data for Stock Prediction," Procedia Comput. Sci., vol. 132, no. Iccids, pp. 956–964, 2018, doi: 10.1016/j.procs.2018.05.111.
37. P. F. Pai and C. H. Liu, "Predicting vehicle sales by sentiment analysis of twitter data and stock market values," IEEE Access, vol. 6, pp.



*Published by: The Mattingley Publishing Co., Inc.*

*January-February 2020*
*ISSN: 0193-4120 Page No. 17702 - 17712*

57655–57662, 2018, doi: 10.1109/ACCESS.2018.2873730.

38. V. Hatzivassiloglou and K. R. McKeown, "Predicting the semantic orientation of adjectives," 1997, doi: 10.3115/979617.979640.
39. J. Rothfels and J. Tibshirani, "Unsupervised sentiment classification of English movie reviews using automatic selection of positive and negative sentiment items," CS224N-Final Proj. Rep., 2010.
40. T. Zagibalov and J. Carroll, "Automatic seed word selection for unsupervised sentiment classification of Chinese text," in Coling 2008 - 22nd International Conference on Computational Linguistics, Proceedings of the Conference, 2008, doi: 10.3115/1599081.1599216.
41. M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment strength detection for the social web," J. Am. Soc. Inf. Sci. Technol., 2012, doi: 10.1002/asi.21662.
42. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up?," 2002, doi: 10.3115/1118693.1118704.



*Published by: The Mattingley Publishing Co., Inc.*