



## SENTIMENT ANALYSIS ON ONLINE SOCIAL NETWORKING SITES USING SLANGS

**Dr.Thupakula Bhaskar**, Associate Professor, Dept. of Computer Engineering, Sanjivani College of Engineering, Kopergaon, Savitribai Phule Pune University, Maharashtra, [bhaskarcomp@sanjivani.org.in](mailto:bhaskarcomp@sanjivani.org.in)  
**Dr.Naresh Thoutam**, Associate Professor, Dept. of Computer Engineering , Sandip Institute of Technology and Research Centre, Nashik, Savitribai Phule Pune University, Maharashtra,[naresh1060@gmail.com](mailto:naresh1060@gmail.com)  
**Dr.B.J.Dange**, Associate Professor, Dept. of Computer Engineering, Sanjivani College of Engineering, Kopergaon, Savitribai Phule Pune University, Maharashtra, [dangebapusaheb@sanjivani.org.in](mailto:dangebapusaheb@sanjivani.org.in)  
**Dr. H. E. Khodke**, Assistant Professor, Dept. of Computer Engineering, Sanjivani College of Engineering, Kopergaon, Savitribai Phule Pune University, Maharashtra, [khodkehecomp@sanjivani.org.in](mailto:khodkehecomp@sanjivani.org.in)

**Abstract** - Numerous social media sites have been created and have grown to be significant components of contemporary living in the information and technology era we live in. Nowadays people share lots of information regarding their life and their opinions and reviews on various social media networking sites. These sites, therefore, provide people with an effective platform for expressing their thoughts, ideas, and sentiments. Such reviews and thoughts of people on various topics can affect them positively or negatively. Therefore, understanding the sentiments of the user is an important task. Numerous social networking sites have amassed a sizable quantity of user-generated data that provide new views for organisations and states. Extraction of meaningful information is difficult from a big amount of data. Sentiment analysis provides a method of analyzing the sentiments, emotions, and opinions of a user. Finding the logical link between mood statements and the topic is crucial for improving sentiment analysis. The study adds to the body of knowledge on sentiment analysis and outlines several benefits and uses for sentiment analysis. A rule-based textual approach is used that is dependent upon the noisy non-grammatical text, the use of unpleasant words, and short language that are widely used today on various social media platforms known as slang. The paper presents the breakdown structure of the sample system where sentiment analysis can be used and ways to achieve it.

**Keywords** - Sentiment Analysis, Tokenization, Slangs, Stanford NLP.

### I. INTRODUCTION

Nowadays the Internet is widely used and is one of the most important mediums of communication and working along with which users can get access to various services and also post their opinions and views regarding a particular subject. These reviews and opinions posted by users are a heterogeneous collection of data, and they can be further then used to identify the sentiments of users. Where sentiment analysis, also referred to as opinion mining, is a method of using Natural Language Processing (NLP) to find out how people feel and what they think. It categorizes and clarifies the opinions of users behind every product, idea, service, or news. Sentiment analysis describes sentiments from any verbal or non-verbal communication. In various applications, there is a need to classify the sentiments into positive, negative, and neutral accordingly, which provides an effective and powerful functionality for marketing analysis, competitive analysis, and detection of inefficient rumors for risk management.

Appropriate techniques are required to classify the sentiments accordingly. Making an analysis of favorable and unfavorable opinions, positive and negative opinions, and efficient and inefficient opinions is a task that requires deep understanding and intelligence of the textual context and knowledge of the domain. Making use of appropriate techniques and mechanisms a large amount of information can be transformed into the information that will be supported for operational, managerial, and strategic decision-making processes.



In today's world of technology and social media, the posts and reviews of users are considered as vital parts for further analysis and recommendations. These sentiments being expressed by the users are sometimes genuine and sometimes even fake and abusive, which can lead to negative results. Today users also make use of short and informal words to express their views which are difficult to evaluate. Therefore there comes a need for analyzing the sentiments and reviews of the user on Online Social Networking Sites using slang and categorizing them into positive, negative, and neutral sentiments. Objectives of the Proposed System are stated below:

1. Emerging a user-friendly scheme for Sentiment analysis.
2. The objective is to provide the user and applications with the capacity to classify the sentiments of users.
3. Analyse the topics on which users express their reviews.
4. Enables individuals, organizations, and government systems to gain benefits from reviews of users.

Sentiment analysis, also known as opinion mining, is the study of opinions, feelings, assessments, judgments, attitudes, and emotions towards various things, such as objects, services, groups, people, issues, events, subjects, and the traits of the people. The terms "sentiment analysis" and "opinion mining," as well as "opinion extraction," "subjectivity analysis," and "emotion analysis," among others, are all used to describe sentiment analysis. There are different levels of sentiment analysis. They are as follows:

1. **Sentence Level:** The challenge at this phrase level is determining whether each sentence conveys a favorable, negative, or indifferent view. This degree of analysis is closely linked to the categorization of subjectivity, which separates words that express true information from those that express biased viewpoints and opinions.
2. **Document Level:** At this stage, the goal is to decide if an entire assessment paper conveys a positive or negative disposition. The program, for instance, can determine whether a product evaluation is usually biased in favour of or against the item. This cycle is generally known as "report level feeling characterization". At this phase of study, it is believed that each paper will communicate concepts regarding a specific object. (e.g., a single product). It therefore does not apply to papers that evaluate or compare multiple organisations.
3. **Entity and Aspect Level:** The studies at the document or phrase levels do not specifically show what individuals enjoyed and hated. The aspect level does a more precise analysis. The aspect level directly examines the opinion itself rather than considering the language construct. It is predicated on the notion that a target and a sentiment make up an opinion. Without knowing who it is intended for, an opinion is of minimal value. Understanding the importance of opinion goals also helps us better understand the problem with mood analysis.

Sentiment Analysis is being used in various applications such as Review related websites, Business and government Intelligence, Text classification for news organizations, email spam filtering, and patent analysis were created for a wide range of applications. Sentiment analysis is necessary due to the difficulties presented by social media, including target identification, negation, contextual information, relevance, volatility over time, and opinion consolidation. One of the challenges of Sentiment Analysis on online social networking sites is the use of slang and informal language. Slangs are words and phrases that are not typically used in formal language and are often used in online communication to convey emotions, humor, or sarcasm. The use of slang can significantly impact the accuracy of sentiment analysis as it may have multiple meanings and can be context-dependent.

## II. LITERATURE REVIEW



**Sentiment Analysis of Users on Social Networks: Overcoming the challenge of the Loose Usages of the Algerian Dialect [1]** The paper demonstrates a method for sentiment analysis on Online social network sites using the reviews posted by users. The writers have addressed the issue of sentiment analysis of remarks left by Facebook users from Algeria on public sites from different businesses. Prior to creating a collection of more than 25000 sentiment-annotated remarks, the researchers first examined the uniqueness of AlgD and the language habits of Algerians on social media. The pre-processing portion of this dataset was the next stage, and a Naive Bayes predictor was used to assess each step's effect on the data's quality. Two various types of profound brain networks have been utilized in their work. The first is a (profound) MLP, and the ideal setting created a precision of 81.6%. The second is a convolutional brain organization, which achieved 89.5% precision.

**Sentiment Analysis in Social Media and Its Application: Systematic Literature Review [2]** This study demonstrates the outcome of a large portion of the articles that involved the assessment vocabulary technique to break down message opinion in virtual entertainment. The information was extricated from microblogging locales, primarily Twitter, and sentiment analysis applications can be seen in global events, healthcare, politics, and business. In the performed thorough literature survey, studies on mood analysis in social media are addressed. The three advances made by the article are as follows. There are many methods that have been developed by researchers, but SentiWordnet and TF-IDF, as well as Nave Bayes and SVM, are the most frequently used in Lexicon-based methods and machine learning, respectively. Combining vocabulary and machine learning techniques is advised in order to increase the result's quality and precision. Second, they specify the most typical category of social media site from which data can be extracted for mood analysis. Twitter is the most famous social media platform from which data can be gleaned. Twitter is most frequently used as the social media setting in the articles we examined. others. showed the use of opinion analysis in social media, third. Sentiment analysis has a wide range of applications and can be used to improve quality and strategy in business, predict election results politically, raise awareness of the value of data security, gauge public opinion of a certain sport, and enhance location and disaster response, among other things. This demonstrates the critical function mood analysis plays in comprehending people's views and assisting in decision-making.

**A Survey of Sentiment Analysis from Social Media Data [5]** In today's world of social media and machines, machines are being used on a large scale to provide accurate interpretations of the opinions of the users. There comes a need of analyzing the sentiments of users for knowing what people think and like. This paper thus provides the result of the process of collecting the data from social media platforms and then detecting the reviews and choices of the users. This paper has surveyed various techniques for communalizing user data. This paper also addresses the various ways through which the sentiments can be categorized and compared. The two primary techniques for mood analysis are covered in the article. Both lexicon-based and classification-based methods are used. Here while calculating the sentiments two theories are being considered, first that sentiments are independent of the context, and the second states that the sentiments can be articulated through numbers. The paper shows the conclusion that the accuracy for the mining of the sentiments was 84.29%.

**Sentiment Analysis of Social Media Posts about Events [9]** The demand for mood analysis is growing daily as a result of advances in technology and the use of social media. The existing approaches have certain limitations such as the analysis of textual content without considering the graph-based approach. The proposed paper presents the solution to overcome the above limitation. It makes use of event detection in order to extract the topics that are interesting and are required for sentiment analysis. Support Vector Machine (SVM), a classification algorithm in machine learning, and Logistic Regression are together being used for the classification of sentiments. As compared to Logistic Regression, Support Vector Machine (SVM) provides better evaluation scores. The proposed solution consists of two classes that are Event



Detection(ED) and Sentiment Analysis(SA). The above approach consists of modules such as Training Engine, Storage Module, WEB Module, etc. The proposed approach provides an accurate event and detection of types of sentiments posted on social media platforms. The paper represents how the sentiments are being analyzed from the social media data and provides high-accuracy results with the help of the above-stated approach.

### III. PROPOSED SYSTEM

The below system consists of four main modules. They are as stated in the below breakdown structure of the system, which describes the working of every module in the system.

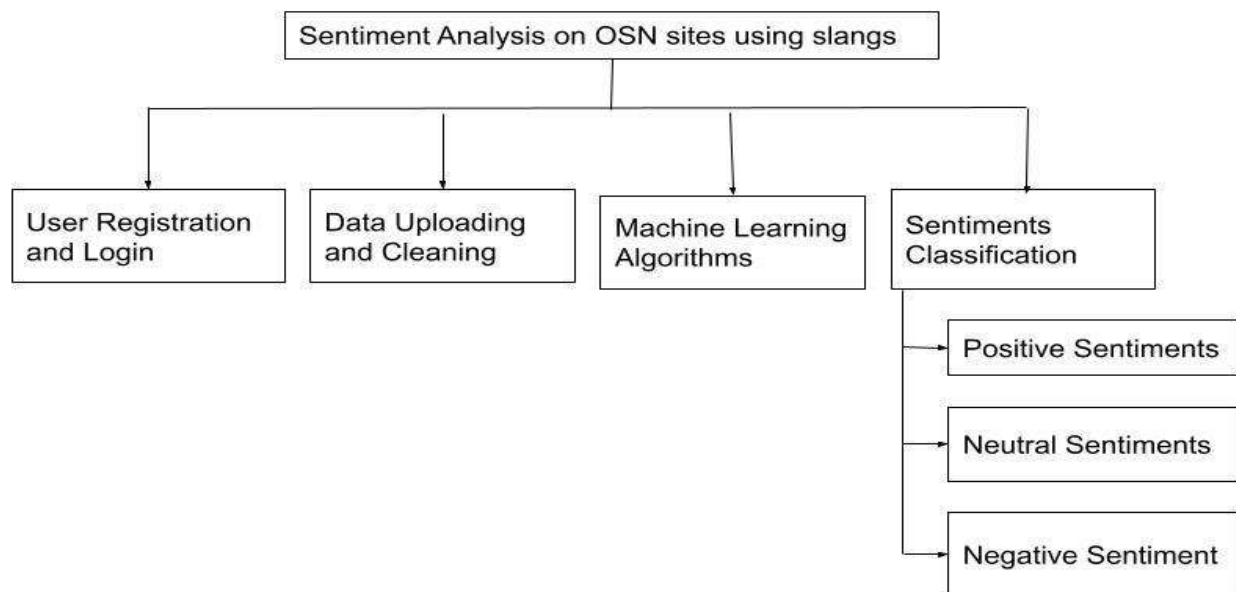


Figure 1: System Breakdown Structure

As described in the above breakdown structure of the system, the four modules of the system are as below:

1. **User Registration and Login:** In this module, the user registers themselves to log in and then perform their tasks. The user posting their reviews and comments needs to register first. For the process of registration, the basic information of the user such as name, email, password and address will be required. During login, the system will then evaluate the login credentials such as username and password. If correct the user will get logged in to the system.

2. **Data uploading and cleaning:** Initially the data of various sentiment expressions are added to the database and then the data is trained to further classify the sentiments of the user. The data according to its category will be added to its respective dataset such as positive, negative, neutral, abuse, and slang words. The newly added dataset will be loaded into the system for further classification.
3. **Machine learning Algorithms:** The Naive Bayes Classification method will be employed here as the machine learning method. The Naive Bayes algorithm will be a previously taught algorithm for conducting mood categorization. The algorithm will be trained in such a way that it will recognize the individual word and compare it with the already existing and loaded data.
4. **Sentiment Classification:** According to the study and the algorithms applied the sentiments will be further classified into positive, negative, and neutral sentiments. The new opinions/reviews posted by the user will be compared with the initial stored and trained data. The sentiments are classified as:
  - i) **Positive Sentiments:** Positive sentiments define the feelings or comments being expressed by the users that are favorable or have an optimistic attitude.
  - ii) **Neutral Sentiments:** Neutral sentiments define the feelings or comments of users that are neither positive nor negative and are therefore defined as neutral tones.
  - iii) **Negative Sentiments:** Negative sentiments define the feelings or comments being expressed by the users that are unfavorable or have a pessimistic attitude.

#### IV. METHODOLOGY

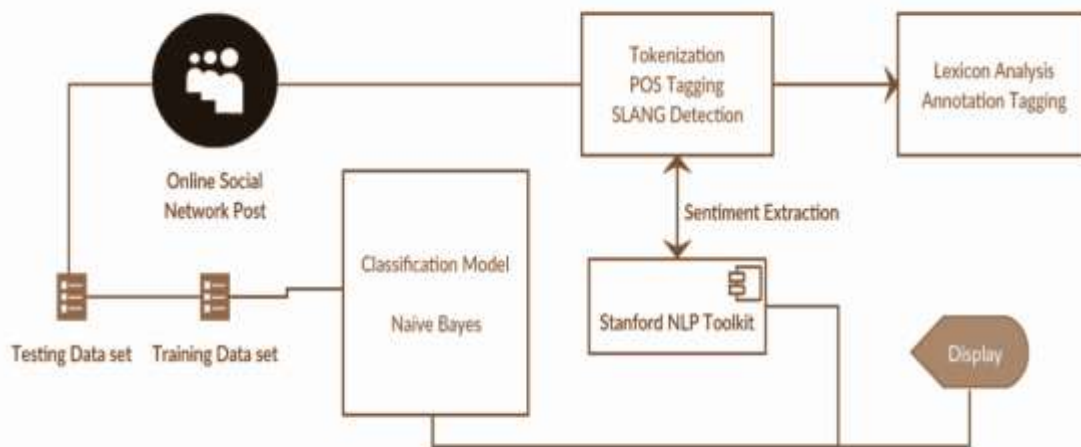


Figure 2: System Architecture

The above diagram represents the system architecture that consists of the main working modules and methodology. They are as follows:



1. **Online Social Network Post:** Initially, the user needs to post their reviews or comments in order to check for further analysis. These posts of users or comments will be analyzed in further classification modules.
2. **Training and Testing Data set:** The data collection must first be educated and evaluated using different kinds of emotions in order to categorise user messages. The newly added dataset of the words will be trained and tested so that further it classifies the sentiments accordingly.
3. **Classification Module:** Classification Module consists of the implementation of the Naive Bayes Classification Algorithm which classifies the reviews or posts of users into positive, negative, or neutral. Later on, if the tested comment turns positive will be posted accordingly, and if the comment is negative and crosses the set threshold value, the post will be discarded and will not be posted by the user.
4. **Tokenization:** The process of tokenization breaks down the stream into sentences, words, and symbols that are known as tokens. This step explores the words in a sentence. It identifies the significant words and then is used for classification accordingly.
5. **POS Tagging:** POS (Part-of-Speech) Tagging is used for sentiment analysis where, after the tokenization, the split tokens are tagged with appropriate parts of speech. This module assumes an imperative part as it recognizes the piece of speech of the split tokens.
6. **SLANG Detection:** TO handle and identify the informal words that are used by the user can be identified with the slang detection module. This module is important because slang words do carry different meanings and these words must get classified according to their sentiments.
7. **Stanford NLP Toolkit:** The Stanford NLP (Natural Language Processing) Toolkit consists of resources and tools for recognizing the entity, parts of speech tagging, and parsing. This module consists of pre-trained algorithms that will process and analyze the text data.
8. **Lexicon Analysis and Annotation tagging:** This module is used to extract the information from the text posted by the user. The words are identified in the lexicon analysis step and the annotation tagging module makes use of the sentiment score for each token and then identifies the sentiment behind the user's posts.

## V. NAIVE BAYES CLASSIFICATION ALGORITHM

An method for guided learning that is built on the Bayes theorem is the naive Bayes categorization algorithm. The guileless Bayes technique is utilized to tackle order issues and is fundamentally utilized in text classification that utilizes a high layered exchange dataset. The naïve Bayes algorithm is one of the most basic and effective Classification algorithms for developing fast machine learning models and generating forecasts quickly and simply. It is a statistical predictor that makes predictions about the outcome based on how likely an item is to be present. The Naive Bayes Algorithm is used for mood analysis, categorising documents, and spam filtering.

The words in the written document that occur in the lists of positive words, negative words, and neutral words can all be identified with the aid of this bag-of-words model. The overall score of the text, for instance, will be revised with +1 if the term occurs on the roster of favourable phrases, and vice versa. Accordingly, if the final total score is positive, the text will be further categorised as positive; if the final total is negative, the text will be categorised as negative; otherwise, it will be categorised as impartial. The Naive Bayes Algorithm is summarized in the below steps:

**Step 1:** Prepare the training data set by separating the data into various classes. In the proposed system, the classes are positive, negative, and neutral sentiments.

**Step 2:** A guided learning algorithm built on the Bayes theorem is the naive Bayes categorization algorithm. The naïve Bayes method is employed to address classification issues, and it is frequently used in text



*categorization when dealing with high-dimensional trading datasets. The naïve Bayes algorithm is one of the most straightforward and effective Classification algorithms for producing fast machine learning models and generating forecasts. It is a stochastic predictor that forecasts the outcome in light of the likelihood of an item. Sentiment analysis, article classification, and spam filtering are a few applications of the Naive Bayes algorithm.*

*With the aid of this bag-of-words model, we can determine which word from a written document occurs in the lists of positive, negative, and neutral terms. For instance, if a word shows on the list of favourable terms, the text's overall score will be changed to +1, and vice versa. So, if the final overall number is positive, the text will be further classified as positive; if it is negative, it will be classified as negative; otherwise, it will be classified as impartial.*

*Make a previous probability calculation for each subgroup. The chance that an article corresponds to a specific class is known as the antecedent probability. A class's prior likelihood is determined by dividing the overall number of training instances by the number of training examples in that class.*

*Step 3: For each class, the conditional odds of each phrase are computed. A word's conditional likelihood measures how likely it is to appear in a specific mood class. In order to calculate this likelihood, we tally the instances of each word in each class and split that number by the total number of terms in that class.*

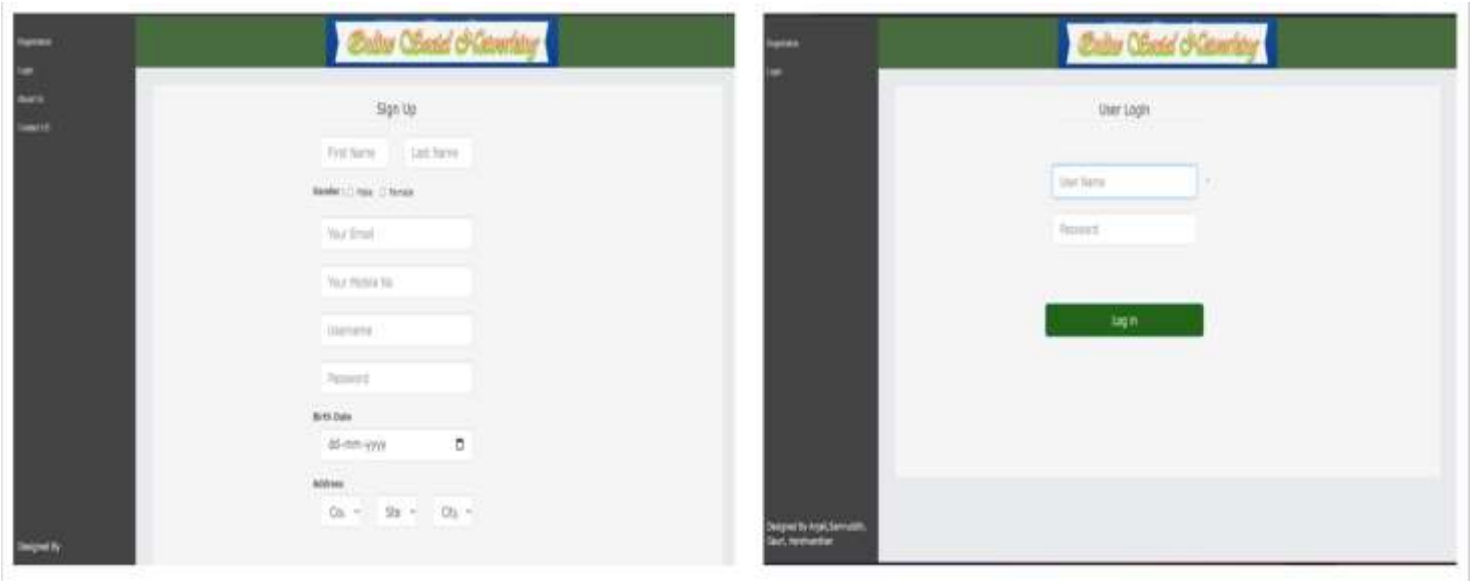
*. We apply some form of smoothing to handle the case where a word appears in one class, but not in another class.*

*Step 4 According to the provided terms in the posts, we determine the likelihood that a particular post corresponds to each class in order to categorise it. We multiply the conditional odds of each word in the provided text, given each class, and then multiply that result by the antecedent probability of the class to arrive at the probability..*

*Step 5: Once the probabilities of each class is being calculated, then the class with the highest probability is chosen as a prediction for the sentiment of the post.*

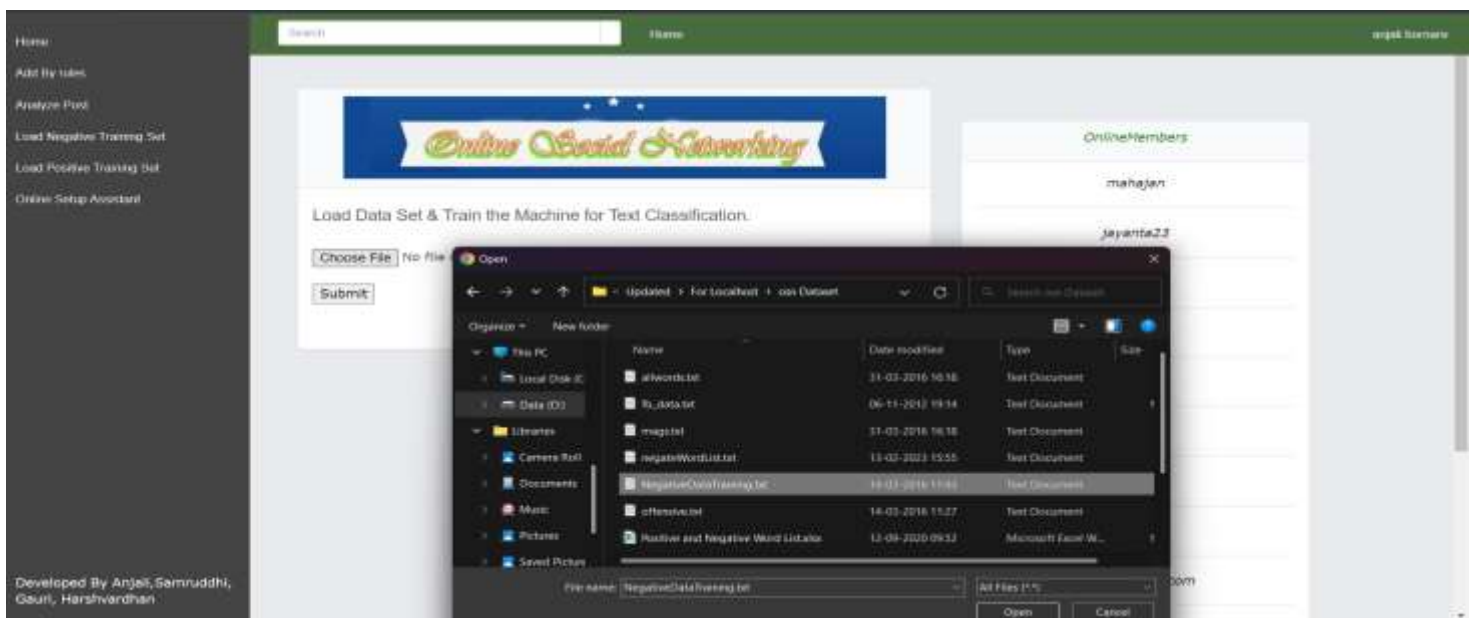
*Step 6: Finally, The evaluation of the performance of the classifier using a test data set is done. The computation of the accuracy of the classifier by comparing the predicted sentiment of each post to the actual sentiment of that post in the test data set is done.*

## V. RESULTS AND DISCUSSIONS



**Figure 3: User Registration and Login Page**

In above **Figure 3**, there are Login and Registration pages for the user. This application is developed to be used by many different people. The Registration page is where users go to create their accounts. This typically includes basic information such as first name, last name, username, password, etc. Once the user enters all the necessary information, they can submit the page and their account will get created. After creating the account, the user can log in using the Login page. This is where the user will enter the username and password to access the Sentiment Analysis System. Once logged in, the user can be able to access the features and functions of the application.

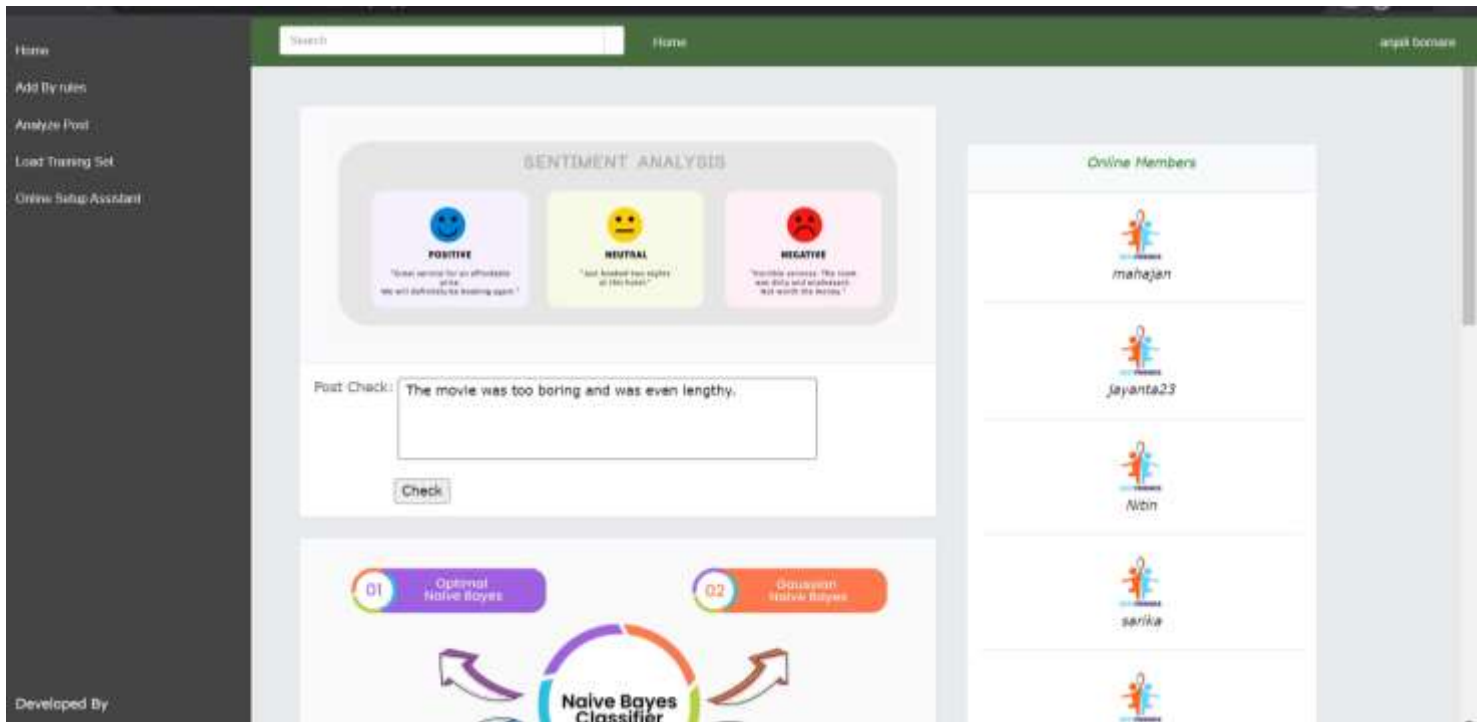


**Figure 4: Loading Training Set**



In **Figure 4**, the feature of Loading Training Set can be used by the user to Load the various types of datasets such as Positive, Negative, Offensive, Slang words, etc. The newly added dataset will be helpful for the system to analyze the type of sentiment with more accuracy.

The entered text will be compared with the data in the loaded dataset. Once the user loads the datasets, the entered text as the post is been compared with the data in the loaded dataset. If the entered text contains words or phrases that match the loaded dataset, the proposed system model can use this information to enhance its classification accuracy.



**Figure 5: Analyzing Post**

In the above **Figure 5**, the user needs to enter the text that needs to be checked, whether the submit check option indicates that the content is favorable, negative, or indifferent. When you select the Check option, the result is indicated as either positive, negative, or indifferent. The output consists of splitting tokens, parts-of-speech tagging, and then analyzing the types of sentiment. As shown in the above figure, the comment to be checked is **“The movie was too boring and was even lengthy”**, The above text consists of negative words such as boring, and lengthy. These words fall under the dataset of negative words list, the other remaining words are neutral words. Therefore the count of negative words is more and gives the output as negative.

## Tokenization

Tokenize Data :

The  
movie  
was  
too  
boring  
and  
was  
even  
lengthy.

## Parts of Speech Tagging and Sentiment Analysis using Stanford Natural Language Processing

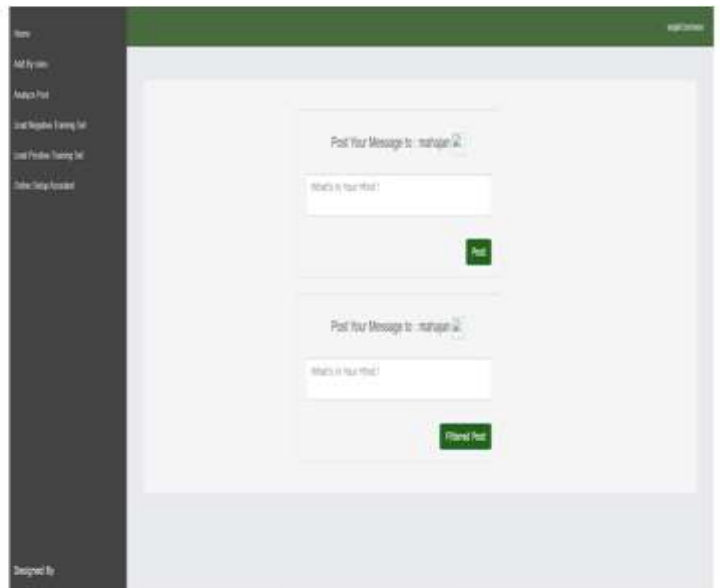
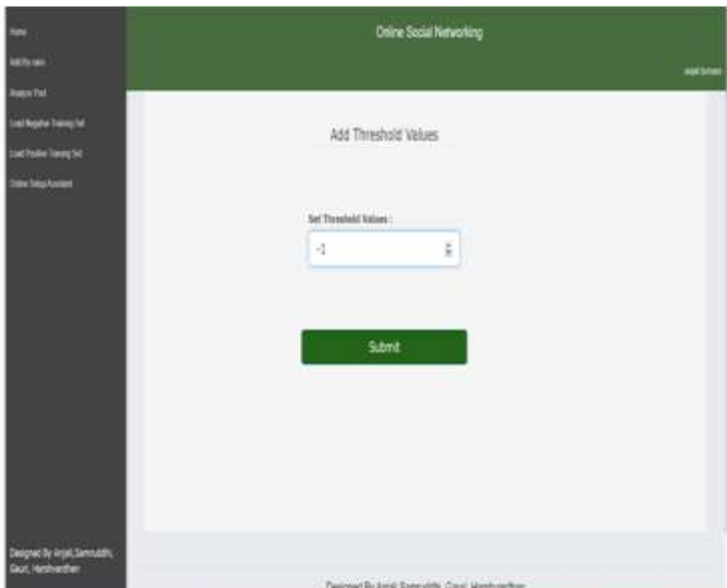
Sentence #1 (10 tokens): The movie was too boring and was even lengthy. [Text="The CharacterOffsetBegin=1 CharacterOffsetEnd=4 PartOfSpeech=DT] [Text="movie CharacterOffsetBegin=5 CharacterOffsetEnd=10 PartOfSpeech=NN] [Text="was CharacterOffsetBegin=11 CharacterOffsetEnd=14 PartOfSpeech=VBD] [Text="too CharacterOffsetBegin=15 CharacterOffsetEnd=18 PartOfSpeech=RB] [Text="boring CharacterOffsetBegin=19 CharacterOffsetEnd=25 PartOfSpeech=JJ] [Text="and CharacterOffsetBegin=26 CharacterOffsetEnd=29 PartOfSpeech=CC] [Text="was CharacterOffsetBegin=30 CharacterOffsetEnd=33 PartOfSpeech=VBD] [Text="even CharacterOffsetBegin=34 CharacterOffsetEnd=38 PartOfSpeech=RB] [Text="lengthy CharacterOffsetBegin=39 CharacterOffsetEnd=46 PartOfSpeech=JJ] [Text=" CharacterOffsetBegin=46 CharacterOffsetEnd=47 PartOfSpeech=.] (ROOT (S (NP (DT The) (NN movie)) (VP (VP (VBD was) (ADJP (RB too) (JJ boring))) (CC and) (VP (VBD was) (ADJP (RB even) (JJ lengthy)))) (. .)) root(ROOT-0, boring-3) det(movie-2, The-1) subj(boring-3, movie-2) cop(boring-3, was-3) advmod(boring-3, too-4) cop(lengthy-9, was-7) advmod(lengthy-9, even-8) conj\_and(boring-3, lengthy-9) The top level annotation [Text=" The movie was too boring and was even lengthy. Tokens=[The-1, movie-2, was-3, too-4, boring-5, and-6, was-7, even-8, lengthy-9, -10] Sentences=[The movie was too boring and was even lengthy.] (ROOT (NP (DT The) (NN movie)) (@5 (VP (@VP (VP (VBD was) (ADJP (RB too) (JJ boring))) (CC and) (VP (VBD was) (ADJP (RB even) (JJ lengthy)))) (. .)) Sentiment is 1 [Text=" The movie was too boring and was even lengthy. CharacterOffsetBegin=1 CharacterOffsetEnd=47 Tokens=[The-1, movie-2, was-3, too-4, boring-5, and-6, was-7, even-8, lengthy-9, -10] TokenBegin=0 TokenEnd=10 SentenceIndex=0 Tree=(ROOT (S (NP (DT The) (NN movie)) (VP (VP (VBD was) (ADJP (RB too) (JJ boring))) (CC and) (VP (VBD was) (ADJP (RB even) (JJ lengthy)))) (. .)) CollapsedDependencies=-> boring/J (root) -> movie/NN (nsubj) -> The/DT (det) -> was/VBD (cop) -> too/RB (advmod) -> lengthy/J (conj\_and) -> was/VBD (cop) -> even/RB (advmod) BasicDependencies=-> boring/J (root) -> movie/NN (nsubj) -> The/DT (det) -> was/VBD (cop) -> too/RB (advmod) -> lengthy/J (conj\_and) -> movie/NN (nsubj) -> was/VBD (cop) -> even/RB (advmod) CollapsedCCProcessedDependencies=-> boring/J (root) -> movie/NN (nsubj) -> The/DT (det) -> was/VBD (cop) -> too/RB (advmod) -> lengthy/J (conj\_and) -> movie/NN (nsubj) -> was/VBD (cop) -> even/RB (advmod) BiannotatedTree=(ROOT (S (NP (DT The) (NN movie)) (@5 (VP (@VP (VP (VBD was) (ADJP (RB too) (JJ boring))) (CC and) (VP (VBD was) (ADJP (RB even) (JJ lengthy)))) (. .)) AnnotatedTree=(ROOT (NP (DT The) (NN movie)) (@5 (VP (@VP (VP (VBD was) (ADJP (RB too) (JJ boring))) (CC and) (VP (VBD was) (ADJP (RB even) (JJ lengthy)))) (. .)) ClassName=Negative] Sentiment of "The movie was too boring and was even lengthy." is: 1 (i.e., Negative)

Consider rewriting your post as it doesnt sound too appealing !

Figure 6: Result of given Post (Negative Sentiment)

In above Figure 6, the output is shown once the comment is tested. The output consists of the tokenization part, where the tokens in the given string are split into individual tokens. Each individual token is tagged with its respective parts of speech. The verb and other respective words that defines the sentiments are being compared with the previously loaded dataset. On comparing with the dataset, the number of words are being calculated, and the words that range with the highest sum are classified into their respective sentiment.

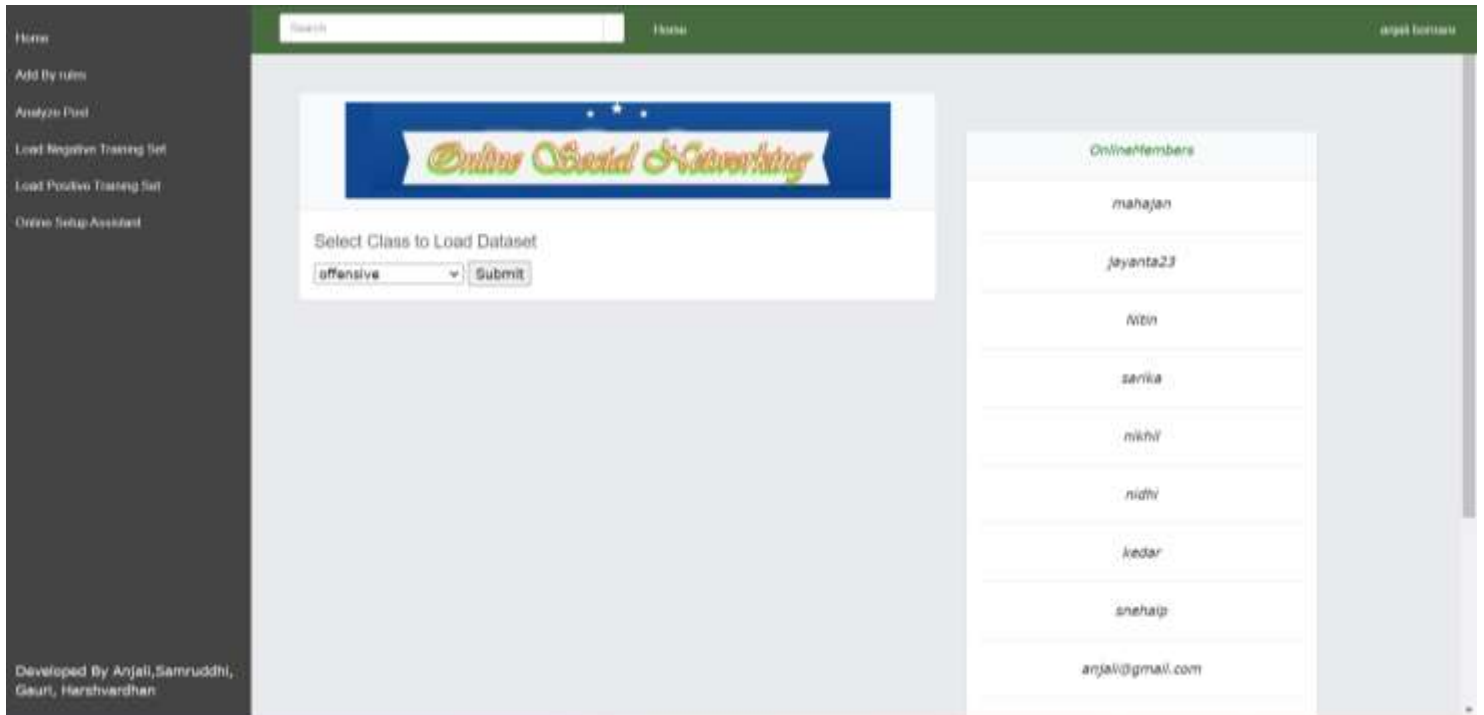
As in above Figure 6, the words from the Negative dataset have the highest sum therefore, the output is given as Negative.





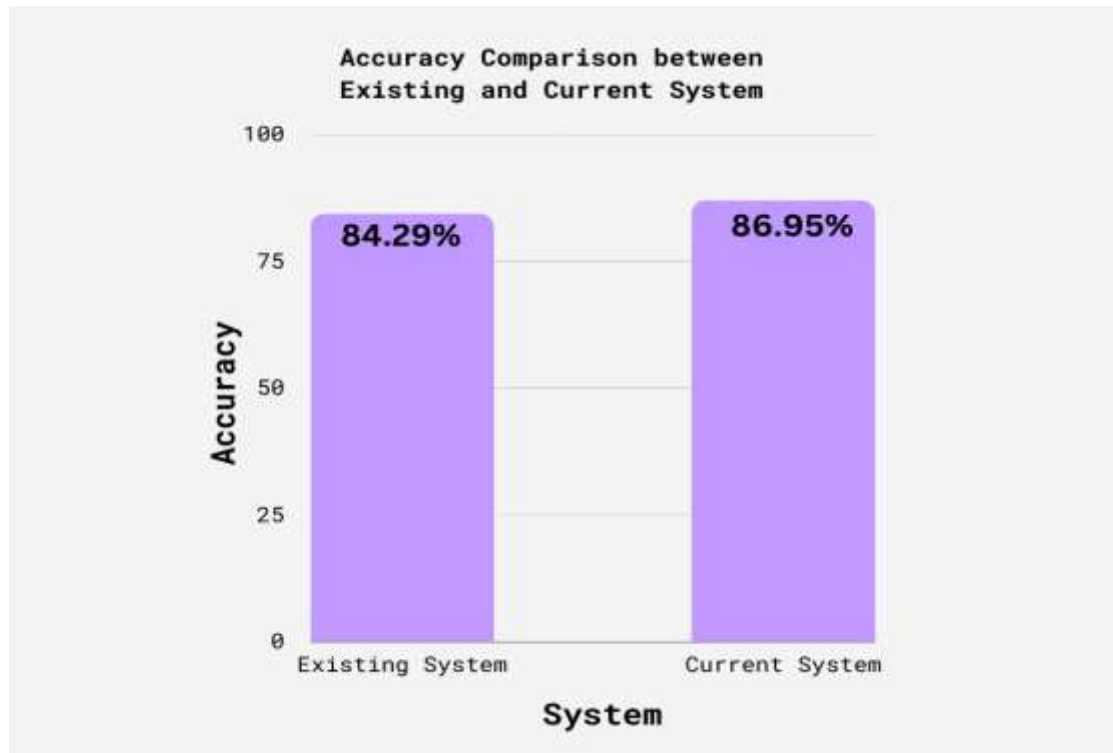
### Figure 7: Adding Rules

As shown in above **Figure 7**, the user can set a threshold value, through which the user can restrict the offensive and negative incoming messages that are above the set threshold value. For example, if the user set its threshold value as -1, once the comment exceeds the set threshold value, the post will be discarded automatically. The next is the option for sending messages to the registered user. The user has the privilege to first check the message that will be sent, if it is positive or less than the threshold value it will get sent to the receiver or else not.



**Figure 8: Selecting Class to Load Dataset**

As shown in above **Figure 8**, the user has the provision to load the dataset other than positive and negative such as offensive, slang words, taboo words, etc. Once this dataset is loaded and submitted, the comments of the user that consists of above stated words will also be checked and accordingly, the result will get generated as positive, negative, or neutral sentiments.



**Figure 8: Accuracy Comparison between Existing and Current System**

As mentioned in above Figure 8, the existing system[5] provides an accuracy of 84.29%, which uses the Naïve Bayes Classifier for identifying the type of tweets. The current developed system provides an accuracy of 86.95% that uses the Naïve Bayes Classifier Algorithm and Stanford NLP module for the classification of sentiments.

## VI. CONCLUSION

In conclusion, sentiment analysis on online social networking sites using slang is a challenging task due to the informal nature of the language used on these platforms. However, with the use of natural language processing techniques, it is possible to accurately identify the sentiment of text written in slang. The biggest venue where people communicate their ideas or, you could say, thoughts in the formats are social media accounts. Social media platforms must be thoroughly examined if sentiment analysis is to be studied in this area.

This paper states that using a combination of machine learning algorithms and pre-trained language models can achieve high accuracy in sentiment analysis on social media platforms. Additionally, the use of slang-specific lexicons and dictionary-based approaches can also improve sentiment analysis results.

### Availability of Data and Materials:

Dataset from kaggle datasets & also used few more datasets & analyzed during the current study. Link of Datasets Used in this Research Article:

[https://drive.google.com/drive/folders/15KViiV7aIBUVs9JLlnAgE8yJT-\\_EHNh](https://drive.google.com/drive/folders/15KViiV7aIBUVs9JLlnAgE8yJT-_EHNh)



## REFERENCES

1. Assia Soumeura, Mheni Mokdadi, Ahmed Guessoum, Amina Daoud, "Sentiment Analysis of Users on Social Networks: Overcoming the challenge of the Loose Usages of the Algerian Dialect," The 4th International Conference on Arabic Computational Linguistics (ACLing 2018), November 17-19 2018, Dubai, United Arab Emirates, doi: 10.1016/j.procs.2018.10.458
2. Zulfadzli Drugs, Haliyana Khalid, "Sentiment Analysis in Social Media and Its Application: Systematic Literature Review," The Fifth Information Systems International Conference 2019, doi: 10.1016/j.procs.2019.11.174
3. M. Kavitha, B. B. Naib, B. Mallikarjuna, R. Kavitha and R. Srinivasan, "Sentiment Analysis using NLP and Machine Learning Techniques on Social Media Data," 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, India, 2022, pp. 112-115, doi: 10.1109/ICACITE53722.2022.9823708.
4. R. Singh and P. Sharma, "An Overview of Social Media and Sentiment Analysis," 2021 5th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2021, pp. 1-4, doi: 10.1109/ISCON52037.2021.9702359.
5. K. Chakraborty, S. Bhattacharyya and R. Bag, "A Survey of Sentiment Analysis from Social Media Data," in IEEE Transactions on Computational Social Systems, vol. 7, no. 2, pp. 450-464, April 2020, doi: 10.1109/TCSS.2019.2956957.
6. N. R. Kasture and P. B. Bhilare, "An Approach for Sentiment Analysis on Social Networking Sites," 2015 International Conference on Computing Communication Control and Automation, Pune, India, 2015, pp. 390-395, doi: 10.1109/ICCUBEA.2015.82.
7. S. Dhawan, K. Singh and P. Chauhan, "Sentiment Analysis of Twitter Data in Online Social Network," 2019 5th International Conference on Signal Processing, Computing and Control (ISPCC), Solan, India, 2019, pp. 255-259, doi: 10.1109/ISPCC48220.2019.8988450.
8. S. Mittal, A. Goel and R. Jain, "Sentiment analysis of E-commerce and social networking sites," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2016, pp. 2300-2305.
9. A. Petrescu, C. -O. Truică and E. -S. Apostol, "Sentiment Analysis of Events in Social Media," 2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP), Cluj-Napoca, Romania, 2019, pp. 143-149, doi: 10.1109/ICCP48234.2019.8959677.
10. S. Dutta, D. Sarkar, S. Roy, D. K. Kole and P. Jana, "A Study on Herd Behavior Using Sentiment Analysis in Online Social Network," 2021 International Conference on Communication, Control and Information Sciences (ICCISc), Idukki, India, 2021, pp. 1-6, doi: 10.1109/ICCISc52257.2021.9484918.
11. E. Ishchukova, V. Salmanov, A. Antonenko and D. Kryuk, "Algorithms for Automated Sentiment Analysis of Posts in Social Networks," 2020 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT), Yekaterinburg, Russia, 2020, pp. 0361-0363, doi: 10.1109/USBREIT48449.2020.9117684.
12. A. K. Goel and K. Batra, "A Deep Learning Classification Approach for Short Messages Sentiment Analysis," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2020, pp. 1-3, doi: 10.1109/ICSCAN49426.2020.9262430.
13. Q. Hou, M. Han, and Z. Cai, "Survey on data analysis in social media: A practical application aspect," in Big Data Mining and Analytics, vol. 3, no. 4, pp. 259-279, Dec. 2020, doi: 10.26599/BDMA.2020.9020006.
14. D. Tanna, M. Dudhane, A. Sardar, K. Deshpande, and N. Deshmukh, "Sentiment Analysis on Social Media for Emotion Classification," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2020, pp. 911-915, doi: 10.1109/ICICCS48265.2020.9121057.
15. V. Tyagi, A. Kumar and S. Das, "Sentiment Analysis on Twitter Data Using Deep Learning approach," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2020, pp. 187-190, doi: 10.1109/ICACCCN51052.2020.9362853.



Industrial Engineering Journal

ISSN: 0970-2555

Volume : 52, Issue 11, November : 2023