IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

Sentiment Analysis of Social Media- A Survey of Methods

Aman Kashyap[1], Alika Abdi[2], Amisha, Yash Bihare[3], Mr. Devesh Garg[4]
Studnets, Department of Computer Science and Engineering[1,2,3]
Associate Professor, Department of Computer Science and Engineering[4]
Raj Kumar Goel Institute of Technology, Ghaziabad, Uttar Pradesh, India

Abstract: *The world is evolving quickly and with amazing inventiveness. Social media sites like Twitter, Facebook, and Google+ are being used more often by users to share and express their opinions on a variety of subjects, participate in online debates with diverse communities, and communicate with a large audience. This paper's primary goal is to review current techniques for social media sentiment analysis and to provide a theoretical comparison of cutting-edge methodologies. Recently, sentiment analysis researchers have focused on looking at attitudes on a range of subjects, including films, products, social media, and common social issues. Users frequently use social media as a venue to voice their ideas. This research focuses mostly on social media sentiment analysis, which may be used to determine whether or not opinions are good or negative. Sentiment analysis, commonly referred to as open mining, is a technique for identifying a text's emotional undertone. This article contrasts and compares open mining techniques for social media sentiment research. We study the sentiment of social media data streams using a variety of machine learning methods, including Linear SVC-, Logistic Regression, and BernoulliNB*

Keywords: Open Mining; Machine Learning; Logistic Regression; Support Vector Machine; Sentiment Analysis; Linear Regression

I. INTRODUCTION

It's in our nature to take into account the opinions of others before making a decision.By taking into account the circumstance from the perspective of another individual, a better choice can be made. Businesses and organisations are interested in finding out what the public thinks about their goods and services. People are curious about what other people think of the things they are considering purchasing. People used to ask their friends and relatives' opinions in the past, and organisations would use focus groups to produce polls. Because of how quickly these networks have grown, individuals and organisations use the information provided by social media sites like Twitter, Facebook, and Instagram to aid in decision-making. People also utilise these sites to read product reviews and to offer their opinions on widely discussed issues[6]. Internet users now have access to enormous amounts of data because of the advancement of web technology, and enormous amounts of data are produced daily as well. Such innovations now have an easy and cost-free platform to advance thanks to the internet. The gathering and analysis of consumer input on goods, services provided by the government, and other topics is substantially aided by social media platforms. This analysis is conducted using sentiment analysis.

Sentiment analysis uses intensely emotional or passionate reactions to a record, communication, or event to try to ascertain the attitude of a speaker, author, or other topic. It is a technique that uses Natural Language Processing (NLP) to extract, convert, and analyse views from a text and divide them into positive and negative sentiment.Sentiment analysis's main objective is to classify text documents or short sentences by determining their polarity. Sentiment polarity can be categorised as "positive", "negative", or "impartial" (neutral).Sentiment mining can be performed on three levels as follows:

At the document level, sentiment analysis provides complete classification of a document as "positive," "negative," or "neutral."

Sentence by sentence classification of emotion: At this level, each sentence is classified as "positive," "negative," or "unbiased."

Emotional classification based on characteristics and aspects: At this level, some attributes of the phrases or archives can classify claims as "positive," "negative," or "non-partisan." This level is sometimes referred to as "perspective-level assessment grouping."[4]

It involves a variety of activities, such as sentiment analysis and categorization, looking for factual information in texts, subjective detection, opinion summaries, and identifying opinion spam, to name a few. It is vital to formalise a viewpoint in order to complete these jobs correctly. A variety of formalisms and mathematical representations for expressing opinions have been created for this purpose. Sentiment analysis offers many prospects for expansion, in large part because of the explosion of data that is now accessible online, such as in social networks, forums, and other places.

As a proof of concept for potential applications, this study seeks to demonstrate the application of sentiment analysis techniques in the context of social media assessments. A logistic regression, a BernoulliNB, and a linear SVC are some of the methods employed.

II. DEFINITION AND MOTIVATION

The emotional undertone of a textbook's main body can be identified using the natural language processing (NLP) technique of sentiment analysis. This is a common method used by organisations to ascertain and categorise opinions regarding a good, service, or concept. It involves mining textbooks for sentiment and private information using Data Mining, Machine Learning and Artificial Intelligence. Sentiment analysis is a method that automatically extracts sentiments, views, and feelings from text, speech, tweets, and database sources using Natural Language Processing (NLP)[21].

Sentiment analysis involves sorting textbook opinions into categories like positive, negative and neutral. It is also known as subjectivity analysis, opinion mining, and the creation of appraisals. Sentiment analysis is a method for assessing the opinions of particular people or groups, such as a subsection of a brand's supporters or a client communicating with a customer service representative. As a method of scoring, sentiment analysis listens to interactions and assesses language and voice intonation in order to determine feelings and passions, particularly those linked to a business, a product or service, or a particular subject[22].

Continual transmission of massive amounts of client-generated web-based social networking activities, including studies, online journals, comments, opinions, images, and recordings, take place. These correspondences supply beneficial chances to comprehend the perspectives of customers on fascinating subjects and give information that is useful for deciphering and foreseeing commercial and social news, like product offers, stock returns, and the outcomes of political choices[16]. Evaluating the ideas that customers articulated in their content exchanges is essential for these assessments.The goal of the popular field of research known as "notion examination" is to enhance computers' ability to recognise the emotions conveyed in data. Inferred data is utilised more skillfully as it becomes more prevalent.

Machine learning and lexicon-based methods have been named as the two main sentiment analysis techniques. As opposed to lexicon-based approaches, which work by counting the positive and negative words associated with the data, machine learning approaches use algorithms to extract and detect sentiment from data. A fresh, precise model for sentiment analysis has been created by academics. The majority of the model is written in English, which presents a problem when developing it. A new study, however, demonstrates that sentiment analysis model designs are accessible in a number of languages, spanning Arabic, Korean, Chinese, and Thai. According to reports, the disciplines of marketing and business, politics, and public policy have all used sentiment analysis. Applications can be used for things like e-commerce, polling software, and global events. The vast majority of the study's data came from social media. Social media is appropriate for sentiment analysis investigations because it provides access to all relevant information about a certain good, service, place, or event. Data from online users is widely available on social media[25].

III. IMPORTANCE AND BACKGROUND

Opinions are the main factors influencing our behaviour, hence they underlie every single human activity. Every time we have to make a decision, we need to understand what other people are thinking. In fact, companies and organisations must constantly monitor public perception of their goods and services. Through a variety of online platforms,Consumers participate in social contact, including on internet-based social networking sites like Facebook

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

and Twitter. Through these online social networks, customer interaction gradually grows.A connection like this offers a fantastic window of opportunity for studying marketing. The internet is used by people from every race, sexual preference, country, and socioeconomic class to share experiences and opinions about practically every aspect of their lives.

In addition to sending messages, blogging, or posting comments on company websites, many people use unofficial business sites to share their thoughts, express their emotions, and get insight into their everyday lives. Almost anything can be the subject of a letter, including books, products, or social events. These journals are online communities where users educate and shape others who circulate through online communities. These logs give marketers useful information about the propensities of consumer behaviour as well as a consistent chance to discover the sentiments and recognitions of consumers as they happen naturally and unprompted. However, despite this, the recent explosion of user-generated content on social media is posing new difficulties for locating, analysing, and translating printed material because the data is dispersed, confused, and fractured[23].

These issues can be overcome by information mining approaches like opinion investigation, which meticulously isolate and examine internet-based data without causing delays. Clients can continuously learn about the feelings and mental states of their customers by applying conclusion analysis, regardless of the obstacles of information, the amount and framework. Two considerations explain the enthusiasm for sentiment analysis as a strategy for promoting research tools in this study.

Sentiment analysis advises businesses to find out what customers like and dislike about their products and brand perception. In order to analyse industry and organisational data and aid in business decision-making, it is also crucial[10].

IV. CLASSIFICATION TECHNIQUES

Machine learning experts have developed classification methods that use a variety of strategies to classify unlabeled data. There is a chance that classifiers require training data. Machine learning models include BernoulliNB, LinearSVC, and Logistic Regression, to name a few. Because they need training data, these are known as supervised machine learning techniques. It is critical to keep in mind that accurate classifier training will support next projections[7].

A. BernoulliNB

Bernoulli's theorem is fundamentally based on Naive Bayes.A supervised machine learning approach called Naive Bayes is used to forecast the probability of various classes based on a variety of attributes. It reflects the propensity for an event to occur. Another name for naive Bayes is conditional probability.

Bernoulli Naive Bayes (BNB) is a variation of the Naive Bayes approach,which is frequently employed in machine learning applications that involve classification.It is very useful when working with text data, such as when classifying emails as spam or not[15].

The Bernoulli distribution, a distribution of probabilities for a single binary random variable, provides the foundation for BNB (i.e., a variable that can have one of two values, such as 0 or 1). When determining whether a certain characteristic is present in a document or not, BNB uses a binary variable.

The method determines the likelihood of each class based on whether each feature is present or absent in the document. It assumes that the features are conditionally independent given the class label, which is why it is regarded as "naive". BNB usually performs effectively in real-world scenarios, especially when there are many characteristics but limited data, despite this oversimplifying assumption.

The algorithm is trained on a labelled data set where each document is marked with its appropriate class so as to employ BNB for classification(for example, spam or not spam). For each class during training, the algorithm calculates the likelihood that each attribute will be present or absent[18].

$$P(A \mid B) = \frac{P(B \mid A) \cdot P(A)}{P(B)}$$

where:- A:event-1, B:event-2

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

P(A|B):Probability of A being true given B is true (posterior probability)
P(B|A):Probability of B being true given A is true (the likelihood)
P(A): Probability of A being true (prior)
P(B): Probability of B being true (marginalization)
The denominator, or marginal probability, is ignored in the Naive Bayes classifier since the maximum posterior probability is the only factor that interests us. The normalisation term is not necessary for Argmax to exist.
Two fundamental presumptions form the foundation of the Naive Bayes classifier:-
(i) Conditional Independence - Each feature operates separately from the others. This suggests that the functionality of one feature has no impact on that of another.
(ii) Feature Importance - Every characteristic is equitably significant. To produce sound forecasts and obtain the most precise outcomes, it is necessary to be familiar with all the features.
Bernoulli Naive Bayes is a member of the Naive Bayes family. It is built on the Bernoulli distribution and only accepts binary input, or values of 0 or 1. If the dataset's characteristics are binary, we can assume that Bernoulli Naive Bayes shall be the appropriate algorithm to use[21].
Example:
(i) Based on the provided data, the Bernoulli Naive Bayes classifier can be used to determine whether or not a person is suffering from an illness. The Bernoulli Naive Bayes algorithm would perform well in this situation because it would be a binary classification problem.
(ii) To identify whether an SMS is "spam" or "not spam," the Bernoulli Naive Bayes classifier can also be employed for text classification.
Mathematical formula of BernoulliNB is demonstrated as:
Let there be a random variable 'X' and let ‘p’ be the probability of success and ‘q’ be the likelihood of failure [1].
Success: p
Failure: q
q = 1 - (probability of Success)
q = 1 - p

$$p(x) = P[X = x] = \begin{cases} q = 1 - p & x = 0 \\ p & x = 1 \end{cases}$$

$$X = \begin{cases} 1 & \text{Bernoulli trial} = \mathbf{S} \\ 0 & \text{Bernoulli trial} = \mathbf{F} \end{cases}$$

As we saw, x can take two values only (binary values), i.e., 0 or 1.

	precision	recall	f1-score	support
0	0.81	0.79	0.80	39989
1	0.80	0.81	0.80	40011
accuracy			0.80	80000
macro avg	0.80	0.80	0.80	80000
weighted avg	0.80	0.80	0.80	80000

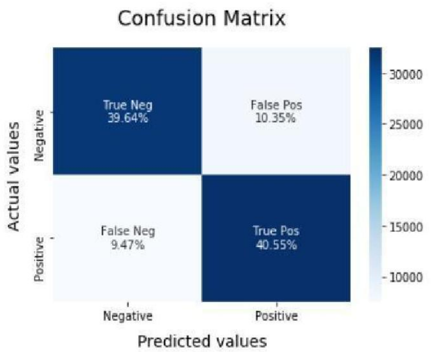


Fig1: BernoulliNB Matrix

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

B. LinearSVC

The goal of the Linear Support Vector Machine (Linear SVC) approach is to find a hyperplane that maximizes the separation between categorized samples.

The goal of a Linear SVC (Support Vector Classifier) is to divide or classify the data provided by returning a "best fit" hyperplane. After obtaining the hyperplane, some features are feeded to the classifier to obtain "predicted" class. This makes the specific algorithm (which may be applied in a number of situations) pretty perfect for our needs[2].

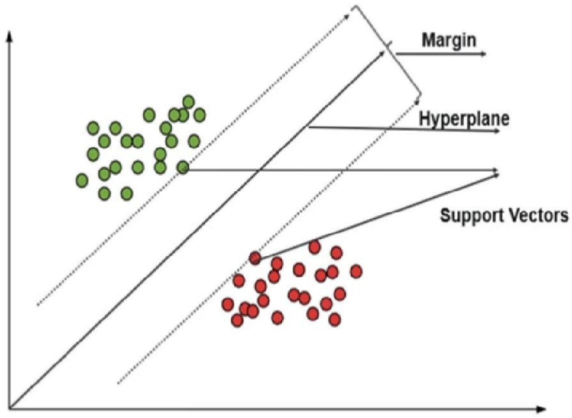


Fig2: Figure of Support Vector Machine

1. SVM – falls within Supervised ML.
2. It can perform both Regression and Classification.
3. Aim–To conveniently place new data points in the appropriate order, create the optimal decision boundary for classifying n-dimensional space.

	precision	recall	f1-score	support
0	0.82	0.81	0.82	39989
1	0.81	0.83	0.82	40011
accuracy			0.82	80000
macro avg	0.82	0.82	0.82	80000
weighted avg	0.82	0.82	0.82	80000

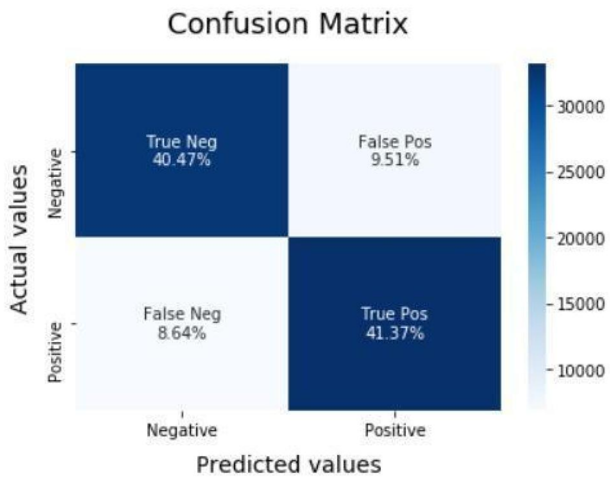


Fig 3: Linear SVC Matrix

The linear SVC (Support Vector Classifier) belongs to the family of classification techniques known as Support Vector Machines (SVMs). A subclass of supervised learning algorithms called SVMs can be used for tasks including outlier detection, regression analysis, and classification.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

The best approach for binary classification problems is linear SVC, which divides the data points into two classes. The method searches for the optimal hyperplane to discriminate between the two groups. This method of selecting the hyperplane maximizes the margin, or the separation between it and the nearest data points of each class.

The best technique is linear SVC when categorising the data points into two groups to address binary classification issues. The algorithm looks for the optimum hyperplane which can separate these two classes. The margin, or the separation between the hyperplane and the nearest data points of each class, is maximised using this method of selecting the hyperplane.

The benefit of linear SVC is that it performs admirably when dealing with high-dimensional data, where the difference between the number of features and samples is noticeably larger. Additionally, it is quick and suitable for large datasets. Overall, anomaly detection, text classification, and picture classification are only a few of the jobs that can be handled by Linear SVC's versatility and strength[14].

C. Logistic Regression

Logistic regression is a concept of supervised learning. Using a predetermined set of independent variables, it is utilized to predict the categorical dependent variable.

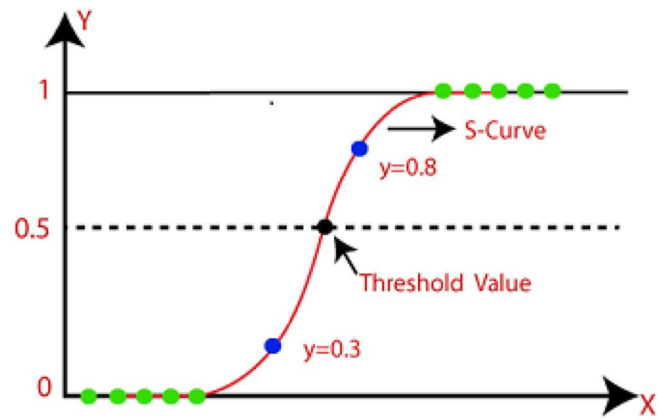
A categorical dependent variable's output can be predicted with logistic regression. As an outcome, the result must be a discrete or categorical value. Instead of the precise numbers between 0 and 1, it delivers the probabilistic values that fall between 0 and 1. Either false or true, 0 or 1, or No or Yes, are probable results.

With the possible exception of how they are applied, linear regression and logistic regression are pretty comparable. While regression problems are handled by linear regression, classification problems are handled by logistic regression.

In logistic regression, we design a "S" shaped logistic function which predicts two highest possible values (0 or 1) rather than a regression line. Numerous possibilities are shown by the logistic function's curve, including whether or not the cells are harmful and whether or not a mouse is obese based on its weight..

Logistic regression is an important machine learning method for classifying new data and may be applied to both discrete and continuous datasets[27].

When sorting observations using numerous sources of data, it is possible to quickly identify the variables which will turn out to be beneficial by using logistic regression. The logistic function is displayed in the illustration below:



-Fig4: Figure of Logistic Regression Curve

Type of Logistic Regression:

On the basis of the categories, there are three categories for logistic regression :

Binomial:Under binomial logistic regression, the dependent variables can only take one of two possible forms, like 0 or 1, Pass or Fail, etc.

Multinomial:Multinomial logistic regression allows the dependent variable to be any one of three or more unordered sorts, like "cats" or "dogs".

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.301
Volume 3, Issue 10, May 2023

Ordinal:Ordinal logistic regression allows for up to three different ordered types of dependent variables, like "High", "Medium" and "Low".

Logistic Regression Equation: The linear regression equation yields the logistic regression equation. Below are the mathematical steps to obtain Logistic Regression equations:

We are aware that the equation for a straight line can be expressed as:

$$y = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n.$$

Let's divide the previous equation by (1-y) because y in Logistic Regression can only be between 0 and 1:

$$\frac{y}{1-y}; \text{0 for y= 0, and infinity for y=1}$$

But we need range between -[infinity] to +[infinity], then on taking logarithm of the equation , it will become:

$$log\left[\frac{y}{1-y}\right] = b_0 + b_1x_1 + b_2x_2 + b_3x_3 + \cdots + b_nx_n$$

The resultant equation is the final equation for Logistic Regression[3].

	precision	recall	f1-score	support
0	0.83	0.82	0.83	39989
1	0.82	0.84	0.83	40011
accuracy			0.83	80000
macro avg	0.83	0.83	0.83	80000
weighted avg	0.83	0.83	0.83	80000

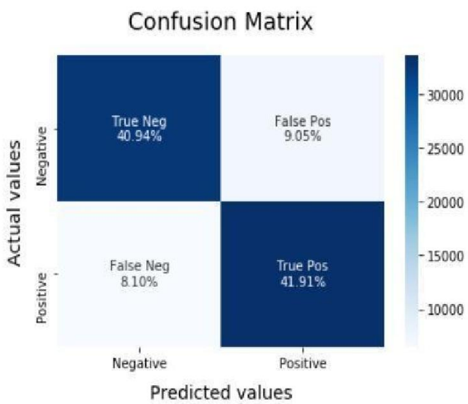


Fig5: Logistic regression Matrix

	precision	recall	f1-score	support
0	0.81	0.79	0.80	39989
1	0.80	0.81	0.80	40011
accuracy			0.80	80000
macro avg	0.80	0.80	0.80	80000
weighted avg	0.80	0.80	0.80	80000

	precision	recall	f1-score	support
0	0.82	0.81	0.82	39989
1	0.81	0.83	0.82	40011
accuracy			0.82	80000
macro avg	0.82	0.82	0.82	80000
weighted avg	0.82	0.82	0.82	80000

	precision	recall	f1-score	support
0	0.83	0.82	0.83	39989
1	0.82	0.84	0.83	40011
accuracy			0.83	80000
macro avg	0.83	0.83	0.83	80000
weighted avg	0.83	0.83	0.83	80000

	text	sentiment
0	I love twitter	Positive
1	May the Force be with you.	Positive
2	Mr. Stark, I don't feel so good	Negative

Fig6: Comparison of the classification techniques

According to the comparison chart obtained as output , it can be concluded that Logistic Regression is the best classification technique for Sentiment Analysis as it provides the maximum accuracy[5].

V. CHALLENGES AND LIMITATIONS OF SENTIMENT ANALYSIS

Businesses deal with an array of issues when it concerns sentiment analysis challenges in order to achieve sentiment analysis accuracy. Natural language processing can make it challenging because machines should be taught to interpret sentiment and feelings in the same manner that the human brain does. This is crucial, along with comprehending the subtleties of various languages. Software for sentiment analysis is more capable of addressing these problems as data science develops. The following is a list of the key problems with sentiment analysis[8].

1.Tone:

Problem: It can be challenging to determine tone when chatting, and it can be far harder when writing. When attempting to examine a sizable amount of data that could include objective as well as subjective responses, things get much more challenging. Finding subjective feelings and appropriately evaluating them for the desired tone can be challenging for businesses.

Solution:In order to choose the right tone, any efficient sentiment analysis algorithm must be able to differentiate between subjective and objective assertions. In spite of being subjective, the statement "The product is glam but not at that cost" communicates the notion that the object is less beautiful because of the cost. Businesses can broadly interpret these subtleties in tone using a clever sentiment API[9].

2.Polarity:

Problem: Words like "hate" and "love" rank highly for positive and negative polarity, respectively. These are simple to comprehend. The term "not so bad," which can also mean "average," is mid-polar (-75), and there are several conjugations that fall in the centre. This kind of language is occasionally omitted, which decreases the sentiment score.

Solution: To give a full study of a statement, sentiment analysis technologies make it simple to distinguish such mid-polar phrases and words. In this situation, subject-based sentiment analysis can provide an in-depth analysis, whereas aspect-based analysis of sentiment can provide a detailed look at several facets of a comment[11].

3.Sarcasm:

Problem: Sarcasm and irony are frequently used in casual discussions and online memes. It may be challenging for sentiment analysis algorithms to comprehend the genuine perspective of what the response actually means when negative sentiment is presented with unintentional praises. As a consequence, the system can see a rise in undesirable "positive" input.

Solution: A high calibre sentiment analysis API will be able to identify the linguistic perspective, other factors, and genuine sentiment when someone uploads something. This requires a precise and sizable linguistic dataset used to train the model for sentiment analysis[26].

4.Emojis:

Problem: The overuse of emojis in text-driven social media posts, like those on Twitter, is a concern. NLP tasks are drilled with language-specific activities. Emojis have their own language, despite the fact that they can convert words from simple graphics. Emojis are typically treated as special characters in the majority of emotion analysis software and removed from data sets during sentiment mining. However, it will make it impossible for businesses to fully understand the data.

Solution: Use of emotion analyst technology that can understand the syntax in emojis and not mix them with special characters such as commas, spaces, or full stops is essential for a company to handle sentiment analysis difficulties like these. Models similar to Repustate's model are trained expressly for this, which is a pretty sophisticated application in and of itself. Data scientists determine if users use emojis more often in positive or negative settings before training the algorithms to detect links between various emoticons and phrases[20].

5.Multilingual sentiment analysis:

Problem: When multiple languages are used, all of the problems with single-language sentiment analysis are exacerbated. To grasp negations, every tongue requires a unique part-of-speech tagger and grammatical constructions. Since each language is unique, it cannot be understood by translating it into a simple language like English. The meaning would have been lost, for example, if the phrase "like a fish takes to water" were translated into, say, German[24].

Solution: Only through going through hardship will these issues with sentiment analysis for multilingual data be solved. This suggests that, similar to Repustate. A platform and a named entity recognition model that has been programmed specifically for each language are required for the sentiment analysis model. Data scientists must personally train the model in each language, so there is no short cut here. It is time-consuming and demands accuracy and diligence. The results are worthwhile since you will acquire the highest sentiment analysis accuracy ratings available..

6. Negations:

Problem: Negative claims made by terms such as "never", "not", "were not," etc. can deceive the ML model. The sentence "I can't not go to my class reunion" is one that conveys an individual's desire to attend the reunion, and a machine algorithm must comprehend this.

Solution: It takes training for a sentiment analysis tool to comprehend that two negatives cancel each other out and make a text positive. This is only possible if the algorithm had been trained on a sufficiently large corpus and contains the greatest number of negation terms, which permits the greatest number of combinations and permutations[12].

VI. CONCLUSION

The use of sentiment analysis techniques in social media has been examined in this article. For the classification process, it has used techniques like BernoulliNB, Linear SVM, and Linear Progression. Logistic Regression provides the best performance.

In the future, it will be necessary to evaluate the efficiency of the algorithm that was used to determine the scores for each aspect. Because it didn't have a defined metric similar to the other procedures assessed, its results were examined by comparing those attained in every assessment and the overall average[19].

The tags for each component are required in order to more accurately assess these outcomes. Although this study's subjectivity is one of its flaws, these should be manually gathered when each review is analysed. As a result, techniques for producing tags automatically for each element could be researched[13].

In future study regarding potential model improvements, individual reviewer bias—the likelihood that a reviewer will give papers a score that is either lower or greater than the average—could be taken into consideration. The existing model would have to be altered to account for this bias. Additional issues that might be addressed include the correct handling of reviews that are written in various languages and the look for a suitable parameterization in this case. Social media sentiment analysis has grown into a crucial tool for determining how the general public feels about a variety of issues. The methodologies and techniques used for sentiment analysis continue to advance, making it a more effective tool for organisations, governments, and individuals even though it has drawbacks and limitations.

Future research on the proposal's applicability may examine the uniformity of reviews and the recognition or rejection of an article by every reviewer over time[17].

REFERENCES

[1]. Sentiment analysis using product review data- Xing Fang* and Justin Zhan
[2]. A Study on Sentiment Analysis Techniques of Twitter Data
[3]. Abdullah Alsaeedi1, Mohammad Zubair Khan2 Department of Computer Science,College of Computer Science and Engineering ,Taibah University,Madinah, KSA
[4]. Sentiment Analysis in Social Media and Its Application: Systematic Literature Review - Zulfadzli Drus, Haliyana Khalid* Azman Hashim International Business School, Kuala Lumpur, 54100, Malaysia
[5]. Kim S-M, Hovy E (2004) Determining the sentiment of opinions. In: Proceedings of the 20th international conference on Computational Linguistics, page 1367. Association for Computational Linguistics, Stroudsburg, PA, USA
[6]. Liu B (2010) Sentiment analysis and subjectivity. In: Handbook of Natural Language Processing, Second Edition. Taylor and Francis Group, Boca

[7]. Liu B, Hu M, Cheng J (2005) Opinion observer: Analyzing and comparing opinions on the web. In: Proceedings of the 14th International Conference on World Wide Web, WWW ’05. ACM, New York, NY, USA. pp 342–351

[8]. Pak A, Paroubek P (2010) Twitter as a corpus for sentiment analysis and opinion mining. In: Proceedings of the Seventh conference on International Language Resources and Evaluation. European Languages Resources Association, Valletta, Malta

[9]. Pang B, Lee L (2004) A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42Nd Annual Meeting on Association for Computational Linguistics, ACL ’04. Association for Computational Linguistics, Stroudsburg, PA, USA

[10]. Pang B, Lee L (2008) Opinion mining and sentiment analysis. Found Trends Inf Retr 2(1-2):1–135

[11]. Turney PD (2002) Thumbs up or thumbs down?: Semantic orientation applied to unsupervised classification of reviews. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, ACL ’02. Association for Computational Linguistics, Stroudsburg, PA, USA. pp 417–424

[12]. AGiachanou and F. Crestani, "Like It or Not: A Survey of Twitter Sentiment Analysis Methods," ACM Comput. Surv., vol. 49, no. 2, pp. 1-41, 2016.

[13]. Abirami and V. Gayathri, "A survey on sentiment analysis methods and approach," in Advanced Computing (ICoAC), 2016 Eighth International Conference on, 2017: IEEE, pp. 72-76.

[14]. J. Read, "Using emoticons to reduce dependency in machine learning techniques for sentiment classification," in Proceedings of the ACL student research workshop, 2005: Association for Computational Linguistics, pp. 43-48.

[15]. AS. Nugroho, A. B. Witarto, and D. Handoko, "Support vector machine," Teori dan Aplikasinya dalam Bioinformatika, Ilmu Komputer. com, Indonesia, 2003.

[16]. Bollegala, D., Weir, D., & Carroll, J.. Cross-Domain SentimentClassification using a Sentiment Sensitive Thesaurus. Knowledge andData Engineering, IEEE Transactions on, 25(8), 1719-1731,2013

[17]. Taboada, M., Brooke, J., Tofiloski, M., Voll, KStede, M..“Lexicon based methods for sentiment analysis”. Computational linguistics, 2011:37(2), 267-307.

[18]. Neethu M,S and Rajashree R,” Sentiment Analysis in Twitter using Machine Learning Techniques” 4th ICCCNT 2013,at Tiruchengode, India. IEEE – 31661

[19]. Bifet and E. Frank, "Sentiment Knowledge Discovery inTwitter Streaming Data", In Proceedings of the 13th InternationalConference on Discovery Science, Berlin, Germany: Springer,2010, pp. 1-15.

[20]. Socher, Richard, et al. "Recursive deep models for semantic compositionality over a sentiment Treebank." Proceedings of theConference on Empirical Methods in Natural Language Processing(EMNLP). 2013.

[21]. Agarwal, B. Xie, I. Vovsha, O. Rambow, R. Passonneau, “Sentiment Analysis of Twitter Data", In Proceedings of the ACL 2011 Workshop on Languages in Social Media,2011 , pp. 30-38

[22]. ZhunchenLuo, Miles Osborne, TingWang, "An effective approach to tweets opinion retrieval", Springer Journal onWorldWideWeb,Dec 2013, DOI: 10.1007/s11280-0130268-7.

[23]. Po-Wei Liang, Bi-Ru Dai, “Opinion Mining on Social MediaData", IEEE 14th International Conference on Mobile Data Management,Milan, Italy, June 3 - 6, 2013, pp 91-96, ISBN: 978-1-494673-6068-5, http://doi.ieeecomputersociety.org/10.1109/MDM.2013.

[24]. K. Makar, S. Goel, P. Kaur, M. Singh, P. Jain and P. K. Aggarwal, "Reliability of Mobile Applications: A Review and Some Perspectives," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021, pp. 1-4, doi: 10.1109/ICRITO51393.2021.9596350.

[25]. P. Chaudhary, S. Goel, P. Jain, M. Singh, P. K. Aggarwal and Anupam, "The Astounding Relationship: Middleware, Frameworks, and API," 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021, pp. 1-4, doi: 10.1109/ICRITO51393.2021.9596088.

[26]. S. Mishra, A. Shukla, S. Arora, H. Kathuria and M. Singh, "Controlling Weather Dependent Tasks Using Random Forest Algorithm," 2020 Third International Conference on Advances in Electronics, Computers and Communications (ICAECC), 2020, pp. 1-8, doi: 10.1109/ICAECC50550.2020.9339508.

[27]. Sharma, A., Singh, M., Gupta, M., Sukhija, N., & Aggarwal, P. K. (2022). IoT and blockchain technology in 5G smart healthcare. Blockchain Applications for Healthcare Informatics, 137–161. https://doi.org/10.1016/b978-0-323-90615-9.00004-9

[28]. H. Garg, M. Singh, V. Sharma and M. Agarwal, "Decentralized Application (DAPP) to enable E-voting system using Blockchain Technology," 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2022, pp. 1-6, doi: 10.1109/ICCSEA54677.2022.9936413.