International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7383
***Article Received**: 30 March 2023 **Revised**: 15 May 2023 **Accepted**: 19 June 2023*

Aspect Based Sentiment Analysis using Various Supervised Classification Techniques: An Overview

Miss. Sushadevi Shamrao Adagale[1], Dr. Praveen Gupta[2]
[1]Department of Computer Engineering Assistant Professor, Trinity Academy of Engineering, Pune
Research Scholar, CSMU,Navi Mumbai
Mumbai,India susha.adagale10@gmail.com
[2]Department of Computer Engineering CSMU,Navi Mumbai
Mumbai,India Praveenymt2014@gmail.com

Abstract— The Sentiment Analysis (SA) work is concerned with identifying aspect terms and categories and categorising emotions (positive, negatively, conflict, and neutral) in ratings and reviews. When it comes to subjectivity, it's typical to divide sentences into objective phrases that include accurate information and subjective statements that include express ideas, beliefs, and perspectives on a given topic. Various existing researchers have already done a lot of work in sentiment analysis with various methods, including aspect extraction. This paper proposed a systematic literature analysis of numerous sentiment analysis using supervised and unsupervised classification techniques. We investigate a few features extraction Natural language Processing (NLP) techniques used to identify aspects of machine learning for the detection of sentiment. An extensive experiment analysis, we discuss the findings of the study, challenges of the current and define the problem statement for the future direction

Keywords : Aspect base sentiment analysis, natural language processing, feature extraction and feature selection, classification, machine learning, deep learning.

I. INTRODUCTION

Sentiment Analysis (SA) aims to uncover and extract personal information for this research using text analysis, natural language processing, and computational linguistics. It's used for several purposes in reviews and social media, including marketing and customer service. As seen in Figure 1,a vast quantity of data is created every day through social networks, healthcare, blogs, and other media.

Figure 1: Various application areas of sentiment analysis

Approaches of Sentiment Analysis

In SA, detecting sentiment from text data is a significant issue.Aspect identification methodologies are categorised as frequency-based methods (FB), natural language processing (NLP), deep learning (DL), syntax-based methods (SB), unsupervised machine learning methods, supervised machine learning (SML), and hybrid approaches, as illustrated in Figure2.

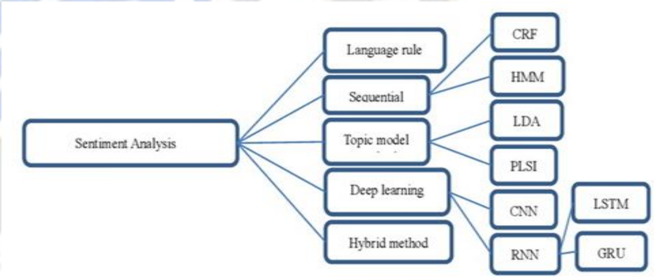

Figure 2 : Different methods to perform Sentiment Analysis

This enormous volume of data provides vital opinion-related data that might be useful to businesses and scientific groups. Because manual tracing and abstraction of this vital information are impossible, SA is required. It divides feelings into three groups: pleasant, negative, and neutral. Consequently, it expresses the writer's general attitude toward the issues at hand.

Words that are often used are considered characteristics in text categorization. Over the whole lexicon, these little clusters of words are favoured. The attributes may be selected depending on frequency after pre-processing. Some methods extract characteristics such as nouns, adjectives, and adverbs. As a result, characteristics with a higher frequency of recurrence are picked above those with a lower frequency of occurrence.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7383
***Article Received**: 30 March 2023 **Revised**: 15 May 2023 **Accepted**: 19 June 2023*

II. LITERATURE SURVEY

The authors of [1] officially articulated their goal and explained how to add social and topical background through into basic estimation method analytically. They looked at the evidenced relationships between the themes and used to quantify them. Scientific method on the Twitter given dataset they developed confirmed their predictions regarding social environment and topical environment. Finally, they performed tests to test the increased organizational, and the findings showed because both cultural context and subject context may aid in improving user subject opinion forecasting accuracy.

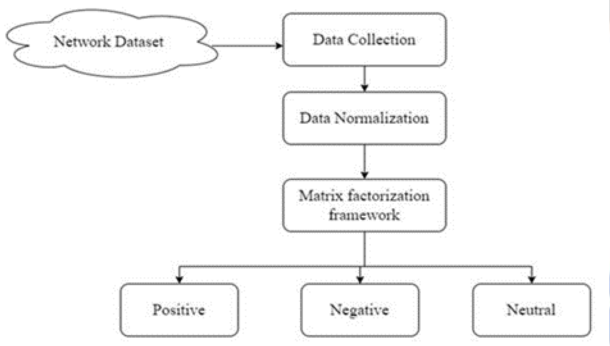

Figure 3: Dual Sentiment Analysis: Considering Two Sides of One Review [2]

Disadvantages:

1: No provision for implicit and explicit aspects for sentiment classification

2: Neutral count is very high

To identify difficulties and faults in students' learning programs, the researchers of [3] focused on technology students' Newsfeed. They started with a subjective examination of roughly 25,000 retweets related to engineer graduates' college experiences. They discovered that students face issues such as a lack of social involvement, a hefty academic load, and insufficient sleep. Findings from this study, we developed a multi-label categorisation to classify tweets about students' issues. They then utilised the technique to train a classifier of student difficulties based on 35,000 tweets sent from the College geo-location. This paper demonstrates how incidental social media marketing may directness into learners' experiences by presenting a strategy and findings.

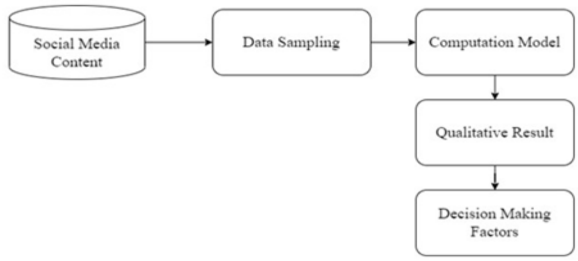

Figure 4: Mining Social Media Data for Understanding Students' Learning Experiences [3]

Disadvantages:

1: Only NLP features has considered that generate high errorrate

Because terms that exist in the train (origin) domain may not present in the testing (targeted) domain, using an emotion classifier developed leads in underperformance for emotion classification. [4] presented an inter classification model employing an autonomously derived sentiment responsive vocabulary to tackle the feature mismatch issue in merge sentiment analysis. They compared the results to the Sent Word vectors, which is a lexical database for word neutrality. They demonstrate that the attitude thesaurus they developed correctly catches terms that represent comparable emotions.

Disadvantages:

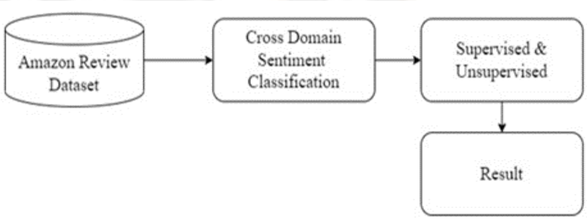

Figure 5: Cross-Domain Sentiment Classification Using a Sentiment SensitiveThesaurus [4]

Because earlier research had mostly concentrated on modelling and measuring public opinion, the researchers went a step beyond that and attempted to analyse sentiment fluctuations. They collaborated on a material undergoes on Twitter. They discovered that during the sentiment fluctuation periods, emergent themes are significantly connected to the true causes for the changes. These prominent issues may aid in deciphering sentiment shifts [5].

Disadvantages:

1: Utilizes very basic machine learning techniques, only applicable for single-label binary classification, only

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7383
***Article Received**: 30 March 2023 **Revised**: 15 May 2023 **Accepted**: 19 June 2023*

considersunigrams with limited POS tag

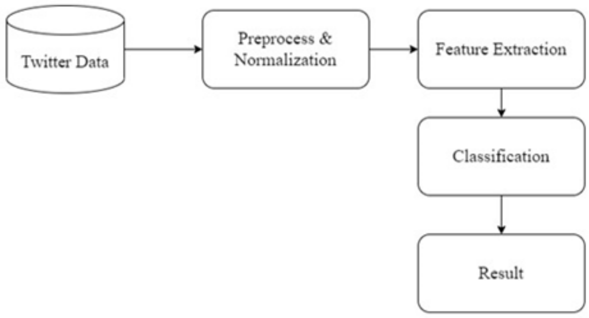

Figure 6: Interpreting the Public Sentiment Variations on Twitter [5]

Though most known techniques to sentiment categorization favour deep networks, the authors introduced a hybrid viewpoint model and a re-parametrized version of JST termed Reverse-JST. Unlike unsupervised emotion classification methods, which almost always fail to generate satisfactory results when applied to domain names, JST's sparse representation element makes it very adaptable to domain names [6].

Disadvantages:

1: Focus only on implicit aspect identification

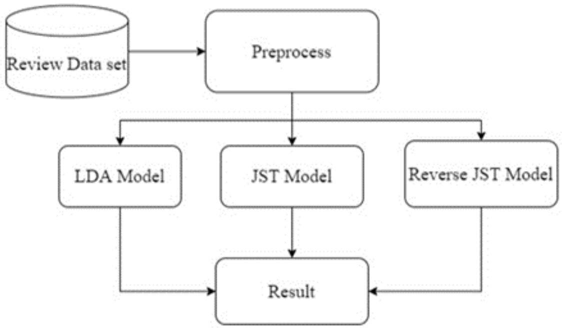

Figure 7: Weakly Supervised Joint Sentiment-Topic Detection from Text [6]

The [7] indicates that two sources of evidence, which include sentiment score of twitter accounts the hash tag, new hashtag founder relation-ship, as well as sentiment polarizability of twitter accounts the hash tag, are useful to quickly create the overarching sentiment score for a specified hash tag in a specific time period, which differs significantly from traditional statement and manuscript sentiment analysis. as well as the literal definition of the term "hashtag" They offer a unique data structure and examine three estimated communal classification methods for inference in order to combine the first two sets of data into a classification model where tweets may be evaluated individually.

Disadvantages:

1: For a small annotated data set this method will be inaccurate

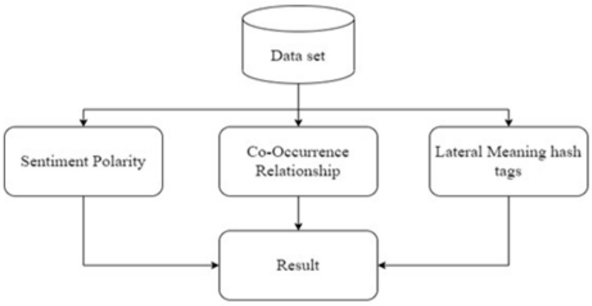

Figure 8: Topic Sentiment Analysis in Twitter: A Graph-based Hashtag Sentiment Classification Approach [7]

The graphical analysis findings obtained by the hybrid technique are superior to the dictionary and acquiring knowledge baselines, according the researchers of [8]. Their usual inflections method achieved excellent accuracy for both tweet sentiment categorization and emotion strength recognition, which is extremely close to the pure attempting to learn system and significantly higher than that of the pure etymological roots approach. This method combines the best of both worlds: a well-built lexicon's consistency and accessibility, as well as a strong supervised machine learning computation good precision.

Disadvantages:

1: Reference resolution and long distance negation detection need to be addressed in the future

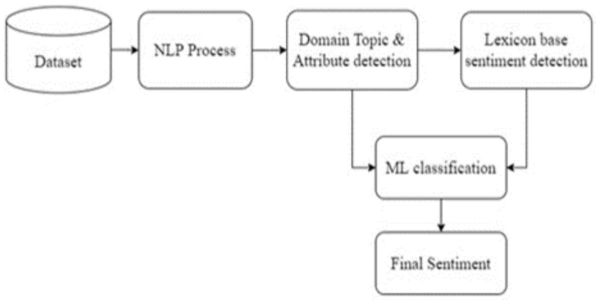

Figure 9: Combining Lexicon and Learning based Approaches for Concept-Level Sentiment Analysis [8]

The predictive potential of reviews has been investigated using movie industry as a research study, as well as the difficulty of forecasting sales success using emotion data gleaned from reviews [9]. To represent its complex personality of sentiments, they offer Emotion PLSA, in which an evaluation is seen as a document formed by a number of parameters sentiment components. Following

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7383
***Article Received**: 30 March 2023 **Revised**: 15 May 2023 **Accepted**: 19 June 2023*

that, they present ARSA, a Vector autoregression Emotion model for sales forecasting.

Disadvantages:

1: The sentiment score of a word is generally dependent ona particular domain and changes when a domain switch occurs.

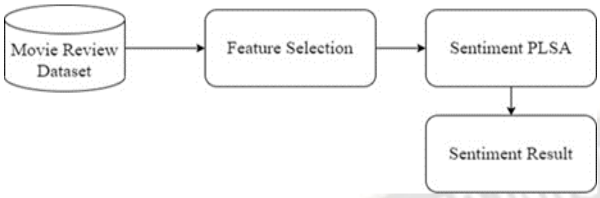

Figure 10: Mining Online Reviews for Predicting Sales Performance: A CaseStudy in the Movie Domain [9]

This approach presented a technique for mining Twitter data to find query results such whether the price of a group of 30 companies that are listed on the NASDAQ as well as the National Stock Exchange can be accurately forecasted using 15million recordings of Twitter chats [10]. this survey for the various kinds of work conducted in the Text Analytics sector. The reference number is R*. The purpose of picking the table's components is to examine the work carried out in the area of sentiment classification.

Disadvantages:

1: Mining process needs to improve to enhance the performance

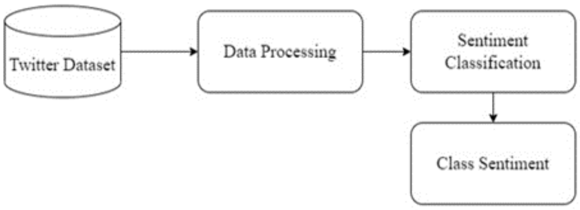

Figure 11: Public Sentiment Analysis in Twitter Data for Prediction of ACompany's Stock Price Movements [10]

It was utilised according to the terminology supplied by SentiStrength [11]. This lexicon includes an emotional vocabulary, an emotes list, and a collection of negation or strengthening words. The suggested method calculates the sentiment meaning of each and every word in the provided text. When phrases are denied, their polarity is reduced; conversely, when keywords are emphasized, their polarity is enhanced. The classifier then identifies the text either positive, negative, or neutral.

Disadvantages:

1: Rules that need to be upgraded , It focus only on aspectextraction tasks

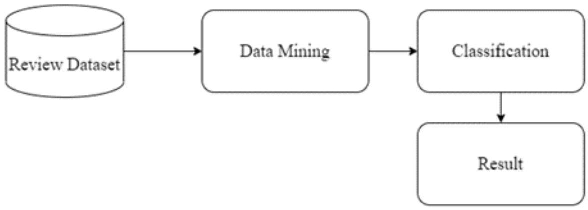

Figure 12: Sentiment Strength Detection in Short Informal Text [11]

In SemEval-2014, the researchers described the NILC technology and presented a trio classification method that includes three different classification algorithms: rule-based methodology, lexicon-based methodology, and computer attempting to learn methodology. There are five stages in the suggested method.

Disadvantages:

1: Lacks the capacity for implicit aspect extraction. It exhibits low accuracy in terms of sentiment classification, doesnot address multiple aspects and the related sentiments that arepresent in a single sentence

Figure 13: NILC USP: A Hybrid System for Sentiment Analysis in TwitterMessages [12]

The [13] devised a other hybrid technique. Based on a conceptual structure, emotional lexicon, and a language modelling technique, this research presented an approach for Facebook and SMS trend analysis. It was discovered that although the language model performed poorly on its own, it performed better when combined with a lexicon-based model. Even when employing n-grams, emotional evaluations, and component, the modelling approach was highly effective.

Disadvantages:

1: Only bi-gram features are not sufficient for detection ofsentiment

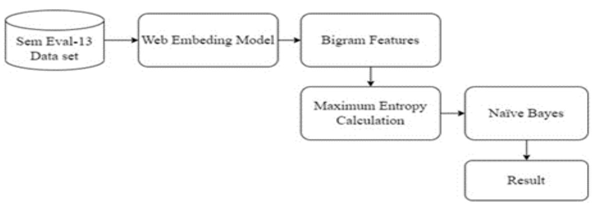

Figure 14: SAIL: A hybrid approach to sentiment analysis [13]

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7383
***Article Received**: 30 March 2023 **Revised**: 15 May 2023 **Accepted**: 19 June 2023*

The [14] demonstrated the pSenti, a construct emotion predictive algorithm that incorporates lexical and acquiring knowledge sentiment categorization algorithms. pSenti exhibited greater precision in sentiment amplitude detection and polarization classification than pure vocabulary techniques.Once compared to simple machine attempting to learn approaches, the tool, from the other hand, provided somewhat poorer accuracy. For the assessment of the suggested technique, substantial tests were run on two separate datasets, namely the CNet Technology Reviews Database and the IMDBSentiment Classification Dataset.

Disadvantages:

1: This model is not working on other domain and sometimes it is not able to recognize complex relationship among opinion and aspects

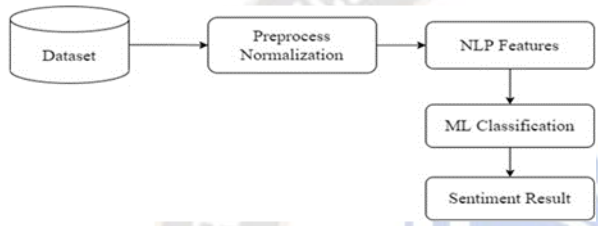

Figure 15: Combining lexicon and learning based approaches for concept-levelsentiment analysis [14]

[15] employed an enriched vocabulary technique for institutiontrend analysis. Secondly, they used the Chi-square analysis on the findings from the vocabulary technique to purchase additional subjective indicators, such as words and symbols. With the use of new opinions indications, more argumentative messages were discovered. A sentiment categorization technique is used to give sentiment polarity ratings to items in the newly detected tweets.

Disadvantages:

1: Cannot address sentences without sentiment word, interrogative, conditional sentences, and sarcastic sentences.

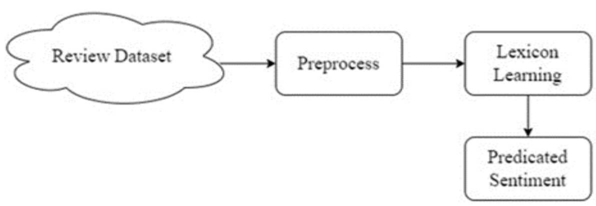

Figure 16: Combining lexicon-based and learning-based methods for Twittersentiment analysis [15]

The lexical technique produces essentially instructional data for the classifiers, and the whole procedure requires no user labelling besides the test set. Jie Li; Lirong Qiu et al. [16] highlight how mood architecture and emotion calculation principles may be used to solve difficulties. The emotion structure is derived from dependencies parsing combined with connection immigration and adjusted proximity, and it contributes significantly to interpreting the attitude of brief text. The sentiment of a brief text is calculated using the varied influences of connections between the modulator and the emotions word, as well as the participation of each phrase to the sentiment computation.

Disadvantages:

1: Multi class classification problem due to future edundancy

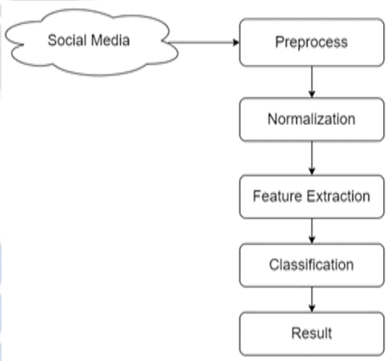

Figure 17: A sentiment analysis method of short texts in microblog [16]

Chae Won Park with Dae Ryong Seo [17] proposed a method for determining whether machine intelligence companion is objectively superior. A dictionary called Up With an interesting Dictionary and emotion Reasoner gathered user views on three artificially intelligent aides from Twitter and rated them as favourable, negative, or neutral (VADER). We also used independent - Sample t, Kruskal-Wallis tests, and Mann-Whitney tests to examine tweets for significance levelbetween groups.

Disadvantages:

1: Relational dependency features has used for classification; some relationship tags generate overhead sometimes.

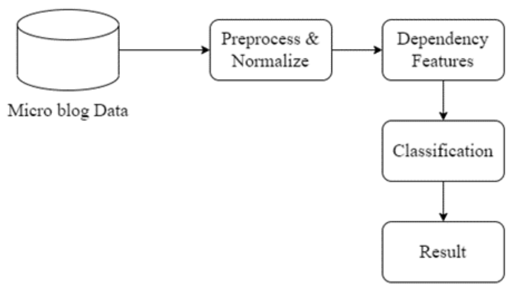

Figure 18: Sentiment analysis of Twitter corpus related to artificial intelligenceassistants [17]

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7383
***Article Received**: 30 March 2023 **Revised**: 15 May 2023 **Accepted**: 19 June 2023*

To capture emotion from social networking sites, Li-Chen Cheng with Song-Lin Tsai [18] proposed an unique emotion analysis method development of deep learning algorithms. They gather information from whom we create a dataset. We want to create a semantics dataset for additional study once we've processed these unique phrases. The data gathered will be beneficial in a variety of potential developments.

Disadvantages:

1: Very high neutral rate on large dataset, due to utilizationlexicon dictionary.

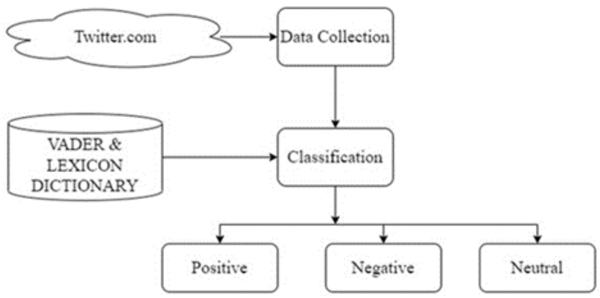

Figure 19: Deep learning for automated sentiment analysis of social media [18]

In this paper, Ou Wu et al. [19] use a novel labelling technique and a usually two bad memory system to develop a classification model to overcome this problem. Linguistic cues were used in traditional experiments and are beneficial for sentiment classification. To represent the polar turning of phrases, a flipping framework is described.

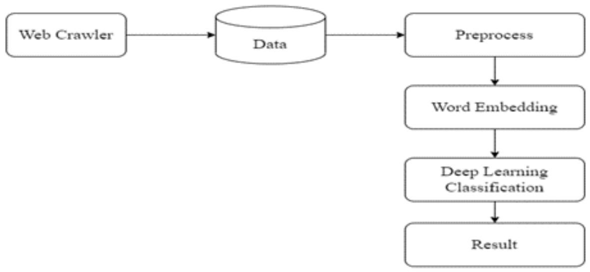

Figure 20: Two-Level LSTM for Sentiment Analysis with Lexicon Embeddingand Polar Flipping. [19]

Disadvantages:

1: word embedding has used that generates overfitting problem for classification.

Hongyu Han et al. [20] suggested an autonomous reviews text analytics approach that incorporates both of the categorization methods. A etymological roots technique is utilised in the training sentence to acquire assurance variables, which are thenused to pick a classifier from the a small-scale labelled dataset. Then there's a training set for a Naive Bayes emotion classifier.

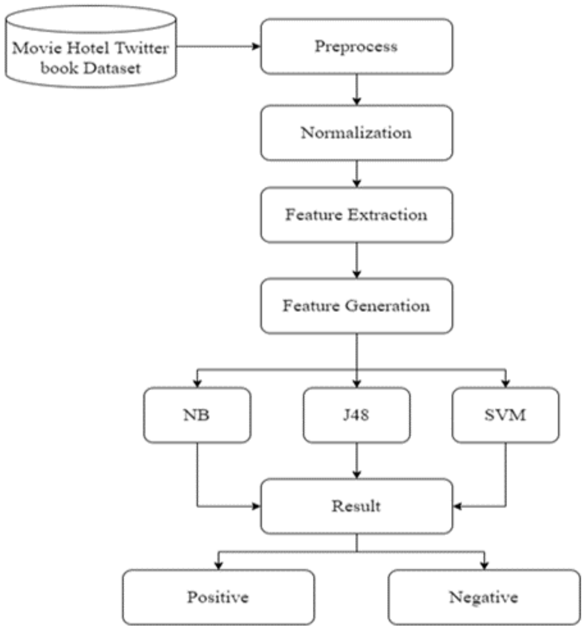

Figure 21: A Hybrid Sentiment Analysis Method [20]

III. FINDINGS AND SUGGESTIONS

We have outlined the most recent advances in picture sentiment using classic data-driven machine learning approaches. This will be beneficial to other researchers. In addition, we conducted a gap analysis based on a recent poll. These omissions will serve as crucial guiding principles for further study. The fundamental outcome we find in our entire study, which is mentioned below.

- The majority of sentiments analysis tools are assessed at the classification stage. For comparing items based on characteristics, this kind of analysis is ineffective.
- The properties of part-of-speech and bigrams are used in many works. It's crucial to extract precise word associationssince they're useful for assessing how people feel about the product's features.
- Unsupervised techniques are inefficient because extracting aspects and determining feelings takes time.
- As features, several systems employ phrases that occur in the text. When writing about the same topic, the user may use various phrases, resulting in an incorrect result. As a result, concept-based aspects must be considered.
- Previously, sentiment analysis tasks were assessed using specialised datasets of online reviews, tweets, and postings. They aren't examined utilising a variety of datasets.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 6
DOI: https://doi.org/10.17762/ijritcc.v11i6.7383
***Article Received**: 30 March 2023 **Revised**: 15 May 2023 **Accepted**: 19 June 2023*

IV. CONCLUSION

This research describes a literature study of aspect-based sentiment analysis using various machine learning and deep learning techniques. The implementation has a done bye by various researchers in different domains such as image sentiment classification, text review classification, polarity identification, aspect classification etc. The primary scenario in existing studies is negation handling issues and a high error rate. On the other hand, our research looked at deep learning approaches in classification systems, such as estimating raw data received straight from text data. Our findings imply that in-depth learning produces good outcomes with a presentation almost equivalent to specific tactics on the characterisation problem, utilising carefully constructed highlights. Photos have a variety of uses, including the automated classification of text sequences into mysteries, comedy, romance, and other genres, as well as the programmed tagging of an object with emotive portions in succession.

REFERENCES

[1] Fuji Ren,Ye Wu,” Predicting User-Topic Opinions in Twitter with Social and Topical Context”, IEEE Trans. on affective computing, vol. 4, no. 4,October-December 2013.
[2] Rui Xia, Feng Xu, Chengqing Zong, Qianmu Li, Yong Qi, and Tao Li,”Dual Sentiment Analysis:Considering Two Sides of One Review”, IEEE Trans.on Knowledge and Data Engineering, 2015
[3] Xin Chen,Mihaela Vorvoreanu, and Krishna Madhavan,” Mining Social Media Data for Understanding Students' Learning Experiences” IEEE trans. on learning technologies, vol. 7, no. 3, JulySeptember 2014.
[4] Danushka Bollegala, David Weir, and John Carroll,” Cross-Domain Sentiment Classification Using a Sentiment Sensitive Thesaurus”, IEEE trans. on knowledge and data engineering, vol. 25, no. 8, August 2013.
[5] Shulong Tan, Yang Li, Huan Sun, Ziyu Guan, Xifeng Yan, Jiajun Bu,
[6] ,Chun Chen, and Xiaofei He,” Interpreting the Public Sentiment Variations on Twitter”, IEEE trans. on knowledge and data engineering, vol. 26, no. 5, May 2014.
[7] Chenghua Lin, Yulan He, Richard Everson, Member, IEEE, and Stefan Ru¨ger,” Weakly Supervised Joint Sentiment-Topic Detection from Text”, IEEE trans. on knowledge and data engineering, vol. 24, no. 6, June 2012.
[8] Xiaolong Wang, Furu Wei, Xiaohua Liu, Ming Zhou, Ming Zhang,” Topic Sentiment Analysis in Twitter: A Graph-based Hashtag Sentiment Classification Approach”, ACM, CIKM'11, October 24–28, 2011, Glasgow, Scotland, UK, 2011.
[9] Andrius Mudinas, Dell Zhang, Mark Levene,” Combining Lexicon and Learning based Approaches for Concept-Level Sentiment Analysis”, WISDOM' 12, August 12, 2012, Beijing, China Copyright 2012, ACM.
[10] Xiaohui Yu,Yang Liu, Jimmy Xiangji Huang, and Aijun An,,” Mining Online Reviews for Predicting Sales Performance: A Case Study in the Movie Domain”, IEEE Trans.On knowledge and data engineering, vol. 24, no. 4, April 2012.
[11] LI Bing, Keith C.C. Chan, Carol OU,” Public Sentiment Analysis in Twitter Data for Prediction of A Company's Stock Price Movements”, 2014 IEEE, 11th International Conference on e-Business Engineering.
[12] Mr. Dharmesh Dhabliya, Mr. Rahul Sharma. (2012). Efficient Cluster Formation Protocol in WSN. International Journal of New Practices in Management and Engineering, 1(03), 08 - 17. Retrieved from http://ijnpme.org/index.php/IJNPME/article/view/7
[13] C, S. L. ., Saxena, S. ., & Kumar, B. S. . (2023). Design Text Mining Classifier for Covid-19 by using the Machine Learning Techniques. International Journal of Intelligent Systems and Applications in Engineering, 11(2s), 240 –. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2622
[14] M. Thelwall, K. Buckley, G. Paltoglou, and D. Cai, “Sentiment Strength Detection in Short Informal Text,” Am. Soc. Informational Sci. Technol., vol. 61, no. 12, pp. 2544–2558, 2010.
[15] P. P. Balage Filho and T. A. S. Pardo, “NILC{_}USP: A Hybrid System for Sentiment Analysis in Twitter Messages,” in Second Joint Conference on Lexical and ComputationalSemantics (*SEM), Volume 2: Proceedings of the Seventh International Workshop on Semantic Evaluation (SemEval 2013), 2013, vol. 2, no. SemEval, pp. 568–572.
[16] N. Malandrakis, A. Kazemzadeh, A. Potamianos, and S. Narayanan, “SAIL : A hybrid approach to sentiment analysis,” vol. 2, no. SemEval, pp. 438–442, 2013.
[17] Ana Silva, Deep Learning Approaches for Computer Vision in Autonomous Vehicles , Machine Learning Applications Conference Proceedings, Vol 1 2021.
[18] Juan Garcia, Guðmundsdóttir Anna, Johansson Anna, Maria Jansen, Anna Wagner. Optimal Decision Making in Supply Chain Management using Machine Learning. Kuwait Journal of Machine Learning, 2(4). Retrieved from http://kuwaitjournals.com/index.php/kjml/article/view/209
[19] A. Mudinas, D. Zhang, and M. Levene, “Combining lexicon and learning based approaches for concept-level sentiment analysis,” Proc. First Int. Work. Issues Sentim. Discov. Opin. Min. - WISDOM '12, pp. 1–8, 2012.
[20] L. Zhang, R. Ghosh, M. Dekhil, M. Hsu, and B. Liu, “Combining lexicon-based and learning-based methods for Twitter sentiment analysis,” Int. J. Electron. Commun. Soft Comput. Sci. Eng., vol. 89, pp. 1–8, 2015
[21] Li J, Qiu L. A sentiment analysis method of short texts in microblog. In2017 IEEE International Conference on Computational Science and Engineering (CSE) and IEEE International Conference on Embedded and Ubiquitous Computing (EUC) 2017 Jul 21 (Vol. 1, pp. 776-779). IEEE.