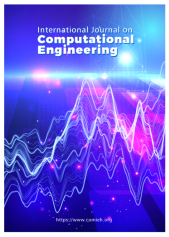
International Journal on Computational Engineering

Journal Homepage www.comien.org/index.php/comien



An Intensive Review on Opinion Mining Techniques: State of The Art, Open Cutting-Edge Challenges, and Future Research Opportunities

Gangfa Ezekiel Magaji [a], Christopher Ifeanyi Eke [b,*], Damap Tobias Yunana [a]

[a] *Department of Computer Science, Faculty of Computing, Federal University, Lafia, Nigeria*
[b] *Department of Information Technology, Faculty of Computing, Federal University of Lafia, Nigeria*
Corresponding author: [*]eke.christopher@science.fulafia.edu.ng

***Abstract*—Opinion mining (OM), a subfield of natural language processing, has gained significant attention in recent years due to its potential to uncover valuable insights from vast amounts of text data. This paper provides an intensive review of the state-of-the-art methodologies in opinion mining, highlighting their strengths and limitations. We discuss various taxonomy, including Machine Learning Technique, Reinforcement Technique, Transfer Learning Technique, Hybrid Technique, and their applications in sentiment analysis. This paper also looked at opinion classification, classification process including feature engineering, model construction and confusion matrix as well as evaluation metrics such as accuracy, precision, F1-core, recall and AUC approximation formulae of ROC curve. Our review also identifies various domain application of opinion mining including government and politics, as well as open cutting-edge challenges, such as context and handling of ambiguity, techniques integration and handling of tools, scalability as well as identification and dealing with phrases. Furthermore, we outline future research opportunities, including Explainable AI (EAI) and beyond sentiment for national security. This review aims to provide a valuable resource for scholars and academic practitioners, helping to advance the field of opinion mining and its applications in various domains without restrictions.**

***Keywords*— Area under the curve; deep learning; explainable AI; machine learning; opinion mining.**

Manuscript received 27 Aug. 2024; revised 29 Nov. 2024; accepted 9 Feb. 2025. Date of publication 31 Mar. 2025.

I. INTRODUCTION

Nowadays, every individual tends to show their feelings and express their opinion on the happenings around them. Public opinion has gotten an exponential growth on online platforms which led to fast expansion of social networks. Social network is a well-known forum for universal communication and discussion of idea among users [1]. This is very important for governments, organizations and businesses, that want to build methods in order to detect market movement for effective response to problems posed by public opinions [2].

Opinion mining (OM), also known as sentiment analysis (SA), has been in use for extracting and interpreting opinion of the public for more than fifty years by academics, research communities, and service industries [3]. [4] said that, analysis of opinion is all about detecting, mining and opinion evaluation on text or other data.

This research scope tends to carry out an extensive review on current opinion mining methods confines within state of the arts, open cutting-edge challenges and imminent research opportunities. The study tries to showcase the power of social networks, social network data, machine learning as a subset of artificial intelligent in evaluating opinions as well as opinion analysis techniques. Currently, some studies and research have been published on analysis of opinion and sentiments. study on textual opinion analysis as well as feeling detection was carried out by [5]. He offers a complete examination of the change on the current movements from text opinion analysis to detection of individual feeling including the problems associated with it. They also looked at the models of feeling classes that are frequently used. Textual detection of feeling has deviated from assessments of keyword to ML and DL techniques with better performance and flexibility. Their study did not provide a clear descriptive taxonomy for the text-based sentiment analysis techniques. [6] for instance, conducted a comprehensive survey of opinion analysis. He also inspects many uses of sentiment analysis in numerous fields like consumer feedback analysis, brand management, market research, social network analysis and political opinion analysis. It serves as an illustration of how opinion analysis impacts decision-making process and provides insights into

customer partialities and opinions. The article concluded by stress possible impending exploration areas and unanswered problems in opinion analysis like fine-grained opinion analysis, multimodal opinion analysis, and the need for models that are explainable and easy to understand. [7] carried out a study on opinion analysis. Their research study points out different techniques in opinion analysis approaches for mining and analyzing sentiments related to positive, negative or neutral polarities based on a selected subject. The study shows different procedures and classifications in opinion analysis with the way to mine and deploy data to produce better outputs. The purpose of the study was to offer idea on opinion analysis methods as well as its methods of classification.

Not withstanding, there are some constrains associated with the previous review papers. First and famous, the studies above did not provide an extensive study of the datasets used for opinion mining approaches. Second, eventhough the steps for data preprocessing is very crucial in any classification challenges, no study has provided a comprehensive and detailed study on current preprocessing methods such as feature engineering for opinion analysis. Moreover, no study could provide a complete research on feature selection and presentation structures for opinion mining. Additionaly, previuos reviews also lacked a review on performance parameters of opinion methods, such as ROC curves and AUC.

The main aim of this review study is to offer detail overview of recent research trends, challenges on opinion mining techniques that focus on datasets, techniques of feature engineering like approaches extracting feature, feature seclection approaches, feature presentation approaches; as well as taxonomy, opinion classification, classification models and score classification models including ROC curves. The objectives of this study help researchers in conducting investigation on opinion analysis and to address the following research queries: (1) What possible up-to-date taxonomies are use for sentiment analysis? (2) what are the cutting-edge sentiment classifications and processes? (3) What are the construction models for opinion classification? (4) How can sentiment analysis contribute to informed decision-making in various sectors by extracting, analyzing, and interpreting public emotions and sentiments expressed on social networks? (5) Can opinion mining have implication on business, marketing, health, education, transport, politics, and Government policies and national security? (6) What are the emerging trends and future research directions in opinion analysis?

Contribution: The novelty of this reserach study to the filed of opinion analysis is (1) To provides an extensive review on opinion mining based on, data acquisition, fearure mining, feature engeering, modelling techniques (by identifying their strengths and weaknesses), evaluation classification models including ROC curve. (2) To carry put Comprehensive review of sentiment analysis methodologies. (3) Identification of the involvement of opinion analysis capabilities across various fields like Social network, Business, politic, Government, Transportation, Health, Entertainment, Education and politics. (4) To spelled out research challenges that focuses on the imminent research direction on sentiment analysis for national security, building on a theoritical foundation that can serve as a solution to the identified challenges.

Organization: The arrangement of the proceeding sections of this reserach paper is detailed as follows: Section 2: Material and Method. Section 3: Result and Discussion. Section 4: Conclusion. This review paper concludes with findings and offers recommendations for imminent research opportunities.

II. Materials and Method

A. Taxonomy of Opinion Mining Techniques

This section deals with all the techniques included in this review used for opinion analysis like Machine Learning (ML) Technique, Reinforcement (RL) Technique, Transfer Learning (TL) Technique and Hybrid Technique.

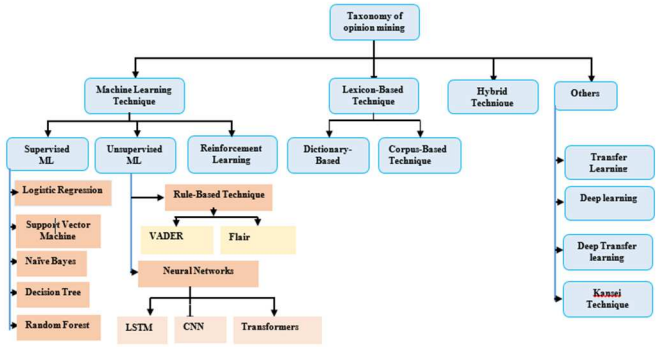


Fig. 1 Taxonomy of opinion mining

B. Machine learning-based Techniques

Ariel *et al* [8] carried out a study where number of ML methods were utilized in dealing with public sentiment classification. The three main ML techniques are Supervised learning, unsupervised learning as well as semi-supervised learning [3].

Supervised learning: [9] asserted that, the most popular ML method used the supervised learning. [10] explained that, supervised learning needs a well labeled dataset with positive and negative labels training, so that algorithm can identify whether words when used in a sequence, are likely to convey positive or negative sentiment.

Unsupervised learning: As discussed by [10], these models take a unique approach than the previous types. The training dataset involves several inputs that have not been assigned a label, so the learning model must determine the nature of inputs. Unsupervised learning techniques separates unlabeled data in groups that look alike. Take for instance, the model may look at data to be alike due to words that are common in a text [9].

C. Reinforcement Learning Techniques (RL)

Reinforcement Learning is machine learning method which permits agent to learn based on trial and error in a collaborative environment using input from its actions and experiences. Base on natural phenomenon and for practical purpose, [11] said RL is like training your dog or cat to do certain trickeries when you choose to in a challenging way. A reward is provided as goodies, if your pet absolutely

performs the desired trick, otherwise, the pet serves a punishment or lemon is given to the pet.

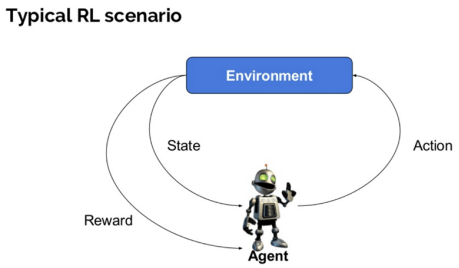


Fig. 2 Reinforcement Learning

The approaches to reinforcement learning (RL) according to [12] are; Critic-only method: This approach is furthermost applied RL in finance. The goal of RL approach is to learn a value-function-based on which agent compare ("criticized") the anticipated outcome of the actions. Actor-only method: Base on this method, action taken by agent is a result of its senses to the environmental condition. Actor-critic method: The actor and the critic are two agents involve in the technique. Action is defined by the actor and formulate the procedural rules, while evaluation of action is done by the critic.

D. Transfer learning (TL) techniques

Transfer learning under machine learning is a technique where pre-existing models are reused to solve current problems. It is a technique employed to train models together. The idea of pre-existing trained datasets is utilized to improve the performance of the ongoing problems, so the solution need not have to be built from start. The transfer learning techniques whose core depends on what, how and when to transfer research questions are: Unsupervised TL, Inductive TL as well as Trans-ductive TL [13].

E. Hybrid Approach

Hamid *et al* [14] said that, hybrid method joins several methods like vocabulary method, ML as well as Rule-Based approaches to explore individual power to minimize individual vulnerability. Take for example, vocabulary-based of the hybrid approach might be used to assign initial sentiment scores, then refine predictions based on context and other features using machine learning model. For example, in a research conducted [15], an opinion classification on Twitter at multiple levels using a hybrid machine learning model was carried out. Their paper presented a multilevel hybrid system based on machine learning (ML) for sentiment classification on Twitter.

III. Result and Discussion

A. Opinion Classification

Opinion classification involves mining and classification, whose goal is to classify texts base on polarity of opinions it contained. [16] said, there are two techniques in which opinion analysis is performed. Rule-based method: this is a well-defined description of opinion like homogeneity, subjectivity, and polarity. Rule-based algorithms include a basic natural language processing pipeline such as Derivation, Tokenization, Voice tagging, Parsing and Lexical analysis Operation. This technique checks through the sentiments and fetch out sentiment which correspond with the given criteria which then computes the dominant sentiment. The text is positively polarized If there are more positive words. Automated opinion mining technique: This type of opinion analysis leverage machine learning to find the gist of the message. In automated method, supervised ML models is use for classification while unsupervised ML methods for data exploration.

B. Opinion Classification Process

Data acquisition: The first opinion analysis step is gathering of data through different mediums, like social networks, news channels and the rest. [7] suggests that, data can be collected from sources such as X (Twitter), Facebook, blogs and marketable websites and others. [17] asserted that, in any data mining task, dataset is very crucial.

Feature Engineering: Feature engineering involves the process of selecting raw data, manipulating raw data, and converting raw data back to features that would be used in supervised learning method. In order for ML to perform very well on new assignment, better features have to be structured and trained [18]. According to [1] ; [19], the important step apply to any classification task is the feature engineering. These steps include extraction of feature, representation of feature as well as selection of feature.

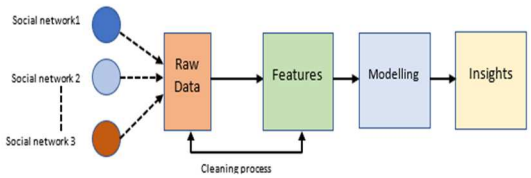


Fig. 3 Feature Engineering

Extraction of Feature: The third step of supervised learning for the task of text classification is feature extraction. This method is used to diminish the number of wanted materials to define a dataset which transforms the input data into several features. Some feature extraction techniques include Term Frequency, Inverse Document Frequency (TF-IDF), BOW (Bag of Words), Negation Processing, Word2Vec, N-Gram, Word2seq, tokenization, stemming, GloVe, POS vector, static word embedding and Term Document Matrix [20].

Feature Selection: This process involves identifying and removing redundant as well as immaterial attributes from feature, and improves the accuracy of opinion classification[21]. Different methods are used in removing wanted and unnecessary features [21]. The wrapper-based method make use of query technique to select feature among dissimilar groupings and performs computation with combinations, while embedded approach assess important features in the process model development [1]. Filter-based method embrace statistical method in assigning scores feature, while selection and removal of feature depends on the score.

Feature representation: In data analysis, features presentation refers to the process of exhibiting and emphasizing traits or features of the data in order to derive significant conclusions. This can involve arranging the data

so that patterns, trends, and relationships within the dataset are communicated clearly using tables graphs, or other visual aids; for analyst and stakeholders to quickly understand the significant aspects of the data and make choices based on the presented information.

C. Construction of Classification Models

This is the phase where classification models are developed on the training dataset using a machine learning technique. Some of these models are described below:

Naïve Bayes (NB): According to [22], a Bayesian classifier in machine learning is a probabilistic classification technique which uses theorem of Bayes. Strong independence assumptions underlie the Bayesian classifier's use of the feature model. This suggests that a class's characteristics exist independently of one another and have no bearing on one another. Independent events are defined as follows: if both E and F have positive probabilities, P(E|F) = P(E) while P(F|E) = P (F) then the two independent events are E and F.

Support Vector Machine (SVM): SVM makes use of a hyperplane, to analyze data and to set decision bounds. One popular non-probability supervised learning method for classification applications is SVM. One objectives of support vector machine are to Find the hyperplane that divides the data into discrete classes. SVM then looks for the hyperplane with the largest possible margin. [21]. According to [23], Support vector classifier is represented by equation;

$$f(x) = \quad_0 + \sum_{j \in S} (j \times K\,(Xj,\ Xj)\,) \qquad (1)$$

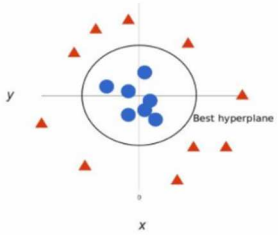


Fig. 4 Support Vector Machine [16].

Decision Tree Classifier (DTC): according to [23], an effective non-parametric supervised method for classifying and regression problems is DT classification. Decision tree consist of several decision nodes, each of which describes a decision rule that, when split to reduce impurity recursively, creates a leaf node that belongs to a class. The classifier accomplishes a recursive double partitioning in order for the feature space to split the dataset into incrementally atomic sets while the decision tree grows increasingly [24]. The idea behind is to split the sample responses into groups that are closely related as possible and are uniquely from each other, and later split the sub-sample into subgroups until it brings out decision making possibilities [11].

Random Forest (RF): RF is an ensemble learning technique where numerous DTs are trained, but each of the tree comprises of bootstrapped sample of observations called available observations. The Random Forest learning algorithm computes an overall score by relating the true value of the observation with the prediction of a subset of trees that were not trained for every observation. This overall score is taken as a measure of the forest's stochastic performance [23].

Ensemble Learning (EL): The deliberate process of bringing together several learners to create an intelligent model is known as ensemble learning. By decreasing undesirable and ineffective selection, it exacerbates the classification issue. It has the skills and experience of a variety of learners, which improves classification precision and lowers prediction mistakes.

D. Evaluation of Classification Models

Evaluation metrics for model performance such as accuracy, precision, recall, and F1 score depends on specific task and availability of real data [14].

Confusion metric (CM): This is a predictive analytic tool in ML for testing the performance of a model. Furthermore, we can say that the CM is a summary table of number of accurate and inaccurate predictions of a classifier for binary classification tasks. Matrix column represent instances of predicted class, while rows represent instances of actual class [22].

TABLE I
CONFUSION METRIX

	Predicted Class	
Actual Class	(TN)	(FP)
	(FN)	(TP)

Where N = number of target class, TN = True Negative, FP = False Positive, FN = False Negative and TP = True Positive

Accuracy: The number of rightful predictions a classifier made across the whole testing data set is known as accuracy. The percentage anticipated cases is provided, said [17]. When assessing the performance of a model, accuracy makes an excellent starting point. Accuracy of any given classifier is computed by;

$$\frac{TP + TN}{TP + T\,N + FP + FN} \qquad (2)$$

Precision: The number of times a system produce accurate result is known as precision. According to this criterion, an operation accuracy increases with number of rightful results it produces as well as number of incorrect results it produces [25].

$$Precision = \frac{TP}{TP + FP} \qquad (3)$$

Precision (PRE) provides the accuracy of the model in case false positives exist. Therefore, the accuracy of a model gives the total of false positive cases and rejected positive cases, [17].

Recall: Another name for recall is sensitivity. This is the percentage of true positive cases out of total positive cases in the classifier. Recall calculates the model classification error.

$$Recall = \frac{TP}{TP + FN} \qquad (4)$$

Karimi, (2021) said that, higher recall indicates that most positive cases (TP+FN) will be categorized as true positive. There will be more FP as a result, which reduce total accuracy. High number of false positives that are mislabeled as negatives, indicates low recall. This implies that, if a positive case is discovered, you can be more confident that it is a true positive case.

F1 Score: The term "F1 Score" means to combine recall and precision together and then calculate as harmonious average. It reaches its maximum value when precision equals recall.

$$F1 = \frac{2}{\frac{1}{recall}+\frac{1}{precision}} = 2 \text{ x } \frac{precisio\ x\ recall}{precision+recall} \quad (5)$$

ROC Curve and AUC: ROC curve stands for receiver operating characteristic curve. The ROC curve and AUC are very important concepts for evaluation in binary classification models. ROC curve used for assessing how well a binary model performs when the threshold is changed. The following are discussed in order to comprehend the ROC Curve clearly.

Sensitivity or the True Positive Rate (TPR)**:** Sensitivity represents fraction of true positives that are accurately classified as such, and then plot on y-axis of ROC curve. Sensitivity is computed by;

$$\frac{True\ Positive}{True\ Positive+False\ Negative} \quad (6)$$

False Positive Rate (FPR): FP rate, is the fraction of true negative that are mistakenly identified as positive, is represented by the curve on the x-axis. It is computed by;

$$\frac{False\ Positive}{False\ Positive + True\ Negative} \quad (7)$$

Plotting sensitivity on the y-axis and false positive rate on the x-axis at different threshold controls, produces the ROC Curve in figure 7.

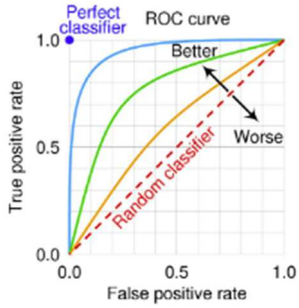


Fig. 5 ROC Curve

Thresholds: A threshold in a classification model is a point were decision is made whether a given instance belongs to a certain class.

AUC (Area Under the Curve): The effectiveness of binary classification model is determined by AUC. The region under the ROC curve from the prediction score is indicated by the area under the curve [23]. AUC with value of 1 represent better classification model, while AUC with value of 0.5 represent worse classifier. The most common method for calculating AUC of ROC is by rule of trapezium, which is approach involves approximating the definite integral function. Below are the steps to find the approximation:

Step1: The area under the curve is divided into trapezoids.

Step2: the area of each trapezoid (height + base)/2 are calculated.

Step3: Sum up the areas of all trapezoids to get the total area.

Mathematically,

$$AUC = (\sum ([y_1 + y_2)/2] * (x_2 - x_2)) \quad (8)$$

Given that y_1 and y_2 are heights of the curve at start to end of each trapezoid, while $x_1$and x_2 are the x-coordinates at start and end of each trapezoid.

E. Application of Opinion Mining Across Various Domains

Poojary [27] said, opinion mining has numerous applications in diverse fields, including journalism, politics, marketing, customer service and healthcare. Though, below are various domain application of opinion mining:

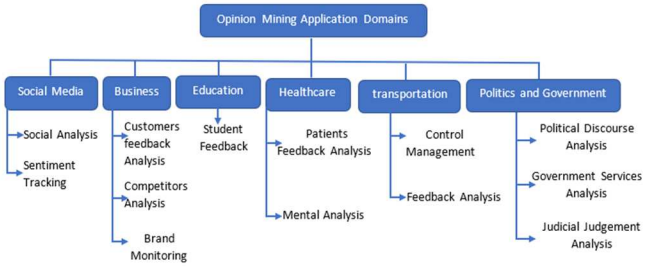


Fig. 6 Domains opinion mining application

F. Social Media Analytics

The method of gathering as well as analyzing data from social networks for measuring the effectiveness of actions taken in response to business decision is known as social media analytics. These methods include, opinion mining of Social Media and Social media sentiment tracking. according to [14], opinion tracking is extremely used in monitoring opinion of the public on social platforms in order to provide understandings of public sentiments, responses to events as well as emerging topics.

G. Business and products review

Analysis of customer feedback on opinion mining assist businesses in analyzing customer reviews, surveys, feedback for proper understanding of customer satisfaction, finding areas for improvement, and to enhance product [14]. In the aspect of competitor analysis, emotional analysis can be used to gather information from multiple platforms, including Twitter, Facebook, and blogs, produce measurable outcomes, and get around challenges with business intelligence. In Market research and brand monitoring, opinion mining can help companies to identify customer needs, preferences, and complaints, and improve their products and services accordingly. Assessing customers view on product and services is an advantage to design companies, promotion agencies, information tech companies, hotels, news networks, and others [21].

1) Politics and Government: In politics, opinion mining can be used to track public opinion on political discourse, monitor election campaigns, and predict election outcomes. [24] carried out investigation into political sentiment orientations. In the area of Judicial Judgement, sentimental analysis or Opinion mining helps in understanding public sentiments towards a Judicial Judgement concerning the nation.

2) Transportation: Kumar [28] said, analyzing the feedback provided by transport users using sentiment analysis report will be maintained for each mode of transportation, in graphical representation.

3) Healthcare: Public sentiment analysis has enables hospitals to analyze patient response, assessments or experiences in order to improve their services. Sentiment analysis to analyze drugs satisfaction on patient using supervised learning was carried out by [29]. In their work, they employed sentiment analysis technique to focus on predicting the level of satisfaction of drug among the rest of the patients who have experience of the drug effect.

4) Education: User feedback can be used in an e-learning system to evaluate academics as well as educational organizations. Academics can know their feelings on programs based on the analysis of opinions offered by students. This may encourage marketing initiatives and enhance the provision of services. By asking students to write online feedback as a prerequisite for the course, unit coordinators can learn what learners think about their instructors and teammates [30].

H. Open Cutting-Edge Challenges and Research Directions of Opinion Mining

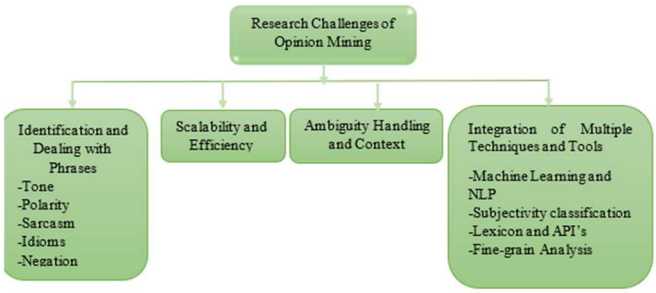


Fig. 7 Problems of opinion mining

I. Identification and Dealing with Sentiments Words and Phrases

Intonation: Although tone might be challenging to understand when spoken, it can be even more so when writing down. The capacity to distinguish between the two types of assertions and then select the appropriate tone in them is necessary for tackling this issue. Polarity: Words with strong positive (+) and negative (-) polarity ratings include "love" and "hate." These are simple to comprehend. However, some words that come between conjunctions, such "not so bad," can also signify "average," placing them in the middle of the polarity spectrum. These kinds of sentences are occasionally omitted, which lowers the sentiment score. These mid-polar phrases and words are readily recognized by sentiment evaluation software to provide a comprehensive picture of a message. Sarcasm: Sarcasm is an emotional expression technique used to humorously critique something or inflict emotional harm on anyone by expressing implicit data that is typically contrary to the intent of the message substance. One of the most difficult problems in the field of natural language processing, is sarcastic recognition in textual data. Because of its significance for enhancing sentiment analysis of social media data, sarcasm identification has recently attracted the attention of researchers [1]. Idioms: machine learning models do not necessarily understand a figure of speech. Therefore, the approach may misinterpret or even overlook a statement that contains a common phrase in an assessment or remark. Negations: Negations can cause confusion for the ML model. Examples of these terms are not, never can't, were not, etc. An emotional analysis program needs to be trained to recognize when two negatives overwhelm one another in order to take negation and convert a sentence into a positive.

J. Ambiguity Handling and Context

A word is said to be ambiguous if it contains multiple meanings that make it difficult to understand the actual intention; an example of this would be "the wrapping is awful, yet the item itself was great." When analysis of sentiment fails to accurately distinguish between these scenarios, it may misrepresent the overall feeling [14]. Opinion mining can be prone to errors due to its inability to precisely capture nuanced details and background.

K. Scalability and Efficiency

To achieve scalability and efficiency in opinion mining, automated tools resources are essential. Machine learning-backed software's and techniques such as lexicon-based approaches have been developed to handle the vast amount of data for sentiment analysis.

L. Integration of Multiple Techniques and Tools

For accurate sentimental analysis, multiple techniques and tools can be integrated. These include, Fine-Grained Analysis: this technique defines the sentiment target and its intensity, addressing challenges such as processing comparative expressions and dual-polarity comments and reviews. Machine Learning and NLP: opinion mining tools reveal the positivity or negativity in written content by utilizing ML, statistics and NLP. This can be done through rule-based sentiment analysis or more advanced algorithm-based methods. Subjectivity Classification: This approach divide text into objective and subjective fragments, allowing for a more nuance understanding of the sentiment expressed. Lexicons and APIs: Various sentiment analysis tools such as Lexalytics and MeaningCloud, offer lexicon and APIs for accurate sentiment analysis. Emotion Detection: some tools can detect specific emotions expressed in text, providing a more detailed analysis of the sentiment conveyed.

IV. CONCLUSION

The study of user-generated text data from social networks platforms to ascertain the views of individuals, feelings, assessments, behaviors, and emotions is known as

opinion mining (OM) research. In this paper, we have thoroughly analyzed opinion mining approaches. The findings demonstrate the widespread use of opinion mining in several well-known industries such as business, social media, transportation, health, education and government. Various taxonomies of opinion mining such as Reinforcement Learning, and Transfer Learning have been explored in this study. This study also looked at various performance evaluation metrices including AUC (Area Under the Curve) of a ROC Curve. Nevertheless, future research direction on opinion mining includes:

Explainable Artificial intelligence (XAI)**:** Another subset of artificial intelligence that can be understood by humans is known as Explainable Artificial Intelligence (XAI). In opinion analysis, explainable AI (XAI) refers to methods for deciphering and elucidating the choices made by AI models.

ACKNOWLEDGMENT

We acknowledge the support of Federal University Lafia in the publication of this review paper. We are grateful for their intellectual assistance, which has significantly enhanced the quality of this work.

REFERENCES

[1] C. I. Eke, A. Anir, and N. Liyana, "Sarcasm identification in textual data: Systematic review, research challenges and open directions," *Artif. Intell. Rev.*, vol. 53, no. 6, pp. 4335–4385, 2019, doi:10.1007/s10462-019-09791-8.

[2] Z. Zhao, W. Liu, and K. Wang, "Research on sentiment analysis method of opinion mining based on multi-model fusion transfer learning," *J. Big Data*, vol. 10, no. 1, p. 68, 2023, doi: 10.1186/s40537-023-00837-x.

[3] N. Afiza, M. Razali, N. A. Malizan, N. A. Hasbullah, and M. Wook, "Opinion mining for national security: Techniques, domain applications, challenges and research opportunities," *J. Big Data*, vol. 8, no. 1, p. 138, 2021, doi: 10.1186/s40537-021-00536-5.

[4] R. Arifiansyah, "Sentiment analysis in business management: A systematic review," *Ilomata Int. J. Manag.*, vol. 5, no. 1, pp. 133–156, 2024.

[5] L. P. Hung and S. Alias, "Beyond sentiment analysis: A review of recent trends in text-based sentiment analysis and emotion detection," *J. Adv. Comput. Intell. Intell. Inform.*, vol. 27, no. 1, pp. 84–94, 2023, doi: 10.20965/jaciii.2023.p0084.

[6] P. Ingale, "A comprehensive survey of sentiment analysis: Techniques, challenges, and applications," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 5, pp. 6496–6505, 2023.

[7] A. Abdullah, Q. Aqlan, B. Manjula, and R. L. Naik, "A study of sentiment analysis: Concepts, techniques, and challenges," in *Proc. Int. Conf. Comput. Intell. Data Sci.*, Singapore: Springer, 2020, pp. 1–12, doi: 10.1007/978-981-13-6459-4_1.

[8] Q. Ariel, V. Chang, and C. Jayne, "A systematic review of social media-based sentiment analysis: Emerging trends and challenges," *Decis. Anal. J.*, vol. 3, p. 100073, 2022, doi: 10.1016/j.dajour.2022.100073.

[9] A. Ligthart, C. Catal, and B. Tekinerdogan, "Systematic reviews in sentiment analysis: A tertiary study," *Artif. Intell. Rev.*, vol. 54, no. 7, pp. 4997–5053, 2021, doi: 10.1007/s10462-021-09973-3.

[10] P. Rita, N. António, A. Patrícia, and P. Rita, "Social media discourse and voting decisions influence: Sentiment analysis in tweets during an electoral period," *Soc. Netw. Anal. Min.*, vol. 13, no. 1, p. 95, 2023, doi: 10.1007/s13278-023-01048-1.

[11] A. A. Iorkaa and M. Barma, "Machine learning techniques, methods and algorithms: Conceptual and practical insights," *IOSR J. Comput. Eng.*, vol. 20, no. 8, pp. 55–64, 2021, doi: 10.9790/9622-1108025564.

[12] V. Singh, S. Chen, M. Singhania, B. Nanavati, and A. Gupta, "How are reinforcement learning and deep learning algorithms used for big data based decision making in financial industries – A review and research agenda," *Int. J. Inf. Manag. Data Insights*, vol. 2, no. 2, p. 100094, 2022, doi: 10.1016/j.jjimei.2022.100094.

[13] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, "Transfer learning: A friendly introduction," *J. Artif. Intell. Mach. Learn.*, vol. 4, no. 2, pp. 45–60, 2022.

[14] H. H. Hamid, S. Bangare, A. Nirmal, and P. S. Bangal, "Sentiment analysis: A comprehensive review of techniques and applications," *Int. J. Eng. Res. Technol.*, vol. 10, no. 8, pp. 775–779, 2023.

[15] A. C. Ojha, P. K. Shah, S. Gupta, and S. Sharma, "Classifying Twitter sentiment on multi-levels using a hybrid machine learning model," *Intell. Syst. Appl. Eng.*, vol. 12, pp. 328–333, 2024.

[16] P. A. Grana, "Sentiment analysis of text using machine learning models," *Int. J. Comput. Sci. Eng.*, vol. 10, no. 5, pp. 2952–2956, 2022.

[17] C. I. Eke, A. A. Norman, and L. Shuib, "Context-based feature technique for sarcasm identification in benchmark datasets using deep learning and BERT model," *IEEE Access*, vol. 9, pp. 123456–123467, 2021.

[18] A. Patil, "A review on sentiment analysis approaches," *J. Data Anal. Inf. Process.*, vol. 9, no. 2, pp. 45–60, 2021.

[19] K. K. Mujtaba, L. Shuib, N. Idris, and W. L. Hoo, "Clinical text classification research trends: Systematic literature review and open issues," *J. Med. Syst.*, vol. 42, no. 1, pp. 21–33, 2018, doi: 10.33094/7.2017.2018.31.21.33.

[20] N. Varghese, B. E. Grace, and M. Kanaga, "Sentiment analysis in social media data for depression detection using artificial intelligence: A review," *SN Comput. Sci.*, vol. 3, no. 1, p. 74, 2022, doi: 10.1007/s42979-021-00958-1.

[21] M. Wankhade, A. Chandra, S. Rao, and C. Kulkarni, "A survey on sentiment analysis: Challenges and applications," *Artif. Intell. Rev.*, vol. 55, no. 7, pp. 5735–5780, 2022, doi: 10.1007/s10462-022-10144-1.

[22] B. Klein, *Machine Learning with Python Tutorial*. O'Reilly Media, 2021.

[23] V. Umarani, A. Julian, and J. Deepa, "Sentiment analysis using various machine learning and deep learning techniques," *J. Data Sci.*, vol. 3, pp. 385–394, 2021.

[24] M. Z. Ansari, M. B. Aziz, M. O. Siddiqui, H. Mehra, and K. P. Singh, "Analysis of political sentiment orientations on Twitter," *Procedia Comput. Sci.*, vol. 167, pp. 1821–1828, 2020, doi: 10.1016/j.procs.2020.03.201.

[25] A. Ullah, S. N. Khan, and N. M. Nawi, "Review on sentiment analysis for text classification techniques from 2010 to 2021," *Multimed. Tools Appl.*, vol. 82, no. 5, pp. 6789–6824, 2023, doi: 10.1007/s11042-022-14112-3.

[26] Z. Karimi, *Confusion Matrix: A Practical Guide for Machine Learning Practitioners*. Springer, 2021.

[27] A. Poojary, S. Kapoor, and M. Studies, "A review of natural language processing techniques," *J. Nat. Lang. Process.*, vol. 8, no. 1, pp. 1–16, 2023.

[28] G. L. Kumar, I. Karan, R. Pati, and S. Koduru, "Sentiment analysis in transportation system," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 11, pp. 618–627, 2021.

[29] V. Gopalakrishnan and C. Ramaswamy, "Patient opinion mining to analyze drugs satisfaction using supervised learning," *J. Med. Syst.*, vol. 41, no. 15, pp. 311–319, 2017.

[30] H. Binali, V. Potdar, and C. Wu, "A state of the art opinion mining and its application domains," in *Proc. IEEE Int. Conf. Ind. Technol.*, 2014, pp. 1–6, doi: 10.1109/ICIT.2009.4939640.