A Survey Paper on Sentiment Analysis : Approches, Methods & Challenges

Dr.K.Anuradha[1], Dr.M.Vamsi Krishna[2], Dr.Banitamani Mallik[3],Prof.B.P.Mishra[4]

[1]*Research Scholar, Centurion University of Technology & Management, Odisha and Assistant Professor, Dr. L. B. College of Engineering for women, Visakhapatnam.*

[2]*Principal, Chaitanya Institute of Science & Technology, Kakinada*

[3]*Associate Professor, Centurion University of Technology & Management , Odisha*

[4]*Director, (R&D), GITA, Bhuvaneswar, Odisha*

Abstract

Sentiment Analysis is the domain of understanding these emotions with software, and it's a must-understand for developers and business leaders in a modern workplace. As with many other fields, advances in Deep Learning have brought Sentiment Analysis into the foreground of cutting-edge algorithms. Today we use natural language processing, statistics, and text analysis to extract, and identify the sentiment of text into positive, negative, or neutral categories. In this paper, an attempt is made to give an overview of different methods available for sentiment analysis, along with different approaches, challenges in sentiment analysis

Keywords: *Sentiment Analysis, Data Mining, SVM, Deep Learning Algorithms*

I. Introduction

Sentiment Analysis or Opinion Mining is a type of contextual mining for text, which identifies and extracts subjective information in source material, and also helps a business to understand the social sentiment of their brand, product or service while monitoring online comments or feedback. However, analysis of social media streams is usually restricted to just basic sentiment analysis and count based metrics. This is just scratching the surface and missing out on those high value insights that are waiting to be discovered. So what should a brand do to capture that low hanging fruit?

With the recent advances in deep learning, the ability of algorithms to analyze text has improved considerably. Creative use of advanced artificial intelligence techniques can be an effective tool for doing in-depth research. We believe it is important to classify incoming customer conversation about a brand based on what are the key aspects pertaining to that brand and what are the reactions and intentions concerning these aspects.

The online medium has become a significant way for people to express their opinion and with social media; there is an abundance of opinion information available. There are more than a hundred websites such as Facebook, Twitter, LinkedIn and Youtube offering a range of services for users who depend on the networks for sharing their interests, photos, videos and blogs. However, the main purpose of online social networks is sharing information. According to the ranking from ebizMBA in January2019, top 10 social networking sites are shown in below table [4].

Table 1

S.NO	WEBSITE	NO OF USERS
1	facebook	1,500,000,000
2	YouTube	1,499,000,000
3	Twitter	3 400,000,000
4	Instagram	275,000,000
5	LinkedIn	250,000,000
6	Reddit	125,000,000
7	VK	120,000,000
8	Tumblr	110,000,000
9	Pinterest	105,000,000
10	Google Plus	100,000,000

II. Sentiment Analysis: This is one of the most common text classification tool which analyses an incoming message and tell whether the underlying sentiment is positive, negative our neutral.

Sentiment Analysis can be used to determine sentiment on a variety of level. It will score the entire document as positive or negative, and it will also score the reaction of individual words or phrases in the document. Sentiment analysis can track a particular topic, many companies use it to track or observe their products, services or status in general. For Example, if someone is attacking your brand on social media, sentiment analysis will score the post as enormously score negative, and you can create alerts for posts with hyper negative sentiment scores. [3]

Intent Analysis: This analysis steps up the game by analyzing the user's intention behind a message and identifying whether it relates an opinion, news, marketing, complaint, suggestion, appreciation or query.Intent analysis basically detects what people want to do with a text rather than what people say with that text. Look at the following examples:

- "Your customer support is a disaster. I've been on hold for 20 minutes".
- "I would like to know how to replace the cartridge".
- "Can you help me fill out this form?"

A human being has no problems detecting the complaint in the first text, the question in the second text, and the request in the third text. However, machines can have some problems to identify those. Sometimes, the intended action can be inferred from the text, but sometimes, inferring it requires some contextual knowledge.

Contextual Semantic Search(CSS): This is one of the interesting approaches. To derive actionable insights, it is important to understand what aspect of the brand a user is discussing about. For example: Amazon would want to segregate messages that are related to: late deliveries, billing issues, promotion related queries, product reviews etc. On the other hand, Starbucks would want to classify messages based on whether they relate to staff behavior, new coffee flavors, hygiene feedback, online orders, store name and location etc.

We introduce an intelligent smart search algorithm called Contextual Semantic Search (a.k.a. CSS). The way CSS works is that it takes thousands of messages and a concept (like Price) as input and filters all the messages that closely match with the given concept. The graphic shown below demonstrates how CSS represents a major improvement over existing methods used by the industry.

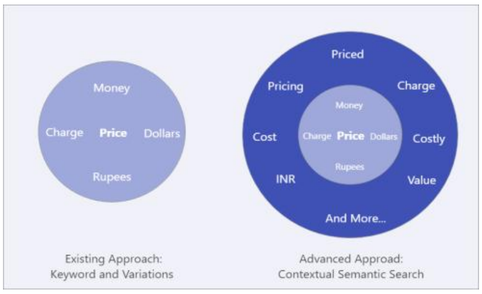

Figure 1. Existing approach vs Contextual Semantic Search

A conventional approach for filtering all Price related messages is to do a keyword search on Price and other closely related words like (pricing, charge, $, paid). This method however is not very effective as it is almost impossible to think of all the relevant keywords and their variants that represent a particular concept. CSS on the other hand just takes the name of the concept (Price) as input and filters all the contextually similar even where the obvious variants of the concept keyword are not mentioned.

For the curious people, we would like to give a glimpse of how this works. An AI technique is used to convert every word into a specific point in the hyperspace and the distance between these points is used to identify messages where the context is similar to the concept we are exploring. A visualization of how this looks under the hood can be seen below:

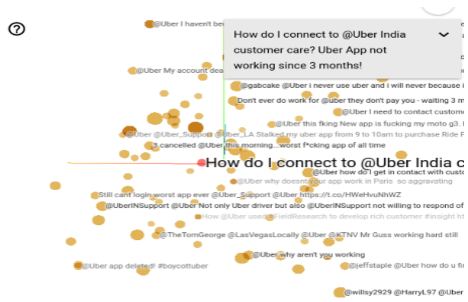

Figure 2. Visualizing contextually related Tweets

III. Scope of sentiment analysis

Sentiment analysis can be applied at different levels of scope by breaking the document into its basic grammatical level. Now some sentiment-bearing phrases such as "Worst-service" are identified using specially designed algorithm. Eachsentiment-bearing phrase is given a score based on a logarithmic scale between -10 and 10. Finally, the scores are summed up tocalculate the overall sentiment of the document or sentence. The scores of the document range between -2 and 2. Most of the researchers have mainly classified the study of sentiment analysis into three basic levels of granularity. [2]

- Document level sentiment analysis obtains the sentiment of a complete document or paragraph. The major hindrance in this level is that the useful insights are hidden so that the clients cannot extract useful information. [1]
  - Sentence level sentiment analysis obtains the sentiment of a single sentence .Here, the sentences related to factual information are separated from the sentences related to subjective views and opinions. [5] Then the opinionated sentences are expressed as either positive or negative opinion, which is called sentiment classification. In this we have three-classification problem, that is, to classify a sentence as positive, negative or neutral. **1.**
  - Sub-sentence level sentiment analysis obtains the sentiment of sub-expressions within a sentence. This can be broadly specified as the fine graining when compared to both the earlier approaches i.e. Document Level and Sentence level. The main goal of this level is to target the features of entities which are called as targets. [6] The aspect level analysis involves few steps which are aspect extraction, entity extraction, and aspect sentiment classification. For example, from the sentence, "At Novotel the food quality is good, but the service is worst", entity extraction should identify "Novotel" as the entity and aspect extraction must identify the "food quality" and "the service" as the two aspects. Aspect classification should classify the sentiment expressed on the "food quality" and "the service" as "positive" and "negative".

IV. Overview of Text Mining and Analytics

Since an enormous amount of data has emerged over the years at a staggering rate, there is a need to incorporate some sort of analytics to gain meaningful insights from the raw and unstructured data in the form of text, images, and videos. Text mining is one of the approaches that are the predecessors to text analytics. Text mining uses natural language processing, knowledge management, data mining, and machine learning techniques to process text documents [8]. Text analytics, while similar to text mining in terms of method, usually deal with a bigger amount of data to extract and generate useful non-trivial information and knowledge [9]. Text mining/analytics are originally conducted for two purposes. The first purpose is to analyze people's sentiment on an issue or phenomenon. Hence, sentiment analysis goes through a huge amount of textual data to identify people's attitudes, thoughts, judgments, and emotions on an issue [10], [11]. The second purpose is to assess people's opinion on a product, person, event, organization, or topic from a user or group of user perspectives. Similar to sentiment analysis, opinion mining is a natural language processing task that employs an algorithmic technique to recognize opinionated content and categorize it into positive, negative, or neutral polarity [12]. Nonetheless, the application of text mining/analytics has been extended to other areas of human computer applications, and the applications are growing with the growth in big data analytics.

V. Overview of Opinion Mining and Sentiment Analysis

An opinion refers to a person's or group's sentiment or views, emotions, and attitudes about a product, service, occasion, or other topic present in the environment. Like sentiment analysis, opinion mining is also grounded on the algorithmic technique [12]. Covering a huge variety of pubic opinions, [13] have argued that opinion can be classified into three main types: regular opinions, which refer to a single entity only; comparative opinions, which compare or contrast more than one entity; and suggestive opinions, which suggest a single or multiple entities. The regular opinion is mainly used to identify a positive or negative outlook of a particular product [14]. On the other hand, comparative opinions help in elucidating the association among multiple entities and are mainly used for competitive intelligence [14]. However, there is a dearth of literature concerning the identification of comparative sentences that is being used for the comparison of multiple entities. Recently, suggestive review has been introduced in the field of opinion mining [15]. The extraction of these types of opinions from text can be utilized for various application areas in the field of business, engineering, medical science, and e-learning. It can be offered to various online communities for their assistance as well [13].

Similarly, private statements of individuals are called sentiments, which comprise thoughts, opinions, attitudes, views, judgments, and feelings. These are commonly gathered by conventional scientific methods [10], [11]. [16] Pronounced the feelings that are expressed in language by using subjective expression. The sentiments can be analyzed through the machine learning technique, which can be further classified into supervised and unsupervised, using a lexicon-based approach, using a keyword, and using a concept-based technique [17]. Recently, research on sentiment analysis has focused on multiple modalities such as in speech and video as opposed to earlier work that focused on unimodality related to text [18], [19]. Sentiment analysis tackles many NLP subtasks, including aspect extraction [20], subjectivity detection [21], named entity recognition, and sarcasm detection [22].

In most cases, the main objective of sentiment analysis is to unearth people's opinions to gain meaningful insight about products or services. Its aim is to exhibit useful information to both customers and manufacturers. It is established that both manufacturer and customers look upon summarized opinions instead of detailed reviews. Hence the opinions that are categorized on positive, negative, or neutral sentiments are useful for both parties in making the right call [23]. Despite the large number of studies on opinion mining and sentiment analysis techniques, the impact they have on people has been less explored. There has been great emphasis on the techniques used and less on how people can benefit from the findings. Hence, this study aims to investigate the human element in opinion mining and sentiment analysis research. To achieve this aim, we will systematically review the relevant literatures that have employed both approaches.

The study offers several contributions. The first and foremost significance of this study is to refocus the study of opinion mining and sentiment analysis to both technical and non-technical challenges. Secondly, it places emphasis on the areas of opportunity by looking at the trends of application coverage that would offer some potential areas for research. Thirdly, the paper presents information on different datasets that were used in opinion mining and sentiment analysis studies that future researchers could use in their research.

VI. Review of Opinion Mining and Sentiment analysis

Literature reviews are rooted in medical science, which has been categorized as a critical approach mainly applied where evidence is important [24]. They involve a stringent process of finding, selecting, and examining secondary data [25], [26]. The synthesis of evidence from the current literature can create new knowledge in the current studies, which is as significant as conducting new research [27].

Rousseau et al. [28] maintained an argument that systematic reviews are different from tradition reviews in that systematic literature reviews are comprehensive in nature, use transparent and unbiased analysis, and apply certain criteria for interpretation of the findings that provided in the previous literature. In addition, systematic literature reviews mainly focus on objectivity and reproducibility of results [26]. The process of review starts with framing the questions and conducting a systematic and step-by-step process and applying a replicable method to answer these questions [26]. Thus, the evidence generated from the rigorous approach of identifying, selecting, and analyzing the data can have a significant impact on the body of knowledge collected, but the supreme concern of this practice is synthesizing the results produced through this systematic process [25], [26].

The methodology used in the study is a five-step process shown in Figure 3, as proposed by [25]. It is systematic in nature, clear and reproducible, and involves identifying, examining, synthesizing, deducing, and reporting the evidence from the existing sentiment analysis and opinion mining literature.

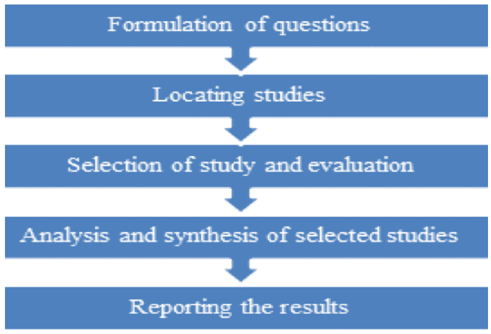

Figure 3. Research methodology of systematic literature review

A. Question Formulation

A deep and insightful literature review should start with the development of a clear understanding about your objectives [27]. Therefore, to ascertain this, we clearly formulated and considered research questions to evade doubts in our study [28]. The purpose of the paper is to discuss the methodological and application side of opinion mining and sentiment analysis and explore whether the intervention of opinion mining and sentiment analysis would be applicable to humans or in an organization as a whole. Hence the purpose of our systematic review is to answer two research questions:

1. What are the trends of opinion mining and sentiment analysis publication from 2000–2018?
2. Which are the areas in which opinion mining and sentiment analysis have been applied?
3. What are the sources of data in the areas in which opinion mining and sentiment analysis have been applied?

B. Locating Studies

The objective of identifying an appropriate journal articles is to develop a list of all related articles to our research questions. We have selected Web of Science as a core database. The study has only focused on peer-reviewed articles that were written in English and published in the Web of Science Journal category. The study did not consider articles in other categories like conference proceeding papers, book chapters, review papers, and theses.

Since the study is based on opinion mining and sentiment analysis, we used different strings to identify relevant papers. The researcher employed different techniques for searching, including separate keywords for "sentiment analysis" and "opinion mining", combining two keywords at the same time through simple operators and Boolean logic. For example, we used the string "opinion mining" and "sentiment analysis" for exploring the papers containing the exact phrases "opinion mining" and "sentiment analysis."

C. Study Selection and Evaluation

In order to ensure and maintain the quality of the paper, we have constrained our selection of articles to only peer-reviewed journals. Peer-reviewed journals have strict quality control and have gone through systematic, rigorous processes and have stringent requirements for publication, which leads to better research output [27]. The process began with scanning of selected articles from the Web of Science database. The timeline we set was from the years 2000 to 2018. The initial criterion of selection was based on choosing the keywords "opinion mining" and "sentiment analysis."

We have further followed the criteria suggested by [29] for the inclusion and exclusion of the articles from the selected list. Basically, we have developed the inclusion criterion as follows:

- Published in peer-reviewed journals
- Within the database of the last 17 years (2001 to 2018).
- In the English language
- Containing at least one keyword

The exclusion criteria are as follows:

- Have very narrow horizon or context
- Do not explicitly focus on application side (human or organizational level) of opinion mining and sentiment analysis

Our search resulted in an initial list of 274 articles for sentiment analysis and 91 articles for opinion mining. Thus, the preliminary results provided a total of 365 articles that were written on opinion mining and sentiment analysis.

The next step involved reading the abstracts to evaluate whether it was relevant to our research topic. Initially a single person read it, but to warrant its rigor, an independent person to improve its objectivity and validity read the same number of articles. Scholarly outputs that did not align with our developed research questions or that seemed irrelevant and non-substantive were excluded. The articles that were included exhibited good fit with the objective of the study. Thus, a total of 99 articles were shortlisted based on the initial evaluation.

In the next step, two authors reviewed the pre-selected articles separately. The total number of peer-reviewed journal papers selected for critical reviews after rigorous assessment were 58, published over a period of 17 years. The selected papers were then examined in detail and synthesized to answer the research questions.

VII. Methods for Sentiment Analysis

In this section we discuss the general process for sentiment analysis. A common process of sentiment analysis loop starts with goal setting for employing the sentiment analysis. This depends on the application of sentiment analysis, and then big data is extracted either from a single or multiple sources. The next step is the application of specific sentiment analysis methods in order to mine this huge data and get insight into the company in making final decisions regarding their products.

A. Keyword-Based Classification

This method classifies text based on the presence of positive or negative polarity words such as happy, joyful, delighted, miserable, sad, terrified, and uninterested [30]. The main drawback of keyword-based classification is the inability to steadfastly classify the negated words and polarity, as this approach depends on surface features [30]. Another drawback is that this approach is based on the obvious presence of positive or negative polarity. However, occasionally, a post may covey sentiment or opinion through underlying meaning rather that obvious polarity words [30].

B. Lexicon-Based Classification

Lexicon-based approaches construct lists of words manually labeled as having positive and negative polarity, and a polarity score for each word is created. This constructed lexicon is used to calculate the overall sentiment score of a given post or text. The notable advantage of the lexicon-based method is that these methods do not need training data (as the supervised machine-learning method does). The lexicon-based method is widely used in conventional text like reviews, forums, and blogs [31], [32]. However, they are less likely to be used for big data extracted from social media websites [32]. The key reason is the unstructured format and nature of social media websites (the data contains textual peculiarities, informal and dynamic nature of language, new slang, abbreviations, and new expressions) [32]. Even though this approach outperforms the keyword-based classification, it still has drawbacks. Since it works at the word level, negated posts and posts with other meanings trick the lexicon polarity score measurement [30]. Second, lexical dictionary and polarity scores are usually biased toward the text of a specific type, dictated by the linguistic corpora source [30]. Therefore, it is challenging to construct a more generalized model regardless of the application domain.

C. Machine Learning-Based Approach

Machine learning research has become a significant task in numerous application areas. Machine learning reaches throughout recent years have magnificently created algorithms for handling volumes of data to unravel real-world issues. Machine learning algorithms are grouped into supervised learning and unsupervised learning algorithms. Supervised learning algorithms will help users train and learn from the training example, which is then tested and evaluated using the test data. The main drawback of supervised machine learning algorithms is the obligation to create a training example. The training example must be comprehensive enough to make the algorithm effective and reliable enough to classify the instance in test data. Another type of machine learning is a unsupervised learning algorithm. The working principle of this algorithm is to identify the hidden associations in unlabeled data. The unsupervised learning methods are based on calculating similarity differences between data. For example, it calculates k-means in which similarity between data is computed based on proximity measures, such as Euclidean distance.

The constructing machine learning-based method involves the following important steps.

1) *Features Extraction*

As shown in Figure 4, feature extraction is an important part of building an effective machine learning method in which the textual posts (P1,P2,P3,…,Pn) are transformed into valuable word features (wf1,wf2,wf3,…wfn) by using various feature engineering approaches. Feature extracting is one the most important steps of constructing effective classifiers [33]–[34][35][36]. The accomplishment or failure of the sentiment classification model is intensely dependent on the features quality. If the extracted features relate well with the sentiment polarity and can provide discrimination power between positive and negative, then classification will be more precise. In contrast, if the extracted features do not relate well with the sentiment polarity and the similar features exist in both positive and negative posts, then the classification task will be more challenging and less precise. The most commonly used features are the first automatic generated features of Bag of Words (BoW), Bag of Phrases (BoP), n -gram, and Bag of Concepts (BoC). The second group of features is based on lexical features such as opinion words, sentiment words, and negation words. The third group of features is varied based on the data source; for example, the data from social media normally adds features such as the number of hashtags and social media-related features such as abbreviations and emojis.

2) *Machine Learning Algorithms*

This subsection briefly describes the most commonly used machine learning algorithms in literature for sentiment analysis.

a: ***Artificial Neural Network (ANN)***

The Artificial Neural Network (ANN) is a mathematical modeling approach that is stimulated by the operative processes of the human mind [37].

It is based on the artificial adaptive system, which has some form of distributive architecture [38]. The system in ANN encompasses closely-knit adaptive processing elements called artificial neurons or nodes which are proficient in carrying out enormous analogous and parallel computations for the purpose of information processing and knowledge representation [38].

These systems can also adapt to their internal structure in relation to a functional purpose. There are several types of artificial neural networks, which have been used for different problems. Generally, their application is appropriate for a problem, which is nonlinear in nature [39]. The nature of the problem can be carrying out pattern recognition, modeling memory, and envisaging the development of a dynamic system [38]. In most cases, these network types perform data modeling which is driven in a supervised or unsupervised fashion. The supervised learning technique is provided with both input and output. Then the network uses this input to generate output, and hence it is compared to desired outputs. On the other hand, in unsupervised training, only input is being provided and the network is used to find natural grouping with a dataset independent of external constraints. That is, the system itself must have the capability to decide which features to incorporate in order to group the input data.

b: ***Random Forest***

Random Forests is a classification and regression method based on the ensemble of a proliferation of decision trees [40]. Recently, attention has been given to ensemble learning, a method which can create several classifiers and produce aggregate results [41]. The two commonly known and used methods for the classification of trees are boosting as proposed by [42] and bagging as proposed by [43]. The latter proposed the general operating mechanism of the Random Forest (RF), which augments the additional layer of randomness to bagging. In RF, each tree is a standard Classification or Regression Tree (CART) that uses the so-called splitting criteria like Decrease of Gini Impurity (DGI). Moreover, it picks the splitting predictor from a randomly chosen subset of predictors. Each tree is fabricated by using a bootstrap sample of the data, and the prediction of all trees are ultimately accumulated by means of majority voting [40].

c: ***Support Vector Machine***

Support vector machines are considered to be universal learners. Generically they learn linear threshold functions. However, with the use of a suitable kernel function plug in, they can be used to learn in different applications in the form of radical basic function and sigmoid neural networks and be trained on polynomial classifiers [44]. SVMs were initially intended for binary classification, but research has extended it into a multiclass classification. There are two commonly used approaches for multiclass SVMs. The first deals with fabricating and conjoining the number of binary classifiers, and the second one directly involves keeping all data in one optimization construction [45].

d: ***Genetic Algorithm***

The idea of developing the Genetic Algorithm was initiated and developed by John Holland. He proposed this idea in the year 1975 in his book "Adaptation in natural and artificial systems". Since then, the Genetic Algorithm (GA) has been increasingly recognized as a popular evolutionary computational research technique [46]. It has gained popularity over the years as an optimization tool in a variety of research domains, including computer science, operational research, engineering, management, and social sciences [47]. A major reason behind the realization of this technique is its diversified applicability, efficiency of operations, and applicability on global scenario [48].

Genetic Algorithms are search and optimization techniques in a multifaceted search space. They are inspired by the concept of genetics and natural selection. Some essential and principal ideas are adopted from the field of genetics and then used artificially to create a kind of algorithm, which is flexible, robust, and efficient in nature [49]. Moreover, it characterizes the emergent technique, which is to be used to understand different relationships in the course of the development of data, in which the data can come in the form of binary strings, formula, program, query, grammar, and images [47].

e: ***Naïve Bayes***

The Naïve Bayes (NB) classifier has long been used in most applications of supervised machine learning. It is considered a tool for the retrieval of data [50]. It is based on a simple theorem of probability for making a probabilistic model of data. The mechanics of the NB algorithm are applied to numeric data [51]. It is simple, easy to understand, and quick for classification. It normally entails a minimal data set for training and then is used to predict the parameters needed for classification purposes.

f: ***Decision Tree***

The decision tree (DT) classifier has been widely used for prediction and classification of tasks. The rules in creating the decision tree are easy to understand. The classifiers built through the decision tree are given in hierarchical representation. The tree is composed of decision nodes, event nodes, edge, and path [52]. A variety of classifiers are used in a variety of applications. Some of the DT classifiers are ID3, C4.5, and C5.0, but a common problem that has been found in the DT classifier is the ability to incorporate all types of variations in data, including noise when trees get bigger and deeper. This problem is commonly known as overfitting. In addition, the structure of the tree is distorted with the addition of data. To avoid this problem, the random forest technique is recommended in which many trees are formed and trained by dividing the training sets, and outcomes are generated through aggregation of all trees.

g: ***K-Nearest Neighbour***

The K–nearest neighbor algorithm performs classification based on instance learning. It works through a non-parametric procedure of storing all inputs and instances and classifies new inputs by means of similarity measures like Euclidean distance [46], [47].

h: ***Ensemble Voted Classifier***

An ensemble of classifiers is the collection of many classifiers in which their decisions are aggregated by means of weighted an unweighted voting mechanism to predict the outcome. The classification in the ensemble voted classifier is based on a voting mechanism in which classification of new instances is done by considering the majority vote of the prediction [43], [48]–[49][50].

VIII. Validation Metrics

To validate the proposed machine learning/ Deep learning models, one has to measure the performance by standard metrics. The standard metrics for performance is derived from the special matrix called confusion matrix. Confusion matrix gives the True positives, False Positives, True Negatives and False Negatives as elements.

True Positive (TP): Number of predictions as Positives, as they are Positive in ground truth
False Positive (FP): Number of predictions as Positives, as they are not positive in ground truth
True Negative (TN): Number of predictions as Negatives, as they are Negative in ground truth
False Negative (FN): Number of predictions as Negatives, as they are not negative in ground truth

True Positive (TN)	False Positive (FP)
False Negative (FN)	True Negative (TN)

Table 2. Confusion Matrix

Accuracy (ACC) is calculated as the number of all correct predictions divided by the total number of the dataset. The best accuracy is 1.0, whereas the worst is 0.0.

Sensitivity (Recall or True positive rate) Sensitivity(SN) is calculated as the number of correct positive predictions divided by the total number of positives. It is also called recall (REC) or true positive rate (TPR). The best sensitivity is 1.0, whereas the worst is 0.0.

Specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives. It is also called true negative rate (TNR). The best specificity is 1.0, whereas the worst is 0.0

Precision (PREC) is calculated as the number of correct positive predictions divided by the total number of positive predictions. It is also called positive predictive value (PPV). The best precision is 1.0, whereas the worst is 0.0.

Negative positive rate (NPR) is calculated as the number of incorrect positive predictions divided by the total number of negatives. The best false positive rate is 0.0 whereas the worst is 1.0. It can also be calculated as 1 – specificity.

The false positive rate usually refers to the probability of falsely rejecting the null hypothesis for a particular test.

The false discovery rate (FDR) is a method of conceptualizing the rate of errors in null hypothesis testing when conducting multiple comparisons.

False omission rate (FOR) is a statistical method used in multiple hypothesis testing to correct for multiple comparisons and it is the complement of the negative predictive value. It measures the proportion of false negatives which are incorrectly rejected.

Mathew's correlation coefficient and F-score can be useful, but they are less frequently used than the other basic measures. F-score is a harmonic mean of precision and recall.

Matthews correlation coefficient (MCC) is a correlation coefficient calculated using all four values in the confusion matrix.

IX. Challenges to Sentiment Analysis

Sentiment Analysis runs into a similar set of problems as emotion recognition does – before deciding what the sentiment of a given sentence is, we need to figure out what "sentiment" is in the first place. Is it categorical, and sentiment can be split into clear buckets like happy, sad, angry, or bored? Or is it dimensional, and sentiment needs to be evaluated on some sort of bi-directional spectrum?[7]

In addition to the definition problem, there are multiple layers of meaning in any human generated sentence. People express opinions in complex ways; rhetorical devices like sarcasm, irony, and implied meaning can mislead sentiment analysis. The only way to really understand these devices are through context: knowing how a paragraph is started can strongly impact the sentiment of later internal sentences.

Most of the current thinking in sentiment analysis happens in a categorical framework: sentiment is analyzed as belonging to a certain bucket, to a certain degree. For example, a given sentence may be 45% happy, 23% sad, 89% excited, and 55% hopeful. These numbers don't add up to 100 – they're individual indications of how "X" a sentence's sentiment is.

To address the context issue, a lot of research surrounding sentiment analysis has focused on feature engineering. Creating inputs to a model that recognize context, tone, and previous indications of sentiment can help increase accuracy and get a better overall sense of what the author is trying to say. For an interesting example, check out this paper in Knowledge-Based Systems that explores a framework for this kind of context focus.

Finally, one more challenge in sentiment analysis is deciding how to train the model you'd like to use. There are a number of pre-trained models available for use in popular Data Science languages. For example, TextBlob offers a simple API for sentiment analysis in Python, while the Syuzhet package in R implements some of research from the NLP Group at Stanford.

These modules can help you get off the ground quickly, but for the best long term results you're going to want to train your own models. Getting access to labeled training data for sentiment analysis can be difficult, but it's key to building models that work for your specific use case. You may execute a workflow where you gather your proprietary data (e.x. customer service conversations) and use a service like CrowdFlower to label and prepare it.

X. Conclusion and future scope

Sentiment Analysis works to analyze the emotion of the people, which is written in the form of text. Further different algorithms and different data mining techniques are to be applied to get accurate emotions. In order to get accuracy, we try to use deep learning algorithm so as to get more accuracy and optimum solutions.

XI. References

[1] B. Pang, et al., "Thumbs up?: sentiment classification using machine learning techniques," in Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10, 2002, pp. 79-86.

[2] Turney, Peter D. "Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews." Proceedings of the 40th annual meeting on association for computational linguistics.Association for Computational Linguistics, 2002.

[3] Rajini Singh, RajdeepKaur, Sentiment analysis on social media and online review.

[4] http://www.ebizmba.com/articles/social-networking-websites

[5] S.-M. Kim and E. Hovy, "Determining the sentiment of opinions," in Proceedings of the 20th international conference on Computational Linguistics, 2004, p. 1367.

[6] T. Wilson, et al., "Recognizing contextual polarity in phrase-level sentiment analysis," in Proceedings of the conference on human language technology and empirical methods in natural language processing, 2005, pp. 347-354.

[7] ShahidShayaa, et al, "Sentiment analysis of Big Data: Methods Applications and Challenges", IEEE Access ,Vol – 6 Issue 2018, Pg: 37807 – 37827

[8] F. R. Lucini et al., "Text mining approach to predict hospital admissions using early medical records from the emergency department", Int. J. Med. Inf., vol. 100, pp. 1-8, Apr. 2017.

[9] Z. Khan, T. Vorley, "Big data text analytics: An enabler of knowledge management", J. Knowl. Manage., vol. 21, no. 1, pp. 18-34, 2017.

[10] T. T. Thet, J.-C. Na, C. S. G. Khoo, "Aspect-based sentiment analysis of movie reviews on discussion boards", J. Inf. Sci., vol. 36, no. 6, pp. 823-848, 2010.

[11] H. Yu,V.Hatzivassiloglou, "Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences", Proc. Conf. Empirical Methods Natural Lang. Process., pp. 129-136, 2003.

[12] R. Piryani, D. Madhavi, V. K. Singh, "Analytical mapping of opinion mining and sentiment analysis research during 2000–2015", Inf. Process. Manage., vol. 53, no. 1, pp. 122-150, 2017.

[13] A. Qazi, A. Tamjidyamcholo, R. G. Raj, G. Hardaker, C. Standing, "Assessing consumers' satisfaction and expectations through online opinions: Expectation and disconfirmation approach", Comput. Hum. Behav., vol. 75, pp. 450-460, Oct. 2017.
[14] N. Jindal, B. Liu, "Identifying comparative sentences in text documents", Proc. 29th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retr., pp. 244-251, 2006.
[15] A. Qazi, R. G. Raj, M. Tahir, E. Cambria, K. B. S. Syed, "Enhancing business intelligence by means of suggestive reviews", Sci. World J., vol. 2014, Jun. 2014.
[16] M. Quigley, Encyclopedia of Information Ethics and Security, Hershey, PA, USA:IGI Global, 2007.
[17] Cambria, "Affective computing and sentiment analysis", IEEE Intell.Syst., vol. 31, no. 2, pp. 102-107, Mar./Apr. 2016.
[18] S. Poria, E. Cambria, N. Howard, G.-B. Huang, A. Hussain, "Fusing audio visual and textual clues for sentiment analysis from multimodal content", Neurocomputing, vol. 174, pp. 50-59, Jan. 2016.
[19] S. Poria, I. Chaturvedi, E. Cambria, A. Hussain, "Convolutional MKL based multimodal emotion recognition and sentiment analysis", Proc. IEEE 16th Int. Conf. Data Mining (ICDM), pp. 439-448, Dec. 2016.
[20] S. Poria, E. Cambria, A. Gelbukh, "Aspect extraction for opinion mining with a deep convolutional neural network", Knowl.-Based Syst., vol. 108, pp. 42-49, Sep. 2016.
[21] Chaturvedi, E. Cambria, S. Poria, R. Bajpai, "Bayesian deep convolution belief networks for subjectivity detection", Proc. IEEE 16th Int. Conf. Data Mining Workshops (ICDMW), pp. 916-923, Dec. 2016.
[22] S. Poria, E. Cambria, D. Hazarika, P. Vij, A deeper look into sarcastic tweets using deep convolutional neural networks, 2016, [online] Available: https://arxiv.org/abs/1610.08815.
[23] Cambria, B. White, "Jumping NLP curves: A review of natural language processing research", IEEE Comput. Intell. Mag., vol. 9, no. 2, pp. 48-57, May 2014.
[24] A. Zaslavsky, C. Perera, D. Georgakopoulos, Sensing as a service and big data, 2013, [online] Available: https://arxiv.org/abs/1301.0159.
[25] D. Denyer, D. Tranfield, "Producing a systematic review" in The Sage Handbook of Organizational Research Methods, Thousand Oaks, CA, USA:Sage, 2009.
[26] D. Tranfield, D. Denyer, P. Smart, "Towards a methodology for developing evidence-informed management knowledge by means of systematic review", Brit. J. Manage., vol. 14, no. 3, pp. 207-222, 2003.
[27] R. J. Light, D. B. Pillemer, Summing Up: The Science of Reviewing Research, Cambridge, MA, USA:Harvard Univ. Press, 1984.
[28] D. M. Rousseau, J. Manning, D. Denyer, "11 evidence in management and organizational science: Assembling the field's full weight of scientific knowledge through syntheses", Acad. Manage. Ann., vol. 2, no. 1, pp. 475-515, 2008.
[29] S. L. Newbert, "Empirical research on the resource-based view of the firm: An assessment and suggestions for future research", Strategic Manage. J., vol. 28, no. 2, pp. 121-146, 2007.
[30] Cambria, B. Schuller, Y. Xia, C. Havasi, "New avenues in opinion mining and sentiment analysis", IEEE Intell. Syst., vol. 28, no. 2, pp. 15-21, Mar. 2013.
[31] M. Taboada, J. Brooke, M. Tofiloski, K. Voll, M.Stede, "Lexicon-based methods for sentiment analysis", Comput.Linguistics, vol. 37, no. 2, pp. 267-307, 2011.
[32] A. Giachanou, F. Crestani, "Like it or not: A survey of Twitter sentiment analysis methods", ACM Comput. Surv., vol. 49, no. 2, pp. 28, 2016.
[33] C. C. Aggarwal, C. Zhai, "A survey of text classification algorithms" in Mining Text Data, Boston, MA, USA:Springer, pp. 163-222, 2012.
[34] P. Domingos, "A few useful things to know about machine learning", Commun.ACM, vol. 55, no. 10, pp. 78-87, 2012.
[35] I. H. Witten, E. Frank, M. A. Hall, C. J. Pal, Data Mining: Practical Machine Learning Tools and Techniques, San Mateo, CA, USA:Morgan Kaufmann, 2016.
[36] D. H. Wolpert, W. G. Macready, "No free lunch theorems for search", 1995.
[37] W. S. McCulloch, W. Pitts, "A logical calculus of the ideas immanent in nervous activity", Bull. Math. Biophys., vol. 5, no. 4, pp. 115-133, 1943.
[38] C. A. L. Bailer-Jones, R. Gupta, H. P. Singh, "An introduction to artificial neural networks" in Automated Data Analysis in Astronomy, New Delhi, India:Narosa Publishing House, vol. 52, pp. 18, Sep. 2001.
[39] E. Grossi, M. Buscema, "Introduction to artificial neural networks", Eur. J. Gastroenterol.Hepatol., vol. 19, pp. 1046-1054, Dec. 2007.
[40] A.-L. Boulesteix, S. Janitza, J. Kruppa, I. König, "Overview of random forest methodology and practical guidance with emphasis on computational biology and bioinformatics", Wiley Interdiscipl. Rev. Data Mining Knowl.Discovery, vol. 2, no. 6, pp. 493-507, 2012.
[41] A. Liaw, M. Wiener, "Classification and regression by randomforest", R News, vol. 2, no. 3, pp. 18-22, 2002.
[42] R. E. Schapire, Y. Freund, P. Bartlett, W. S. Lee, "Boosting the margin: A new explanation for the effectiveness of voting methods", Ann. Stat., vol. 26, no. 5, pp. 1651-1686, 1998.
[43] L. Breiman, "Bagging predictors", Mach. Learn., vol. 24, no. 2, pp. 123-140, 1996.
[44] T. Joachims, "Text categorization with support vector machines: Learning with many relevant features", Proc. Eur. Conf. Mach. Learn., vol. 98, pp. 137-142, 1998.
[45] C.-W. Hsu, C.-J. Lin, "A comparison of methods for multiclass support vector machines", IEEE Trans. Neural Netw., vol. 13, no. 2, pp. 415-425, Mar. 2002.
[46] S. N. Sivanandam, S. N. Deepa, "Genetic algorithms" in Introduction to Genetic Algorithms, New York, NY, USA:Springer-Verlag, 2007.
[47] M. Mitchell, "An introduction to genetic algorithms", Sadhana, vol. 24, no. 4, pp. 293-315, Aug. 1999.
[48] D. E. Goldberg, Genetic Algorithms in Search Optimization and Machine Learning, Reading, MA, USA:Addison-Wesley, 1989.
[49] A. Kumar, R. M. Pathak, Y. P. Gupta, "Genetic algorithm based approach for file allocation on distributed systems", Comput. Oper. Res., vol. 22, no. 1, pp. 41-54, 1995.
[50] D. D. Lewis, "Naive (Bayes) at forty: The independence assumption in information retrieval", Proc. Eur. Conf. Mach. Learn., pp. 4-15, 1998.
[51] M. Sahami, S. Dumais, D. Heckerman, E. Horvitz, "A Bayesian approach to filtering junk e-mail", Proc. Learn. Text Categorization Papers Workshop, vol. 62, pp. 98-105, 1998.
[52] J. R. Quinlan, "Induction of decision trees", Mach. Learn., vol. 1, no. 1, pp. 81-106, 1986.