Sentimental Analysis of a Sentence

Fazil Amin Mir[1], and Ravinder Pal Singh[2], and Dr. Monika Mehra[3]

[1] M.Tech, Department of Electronics and Communication Engineering, RIMT University, Mandi Gobindgarh, Punjab, India
[2] Technical Head Department of Research, Innovation & Incubation, RIMT University, Mandi Gobindgarh, Punjab, India
[3] Head of Department, Department of Electronics and Communication Engineering, RIMT University, Mandi Gobindgarh, Punjab, India

Correspondence should be addressed to Fazil Amin Mir; mirfazil88@gmail.com

ABSTRACT-Sentimental examination or assessment mining is a procedure of regular language handling which is utilized to decide the extremity of the content (i.e, regardless of whether the content is good , negative , or impartial)and furthermore used to decide sentiments or feelings (furious, cheerful, miserable and so forth) .Nostalgic investigation is frequently performed on text based information to assist organizations with observing brand and item slant in client criticism and comprehend client needs. Nostalgic investigation acts like an incredible asset for clients to separate the needful data just as to total the aggregate assessments of the surveys. Nostalgic investigation is likewise used to extricate information from web-based media stages, for example, twitter and so forth and dissect the content.

KEYWORDS- Sentiment investigation, Artificial Intelligence, Common language handling, Machine Learning, Text examination.

I. INTRODUCTION

Sentiment Analysis analyses and processes the emotions in textual data. In other words, Sentiment Analysis determines whether the given information is favourable, negative, or neutral regarding a particular topic or product.This circumstance is often described as "opinion mining. Sentiment Analysis uses NLP *(natural language processing)*, AI (*artificial intelligenc*e), and machine learning *(Machine Learning)* to process textual data. Conclusively, sentiment analysis is concerned with people's opinions on a given issue or product. Sentiment Analysis is capable of comprehending people's opinions and providing a foundation of evidence and logic on a certain product. Sentiment Analysis is capable of comprehending people's opinions and providing a foundation of evidence and logic on a certain product. Sentiment Analysis examines a large amount of textual data on the internet in order to provide future insight for the organization and to assist the public in making purchasing decisions. People buy products on an e-commerce website and leave reviews with or without using the goods. These reviews must be beneficial to people who want to buy the product or item[2]. However, when the quantity of reviews increases, customers may find it difficult to read them all. As a result, an automated method is required to provide an acceptable response. Sentiment Analysis is a task that requires an automated procedure to provide a suitable conclusion for a specific product or topic. Sentiment analysis is a technique for convertingunstructured data into structured data that categorizes text properties such as sentiment orientation *(positive, negative, or neutral)* of opinions, the topic under discussion, and the entity expressing the thoughts.

II. LEVELS OF SENTIMENT ANALYSIS

Sentiment analysis has primarily been studied at the five levels:

A. *Sentence level*

Sentence level analysis, which seeks to determine a sentence's polarity, is extensively used in micro-blogging posts. This type of analysis distinguishes between objective sentences that include concrete information and subjective sentences that contain subjective ideas and opinions[1]. Subjectivity classification and sentiment classification are the two tasks that make up sentence level analysis(as in figure 1).
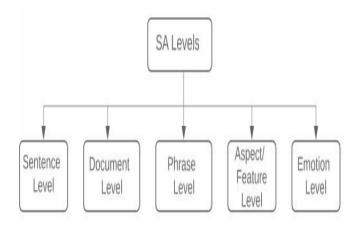

Figure 1: Levels of Sentiment Analysis

B. Document Level

Document-level analysis examines the entire "emotional" document and determines the document's overall polarity *(positive/negative)*. This type of analysis has proven to be incredibly useful for recommended systems in gaining an understanding of public sentiment. The primary sources for document level analysis are online reviews, blogs, articles, editorial sites, and forum conversations[5]. Though fine-grained tasks are limited, it is the most straightforward technique to classify a text's overall polarity. Because it focuses on a single document, document-level analysis is not applicable to documents expressing opinions about several entities.

C. PhraseLevel

The contextual polarity of a text is the focus of phrase-level analysis. The overall polarity of the phrase in which the word appears can change, but it has a positive sense. For example ,"Without the centralized heating systems, this place is excessively cold," . Now, "cool" usually connotes a positive attitude, but in this case, the polarity of the statement is negative. As a result, the polarity of the phrase is determined not only by the opinion words, but also by the word's contextual polarity[10].By first detecting if the expression is polar or neutral, and then discombobulating the polar ones, we established a methodology to automatically determine the contextual polarity for emotional expressions.

D. Aspect Level

The focus switches from linguistic constructions (documents, paragraphs, sentences, clauses, or phrases) to identifying features of an entity and associated sentiment words in feature-based, topic-based, entity-based, and target-based analysis[3]. The identification of the entity under observation is the first step in the analysis. The polarity is used to identify and categorize sentiment words connected to the entity (positive, negative and neutral). As a result, the entity's aspects are vividly found and scored according to polarity, resulting in fine-grained analysis.

E. Emotion Level

Emotion level classifies the text according to the underlying emotions such as happiness, sadness, rage, or joy. This form of investigation is more detailed and precise..At this level, sentiment analysis using emoticons (Emojis) is effective. The emoticons (emojis) and hence the emotions or sentiments are easier to categorize. The Lexicon-based technique is frequently used in emotion level analysis. In addition, the methods for determining sentiment strength can be classified depending on the rating levels - one for identifying elements of a product or service, and the other for rating a review on a global level that only analyses the polarity of the Review (positive/negative)[15].

III. SENTIMENT ANALYSIS CLASSIFICATION

The classification of sentimental analysis involves following steps (as in figure 3):

A. Preprocessing

Preprocessing is the process of handling unanalyzed data in order to extract features. It is further broken down into the following steps:

B. Tokenisation

Tokenisation is the process of turning text into tokens before translating them to vectors. It's also simple to filter out tokens that aren't needed. Raw text is broken down into words and sentences, which are referred to as tokens. These tokens aid in the comprehension of the context orthe development of a Natural Language Processing model. .Tokenization aids in interpreting the meaning of the text by evaluating the sequence of words.

C. Stop Word Elimination

Stop words are abundant in any human language. By removing these words, we are removing low-level information from our text, allowing us to focus more on the important information, such as articles.

D. Stemming

Stemming is the process of removing a word's suffix and reducing it to its underlying term. Tokens are boiled down to their simplest forms.

E. Case Normalization

When a word, statement, or document is case normalized, the entire word, statement, or document is changed to lower case or higher case[14].

F. Feature Extraction

Feature extraction is responsible for the following tasks:

G. Feature Type

Features such as term frequencies, term co-occurrences, Opinion words, OS information, and Negation Syntactic Dependencies are identified in this step[12].

H. Feature Selection

The following methods are used to select good features for classification: Information Gain, Document Frequency, Odd Ratio, and Mutual Information.
The attributes are prioritized based on the existence of terms, term frequency, and inverse document frequencies

I. Lowered Feature Size

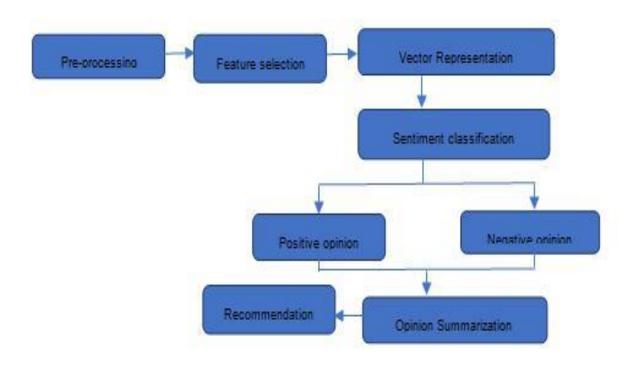The vector size is reduced to increase the classifier's performance.

Figure 2: Sentimental Analysis classification

IV. SENTIMENTAL ANALYSIS APPROACH

Sentiment analysis may be done in a variety of ways (see Figure 3). Because of the importance of sentiment analysis in today's global economy, research is still ongoing to identify better solutions.
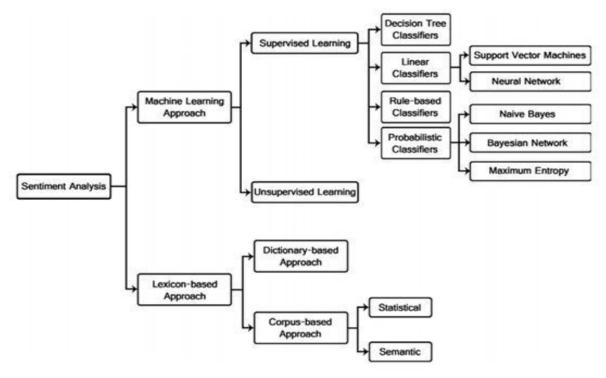

Figure 3: Sentimental Analysis Approach

A. *Machine Learning Approach*

Data processing processes that were previously time-consuming have been sped up thanks to machine learning techniques. It works by supplying an algorithm a training data set of instances and experiences. This implies that the algorithm learns automatically from the data supplied, which is referred to as the training phase of the algorithm, and then we give the system some test data to acquire the projected result, which is referred to as the testing phase of the method. Unsupervised and supervised techniques are the two sorts of techniques[6].

1) *Supervised Machine Learning*

Given a sample of training data, the algorithm predicts the outcome in real time until an acceptable level of performance is reached. The outcome is known ahead of time. When mapping categorical input data into labelled classes or regression, where input data is translated to continuous output rather than discrete output, supervised learning is typically used[11]. The overarching objective is to identify and to anticipate the correct conclusion using a certain co-relation in the input dataset.

2) *Unsupervised Machine Learning*

When dealing with real-world situations, the data will not have predefined labels, which is a requirement for Supervised machine learning methods[8]. Unsupervised machine learning models are employed in this scenario to classify data based on some feature commonality, which is then used to forecast the classes on new data. The goal is to find the natural organization or distribution of data within a set of data points without requiring labels that are expressly provided. Because no labels are provided in unsupervised learning approaches, comparing model performance is difficult. Nonetheless, it has uses in dimensionality reduction and exploratory analysis.

B. *Lexicon Based Approach*

External lexical resources that map words in the lexicon to a categorical *(neutral, negative, positive)* or numerical sentiment score are used in these approaches. This score is used by the algorithms to assess the text's overall sentiment[7].
 The approach's efficacy is based on-depending on the lexical resource's quality It does not require any training data, unlike the Machine Learning methodology, but due to the flexible and ever-changing nature of expressions and natural language, coming up with an extensive vocabulary is always a difficulty. SentiWordNet, WordNet-Affect, MPQA, and SenticNet are the most popular lexical resources *(concept-level SA)*. Lexicon-based techniques come in two varieties: dictionary-based and corpus-based. In a dictionary-based technique, the opinion words from the review text are first identified, followed by the identification of their synonyms and antonyms in the dictionary. For mapping and scoring, dictionaries such as WordNet, SentiWordNet, and SenticNet can be used[13]. The corpus-based method aids in the discovery of opinion words in a document. Orientation in relation to the situation Many approaches rely on the corpus-based approach, which

begins with a list of opinion terms and searches a vast corpus for more opinion terms.Each aspect has a strength in the global level categorization, which seeks to determine a review rate. The methods for classifying global reviews rely on machine learning techniques and only consider the polarity of the review. More linguistic features, such as negation, intensification, discourse structure, and modality, must be included in a complete classification of reviews.

V. CONCLUSION

For any form of business, sentiment analysis is critical. Several studies are being conducted on this subject, and it would be extremely beneficial if software could be developed that could recognize human sentiment or opinion that would help keep businesses, blogs, and social media sites running properly. In today's environment, we need to know how people react to every form of business. We cannot determine whether there is a need for change without people's opinions. We devised a text-based sentence analysis approach to assess the polarity of the statements, i.e. whether the evaluation offered by individuals is positive, negative, or neutral in nature. Sentiment Analysis shows how technology is being used in the real world to gain a competitive advantage and improve the overall effectiveness of business processes[9]. Sentiment Analysis will likely become more integrated into businesses in the future, particularly in marketing and advertising, where understanding customer opinions and market sentiment is extremely valuable.

REFERENCES

[1]. Mitali Desai, Mayuri Mehta, "Techniques for Sentiment Analysis of Twitter Data- A Comprehensive Survey", IEEE, pp.149-154,2016

[2]. Jatinder Kaur, "A Review paper on Twitter Sentiment Analysis Techniques", International Journal for analysis in subject field & Engineering Technology, vol.4, pp.137-141, October-2016

[3]. RishabhSoni, K. James Mathai, "Effective Sentiment Analysis of a Launched Product victimisation cluster and call Tree", International
 Journal of Innovative analysis in laptop and Communication Engineering, vol.4, pp.884-891, Jan 2016.

[4]. Hima Suresh, Dr.Gladston rule. S, "An unattended Fuzzy cluster methodology for Twitter Sentiment Analysis", IEEE,2016.

[5]. Roshan Fernandes, Dr. urban center D'Souza, "Analysis of Product Twitter knowledge although Opinion Mining", IEEE, 2016

[6]. Nidhi Grover, "A Study of assorted Fuzzy cluster Algorithms", International Journal of Engineering analysis, vol.3, pp.177-181, 2016

[7]. Govin Gaikwad, Prof. Deepali Joshi, "Multiclass Mood Classification on Twitter victimisation Lexicon lexicon and Machine Learning", IEEE,2015

[8]. Matteo Cristani, C. T. (2018). creating sentiment analysis algorithms ascendable. analysis Gate.

[9]. Sharma, S. (2018). A Context based mostly formula for Sentiment Analysis. analysis Gate Article in International Journal of process Vision and AI .

[10]. Theodoros Giannakopoulos, M. P. (2015). Visual sentiment analysis for complete observation improvement. ninth International conference on Image and Signal process and Analysis (ISPA 2015).

[11]. WalaaMedhat, A. H. (2014). Sentiment analysis algorithms and applications: A survey. own Shams Engineering Journal , 1093-1113.

[12]. Wankhede Rohit, R. J. (2018). Associate in Nursing Approach to Sentiment Analysis. JSRST , 1508-1513.

[13]. YenierCastañeda, A. C. (2014). UMCC_DLSI_Prob: A Probabilistic Automata for facet based mostly Sentiment Analysis. Proceedings of the eighth International Workshop on linguistics analysis (SemEval 2014) , 722–726.

[14]. listing (n.d.). Retrieved Apr 2019, from EnglishGrammarToday:https://dictionary.cambridge.org/grammar/englishgrammar-today/

[15]. KeltGuibon, M. O. (2016). From Emojis to Sentiment Analysis. WACAI 2016 , .