Detailed Emotion Detection and Sentiment Analysis of Social Media

Arvind Singh
Assistant Professor
Buddha institute of Management
Email Id: avirsingh111215@gmail.com

Abstract:--- Perhaps the most abundant source for human-generated text is social media. Internet users' views, comments, and critiques reveal attitudes and sentiments regarding particular subjects, goods, or services. It is practically difficult for any group of people to read through the sheer quantity of such material. As a result, social media sentiment analysis has emerged as a crucial field of study to understand social media discourse. Nevertheless, the majority of sentiment analysis methods currently in use only considers sentiment at the aggregate level, broadly categorizing sentiments as positive, neutral, or negative, and is unable to conduct extremely sophisticated sentiment analysis. This study presents an analytics tool for social media that automatically classifying text messages into sentiment groups (mixed, neutral, negative, and positive) using a social adaptive fuzzy similarity-based classification method. It can also determine the predominant emotion categories of the messages, such as joy, excitement, rage, sadness, nervousness, and gratification. Additionally, it is integrated into a comprehensive a mechanism for social media analysis that can gather, filter, categorize, and examine text data from social media platforms in addition to present an instrument panel of informative and foretelling metrics for a particular idea. The suggested approach has been created and is prepared for user licensing.

Keywords— Social media, analysis of sentiment, sentiment categorization, mining of opinions, social adaptive fuzzy coherence, and emotions.

I. INTRODUCTION

Internet users most often utilize social media sites like Facebook, Twitter, and Chinese Weibo to express their opinions or experiences about particular goods, services, or regulations. For people who see the importance of comprehending public opinion, it is a treasure trove.

Sentiment analysis on social media has many compelling applications, such as when consumers use online reviews to help them make better decisions about what to buy, when businesses seek out greater insight into market preferences to improve their products, or when politicians want to know how the public is reacting to their speeches or strategies. It should come as no surprise that Sentiment analysis is among the most often used fields in social analytics research.

Sentiment analysis's objective is to determine the degree of data sentiment [1]. Numerous social media analysis tools are now accessible to carry out this type of study, including TweetStats on Twitter, Facebook Insights on Facebook, and Stanford NLP's natural language processing tool [2]. Nevertheless, the current analysis tools concentrate on determining the sentiment at the aggregate level, with the polarity of sentiment usually falling into one of 2 groups ("positive" and "negative") or 3 groups (including "neutral") [2]. More specific and useful findings with detailed negative emotion subclasses like enmity, despair, and anxiety or positive feelings subclasses like happiness and enthusiasm will be obtained if finer-grained sentiment analysis is possible [3].

In this work, we present a novel approach to the fine-grained categorization of emotion on social media. According to the demands of the industry, the true sentiments and specific emotions were determined. A number of patents submitted in [4], [5], and [6] serve as the foundation for this approach.

The paper is organized in this manner. The current technologies for sensing are covered in Section-II. The suggested extremely fine sentiment analysis procedure is presented in Section-III. Using actual social media data, Section-IV analyzes the suggested method's performance. Finally, we wrap up this investigation in Section-V.

II. CURRENT TECHNOLOGY

There are two primary types of sentiment analysis methods: lexical-based and learning-based [7] [8]. A learning-based method uses known information to anticipate unlabeled fresh data characteristics obtained training data that has been labeled. It determines the connection between the text segment's properties in text data. The Extreme Learning Machine (ELM) [14] [15], Maximum Entropy (MaxEnt) classifier [11], Naïve Bayes (NB) classifier [9] [10], and support vector machine (SVM) [12] [13] are a few instances of learning-based techniques.

A sizable labeled training dataset is usually necessary for models employing such learning-based techniques to function well and attain respectable classification accuracy [16] [17]. However, because the diversity of the social conversation is unknown beforehand, it is challenging to ascertain what quantity of labeled dataset is suitable in the majority of social media scenarios [3] [12]. Not only would the labeling task be wasted because the training results could not be easily transferred to new datasets, but it would also be expensive or perhaps prohibitive [3] [7] [12].

Conversely, lexical-based approaches usually look for sentiment or emotion cues in a text that are listed in the lexicons that are currently in use [7] [18] [19] [20]. The impacts of the signs are then added up to identify the text's primary polarity. Lexical-based methods are simpler to implement across various datasets compared to learning-based methods, and they don't require costly labeling assignments because they don't require training. Nevertheless, the existing lexicon-based approaches have drawbacks. Making a unique lexical-basic dictionary that can be tested in many applications is challenging. As a result, the current techniques employ manually generated cleaned samples. Such information, however, differs from actual social media data, and only actual social media data may give businesses meaningful insights. The absence of extremely fine sensing capabilities [3] that offers detailed emotion detection is the other drawback, as was previously discussed in Section I. The present learning-based approaches are likewise constrained by these lexicon-based methods' drawbacks.

The study of emotion has a lengthy evolutionary history, and during the last 20 years, there has been a notable surge in emotion research activity. Shaver et al.'s work was among the first attempts to study emotion [21]. Based on the idea that various components of emotion knowledge typically form an ordered whole, Shaver et al. categorized emotions into prototypes [21]. They started their experiment by choosing a set of words and asking participants to score them according to whether or not they were emotional. They were able to create an abstract-to-concrete hierarchy of emotions by using the standard prototyping technique.

The psychologists Turner and Ortony contested the idea that fundamental emotions are psychologically archaic [22]. They suggested that every emotion has an ordered framework and is distinct and independent from the others.

The foundation of Ekman's emotion model is the claim that different gestures exist [23]. According to this approach, emotions are separate, quantifiable, and biologically different. In accordance with Shaver's approach, each emotion is a family of connected states [21].

Plutchik created the "wheel of emotions" and improved Ekman's biologically based viewpoint [24]. To depict the fundamental emotions, he created a graphic that resembled a wheel and categorized the main emotions into positive and negative categories, such as surprise and anticipation, wrath and anxiety, happiness and disgust, and faith and disgust [24] [25].

However, Alena et al. also adopted the conventional lexicon method, which benefited from and improved upon the aforementioned emotion models [26]. They assembled the words into an emotion dictionary after having professional annotation experts annotate each emotion word [26].

Although the aforementioned efforts have made significant contributions to the study and detection of emotions, only a small number of studies have used them to improve the capabilities of sensing technologies and include emotion analysis into sentiment assessment.

In order to address the limitations of the current technologies, we use the emotion research mentioned above to create employ a fine-grained emotion detection technique with fine-grained sentiment assessment technique. Furthermore, we used the approach in a real-world scenario that answers the crucial query of whether public feelings can help with efficient social sensing and policy management. Decision-makers can develop plans and improve the caliber of their goods, services, and policies with the help of the ensuing insights.

III. SUGGESTED METHODS

The goal of sentiment analysis is to identify a user's attitude or feelings toward particular subjects or domains. [2] [4] [5] [6]. The suggested approach is designed using a variety of methodologies [3]. These strategies include fuzzy logic [29], emotion theories [21] [22] [23] [24], Emotional standards for English words (ANEW) methodology for determining normative assessments of text's emotional content [28], and the linguistic inquiry and word count (LIWC) method [27].

The suggested approach gives particular consideration to the difficulties presented by real-world datasets. With linguistics interpreters made to reduce semantic ambiguity, it employs a novel social responsive fuzzy rule inference approach. In order to determine dominant valence (positive, negative, neutral, mixed) and predominant emotions (such as frustration; sorrow, nervousness, fulfillment, joy, and thrill), this is coupled with the integration and creation of multi-source lexicon.

A. Design Elements and Elements of the Suggested Approach

An algorithm for socially adaptive fuzzy inference that mimics how people perceive behaviors and feelings are expressed in online social network environments forms the basis of the suggested approach. Sentence decomposers, negation handlers, amplifiers, diminisher handlers, and other sub-modules are part of the integrated advanced language processing unit. [4], [5,] [6] Furthermore, a list of words and idioms pertaining to emotions from local languages, Internet/social media slang terms, and standard English are among the built-in linguistic lexicons that support the suggested approach. Moreover, emoticons are included. It may therefore reach the same degree of measurement precision with fewer humans input as straightforward lexicon-based and learning-based techniques because it tackles sentiment classification using more linguistics-enhanced fuzzy similarity rules and doesn't rely on any training data.

In order to create domain lexicon dictionaries, the domain lexicon extraction of information approach was used to obtain the domain knowledge [30]. Additionally, by specifying a seed lexicon, an expert user can further tune the domain knowledge to improve domain adaptation. For instance, in the area of corporate reviews, the competent user may add the word "salary lower" to the vocabulary.

In the domain of smart phones, they can delete the word "smart" from the lexicon, as in "smart watch." This can outperform straightforward lexicon-based and learning-based approaches in terms of measurement accuracy.

B. A Study of the System for Social Media

We include the suggested approach into an end-to-end social media analysis system in order to make it applicable to real-world datasets. Social data collectors, noise filters, outcomes, prediction analyzer, and sentiment and emotion analysis engine module display, and database are the six modules that make up the system. The architecture of the system is displayed in Fig. 1.

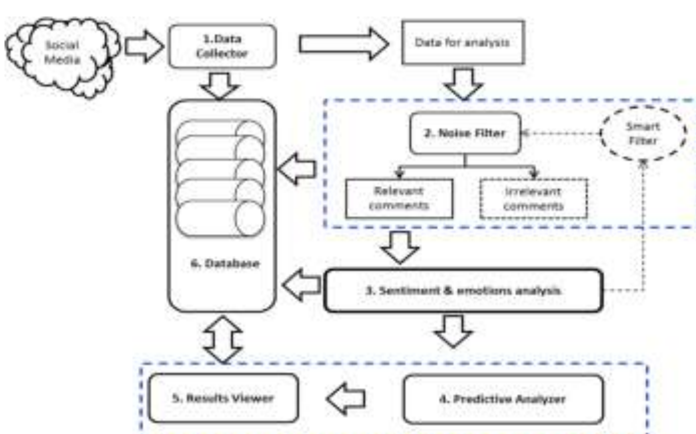

Fig. 1. Analysis System for Social Media [4] [5]

Raw data is crawled by the "Data Collector" from a variety of websites, such as blogs, Twitter, and forums. The availability of a programmatic gateway in the data sources will determine for reading data (like Twitter's keyword-based REST API and streaming API that continuously reads data), the module is made up of a number of scripts that gather data and forward it to the "Noise Filter" module for processing.

The "Noise Filter/Smart Filter" gets rid of noisy, pointless data," including commercials, useless content with no commentary, and other noises unique to a particular bit of information. To identify whether raw data is relevant or irrelevant, a "Noise Filter" preprocesses the data. An alternative sub-module called "User-defined filter" receives the pertinent data and lets the user create rules to further filter out particular information. The information sent to the "sentiment analysis engine" module is ensured to be relevant to the concept by these filters being studied further.

In order for the "Predictive Analyzer" to be utilized for critical business tasks like forecasting, monitoring, and action planning, it must do predictive analysis of significant outcomes like sales figures and reputation crises. The predictive algorithm pool and the feature set/predictor are its two primary constituents. The findings of sentiment and emotion assessment, including mixed, neutral, negative, and favorable emotions along with feelings like fear and despair, and rage, are used as new predictors or features on top of preexisting ones.

Applying the sentiment and emotion analysis engine's output in conjunction with the additional results obtained from the pool of predictive algorithms will enable time-series analysis for projections of sales, anomaly detection, and customer preferences study.

IV. EXPERIENCE IN THE REAL WORLD USING THE SOCIAL MEDIA ANALYSIS SYSTEM

Understanding emotion, especially negative emotions that demand concern from crisis managers and decision-makers, enhances the appraisal of the situation even more than knowing the valence of sentiments, which aids in gauging public reactions generally.

Sentiment categories and fine-grained feelings are the ultimate results of any text, as illustrated in Fig. 2 [2] [23] [24] [25]. Figure 2 (a) displays the sentiments, while Figure 2 (b) displays the system's output of very fine feelings.

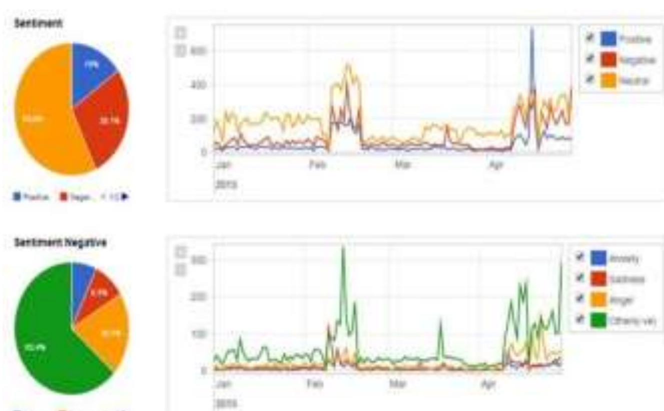

Fig. 2 shows how sentiment, whether positive or negative, can be further subdivided into more specific emotions. (a) Feelings; (b) Dissection of negative feelings into more specific feelings

Fig. 3 shows the real-time data analysis interface for testing the suggested method in real time. The tweets serve as a test case to demonstrate how data is collected, analyzed, and visualized in real time. A map is used to illustrate the data that contains geographical details.

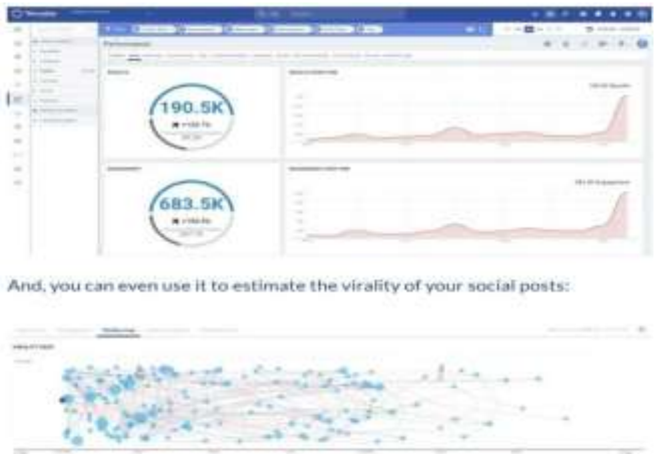

Fig. 3: A section of the social media analytics system's interface

V. CONCLUSION

This study outlines a social media analytics technique that can do in-depth sentiment and emotion analysis. This study provides fresh approaches for developing a reliable approach that handles fine-grained sensing categorization （sentiments and emotions） in textual datasets by utilizing fuzzy logic, adaptive learning capabilities, and social science principles. The suggested approach can be used in many different industries, like healthcare, business, leisure, and the public and private sectors, to help them better understand their clients, detect pertinent risks, and enhance their goods and services.

REFERENCES

[1] A. Trilla and F. Alías, ―Sentence-based sentiment analysis for expressive text-to-speech,‖ Audio, Speech, Lang. Process. IEEE Trans., vol. 21, no. 2, pp. 223–233, 2013.

[2] M. S. Neethu andR. Rajasree, Sentiment analysis in Twitter using 795 machine learning techniques, inProc.4thInt.Conf.Comput.,Commun. 796 Netw. Technol.(ICCCNT),2013,pp.15. 797

[3] S.AsurandB. A.Huberman, Predicting the future with social media, 798 inProc. IEEE/WIC/ACMInt.Conf. Web Intell. Intell. Agent Technol.(WI-799 IAT),vol.1,Aug./Sep.2010,pp.492499.

[4] Z. Wang, R. S. M. Goh, and Y. Yang, ―A method and system for sentiment classification and emotion classification,‖ Patent Cooperation Treaty (PCT) Application, PCT/SG2015/050469, 2014.

[5] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, Generative adversarial nets, in Proc. Adv. Neural Inf. Process Syst., 2014, pp. 26722680.

[6] S. Kusal, S. Patil, K. Kotecha, R. Aluvalu, and V. Varadarajan, ‗‗AI based emotion detection for textual big data: Techniques and contribution,'' Big Data Cognit. Comput., vol. 5, no. 3, p. 43, Sep. 2021.

[7] B. Yuan, Y. Liu, and H. Li, ―Sentiment classification in Chinese microblogs: Lexicon-based and learning-

based approaches,‖ Int. Proc. Econ. Dev. Res., vol. 68, pp. 1–6, 2013.

[8] T.Cassidy, Environmental Psychology: Behaviour and ExperienceinC

[9] L. Yu, W. Zhang, J. Wang, and Y. Yu, SeqGAN: Sequence generative adversarial nets with policy gradient, in Proc. 31th AAAI Conf. Artif. Intell., 2016, pp. 28522858.

[10] M. Mirza and S. Osindero, Conditional generative adversarial nets, 2014, arXiv:1411.1784. [Online].

Available: http://arxiv.org/ abs/1411.1784

[11] P.-Y. Zhang and C.-H. Li, Automatic text summarization based on sentences clustering and extraction, in Proc. 2nd IEEE Int. Conf. Comput. Sci. Inf. Technol., 2009, pp. 167170.

[12] I. Sutskever, O. Vinyals, and Q. Le, Sequence to sequence learning with neural networks, in Proc. Adv. Neural Inf. Process. Syst, 2014, pp. 31043112.

[13] P. Liu, X. Qiu, and X. Huang, Recurrent neural network for text classification with multi-task learning, in Proc. 25th Int. Joint Conf. Artif. Intell., 2016, pp. 28732879.

[14] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, ―Extreme learning machine: Theory and applications,‖ Neurocomputing, vol. 70, no. 1–3, pp. 489–501, Dec. 2006.

[15] Z. Wang and Y. Parth, ―Extreme Learning Machine for Multi-class Sentiment Classification of Tweets,‖ Proc. ELM-2015, Springer Int. Publ. 2016, vol. 1, pp. 1–11, 2016.

[16] B. Pang, L. Lee, and S. Vaithyanathan, ―Thumbs up? sentiment classification using machine learning techniques,‖ Proc. ACL-02 Conf. Empir. methods Nat. Lang. Process. Assoc. Comput. Linguist., vol. 10, pp. 79–86, 2002.