

## Sentimental analysis for user opinion- A survey

Prof. K. Khandait<sup>1</sup>, Ms. P. Girhepunje<sup>2</sup>, Ms. S. Patle<sup>3</sup>, Ms. P. Barbate<sup>4</sup>, Ms. P. Chavhan<sup>5</sup>

<sup>1</sup>Assistant Professor Prof. K. Khandait  
MIET, Bhandara, India, 441906

Ms. P. Girhepunje<sup>2</sup>, Ms. S. Patle<sup>3</sup>, Ms. P. Barbate<sup>4</sup>, Ms. P. Chavhan<sup>5</sup> Student  
MIET, Bhandara, India, 441906

**Abstract** – Opinions are shared openly and freely on social media, a process that provides a rich source for trend analyses. Sentiment analysis is one of the prominent research areas in data mining and knowledge discovery, which has proven to be an effective technique for monitoring public opinion. The big data era with a high volume of data generated by a variety of sources has provided enhanced opportunities for utilizing sentiment analysis in various domains. Our proposed sentiment analysis approach builds a binary classification model based on two feature selection techniques: an entropy-based metric and an evolutionary algorithm. The proposed feature selection model is shown to achieve significant performance improvements in both datasets, increasing classification accuracy for all utilized machine learning and text representation technique combinations. Sentiment Analysis is an important research direction of natural language processing, and it is widely used in politics, news and other fields. Word embedding play a significant role in sentiment analysis. . In the era of big data, mining the emotional tendency of comments through artificial intelligence technology is helpful for the timely understanding of network public opinion. It is so much indispensable to discover, analyze and consolidate their views for enhanced decision making. Sentiment analysis presents an effective and efficient opinion of consumers in real time which can greatly affect the decision making process for business domain. Several classifiers are trained to extract basic moods such as Joy, Trust, Fear, Surprise, Sadness, Disgust, Anger, and Anticipate from text. Finally, this paper discusses the challenges faced during sentiment and emotion analysis.

**Keywords-** Sentiments, Opinion, Social Media, Feelings

### I-INTRODUCTION

In recent years, an increasing number of people have used social media to share their opinions or to leave reviews of specific services. As of January 2018, 3 billion people were on social media, and 130 million of them were Arabs. Due to the influence of social media content on government, diplomacy, and business, sentiment analysis is required for social media research. Sentiment analysis is the computational assessment of people's opinions towards policies, products, services, or news [1]. Sentiment analysis, which provides automated extraction of opinions or feelings, is one of the techniques that play an essential role in decision making processes. It is also known as opinion mining, since it aims to extract subjective opinion from a piece of text. Sentiment analysis has been gaining more attention recently, as it is a significant element of many real-world applications, including recommendation systems, analysis of product reviews, terrorist organization tracking, detection and

analysis of critical events, real-time observation of public opinion, finance and healthcare systems [2]. Sentiment analysis faces challenges due to the existence of slang words, spelling mistakes and ironic remarks in documents. Sentiment Analysis is a technology that automatically extracts sentiment information from unstructured texts.

Sentiment Analysis is applied to many fields such as natural language processing (NLP), data mining and machine learning. Word vector representation is a key step in Sentiment Analysis. Nowadays, the widely used word embedding technology is Word2Vec and Glove, which based on distributed representation. The idea is that words with similar contexts have similar vector representations. It is very useful for many tasks which related to semantic similarity because it can capture lots of contextual features to represent texts [2-3].

The main contributions of this paper are as follows: proposing the sentiment concept to achieve the accurate embedding of sentiment information and provide more precise semantics and sentiment representations for words,

(2) constructing a sentiment intensity lexicon containing single-semantics and multi-semantics sentiment words through the multi-semantics integration of six representative sentiment intensity lexicons, to provide more accurate sentiment information for words with different semantics, (3) Refined-Word2Vec and Refined-Glove we improved are averaged to obtain Refined Global Word Embedding [4].

One important challenge in aspect-level sentiment analysis is how to model the semantic relationship between aspect terms and sentences. Many methods have been proposed to solve the problem. The traditional NLP approach is a method based on sentimental knowledge, which relies on the existing sentiment lexicon or domain lexicon in subjective text and the combined evaluation unit with sentiment polarity. Some of them use the features to train a sentiment classifier, such as Support Vector Machines [9]. Social Networking websites and micro blogging websites in today's world has become the biggest web destinations for people to communicate with each other, to express their thoughts about products or movies, share their daily experience and communicate their opinion about real time and upcoming events, such as sports or political elections, etc. [11-12].

Unlike traditional social media, such as newspapers, online social media contains a large amount of multimodal data that can provide a considerably large number of clues for estimating sentiments when compared with that provided by words alone. Further, the tweets posted on social media usually present abundant contextual information, such as the timelines of the users and the comments of other people. This contextual information is helpful for conducting sentiment analysis because it can comprehensively characterize the contextual attributes of tweet streams [13-14]. In contrast, Emotion detection is a means of identifying distinct human emotion types such as furious, cheerful, or depressed. People are using social media to communicate their feelings since Internet services have improved. On social media, people freely express their feelings, arguments, opinions on wide range of topics [15].

In this review paper, Sect. 1, introduces sentiment analysis and its various levels, emotion detection, and psychological models. Section 2 discusses what is sentiment analysis, multiple steps involved in sentiment and emotion analysis. Section 3 discusses motivation and contribution addresses multiple challenges faced by researchers during sentiment and emotion analysis. Section 4 is of survey of sentimental analysis and about proposed method. Finally; Sect. 5 concludes the work.

## II-SENTIMENT ANALYSIS

### A. What is sentiment analysis?

Sentiments are feelings that reflect attitude, emotions and opinion. Sentiment Analysis can be classified as computation study of opinions, evaluations, attitudes, subjectivity, and views expressed in text. An emotive valuation of a condition is a general assessment of that state that can be either positive or negative depending on bodily or mental responses [1-2]. The importance of sentiments is not new to business. To a business process customer responses are indirect motivators of purchase behavior. Sentiment Analysis addresses the issues by systematically collecting and analyzing online sentiments from variety of sources and from a very huge sample of customer in real time [4].

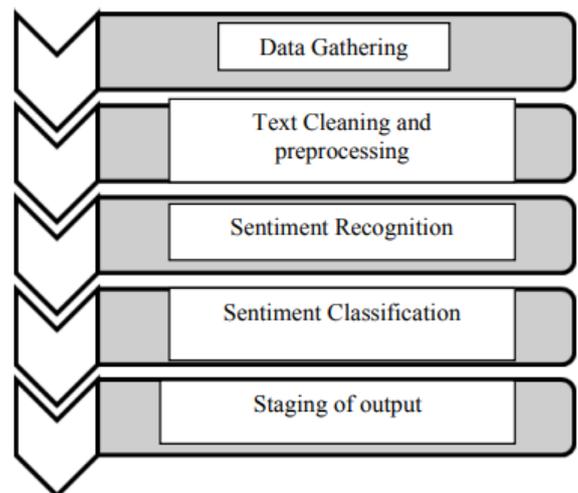


Fig-1 Sentiment Analysis Process

Sentiment analysis is a methodical examination of online expressions. Precisely, sentiment analysis emphasizes on weighing attitudes and opinions on a topic of interest using machine learning techniques. The description of sentiment analysis in data mining can be pronounced from two standpoints: functional and operational. The functional aspects focus on practical uses of the technique. For illustration, sentiment analysis is described as a process that categorizes a body of textual information to determine feelings, attitude and emotions towards a particular issue or object [8]. The definition points to the way sentiment analysis works and describes the outcome of polar classification. Another facet of sentiment analysis in data mining focuses on the operations of the technique as a sub-field of computational linguistics. Sentiment analysis is

described as automated subjectivity analysis comparable to opinion mining and assessment withdrawal which focuses on extracting and classifying texts with machine language and computer programming [7-8]. Despite the differences in both perspectives, the general narrative is the same. In other words, sentiment analysis is a data mining technique that practices natural language processing, computational linguistic and text analytics to recognize and excerpt content of interest from a body of textual data.

### III-MOTIVATION AND CONTRIBUTION

Sentiment analysis becomes difficult when conducted on Arabic social media texts due to the limitations of the existing natural language processing tools and resources available for the Arabic language. The main challenges of sentiment analysis and opinion mining arise from the use of colloquial words, merging words, repeated letters, and spelling errors. The objective of this study is to find suitable techniques and models to automatically determine the sentiments of tweets posted in particular domains (news, sports, and politics). It specifically aims to develop a classifier that can be used to automatically classify Gulf dialect comments as positive, negative, or neutral [1]. In scientific literature, several works, such as have applied Collaborative Filtering (CF) to the recommendation problem. The CF-based approaches intend to generate recommendations based on the similarity in actions and preferences of users. A hybrid context-aware recommendation framework that performs recommendations based on mining and analysis of users' textual feedbacks. We attempt to address some of the inherent issues of recommendation systems [7]. The contributions of this study are summarized as follows: First, we propose the stacked gated recurrent unit (SGRU) and stacked bidirectional gated recurrent unit (SBI-GRU) models and compare them with the support vector machine (SVM) and the Arabic bidirectional encoder representations from transformers (AraBERT) developed by Antoun et al. to investigate their performance in analyzing Arabic using effective preprocessing techniques. Second, we implement an ensemble approach using multiple models (SGRU, SBI-GRU, AraBERT) to generate the best suited sentiment analysis model for the Arabic language. Based on the proposed model, increasing the depth of the network provides an alternate solution that requires fewer neurons and is faster to train. Ultimately, adding depth is a type of representational optimization [9].

### IV-LITERATURE SURVEY

In this section, we present recent studies that use the notion of stacking. We present the first study that used the RNN stacking method in the sentiment analysis field. In addition, we present to compare the performance of stacked bidirectional long short-term memory (SBI-LSTM) with our proposed models [1]. Sentiment analysis has been a popular research topic due to its wide scope of applications, ranging from recommendation systems to finance. Although sentiment analysis has been extensively studied in the literature, new studies continue to emerge as available data continually grow and become more complex. It is crucial to select the optimal feature subset for sentiment analysis to achieve high performance. Therefore, feature selection is an indispensable preprocessing step, alleviating the burden caused by the high-dimensional data [2]. Aspect-based sentiment analysis is a special type of sentiment analysis that aims to detect opinion toward fine-grained aspects. During the first step, Aspect Category Detection (ACD), a system aims to detect a set of the pre-defined aspect categories that are described in the given text. During the second step, Aspect Category Polarity (ACP), a system aims to classify a text into one of sentiment polarity labels (i.e., Positive, Negative, Neutral, etc.) given a pair of text and aspect categories [5]. Recently, many scientific literatures have applied text mining approaches to improve performance of recommendation systems by augmenting numerical rating with sentiment analysis of textual reviews [7]. Machine learning is also a mainstream sentiment classification method. These methods mainly involve text representation and feature extraction, such as bag-of-words models and sentiment lexicon features, then training a sentiment classifier, such as support vector machine (SVM) and logistic regression [9]. Sentiment Analysis is a major tool for a machine to understand human psychology. This technology is being studied extensively in order to implement in the fields where humans were needed to detect sentiment or feeling. It is of major importance in assistant chatbots and combined with speech recognition technology, it can also be used to substitute humans in call centers [15].

#### A. Word Embedding

Word embeddings focus on learning distributed vector representations of words by leveraging the contextual information in large corpora using neural network

architectures. The dependency-based word embeddings addressed the limitation of linear contexts by introducing syntactic contexts derived from a dependency parser, while the Glove model addressed the limitation of local contexts by accounting for global word-word co-occurrence statistics. In addition to different types of contextual information, character-level sub words are also useful information for learning word embeddings. Word embedding learned from a particular context tend to generate similar vector representations for words with similar contexts. This property works well for semantic-oriented applications, but it is problematic for sentiment analysis because words with similar vector representations due to similar contexts may have an opposite sentiment polarity, as in the example of happy-sad mentioned in and good-bad in Such examples are not rare [6]. The future scope is to rate business based on different features, to use better and more dataset to train and to detect sarcasm. Some limitations of the project include accuracy and less diversity of categorization. The quality of the training data set and testing data set is a major aspect in determining the efficiency of any such algorithms [15].

### B. Sentiment Embeddings

A common characteristic of existing sentiment embeddings is the use of positive and negative polarity labels provided by labeled corpora to guide the learning process through an objective function. Recent work has focused on developing neural network models to learn sentiment embeddings from tweets automatically collected by the distant supervision method for Twitter sentiment classification. [6]

## V-CONCLUSION

In this paper, first we proposed extraction of opinions or feelings, is one of the techniques that play an essential role in decision making processes. Sentiment Analysis has been applied in many fields, and the effect of sentiment analysis depends more on the quality of word embedding, so it is necessary to study word embedding methods in Sentiment Analysis tasks.

Second, we proposed the importance of sentiments is not new to business. To a business process customer responses are indirect motivators of purchase behavior.

Third, we proposed that Sentiment analysis becomes difficult when conducted on Arabic social media texts due to the limitations of the existing natural language processing tools and resources available for the Arabic language.

Our future work involves the use of multi-objective optimization to find out an optimal solution from multiple conflicting objectives, thus improving the overall quality for recommendations

## REFERENCES

- [1] Asam Al Wazrah and Sarah Alhumoud, "Sentiment Analysis Using Stacked Gated Recurrent Unit for Arabic Tweets", Saudi Arabia, September 2021.
- [2] Ayca Deniz 1, (Member, IEEE), Merih Angin 2, and Pelin Angin 1, (Member, IEEE), "Evolutionary Multiobjective Feature Selection for Sentiment Analysis", Ankara, İstanbul, Turkey, October 2021.
- [3] Yabing Wang, Guimin, Huang, Jun Li, Hui Li, Ya Zhou, and Hua Jiang, "Refined Global Word Embedding Based on Sentiment Concept for Sentiment Analysis", in IEEE, Guilin, China, March 2021.
- [4] Guixian Xu 1, Yueting Meng 1, Xiaoyu Qiu 2, Ziheng Yu 1, and Xu Wu 1, "Sentiment Analysis of Comment Texts Based on BiLSTM", April 2019, Beijing, Jinan, China.
- [5] Heereen Shim 1, Dietwig Lowet 2, Stijn Luca 3, and Bart Vanrumste 1, (Senior Member, IEEE), "LETS: A Label-Efficient Training Scheme for Aspect-Based Sentiment Analysis by Using a Pre-Trained Language Model", Leuven, Leuven, August 2021.
- [6] Liang-Chih Yu, Member, IEEE, Jin Wang, K. Robert Lai, and Xuejie Zhang, "Refining Word Embedding Using Intensity Scores for Sentiment Analysis", March 2018.
- [7] Rizwana Irfan 1, Osman Khalid 1, Muhammad Usman Shahid Khan 1, Faisal Rehman 1, Atta ur Rehman Khan 2, (Senior Member, IEEE), and Raheel Nawaz, "SocialRec: A Context-Aware Recommendation Framework with Explicit Sentiment Analysis", Islamabad, Pakistan, Sohar, Oman, August 2019.
- [8] Saumya Chaturvedi Ph.D. scholar AKTU Lucknow, INDIA 1, Vimal Mishra Director IERT Allahabad, INDIA 2, Nitin Mishra, "Sentiment Analysis Using Machine Learning for Business Intelligence", 2017.
- [9] Zhiying Ren 1,3, Guangping Zeng 1,3, Liu Chen 1,3, Qingchuan Zhang 2, Chunguang Zhang 1,3, and Dingqi Pan 1,3, "A Lexicon-Enhanced Attention Network for Aspect-Level Sentiment Analysis", Beijing 100083, China, May 2020.

[10] Widodo Budiharto and Meiliana Meiliana, “Prediction and analysis of Indonesia Presidential election from Twitter using sentiment analysis”, Computer Science Department, School of Computer Science, Bina Nusantara University, Jakarta 11480, Indonesia, 2018.

[11] Lutfun Nahar Assistant Professor Dept. of Computer Science 1, Zinnia Sultana Assistant Professor Dept. of Computer Science 2, Farzana Tasnim B.Sc. in Computer Science 3, Afifa Chowdhury B.Sc. in Computer Science 4, “Sentiment Analysis & Emotion Extraction: A Review of Research Paradigm”, May 2019.

[12] Ms. Farha Nausheen Dept. of IT MJCET Hyderabad, India 1, Ms. Sayyada Hajera Begum Dept. of IT MJCET Hyderabad, India 2, “Sentiment Analysis to Predict Election Results Using Python”, 2018.

[13] Bo Liu, Shijiao Tang, Xiangguo Sun, Qiaoyun Chen, Jiuxin Cao, Junzhou Luo, and Shanshan Zhao, “Context-Aware Social Media User Sentiment Analysis”, August 2020.

[14] Pansy Nandwani 1, Rupali Verma1, “A review on sentiment analysis and emotion detection from text”, July 2021.

[15] Hemalatha S 1, Ramathmika2 Department of Computer Science, “Sentiment Analysis of Yelp Reviews by Machine Learning”, Manipal, Karnataka, India, 2019.