Sentiment Analysis: A Review

Sachin J A, Suloni Praveen, Swathi Meghana K R and Yuvika Balaji
Department of Computer Science and Engineering,Vidyavardhaka College of Engineering

Abstract—**Over the past ten years, sentiment analysis and opinion mining have seen significant growth. This field of study makes a shot to determine, among other things, what people think and feel about something or someone. Natural language processing methods and machine learning algorithms are employed during this analysis. Sentiment analysis is the systematic identification, extraction, quantification, and study of affective states and subjective data. It uses linguistic communication processing, text analysis, linguistics, and biometrics.**

Keywords—Opinion mining, sentiment analysis, paper reviews, hybrid methods, NLP, machine learning, lexical-based.

I. INTRODUCTION

Data Mining is the field where a large amount of data gets processed and outputs the information that is relevant. Basically, it consists of filtering and analysis methods. The extracted information is useful for other purposes in life like solving some problem by the help of patterns detected during data mining. Opinion Mining (OM) is the process of deriving knowledge from large data set. Data Mining is a very important and crucial step in OM. Data Mining consists of Data Pre-processing before it and Data Post-processing after it [1].

Finding out what other people think has always been a significant component of our information-gathering activity. People may actively utilize information technology to seek out and comprehend the ideas of others, which has led to new opportunities and challenges as these opinion-rich resources, such as personal blogs and online review sites, become more widely available and more well-liked. Thus, at least in part, the unexpected flurry of effort in the field of opinion mining and sentiment analysis—which deals with the computational handling of opinion, sentiment, and subjectivity in text—has been a direct reaction to the increase in interest in new systems that deal with opinions as a first-class object.

In contrast to the problems that more conventional fact-based analysis currently faces, the emphasis is on approaches that try to solve the new problems brought on by sentiment-aware applications. The survey lists the many uses, examines general difficulties, and talks about classification, extraction, and summarization. Last but not least, it goes beyond only focusing on the technological challenges, giving careful consideration to the wider ramifications of the growth of opinion-oriented information-access services, including concerns about privacy, manipulation risk, and whether or not reviews may actually have a discernible impact on the economy.

A description of the resources that are available, benchmark data sets and assessment campaigns are also offered to aid future study. Anyone with an interest in opinion-oriented information-seeking systems will find Opinion Mining and Sentiment Analysis to be the first thorough assessment of this dynamic and significant research subject.

NLP stands for Natural Language Processing, which is a branch of artificial intelligence (AI). It enables robots to analyze and understand human speech, allowing them to perform repetitive activities without human intervention. Machine translation, summarization, ticket categorization, and spell checking are some examples for NLP.

Consider sentiment analysis, which uses natural language processing to analyze text for emotions. One of the most popular NLP jobs is the classification task, which is often used by companies to automatically detect brand sentiment on social media. Brands can monitor overall customer satisfaction or identify significant customer issues that require immediate attention by analyzing these interactions made by the customer with the brand.

A. *Natural Language Understanding (NLU)*

Natural Language Understanding (NLU), which uses text and speech syntactic and semantic analysis to interpret sentences, is a subset of natural language understanding. The grammatical form of a sentence is referred to as its syntax, while its intended meaning is referred to as its semantics. Additionally, NLU creates a pertinent ontology, which is a data structure that describes the connections between words and sentences. The combination of these analyses is necessary for a machine to comprehend the intended meaning of various texts, even if humans naturally do this in conversation. The subtleties of language are vividly illustrated by our capacity to discriminate between homonyms and homophones.

B. *Natural Language Generation (NLG)*

Another division of natural language processing is natural language generation. Natural language generation, in contrast to natural language understanding, focuses on enabling computers to write. NLG is the process of generating a written answer in human language from some data input. Through text-to-speech services, this text can also be turned into a spoken format.

As part of its capabilities, NLG also includes text summarization tools that produce accurate summaries from input materials.
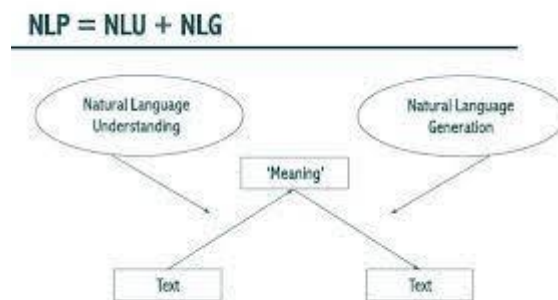
Similar to NLU, NLG applications must take into account morphology, lexicons, syntax, and semantics-based language norms when deciding how to phrase responses. Three steps are taken to address this:

Text Planning: Planning the text entails formulating and logically arranging the overall material.

Sentence Planning: Planning sentences involves organizing the content into paragraphs and sentences, paying attention to grammar and text flow, and, when necessary, using pronouns or conjunctions.

Realization: At this stage, grammar accuracy is taken into consideration, and regulations regarding punctuation and verb conjugations are strictly adhered to. The past tense of the word "eat," for instance, is "ate," not "eated".

III. SENTIMENT ANALYIS

Sentiment Analysis (SA) has emerged as one of the key computational disciplines supporting Natural Language Processing (NLP). In general, SA deals with the data mining of attitudes, opinions, or recommendations from a group of people regarding a given topic. The document level comments for some applications are also received.

Sentiment analysis is becoming increasingly popular in domains like politics (to forecast election results), marketing, and even business (to analyze online sentiments in the social media for stock market prediction).

The intricacies of each review can be understood by a sentiment classifier with natural language processing capabilities, which can then automatically identify the first review as negative and the second as positive. The sentiment of the sentences can be neutral, negative, or positive depending on the circumstances.

Think about a rise in negative comments about your brand on social media. Sentiment analysis tools would be able to see this immediately away, giving you the opportunity to act appropriately before the problem worsens.

Aspect or object-based extraction, which is the main emphasis of SA, is one of the most crucial activities. Therefore, sentences are extracted, categorized, and used in the supplied analysis depending on whether they are subjective or objective.
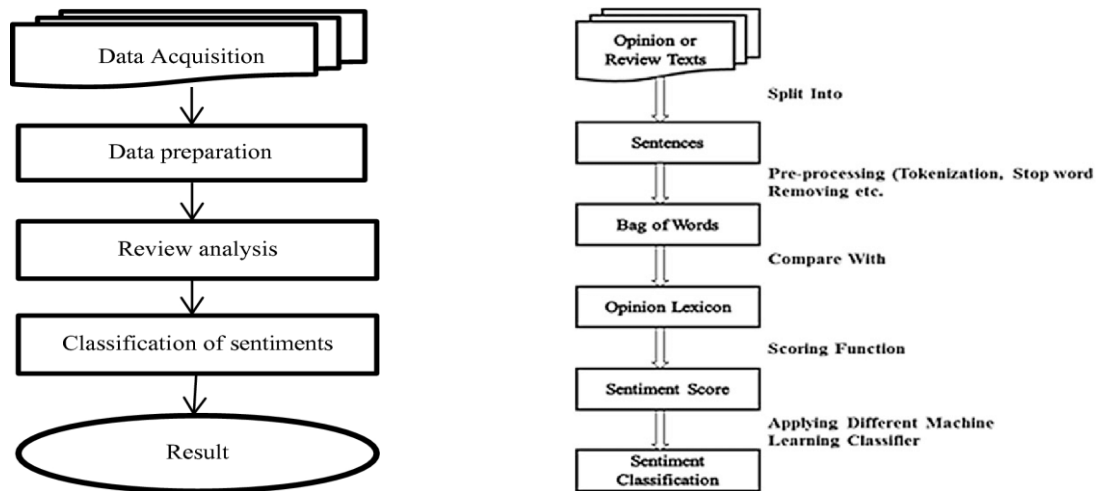
The main task in SA is subjectivity classification, which involves categorizing sentences as objective or subjective. Sentiment analysis is applied to the subjective sentences. The objective sentences that have been identified are discarded. This is done because objective sentences contain factual information that contributes nothing to the overall SA process. Subjective sentences, on the other hand, are very important because they contain people's opinions, views, claims, thoughts, observations, and personal beliefs. These sentences are divided into three categories: positive, negative, and neutral. Neutral opinions are those that cannot be classified as positive or negative.

For example, "I think this color is very bright." In this example, the opinion holder is making an observation or giving an opinion about a certain color. The opinion holder claims that the color is very bright but once this sentence is sentimentally analyzed we cannot determine if the opinion holder is giving a positive sentiment (e.g., the color is pretty) or a negative sentiment (e.g., the color is harsh and blinding). Therefore this sentence is considered as neutral.

CONCLUSION

As a result, it is now a crucial research issue to apply sentiment analysis or opinion mining to mine a variety of unstructured data. Sentiment analysis results in better product and service development as well as effective business management. The review paper reviewed the relevant work that had been done up to that point in the Sentiment Analysis discipline. According to the research presented in the articles, opinion mining or the enhancement of sentiment classification algorithms is still a relatively unexplored field of study.

Social media platforms like micro blogs, forums, blogs, and news sources provide a wealth of information on people's attitudes and feelings toward a certain topic or item. But there is still a requirement for extensive deeper study and investigation when employing these social media in networking or micro blogging sites as a source of data for OM or SA tasks. The key difficulties include negations, generalizing language, and classification difficulty. Additionally, sentiment analysis researchers have recently become interested in natural language processing tools, however, these tools still require refinement. In contrast, several algorithms that have been employed in OM or SA produce good outcomes.

REFERENCES

[1] L. Qu, G. Ifrim and G. Weikum." The bag-of-opinions method for review rating prediction from sparse text patterns". In Proceedings of the 23rd International Conference on Computational Linguistics, pp. 913-921.
[2] Sentiment Analysis or Opinion Mining: A Review
October 2017 International Journal on Advanced Science Engineering and Information Technology by Saidah Saad from Universiti Kebangsaan Malaysia
[3] Aarti Sharma, Rahul Sharma,Vivek Kr. Sharma,Vishal Shrivatava, "Application of Data Mining – A Survey Paper"/International Journal of Computer Science and Information Technologies,ISSN:0975-9646, Vol. 5 (2),pp. 2014, 2023-2025,year 2014.
[4] P.Lissy,P. Sudha "Business Intelligence in Social Network-Survey"/ International Journal of Modern Computer Science (IJMCS) ISSN: 2320-7868, Vol 4, Issue 4,pp. 148-151, Aug 2016
[5] M. M. Altawaier & S. Tiun "Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis". International Journal on Advanced Science, Engineering and Information Technology, Vol. 6 (2016) No. 6, pages: 1067-1073
[6] K. Dave, S. Lawrence and D. M. Pennock. "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews". In Proceedings of the 12th international conference on World Wide Web, pp. 519-528.
[7] A. Balahur, R. Mihalcea and A. Montoyo "Computational approaches to subjectivity and sentiment analysis": Present and envisaged methods and applications. Computer Speech & Language, 28(1): 1-6, 2014.
[8] E. Boiy and M. F. Moens "A machine learning approach to sentiment analysis in multilingual Web texts". Information retrieval, 12(5): 526-558, 2009.
[9] NLP vs. NLU vs. NLG: the differences between three natural language processing concepts By Eda Kavlakoglu | 5 minute read | November 12, 2020
[10] M. Ghiassi, J. Skinner and D. Zimbra. "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network". Expert Systems with applications, 40(16): 6266-6282, 2013