

A SURVEY ON SENTIMENT ANALYSIS

Arti Pandey

Assistant Professor, Ajay Kumar Garg Engineering College, Ghaziabad, U.P. India
pandeyarti@akgec.ac.in

Abstract: Sentiment means feelings, emotions or opinion that is held or expressed. Sentiment analysis is an application of Artificial Intelligence and Machine Learning. It is textual conceptual mining that discovers and extracts subjective information from sources. It helps in determining or assessing the positive, negative, and neutral aspects of a text. In the discipline of computer science, it is the fastest-growing research area. This paper shows comparison between the various methods and algorithm used in the field of sentiment analysis.

Keywords: Machine Learning, sentiment analysis, mining, Artificial Intelligence.

I. INTRODUCTION

The Internet has revolutionised the way people share their thoughts and opinions today. Blog entries, product review websites, internet forums, social media, and other social networking platforms are commonly used to express opinions and perspectives. Millions of individuals use social media sites such as Twitter, Facebook, Google Plus, and others to express their thoughts, emotions, and share information about their daily life. Consumers obtain an interactive media through these online platforms, where they may inform and influence others through forums. Tweets, blog posts, comments, status updates, reviews, and other forms of social media generate a tremendous volume of sentiment-rich data. Businesses benefit from social media because it gives them a platform to connect with their customers for advertising or to obtain feedback on the services they provide. People rely heavily on other users' generated content (feedback) while making decisions about online purchases. For example, if someone wants to use a service or buy a product, they would first read online reviews and discuss it on social media before making a decision. The volume of content published by users on social media sites is far too massive for a common person to examine [1]. As a result, there is a need to automate this process, and a variety of sentiment analysis approaches are frequently used. Sentiment analysis (SA) primarily informs the user whether the product's information is suitable or not before to purchase. Marketers and businesses use this data to better understand their services or goods so that they may be tailored to the needs of their customers, with minor tweaks if necessary.

Sentiment classification can be performed in 3 levels. These are:

- Document level
- Sentence level
- Feature level

Sentiment analysis only employs one object and gets a single opinion from a single opinion holder at the document and phrase level. In many circumstances, however, such assumptions are incorrect. Extracting sentiment for a complete document/blog is inefficient compared to extracting sentiment for each subject in a single line [2].

II. APPLICATIONS

There are the following applications of sentiment analysis:

1. **Social Media Monitoring:** Consumer stories and opinion data abound on social media. However, social media posts are littered with abbreviations, acronyms, and emoticons. Many social analytics tools are unable to deal with these discrepancies. The sheer volume is also an issue. Some social media monitoring technologies have a hard time scaling up. Meanwhile, data analysts squander time manually parsing mountains of social data.
2. **Customer Experience Management:** Customer feedback play an important role behind the success of any business. Sentiment analysis is helpful to classify between positive and negative feedback. It will clear the picture in front of company or organization.
3. **Robotic Process Automation:** Customers are increasingly demanding larger, transformational RPA and deeper analytics integrations from robotic process automation (RPA) vendors. However, many companies are trailing behind in terms of offering popular text analytics use cases. Others have better text analytics but lack the ability to handle "unstructured document use cases" such as PDFs. Others are experiencing difficulty integrating text analytics and natural language processing into their overall environment [2].
4. **Market and competitor research:** It's another example of sentiment analysis in action. It is beneficial in studying the most recent product demands in the industry, as well as product reviews. It also aids in determining who your competitors are mentioning and how your marketing

efforts compare. From the ground up, get a thorough view of your and your competitors' consumer bases. Analyzing the content of your competitors can help you figure out what works with the general public that you hadn't thought of [3]. You'll be able to evaluate your own strengths and weaknesses, as well as how they compare to those of your rivals.

5. Brand monitoring and management: Bad reviews build quickly on the internet, and the longer you wait to respond, the worse the problem will get. Sentiment analysis technology will be used to quickly detect negative brand remarks. Not only that, but you can monitor your brand's image and reputation over time or at any one point in time, allowing you to monitor your progress. You may turn information about your brand found in news stories, blogs, forums, or social media into relevant data and analytics. Machine learning may also be used to spot trends and predict outcomes, allowing you to stay ahead of the game and shift from reactive to proactive mode.

III. APPROACHES OF SENTIMENT ANALYSIS

There are the following approaches in sentiment analysis:

A. Machine Learning

For sentiment categorization of categorised text into positive, negative, or neutral categories, machine learning techniques are the most useful strategies. The following are examples of supervised machine learning techniques:

- i. Support Vector Machine (SVM): For classification and regression problems, it is a popular Supervised Learning technique. However, it is mostly utilised to address classification problems in Machine Learning. The goal of this algorithm is to find the best line or decision boundary for classifying n-dimensional space into classes so that subsequent data points can be easily classified. A hyperplane is the best possible decision boundary. The extreme points/vectors that assist create the hyperplane are chosen using SVM. Face detection, image classification, text categorization, and other tasks can all benefit from this method.
- ii. Naïve Bayes: It is a basic classifier that's one of the most often utilised in the field of text classification [4]. The Bayes' Theorem determines the probability of a subsequent event based on the probability of a previous event. The following equation mathematically shows Bayes' theorem:
$$P(A|B) = P(B|A) P(A)/P(B)$$
where A and B are events and $P(B) \neq 0$.
- iii. Bayesian Network: It is also known as a Bayes network or belief network. It is a probabilistic graphical model that uses a directed acyclic graph to represent a set

of variables and their conditional relationships. This network is used to represent the joint probability distribution.

- iv. Maximum Entropy: Maximum Entropy classification is a natural language processing technique that has shown to be effective in a variety of applications. Because it makes no assumptions about the relationships between characteristics, it may perform better when conditional independence assumptions are violated.

B. Deep Learning

- i. Deep Neural Network: A artificial neural network having multiple hidden layers between the input and output layers is known as a deep neural network (DNN) [8]. DNNs can represent complex non-linear relationships in the same way as shallow ANNs can. A neural network's basic goal is to take a collection of inputs, conduct increasingly sophisticated computations on them, and then output results to solve real-world issues like categorization.
- ii. Recurrent Neural Network: These networks are extremely powerful because they combine two characteristics: First is, non-linear dynamics that allow them to update their hidden state in complex ways. Second property states that the distributed hidden state that allows them to store a lot of knowledge about the past efficiently [5].
- iii. Convolutional Neural Network: CNN is a type of deep learning and neural network that is frequently used to evaluate visual data. Convolutional neural networks use minimum preprocessing compared to other image classification algorithms, which means the network learns the filters that are generally hand-engineered in other systems [6,7]. CNN consists of pooling layer and fully connected layer.

IV. CONCLUSION

Sentiment analysis has been a prominent study topic, and many great studies have been conducted in this area. In this survey, a series of the state-of-the-art literatures have been reviewed. Many researchers used different types of models and techniques in the paper to improve the accuracy. This field is still open for the future advancement. Future sentiment analysis and opinion mining techniques will require more common-sense approaches that are inspired by human cognitive processes, resulting in more honest, accurate, and psychologically relevant outcomes.

REFERENCES

- [1] A. R. Pathak, M. Pandey, and S. Rautaray, "Topic-level sentiment analysis of social media data using deep learning," *Applied Soft Computing*, vol. 108, pp. 1-17, 107440, 2021. Elsevier. <https://doi.org/10.1016/j.asoc.2021.107440>.

- [2] A. Naresh and P. Venkata Krishna, "An efficient approach for sentiment analysis using machine learning algorithm," *Evolutionary Intelligence*, vol. 14, pp. 725-731, 2021. Springer. <https://doi.org/10.1007/s12065-020-00429-1>
- [3] LC Chen, CM Lee, MY Chen, "Exploration of social media for sentiment analysis using deep learning", *Soft Computing* vol. 24, pp. 8187–8197, 2020. Springer. <https://doi.org/10.1007/s00500-019-04402-8>
- [4] N. Zainuddin, A. Selamat and R. Ibrahim, "Hybrid sentiment classification on twitter aspect-based sentiment analysis", *Applied Intelligence*, Volume 48, pp :1218–1232,2018. Springer. <https://doi.org/10.1007/s10489-017-1098->
- [4] J. Singh, G. Singh, R. Singh, and P. Singh, "Morphological evaluation and sentiment analysis of Punjabi text using deep learning classification", *Journal of King Saud University - Computer and Information Sciences*, Volume 33, Issue 5, June 2021, Pages 508-517
- [6] Ruz, G. A., Henríquez, P. A., & Mascareño, A. (2020). Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. *Future Generation Computer Systems*.
- [7] Abd El-Jawad, M. H., Hodhod, R., & Omar, Y. M. K. (2018). Sentiment Analysis of Social Media Networks Using Machine Learning. 2018 14th International Computer Engineering Conference (ICENCO).
- [8] Bhuiyan, H., Ara, J., Bardhan, R., & Islam, M. R. (2017). Retrieving YouTube video by sentiment analysis on user comment. 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). doi:10.1109/icsipa.2017.8120658

ABOUT THE AUTHOR



Arti Pandey received the Master's degree in Information Technology from Banasthali Vidyapith, Rajasthan. She is currently work as an Assistant Professor in Ajay Kumar Garg Engineering College, Ghaziabad. Her research interest includes Information Retrieval, Artificial Intelligence, Cloud Computing and Big data.