Sentiment Analysis on Social Media

Abstract

Interpretation of an opinion has been a challenging task for the ages since it is not only dependent on speaker or writer's appeal to subject, but standpoint and perception of the audience also matters. Today with social media giving an opportunity to masses, to have their own voice; marketers, politicians and debaters have an instant chance to find out trivial question: "How do people feel?". Number of mentions and comments might reveal popularity whereas a thorough sentiment analysis might enable its performer to track online conversations to withdraw quantifiable insights on perception. In a world where every action has a reaction, gaining an information regarding perception is critically important for people / brands / authorities especially in a democratic environment. This paper provides a deep dive analysis into complex process of sentiment analysis; what are the merits and limitations, tools, and quantification of the social sentiment.

Introduction

Social media sentiment is referred as "opinion mining" since it is about digging into the words and context of social posts to understand the opinions underneath [1]. Process of collecting and analyzing how people talk and ideate on a subject with AI capabilities is sentiment analysis. Thinking this process as a layer of understanding to the rest of analytical capabilities of your organization is crucial since it provides a contextual meaning for the rest of the data. It categorizes consumer emotions and market movements by type and intensity, enables market players to innovate, communicate in a way to maintain the competitive edge. Social listening, social monitoring, image analytics and customer experience analysis rely on sentiment analysis for accuracy and usefulness hence it is not limited to understand feelings and opinions but interested in objective facts and neutral statements which are relevant to the subject [2].

How Sentiment Analysis Works?

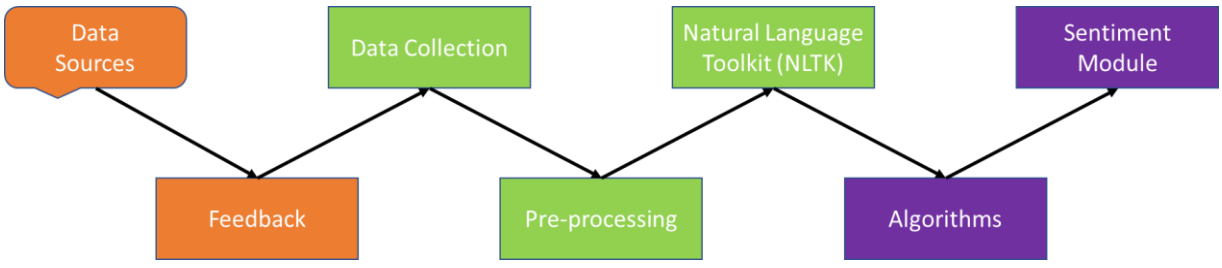

Figure 1. Pipeline of sentiment analysis process [3]

Digging into sentiment analysis which has a pipeline of steps during process. Inputs of the process is obtained by various data gathering methods. For example, Twitter is a critical resource and by using scraping tools, APIs, customers' data feed. User reviews on any sales channel platform mentioning company or brand over a specific period of time is worth listening, and this practice is a common attribute of social media listening forms [3].

Text cleaning allows data to be analyzable. For this process, various tools may be employed to removing stop words, punctuation, and reducing words to their stem. Intermediate product is “cleaned / stripped texts” from anything that might be irrelevant to the analysis [3].

To analyze your sentiments, you’ll need to define the good, the bad, and the ugly. Realistically, you’re looking for words like love, thanks, perfect, incredible for the positives, and worst, hate, avoid for the negatives. Divide your sentiment terms into their own emotion camps to shortlist them. As well as words, scan emojis for a clear understanding of your audience’s feelings about your brand [5]. A classification tool can identify more complex sentiments which yields more evenly distributed data on positive, neutral, negative spectrum.

The final step takes the quantified sentiment data and turns it into actionable information. A timestamp attained to each data will show how sentiment towards a subject change over time with peaks and valleys [3]. A “normal” can be determined in order to monitor chatter differentiating both positive and negative way. This is crucial for finding out baseline in terms of sentiment for a subject. Sentiment analysis also takes product features into account since the only way to sort product quality is the key features. Also, the volume of sentiment around said features, which lets you judge which topics will please the most customers rather than indirectly [4].

Examining emotional map for a 12-cup coffee machine and the discussion that has been going on, might yield a better understanding on the feature sentiments.

Figure 2. Emotional map for a 12-cup coffee [3]

The product's functionality is the most frequently discussed subject, as shown in the graph above. Given the overall unfavorable opinion, this should be at the top of the list of changes. Looking at the topics that generate the most negative conversation, we can say that the coffee maker's illumination, durability, and water reservoir capacity are among the worst. The amount of conversation about the machine in each of those areas is far less than the amount of talk about functionality. The overall numbers may not result in as much of a gain in customer satisfaction generally, even though changing them will produce the best feeling in individuals who were dissatisfied [3].

Reach the right audiences with right tone and message is the ultimate dream, and with a sentiment analysis on a subject people / companies might gain critical insights to alter the views in their favor. You may discover what your audience wants by observing patterns and looking into peaks in positive, valleys of negative, and neutral emotion. This can help you understand the type of messaging you should post on each social media platform [5].

Methodologies

Sentiment analysis can be performed with various approaches based on amount of data, platform, and budget available. A simple dictionary-based methods might perform better in analyzing information gathered from a formal ideation environment. However, with a vast amount of data, a variety of natural language processing techniques (NLP) ranging from simple to complex and developed can be utilized to achieve accurate and insightful results. Focusing on strengths and limitations of each approach is critical for picking up the right one for the next sentiment analysis.

Dictionary-Based Approaches

Simple yet effective and most-widely used approaches to sentiment analysis on social media is the dictionary-based method. A non-topic-based text categorization of words each attaining a sentiment value is pre-defined along with phrases associated with positive and negative value. An easy classification example might be classifying a tweet which includes “love” as positive sentiment whereas another tweet with “hate” is classified as negative.

Advantages of dictionary-based sentiment analysis are easy implementation, domain customizability and language interpretability. As long as a suitable dictionary is buildable, it is easy to construct one for any language. A diligent dictionary preparation is critical for successful sentiment analysis. Additionally, for specific domain or topic where the needs and opinions are clustered, a dictionary-based approach is handy since it is customizable to fit the needs of the matter.

Off-the shelf dictionaries and existing dictionaries are readily available for use of sentiment analysis. One well-known example is the Linguistic Inquiry and Word Count dictionary which has 5000+ words and phrases which are presented in a categorically organized manner [6]. Sentiment 140 dictionary and the General Inquirer dictionary are other dictionaries that have been used for sentiment analysis on social media [7][8].

On the other hand, systems that are based on dictionaries also have a number of drawbacks. They are prone to making mistakes, especially when it comes to dealing with sarcasm or irony [9]. They also rely on the notion that the tone of a social media post can be effectively conveyed by a single word or phrase, which is not necessarily the case in all situations. [10]. For example, a tweet containing the word "not" followed by a positive word (e.g. "not bad") might be classified as neutral or negative by a dictionary-based approach, despite the overall sentiment being positive.

To overcome these limitations, some researchers have recommended the use of more advanced techniques, such as the incorporation of background information or the use of idioms that convey emotion [11]. Adding information about the context in which a word or phrase is used (such as the existence of negation or intensifiers) might increase the accuracy of dictionary-based techniques for sentiment analysis [12]. In addition, several academics have advocated using machine learning techniques to improve the precision of dictionary-based algorithms [13]. Using a dictionary as one of the features, for instance, a classifier may be trained on a dataset of annotated social media posts and then used to predict the sentiment of fresh, unannotated social media posts. By combining the simplicity and adaptability of dictionary-based approaches with the strength of machine learning, it may be able to increase the performance of social media sentiment analysis.

Machine Learning Approaches

With the advancement of high-level languages and machine learning algorithms, inevitably market players demanded more sophisticated approach to sentiment analysis. Machine learning algorithms rely on training of artificial intelligence model with a dataset of labeled opinions and utilizing the trained model to predict label for unannotated data. For sentiment analysis purposes of social media content, a specially purposed “classifier” is trained on social media posts with human evaluated sentiment value annotations. Later this classifier is used for predicting actual sentiment on a social media regarding a topic or a product. Numerous machine learning algorithms are used for sentiment analysis. SVM, Naïve-Bayes classifiers, and neural networks are the common examples.

- Support vector machines (SVMs): SVMs are supervised learning algorithms used for classification tasks. They function by locating the hyperplane in a high-dimensional space that maximally divides distinct data types. SVMs have been utilized for social media sentiment analysis, with some research indicating good results. [14].
- Naive Bayes classifiers: The Naive Bayes classifier is a form of supervised learning method focused on the maximization of a posteriori probability. They assess the likelihood of a class given a set of features by applying Bayes' theorem. Using Naive Bayes classifiers for social media sentiment analysis, with some studies finding good results [15].
- Neural networks: Neural networks are a method for machine learning that are modeled after the structure and function of the human brain. They comprise of numerous layers of interconnected nodes and can be applied to a variety of tasks, such as sentiment analysis. The use of neural networks for sentiment analysis on social media has been reported with high F1 scores on several research. [16].

When conducting sentiment analysis on social media, adopting methodologies informed by machine learning offers a number of important benefits. One of the most significant benefits is that they are able to handle more complicated language structures and situations than systems that are based on dictionaries, and they may also be more accurate in determining emotions. [17]. Additionally, machine learning approaches can be fine-tuned and improved over time through the use of additional training data.

However, approaches to sentiment analysis that are based on machine learning also have their own set of constraints. The necessity for a large dataset that has been annotated in order to train the classifier is one of the most significant drawbacks. Acquiring such a dataset may be time-consuming and resource-intensive. [18]. In addition, systems based on machine learning may be less interpretable than approaches based on dictionaries, making it more challenging to comprehend the rationale behind the assignment of a particular emotion to a specific social media post. [19].

Commercial Applications

Businesses are heavily dependent on customer profile, opinions around their product and market positioning. Pricing power, market share and e-com sales have huge impact on profitability which drives managers to perform sentiment analysis continuously. Customer feedbacks are crucial for planning the next iteration of product development. Market research can yield customer perception which can be altered in a positive way to increase sales and gain competitive advantage. Regular checks and reports on brand monitoring incentives insight driven business actions and results. Drawbacks on the other hand are these efforts are limited with customers who have reach and active on social media. A product which targets silent generation might not be suitable for conducting and making business decisions based on sentiment analysis [3].

- Improved customer satisfaction: By analyzing customer sentiment, businesses can identify areas of their product or service that are particularly well-received or poorly received, and take steps to address any issues or concerns. This can lead to improved customer satisfaction and loyalty.
- Increased sales: By understanding the sentiment of their customers, businesses can tailor their marketing and sales efforts to better align with the needs and preferences of their target audience. This can lead to increased sales and revenue.
- Competitive advantage: By regularly analyzing customer sentiment, businesses can gain a better understanding of their market and identify opportunities for innovation and differentiation. This can give them a competitive advantage over their rivals.
- Improved brand reputation: By monitoring and responding to customer sentiment, businesses can proactively address any negative perceptions and build a more positive reputation for their brand.

Political Applications

Sentiment analysis on social media also used by political actors who are in need to solidify their base and gain more supporters. Understanding overall sentiment on a topic help political campaigners to tailor their message and come up with a better strategy [2]. Sentiment analysis can provide insights into attitudes and opinions of the general community.

- Track the sentiment of voters around a particular candidate or issue: By analyzing social media posts related to a particular candidate or issue, political campaigns and

organizations can get a sense of how people feel about them. This can help them identify areas of support or concern and adjust their messaging accordingly.

- Identify trends and patterns in public opinion: By regularly analyzing social media sentiment, political campaigns and organizations can identify trends and patterns in public opinion over time. This can help them understand the changing attitudes and concerns of the public, and respond accordingly.

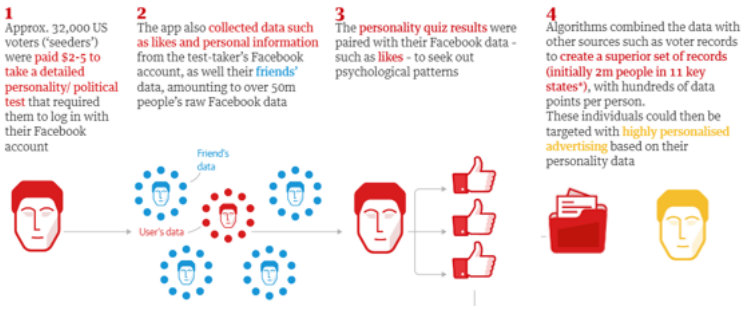

Figure 3. Political Sentiment Analysis [20]

- Monitor and respond to negative sentiment: By tracking negative sentiment on social media, political campaigns and organizations can proactively address any concerns or issues that may arise. This can help them maintain a positive reputation and address any potential negative impacts on their campaign.
- Engage with voters and the public: By analyzing social media sentiment, political campaigns and organizations can identify opportunities to engage with voters and the general public on particular issues or concerns. This can help them build relationships and establish trust with their constituents.

Sentiment analysis in the hands of greedy politicians became a powerful tool and many ethical considerations and debates has been going on since infamous "Cambridge Analytica" scandal. Facebook data along with a survey has given democratic party chance to profile electors and act in order to address fears and hopes of larger society strategically [20].

Conclusion

In this paper, we reviewed the current state of the art in sentiment analysis on social media, describing process, methodology, common practices with advantages and limitations. Sentiment analysis on social media is a complex and multifaceted process that can provide valuable insights into public opinion and sentiment. Through a combination of dictionary-based, machine learning, and hybrid approaches, organizations can gain a more comprehensive understanding of how their customers, clients, or stakeholders feel about a particular topic or product. However, it is important to carefully consider the limitations and potential biases of different approaches, as well as ethical concerns such as privacy and the potential for manipulation or exploitation of data. By using a variety of methods and sources and being mindful of these issues, organizations can use sentiment analysis on social media to inform and improve their decision-making and communication strategies, and to better understand and respond to the needs and desires of their audiences.

References

[1] Newberry, Christina. “Social Media Sentiment Analysis: Tools and Tips for 2022.” Social Media Marketing & Management Dashboard, 27 Oct. 2022, https://blog.hootsuite.com/social-media-sentiment-analysis-tools/#:~:text=sentiment%20over%20time.-,What%20is%20social%20media%20sentiment%20analysis%3F,analysis%20considers%20emotions%20and%20opinions.

[2] Oliveri, Sergio. “What Is Social Sentiment Analysis and Why Is It Important?” NetBase Quid, 12 Oct. 2022, https://netbasequid.com/blog/what-is-social-sentiment-analysis/#WhatSA1.

[3] Spencer, Emily Louise. “Sentiment Analysis for Brand Building: A Comprehensive Guide.” Blog, 9 Sept. 2022, https://www.revuze.it/blog/sentiment-analysis/.,

[4] “Social Media Sentiment Analysis – Explained.” Oktopost, 16 Nov. 2022, https://www.oktopost.com/blog/social-media-sentiment-analysis/.

[5] Sitbon, Jonathan. “Ethos, Pathos, Logos: What Are They and How to Use Them.” Content-Writing, Content-Writing, 24 July 2022, https://www.wix.com/wordsmatter/blog/2020/12/ethos-pathos-logos/.

[6] Pennebaker, J. W., Francis, M. E., and Booth, R. J. (2001). Linguistic Inquiry and Word Count: LIWC2001. Mahwah, NJ: Lawrence Erlbaum Associates. https://www.liwc.app/

[7] Go, A., Bhayani, R., and Huang, L. (2009). CS224N Project Report, Stanford University. http://help.sentiment140.com/for-students

[8] Stone, P. J., Dunphy, D. C., Smith, M. S., and Ogilvie, D. M. (1966). The General Inquirer: A Computer Approach to Content Analysis. Cambridge, MA: MIT Press. https://mitpress.mit.edu/9780262690119/the-general-inquirer/

[9] Kim, Y. (2017). Convolutional neural networks for sentiment analysis of short texts. In Proceedings of the International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2017, pp. 264-277.

[10] Mohammad, S., Kiritchenko, S., and Zhu, X. (2015). NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets. In Proceedings of the NAACL-HLT 2015 workshop on computational approaches to subjectivity, sentiment and social media analysis, WASSA ’15, pp. 70-79.

[11] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies, 5(1), 1-167.

[12] Pang, B., and Lee, L. (2008). Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval, 2(1-2), 1-135.

[13] Wiegand, M., and Pletka, J. (2013). Enhancing dictionary-based sentiment analysis with machine learning techniques. In Proceedings of the International Conference on Computational Linguistics and Intelligent Text Processing, CICLing 2013, pp. 482-493.

[14] Zainuddin, Nurulhuda & Selamat, Ali. (2014). Sentiment analysis using Support Vector Machine. I4CT 2014 - 1st International Conference on Computer, Communications, and Control Technology, Proceedings. 333-337. 10.1109/I4CT.2014.6914200.

[15] Sharma, Manish. "Sentiment Analysis (Introduction to Naive Bayes Algorithm)." Medium, Towards Data Science, 11 May 2020, https://towardsdatascience.com/sentiment-analysis-introduction-to-naive-bayes-algorithm-96831d77ac91.

[16] Oyinlola, Salim. "How to Train a Neural Network for Sentiment Analysis." DigitalOcean, DigitalOcean, 28 Nov. 2022, https://www.digitalocean.com/community/tutorials/how-to-train-a-neural-network-for-sentiment-analysis.

[17] Stewart, Matthew. "The Limitations of Machine Learning." Medium, Towards Data Science, 29 July 2020, https://towardsdatascience.com/the-limitations-of-machine-learning-a00e0c3040c6.

[18] Boiy, Erik & Moens, Marie-Francine. (2009). A Machine Learning Approach to Sentiment Analysis in Multilingual Web Texts. Inf. Retr.. 12. 526-558. 10.1007/s10791-008-9070-z.

[19] Creating and Evaluating Resources for Sentiment Analysis in the Low-resource Language: Sindhi (Ali et al., WASSA 2021)

[20] "Cambridge Analytica: How Did It Turn Clicks into Votes?" The Guardian, Guardian News and Media, 6 May 2018, https://www.theguardian.com/news/2018/may/06/cambridge-analytica-how-turn-clicks-into-votes-christopher-wylie.