Three Executable Applications Based on Sentiment Analysis

Yitong Peng

Abstract:

Sentiment analysis can be said to be one of the hottest topics in the world, and the discussion of scholars on it involves different professions and fields. In digital society, sentiment analysis as a powerful tool has gradually penetrated into various fields from theory. The purpose of this article is to propose three industries and applications where sentiment analysis can shine. The main contribution of this article is to provide enterprises with information including project feasibility, disadvantages and possible solutions.

Keywords: Sentiment analysis; Opinion mining; NLP; Machine Learning

1. Introduction

Sentiment analysis is devoted to the study of opinions about task-specific entities (Feldman, 2013). Whether in business, media or politics, the decision-making process is influenced by the opinions of users and society. Users' emotions exploded from social media including Twitter, Facebook, Tiktok. Nowadays, emotions, like other data, are happening in huge quantities and updated in real time. Sentiment analysis, or opinion mining, is the computational study of an entity's opinions, attitudes, and emotions. The definitions and functions of sentiment analysis or opinion mining are considered primarily interchangeable, although some scholars have argued that there is a slight difference between them (Tsytsarau & Palpanas, 2012). Sentiment analysis can extract people's emotional tendencies from large amounts of text, which not only helps companies understand customer needs, but also provides an effective way to manage content for social media platforms.

2. Thesis statement

From my perspective, sentiment analysis can develop in business strategy, social media, and maintain social harmony. Since emotions themselves have no precise criteria to measure, any numerical values and thresholds are artificial interventions and assumptions that may not be appropriate in some situations. Therefore, it is necessary to constantly update and evolve to optimize. The combination of sentiment analysis with other mature technologies can achieve synergies and further expand the impact. In other words, sentiment analysis is a process that relies on various support to quantify subjective emotion expression and tone.

Specifically, the role of sentiment analysis can be embodied in the following three applications. First, optimizing business decisions can improve the user experience and achieve business monetization. Second, sentiment analysis can enhance mental health monitoring and intervention. Third, this aspect of technology can promote the progress and application of human-machine interaction technology. Some of these applications are already efficient business or

social models, but others are still at the stage where they are theoretically considered executable.

3. Discussion

3.1 The principles and processes on which the three applications depend

To a certain extent, sentiment analysis can serve as a task formulated by NLP to retrieve the emotions expressed in the text (Tsytsarau & Palpanas, 2012). After classification based on positive, negative, and neutral, the classification results can be annotated as supervised learning labels (Goldberg & Zhu, 2006). In other words, the acquisition of emotions in messages and text provides a good source of classifier training data in machine learning.

Lexicon based approach is one of the main sentiment analysis methods, including corpus based approach and dictionary based method (Wankhade et al., 2022). In practice, specific operations need to be carried out with the help of lexicon based method available dictionaries, and whether the emotional orientation of words will change with the context can be judged by semantic analysis.

A dictionary is a collection of tokens in which numerous words are assigned a predefined score (Kiritchenko et al. 2014). Neutral, positive, and negative properties are assigned based on polarity by numerical notation such as + 1,0, −1 for positive, neutral, negative, or fractions can be assigned according to the strength of the polarity and its value with the range [+1, −1], where + 1 means extremely positive and −1 means extremely negative. This process depends the operator's own definition of the score without precise standard, that is, different dictionaries may be the same word for the number of different. The next step is dividing a specific document or a paragraph or a sentence into words, so that each word becomes an independent unit. Remove meaningless words such as how, to, also, etc., and then match the remaining words against a dictionary to determine the score of each word in the target field.

Once the score has been assigned, it can be processed in two ways, either as an integration or as a data source. If the results of sentiment analysis are integrated, the numerical values of some words can be determined according to the semantics. The sum of the values of a sentence, paragraph, or article can be weighted to make the final result reasonable. Because the emotional positives and negatives of words are artificially set, even the integrated judgment results rely on artificially set thresholds. Therefore, training as a label in the data can make the model have a general decision, which facilitates the disposal of large-scale information decision. I believe that this way of processing and discerning emotion in words can serve the three practical applications mentioned above.

3.2 Business decisions and market analysis

First, opinion mining helps with business decision-making and obtaining market feedback, such as product reviews, market research, reputation management, and social media policing. Enterprises and factories can collect a large number of user experience feedback and comments from users and buyers from multiple channels, and judge whether the launched products meet the market demand and meet consumer expectations through analysis and quantification. Adjusting the marketing strategy or directly updating the product, according to the results of the judgment, is an effective measure to improve the competitiveness of the company.

This application is already being used in some enterprises. On Aug. 20, BMW said it was recalling 720,796 vehicles in North America because they were equipped with electric pumps that could short circuit due to faulty installation, posing a fire hazard (Lee, 2024). Although BMW does not disclose its exact sentiment analysis process, its awareness and response to negative reviews do exist. According to BMW, 18 negative comments posted on the company's website in August were detected by the company, all about short circuits in electric pumps (Lee, 2024). If I were a decision maker at BMW, I would inevitably gather further comments to determine the authenticity and urgency of the failure, and then examine these vehicles to come up with solutions to reduce the financial and reputation damage.

Future innovations could focus on improving the accuracy of sentiment analysis results. Because of the fierce market competition, the illegal behavior in the network ecology is widespread. Rival companies may hire some users to comment on untrue information, which may cause sentiment analysis results to deviate from reality. Optimization can start by identifying the characteristics of fake reviews, such as comments that suddenly appear in a short period of time with plenty of new registered users, or comments that are highly similar or that point to extremely specific words. When suspicious information arises, decision makers can choose to optimize dictionary and contextual patterns or to artificially interfere with the analysis process to screen out abnormal data.

3.3 Monitor the platform and maintain network harmony

Additionally, some argue that emotional expression is not necessarily subjective in nature (Wilson et al., 2005). This method of emotion definition and analysis framed by stan-

dards and definitions can be used to detect and constrain behaviors that violate the harmony of network society. Opinion mining is to describe the emotions in words as objectively as possible. The process that it describes is to concretize and visualize the criteria of measurement. For example, lexicon based approach is based on predefined dictionaries and semantics, and these materials change with expressions in the real world. Sarcasm, that is, a sentence with a high level of positive after the integration score but a high level of negative in a special context, there are also keywords to indicate the appearance of ironic context.

As a result, it is possible to accurately judge emotions in most cases. If law enforcement agencies want to identify inappropriate comments and statements, they can combine technology such as crawling to find users who touch sensitive words across the entire platform and network. Once a user who is defined as breaking the law defends his or her remarks or disobeys the results, the mention of certain socially unethical words is evidence of guilt. Thus, these words do cause discomfort to other users or damage the harmonious atmosphere of online discussion, even if they are not aware of it, human moderators can be added to avoid misjudgments. Some social media should already be using this technology to avoid overheated comments. For example, some platforms will detect illegal topics before users post posts and videos, which can prevent inappropriate comments from breeding in the first place.

At present, the detection of sensitive words is mainly the responsibility of employees since the technology is not mature and comprehensive. Context and dictionaries have key words that can be set to identify sarcasm, but it is difficult to correctly identify emotions with complex sentence patterns such as double negative and rhetorical question. It does not rule out that many users have language problems when commenting, which makes sentiment analysis need to consider more factors and add more conditions.

3.4 Personalized recommendation

The main existing uses of sentiment analysis, including the two previously mentioned applications, are basically to determine the general opinion of the community on some specific issues. In my opinion, focusing on individual users also enables opinion mining to effectively improve user satisfaction. For example, recommendation algorithms mostly rely on information such as the time users spend on videos related to a certain topic or the keywords users search for. Theoretically, relying on the results of sentiment analysis can also help the model to cluster and recommend. Specifically, the information related to the topic that the user comments have a positive attitude is recommended, and the opposite is not recommended.

It is also feasible to approach the community, but is essentially user-based. Analyzing the interests of each user and considering the similarities between users. For instance, if two users commented that a blogger's share is intriguing, the other blogger that one user likes may attract the other user. However, the reason why this application is not widely implemented may be that it is not precise and concise enough. Data like viewing time and number of likes are easy to read and impossible to disambiguate, while positive and negative attitudes can be misjudged and require a complicated process of matching dictionaries. As a result, many platforms find sentiment analysis inefficient. In the final analysis, the technology of sentiment analysis is immature and high cost, which cannot meet the needs of the current market.

4. Conclusion

This article introduces the process and essential theoretical basis of sentiment analysis from its inception to its rapid development. Over the past few years, we have witnessed an increase in interest in textual data. Many seemingly subjective and abstract entities can be described using quantitative criteria and objective definitions. Even though opinion mining has great potential and economic value, there are limits to its development and innovation due to its immaturity. According to the needs of society and enterprises, three types of scenarios and industries that can widely implement sentiment analysis are proposed. The first application is how to help companies improve their competitiveness and reputation through market research and user reviews. It also gives an example to show that BMW has been using sentiment analysis to reduce losses, but the method is not systematic enough. The problem with fake reviews and what the individual considers feasible solutions are involved. Next, the article discussed how dictionaries and context Settings can identify illegal behavior of online comments. However, because of the complexity and ambiguity of sentence patterns, the judgment cannot be completely precise, and it needs to be combined with manual labor. This last application could theoretically be used by popular recommendation algorithms. Due to the complexity and uncertainty of OM itself, most enterprises and platforms do not consider it. Future work will include simplifying the sentiment analysis process and optimizing the theory so that its accuracy is not constrained by dictionaries, contexts, and sentence patterns.

References

Feldman, R. (2013). Techniques and applications for sentiment analysis. *Communications of the ACM*, *56*(4), 82-89.

Tsytsarau, M., & Palpanas, T. (2012). Survey on mining subjective data on the web. Data Mining and Knowledge Discovery, 24, 478-514.

Goldberg, A. B., & Zhu, X. (2006). Seeing stars when there aren't many stars: Graph-based semi-supervised learning for sentiment categorization. In Proceedings of TextGraphs: The first workshop on graph based methods for natural language processing (pp. 45-52).

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review, 55(7), 5731-5780.

Kiritchenko S, Zhu X, Mohammad SM (2014) Sentiment analysis of short informal texts. J Artif Intell Res 50:723–762

Lee, A. M. (2024). BMW recalls more than 720,000 cars because electric water pump may catch fire. CBS News. https://www.cbsnews.com/news/bmw-recall-720000-vehicles-electric-water-pump-fire-risk-2024/

Wilson, T. (2005). Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis. In Proceedings of HLT/EMNLP.