



A Comprehensive Analysis of Opinion Mining and Sentiment Analysis for Social Impact

Dr Meenakshi Garg

Vivekanand Education Society's Institute of Technology, Mumbai , India, Email: meenakshi1404@gmail.com

Abstract

Due to the explosive increase of user-generated material on numerous online platforms, sentiment analysis, also known as opinion mining, has become a crucial area of study. This study provides a thorough analysis of the challenges and applications of sentiment analysis. Recent years have seen a substantial increase in interest in sentiment analysis and opinion mining as a result of the proliferation of user-generated material on numerous internet platforms. By analyzing their effects across several sectors, this research investigates how opinion mining and sentiment analysis tools might be used to benefit society. The study focuses on the role that these methods can play in social media activism, public sentiment analysis, brand reputation management, and decision-making processes. The research intends to shed light on the possible advantages of utilizing sentiment analysis and opinion mining for societal well-being and to offer suggestions for the field's future developments. This survey attempts to provide insights into the current state-of-the-art in sentiment analysis and provide guidelines for future research in this dynamic field by examining the existing literature.

Keywords: Sentimental Analysis, languages, text, social media, opinion.

1. Introduction

Not only among researchers but also among enterprises, governments, and organizations in recent years, sentiment analysis has gained substantial attention and implementation (Sánchez- Rada and Iglesias 2019). The Internet has risen to prominence as the main source of information on the planet. Numerous users use different internet platforms to share their thoughts and opinions. Utilizing user-generated data and automatically analyzing it has become essential for monitoring public mood and promoting informed decision-making. As a result, sentiment analysis is becoming increasingly well-liked among research communities. Opinion analysis and opinion mining are other names for this discipline. Sentiment analysis has seen substantial growth recently, with numerous research studies concentrating on various facets of this endeavor. Piryani et al. (2017) studied the subject from 2000 to 2015 and proposed a framework for computationally processing unstructured data to extract views and identify their attitudes. Lighthart et al. (2021) gave an overview of opinion mining in its early phases. Several surveys have offered insights into the difficulties and prospective directions in sentiment analysis, including those by Yousif et al. (2019) and Birjali et al. (2021). [13]

Research on sentiment classification has been done by Soleymani et al. (2017), Yadav and Vishwakarma (2020), Yue et al. (2019), and Liu et al. (2012), including the detection of opinion spam and false reviews. Online reviews can be used for sentiment categorization, predictive decision-making, and the identification of fraudulent reviews using machine learning systems, according to Jain et al. (2021b). Balaji et al. (2021), using powerful machine learning algorithms, thoroughly investigated a variety of social media analysis applications. [4] A succinct review of the machine learning techniques used in social media analysis is also provided by Hangya and Farkas (2017).

As social networking sites have grown in popularity, other areas have emerged that are devoted to studying these networks and gleaning useful information from their content. Sentiment analysis is one such area, with its emphasis on identifying the sentiments represented in text based on its content. Sentiment analysis is a branch of natural language processing (NLP) that has attracted a lot of interest since it has the ability to guide judgments based on public opinion. Even in the new millennium, sentiment analysis is still evolving and developing, despite the fact that there have been early publications on the topic.

Sentiment analysis is useful in a variety of real-world contexts, including product analysis, where it helps to pinpoint the features or attributes of a good that appeal to consumers. A method for developing adaptable aspect-based lexicons for sentiment categorization is put forth by Mowlaei et al. in 2020. They outline two methods for building dynamic lexicons that provide sentiment categorization based on various factors, one of which is based on statistics and genetic algorithms.

Hotels, airlines, healthcare, and the stock market are just a few of the industries that have used sentiment analysis in the reviews space (Zvarevashe and Olugbara 2018). To better understand client preferences and dislikes, hotel reviews have been studied. Sentiment analysis has been used to identify market trends in the context of the stock market and cryptocurrencies (Valencia et al. 2019).. Ahmad et al. (2019a) analyze sentiments of tweets across different domains. [1] In the healthcare domain, sentiment analysis has seen increased application, including customer opinion analysis, customer satisfaction analysis, and other healthcare-related applications. [8] The business sector has long leveraged sentiment analysis for various purposes, including reputation management, market research, competitor analysis, product analysis, and understanding customer voices (Ahmad et al. 2019a). Sentiment analysis continues to play a vital role in improving business operations and decision-making processes.

The obstacles faced by sentiment analysis and natural language processing include sarcasm, irony, informal writing styles, and language-specific problems. Individuals' informal writing practices might make sentiment analysis more

difficult because of the use of slang, acronyms, and irregular syntax. As they involve the opposite or distinct meaning from what is openly expressed, irony and sarcasm provide considerable challenges.

The differences in word meanings and orientations between various languages, contexts, and domains give rise to language-specific difficulties. These difficulties are made even more difficult by the lack of tools and resources for sentiment analysis across all languages. Sarcasm and irony identification in text, which are regarded as major obstacles in sentiment analysis, have been actively researched. [11]

This research attempts to examine the numerous difficulties in sentiment analysis as well as the approaches, tools, and algorithms used in this area. To give the task complications a clear grasp, comparative data analysis will be presented utilizing tables, flow charts, and graphs. These visual representations will make it easier to understand the analysis's conclusions and key insights.

SENTIMENTAL ANALYSIS LEVEL

Depending on the level of detail used in the analysis, sentiment analysis can be divided into different categories. Understanding the breadth and depth of sentiment analysis is aided by these levels. The typical sentiment analysis levels are as follows:

Sentiment analysis at the document level seeks to ascertain the overall attitude indicated in a textual component, such as a complete document, article, or review. It offers a broad grasp of the sentiment's positivity, negativity, or neutrality. [5]

Sentiment analysis at the sentence level: Sentiment analysis at the sentence level is concerned with examining the sentiment expressed inside specific sentences. It seeks to categorize each statement as having a positive, negative, or neutral sentiment, depending on its polarity. This degree of research offers more detailed perceptions into the sentiment included inside a piece.

Sentiment analysis that is focused on particular traits or aspects of a good, service, or subject goes deeper by locating and examining these sentiments. It seeks to comprehend the feelings connected to each individual feature. Aspect-based sentiment analysis, for instance, can discover feelings for features like price, usability, customer service, etc. in a product review.

Entity-level Sentiment Analysis: Entity-level sentiment analysis is concerned with figuring out how readers feel about certain things that are referenced in the text, such individuals, groups, things, or brands. Each sentiment connected to each entity is identified and examined separately.

Sentiment analysis with finer granularity: Sentiment analysis with finer granularity goes beyond straightforward positive, negative, or neutral classifications. To represent a more complex sentiment analysis, it seeks to assign sentiment scores or intensities. This level of analysis offers a thorough comprehension of the many sentiment levels, including minute details and nuances of thought. These sentiment analysis levels offer several viewpoints and insights, enabling a more thorough comprehension of the sentiments portrayed in text data. The suitable level selection is based on the application's specific analysis goals and needs.

DATA COLLECTION AND FEATURE SELECTION

Data Collection

Data Collection in Sentiment Analysis: Gathering the essential text data for analysis is a necessary first step in sentiment analysis. To ensure accuracy, the data should be comprehensive and representative of the target subject or issue the effectiveness of the sentiment analysis model. There are several approaches to data collection:

- **Manual Data Collection:** Using this method, human annotators manually assign sentiment labels to text data. Although it takes some time, this method enables precise sentiment annotations.
- **Crowdsourcing:** Data that has been sentiment-labeled can be gathered from a larger pool of annotators using crowdsourcing services like Amazon Mechanical Turk. This method speeds up the collection of a huge amount of data, although the quality of the annotations may vary.
- **Pre-existing Datasets:** Researchers can also make use of pre-existing datasets that have already been sentimentally annotated. These datasets, which frequently span a wide range of topics including movie reviews, product reviews, social media posts, etc., are readily accessible to the general public.

Feature Selection

Choosing the most pertinent and instructive features from the text data to represent sentiment is known as feature selection in sentiment analysis. Enhancing the precision and effectiveness of sentiment analysis algorithms requires effective feature selection. Typical methods for feature selection include [2]:

Bag-of-Words (BoW): Using the BoW technique, text data is represented as a group of distinct words or tokens. The choice of features depends on how frequently they appear in the corpus of text. When choosing features, stop words (common words like "the," "is," etc.) may not be included. **N-grams:** In a text, an n-gram is a group of n consecutive words. They accurately depict word placement and context. Unigrams (single words), bigrams (two consecutive words), and trigrams (three consecutive words) are examples of frequently seen n-grams.

Lexicon-based Features: Sentiment dictionaries or lexicons are pre-defined sets of words that have corresponding sentiment scores. Lexicon-based features use the sentiment scores as features and match terms in the text to the

sentiment lexicon. [10]

Word embeddings: By portraying words as dense vectors in a high-dimensional space, word embeddings are able to capture the semantic meaning of words. Word2Vec, GloVe, or FastText are examples of popular word embedding techniques that can be used to extract useful features.

Advanced approaches: To choose informative features, advanced feature selection approaches can be used, including TF-IDF (Term Frequency-Inverse Document Frequency), word frequencies, and statistical measurements like chi-square or mutual information.

The nature of the data, the difficulty of the sentiment analysis task, and the computational resources available all influence the feature selection technique that is used. To find the best combinations for precise sentiment analysis, it is crucial to experiment with various feature selection techniques. [9]



Fig 1: Sentimental Analysis Process

NEED OF SENTIMENTAL ANALYSIS

Sentiment research is very beneficial since it gives them an understanding of how their customers view their brand. Organizations can make wise choices by automatically classifying the emotions displayed in social media interactions, reviews, and other types of data. Sentiment analysis includes a number of techniques and tactics that let businesses examine the opinions of their clientele with reference to a certain good or service. It aids in determining the polarity of feelings (positive, negative, or neutral), the topic of conversation, and the people or things expressing the feelings. Businesses may efficiently comprehend and respond to client feedback by utilizing sentiment analysis, which improves customer happiness and helps them make wise decisions. The following explanations explain why sentiment analysis has grown in significance:

Customer feedback: Whether it comes from social media, online reviews, polls, or customer support contacts, sentiment analysis aids firms in understanding and analyzing customer feedback. Businesses can find areas for improvement, respond to consumer issues, and improve their products or services by understanding the mood underlying client feedback.

Sentiment analysis makes it possible for businesses to track and manage their brand reputation. Businesses may spot potential problems, quickly address negative sentiment, and take proactive steps to safeguard and improve their brand image by evaluating the sentiment expressed in social media posts, news stories, or online forums.

Sentiment analysis gives businesses insight into consumer preferences, attitudes, and trends. Market research and competitive analysis. Businesses can identify emerging market trends, comprehend customer sentiment toward their own products as well as those of competitors, and make educated decisions about product development, marketing strategies, and market positioning by analyzing sentiment across various demographics or customer segments.

Sentiment analysis is used in political campaigns and administration to monitor voter sentiment, gauge public opinion, and evaluate the efficacy of political communications. Political groups can modify their messaging, concentrate on important subjects, and comprehend the attitudes of voters by evaluating sentiment expressed on social media, in news stories, or in public forums.

client service and support: Customer service and support systems can employ sentiment analysis to rank and classify incoming client questions or complaints depending on their sentiment. This aids businesses in identifying pressing problems, offering prompt solutions, and effectively allocating resources to meet client concerns.

Sentiment analysis assists companies in gathering feedback on their goods and services from a variety of sources, such as client reviews, surveys, and internet discussions. Businesses can improve their services by evaluating sentiment to spot trends, find areas for improvement, and make data-driven decisions.

Sentiment analysis is an effective method for deciphering and harnessing the power of human emotions as they are conveyed in text. It makes it possible for companies, organizations, and governments to obtain knowledge, make wise choices, and take proactive steps to enhance their goods, services, and overall consumer experience.

APPLICATIONS OF SENTIMENT ANALYSIS

Sentiment analysis is a flexible technique with many uses in a variety of fields. It can be applied to examine consumer feedback, judge the mental health of patients based on social media posts, and much more. The potential uses of sentiment analysis have also been increased by developments in technologies like Blockchain, IoT, Cloud Computing, and Big Data, making it useful in practically any field. Sentiment analysis can now handle massive amounts of data, assure data security, take use of real-time data streams, and use cloud-based infrastructure thanks to these technical breakthroughs. In order to provide data-driven insights and decision-making in virtually any sector or domain, sentiment analysis can be applied into a variety of fields.

Product Review

Both the quantity of products sold and the number of customer evaluations have significantly increased as a result of e-commerce's explosive expansion. Customers can make far more educated selections about their purchases by analyzing the sentiment indicated in these reviews (Paré, 2003). According to Schouten and Frasincar (2015), sentiment analysis can be carried out at either the phrase level or the aspect level. Businesses can learn more about what customers think about their newest items, whether it be after launch or through the analysis of comments and reviews, by performing sentiment analysis on product reviews. Relevant keywords can be used to concentrate on particular features of a product, such as food quality, customer service, or cleanliness. by educating a framework for sentiment analysis. [12]

Business analysis

Business intelligence sentiment analysis has many benefits, allowing companies to use the data for product development, consumer feedback analysis, and creative marketing tactics. The main use of sentiment analysis in business intelligence is to look at how customers feel about certain goods or services. However, customers can also use sentiment analysis to review products and make informed judgments. This approach is not just used by product manufacturers. In business intelligence, sentiment analysis has many advantages. For example, businesses can use the data to improve their products, assess client feedback, and create successful marketing strategies (Han et al., 2019). Though sentiment analysis is frequently used employed to analyze customers' perceptions of products or services, it extends beyond the purview of product manufacturers. In order to assess different items and make wise decisions, consumers themselves might use sentiment analysis.

Bose et al. (2020) used an emotion lexicon to classify attitudes into eight different emotions and two moods (positive and negative) in an analysis of restaurant service reviews on Amazon.com over an eight-year period. They discovered that sentiment analysis may spot customer tendencies and potential dangers, which eventually results in higher client satisfaction.

Review analysis

Sentiment analysis is heavily used in the entertainment industry. It is frequently used to examine movie, television, and short film reviews in order to ascertain the reactions of the audience (Kumar et al., 2019). This not only helps viewers make informed decisions, but it also promotes the visibility and acceptance of high-quality material. To precisely ascertain the general sentiment represented in the reviews, sentence-level sentiment analysis (Lin and He, 2009) is frequently used in this field. [6]

Similar to other industries, the travel sector has been making efforts to improve customer experiences by deploying machine learning and online consumer recommendation systems that make use of smart, data-driven decision-making processes (Jain et al., 2021f). Additionally, Jain et al. (2021e) talked on how customer decisions might be classified as positive or negative based on web information on online reviews provided by valuable consumers. This approach helps the travel industry gain insights into customer feedback and improve their services accordingly.

Customer reviews

Customers and owners can benefit from sentiment analysis of hotel and restaurant reviews by using it to make better decisions and grow their businesses. The goal of aspect-based sentiment analysis is to pinpoint the parts of reviews that attract the majority of favorable or negative comments. According to Sann and Lai (2020) and Al-Smadi et al. (2018), this study enables hotels and restaurants to pinpoint areas for improvement, enhancing both their reputation and client happiness. [3]

Due of its potential advantages, sentiment analysis has found particular favor in the hospitality sector. Customers can choose better options by reading reviews and considering the opinions of past visitors. Additionally, ABSA offers owners and service providers information into particular elements that elicit favorable or unfavorable opinions. This enables them to focus on areas that require attention and make necessary improvements.

In the end, service providers stand to benefit the most from sentiment analysis because they can spot the components getting bad reviews and take action to improve them. Hotels and restaurants can improve client experiences, build their reputations, and eventually increase profitability by addressing customer complaints and fixing flaws.

Social media monitoring

Customer sentiment can be continuously monitored in real-time, around-the-clock, seven days a week, thanks to social data sentiment analysis. This enables companies to quickly respond to unfavorable remarks and enhance their brand image while also monetizing positive mentions. Additionally, this research offers businesses a continuous and trustworthy source of information about clients, allowing them to monitor their success across several seasons and base choices on the data.

Social media posts frequently feature some of the most frank opinions on goods, services, and businesses since users freely voice their opinions without being asked. Customers feel obligated to express their emotions to the world, making social media a wealth of candid reviews. [7]

CHALLENGES IN SENTIMENT ANALYSIS

Despite its useful applications, sentiment analysis confronts a number of difficulties that may compromise its accuracy and efficacy. Here are quick summaries of the principal difficulties in sentiment analysis:

Ambiguity and Subjectivity: Dealing with subjective language, which varies from person to person, is a part of sentiment analysis. It might be difficult to comprehend the subtleties and various degrees of sentiment portrayed in

language. Accurate sentiment analysis is further complicated by ambiguity in language, such as sarcasm, irony, or figurative terms.

Understanding Context: Words and phrases can have different meanings based on the situation in which they are employed. To prevent misunderstandings and enhance sentiment categorization, sentiment analysis models must accurately capture contextual information.

Data noise and preprocessing: Spelling, grammar, abbreviation, and slang errors are common in text data gathered for sentiment analysis. Effective preprocessing methods are needed to deal with these noise components since they can negatively impact the performance of sentiment analysis algorithms.

Sentiment analysis models that have been trained on one domain may not function well when used in another area or genre. Researchers are attempting to overcome the difficulty of adapting sentiment analysis algorithms to particular domains or genres, such as product evaluations, social media posts, or news stories.

Handling Negation and Contrast: Words and phrases used in negative sentences can shift the emotional polarity of a sentence. For a sentiment analysis to be accurate, it is crucial to understand negations and how they affect sentiment. Treating contrasting statements in a similar manner conflicting sentiments within a text is another challenge that needs to be addressed.

Data Imbalance: Sentiment analysis datasets often suffer from class imbalance, where one sentiment class has a significantly higher number of samples compared to others. This can lead to biased models that perform poorly on minority sentiment classes.

Language and Multilingual Analysis: Sentiment analysis becomes more complex when dealing with multiple languages. Each language may have its own sentiment expressions, idiomatic phrases, and cultural nuances that require language-specific sentiment analysis techniques.

Lack of Annotated Data: Training accurate sentiment analysis models requires a significant amount of annotated data. However, creating large-scale annotated datasets is time-consuming and expensive, making it challenging to have diverse and representative training data for sentiment analysis.

Evaluation Metrics: Choosing appropriate evaluation metrics to assess the performance of sentiment analysis models is a challenge. The intricacies of sentiment analysis may be missed by popular metrics like accuracy, precision, recall, and F1-score, providing only limited insights into model performance.

Sentiment analysis models may inherit biases from the training set, which could result in predictions that are unfair or biased. The ethical issues surrounding sentiment analysis algorithms and guaranteeing justice in them are a significant difficulty that researchers are attempting to solve.

These difficulties offer chances for additional study and development of sentiment analysis methodology and procedures. The accuracy and usefulness of sentiment analysis in numerous domains are still being improved as researchers continue to investigate novel solutions to these problems.

CONCLUSION

Opinion mining for social good has enormous potential in many areas. It is possible to glean important insights from the analysis of public attitudes and opinions expressed on social media, in online reviews, and on other platforms, helping to shape policy decisions and enhance society as a whole. The purpose of this study article was to examine the uses and effects of opinion mining for societal good.

The research's findings emphasize a number of important topics. First of all, opinion mining can offer useful data that businesses and organizations can use to better understand client feedback, enhance customer satisfaction, and improve their goods and services. Companies may make data-driven decisions that are in line with customer demands and preferences by evaluating sentiment trends and finding areas of concern or improvement. Second, opinion mining can be quite useful in the public sector, helping decision-makers comprehend the attitudes and concerns of the general populace. Governments can better understand the goals and preferences of citizens by examining public opinion on social and political topics. This information helps them create effective policies and programs that satisfy societal needs and ambitions.

Opinion mining can also be used in a variety of industries, including healthcare, education, and public safety. Emerging trends, public health issues, educational difficulties, and problems with public safety can all be identified with the use of this, enabling stakeholders to act promptly and pro-actively. Additionally, opinion mining has the potential to advance accountability and openness. By analyzing opinions and sentiments expressed by the public, it becomes possible to hold organizations, institutions, and individuals accountable for their actions and decisions. This can lead to greater trust and confidence in various sectors of society. It's crucial to recognize, nevertheless, that opinion mining has certain difficulties. A few of the challenges that need to be overcome are privacy issues, biases in data gathering, and the requirement for sophisticated algorithms to manage large-scale data processing. To ensure the ethical and appropriate use of opinion mining tools, researchers, policymakers, and technology developers must collaborate.

In conclusion, opinion mining for societal good has enormous potential to have a constructive influence on many facets of society. Businesses may improve their offers, governments can create policies that are in line with public needs, and decision-makers can make well-informed decisions by leveraging the power of popular mood and opinions. To fully realize the benefits of opinion mining for society, more research, innovation, and cooperation are required.

References

1. Ahmad S, Asghar MZ, Alotaibi FM, Awan I (2019) Detection and classification of social media-based extremist affiliations using sentiment analysis techniques

2. Ahmad SR, Bakar AA, Yaakub MR (2019) A review of feature selection techniques in sentiment analysis
3. Akhtar N, Zubair N, Kumar A, Ahmad T (2017) Aspect based sentiment oriented summarization of hotel reviews
4. Balaji T, Annavarapu CSR, Bablani A (2021) Machine learning algorithms for social media analysis: a survey.
5. Behdenna S, Barigou F, Belalem G (2018) Document level sentiment analysis: a survey.
6. Annett M, Kondrak G (2008) A comparison of sentiment analysis techniques: Polarizing movie blogs. In: Conference of the Canadian Society for Computational Studies of Intelligence
7. Arora A, Chakraborty P, Bhatia M, Mittal P (2021) Role of emotion in excessive use of twitter during COVID-19 imposed lockdown in India.
8. Baashar Y, Alhussian H, Patel A, Alkawsi G, Alzahrani AI, Alfarraj O, Hayder G (2020) Customer relationship management systems (CRMS) in the healthcare environment: a systematic literature review
9. Duric A, Song F (2012) Feature selection for sentiment analysis based on content and syntax models.
10. Baccianella S, Esuli A, Sebastiani F (2010) Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining.
11. Eke CI, Norman AA, Shuib L, Nweke HF (2020) Sarcasm identification in textual data: systematic review, research challenges and open directions.
12. Bai X (2011) Predicting consumer sentiments from online text.
13. Birjali M, Kasri M, Beni-Hssane A (2021) A comprehensive survey on sentiment analysis: approaches, challenges and trends.