

Anatomy of Sentiment Analysis in Ontological, Epistemological, and Axiological Perspectives

M. Noer Fadli Hidayat¹, Didik Dwi Prasetya², Triyanna Widiyaningtyas³, Syaad Patmanthara⁴

^{1,2,3,4} Department of Electrical Engineering and Informatics, Universitas Negeri Malang, Indonesia

¹ Department of Informatics Engineering, Universitas Nurul Jadid Probolinggo, Indonesia

Article Info

Article history:

Received November 11, 2023

Revised September 21, 2024

Accepted November 15, 2024

Published May 10, 2025

Keywords:

Axiological

Epistemological

Ontological

Sentiment Analysis

ABSTRACT

The aim of this article was to examine sentiment analysis methods from the perspective of the philosophy of science with three approaches, ontological, epistemological and axiological. This research used a qualitative research method (descriptive-analysis) with an ontological, epistemological and axiological approach that uses library research and document studies of previous research results. Data collection was carried out through books and reputable scientific journals on Scopus, ScienceDirect, IEEEExplore and Springer Link. The results of this research showed that sentiment analysis from an ontological perspective describes the definition, development and relationship of sentiment with social reality. Meanwhile, from an epistemological perspective, sentiment analysis is viewed from how the source of knowledge is obtained, explaining the production of sentiment analysis knowledge, and several ways of working that can be applied in studies. Axiologically, sentiment analysis can see the function and value resulting from sentiment analysis, as well as discussing the results of interpretation from sentiment analysis studies. These findings showed the development of sentiment analysis in answering various problems to improve the quality of sustainable services in various fields.

Corresponding Author:

M. Noer Fadli Hidayat,

Department of Electrical Engineering and Informatics, Universitas Negeri Malang

Jl. Semarang 5 Malang 65145

Email: fadli@unuja.ac.id

1. INTRODUCTION

The emergence of digitalization has accelerated the expansion of the public sphere in expressing human feelings or opinions on various social media and online platforms [1]. Currently, both experts and the general public try to reach optimal decisions or opinions by utilizing available opinion data. Every online platform, such as an online course platform or social media site, continues to maintain open information so that the level of user satisfaction is widely known [2]. However, one topic on an online course platform has millions of varying opinions. The opinions or feelings expressed may include brief details or even reflect general views, thus requiring more in-depth study [3]. This is the basis of the principles of sentiment analysis which makes it easier to get details about the sentiment of opinion data on websites and classify sentiment patterns in various fields.

The description above shows that sentiment analysis has a very close influence and relationship between various scientific disciplines and developments over time. Therefore, by studying in-depth sentiment analysis, it is hoped that in the end it will be possible to reveal the truth according to the tests carried out. In this case, it is very important to explore sentiment analysis using the perspective of the philosophy of science, as stated by Bakhtiar [4] that the objectives of the philosophy of science include

exploring the basic elements of science, so that we can thoroughly understand the sources, nature and objectives of science, as well as understanding the history of the growth, development and progress of science in various fields, so that we get an overview of the historical process of science.

LO Kattsoff also believes that, "the role of the Philosophy of Science is to explain the nature of science which has many limitations, so that a unified understanding can be obtained regarding various natural phenomena which have become the object of science itself, apart from that the philosophy of science can also train ways of thinking to be more critical" [5]. It is hoped that a basic understanding of the Philosophy of Science will be useful in providing direction and a basis for determining policies that regulate the interests of society in general, as well as those related to the development of science in the future.

The urgency of studying the Philosophy of Science is the main basis for discussing the nature of science, the application of various philosophical methods in an effort to find the roots of problems and find the principles of reality that are questioned by certain fields of science to obtain more definite clarity. In this article, the author will explain the study of the philosophy of science in the context of Sentiment Analysis. Ontological, epistemological, and axiological discussions will be linked to sentiment analysis.

2. METHOD

This study used a descriptive-analytical method with an ontological, epistemological, and axiological approach. This study used previously published research related to the philosophy of science and sentiment analysis. In collecting data, this study used Giuntini's approach model [5]. First, develop criteria for determining which topics to analyze. Second, design a literature search strategy. Third, develop a plan to document and analyze the selected topic.

Literature collection was carried out from 4 October to 7 November 2023 by searching books, reputable scientific journals throughout Scopus, ScienceDirect, IEEEExplore, Jstor, Sage, and Google Scholar published from 2013 to 2023. The collected data were then filtered according to topic, followed by analyzing the relationship between philosophy of science and sentiment analysis from an ontological, epistemological, and axiological perspective, and the final stage was to provide a conclusion. From the results of the initial literature collection, 854 articles were found. After identification, 71 articles were selected that were relevant to the topic of this research, as shown in Figure 1.

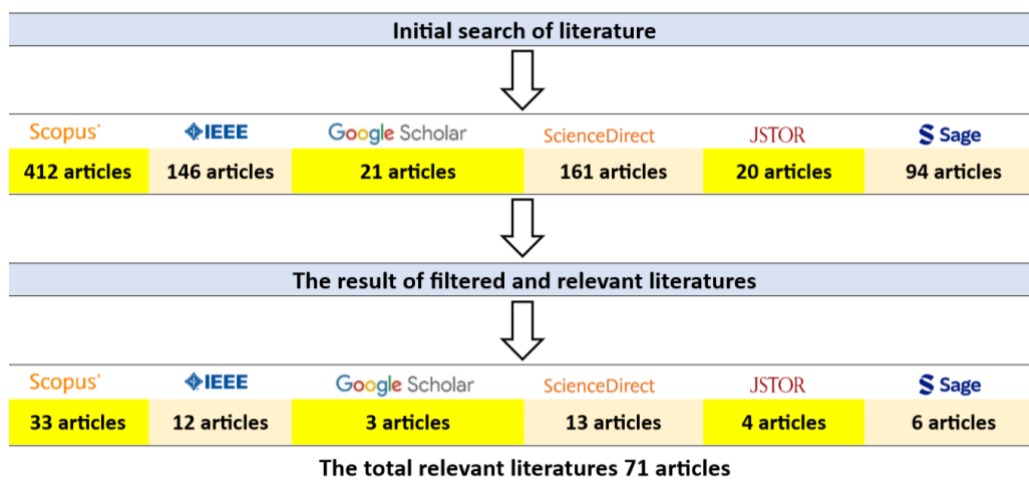


Figure 1. Source platforms

3. RESULT AND DISCUSSION

3.1 Sentiment Analysis in an Ontological Perspective

Ontology is a series of two words, namely On and Logos. On means being and Logos means knowledge. Thus, the meaning of ontology can simply be interpreted as a science that studies the nature of something that is tangible (being) [4]. According to Jujun S. Suriasumantri, ontology discusses what we want to know, how much we want to know [6]. Ontology studies can enter every branch of science, including the study of sentiment analysis.

In relation to sentiment analysis, the ontological approach in this case will reveal about; definition of sentiment analysis, development of sentiment analysis studies, and the relationship between sentiment analysis and social reality.

a. Definition of Sentiment Analysis

The emergence of digitalization has accelerated the expansion of the public's scope for expressing their feelings or opinions on various social media and platforms online [1]. Currently, both experts and the general public try to reach optimal decisions or opinions by utilizing available opinion data. Every online platform, such as an e-commerce site or a social media site, maintains a level of transparency, which in turn increases the likelihood of its influence on other users [2]. However, a single topic or item can have millions of varying opinions on one platform. The opinions or feelings expressed may include brief details or even reflect general views, thereby increasing the research community's interest in conducting more in-depth studies. All of this is the basis of the principles of sentiment analysis, also known as opinion mining. Sentiment analysis makes it easier to get details about sentiment, analyze opinion/sentiment data on the web, and classify sentiment patterns in areas, such as business [7], [8], [9], banking [10], politic [11], as well as the government sector [12], [13].

Definition of sentiment analysis according to Bordoli [14] is a method that can make it easier to obtain sentimental details, analyze web opinion/sentimental data, and classify sentimental patterns in various situations. Ghosh argues that sentiment analysis is a representation of various opinions from various entities, such as events, problems, aggression, anger, attitudes, and so on. Sentiment Analysis attempts to classify an individual's feelings and opinions into three main categories: positive, negative, and neutral [15]. Xiong, in his research, revealed that sentiment analysis is a science developed in the field of natural language processing (NLP), with the aim of extracting user attitudes towards products from text comments that reflect the user's subjective awareness, as well as providing more accurate emotion classification results [16].

Sentiment analysis can be explained as a procedure for identifying, recognizing, and/or categorizing users' emotions or opinions towards various services, such as films, products, events, or certain attributes, either as positive, negative, or neutral [17]. Cortis has referred to various studies in which sentiment analysis is considered as a classification or regression task. When analyzing feelings by attributing feeling scores into the range [-1, 0, 1], it was found that there are situations where prediction is sometimes treated as a classification task, while at other times as a regression [18].

From these definitions, it can be concluded that sentiment analysis is a field of science developed in natural language processing (Natural Language Processing) which is used to analyze feelings from various data in the form of text, video, icons, sounds, or other data in order to obtain accurate results. When sentiment is expressed as polarity in the context of computational linguistics, it is generally treated as a classification task. When the sentiment score is within a certain range and is used to express emotions, this task is considered a regression problem. Because sentiment analysis has become a very important field of science in this digital era, this science continues to evolve very quickly.

b. Development of Sentiment Analysis Studies

Since 2008, sentiment analysis has become an increasingly popular research field, as can be seen from the number of articles published and documented in various research databases such as Elsevier's ScienceDirect, IEEE Xplore Digital Library, Springer Link, or Wiley Online Library. In the last five years, more than 8,000 patents have been issued, and more than 2,300 articles have been published over the last fifteen years [19].

In this section, the development of sentiment analysis studies is focused on the concept of sentiment analysis in social networks, which is growing very significantly with a growth rate of around 34% from year to year. From various research conducted with a focus on social networks until 2023, the reviews discussed are divided into two main issues, namely: the techniques used, namely machine learning (deep learning) or lexicon-based methods; and problems in specific application areas such as

education, marketing, election outcome prediction, etc. Scheme and classification of the structure of social network analysis sentiment articles as presented in Figure 2.

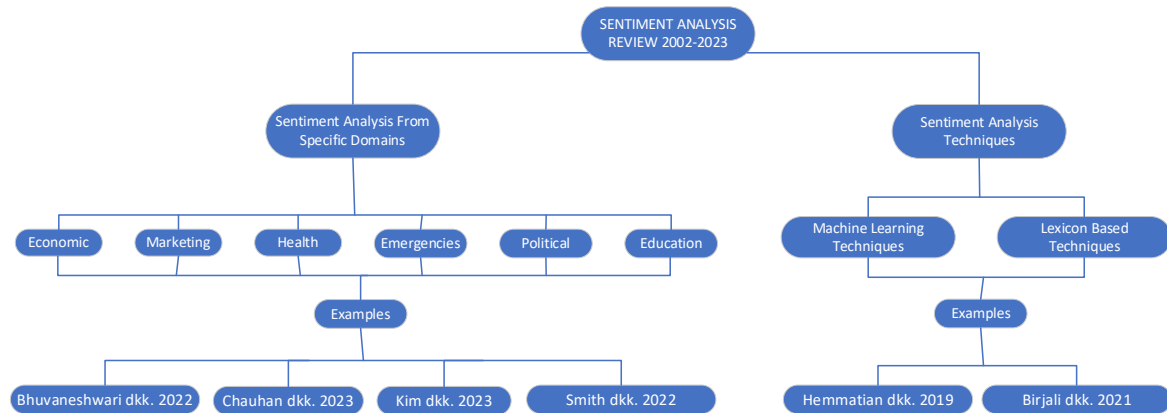


Figure 2. Development of sentiment analysis in 2002 – 2023

Preethi, in her research, has summarized more than a hundred articles published between 2002 and 20015. From the results of this summary it is stated that sentiment analysis research focuses more on discussing various analytical approaches and general issues [20]. In 2016, Balazs and Velazquez [21] highlights the relevance of opinion mining and the incorporation of information that can be found in sentiment analysis.

Birjali et al in 2017 conducted research that focused on sentiment analysis methods, application domains, and emerging challenges [22]. Apart from that, subjectivity analysis has been analyzed by Chaturvedi in 2018 [23]. Hemmatian and Sohrabi [24], in 2019, conducted a more specific review regarding classification techniques for opinion mining, sentiment analysis, and characteristic extraction. In contrast, Ashima et al [23] in 2019 summarized various significant and extensive approaches in the field of sentiment analysis using deep-learning approaches. The study compiled a survey of 130 research articles and provided a detailed overview of the most popular deep-learning techniques at the time.

In 2020, Morone Birjali et al [22] conducted sentiment analysis research that reviews commonly used methods, software, or techniques used in sentiment analysis (machine learning, lexicon-based, hybrid, etc.). This article also describes in detail the advantages and disadvantages of these natural language processing techniques, focusing on sentiment analysis in the context of Twitter datasets. Apart from that, there are also other studies that focus on various models, such as polarity on Twitter by Singh [25], predictions about the opinion of the UK and EU parliament regarding Brexit by Chandio & Sah, and predictions about the price movements of nine cryptocurrencies by Kraaijeveld & Smedt [26].

And since 2020, as a result of the Covid-19 outbreak, sentiment analysis in the world of education has also developed rapidly, especially regarding online learning platforms as a new revolution in the world of education and continues to develop rapidly during the pandemic [27] in various European and Asian countries in Europe and Asia [28], [29], [30] to date. Online learning platforms that are currently popular in the world include Coursera, edX, Udacity, and Udemy [31]. Meanwhile, in Indonesia, the Ruangguru platform is considered one of the more innovative learning platforms in technology-based educational content [32]. Satria [33] developed an efficient and effective method for gathering opinions and understanding people's perceptions of online learning applications in Indonesia using the programming languages Logstash, Kibana, and Python (ELK stack) and the Naïve Bayes algorithm with the result that 76.6% of tweet reviews had positive feelings towards learning on line. Then, Su [34] developed the Hierarchy Attention Mechanism method to identify the polarity of sentiment of users of online course platforms at universities in China to assist teachers in improving their teaching models. And recently, Kim et al [35] proposed the Valence Aware Dictionary for Sentiment Reasoning (VADER) method to analyze students' feelings through the Student Assessment of their Learning Gains (SALG) survey to make it more concise and efficient.

Sentiment analysis has a close relationship with social reality because it involves understanding and measuring human feelings, opinions, and sentiments towards various aspects of social life. Some evidence that sentiment analysis is related to social reality can be seen from several studies that develop sentiment analysis in the business sector to improve product commercialization and increase customer satisfaction. Like research conducted by Jung et al [36] in developing a product quality evaluation model that combines deep learning, word vector conversion, keyword clustering, and word feature extraction technology. The model improves product features based on consumer reviews and then calculates customer satisfaction by paying attention to short text comments using sentiment tags. Bhuvaneshwari et al [37] proposed a Bi-LSTM Self-Attention based CNN (BAC) model for user review analysis using pre-learned word embeddings. Wang et al. [38] designed a multi-attention bi-directional LSTM (BLSTM) and used Latent Dirichlet Allocation (LDA) modeling to perform multimodal fusion in product review sentiment analysis. Alantari et al. [39] analyzed 260,489 reviews from platform users covering 25,241 products from nine different product categories. In this research, it was found that previously studied neural network-based machine learning techniques, in particular, provided the most accurate predictions, while topic models such as Latent Dirichlet Allocation (LDA) provided a more thorough diagnosis.

Sentiment analysis also has a close relationship with the social reality of politics and government. Recently, researchers have started to openly share their views or opinions about various political parties, election candidates, government policies, and regulations on various public platforms such as Twitter, Facebook, YouTube, and blogs. This has the potential to have a significant impact on their followers. Therefore, these platforms are used by many experts to predict the results of previous elections, monitor people's feelings about various political movements, or analyze the public's views on government regulations that are being drafted. In one study, Antypas et al. [40] used multilingual sentiment analysis models and Chi-square tests to analyze less popular politician tweets than Greek, Spanish, and English parliamentarians. Their study showed that tweets with negative sentiment spread faster than positive tweets. In another study, Chauhan et al. [41] used the Valence Aware Dictionary and Sentiment Reasoner (VADER) to analyze two million tweets related to India's 2019 Lok Sabha Elections to understand Twitter users' sentiments towards Indian political parties. Meanwhile, Yavari et al. [42] designed indicators of future election results using an aging estimation method based on the ratio of positive messages to the ratio of negative news. All of this research has proven how sentiment analysis can be used to understand social reality and people's feelings in responding to various political issues.

Not only that, the relationship between social analysis and social reality is also clearly illustrated by the use of sentiment analysis, which has been used to analyze the expression of people's opinions about various social events, which has increased recently. This has expanded the scope of developing public sentiment analysis, monitoring, and predicting various outcomes related to events or their causes, as well as helping in formulating steps that need to be taken in the future, especially when facing situations full of uncertainty. One example of the development of a sentiment analysis method to analyze events and situations using social media data was carried out by Ouyang et al. [43]. This research was conducted to develop a multi-grained feeling analysis method and event summary using social media data related to incidents such as accidents and explosions. The system created is able to identify event components that attract users' attention, recognize which reviews play a role in changes in public sentiment, and detect aspects of events that influence user ratings. Meanwhile, research conducted by Smith and Cipolli [44], used a data corpus consisting of 8,013 tweets to analyze policies with statistical models and extract topics that were still considered abstract in the review. These two studies illustrate how sentiment analysis plays an important role in understanding people's feelings and views regarding various social issues and events, and can assist in future decision-making.

3.2 Sentiment Analysis in an Epistemological Perspective

Epistemology comes from the Greek words episteme and logos. Episteme can be defined as knowledge or truth, and logos is defined as thoughts, words, or theories. Etymologically, epistemology can be interpreted as the theory of true knowledge and is usually called the theory of knowledge. The philosophy of science can also be referred to as the Theory of Knowledge, namely the theory of gaining knowledge or processing knowledge to become in-depth and systematic [45]. Epistemology, in general, discusses all processes in an effort to obtain the truth of knowledge [46]. Apart from that, epistemology can also be explained as a branch of philosophy that deals with gaining knowledge and human ways of

thinking in determining and gaining knowledge by using various abilities embedded within a person, for example the abilities of the senses, intuition and also reason. The following will explain how epistemology works as a branch of the philosophy of science in the context of sentiment analysis studies [47].

a. Sources of Knowledge in Sentiment Analysis

Knowledge sources in sentiment analysis refer to data or information used to analyze feelings and opinions in texts or social contexts. Data and information are the main sources in sentiment analysis, and can be in the form of text, symbols, images, sound, video, or other forms obtained from various sources such as product reviews, social media posts, news articles, blogs, and others. This data can be used to identify and analyze the feelings and opinions contained therein. The data collected in large quantities can then be used as a dataset. In sentiment analysis, datasets are a source of basic knowledge for training and evaluating various sentiment analysis techniques, models, or methods. Although some studies use datasets collected themselves, for example, using scraping techniques to collect reviews and the Ruangguru online course platform. A number of sentiment analysis datasets have also become publicly available and have proven to be valuable resources. Some examples of these datasets include the Internet Movie Database (IMDb), Twitter US Airline Sentiment, Sentiement140, and the SemEval-2017 Task 4 dataset.

The Internet Movie Database (IMDb) is a comprehensive and balanced collection of film reviews, with over 50,000 film reviews. This dataset has been divided into two equal parts; 25,000 reviews are classified as positive reviews, and half are labeled as negative reviews [48]. This dataset is interesting to apply to sentiment analysis models because the reviews in IMDb include film stories and personal opinions, which makes it have its own complexity if applied to sentiment analysis models, so as to produce accurate analysis. This dataset consists of two columns, namely "review" and "sentiment", and the complexity of the language used in reviews adds a level of difficulty to the sentiment analysis model [49].

The US Airline Sentiment Twitter dataset, collected in 2017 by CrowdFlower, provides a collection of customer reviews from six major American airlines. This dataset is a valuable source of knowledge for evaluating the performance of sentiment analysis models in real-world situations. Therefore, it is of great interest to researchers in this field. Sentiment classes in this dataset include positive, negative, and neutral, with sample sizes of 2,363, 9,178, and 3,099, respectively [50]. One of the main challenges facing this dataset is label imbalance, where the majority of samples fall under negative sentiment labels, which can affect the accuracy of sentiment analysis models. In addition to the label imbalance problem, tweets in the dataset are often short and use informal language, which poses additional challenges for sentiment analysis models. Tweets tend to have an informal writing style, so context and keywords related to sentiment can often be overlooked [51]. These challenges can result in sentiment analysis models misclassifying Tweets, making it difficult to determine the true feelings expressed in these texts.

The Sentiment140 dataset collected by Stanford University from Twitter is a large and comprehensive collection of customer sentiment data. This dataset consists of 1.6 million samples evenly split between positive and negative sentiment labels. The short and informal nature of Tweets in the Sentiment140 dataset can pose challenges for sentiment analysis models [52]. Tweets are typically written in a more random and concise style compared to other forms of text, which can lead to ambiguity in sentiment polarity. This makes it difficult for sentiment analysis models to identify appropriate and accurate feelings, because keywords that are important for expressing sentiment or information context are often overlooked. Apart from being short and informal, the Sentiment 140 dataset also has special complexity in sentiment analysis models because it reflects real-world situations [53]. Customer sentiment data collected from social media platforms like Twitter provides valuable insight into how customers view and interact with companies and their products [54]. The Sentiment 140 dataset provides an opportunity for researchers and practitioners to evaluate the performance of various sentiment analysis methods.

Some examples of datasets above are sources of knowledge in sentiment analysis, and there are many other sources of knowledge that come from various sources, for example: 1) Social media sources

such as Twitter [55], YouTube [38], Facebook [56], and platforms that are also frequently used to understand public feelings and opinions on various topics, including politics, brands, or important events; 2) Surveys and questionnaires are used to collect opinions and feelings from respondents [57]. The results of this survey can be used in sentiment analysis to gain an understanding of people's views on a topic; 3) Social network analysis involves the study of relationships and interactions between social entities [58]. This data can be used to see how influence and opinions spread through social networks; 4) Sensor Data, such as Internet of Things (IoT) devices, can be used to collect data about feelings; 5) Geographic Data includes location or geolocation information. It can be used to analyze feelings that vary by geographic location, as well as to understand the relationship between geography and feelings; 6) Multi-Modal Data combines text with other data such as images, sound, or video [59]. This enables richer and more contextual sentiment analysis; 7) Historical Data includes data that has been collected over a period of time, and this can be used to see time trends in feelings. This helps in understanding how feelings have changed over time; 8) The source of the Sentiment Lexicon is a collection of words or phrases that have sentiment polarity [60]. This is used as a reference in analyzing text sentiment by detecting words contained in the sentiment dictionary; 9) Public data is information that is publicly available, including government data, business data, and other open resources that can be used in sentiment analysis [61]. Table 1 shows the use of datasets in sentiment analysis studies.

Table 1. The use of datasets in sentiment analysis studies

Article	Year	Dataset	Model	Accuracy	Precision	Recall	F1 Score	AUC
Dahir et al. [49]	2023	IMDB Movie, 10,000 film reviews	SVM, TF-IDF	89,20%	88,50%	90,2%	-	89,30%
Tusar et al. [50]	2021	Twitter US Airline Sentiment, 14,640 tweet	Random Forest, BoW	74%	73%	74%	73%	-
Hasib et al. [51]	2021	CrowdFlower Twitter, 14,640 tweet	DNN, CNN	91%	85,66%	87,33	87,66 %	-
Tan et al. [54]	2023	Sentiment140	BERT, GRU	89,59%	90%	90%	90%	-
		IMDb (50,000 film reviews)	BERT, GRU	94,63%	95%	95%	95%	-
		US Airline Twitter (14,160 tweets)	BERT, GRU	91,52%	91%	91%	91%	-
Wang et al. [38]	2022	2050 comments from 93 youtube videos	Multi-attention bi-directional LSTM	81,5%	-	-	80,6%	-
BryanSmit h et al. [55]	2023	Twitter data about floods and weather totalling 795,065 tweet	Contrastive Language-Image Pretraining (CLIP)	84,1	81,7%	79,1	80,2%	-
Haque et al. [56]	2023	42,036 Facebook comments 282,000	CNN, LSTM	85,5%	86%	85%	85,85 %	-
El Barachi et al. [58]	2021	responses to the 40 most popular tweets	Bi-directional LSTM	89,90%	90,34%	90,54 %	90,44 %	-
Asif et al. [60]	2020	1,800 positive/negative English lexicon	Multimodal Naïve Bayes, Linear Support Vector Classifier	88%	87%	93%	90%	-

b. Knowledge Processes in Sentiment Analysis

The evolution of the science of sentiment analysis is marked by the development of various models from experts and researchers. After more than 500 proposed sentiment analysis models, the general knowledge process in sentiment analysis has been defined in four modules, along with additional optional modules that help in understanding and interpreting sentiment contained in text or data. Each module has the following functions: data collection and standardization, pre-processing of the dataset, extraction of features or keywords that represent the entire dataset, prediction or

classification of feelings related to keywords, entire sentences or documents according to needs, as well as a summary of feelings comprehensively connected to the dataset, as in Figure 3.

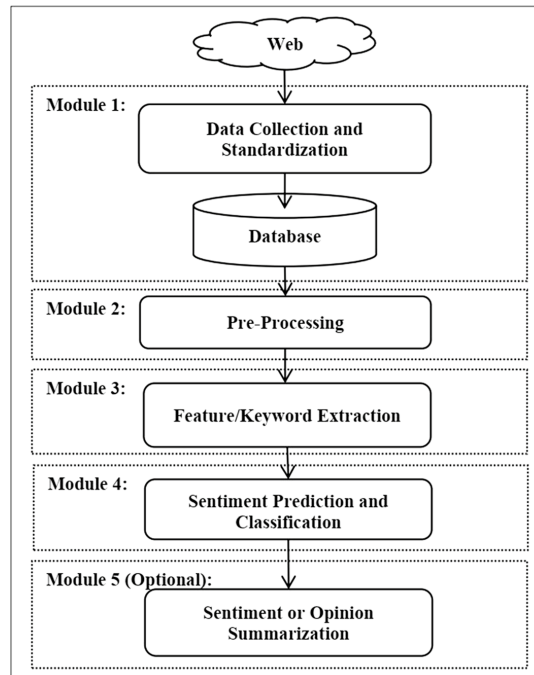


Figure 3. Sentiment Analysis Knowledge Process

Module 1: data collection and standardization

The data collected is critical to the performance of the sentiment analysis model. Many studies rely on a variety of public data sources to design sentiment analysis models. Like the Blitzer multi-domain sentiment dataset [62] used to classify sentiment from reviews of products used by users [63]. Epinions (<http://www.epinions.com/>) is also the choice of several researchers in analyzing product reviews [64], [65]. The UCI Machine Learning Repository provides standard datasets for sentiment analysis, such as Twitter data for Arabic sentiment analysis, sentences labeled with sentiment, journal reviews, sentiment analysis of distance education in Saudi Arabia during the Covid-19 Pandemic, and so on. Due to the high production of data, it is necessary to design a system that is capable of updating the database periodically to avoid generalizations or biases related to certain times. Therefore, automating large data collection is a very important thing to do. Currently, several applications in the form of Application Programming Interface (API) have emerged to help collect data from social media or e-commerce platforms, such as Twitter [66], youtube [67] and others. Most API applications today facilitate real-time data collection.

For data sets representing specific topics, proper standardization of type, format, and context is critical to improving overall analysis results. Appropriate labels can also improve the performance of sentiment analysis models. Currently, there are various methods for labeling data including Word Splitting, Pattern Analysis, Language Modeling, Emotion Analysis, Aspect-Based Analysis, Entity-Based Sentiment, Machine Learning or combining these methods. However, keep in mind that data labeling methods often suffer from noise problems, which can reduce accuracy.

Module 2: Pre-processing

Pre-processing is removing all types of noise from text datasets and preparing clean, relevant, and structured datasets for sentiment analysis purposes. Proper pre-processing of any dataset has a significant impact on the sentiment analysis results. In the Indonesian language text dataset Antonius [68] carry out pre-processing to explore, clean and prepare the dataset so that it is ready for classification by means of; 1) remove words that do not have meaning such as dates (e.g., "11/16/20"),

special characters (@, #), and words without meaning (e.g., "a+," "a-," "b+"); 2) perform case folding, where all alphanumeric characters are changed to lowercase; 3) carry out tokenization, namely parsing sentences into words using spaces as separators. Tokenization is carried out in two forms, namely 1 gram and 2 grams; 4) perform text normalization, all tokens are converted to "normal" tokens. This process is managed using the Sklearn library; 5) remove filler words (stopwords) from the dataset using the Indonesian stopword list.

To further increase the accuracy value in Indonesian language sentiment analysis, other researchers added a text tagging system in the tokenization process using the Big Indonesian Dictionary (KBBI) as a lexical dataset [69]. Ghulam [70] also pre-processed the Indonesian language dataset using a parsing model to break down documents into word strings and then analyzed the collection of words by separating and determining the syntactic structure of each word. In this pre-processing, various algorithms are used, including Python, including Pandas and OpenPyXL for dataset manipulation, Matplotlib and Seaborn for graphic visualization and graphs, Tqdm for bar displays, and Sklearn and NLTK for text manipulation.

Module 3: Feature/Keyword Extraction

Not every word contained in the dataset is always have relevant in the context of sentiment analysis. The ability to perform accurate sentiment classification has been enhanced through ongoing research and knowledge. Even though the dataset has been cleaned with various pre-processing steps, the use of data in the dataset always raises problems, such as dimensional problems, longer computing time, and the use of irrelevant or less significant features or terms, especially if the data has high dimensions and is multivariable, the problems that arise will become more prominent. So, to build sentiment analysis dynamically based on the required words, a keyword extraction process is needed [71].

According to Benghuzzi and Elsheh [72], Keyword extraction is a technique used to extract important features or terms from text data by finding certain words, phrases, or terms in a document, so that the presentation of the document becomes more concise. The subject of the text can be thoroughly studied and evaluated once the keywords of the text are extracted correctly. This can help make better decisions about texts. In recent years, many researchers have used automatic keyword extraction as a research tool because the manual process of extracting keywords from large databases is time-consuming, expensive, and repetitive. There are three categories of automatic keyword extraction methods: supervision-based methods, semi-supervision-based methods, and unsupervised methods [73].

Module 4: Sentiment Prediction and Classification

So far, various techniques have emerged to support sentiment prediction and classification purposes in sentiment analysis. Some researchers choose techniques based on their availability, challenges faced, or even general themes of sentiment analysis. Affective computing can be carried out either by utilizing knowledge-based techniques, statistical methods, or hybrid approaches [74]. Knowledge-based techniques are used to categorize texts into influencing categories by referring to sources of popular words that have an emotional impact or multi-word expressions. This is based on the appearance of words that influence sentiment, such as "happy," "disappointed," "regret," and so on. Meanwhile, statistical methods use an effectively annotated training corpus to determine the validity of keywords through word co-occurrence frequency, as well as the validity of keywords with other random model approaches, and so on. A hybrid approach, combining knowledge-based linguistic patterns and statistical methods to infer polarity from texts in a more comprehensive way.

Medhat [75] has reviewed classification techniques in sentiment analysis in a systematic and illustrative manner. This research also explains current sentiment predictions and various types of classification methods. Through the development of sentiment analysis, many models have been proposed by experts, and in general, the sentiment analysis framework is presented in Figure 4.

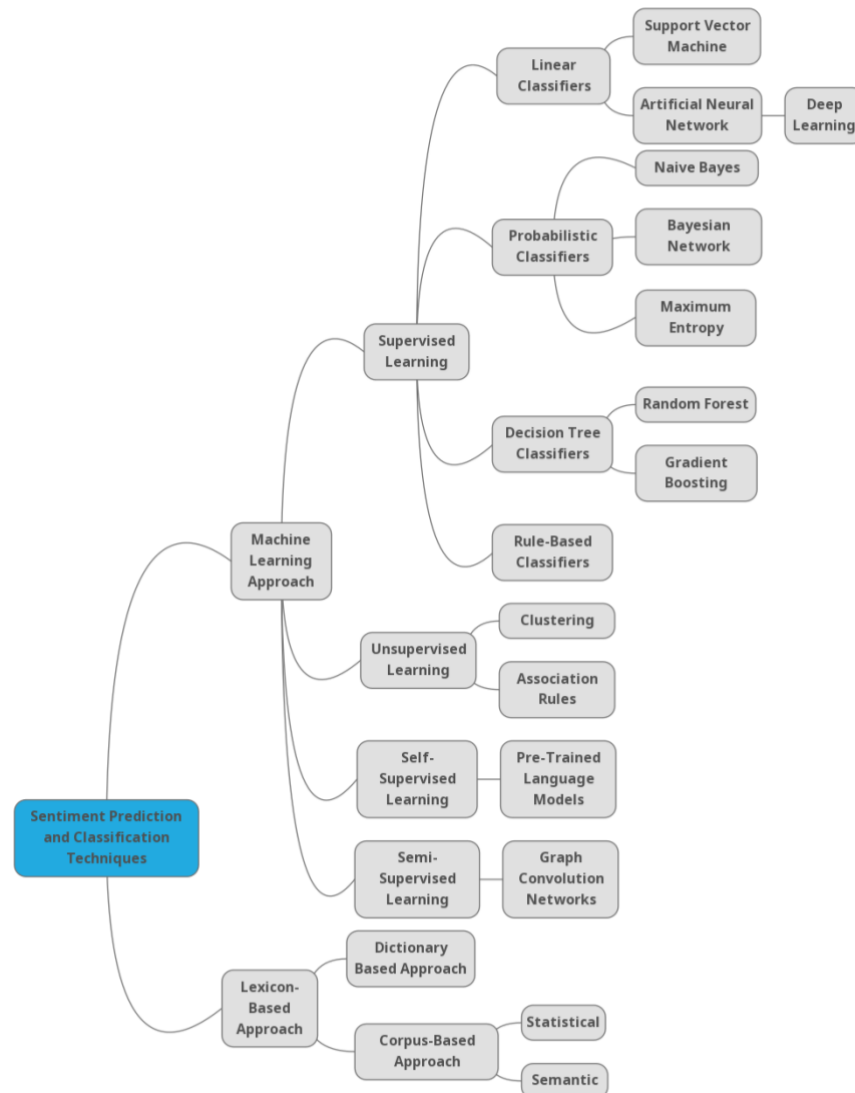


Figure 4. Framework and Development of the Sentiment Analysis Model

Module 5: Sentiment or Opinion Summarization

Sentiment or Opinion summarization aims to provide an overview of the influence of polarity reflected in the dataset by summarizing the polarity of all words, aspects, sentences, or documents. Keep in mind that Sentiment summarization is different from text summarization, although the two are related. Text summarization aims to provide a summary of the contents of the dataset, while sentiment summarization provides an overview of the polarity expressed by the entire dataset.

Researchers have discussed various types of dataset extraction models to get an even picture of sentiment. For example, Pang and Lee [76] take the approach of extracting all subjective sentences and summarizing them. Blair-Goldensohn et al. [69] used a tool to select a few representative documents from a large number of documents, then used these documents to summarize emotions based on aspects (elements, topics, or categories). Hu and Liu [77], instead, suggest an aspect-based sentiment extraction model for online consumer reviews by extracting product rating feedback opinion features.

Besides that, Antonius [78] detected spam comments in Indonesian using the Emoji-Text Feature with the SpamID-Pair data set and was able to increase the detection of Indonesian language spam comments on the 14 latest ML algorithms by an average of 4% to 6%. Naglaa [79] applies emojis to the largest lexicon in Arabic to understand people's feelings towards the pandemic. Meanwhile,

Deepak uses The Amalgam Neural Architecture which is trained using word and emoji embedding to detect sarcasm with the Indian language Sarc-H dataset which is rich in different morphological and linguistic characteristics [80]. Liu also developed an improved emoji-embedding model based on Bi-LSTM architecture called Cemo-LSTM. This model is used to analyze Chinese language texts which have diversity in sentence structure and complex semantics like Indonesian.

c. Limitations of Sentiment Analysis studies

Like scientific fields in general, sentiment analysis also has limitations which are considerations for continuing to develop this scientific field, including:

Accuracy: although sentiment analysis models are improving, they are not always perfect. Sentiment classification may in some cases be inaccurate, especially if the text is very ambiguous or uses figurative language. For example, machine learning (ML) techniques offer superior accuracy and classification results. However, it still has significant limitations, such as requiring a long processing time, being dependent on the domain, and requiring human involvement and information labels [81]. Deepak's research results use The Amalgam Neural Architecture (LSTM-SVM) which is trained on word and emoji embeddings to detect sarcasm with the Indian language Sarc-H dataset which is rich in different morphological and linguistic characteristics. This research shows better accuracy performance results, namely 97.01% than using an Artificial Neural Network (ANN) with an accuracy value of only 92.72%, while the Long Short-Term Memory (LSTM) model produces an accuracy performance of 96.03% [80].

Language and Culture: A sentiment analysis model trained in one language or culture may not be applicable to another. Local language, metaphors, or culture can influence sentiment analysis. Like sentiment analysis in Indonesian, it has different rules, forms and changes in word formation and morphological structure when compared to English. The use of prefixes and suffixes in Indonesian can change the meaning [12]. Therefore, recent sentiment analysis research utilizes emoticons or emojis from language reviews that have complexity in word structure and morphology to increase accuracy, identification and sentiment classification [82]. Naglaa [79] also developed sentiment analysis in Arabic using emojis in the lexicon to understand people's feelings towards the pandemic. Liu developed an improved emoji-embedding model based on Bi-LSTM architecture called Cemo-LSTM. This model is used to analyze Chinese language texts which have diversity in sentence structure and complex semantics like Indonesian. Limitations of sentiment analysis are presented in table 2.

Table 2. Limitations of sentiment analysis in several studies

Article	Year	Language	Method	Limitations
Basiri et al. [81]	2021	English	Attention-based Bidirectional CNN-RNN Deep Model (ABCDM)	The accuracy of long tweets is higher than short tweets
Jain et al. [80]	2022	Indian	LSTM-SVM	Ignore hashtags and sarcasm in reviews
Madyatmadja et al. [12]	2022	Indonesian	Naïve Bayes, SVM-Polynomial	Labeling is done manually
Abdelhady et al. [79]	2023	Arabic	Random Forest (RF) and Support Vector Machine (SVM)	Labeling is done manually

In an epistemological perspective, the source of sentiment analysis knowledge is a dataset in the form of text, symbols, images, sound, video or other forms from various sources of product reviews, social media posts, news articles, blogs, which are collected independently or are publicly available Internet Movie Database (IMDb), Twitter US Airline Sentiment, Sentiement140, and the SemEval-2017 dataset Task 4. The sentiment analysis knowledge process has generally been defined in 5 modules, namely data collection and standardization, pre-processing, feature extraction, sentiment prediction or classification, and sentiment summary. Meanwhile, the limitations in sentiment analysis lie in the imperfect accuracy values, and models applied in one language cannot be applied to other languages.

3.3 Sentiment analysis in an axiological perspective

The word axiology comes from Greek, which consists of two words, axios, which means worthy or appropriate, and logos, which means science or study [47]. Apart from that, value also comes from the Latin word Valere, which means useful, capable, powerful, applicable, or strong, which means

the quality of something that makes it likeable, desirable, useful, or an object of interest. But it can also mean something that is appreciated, valued highly, or appreciated as a good thing [83]. Thus, axiology is a branch of philosophy that studies the values or norms of a science, knowledge that investigates the nature of values, generally viewed from a philosophical point of view. The following will explain how axiology works as a branch of the philosophy of science in the context of sentiment analysis studies.

a. Functions/benefits of sentiment analysis studies

Technological advances such as Blockchain, IoT, Cloud Computing, and Big Data have expanded the scope of applications of sentiment analysis, making it possible to use it in almost all scientific disciplines, as shown in Figure 5.

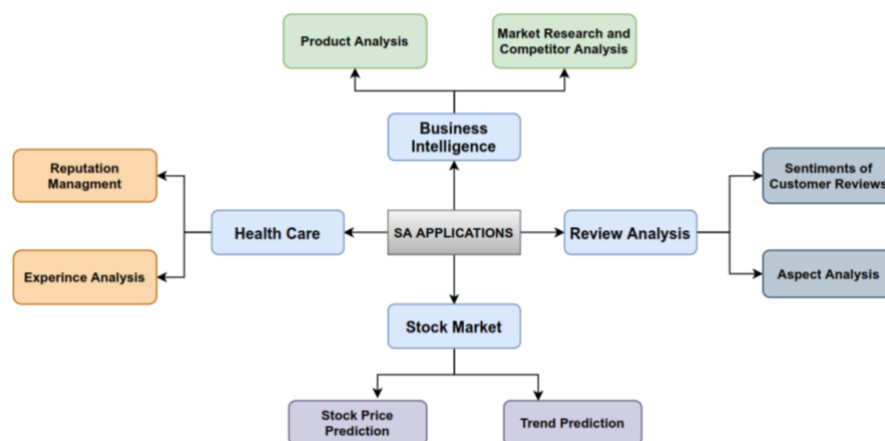


Figure 5. Scope of Sentiment Analysis Applications

Sentiment analysis in the field of business intelligence has provided a number of significant benefits. In this context, companies can leverage data from sentiment analysis to improve their products, explore customer feedback, and design innovative marketing strategies. One of the most common uses of sentiment analysis in the field of business intelligence is to analyze how customers rate the services or products they offer [84]. It should be noted that this analysis is not limited to product manufacturers; Consumers can also use it to evaluate products and make more informed decisions. Hassan et al [85] observing reviews of food services on the Amazon.com site, the results showed that sentiment analysis can be used to identify customer behavior and risks, while increasing customer satisfaction levels.

Sentiment analysis in the health care sector continues to increase in use, this application has helped health service providers in collecting and evaluating patient moods, epidemics, adverse drug reactions, and disease understanding to improve health services [86]. On the other hand, research by Clark et al [87] used Twitter tweets containing patient experiences as additional data to analyze public health. Over the course of one year, they managed to collect around five million tweets related to breast cancer using the Twitter Streaming API. After pre-processing, these tweets are classified using standard LR classifiers and CNN models. The results show that positive experiences in care, social support, and increased public awareness are all interrelated. Overall, applying sentiment analysis to analyze data generated by patients on social media can help in determining patients' needs and views on healthcare.

Sentiment analysis in the entertainment sector has also been widely applied. Reviews of a film or show can be analyzed to determine the response from the audience [88]. This not only helps viewers make better choices, but also helps quality content gain high popularity. Sentence level is one model that is often used in sentiment analysis in this field, to determine the overall feeling of an accurate review [89]. The travel industry also discusses categorizing consumer decisions into positive or negative scores based on online reviews from consumers [90].

Stock market sentiment analysis is used to predict stock prices by analyzing all news related to the stock market to predict stock price trends. Data can be obtained from various sources, including Twitter, news articles, blogs, and so on. The results of research conducted by Xing et al [91], This method

is used to determine whether the stock price trend will rise or fall. The results show that positive news tends to increase stock price trends, while negative news tends to influence stock price trends to decrease. Sentiment analysis is also used to analyze the influence of news sentiment on the volatility, volume and returns of cryptocurrencies such as Bitcoin and digital currencies [92].

b. Results of interpretation of sentiment analysis

The result of interpreting sentiment analysis is an understanding of feelings, opinions or sentiments extracted from certain texts or data using sentiment analysis techniques [88]. This interpretation process produces an understanding of how people respond or feel about a topic or entity. Some important points related to the results of interpreting sentiment analysis include understanding the type of sentiment present in the text or data. These sentiments can be positive, negative, neutral, or even specific such as happy, angry, sad, or happy. Interpretation will help classify these sentiments [52]. Apart from simply determining the type of sentiment, interpretation can also measure the intensity of the sentiment. For example, to understand the extent to which people feel positive or negative about a topic [44].

Interpretation results also help in understanding how society or a particular audience responds to a topic or event. This is an important insight in many aspects, including business, politics, and academic research. So that the results of the interpretation of sentiment analysis can provide valuable feedback to companies, governments, or individuals in making better decisions or can help in marketing planning or product improvements based on understanding public feelings [93].

4. CONCLUSION

Sentiment analysis in an ontological perspective describes the definition, development and relationship of sentiment with social reality. Meanwhile, from an epistemological perspective, sentiment analysis is viewed from how the source of knowledge is obtained, explaining the production of sentiment analysis knowledge, and several ways of working that can be applied in studies. Axiologically, sentiment analysis can see the function and value resulting from sentiment analysis, as well as discussing the results of interpretation from sentiment analysis studies.

From 2002 to 2023, sentiment analysis is developing very rapidly in various fields such as economics, marketing, politics, health and education with increasingly diverse methods (machine learning, lexicon-based, hybrid and others), while in an epistemological perspective, the source of sentiment analysis knowledge is datasets in the form of text, symbols, images, sounds, videos or other forms from various sources of product reviews, social media posts, news articles, blogs, that are self-collected or publicly available, e.g., Internet Movie Database (IMDb), Twitter US Airline Sentiment, Sentiement140, and SemEval-2017 Task 4 datasets.

The sentiment analysis knowledge process has generally been defined in 5 modules, namely data collection and standardization, pre-processing, feature extraction, sentiment prediction or classification, and sentiment summary, while the limitations in sentiment analysis lie in the imperfect accuracy value and the model applied in one language cannot apply to other languages.

The results of sentiment analysis interpretation can also provide people with an understanding of the type of sentiment present in text or data, measure sentiment intensity, and provide valuable feedback to companies, governments, or individuals in making better decisions. This research still does not fully discuss the anatomy of sentiment analysis from an ontological, epistemological and axiological perspective. For example, discussions about theories of sentiment analysis from an ontological perspective, design of sentiment analysis algorithms from an epistemological perspective, as well as ethics and academic honesty in the use of sentiment analysis from an axiological perspective. This can be a space for development for other researchers who want to study similar topics by referring to the results of this research from different points of view.

ACKNOWLEDGEMENTS

Thank you to the State University of Malang for its support to researchers to carry out this research. I would also like to thank the lecturers who have supported this research until completion.

REFERENCES

- [1] D. Antypas, A. Preece, and J. Camacho-Collados, "Negativity Spreads Faster: A Large-Scale Multilingual Twitter Analysis on the Role of Sentiment in Political Communication," Feb. 2022, doi: 10.1016/j.osnem.2023.100242.
- [2] L. Yang, Y. Li, J. Wang, and R. S. Sherratt, "Sentiment Analysis for E-Commerce Product Reviews in Chinese Based on Sentiment Lexicon and Deep Learning," *IEEE Access*, vol. 8, pp. 23522–23530, 2020, doi: 10.1109/ACCESS.2020.2969854.
- [3] Z. Kastrati, A. S. Imran, and A. Kurti, "Weakly Supervised Framework for Aspect-Based Sentiment Analysis on Students' Reviews of MOOCs," *IEEE Access*, vol. 8, pp. 106799–106810, 2020, doi: 10.1109/ACCESS.2020.3000739.
- [4] Amsal Bakhtiar, *Filsafat Ilmu*. Raja Grafindo, 2017.
- [5] F. T. Giuntini, M. T. Cazzolato, M. de J. D. dos Reis, A. T. Campbell, A. J. M. Traina, and J. Ueyama, "A Review on Recognizing Depression in Social Networks: Challenges and Opportunities," *J Ambient Intell Humaniz Comput*, vol. 11, no. 11, pp. 4713–4729, Nov. 2020, doi: 10.1007/s12652-020-01726-4.
- [6] Jujun S. Suriasumantri, *Tentang hakikat ilmu, dalam Ilmu dalam Perspektif*. Jakarta: Gramedia, 1985.
- [7] A. T. Haryono, R. Sarno, and K. R. Sungkono, "Transformer-Gated Recurrent Unit Method for Predicting Stock Price Based on News Sentiments and Technical Indicators," *IEEE Access*, vol. 11, pp. 77132–77146, 2023, doi: 10.1109/ACCESS.2023.3298445.
- [8] T. Hariguna and A. Ruangkanjanes, "Adaptive sentiment analysis using multioutput classification: a performance comparison," *PeerJ Comput Sci*, vol. 9, 2023, doi: 10.7717/peerj-cs.1378.
- [9] L. Kurniasari and A. Setyanto, "SENTIMENT ANALYSIS USING RECURRENT NEURAL NETWORK-LSTM IN BAHASA INDONESIA," 2020.
- [10] B. Andrian, T. Simanungkalit, I. Budi, and A. F. Wicaksono, "Sentiment Analysis on Customer Satisfaction of Digital Banking in Indonesia," 2022. [Online]. Available: www.ijacsa.thesai.org
- [11] G. A. Buntoro, R. Arifin, G. N. Syaifuddiin, A. Selamat, O. Krejcar, and H. Fujita, "Implementation of a Machine Learning Algorithm for Sentiment Analysis of Indonesia's 2019 Presidential Election," *IJUM Engineering Journal*, vol. 22, no. 1, pp. 78–92, 2021, doi: 10.31436/IJUMJ.V22I1.1532.
- [12] E. Di. Madyatmadja, B. N. Yahya, and C. Wijaya, "Contextual Text Analytics Framework for Citizen Report Classification: A Case Study Using the Indonesian Language," *IEEE Access*, vol. 10, pp. 31432–31444, 2022, doi: 10.1109/ACCESS.2022.3158940.
- [13] E. Sutoyo and A. Almaarif, "Twitter sentiment analysis of the relocation of Indonesia's capital city," *Bulletin of Electrical Engineering and Informatics*, vol. 9, no. 4, pp. 1620–1630, Aug. 2020, doi: 10.11591/eei.v9i4.2352.
- [14] M. Bordoloi and S. K. Biswas, "Sentiment analysis: A survey on design framework, applications and future scopes," *Artif Intell Rev*, Nov. 2023, doi: 10.1007/s10462-023-10442-2.
- [15] A. Ghosh, B. C. Dhara, C. Pero, and S. Umer, "A multimodal sentiment analysis system for recognizing person aggressiveness in pain based on textual and visual information," *J Ambient Intell Humaniz Comput*, vol. 14, no. 4, pp. 4489–4501, Apr. 2023, doi: 10.1007/s12652-023-04567-z.
- [16] G. Xiong, K. Yan, and X. Zhou, "A distributed learning based sentiment analysis methods with Web applications," *World Wide Web*, vol. 25, no. 5, pp. 1905–1922, Sep. 2022, doi: 10.1007/s11280-021-00994-0.
- [17] P. Mehta and S. Pandya, "A Review On Sentiment Analysis Methodologies, Practices And Applications," *INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH*, vol. 9, p. 2, 2020, [Online]. Available: www.ijstr.org
- [18] K. Cortis *et al.*, "SemEval-2017 Task 5: Fine-Grained Sentiment Analysis on Financial Microblogs and News," 2017. [Online]. Available: <http://stocktwits.com/>
- [19] M. Rodríguez-Ibáñez, A. Casáñez-Ventura, F. Castejón-Mateos, and P. M. Cuenca-Jiménez, "A Review on Sentiment Analysis from Social Media Platforms," Aug. 01, 2023, *Elsevier Ltd*. doi: 10.1016/j.eswa.2023.119862.
- [20] P. G. Preethi, V. Uma, and A. Kumar, "Temporal sentiment analysis and causal rules extraction from tweets for event prediction," in *Procedia Computer Science*, Elsevier B.V., 2015, pp. 84–89. doi: 10.1016/j.procs.2015.04.154.
- [21] J. A. Balazs and J. D. Velásquez, "Opinion Mining and Information Fusion: A survey," *Information Fusion*, vol. 27, pp. 95–110, Jan. 2016, doi: 10.1016/j.inffus.2015.06.002.
- [22] M. Birjali, M. Kasri, and A. Beni-Hssane, "A comprehensive survey on sentiment analysis: Approaches, challenges and trends," *Knowl Based Syst*, vol. 226, Aug. 2021, doi: 10.1016/j.knosys.2021.107134.
- [23] I. Chaturvedi, E. Cambria, R. E. Welsch, and F. Herrera, "Distinguishing between facts and opinions for sentiment analysis: Survey and challenges," *Information Fusion*, vol. 44, pp. 65–77, Nov. 2018, doi: 10.1016/j.inffus.2017.12.006.
- [24] F. Hemmatian and M. K. Sohrabi, "A survey on classification techniques for opinion mining and sentiment analysis," *Artif Intell Rev*, vol. 52, no. 3, pp. 1495–1545, Oct. 2019, doi: 10.1007/s10462-017-9599-6.
- [25] Sri Shakthi Institute of Engineering and Technology, Institute of Electrical and Electronics Engineers. Madras Section, All-India Council for Technical Education, and Institute of Electrical and Electronics Engineers, *2020 International Conference on Computer Communication and Informatics : January 22-24, 2020, Coimbatore, India*. 2020.
- [26] M. M. Chandio and M. Sah, "Brexit Twitter Sentiment Analysis: Changing Opinions About Brexit and UK Politicians," 2020, pp. 1–11. doi: 10.1007/978-3-030-38501-9_1.
- [27] A. Ray, P. K. Bala, S. Chakraborty, and S. A. Dasgupta, "Exploring the impact of different factors on brand equity and intention to take up online courses from e-Learning platforms," *Journal of Retailing and Consumer Services*, vol. 59, Mar. 2021, doi: 10.1016/j.jretconser.2020.102351.
- [28] N. Songkram, S. Chootongchai, H. Osuwan, Y. Chuppunnarat, and N. Songkram, "Students' adoption towards behavioral intention of digital learning platform," *Educ Inf Technol (Dordr)*, Sep. 2023, doi: 10.1007/s10639-023-11637-4.

- [29] M. Al-Hail, M. F. Zguir, and M. Koç, "University students' and educators' perceptions on the use of digital and social media platforms: A sentiment analysis and a multi-country review," *iScience*, vol. 26, no. 8, Aug. 2023, doi: 10.1016/j.isci.2023.107322.
- [30] S. Alam, I. Mahmud, S. M. S. Hoque, R. Akter, and S. M. Soheli Rana, "Predicting students' intention to continue business courses on online platforms during the Covid-19: An extended expectation confirmation theory," *International Journal of Management Education*, vol. 20, no. 3, Nov. 2022, doi: 10.1016/j.ijme.2022.100706.
- [31] A. Liapis, V. Maratou, T. Panagiotakopoulos, C. Katsanos, and A. Kameas, "UX evaluation of open MOOC platforms: a comparative study between Moodle and Open edX combining user interaction metrics and wearable biosensors," *Interactive Learning Environments*, 2022, doi: 10.1080/10494820.2022.2048674.
- [32] R. Rahmawati, Sukidin, and P. Suharto, "Factor analysis of ruangguru application use on high school students in Jember," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, May 2021. doi: 10.1088/1755-1315/747/1/012026.
- [33] S. F. Persada, A. Oktavianto, B. A. Miraja, R. Nadlifatin, P. F. Belgiawan, and A. A. N. P. Redi, "Public perceptions of online learning in developing countries: A study using the ELK stack for sentiment analysis on twitter," *International Journal of Emerging Technologies in Learning*, vol. 15, no. 9, pp. 94–109, 2020, doi: 10.3991/ijet.v15i09.11579.
- [34] B. Su and J. Peng, "Sentiment Analysis of Comment Texts on Online Courses Based on Hierarchical Attention Mechanism," *Applied Sciences (Switzerland)*, vol. 13, no. 7, Apr. 2023, doi: 10.3390/app13074204.
- [35] H. Kim and G. Qin, "Summarizing Students' Free Responses for an Introductory Algebra-Based Physics Course Survey Using Cluster and Sentiment Analysis," *IEEE Access*, vol. 11, pp. 89052–89066, 2023, doi: 10.1109/ACCESS.2023.3305260.
- [36] J. Joung and H. M. Kim, "Approach for Importance-Performance Analysis of Product Attributes from Online Reviews," *Journal of Mechanical Design*, vol. 143, no. 8, pp. 1–14, Aug. 2021, doi: 10.1115/1.4049865.
- [37] P. Bhuvaneswari, A. N. Rao, Y. H. Robinson, and M. N. Thippeswamy, "Sentiment analysis for user reviews using Bi-LSTM self-attention based CNN model," *Multimed Tools Appl*, vol. 81, no. 9, pp. 12405–12419, Apr. 2022, doi: 10.1007/s11042-022-12410-4.
- [38] Z. Wang, P. Gao, and X. Chu, "Sentiment Analysis from Customer-Generated Online Videos on Product Review Using Topic Modeling and Multi-Attention BLSTM," *Advanced Engineering Informatics*, vol. 52, pp. 1–11, Apr. 2022, doi: 10.1016/j.aei.2022.101588.
- [39] H. J. Alantari, I. S. Currim, Y. Deng, and S. Singh, "An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews," *International Journal of Research in Marketing*, vol. 39, no. 1, pp. 1–19, Mar. 2022, doi: 10.1016/j.ijresmar.2021.10.011.
- [40] D. Antypas, A. Preece, and J. Camacho-Collados, "Negativity spreads faster: A large-scale multilingual twitter analysis on the role of sentiment in political communication," *Online Soc Netw Media*, vol. 33, Jan. 2023, doi: 10.1016/j.osnem.2023.100242.
- [41] P. Chauhan, N. Sharma, and G. Sikka, "Application of Twitter sentiment analysis in election prediction: a case study of 2019 Indian general election," *Soc Netw Anal Min*, vol. 13, no. 1, Dec. 2023, doi: 10.1007/s13278-023-01087-8.
- [42] A. Yavari, H. Hassanpour, B. R. Cami, and M. Mahdavi, "Election Prediction Based on Sentiment Analysis using Twitter Data," *International Journal of Engineering, Transactions B: Applications*, vol. 35, no. 2, pp. 372–379, Feb. 2022, doi: 10.5829/ije.2022.35.02b.13.
- [43] Y. Ouyang, B. Guo, J. Zhang, Z. Yu, and X. Zhou, "SentiStory: multi-grained sentiment analysis and event summarization with crowdsourced social media data," *Pers Ubiquitous Comput*, vol. 21, no. 1, pp. 97–111, Feb. 2017, doi: 10.1007/s00779-016-0977-x.
- [44] H. Smith and W. Cipolli, "The Instagram/Facebook ban on graphic self-harm imagery: A sentiment analysis and topic modeling approach," *Policy Internet*, vol. 14, no. 1, pp. 170–185, Mar. 2022, doi: 10.1002/poi3.272.
- [45] J. S. Suriasumantri, *Filsafat Ilmu*. Jakarta: Pustaka Sinar Harapan, 2007.
- [46] W. Mays, *The philosophy of Whitehead*. Routledge, 2014.
- [47] S. Endraswara, *Filsafat Ilmu*. Media Pressindo, 2021.
- [48] A. L. Maas, R. E. Daly, P. T. Pham, D. Huang, A. Y. Ng, and C. Potts, "Learning Word Vectors for Sentiment Analysis," 2011.
- [49] U. M. Dahir and F. K. Alkindy, "Utilizing Machine Learning for Sentiment Analysis of IMDB Movie Review Data," *International Journal of Engineering Trends and Technology*, vol. 71, no. 5, pp. 18–26, May 2023, doi: 10.14445/22315381/IJETT-V71I5P203.
- [50] M. T. H. K. Tusar and M. T. Islam, "A Comparative Study of Sentiment Analysis Using NLP and Different Machine Learning Techniques on US Airline Twitter Data," in *Proceedings of International Conference on Electronics, Communications and Information Technology, ICECIT 2021*, Khulna, Bangladesh: Institute of Electrical and Electronics Engineers Inc., Sep. 2021, pp. 1–4. doi: 10.1109/ICECIT54077.2021.9641336.
- [51] K. M. Hasib, M. A. Habib, N. A. Towhid, and M. I. H. Showrov, "A Novel Deep Learning based Sentiment Analysis of Twitter Data for US Airline Service," in *2021 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2021 - Proceedings*, Dhaka, Bangladesh: Institute of Electrical and Electronics Engineers Inc., Feb. 2021, pp. 450–455. doi: 10.1109/ICICT4SD50815.2021.9396879.
- [52] A. Go, R. Bhayani, and L. Huang, "Twitter Sentiment Classification using Distant Supervision," 2009. [Online]. Available: <http://tinyurl.com/cvvg9a>
- [53] S. Reddy and Damodar, "Accuracy of prediction by machine learning algorithms," *Int J Eng Adv Technol*, vol. 8, no. 6 Special Issue 3, pp. 1929–1933, Sep. 2019, doi: 10.35940/ijet.F1371.0986S319.
- [54] K. L. Tan, C. P. Lee, and K. M. Lim, "RoBERTa-GRU: A Hybrid Deep Learning Model for Enhanced Sentiment Analysis," *Applied Sciences (Switzerland)*, vol. 13, no. 6, pp. 2–16, Mar. 2023, doi: 10.3390/app13063915.
- [55] L. Bryan-Smith, J. Godsall, F. George, K. Ego, N. Dethlefs, and D. Parsons, "Real-time Social Media Sentiment Analysis for Rapid Impact Assessment of Floods," *Comput Geosci*, vol. 178, pp. 1–13, Sep. 2023, doi: 10.1016/j.cageo.2023.105405.
- [56] R. Haque, N. Islam, M. Tasneem, and A. K. Das, "Multi-class sentiment classification on Bengali social media comments using machine learning," *International Journal of Cognitive Computing in Engineering*, vol. 4, pp. 21–35, Jun. 2023, doi: 10.1016/j.jcce.2023.01.001.

-
- [57] A. Qazi, N. Hasan, C. M. Owusu-Ansah, G. Hardaker, S. K. Dey, and K. Haruna, "SentiTAM: Sentiments Centered Integrated Framework for Mobile Learning Adaptability in Higher Education," *Heliyon*, vol. 9, no. 1, pp. 1–14, Jan. 2023, doi: 10.1016/j.heliyon.2022.e12705.
- [58] M. El Barachi, M. AlKhatib, S. Mathew, and F. Oroumchian, "A Novel Sentiment Analysis Framework for Monitoring the Evolving Public Opinion in Real-time: Case Study on Climate Change," *J Clean Prod*, vol. 312, pp. 1–12, Aug. 2021, doi: 10.1016/j.jclepro.2021.127820.
- [59] Y. Liu *et al.*, "Scanning, attention, and reasoning multimodal content for sentiment analysis," *Knowl Based Syst*, vol. 268, pp. 1–11, May 2023, doi: 10.1016/j.knosys.2023.110467.
- [60] M. Asif, A. Ishtiaq, H. Ahmad, H. Aljuaid, and J. Shah, "Sentiment analysis of extremism in social media from textual information," *Telematics and Informatics*, vol. 48, pp. 1–20, May 2020, doi: 10.1016/j.tele.2020.101345.
- [61] K. L. Tan, C. P. Lee, and K. M. Lim, "A Survey of Sentiment Analysis: Approaches, Datasets, and Future Research," Apr. 01, 2023, *MDPI*. doi: 10.3390/app13074550.
- [62] J. Blitzer, M. Dredze, and F. Pereira, "Biographies, Bollywood, Boom-boxes and Blenders: Domain Adaptation for Sentiment Classification," Association for Computational Linguistics, 2007. [Online]. Available: <http://ida>.
- [63] Y. Dang, Y. Zhang, and H. Chen, "A Lexicon-Enhanced Method for Sentiment Classification: An Experiment on Online Product Reviews," 2010. [Online]. Available: www.computer.org/intelligent
- [64] A. Fahrni, "Old Wine or Warm Beer: Target-Specific Sentiment Analysis of Adjectives," 2008, doi: 10.5167/uzh-8810.
- [65] V. A. Kharde and S. S. Sonawane, "Sentiment Analysis of Twitter Data: A Survey of Techniques," 2016. [Online]. Available: <http://ai.stanford>.
- [66] Y. Guo, S. Das, S. Lakamana, and A. Sarker, "An aspect-level sentiment analysis dataset for therapies on Twitter," *Data Brief*, vol. 50, pp. 1–5, Oct. 2023, doi: 10.1016/j.dib.2023.109618.
- [67] H. Al-Samarraie, S. Muthana Sarsam, and A. I. Alzahrani, "Haptic Technology in Society: A Sentiment Analysis of Public Engagement," *Comput Human Behav*, vol. 147, pp. 1–10, Oct. 2023, doi: 10.1016/j.chb.2023.107862.
- [68] A. R. Chrismanto, A. K. Sari, and Y. Suyanto, "Enhancing Spam Comment Detection on Social Media With Emoji Feature and Post-Comment Pairs Approach Using Ensemble Methods of Machine Learning," *IEEE Access*, vol. 11, pp. 80246–80265, 2023, doi: 10.1109/ACCESS.2023.3299853.
- [69] A. M. Tri Sakti, E. Mohamad, and A. A. Azlan, "Mining of opinions on COVID-19 large-scale social restrictions in indonesia: Public sentiment and emotion analysis on online media," *J Med Internet Res*, vol. 23, no. 8, Aug. 2021, doi: 10.2196/28249.
- [70] G. A. Buntoro, R. Arifin, G. N. Syaifuddin, A. Selamat, O. Krejcar, and H. Fujita, "Implementation of a Machine Learning Algorithm for Sentiment Analysis of Indonesia's 2019 Presidential Election," *IJUM Engineering Journal*, vol. 22, no. 1, pp. 78–92, 2021, doi: 10.31436/IJUMJ.V22I1.1532.
- [71] Y. Li, Q. Pan, T. Yang, S. Wang, J. Tang, and E. Cambria, "Learning Word Representations for Sentiment Analysis," *Cognit Comput*, vol. 9, no. 6, pp. 843–851, Dec. 2017, doi: 10.1007/s12559-017-9492-2.
- [72] M. Elsheh, H. Benghuzzi, and M. M. Elsheh, "An Investigation of Keywords Extraction from Textual Documents using Word2Vec and Decision Tree," 2020. [Online]. Available: <https://sites.google.com/site/ijcsis/>
- [73] S. Beliga and A. Meštrović, "An Overview of Graph-Based Keyword Extraction Methods and Approaches," 2015.
- [74] E. Cambria, "Affective Computing and Sentiment Analysis," *IEEE Intell Syst*, vol. 31, no. 2, pp. 102–107, Mar. 2016, doi: 10.1109/MIS.2016.31.
- [75] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: A survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, Dec. 2014, doi: 10.1016/j.asej.2014.04.011.
- [76] B. Pang and L. Lee, "A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts," 2004. [Online]. Available: www.cs.cornell.edu/people/pabo/movie-
- [77] Y. Hu and W. Li, "Document sentiment classification by exploring description model of topical terms," *Comput Speech Lang*, vol. 25, no. 2, pp. 386–403, Apr. 2011, doi: 10.1016/j.csl.2010.07.004.
- [78] A. R. Chrismanto, A. K. Sari, and Y. Suyanto, "Enhancing Spam Comment Detection on Social Media With Emoji Feature and Post-Comment Pairs Approach Using Ensemble Methods of Machine Learning," *IEEE Access*, vol. 11, pp. 80246–80265, 2023, doi: 10.1109/ACCESS.2023.3299853.
- [79] N. Abdelhady, I. E. Elsemman, M. F. Farghally, and T. H. A. Soliman, "Developing Analytical Tools for Arabic Sentiment Analysis of COVID-19 Data," *Algorithms*, vol. 16, no. 7, Jul. 2023, doi: 10.3390/a16070318.
- [80] D. K. Jain, A. Kumar, and S. R. Sangwan, "TANA: The amalgam neural architecture for sarcasm detection in indian indigenous language combining LSTM and SVM with word-emoji embeddings," *Pattern Recognit Lett*, vol. 160, pp. 11–18, Aug. 2022, doi: 10.1016/j.patrec.2022.05.026.
- [81] M. E. Basiri, S. Nemati, M. Abdar, E. Cambria, and U. R. Acharya, "ABCDM: An Attention-based Bidirectional CNN-RNN Deep Model for sentiment analysis," *Future Generation Computer Systems*, vol. 115, pp. 279–294, Feb. 2021, doi: 10.1016/j.future.2020.08.005.
- [82] L. Li and X. T. Wang, "Nonverbal communication with emojis in social media: dissociating hedonic intensity from frequency," *Lang Resour Eval*, vol. 57, no. 1, pp. 323–342, Mar. 2023, doi: 10.1007/s10579-022-09611-6.
- [83] W. Mays, *The philosophy of Whitehead*. Routledge, 2014.
- [84] T. Han, C. Liu, W. Yang, and D. Jiang, "A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults," *Knowl Based Syst*, vol. 165, pp. 474–487, Feb. 2019, doi: 10.1016/j.knosys.2018.12.019.
- [85] R. Hassan and M. R. Islam, "Impact of Sentiment Analysis in Fake Online Review Detection," in *2021 International Conference on Information and Communication Technology for Sustainable Development, ICICT4SD 2021 - Proceedings*, Institute of Electrical and Electronics Engineers Inc., Feb. 2021, pp. 21–24. doi: 10.1109/ICICT4SD50815.2021.9396899.
- [86] A. Ebadi, P. Xi, S. Tremblay, B. Spencer, R. Pall, and A. Wong, "Understanding the temporal evolution of COVID-19 research through machine learning and natural language processing," *Scientometrics*, vol. 126, no. 1, pp. 725–739, Jan. 2021, doi: 10.1007/s11192-020-03744-7.
-

- [87] E. M. Clark *et al.*, "A Sentiment Analysis of Breast Cancer Treatment Experiences and Healthcare Perceptions Across Twitter," May 2018, [Online]. Available: <http://arxiv.org/abs/1805.09959>
- [88] S. Kumar, M. Yadava, and P. P. Roy, "Fusion of EEG response and sentiment analysis of products review to predict customer satisfaction," *Information Fusion*, vol. 52, pp. 41–52, Dec. 2019, doi: 10.1016/j.inffus.2018.11.001.
- [89] A. Osmani and J. B. Mohasefi, "Weighted Joint Sentiment-Topic Model for Sentiment Analysis Compared to ALGA: Adaptive Lexicon Learning Using Genetic Algorithm," *Comput Intell Neurosci*, vol. 2022, 2022, doi: 10.1155/2022/7612276.
- [90] P. K. Jain, E. A. Yekun, R. Pamula, and G. Srivastava, "Consumer recommendation prediction in online reviews using Cuckoo optimized machine learning models," *Computers and Electrical Engineering*, vol. 95, Oct. 2021, doi: 10.1016/j.compeleceng.2021.107397.
- [91] F. Z. Xing, E. Cambria, and R. E. Welsch, "Natural language based financial forecasting: a survey," *Artif Intell Rev*, vol. 50, no. 1, pp. 49–73, Jun. 2018, doi: 10.1007/s10462-017-9588-9.
- [92] L. Rognone, S. Hyde, and S. S. Zhang, "News sentiment in the cryptocurrency market: An empirical comparison with Forex," *International Review of Financial Analysis*, vol. 69, May 2020, doi: 10.1016/j.irfa.2020.101462.
- [93] S. Hu, A. Kumar, F. Al-Turjman, S. Gupta, S. Seth, and Shubham, "Reviewer Credibility and Sentiment Analysis Based User Profile Modelling for Online Product Recommendation," *IEEE Access*, vol. 8, pp. 26172–26189, 2020, doi: 10.1109/ACCESS.2020.2971087.