Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14070
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Summarizer Precision Value on Tribunnews Gorontalo in the Implementation of Online Discourse Sentiment Analysis

Rahmat Taufik R.L Bau[1)*]
[1)]Universitas Negeri Gorontalo, Indonesia
[1)]rahmattaufik@ung.ac.id

Submitted :Aug 28, 2024 | **Accepted** : Sep 16, 2024 | **Published** : Oct 10, 2024

Abstract: This research investigates the precision of a summarization-based sentiment analysis framework applied to online discourses, specifically from Tribunnews Gorontalo. This study aims to develop and evaluate a sentiment analysis framework that accurately parses complex meanings and nuances in online discourse. The research process begins with summarizing the content using Python, followed by tokenization and sentiment analysis using the BERT model. The precision of the sentiment analysis was meticulously measured. Results indicate that the precision analysis demonstrates that the Python-implemented model achieved a 86% precision rate when applied to ten online discourses from Tribunnews Gorontalo. This research contributes significantly to understanding public sentiments in online content, offering deeper and more accurate insights.

Keywords: Precision; summarizer; online discourse; sentiment analysis; BERT

INTRODUCTION

A better understanding of the role of digital literacy and the effective use of technology can provide important insights into technology education and use in Indonesia and similar contexts (Lee & Hidayat, 2019). In an era where technology permeates every aspect of life, digital literacy is not just about having access to devices and an internet connection but also about the ability to understand, analyze, and critique digital information wisely.

Amidst the ongoing digital transformation, sentiment in online content increasingly shapes people's opinions on various issues, including education and technology utilization. Analyzing sentiments from different sources of online content provides valuable insights into how digital literacy and technology impact people's views. Several previous studies have demonstrated the effectiveness of sentiment analysis in understanding opinions through online content. For example, Rahmi Hajrianti (Hajriyanti & Akbar, 2021) revealed the substantial impact of the coronavirus pandemic on MSME businesses in Indonesia through an investigation of online literature. Similarly, Indri Tri Julianto (Julianto & Lindawati, 2022) analyzed sentiment to assess user satisfaction with the Android-based Student Academic Information System (SIAM) at Institut Teknologi Garut (ITG) by examining user comments on Google Play. Christiany Juditha's study on the Ahok case using news portals highlights the relevance of sentiment analysis in the context of actual events (Juditha, 2017).

In Indonesia, where social, economic, and cultural differences create unique dynamics, understanding the sentiments in online content helps recognize trends and mindsets. Tracing sentiments from online content allows us to read between the lines of people's perceptions, expectations, and tendencies, revealing hidden messages capable of influencing public views. (Barachi at al., 2021) concluded that sentiment analysis frameworks in monitoring public opinion in real-time have significant implications, providing immediate evaluations of collective opinion development. Shi W. et al.'s research supports the view that sentiment analysis is effective in reflecting changes in public opinion through various emotional values (Shi at al., 2013).

However, the challenge of extracting meaning and sentiment from online content has intensified with rapid digital development. Pushmann (Puschmann & Powell, 2018) argues that sentiment analysis, as a method, has hidden assumptions and constraints, including technical limitations in achieving precise results and conceptual limitations in representing the complexity of sentiments. These limitations underscore the need for a more refined approach to sentiment analysis.

Based on these challenges and considerations, this research aims to develop a sentiment analysis framework capable of parsing the meaning in online content with high accuracy. By delving deeper into the sentiments and meanings emerging from online content, we can gain insights into people's perceptions and views on various issues. Recent research has explored its applications in diverse fields. Gunter (Gunter et al., 2014) examined the reliability of sentiment analysis tools in measuring public opinion, highlighting their potential for guiding political and business decisions. Feng (Feng, 2024) focused on deep learning-based approaches for social media sentiment

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14070
e- ISSN : 2541-2019
p- ISSN : 2541-044X

analysis, emphasizing multimodal and aspect-based methods to uncover emotional nuances in digital communication. Hu (Hu et al., 2019) applied semantic and sentiment analysis to online neighborhood reviews, extracting knowledge about residents' perceptions of their living environments. These papers collectively demonstrate the growing importance of sentiment analysis in understanding public sentiment. Therefore, This study seeks to develop a sentiment analysis framework using Python that not only recognizes positive, negative, or neutral sentiments but also accurately parses the nuances and complexities of meanings in online content. The application of this framework to Tribunnews Gorontalo aims to demonstrate its precision, contributing to a more in-depth and accurate sentiment analysis of public opinion

LITERATURE REVIEW

Sentiment analysis is a method used to evaluate the attitudes, opinions, and emotions contained in a text or discourse. In this research context, several models have been developed and applied to carry out this sentiment analysis. Each model has its own advantages and disadvantages, allowing researchers to choose the approach that best suits their research objectives.

Sentiment analysis models cover a wide range of techniques, Various sentiment analysis techniques have been explored in the literature. (Vaghela & Jadav, 2016) emphasized the importance of careful feature selection and existing classification approaches for accuracy. (Cambria et al., 2015) introduced a concept-level sentiment analysis (CLSA) model, which considers various natural language processing tasks. (Abirami & Gayathri, 2017) discuss the limitations of existing methods, such as polarity shifts and binary classification problems, and the use of machine learning algorithms such as Naive Bayes and Support Vector Machine. (Attri & Dutta, 2020) provides an overview of sentiment analysis methods, including supervised and unsupervised learning approaches. From rule-based approaches to machine learning-based models by (Pathak & piyush Rai, 2023). Rule-based approaches typically use a collection of rules or keywords to identify sentiment, while machine learning models involve training algorithms on training data to automatically understand sentiment patterns.

The advantages of rule-based approaches include high interpretability and the ability to handle special cases that have been explicitly defined. However, its drawbacks lie in its limitations in handling language nuances and its inability to handle changes in writing style (Berka, 2020). On the other hand, (Shen et al., 2019) says machine learning models can be more flexible in recognizing complex patterns and adapting to variations in language use. Nonetheless, drawbacks may include lack of interpretability and dependence on large amounts of training data for optimal performance.

Taking these differences into account, the selection of a sentiment analysis model should be tailored to the specific objectives of the research and the characteristics of the data used. Researchers need to consider the trade-off between interpretability and model performance, and choose the most suitable model to address the specific challenges in sentiment analysis in a particular domain or context.

Sentiment analysis has become a critical aspect of understanding human opinions and responses to various topics scattered in digital texts. In the course of its evolution, traditional methods such as Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) have paved the way for more sophisticated models, providing deeper insight into the complexity of modern discourse (Lan, 2022).

BoW, as the first method in sentiment analysis, converts text into a vector based on word frequency without regard to its order or sentence structure. Along with that, TF-IDF assigns weights to words based on their frequency in the document and how rarely they appear throughout the document. Although simple and easy to implement, both have limitations in understanding context and increasingly complex sentence structures (Lan, 2022).

Further developments in sentiment analysis led to the use of models such as Latent Dirichlet Allocation (LDA) and Word2Vec. LDA, as a topic model, is not only capable of classifying topics in text but can also be used to identify positive and negative sentiments. However, (Maier et al., 2018) found that the use of LDA still requires manual determination of the topics, limiting scalability and reliance on human interpretation.

Word2Vec, on the other hand, brought a revolution in natural language processing. The model is capable of generating word vectors from text, improving the understanding of context and relationships between words. Word2Vec can identify complex patterns in text, enriching sentiment analysis with the recognition of more subtle nuances and subtleties. Word2Vec has significantly impacted natural language processing, with its optimal hyperparameters playing an important role in the quality of its embedding (Adewumi et al., 2022). However, the best models are often task-specific, and increasing the size of the embedding dimensions can lead to poor quality (Adewumi et al., 2022)

Although Word2Vec provides significant improvements in understanding language, new challenges arise as discourse becomes more complex and technology advances. In this case, Long Short-Term Memory (LSTM), as a type of Recurrent Neural Network (RNN), is a solution that makes an important contribution. LSTM is able to learn from word sequences, recognize patterns in text in a more adaptive way, and consider context and word order.

However, as the complexity of sentiment analysis increases, there is a need for models that can better capture the context and nuances of language. Therefore, Bidirectional Encoder Representations from Transformers (BERT) entered the scene as the most advanced and recommended solution. (Chen et al., 2019) demonstrated the

*name of corresponding author

e- ISSN : 2541-2019
p- ISSN : 2541-044X

effectiveness of BERT in session-based recommendation, while (Coghill & Reis, 2021) explored its use in database search algorithms. (Gogoulou, 2019) further extended its application to conversational machine understanding. (Chaudhry, 2022) used BERT for sentiment analysis in stock price modeling, demonstrating its versatility in various domains. These studies collectively underline the advanced nature of BERT and are recommended in various tasks. As a deep learning model, BERT offers a more effective and efficient understanding of the context of words in text.

BERT processes data at scale and understands increasingly complex nuances and contexts. By utilizing a transformer approach, BERT is able to overcome the limitations of understanding sentence context and relationships between words. BERT's ability to capture information from words before and after a word in a sentence, known as "bidirectional" capability, enables more holistic and contextual sentiment analysis. Thus, in the growing context of sentiment analysis, BERT can be said to be the first choice for researchers and practitioners. This model not only improves accuracy in recognizing sentiment, but also provides a deeper understanding of the complexity of language expression in digital discourse.

With the continuous development of technology, the possibility of adopting new models that are more sophisticated and intelligent in analyzing sentiment is wide open. The ability to understand and respond to human opinions, attitudes, and emotions through text is critical for a variety of applications, including policy, business, and human-to-human interaction (Chatterjee et al., 2019). Emotion detection in textual dialog can be achieved through a combination of semantic and sentiment-based representations, as shown by Chatterjee. Thus, the evolution in sentiment analysis is not only a reflection of technological advancement, but also a driver of change in the way we understand and interact with the world around us.

In line with technological advancements and continuous research in sentiment analysis, the application of the BERT model as the primary recommendation for understanding and evaluating sentiment in text is not only a step forward, but also a significant breakthrough. The use of state-of-the-art models such as BERT reflects an awareness of the complexity of human language and the need to address the growing challenges of sentiment analysis. Therefore, these steps are not only a representation of technological development, but also a proactive strategy in the face of change and complexity in digital discourse.

METHOD

The method used in this research is a descriptive method with an artificial intelligence approach: Python-based Natural Language Processing to develop sentiment analysis methods in online content in the form of discourse. The outline of the stages of this research can be explained as follows:

Article downloading and summarization

In the first phase of our research, we utilize the `newspaper` library to download, parse, and summarize news articles, which then serve as the text for sentiment analysis. Initially, we define a `url` variable containing the link (URL) to the specific news article we wish to analyze. Using the `Article` class from the `newspaper` library, we create an object to access and manage the article. This object allows us to download the article content, parse its structure, and extract a summary.

Tokenization

Before conducting sentiment analysis, the text must be converted into a numerical representation that the model can interpret. This process, known as tokenization, involves dividing the text into small units called tokens. For the BERT model, tokenization entails breaking the text into tokens, which can be words or subwords. Each token is then mapped to a numerical vector that BERT can process. This numerical representation allows the model to understand and analyze the textual data effectively, paving the way for accurate sentiment analysis.

Sentiment analysis

In this phase, we perform sentiment analysis on the tokenized text using the BERT model. BERT, a state-of-the-art natural language processing model, processes the numerical vectors representing the tokens to determine the sentiment expressed in the text. The model analyzes the context and semantic nuances within the text to classify the sentiment as positive, negative, or neutral. This analysis provides insights into the underlying sentiments conveyed in the summarized article, helping to understand public opinion as reflected in the online discourse from Tribunnews Gorontalo.

Precision measurement

The final phase involves measuring the precision of the sentiment analysis to evaluate the accuracy of our model. Precision measurement compares the model's sentiment analysis results against a set of manually annotated benchmarks to determine how accurately the model identifies the correct sentiment. This involves calculating the proportion of true positive results (correctly identified sentiments) out of all positive results identified by the model. By assessing precision, we ensure the reliability and effectiveness of our sentiment analysis framework,

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14070
e- ISSN : 2541-2019
p- ISSN : 2541-044X

confirming that the model accurately captures the sentiments expressed in the online discourse from Tribunnews Gorontalo..

RESULT

An article from the news website "https://gorontalo.tribunnews.com/2023/11/16/hmj-bahasa-inggris-ung-gorontalo-buka-layanan-curhat-untuk-mahasiswa-soal-kuliah-hingga-finansial" will be downloaded and sentiment analysed.

Figure 1. View of the news page used as a case study.

```python
url="https://gorontalo.tribunnews.com/2023/11/16/hmj-bahasa-inggris-ung-gorontalo-buka-layanan-curhat-untuk-mahasiswa-soal-kuliah-hingga-finansial"
article = Article(url)
article.download()
article.parse()
article.nlp()
text = article.summary
```

Using the `newspaper` library, news articles are downloaded, parsed, and analyzed to produce a summary of the article as the text to be sentiment analyzed. In this section, we first define a URL variable that contains the link (URL) to the news article that we want to download and analyze for sentiment. Next, we use an Article from the newspaper library to create an object that will be used to access and manage the article.

The next step is to download the content of the news article. The download() function is used to retrieve the HTML content from the URL provided. After downloading, we perform parsing to extract important information from the HTML content. This includes the article title, publication date, and of course, the content itself. Then, we use the text analysis module of the newspaper to perform additional natural language processing on the article. This process involves tokenization, entity recognition, and other NLP processes that help in understanding and extracting information from the article text.

The final step is to retrieve the summary of the article as the text to be sentiment analyzed. This summary is generated by the previous text analysis module and often contains the main points or gist of the article. With these steps, we have prepared the article text from the given URL for sentiment analysis using the BERT model or other sentiment analysis methods. The article can cover a wide range of topics, and sentiment analysis will help us understand the views or feelings contained in the article.

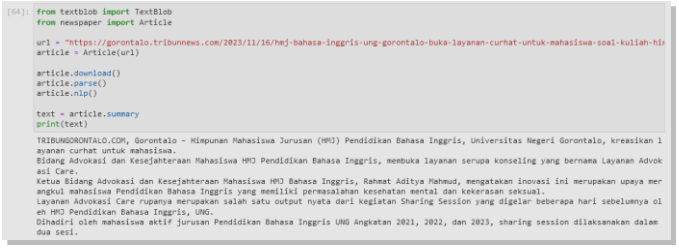

Figure 2. Results of news summaries obtained from the news page

Before analyzing sentiment, the text needs to be converted into a numerical representation that the model can understand. This process is called tokenization. In the case of BERT, tokenization involves dividing the text into small pieces called tokens. Each token is then represented by a numerical vector that the model can understand.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14070
e- ISSN : 2541-2019
p- ISSN : 2541-044X

```python
  tokenizer= AutoTokenizer.from_pretrained('nlptown/bert-base-multilingual-uncased-sentiment')
```

In this section, we use AutoTokenizer from the transformers library to initialize a tokenizer. This tokenizer has been trained on the multilingual BERT model ('nlptown/bert-base-multilingual-uncased-sentiment'). This tokenizer is responsible for converting text into numerical representations (tokens) that the model can understand.

```python
 tokens = tokenizer.encode(text, return_tensors='pt', max_length=512, truncation=True)
```

Once a tokenizer is set up, we can use the **encode** method to convert the text into tokens. The key arguments in this method include **text**, which is the variable containing the text we want to tokenize. We can also specify **return_tensors='pt'**, which ensures the tokenization result is returned in PyTorch tensor format. Additionally, the argument **max_length=512** sets the maximum token length to 512, which is the limit that the BERT model can handle. To ensure the tokens stay within this limit, **truncation=True** is used, which truncates any tokens if the text exceeds the specified maximum length. The result of tokenizer.encode is a PyTorch tensor that contains a numerical representation of the tokenized text. This tensor can later be used as input for the BERT model. After tokenization, the next step is sentiment analysis. BERT models that have been trained for sequence classification tasks can be used to predict sentiment from text.

```python
  model=  AutoModelForSequenceClassification.from_pretrained('nlptown/bert-base-multilingual-uncased-sentiment')
```

In this section, we use AutoModel from the transformers library to initialize a BERT model that has been trained for sequence classification tasks, such as sentiment analysis. The initialized model corresponds to the multilingual BERT model ('nlptown/bert-base-multilingual-uncased-sentiment'). This model is pre-trained in classification tasks, so it has an understanding of the relationships between words in a text. To get the sentiment score from the news summary, we use the `sentiment_score` function.

```python
  text = sentiment_score(text)
```

The variable text is converted to the sentiment label generated by the sentiment_score function. This means that the text variable now contains the sentiment label of the analyzed article text. This label can be a number, such as 1 to 5.

Figure 3. Sentiment scores contained in news stories

In the context of applying label conventions to the BERT model, a sentiment scale from 1 to 5 is used, where each number corresponds to a specific level of sentiment. Starting from 1, **Label 1** represents "Very Negative Sentiment," indicating that the text expresses a highly unfavorable or strongly negative viewpoint. **Label 2**, or "Negative Sentiment," suggests that while the sentiment is critical or unfavorable, it is not as extreme as in Label 1. **Label 3** stands for "Neutral Sentiment," meaning the text presents information without expressing a strong opinion, either positive or negative. Moving up the scale, **Label 4** signifies "Positive Sentiment," indicating that the text conveys a favorable or pleasant outlook. Lastly, **Label 5** represents "Very Positive Sentiment," where the content is seen as highly favorable or containing an exceptionally positive perspective on the topic. With this step, we have successfully applied the BERT model to analyze the sentiment of the article text retrieved from the URL. The result, sentiment labels, can be used to gain a high-level understanding of how the article can be interpreted sentimentally. BERT provides more contextualized sentiment analysis results. This means that the model takes into account the relationship between words in a text, producing more accurate and informative results. This contrasts with the approach of previous models which may only consider words in isolation.

In the final phase, we assess the precision of the sentiment analysis to evaluate the accuracy and reliability of our model. Initially, the sentiment analysis of a specific article was assessed as very positive, indicating that the summarizer effectively captured the essence of the article content. However, to ensure that this result is consistent and not an anomaly, we extend our assessment to multiple articles. We use precision measurement, which involves comparing the model's sentiment analysis results against a set of manually annotated benchmarks for each article. Precision is calculated as the proportion of true positive results (correctly identified sentiments) out of all positive results identified by the model. By evaluating multiple articles, we can confirm the robustness and effectiveness

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14070
e- ISSN : 2541-2019
p- ISSN : 2541-044X

of our sentiment analysis framework, ensuring it accurately captures the sentiments expressed in the online discourse from Tribunnews Gorontalo.

Table 1. The precision of 20 articles

Source	Prediction	Result	Source	Prediction	Result
Article 1	Positive	Neutral (3)	**Article 11**	Positive	Negative (1)
Article 2	Positive	Negative (2)	**Article 12**	Positive	Positive (5)
Article 3	Positive	Positive(4)	**Article 13**	Positive	Positive (4)
Article 4	Positive	Negative(1)	**Article 14**	Positive	Neutral (3)
Article 5	Positive	Positive(4)	**Article 15**	Positive	Netural (3)
Article 6	Positive	Positive(4)	**Article 16**	Positive	Positive (5)
Article 7	Positive	Positive(4)	**Article 17**	Positive	Positive (5)
Article 8	Positive	Neutral(3)	**Article 18**	Positive	Negative (1)
Article 9	Positive	Positive(4)	**Article 19**	Positive	Positive (4)
Article 10	Positive	Positive(4)	**Article 20**	Positive	Positive (4)

With these results, we can calculate the precision:

$$\text{Precision} = 12 / (12+2) = 0.86$$

So, in the development of this sentiment analysis model using Python language, the precision is about 0.86 or 86%. This means that of the texts predicted as positive, about 86% of them are positive. Precision gives an idea of how precise the model is in predicting positive sentiment. The higher the precision value, the better the model is at avoiding giving false positive predictions

CONCLUSION

The analysis revealed that the summarizer effectively captured the essence of the articles, enabling accurate sentiment analysis. Despite the variability in sentiment across the articles analyzed, with some expressing positive sentiments and others negative or neutral sentiments, the overall precision of the sentiment analysis model was calculated to be approximately 86%. This indicates that of the texts predicted as positive, around 86% of them indeed exhibited positive sentiment.

Future work may focus on enhancing the model's accuracy and expanding its applicability to a broader range of topics and contexts, further solidifying its utility in sentiment analysis of online content.

ACKNOWLEDGMENT

This research was funded by LPPM Universitas Negeri Gorontalo through Penelitian Dosen Pemula with the Decree of Rector Universitas Negeri Gorontalo No: 1128/UN47/HK.02/2023.

REFERENCES

Abirami, A. M., & Gayathri, V. (2017). A survey on sentiment analysis methods and approach. *2016 Eighth International Conference on Advanced Computing (ICoAC)*, 72–76. https://api.semanticscholar.org/CorpusID:23311975

Adewumi, T. P., Liwicki, F. S., & Liwicki, M. (2022). Word2Vec: Optimal hyperparameters and their impact on natural language processing downstream tasks. *Open Computer Science*, *12*, 134–141. https://api.semanticscholar.org/CorpusID:247617981

Attri, I., & Dutta, M. (2020). *Review of Various Sentiment Analysis Approaches*. https://api.semanticscholar.org/CorpusID:216225816

Berka, P. (2020). Sentiment analysis using rule-based and case-based reasoning. *Journal of Intelligent Information Systems*, *55*, 51–66. https://api.semanticscholar.org/CorpusID:254602489

Cambria, E., Poria, S., Bisio, F., Bajpai, R., & Chaturvedi, I. (2015). The CLSA Model: A Novel Framework for Concept-Level Sentiment Analysis. *Conference on Intelligent Text Processing and Computational Linguistics*. https://api.semanticscholar.org/CorpusID:2700993

Chatterjee, A., Gupta, U., Chinnakotla, M. K., Srikanth, R., Galley, M., & Agrawal, P. (2019). Understanding Emotions in Text Using Deep Learning and Big Data. *Comput. Hum. Behav.*, *93*, 309–317. https://api.semanticscholar.org/CorpusID:59605749

Chaudhry, P. (2022). Bidirectional Encoder Representations from Transformers for Modelling Stock Prices. *International Journal for Research in Applied Science and Engineering Technology*. https://api.semanticscholar.org/CorpusID:247043641

*name of corresponding author

Chen, X., Liu, D., Lei, C., Li, R., Zha, Z., & Xiong, Z. (2019). BERT4SessRec: Content-Based Video Relevance Prediction with Bidirectional Encoder Representations from Transformer. *Proceedings of the 27th ACM International Conference on Multimedia*. https://api.semanticscholar.org/CorpusID:204837442

Coghill, J. G., & Reis, H. (2021). Hey BERT! Meet the Databases: Explorations of Bidirectional Encoder Representation from Transformers Model Use in Database Search Algorithms. *Journal of Electronic Resources in Medical Libraries*, *18*, 112–118. https://api.semanticscholar.org/CorpusID:237548412

Feng, B. (2024). Deep learning-based sentiment analysis for social media: A focus on multimodal and aspect-based approaches. *Applied and Computational Engineering*, *33*(1), 1–8. https://doi.org/10.54254/2755-2721/33/20230224

Gogoulou, E. (2019). *Using Bidirectional Encoder Representations from Transformers for Conversational Machine Comprehension*. https://api.semanticscholar.org/CorpusID:230857546

Gunter, B., Koteyko, N., & Atanasova, D. (2014). Sentiment Analysis: A Market-Relevant and Reliable Measure of Public Feeling? *International Journal of Market Research*, *56*(2), 231–247. https://doi.org/10.2501/IJMR-2014-014

Hu, Y., Deng, C., & Zhou, Z. (2019). A Semantic and Sentiment Analysis on Online Neighborhood Reviews for Understanding the Perceptions of People toward Their Living Environments. *Annals of the American Association of Geographers*, *109*(4), 1052–1073. https://doi.org/10.1080/24694452.2018.1535886

Lan, F. (2022). Research on Text Similarity Measurement Hybrid Algorithm with Term Semantic Information and TF-IDF Method. *Advances in Multimedia*. https://api.semanticscholar.org/CorpusID:248372311

Maier, D., Waldherr, A., Miltner, P., Wiedemann, G., Niekler, A., Keinert, A., Pfetsch, B., Heyer, G., Reber, U., Häussler, T., Schmid-Petri, H., & Adam, S. (2018). Applying LDA Topic Modeling in Communication Research: Toward a Valid and Reliable Methodology. *Communication Methods and Measures*, *12*, 118–193. https://api.semanticscholar.org/CorpusID:148860329

Pathak, U., & piyush Rai, Er. (2023). Sentiment Analysis: Methods, Applications, and Future Directions. *International Journal for Research in Applied Science and Engineering Technology*. https://api.semanticscholar.org/CorpusID:257320914

Shen, J., Baysal, O., & Shafiq, M. O. (2019). Evaluating the Performance of Machine Learning Sentiment Analysis Algorithms in Software Engineering. *2019 IEEE Intl Conf on Dependable, Autonomic and Secure Computing, Intl Conf on Pervasive Intelligence and Computing, Intl Conf on Cloud and Big Data Computing, Intl Conf on Cyber Science and Technology Congress (DASC/PiCom/CBDCom/CyberSciTech)*, 1023–1030. https://api.semanticscholar.org/CorpusID:207902038

Vaghela, V. B., & Jadav, B. M. (2016). Analysis of Various Sentiment Classification Techniques. *International Journal of Computer Applications*, *140*, 22–27. https://api.semanticscholar.org/CorpusID:18949360

*name of corresponding author