JOURNAL LA MULTIAPP
VOL. 07, ISSUE 02 (315-328), 2026
DOI: 10.37899/journallamultiapp.v7i2.3046

Systematic Review of Hyperparameter Adjustment and Evaluation Metrics in Bert-Based Sentiment Analysis

Aris Rifki Setiya Bahari[1], Fandy Setyo Utomo[1], Berlilana[1]

[1]*Amikom Purwokerto University, Banyumas, Indonesia*

**Corresponding Author: Aris Rifki Setiya Bahari*
Email: arisrifkisb11@gmail.com

Article Info

Article history:
Received 17 December 2025
Received in revised form 9 January 2026
Accepted 2 February 2026

Keywords:
Aspect-Based Sentiment Analysis
Bibliometric Analysis
BERT
Systematic Literature Review
Hyperparameter Adjustment

Abstract

The development of sentiment analysis towards Aspect-Based Sentiment Analysis (ABSA) has made significant progress thanks to deep learning technology, especially through the Bidirectional Encoder Representations from Transformers (BERT) architecture. Despite its increasing popularity, a comprehensive synthesis of global research patterns and optimal model configurations is still urgently needed. This study presents a Systematic Literature Review (SLR) combined with bibliometric analysis to examine BERT-based ABSA research indexed in Scopus. Using the PRISMA and VOSviewer frameworks for visualization, a total of 62 eligible articles up to mid-2025 were analyzed. The results of the study show a strong upward trend of publications with a peak in 2024, where China, India, and Indonesia emerged as the major contributors in this domain. Further, the review identified a critical technical standard for effective model training: the Adam optimizer was the most dominant choice, typically paired with a learning rate between 1e-5 to 2e-5 and a batch size of 16. Regarding performance evaluation, Accuracy and F1-Score are set as de facto standard metrics. These findings provide strategic guidance for researchers to optimize BERT implementation and identify future directions in more in-depth sentiment analysis tasks.

Introduction

In recent years, the global digital landscape has undergone a drastic transformation driven by the massive growth of digital data. This phenomenon has directly driven the urgent need for increasingly sophisticated text analysis techniques capable of handling large volumes of data efficiently. In the midst of this unstoppable flow of information, one of the computational approaches that has emerged as a vital instrument is sentiment analysis. Fundamentally, sentiment analysis is defined as a method for identifying, extracting, and categorizing users' subjective opinions of various entities, ranging from products, services, to social issues. The application of sentiment analysis has expanded significantly in various strategic domains, such as business to monitor customer satisfaction, politics to gauge public opinion, as well as the education and health sectors, where this method serves as a key supporting tool in data-driven decision-making (Karakolias, 2024; Hossain et al., 2024; Onesi-Ozigagun et al., 2024; Chao et al., 2025).

Nonetheless, traditional approaches to sentiment analysis that focus only on assessing the general polarity (positive, negative, or neutral) of a text often come to a dead end when faced with complex reviews. These conventional methods are often considered inadequate in capturing the nuances and complexity of opinions directed at specific aspects of a product or service. In response to these fundamental limitations, the research paradigm has shifted

ISSN: 2716-3865 (Print), 2721-1290 (Online)

towards the Aspect-Based Sentiment Analysis (ABSA) approach (Marutho & Rustad, 2024; Dhanal & Ghorpade, 2024). This approach offers higher granularity by allowing for the analysis of opinions based on specific aspects, such as product features, quality of service, or specific thematic issues. With this capability, ABSA is able to provide much more in-depth, detailed, and contextual analysis than sentence-level sentiment analysis or ordinary documents (Ahmad et al., 2025; Hua et al., 2024; Yang et al., 2025; Aziz et al., 2024).

The most significant breakthrough that accelerated the development and accuracy of ABSA was the adoption of deep learning-based methods, specifically through the utilization of Bidirectional Encoder Representations from Transformers (BERT) architecture (Devlin et al., 2019; Feng et al., 2024; Geetha & Renuka, 2021; Nusantara, 2022). The main advantage of the BERT model lies in its extraordinarily effective ability to capture the meaning of words based on a bidirectional context in a sentence. This ability becomes particularly relevant and crucial in the task of aspect extraction and sentiment classification that requires a deep syntactic and semantic understanding (Sirisha & Chandana, 2022; Wang et al., 2024; Yulianti & Nissa, 2024). Various empirical studies, including research by Chauhan and Mohana (2025), have proven that the integration of BERT in ABSA tasks consistently results in significant improvements in accuracy across various datasets and application domains (An et al., 2025; Alshaikh et al., 2024).

Technically, BERT uses the Transformer architecture and operates as a pre-trained language model. This model works by extracting contextual representations of words from large volumes of unlabeled text data (Alshattnawi et al., 2024; Deb& Chanda, 2022; Patil et al., 2023). Through the application of innovative training techniques, including Masked Language Modeling (MLM) and Next Sentence Prediction (NSP), BERT has demonstrated excellent performance on various Natural Language Processing (NLP) tasks, one of which is ABSA, as well as setting a new benchmark in language computing excellence. This phenomenal success of BERT then sparked a huge wave of interest among researchers, which resulted in the development of many derivative versions and adaptations of specific models of BERT for a variety of needs.

In the Transformer-based model ecosystem, BERT along with Generative Pre-trained Transformers (GPT) have become the primary base model options to fine-tune ABSA tasks. However, there are fundamental differences that make BERT superior for contextual understanding tasks. Its bidirectional framework makes BERT have context awareness at the word level, in contrast to GPT which generally only has awareness at the token level unidirectionally. Another crucial factor driving BERT's massive popularity over GPT is its open-source nature. This provides the freedom for the scientific community to utilize, modify, and further develop it without strict licensing barriers (Jazuli & Kusumaningrum, 2025. Considering these technical and availability aspects, this literature review specifically focuses its discussion on BERT-based models rather than other more common Transformer-based models.

Although BERT has been extensively studied and applied intensively within the ABSA domain, the current literature is still fragmented. There is still an urgent need to synthesize and critically analyze existing studies in order to gain a comprehensive understanding of the performance, limitations, and potential areas for improvement of this model. Indeed, BERT's ability to capture contextual information and its impressive performance on ABSA tasks has made it a game-changer in the NLP sector. However, despite all its successes, BERT is not without limitations. One of the main challenges associated with implementing BERT is its computational complexity and large resource requirements, which can make it difficult to deploy it in resource-constrained environments or on edge devices. Additionally, like many other large language models, BERT also has the potential to exhibit bias and inconsistencies,

ISSN: 2716-3865 (Print), 2721-1290 (Online)

which can be a serious problem in real-world applications (Qiu et al., 2020; Samant et al., 2022).

Furthermore, diversity in technical configurations such as optimizer selection, learning rate determination, and batch size often varies between studies, leading to difficulties in standardizing the most effective training methods. This is where the important role of a systematic literature review (SLR) comes in. SLR can provide new insights into the current state of research, identify research gaps, and inform future research directions in a more structured manner.

This study presents a Systematic Literature Review that intends to structurally analyze the current research on BERT-based models in order to describe the variation in its characteristics on the ABSA task. This research not only summarizes, but also incorporates bibliometric analysis to map global publication trends. By examining and combining findings from relevant research studies, this review aspires to present a clear understanding of the proper implementation of BERT-based models in ABSA tasks, including recommendations for optimal hyperparameter configuration and evaluation metrics (Shaik & Oussalah, 2024). It is hoped that the results of this synthesis can be a strategic guide for researchers and practitioners to optimize the implementation of BERT in the future.

Methods

This study adopts a mixed-method approach that integrates Systematic Literature Review (SLR) with Bibliometric Analysis. The integration of these two approaches was chosen to provide a comprehensive overview of the development, historical flow, and future direction of this field of research.

A bibliometric approach is used to quantitatively map the landscape of research activities, including identifying key contributors, international collaborations, and emerging trends in the relevant literature. Meanwhile, SLR is used to synthesize empirical findings qualitatively, especially related to the technical configuration of the BERT model. The entire analysis process was carried out using the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) standard framework (Kotei, 2023). The use of the PRISMA framework is crucial to improve understanding of the research gap and guide future research to make a significant contribution to academic discourse.

At this stage, the review protocol for conducting studies is designed with the aim of having a systematic and structured literature review. The protocol consists of research questions, search tools, search terms, inclusion criteria, exclusions, and quality assessments. The protocol also establishes a data analysis and synthesis strategy as a guide in extracting and summarizing the collected papers. The protocol is described in Table 1.

Table 1. Systematic *Literature Review Protocol*

Elements	Description
Research Questions	1. **RQ1.** What are the development trends of Aspect-based Sentiment Analysis (ABSA) research using the BERT approach over time, including the year of publication, number of publications, and the main application area? 2. **RQ2.** How is the allocation of research investigations related to Aspect-based sentiment analysis using the BERT approach? 3. **RQ3.** What types of hyperparameters are most considered to be properly set in training BERT-based models? 4. **RQ4**. What types of evaluation matrices are typically used for ABSA tasks?
Search Tools	Scopus AI, VOSViewer

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Search Terms	"BERT"; "Aspect Based Sentiment Analysis"
Inclusion Criteria	1. Papers from scientific journals published between 2020-2025. 2. Publications are indexed by Scopus. 3. Publication in the form of a journal. 4. The theme of the paper on ABSA using a BERT-based model. 5. The paper is written in English.
Exclusion Criteria	1. Unpublished scientific papers between 2020-2025. 2. Publications are not indexed by Scopus. 3. Publication is not in the form of a journal. 4. The theme of the paper is not about ABSA which uses a BERT-based model. 5. The paper is not written in English.
Quality Assessment Criteria	1. Clarity of the purpose of the research. 2. Contains a literature review, background, and research results. 3. Provide relevant conclusions. 4. Explain the method of development or optimization and evaluation of BERT-based models.
Data Analysis Strategy	Based on the selected paper, each variable that is considered appropriate and relevant to answer the research question, will be identified, collected, analyzed, and researched based on the quality of the study.
Data Synthesis Strategy	A data-based methodology, established on the results of data analysis from the paper, is used in the data synthesis. To answer the research question, a table containing a comparison list of each variable was compiled for each selected paper using a data-driven approach.

The initial phase of this scientific examination involves keyword selection, which can be achieved through a top-down macro methodology, evolving from broad keywords to more narrowly defined studies and topics. As a result, after evaluating the inherent limitations of previous research and the scarcity of studies addressing Aspect Based Sentiment Analysis, this investigation included the keyword "Aspect Based Sentiment Analysis" as a focal point in the article title, abstract, and keyword section. Furthermore, researchers use the scopus database for a variety of investigative purposes, including conducting a literature review, identifying subject matter experts, and monitoring research trends.

According to the search results taken on June 20, 2025, from the Scopus database using article titles, abstracts, and keywords: "ABSA AND BERT" across various computer science disciplines, stretching from initial publication in 2020 to the most recent in 2025, the total number of articles on Aspect-Based Sentiment Analysis Using the BERT Model Approach is 270 documents (see Figure 1) (Pranatawijaya et al., 2024; Santiago & Ludeña, 2024; Bai et al., 2020; Chauhan & Mohana, 2023). Following these findings, the process of filtering documents based on their classification. Articles were eliminated by document type: review (3), conference papper (122), non-English (4), access articles gold (47), green (13), bronze (7), hybrid gold (7) resulting in a total of 41 documents. The filtering results, categorized by document type, resulted in 62 articles. This document was then further analyzed in this study to answer RQ1. What is the development trend of Aspect-based Sentiment Analysis (ABSA) research using the BERT approach over time, including the year of publication, number of publications, and its main application area?, RQ2. How is the allocation of research investigations related to Aspect-based sentiment analysis using the BERT?, RQ3 approach. What types of hyperparameters are most considered to be properly set in training BERT-based models? RQ4. What types of evaluation matrices are typically used for ABSA tasks

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Results and Discussion

The results of this study focus on the findings of 62 articles in the Scopus database on aspect-based sentiment analysis using the BERT approach. This data is sourced from the identification of the number of articles published, publications throughout the year, and journal sources. The study will also highlight the most influential elements of aspect-based sentiment analysis using the BERT approach, including the authors, affiliates, and countries involved

This section presents the results and discussion of the research, which begins with a straightforward presentation of the application of the method based on the field data that has been collected.

RQ1: What are the trends in the development of Aspect-based Sentiment Analysis (ABSA) research using the BERT approach over time, including the year of publication, number of publications, and its main application areas

According to data taken from the Scopus database, this "Documents by year" graph presents the distribution of the number of research documents published from 2020 to early 2025. From 2020 to 2024, there is a significant upward trend in the number of publications. However, there will be a drastic decline in 2025 (considering that the 2025 data is likely to be incomplete).

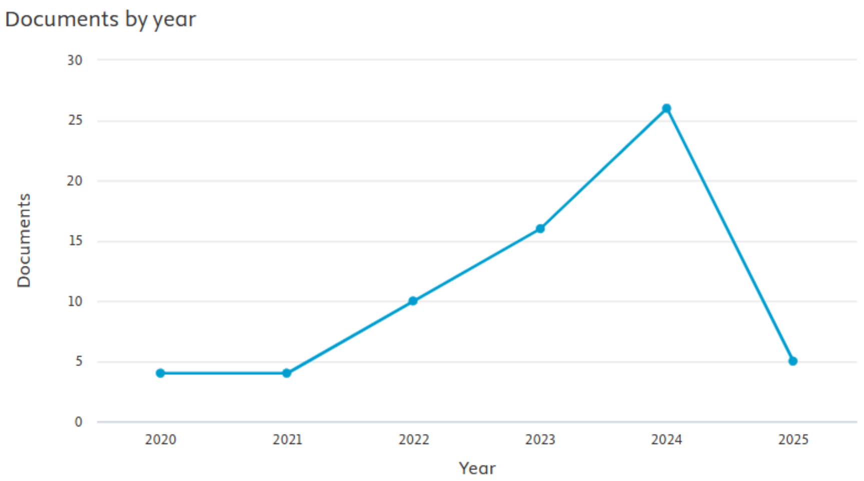

Figure 1. Document graph by year

Details Per Year, starting from 2020 There are 4 published research documents. The number of documents remained stable with 4 publications in 2021. Then in 2022 there was a clear increase, reaching 10 documents. The positive trend continues, with the number of publications increasing to 16 documents in 2023. The year 2024 shows the highest peak of publications in the period observed, with 26 documents. This indicates a strong increase in interest in the topic of Aspect Based Sentiment Analysis using the BERT method in that year. At the beginning of 2025 (until January 31, 2025, as mentioned in the text), only 5 documents were recorded. This decline is most likely due to the fact that the data for 2025 is still incomplete or only covers the early period of the year, rather than the entire year.

This graph indicates that research on aspect-based sentiment analysis using the BERT approach has shown strong growth momentum, especially between 2022 and 2024. The peak of publication in 2024 highlights that this topic is gaining significant attention in the scientific community. The decline in 2025 needs to be interpreted carefully, as it most likely reflects

ISSN: 2716-3865 (Print), 2721-1290 (Online)

unfinished data collection for the year, rather than an actual decline in interest. Complete data for 2025 will provide a more accurate picture of long-term trends.

RQ2: How to allocate research investigations related to Aspect-based sentiment analysis using the BE approach

The distribution analysis of the aspect based sentiment analysis using the bert method in 62 articles was carried out by categorizing articles based on classifications such as country, region, affiliation, source, and author, with a limit of only the top 10 articles in each classification. Insights into the allocation of relevant knowledge related to research-related aspect-based sentiment analysis using the BERT approach will be useful for scholars and practitioners in explaining the future research agenda, especially in the continuous advancement of the aspect-based sentiment analysis paradigm using the BERT approach.

First, the allocation of scientific research relevant to aspect-based sentiment analysis using the BERT approach which is categorized by country or geographical region is dominated by China with 22 articles, India with 11 articles, Indonesia with 10 articles, Saudi Arabia with 7 articles, and Iraq, Japan, Malaysia, the Netherlands, Pakistan and Yemen with 2 articles each

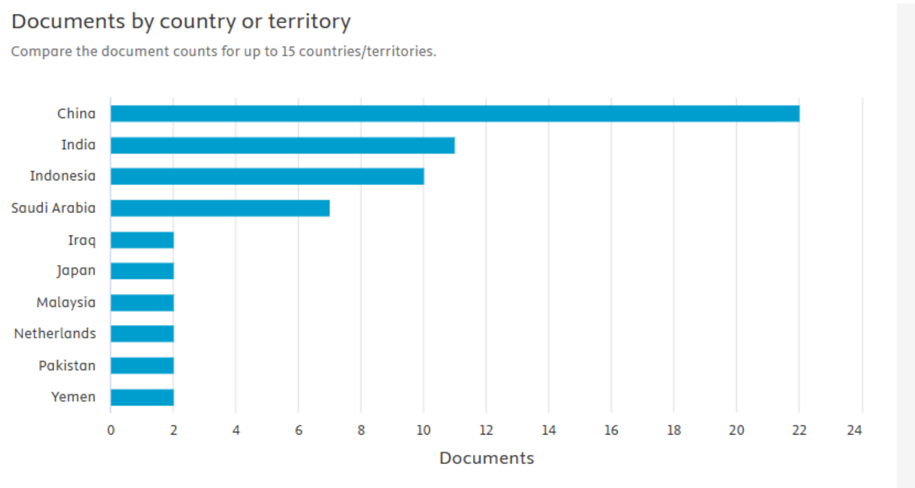

Figure 2. Document graph by country

The allocation of scientific investigations relevant to aspect-based sentiment analysis categorized by country or region shows China's dominance with 22 documents, followed by India and Indonesia as Other Major Contributors. India occupies the second position with a total of about 11 documents, while Indonesia is slightly below it with around 10 documents. These two countries also make substantial contributions in the area of research. Saudi Arabia's important role occupies the fourth position with a total of around 7 documents. This indicates that there are quite active research activities from the country. Countries with Moderate contributions such as Iraq, Japan, Malaysia, the Netherlands, Pakistan, and Yemen show a smaller number of documents, which is about 2 documents per country. Although they are not as numerous as the top countries, their contribution is still present and important. Data shows that aspect-based sentiment analysis research using BERT is not only concentrated in one region, but also spread across various countries, especially in Asia (China, India, Indonesia, Saudi Arabia, Iraq, Japan, Malaysia, Pakistan, Yemen) and also in Europe (Netherlands). This graph compares the number of documents for up to 15 countries/regions, although only 10 countries are clearly shown in the given image. This indicates that there is an effort to identify key contributors globally.

Overall, this graph provides a clear picture of the landscape of aspect-based sentiment analysis research with the BERT approach in various countries, with China as the main leader followed

ISSN: 2716-3865 (Print), 2721-1290 (Online)

by India and Indonesia. The researcher will also analyze the relationship between countries involved in aspect-based sentiment analysis research with the BERT approach using VOSViewer software. This phase is the main key to designing a prospective research plan systematically. The examination of VOSViewer's findings shows the inter-country relationship in investigating the subject of aspect based sentiment analysis using the bert method (see Figure 4).

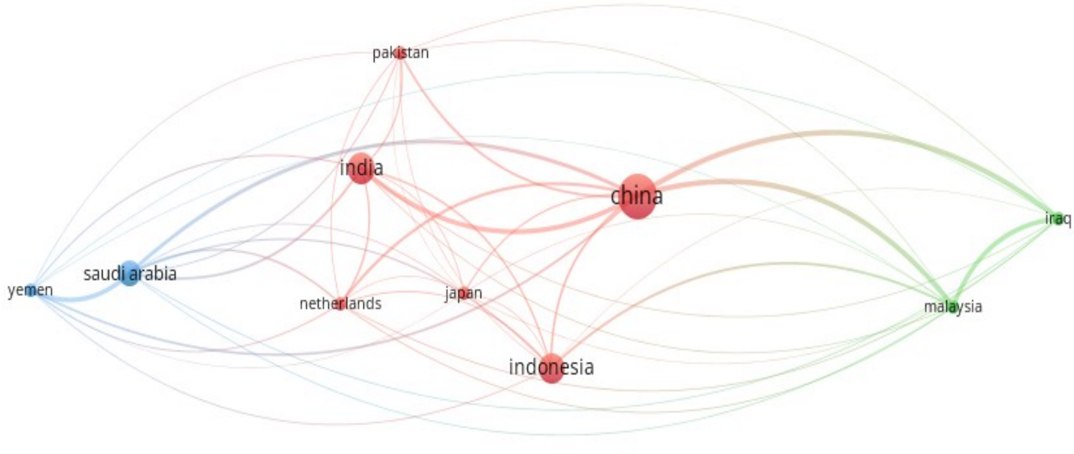

Figure 2. Country network visualization

This network graph shows global collaboration in aspect-based sentiment analysis (ABSA) research using the BERT model. China is a major hub, strongly connected to other countries such as India, Indonesia, and Pakistan, signaling its dominance in BERT research and implementation for ABSA. The involvement of many countries from different continents (Asia, Europe, the Middle East) confirms that the BERT model is relevant and adaptable for cross-linguistic and cultural sentiment analysis. The network demonstrates an active exchange of knowledge, proving that this approach has been tested and has the potential for widespread implementation in different countries for diverse sentiment analysis.

The college's contribution to the aspect-based sentiment analysis (ABSA) research using the BERT approach shows that King Abdulaziz University and the Faculty of Computing and Information Technology are leading with about 5 documents each, demonstrating the strong dominance of the institution in the Middle East. From Indonesia, the Sepuluh Nopember Institute of Technology (ITS) and Dian Nuswantoro University (UDINUS) also made significant contributions with around 3 documents, indicating the active role of campuses in Indonesia. In addition, various international affiliates such as South China Normal University (China), King Saud University (Saudi Arabia), Hangzhou Dianzi University (China), Erasmus Universiteit Rotterdam (Netherlands), Jaypee University of Information Technology (India), and Christ University (India) contributed about 2 documents each. This geographical diversity, ranging from the Middle East, Asia (including Indonesia and China), to Europe, indicates a global interest in this topic and the potential for collaboration between institutions for further development in BERT applications for aspect-based sentiment analysis. Overall, this data illustrates an active and diverse research landscape, dominated by academic environments in different parts of the world.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

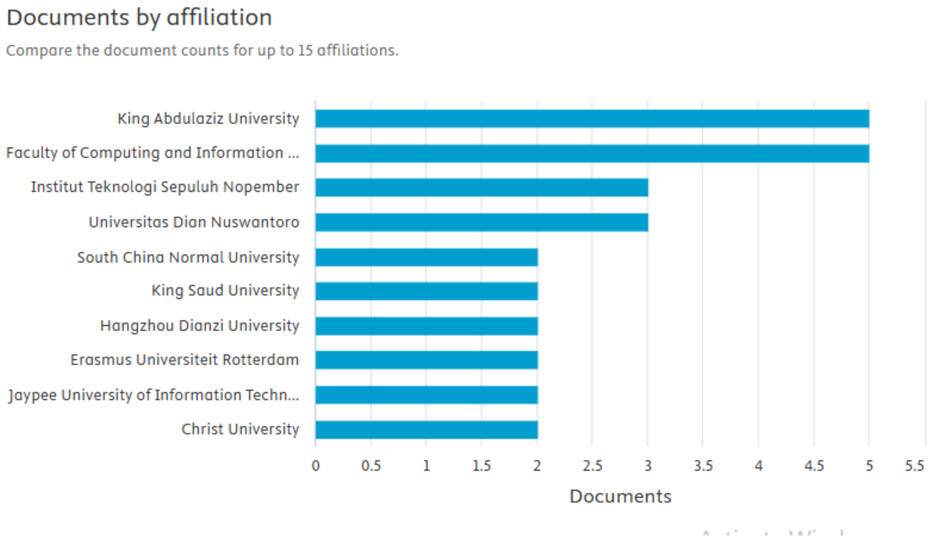

Figure 3. Document graph by college affiliation

This graph effectively identifies leading academic institutions that are active in publishing research on Aspect-based Sentiment Analysis using the BERT approach. King Abdulaziz University (and possibly its related Faculties) is the most prolific, followed by institutions from Indonesia that have made substantial contributions. The spread of affiliation also indicates that research in this area is a collaborative effort and attracts interest from various universities around the world.

In 2020, IEEE Access was the most dominant source with 3 documents. This shows the early role of IEEE Access as an important platform for publications in this field. Diversification and Increase of Activities In 2021, Electronics Switzerland began to emerge with 1 document. The year 2022 shows an increase in activity from several other journals. The International Journal of Advanced Computer Science and Applications and the International Journal of Intelligent Engineering and Systems each contributed 2 papers, while IEEE Access was reduced to 2 papers. In 2023, IEEE Access and the International Journal of Intelligent Engineering and Systems decreased to 1 document, while the International Journal of Advanced Computer Science and Applications remained stable at 2 documents. Electronics Switzerland also remained stable in 1 document. The year 2024 shows an interesting shift with Expert Systems with Applications appearing with 2 documents, while IEEE Access continues its decline to 1 document. The International Journal of Advanced Computer Science and Applications is still on 2 documents. Recent Trends (2025 (which is likely not yet complete) shows a general decline in all journals shown, with only IEEE Access recorded with 1 document. This is consistent with the incomplete pattern of data in the current year that is often seen in bibliometric analysis. (see Figure 6)

Throughout the observed period, IEEE Access became a consistent contributor despite fluctuations. The International Journal of Advanced Computer Science and Applications shows stability in its contributions. The emergence of Expert Systems with Applications in 2024 indicates that specific journals in the field of intelligent systems are also becoming important for this topic. Overall, this graph shows that while there are several journals that consistently publish BERT-based ABSA research, none of them dominate in absolute terms throughout the observed time period. Instead, there is a distribution of contributions among several journals, reflecting the interdisciplinary nature of this topic that may attract publications from different fields of computing and applications.

ISSN: 2716-3865 (Print), 2721-1290 (Online)

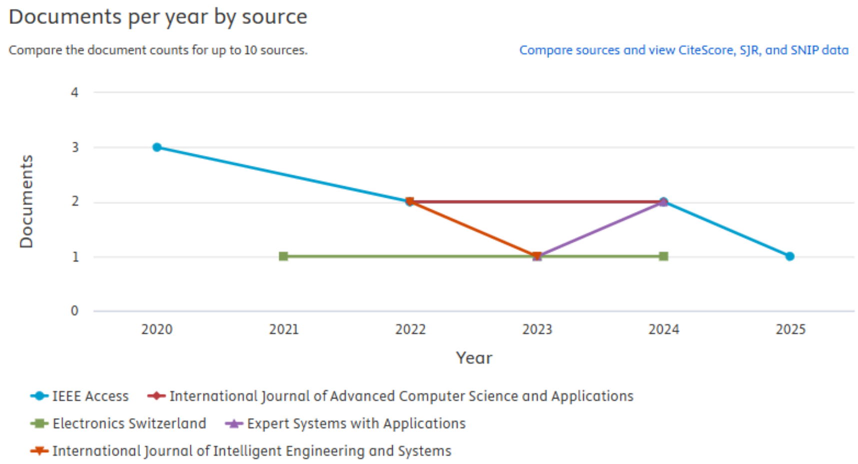

Figure 4. Document graphs by year and source

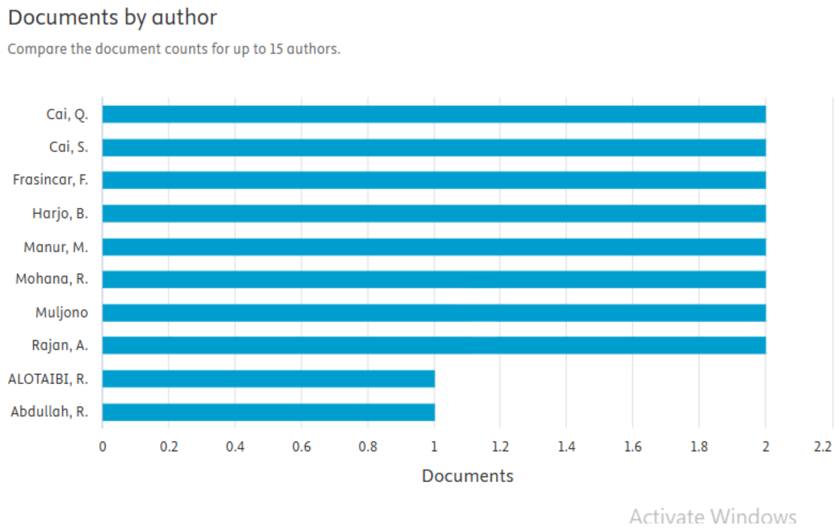

Figure 5. Document graphs by author

Most of the writers featured have equal contributions. Top eight authors (Cai, Q.; Cai, S.; Frasincar, F.; Harjo, B.; Manur, M.; Mohan, R.; Muljono; and Rajan, A.) Each has published 2 documents. Meanwhile, ALOTAIBI, R. and Abdullah, R. each have 1 document. This indicates that contributions in this area tend to be scattered among many researchers, with no single author being predominantly the most productive based on the data presented.

RQ3: What types of hyperparameters are most considered to be properly set in training BERT-based models?

BERT is the latest modification of a series of neural models that make extensive use of pre-training and has proven useful in a variety of NLP tasks, including the classification of texts from datasets. To study word representation, BERT uses the Transformer architecture, which consists of multiple layers of encoder. He analyzes sequential inputs, such as words in text, using an attention mechanism. The attention mechanism helps the model better understand the context of the text by allowing it to capture the connections between distant words. During the training process, many parameter variables are applied to organize the process, resulting in a lot of variation in the results. The setting of certain hyperparameters plays a crucial role in optimizing the training process and the performance resulting from deep learning models such as BERT. The right settings in setting hyperparameters can have a noticeable effect on the training process and the performance of the resulting model on certain tasks. Learning rate, batch size, and optimizer type are the three hyperparameters that are most considered to

ISSN: 2716-3865 (Print), 2721-1290 (Online)

be properly regulated in BERT-based model training. Understanding the usefulness and influence of each hyperparameter is essential to achieve optimal results. Table 1 lists the different types of hyperparameter settings applied in previous studies.

Optimizer is one of the most considered hyperparameters to be set correctly for training deep learning models such as BERT. Optimizer is an algorithm that updates the network weights during the model training process. Efficient optimizers can contribute to faster convergence and better model performance outcomes. Adam is an optimizer that is often applied to BERT-based model training. Adam is an adaptive and efficient optimizer that uses momentum estimation and RMSProp. Interestingly, all of the studies reported in the peer-reviewed paper applied the Adam optimizer, except for one study as reported in [7] that implements SGD for optimizers. As mentioned in their paper, the authors' reason for using SGD is "that SGD performance is tuned to identify ideal parameters to improve relevance through Grid Search and Random Search". Nonetheless, the application of the Adam optimizer outperformed the SGD in some studies as reported in the reviewed paper. Based on the literature, we found that the average Optimizer applied is an Adam Optimizer.

Table 2. Table of use of previous research hyperparameters

Yes	Hyperparameter	Varaint	Quantity
1	Optimizer	Adam	48
		Adam	2
		SGD	1
		Adadelta	1
2	Learning Rate	2nd-5 (0.00002)	26
		1st-5 (0.00001)	18
		3rd-5th (0.00003)	12
		5th-5 (0.00005)	11
		1st-3 (0.001)	8
		1st-4 (0.0001)	3
3	Batch Size	16	29
		32	26
		8	10
		64	6
		24	4
		128	2
4	Dropout Rate	0.1	20
		0.5	10
		0.2	4
		0.3	3
	Regularization	L2 Regularization	15

The second hyperparameter that is most considered to be regulated is the Learning Rate.The learning rate determines how much the weight of the network changes in each iteration of the training. The pace of high-value learning can speed up the training process but has the potential to cause the training process to miss out on the acquisition of optimally performing models. Meanwhile, too low a learning rate has the potential to lead to the acquisition of a model with optimal performance but at the cost of slowing down the training process. Based on the literature, we found that the average applied learning rate was between 2.00E-05 and 1.00E-05. Most of these studies applying the 2.00E-05 and 1.00E-05 learning rate levels may be due to original research on BERT as reported in a paper titled "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", where the authors mention that the optimal hyperparameter values are actually specific to a particular training context, however

ISSN: 2716-3865 (Print), 2721-1290 (Online)

some value ranges generally work well in all types of tasks. For their study, however, the authors specifically set their learning rate values to 2.00E-05, 1.00E-05, and 3.00E-05.

The third hyperparameter most considered to be set is batch size, which refers to the number of data samples handled simultaneously in a single iteration of the training model. Larger batch sizes can speed up the training process. However, they can degrade the quality of models trained in recognizing patterns in new (invisible) data. Smaller batch sizes, on the other hand, will lead to a trained model with higher performance due to better pattern recognition capabilities against new (invisible) data but will take longer to train. Based on the reviewed papers, the most widely applied batch size is 16. The main consideration for selecting a specific batch size level for model training is usually to take into account the available graphical processing unit (GPU) or central processing unit (CPU) memory to avoid memory exhaustion (OOM) issues. Next, we need to examine the trade-off between the training speed and the expected performance of the trained model when selecting a batch size. Small batch sizes can be slower per epoch but can result in better model performance.

According to several studies, proper hyperparameter regulation has a significant impact on BERT-based models. Based on the paper we reviewed, the three hyperparameters that were most considered for training the model were optimizer, learning rate, and batch size. Most studies also apply certain types of settings to these three hyperparameters. The type of arrangement practiced in some of these studies has provided valuable insights for future work in the development and optimization of BERT-based ABSA. It is highly recommended to identify the objectives of the development or optimization of existing models before deciding on certain types of hyperparameter settings in our work. Several factors should be considered when deciding on the setting of certain hyperparameters, including the hardware capacity used, the size of the dataset (corpus), the type of NLP task, and the type of model intended to be developed or optimized (i.e. generic-base models or specific-specific models).

RQ4: What type of evaluation matrix is typically used for ABSA tasks?

A number of evaluation metrics for evaluating BERT-based models for ABSA have been mentioned in the paper we reviewed. The metrics, as described in Table 5, consist of F1-Score, Accuracy, Precision, Recall, Area Under the ROC Curve (AUC), and Root Mean Squared Error. Accuracy is used in most studies, where it is usually combined with F1-Score, Precision, and Recall. This is due to the fact that most of the papers we reviewed reported studies on text classification. For binary, multi-class or multi-label text classification types, this type of metric calculated based on the results of the confusion matrix has become the de facto standard for evaluating model performance. The selection of evaluation metrics should be based on the type of NLP task at hand.

The Accuracy, F1-Score, Precisian, Recall, and ROC (AUC) metrics provide an informative overview of the model's performance in predicting (classifying) the correct labels. This score is widely used in a variety of "discriminatory" AI NLP tasks, such as sentiment analysis and information extraction. This is reflected in the number of studies using this type of metric shown in table 2. Table of users

Table 3. Table of Usage of Evaluation Matrix

Yes	Evaluation matrix	Quantity
1	Accuracy	56
2	F1 Score	53
3	Recall	30
4	Accuracy	29
5	AUC (Area Under Curve)	3
6	RMSE (Root Mean Squared Error)	1

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Conclusion

Based on an extensive review of the literature that has gone through a rigorous selection protocol, this study formulates three main conclusions regarding the application of BERT in Aspect-Based Sentiment Analysis (ABSA). First, the effectiveness of a variant of a BERT-based model depends largely on its suitability to a specific task. Researchers need to consider three crucial aspects in choosing a model: the wide range of method exploration opportunities, the specificity of the domain (such as finance, medical, or legal), and the target language used. This emphasizes that there is no one universal model, but must be adapted to the context of the application.

Second, from the technical side, hyperparameter regulation plays a vital role in optimizing model performance. This review identifies that optimizer, learning rate, and batch size are the three most critical parameters to adjust. The optimal configuration of these parameters does not stand alone, but rather must consider the capacity of the hardware, the size of the data corpus, and the purpose of model development whether intended as a generic or specific model. Third, in terms of performance evaluation, the research community has adopted consistent measurement standards. Accuracy, F1-score, precision, and recall metrics are the most widely used key indicators to validate the success of the model in ABSA tasks. These findings are expected to serve as a strategic guide for the development of BERT-based ABSA research in the future.

References

Ahmad, W., Khan, H. U., Alarfaj, F. K., & Alreshoodi, M. (2025). Aspect-Base Sentiment Analysis: A Comprehensive Review and Open Research Challenges. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3555744

Alshaikh, K. A., Almatrafi, O. A., & Abushark, Y. B. (2024). BERT-based model for aspect-based sentiment analysis for analyzing Arabic open-ended survey responses: A case study. *IEEE Access, 12*, 2288–2302. https://doi.org/10.1109/ACCESS.2023.3348342

Alshattnawi, S., Shatnawi, A., AlSobeh, A. M., & Magableh, A. A. (2024). Beyond word-based model embeddings: Contextualized representations for enhanced social media spam detection. *Applied Sciences*, *14*(6), 2254. https://doi.org/10.3390/app14062254

An, Z., Wan, H., Xiong, T., & Wang, B. (2025). LePB-SA4RE: A lexicon-enhanced and prompt-tuning BERT model for evolving requirements elicitation from app reviews. *Applied Sciences, 15*(5), Article 2282. https://doi.org/10.3390/app15052282

Aziz, K., Ji, D., Chakrabarti, P., Chakrabarti, T., Iqbal, M. S., & Abbasi, R. (2024). Unifying aspect-based sentiment analysis BERT and multi-layered graph convolutional networks for comprehensive sentiment dissection. *Scientific reports*, *14*(1), 14646. https://doi.org/10.1038/s41598-024-61886-7

Bai, P., Xia, Y., & Xia, Y. (2020). Fusing knowledge and aspect sentiment for explainable recommendation. *IEEE Access, 8*, 137150–137160. https://doi.org/10.1109/ACCESS.2020.3012347

Chao, K., Sarker, M. N. I., Ali, I., Firdaus, R. R., Azman, A., & Shaed, M. M. (2023). Big data-driven public health policy making: Potential for the healthcare industry. *Heliyon*, *9*(9). https://doi.org/10.1016/j.heliyon.2023.e19681

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Chauhan, A., & Mohana, R. (2023). Aspect-level sentiment modeling using transformer architectures. *Scalable Computing: Practice and Experience, 24*(3), 277–286. https://doi.org/10.12694/scpe.v24i3.2444

Chauhan, A., Sharma, A., & Mohana, R. (2025). An enhanced aspect-based sentiment analysis model based on RoBERTa. *Computer Systems Science and Engineering, 49*, 193–202.

Deb, S., & Chanda, A. K. (2022). Comparative analysis of contextual and context-free embeddings in disaster prediction from Twitter data. *Machine Learning with Applications*, *7*, 100253. https://doi.org/10.1016/j.mlwa.2022.100253

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of deep bidirectional transformers for language understanding*. Cornell University. https://arxiv.org/abs/1810.04805

Dhanal, R. J., & Ghorpade, V. R. (2024). Aspect-based sentiment analysis using topic modelling and machine learning. *International Journal of Electrical and Computer Engineering, 14*(6), 6689–6698. https://doi.org/10.11591/ijece.v14i6.pp6689-6698

Feng, A., Liu, T., Li, X., Jia, K., & Gao, Z. (2024). Dual syntax-aware graph attention networks with prompt for aspect-based sentiment analysis. *Scientific Reports, 14*(1), Article 74668. https://doi.org/10.1038/s41598-024-74668-y

Feng, J., Cai, S., Li, K., Chen, Y., Cai, Q., & Zhao, H. (2023). Fusing syntax- and semantics-based graph convolutional network for aspect-based sentiment analysis. *International Journal of Data Warehousing and Mining, 19*(1). https://doi.org/10.4018/IJDWM.319803

Geetha, M. P., & Renuka, D. K. (2021). Improving the performance of aspect-based sentiment analysis using fine-tuned BERT base uncased model. *International Journal of Intelligent Networks, 2*, 64–69. https://doi.org/10.1016/j.ijin.2021.06.005

Hossain, Q., Yasmin, F., Biswas, T. R., & Asha, N. B. (2024). Data-Driven Business Strategies: A Comparative Analysis of Data Science Techniques in Decision-Making. *Sch J Econ Bus Manag*, *9*, 257-263. https://doi.org/10.36347/sjebm.2024.v11i09.002

Hua, Y. C., Denny, P., Wicker, J., & Taskova, K. (2024). A systematic review of aspect-based sentiment analysis: domains, methods, and trends. *Artificial Intelligence Review*, *57*(11), 296. https://doi.org/10.1007/s10462-024-10906-z

Jazuli, A., & Kusumaningrum, R. (2025). Optimizing aspect-based sentiment analysis using BERT for comprehensive analysis of Indonesian student feedback. *Applied Sciences, 15*(1), Article 172. https://doi.org/10.3390/app15010172

Karakolias, S. (2024). Mapping data-driven strategies in improving health care and patient satisfaction. *World Journal of Advanced Engineering Technology and Sciences*. https://doi.org/10.30574/wjaets.2024.13.1.0444

Kotei, E., & Thirunavukarasu, R. (2023). A systematic review of transformer-based pre-trained language models through self-supervised learning. *Information*, *14*(3), 187.

Marutho, D., & Rustad, S. (2024). Optimizing aspect-based sentiment analysis using sentence embedding transformer, Bayesian search clustering, and sparse attention mechanism. *Journal of Open Innovation: Technology, Market, and Complexity, 10*(1), Article 100211. https://doi.org/10.1016/j.joitmc.2024.100211

ISSN: 2716-3865 (Print), 2721-1290 (Online)

Nusantara, U. M. (2022). Comparison of IndoBERT-Lite and RoBERTa in text mining for Indonesian language question answering application. *International Journal of Innovative Computing, Information and Control, 18*(6), 1719–1734. https://doi.org/10.24507/ijicic.18.06.1719

Onesi-Ozigagun, O., Ololade, Y. J., Eyo-Udo, N. L., & Oluwaseun, D. (2024). Data-driven decision making: Shaping the future of business efficiency and customer engagement. *International Journal of Multidisciplinary Research Updates*, *7*(2), 19-29. https://doi.org/10.53430/ijmru.2024.7.2.0031

Patil, R., Boit, S., Gudivada, V., & Nandigam, J. (2023). A survey of text representation and embedding techniques in nlp. *IEEe Access*, *11*, 36120-36146. https://doi.org/10.1109/ACCESS.2023.3266377

Pranatawijaya, V. H., Sari, N. N. K., Rahman, R. A., Christian, E., & Geges, S. (2024). Unveiling user sentiment: Aspect-based analysis and topic modeling of ride-hailing and Google Play app reviews. *Journal of Information Systems Engineering and Business Intelligence, 10*(3), 328–339. https://doi.org/10.20473/jisebi.10.3.328-339

Qiu, X., Sun, T., Xu, Y., & Huang, X. (2020). Pre-trained models for natural language processing: A survey. *Science China Technological Sciences, 63*(10), 1872–1897.

Rajan, A., & Manur, M. (2024). Aspect-based sentiment analysis using a novel ensemble deep network. *IAES International Journal of Artificial Intelligence, 13*(2), 1668–1678. https://doi.org/10.11591/ijai.v13.i2.pp1668-1678

Samant, R. M., Bachute, M. R., Gite, S., & Kotecha, K. (2022). Framework for deep learning-based language models using multi-task learning in natural language understanding: A systematic literature review and future directions. *IEEE Access, 10*, 17078–17097. https://doi.org/10.1109/ACCESS.2022.3149798

Santiago, M., & Ludeña, V. (2024). Aspect-based sentiment analysis using social media data. *Advances in Hospitality and Tourism Research*. https://doi.org/10.30519/ahtr.1436175

Shaik, F. A., & Oussalah, M. (2024). Mining mobile emergency communication applications in Nordic countries. *International Journal of Disaster Risk Reduction, 108*, Article 104566. https://doi.org/10.1016/j.ijdrr.2024.104566

Sirisha, U., & Chandana, B. S. (2022). Aspect-based sentiment and emotion analysis. *International Journal of Engineering Research and Technology, 13*(11).

Wang, Y., Yang, N., Miao, D., & Chen, Q. (2024). Aspect-guided multi-graph convolutional networks for aspect-based sentiment analysis. *Data Intelligence, 6*(3), 771–791. https://doi.org/10.3724/2096-7004.di.2024.0052

Yang, R., Ju, X., Guo, C., Ding, L., Li, M., & Zhang, B. (2025). A unified review of aspect sentiment triplet extraction methods in aspect-based sentiment analysis: R. Yang et al. *Knowledge and Information Systems*, *67*(11), 9779-9820. https://doi.org/10.1007/s10115-025-02519-x

Yulianti, E., & Nissa, N. K. (2024). Aspect-based sentiment analysis of Indonesian customer reviews using IndoBERT: Single-sentence and sentence-pair classification approaches. *Bulletin of Electrical Engineering and Informatics, 13*(5), 3579–3589. https://doi.org/10.11591/eei.v13i5.8032

ISSN: 2716-3865 (Print), 2721-1290 (Online)