Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15888

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

From Methodologies to Metrics: A Review of Aspect-Based Sentiment Analysis Approaches

Marwa Esmaeel[1)], **Alaa Taqa**[2)]
[1,2)]University of Mosul, Iraq
[1)] marwa.24esp8@student.uomosel.edu.iq, [2)] alaa.taqa@uomosel.edu.iq

Submitted :Feb 7, 2026 | **Accepted** : Mar 9, 2026 | **Published** : Apr 2, 2026

Abstract: Aspect-based sentiment analysis allows for a refined interpretation of sentiments as it correlates opinions with specific characteristics of the text. To this end, this review aggregates the latest advancements and addresses methodological and evaluation shortcomings in aspect-based sentiment analysis. A systematic research procedure that has been applied through searching on Institute of Electrical and Electronics Engineers Digital Library, ScienceDirect, SpringerLink, the Association for Computing Machinery Digital Library and Google Scholar. The period range from 2015 through 2025, and the inclusion criteria included having specific context of aspect-based sentiment analysis, coverage in the core subtasks, benchmark data or credible evaluations, and exclusion criteria removed non-peer-reviewed papers, non-English texts without full text and studies without evaluative evidence. A total of 50 studies are retained after narrowing and editing. Study distribution encourages a systematic review, with common deficiencies in implicit aspect extraction, cross-domain and cross-lingual generalization, data imbalance, and interpretability identified. As such, they predominantly focus on transformer models and generative language models, evaluate performance via F1-score and accuracy with the exception of relatively little work on multilingual, low-resource settings. The unique contribution of this systematic review is an integrated, task-aware framework that identifies subtasks, models, datasets, and assessment practices, providing a review methodology to orient robust and interpretable model development through transparent synthesis and reproducibility.

Keywords: aspect-based sentiment analysis; generative language models; review methodology; systematic review; transformer models.

INTRODUCTION

Sentiment Analysis (SA) is an integral part of NLP approach, with a focus on sentiment assessment of text. Traditional SA deals with document/sentence-level and gives just one polarity label, which is suitable for coarse-grained judgements, but this type of process cannot cope with the contradictory and complex user opinion in real-world content, namely reviews/surveys (Wankhade et al., 2022) (Cui et al., 2023).

As online reviews grew rapidly, the need for fine-grained analyses has intensified, and Aspect-Based Sentiment Analysis (ABSA) has gained popularity as, in it, elements are filtered out and sentiment is rated based on the elements to achieve more context-sensitive interpretation (Ahmad et al., 2025; D'Aniello et al., 2022).

ABSA is being increasingly adopted in education, healthcare, and public systems for quality assessment and policy evaluation (Chang et al., 2022). It is backed by extensive resources and multilingual datasets, including Arabic and Bangla, which is what makes it of global significance (Alshaikh et al., 2024; Eid et al., 2024; Hasan et al., 2024). In terms of methodology, ABSA has already progressed from lexicon-based and traditional ML methods to more advanced deep learning and

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15888
e-ISSN : 2541-2019
p-ISSN : 2541-044X

transformer approaches that have improved contextualization (Alshaikh et al., 2024). It also employs generative and large language models in sequence-to-sequence and prompt-based forms for clustering subtasks (Zhang et al., 2024; Florindi et al., 2024).

Although there has been substantial progress in techniques used in this work, ABSA still faces challenges like implicit aspect/opinion extraction, strong cross-domain and cross-lingual transfer, dataset imbalance, and limited interpretability of high-capacity neural models (Hua et al., 2024) (Islam et al., 2024) (Ahmed et al., 2025). The previous surveys have often explored isolated modeling paradigms, subtasks, or application domains, leading to fragmented insights and poor capture of interdependencies among tasks, methods, datasets, and evaluation particularly in the transformers and large language models era (Chen et al., 2024) (Chen et al., 2025).

Recent surveys provide useful baselines but tend to focus on specific dimensions. Hua et al. (2024) map domains and methods, Chen et al. (2024, 2025) focus on bibliometric trends, Zhang et al. (2024) address implicit aspect and quadruple extraction, and Ahmad et al. (2025) consolidate the shortcomings of current models. These reviews do not offer a comprehensive task-aware mapping that identifies and explicitly connects subtasks to modeling paradigms, datasets, and evaluation methods in a single framework. The highlighted gap is the lack of a unified synthesis linking tasks, methods, benchmarks, and measures, and explaining methodological choices and design limitations across languages, regions, and domains.

To address these weaknesses, we present an integrated task-aware summary of the current state, diagnose common issues, and propose prospective directions. The following systematic procedures are used in the selection of the study: limiting our selection to include only ABSA, sampling from different methodological paradigms such as lexicon-based and classical ML, deep learning, transformers, and generative LLMs, benchmarking datasets across languages and domains, surveying broad domains (for more relevant studies) and the most recent studies (2022–2025, maintaining the basic benchmark from 2015–2021) and in emerging topics such as explainable ABSA, multilingual/low-resource modeling, and integration of LLMs. Multiple domains and languages were considered to reduce the influence of English as a prevailing language bias (for example, we treated Arabic and Bangla as minority languages). Table 1 presents a summary of the 50 papers included in this review.

Table 1: Classification of Selected Papers by Primary Contribution Focus.

Category	Number of Papers	Research Distribution by Discipline	Examples
Methodological Approaches	20	40%	Gogineni (2023), Alshaikh (2024), Gu (2024), Gupta (2024), Apostol (2025)
Benchmark Datasets	8	16%	Pontiki (2015), Hercig (2016), Jiang (2019), Obiedat (2021), Eid (2024), Hasan (2024)
Comprehensive Surveys/Reviews	7	14%	Chen (2024, 2025), Hua (2024), Zhang (2024), Ahmad (2025).
Explainability and Interpretability	5	10%	Ribeiro et al. (2016), Lundberg & Lee (2017), Diwali (2023), Perikos (2024), Toms (2025).
Multilingual and Low-Resource	4	8%	Obiedat (2021), Abdulkader & Muhammad (2022), Alshaikh (2024), Eid (2024).
Domain-Specific Applications	5	10%	Shaik (2022), Awadh (2025), Han (2025), Taj (2025),
Theoretical Foundations	1	2%	D'Aniello et al. (2022)
Total	50	100%	

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15888
e-ISSN : 2541-2019
p-ISSN : 2541-044X

The review's remainder is organized as follows. The contributions of this work will be presented in Section 2. Followed by Section 3 which introduces the fundamentals of Aspect-Based Sentiment Analysis, including core semantic components and interconnected subtasks. Section 4 provides a comprehensive analysis of methodological paradigms, from lexicon-based approaches to generative large language models. Section 5 examines benchmark datasets across multiple languages and domains, while Section 6 discusses evaluation metrics and practices. Section 7 identifies persistent challenges and open research questions. Section 8 highlights future research opportunities and emerging directions. Finally, Section 9 synthesizes the review with a summary of main findings and recommendations.

Contributions

An integrated review aims to bridge the gaps of the ABSA literature through a task aware synthesis that maps various task constructions on models, datasets, evaluation strategies and positions them within prevailing research agendas. It contributes:

(1) A single, task aware synthesis that encompasses subtasks, methods and evaluation into a coherent framework.

(2) Temporal methodological analysis from lexicon-based methods through deep learning and transformer models to recent LLMs with comparative tables and figures summarizing strengths, limitations, and task suitability.

(3) A review of the benchmark datasets and evaluation practices in different languages and domains and, in particular, design gaps and explainability-oriented evaluation.

(4) Overall perspectives on ongoing issues (implicit aspect extraction, domain adaptability, dataset imbalance) as well as trends and new developments in the field (e.g., explainable ABSA, meta-learning, multilingual modeling, and LLM integration).

(5) A formal guide for building sound, interpretable, and generalizable ABSA systems, which relates tasks, methods, and evaluation strategies.

Research Objective

The objective of this review is to offer a brief, task-aware synthesis of ABSA work by linking task descriptions with modeling approaches, datasets, and evaluation practices, tracing the evolution in methodology from lexicon-based and classical ML approaches towards deep learning, transformers, and recent LLMs. It summarizes key benchmarks across languages and domains to identify design gaps and explainability needs, and consolidates persistent challenges – such as implicit aspect extraction, domain adaptability, and dataset imbalance – alongside emerging directions including explainable ABSA, meta-learning, multilingual modeling, and LLM integration.

METHOD

Review Methodology

The follow-up in this review is performed to standard PRISMA-aligned systematic literature review (SLR), to assure methodological traceability and reproducibility. A search and screening methodology was used. Databases queried for this study were the Institute of Electrical and Electronics Engineers Digital Library, ScienceDirect, SpringerLink, the Association for Computing Machinery Digital Library, and Google Scholar. The search strategy utilized keyword combinations: "aspect-based sentiment analysis," "aspect term extraction," "aspect category detection," "sentiment polarity," "Transformer models," "generative language models," and "explainable sentiment analysis." The period from 2015 to 2025 was used to preserve basic benchmarks and present more recent developments. Inclusion criteria were: direct relevance to aspect-based sentiment analysis, focus on core subtasks, benchmark data, or good evaluation; while, exclusion criteria were non-peer-reviewed sources, papers without available full text, non-English papers without translation, and works with no evaluative evidence. Screening took place in two stages (title/abstract, then full text), yielding a total of fifty studies to be synthesized.

Study selection was established based on the guidelines created by PRISMA framework, to encourage transparent and traceable methodology. An initial set of records resulted from a search of five databases that was subsequently removed by duplicates before it moved on to the title and abstract

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15888
e-ISSN : 2541-2019
p-ISSN : 2541-044X

screening phase. Following the first round of exclusion of studies that misinformed the aspect-based sentiment analysis, the full source text of the rest of the studies to be studied for methodological and empirical appropriateness. We excluded studies that did not provide rigorous criteria for ABSA tasks or that lacked precise evaluation information and full methodological information, leaving fifty studies in the final sample on which the analysis was based.

In order to reduce theoretical ambiguity, in this study we classify models into three main types: pre-trained language models derived from transformational structures specifically oriented towards ABSA tasks, for example to perform the task according to the BERT framework; large language models of a generative nature capable of executing instructions; and generative ABSA formulations that reproduce the task as a sequential generation that can be done using either of the preceding two types. Moreover, the survey makes use of standardized definitions of ABSA subtasks (ATE, OTE, ACD, ASC, AOPE, ASTE, and ACOS) to avoid any comparison mismatch. Besides, the meaning of implicit information is standardized where implicit content is distinguished from implicit opinion; this way it is possible to use terms that further increase terminological precision in both analysis and discussion.

In order to comprehend Aspect-Based Sentiment Analysis (ABSA), it becomes important to decompose the problem into its core semantic components and subtasks. Recent studies indicate that ABSA is moving away from rule-based methods toward unified neural frameworks that can model multiple subtasks simultaneously (Hua et al., 2024).

A good conceptual framework is needed to interpret methodological advances and keep terminologies consistent. ABSA is recognized as a fine-grained sentiment analysis approach that identifies specific features or entities in text and analyzes the sentiment expressed towards each, allowing deep-text opinion mining beyond just document or sentence-level classification (D'Aniello et al., 2022; Zhang & Shafiq, 2024).

From a practical perspective, Aspect-Based Sentiment Analysis (ABSA) aims to identify and relate four key semantic elements expressed in a text: Aspect term an aspect term refers to the explicit word or phrase that denotes the subject of the opinion. For example, in the sentence "This mobile phone's battery life is great," the phrase "battery life" represents the aspect term. Aspect categories an aspect category abstracts the aspect term into a predefined, higher-level category – such as "battery" – making it easier to generalize to semantically related expressions. Opinion terms an opinion term embodies the linguistic expression that conveys the emotion toward the aspect, as in the word "great" in the same sentence mentioned above. Sentiment polarity Finally, sentiment polarity reflects the orientation of the expressed opinion, which is usually categorized as positive, negative, or neutral (D'Orazio et al., 2025) (Ahmed et al., 2025). Together, these elements enable an accurate interpretation of emotions that goes beyond document or sentence level analysis. These four elements are not extracted in isolation, but are linked to each other through a set of increasingly complex subtasks that define the operational scope of ABSA systems (Chebolu et al., 2023).

ABSA Subtasks and Task Formulations

Early approaches addressed ABSA subtasks sequentially within pipeline architectures, whereas contemporary models increasingly favor joint or generative formulations to reduce error propagation and better capture interdependencies (Hua et al., 2024); (Zhang et al.,2023). Table 2 summarizes the most widely studied ABSA subtasks, together with their objectives and typical modeling paradigms.

Table 2: Core ABSA Subtasks and Typical Modeling

Subtask	**Description**	**Reference**
Aspect Term Extraction (ATE)	Identifies explicit opinion targets within the text (e.g., extracting *"battery life"* from *"battery life is great"*)	((Hua et al., 2024) (Awadh et al., 2025)
Opinion Term Extraction (OTE)	Extracts words or phrases expressing sentiment toward identified aspects (e.g., *"great"*)	(Aziz et al., 2024)

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15888
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Aspect Category Detection (ACD)	Assigns aspect terms or phrases to predefined semantic categories (e.g., mapping *"battery life"* to *"battery"*)	(Zhang & Shafiq, 2024)
Aspect Sentiment Classification (ASC)	Establishes the sentiment polarity associated with a certain aspect or category	(Hua et al., 2024)
Aspect–Opinion Pair Extraction (AOPE)	Jointly identifies and links aspect terms with their corresponding opinion terms	(Awadh et al., 2025)
Aspect–Sentiment Triplet Extraction (ASTE)	Extracts integrated triplets {aspect, opinion, polarity} in a unified manner	(Hua et al., 2024)
Aspect–Sentiment Quadruple Extraction (ASQE / ACOSQE)	Extracts complete quadruples {aspect term, aspect category, opinion term, polarity} as a unified structure for end-to end ABSA	(Chang et al., 2022) (Zhang et al., 2023)

This section frames the methodological analysis in the subsequent sections by formalizing the key principles and task formulations of ABSA.

RESULT

The development of ABSA methodology has changed dramatically within the last 20 years, progressing from lexicon- and grammar-based approaches to conventional machine learning techniques, deep learning architectures (DL), and transformer-based methods to the current massive generative language models. This evolution is the continuation of the progress towards natural language processing and representational learning with its own distinctive features and limitations of different ABSA task models. As visualized in Table 3, Aspect-Based Sentiment Analysis comprises several interrelated subtasks, and this sophistication has facilitated the transition from isolated and line-of-process solutions to unified and comprehensive models (Ahmed et al., 2025) (Hua et al., 2024).

Table 3: Comparative Analysis of Methodological Paradigms in Aspect-Based Sentiment Analysis

Methodologies	Core Principles	Key Strengths	Main Limitations	Supported Tasks	Key References
Lexicon-Based and Rule-Based	Sentiment lexicons; Handcrafted linguistic rules; Domain-specific dictionaries; Syntactic patterns	Interpretable; Computationally efficient; Transparent decision-making; No training data required	Limited lexicon coverage; Poor cross-domain generalization; Cannot handle implicit aspects; Context-dependent sentiment struggles	Coarse-grained ABSA; Domain-constrained applications; Simple aspect–opinion extraction	(Wankhade et al., 2022) (Cui et al., 2023) (Raghunathan & Saravanakumar, 2023).

*name of corresponding author

Traditional Machine Learning	Supervised classification; CRF sequence labeling; Manual feature engineering	Improved flexibility; Better performance than lexicon-based; Structured learning	Extensive feature engineering; Limited contextual modeling; Pipeline error propagation	ATE; Aspect Sentiment Classification; OTE; Pipeline-based ABSA	(Isnan et al., 2023) (Diekson et al., 2023) (Hua et al., 2024).
Deep Learning-Based	CNN, RNN, BiLSTM architectures; Automatic feature learning	Automatic feature extraction; Improved performance; Contextual modeling	Limited long-range dependency modeling; Mostly pipeline-based	ATE; Aspect Sentiment Classification; Sequential ABSA tasks	(Huang et al., 2023) (Gogineni et al., 2023) (Gupta et al., 2024).
Transformer-Based	Pre-trained LMs; Self-attention; Graph-based representations; Ensemble learning	Strong contextual modeling; High performance; Robust joint learning	Computationally expensive; Interpretability issues	All ABSA subtasks; Joint and multi-task learning	(Liang et al., 2022) (Shaik et al., 2022) (Gu et al., 2024) (Diwali et al., 2024) (Apostol et al., 2025).
Generative and Large Language Models	Sequence-to-sequence modeling; Prompt-based learning; Unified generation	Unified task formulation; Implicit aspect handling; Low-resource capability	Hallucination; Controllability; Reproducibility issues	Unified and end-to-end ABSA; Implicit aspect extraction	(Chang et al., 2022) (Florindi et al., 2024) (Ahmad et al., 2025)

Table 3A: Quantitative Synthesis of the 50 Included Studies

Dimension	Count	Percentage	Note
Primary window studies (2022–2025)	42	84%	Core synthesis window
Baseline studies (pre-2022)	8	16%	Retained for benchmark/task grounding
Transformer-PLM dominant studies	19	38%	Encoder/decoder transformer fine-tuning
LLM or generative framing studies	11	22%	Prompting, instruction tuning, or generation
Classical/deep non-transformer studies	12	24%	ML/CNN/RNN/BiLSTM-centric designs
Dataset-focused benchmark studies	8	16%	New datasets or benchmark contributions

ABSA methodologies have transitioned from pipeline-based architectures to unified and generative modeling paradigms as shown in Figure **1**, indicative of growing importance placed on joint task formulation and end-to-end learning.

*name of corresponding author

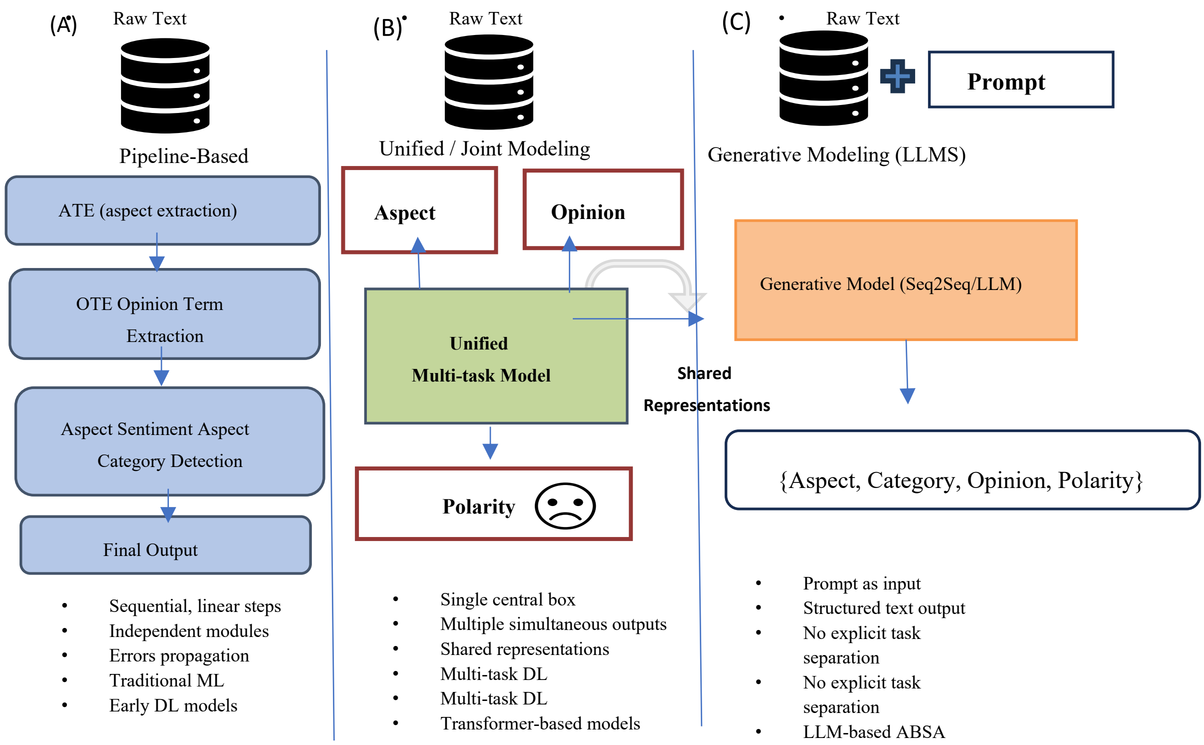

Figure 1: Conceptual comparison of major modeling paradigms in ABSA.

ABSA models yield varying efficacy on the grounds of task complexity, data availability, and modeling objectives. Pipeline-based and early deep learning models are good for isolated subtasks, but they have error propagation. Transformer-based models are well suited for earlier paradigms, as they provide strong contextual representations and are hence effective for joint learning when there is enough annotated data at hand. However, they fail in low-resource, cross-domain, and implicit sentiment situations.

When ABSA is formulated as a structured generation problem, generative and large language models are very strong, especially in extracting entire aspect–opinion–sentiment structures or the use of implicit aspects. Their adaptability and large-scale pretraining make them more capable of handling sparse annotation and multilingual settings, but they have challenges concerning controllability, evaluation consistency, and reproducibility.

This section reviews the methodological paradigms, and it shows a coherent tendency towards unified and generative modeling strategies. The following section highlights benchmark datasets and evaluation processes, revealing how they informed methodological choices by demonstrating ongoing issues like data availability, data imbalance, and domain coverage.

Aspect-Based Sentiment Analysis Datasets

Benchmark datasets have been crucial to the development of Aspect-Based Sentiment Analysis (ABSA) since it was first developed from pipeline approaches to unified and generative models. Early benchmarks, such as SemEval Task 5, defined standards for English restaurant and laptop reviews and focused on Aspect Term Extraction (ATE), Aspect Category Detection (ACD), and Aspect Sentiment Classification (ASC), augmenting
pipeline methods. Unfortunately, these systems do not have explicit structural relations between aspects, opinions, and polarity, resulting in suboptimal end-to-end learning (Hercig et al., 2016) (Toh & Su, 2016) (D'Aniello et al., 2022) (Hua et al., 2024).

To overcome these limitations, new datasets come with multi-aspect and contradictory sentiment settings — e.g., MAMS — that can better reflect the complexity of the real world. Such domain-specific datasets (education, MOOCs, airline reviews, and social media) allow for a more realistic cross-domain

*name of corresponding author

assessment and encourage joint modeling (Jiang et al., 2019) (Shaik et al., 2022). Multilingual and low-resource data sets – particularly in Arabic and Bangla – have extended language diversity and showed challenges related to morphological richness, dialectal diversity, as well as limited pre-trained resources (Obiedat et al., 2021) (Alshaikh et al., 2024) (Eid et al., 2024) (Hasan et al.,2024).

Nonetheless, there are still relatively small non-English datasets that reduce robust evaluation and cross-lingual comparison of large transformer and generative models (Ahmad et al., 2025) (Chen et al., 2024). Table 4 summarizes the common ABSA benchmarks under consideration, the tasks they support, the languages available, and the evaluation metrics.

Table 4: Representative Benchmark Datasets in ABSA.

Dataset Name	Source	Supported Tasks	Language	Domain	Download Link
SemEval-2015/2016	(Pontiki et al., 2016.)	ATE, ACD, ASC	English	Restaurants, Laptops	https://www.kaggle.com/datasets/fouadaurag20/semeval-2016-absa-task5
MAMS	(Jiang et al., 2019)	ATE, ASC	English	General Reviews	https://github.com/siat-nlp/MAMS-for-ABSA
Airline Reviews	(Chang et al., 2022)	ASC, ACOS	English	Airlines	https://www.kaggle.com/datasets/crowdflower/twitter-airline-sentiment
MOOCs Dataset	(Awadh et al., 2025)	ATE, AOPE, ASTE	English	Education	https://www.kaggle.com/search?q=MOOCs+ABSA
Arabic ABSA	(Awadh et al., 2025)	ATE, ASC	Arabic	General Reviews	https://huggingface.co/datasets/eastwind/semeval-2016-absa-reviews-arabic
A-MASA	(Eid et al., 2024)	ATE, ASC	Arabic	Multi-domain	https://github.com/search?q=A-MASA+ABSA
Bangla ABSA	(Hasan et al., 2024)	ATE, ASC	Bangla	General Reviews	https://pmc.ncbi.nlm.nih.gov/articles/PMC11617299/

Evaluation Metrics on Aspect-Based Sentiment Analysis

Benchmark datasets define conditions for evaluation of ABSA, and the changing task formulations have moved the assessment to be structured and made up of explainability related protocols for joint and end-to-end assessment (Florindi et al., 2024). The basic attributes of these metrics are represented by the confusion matrix elements: True Positive (TP) is the number of positive instances correctly classified; True Negative (TN) is the number of negative instances correctly classified as negative; False Positive (FP) is the number of negative instances wrongly classified as positive; False Negative (FN) is the number of positive instances wrongly identified as negative. In light of these core meanings, the evaluation metrics are structured as follows:

Precision: the proportion of predicted positives which are also correct.

$$Precision = TP / (TP + FP) \qquad (1)$$

Recall (Sensitivity): percentage of the real positives correctly named.

$$Recall = TP / (TP + FN) \qquad (2)$$

F1: Harmonic mean of Precision and Recall that balances false positive and false negative, robust under class imbalance.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15888
e-ISSN : 2541-2019
p-ISSN : 2541-044X

$$F1 = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (3)$$

Accuracy: the proportion of all the predictions that turned out to be true.

$$Accuracy = (TP + TN) / (TP + FP + FN + TN) \quad (4)$$

Macro-F1: the unweighted mean of F1 across classes (N classes).

$$Macro - F1 = (\Sigma_{i=1}{}^{N} F1_i) / N \quad (5)$$

(Where N is the number of classes and the corresponding value of the F1-score for class i.)
Exact Match (EM): only if elements (Aspect, Opinion, Category, Sentiment) match the gold annotation does a structured prediction return the correct value (Zhang et al., 2023).

$$Exact\ Match\ (EM): All\ elements\ must\ match \quad (6)$$

Tuple/Quadruple-Level F1: Precision/Recall calculated as the whole of structured outputs and accurate structurally strict (Zhang et al., 2024).

$$Tuple/Quadruple - Level\ F1: Precision/Recall\ on\ complete\ structures \quad (7)$$

Figure 2 presents an integrated study of evaluation metrics in 45 academic articles on Aspect-Based Sentiment Analysis (ABSA), showcasing divergent trends in metric use that mimic ABSA task formulations and measurement trends. The histogram shows that F1-score emerges as the most widely adopted metric.

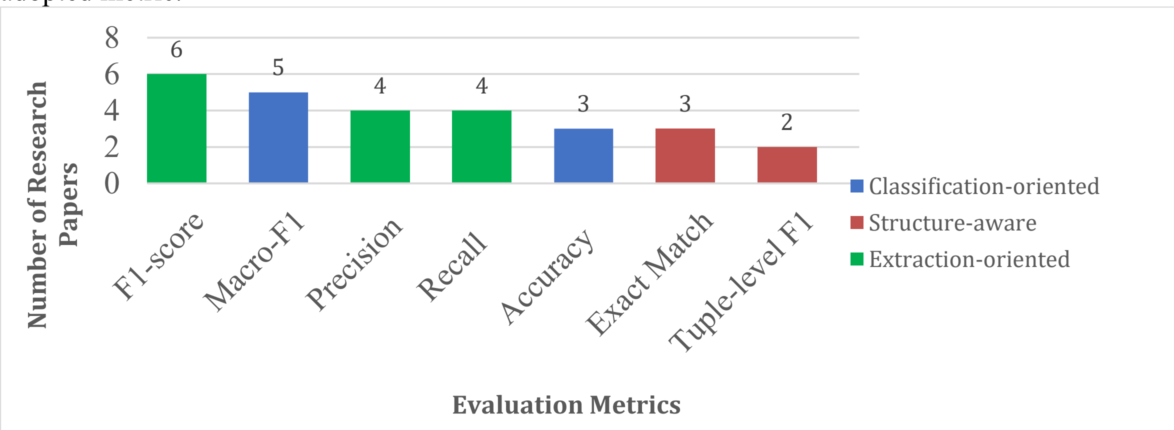

Figure 2: Distribution of evaluation metrics usage in aspect-based sentiment analysis studies.

Model performance in ABSA is highly sensitive to dataset architecture and evaluation parameter selection. This gives us the motivation for the discussion presented in the next part of our paper, focusing on adaptation of modern architectural and learning strategies of the design and approach of ABSA models in order to capture data complexity, multilingually, and evaluation limitations of a modern approach to complex datasets.

RESULTS

The review aggregated fifty studies. As summarized in Table 1, methodological approaches comprise 40% of the papers, benchmark dataset contributions constitute 16%, comprehensive surveys contribute 14%, explainability and interpretability 10%, multilingual and low-resource work 8%, domain-specific applications 10%, and theoretical foundations 2%, respectively. As can be concluded this distribution suggests that modelling is prioritized and there are comparatively fewer studies on explainability and low-resource projects. It is clearly illustrated in Figure 3 that the F1-score is the one most frequently used evaluation metric in recent works, showing that authors prefer a balanced performance reporting when processing class-imbalanced data.

DISCUSSIONS

Aspect-Based Sentiment Analysis (ABSA) has also adapted to transformer designs. Nevertheless, a number of fundamental issues concerning linguistic incomprehension and the variability presented by raw data remain open-ended. This section will describe four of the common problems that arise concerning robustness and generalizability.

One significant challenge to recovering implicit aspects, implicit opinions, and implicit aspect–opinion relations is when sentiment is inextricably context-dependent and is better understood through

*name of corresponding author

common sense than through explicit lexical information (Zhang et al., 2024) (Aziz et al., 2024) (Ahmed et al., 2025).

Domain adaptation and cross-lingual transfer efforts are still weak: Sentiment expressions are domain-specific (e.g., "light" changes depending on context) and cross-linguistic/cross-cultural mismatches lead to transfer stability problems significantly, more severely in English-dominant datasets (Hua et al., 2024) (Ahmad et al., 2025).

The linguistic phenomena of sarcasm, negation, ambiguity, and mixed sentiment still need to be better modeled with a more extensive semantic representation, although these methodologies are not yet sufficient to systematically model these phenomena (Aziz et al., 2024).

Trainings and tests are restricted by imbalance in datasets and lack of data in specialized domains, which leads to biased models and lower minority-class performances (Hua et al., 2024). All these barriers combine to motivate work on knowledge integration, domain adaptation, language understanding, and dataset expansion with further development as information reveals itself.

Future Directions

Recent works in aspect sensitive sentiment analysis focus on three fields. Some applications of explainable artificial intelligence exist in high-stakes fields, including LIME (Raghunathan & Saravanakumar, 2023), model-agnostic explanation surveys (Diwali et al., 2024), and attention visualization (Perikos & Diamantopoulos, 2024) whose goal is to enhance transparency and bias identification; however, the standardization and integration of artificial intelligence technologies without sacrificing performance remain open issues.

Multi-lingual models improve on few-shot learning, unified task formulations, more efficient implicit treatment and multilingual transfer (Mughal et al., 2024; Ahmed et al., 2025; Perikos & Diamantopoulos, 2024; Zhang & Shafiq, 2024), but the controllability, consistency of evaluation, hallucination, computational cost and reproducibility, among others, are still questions.

Conversational and multimedia aspect-based sentiment analysis extend analysis to conversational context and multimodal signals to enhance more robust and realistic opinion modeling (Zhang et al., 2024) (Jim et al., 2024). Together those directions point to much more interpretable, adaptable, efficient systems.

CONCLUSION

This review presents a task-aware, PRISMA-aligned synthesis of 50 ABSA studies, with a clear separation between the main 2022–2025 evidence window and baseline reference data prior to 2022. These model trends are directly supported within the quantitative synthesis (Table 3A) which shows a peak in transformer PLM studies (38%) while continuing growth in LLM/generative adoption (22%) and ongoing gaps in the consistency and explainability of the evaluations. Hence claims of methodological robustness derive from decisions made regarding traceable screening, standardized terminology, and explicit evidence aggregation rather than narrative impression.

By connecting subtasks to the modeling paradigms, datasets, and evaluation practices, this review provides a task-aware synthesis of aspect-based sentiment analysis, and highlights lingering design limitations of implicit aspect extraction, cross-domain generalization, data imbalance, and interpretability. The synthesis explains how methodological selections correspond to given subtasks and creates a unified basis for benchmarks and measures selection. Recommendations for the future should be: 1) make modeling more explainable and trustworthy in high-stakes domains, 2) adopt a robust cross-lingual and cross-domain transfer with limited supervision, or 3) enforce evaluation protocols consistent with end-to-end task structure. These directions are targeted to enhance reliability, transparency and generalizability during real world deployments.

REFERENCES

Ahmad, W., Khan, H. U., Alarfaj, F. K., & Alreshoodi, M. (2025). Aspect-Based Sentiment Analysis: A Comprehensive Review and Open Research Challenges. *IEEE Access*, *13*, 65138–65182. https://doi.org/10.1109/ACCESS.2025.3555744

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 2, April 2026
DOI : https://doi.org/10.33395/sinkron.v10i2.15888
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Ahmed, K., Nadeem, M. I., Wang, G., Zuo, F., & Han, Z. (2025). Instruction-tuned ABSA with auxiliary sentences and knowledge-enhanced graphs for implicit aspect detection. *Expert Systems with Applications*, *289*, 128284. https://doi.org/10.1016/j.eswa.2025.128284

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

Apostol, E.-S., Pisică, A.-G., & Truică, C.-O. (2025). ATESA-BÆRT: A heterogeneous ensemble learning model for Aspect-Based Sentiment Analysis. *Knowledge-Based Systems*, *326*, 113987. https://doi.org/10.1016/j.knosys.2025.113987

Awadh, W. A., Sulaiman, R. B., & Mahmoud, M. A. (2025). Aspect-based sentiment analysis in MOOCs: A systematic literature review introducing the MASC-MEF framework. *Journal of King Saud University Computer and Information Sciences*, *37*(1–2), 2. https://doi.org/10.1007/s44443-025-00018-1

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

Chang, Y.-C., Ku, C.-H., & Nguyen, D.-D. L. (2022). Predicting aspect-based sentiment using deep learning and information visualization: The impact of COVID-19 on the airline industry. *Information & Management*, *59*(2), 103587. https://doi.org/10.1016/j.im.2021.103587

Chebolu, S. U. S., Dernoncourt, F., Lipka, N., & Solorio, T. (2023). A Review of Datasets for Aspect-based Sentiment Analysis. *Proceedings of the 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (Volume 1: Long Papers)*, 611–628. https://doi.org/10.18653/v1/2023.ijcnlp-main.41

Chen, X., Xie, H., Qin, S. J., Chai, Y., Tao, X., & Wang, F. L. (2024). Cognitive-Inspired Deep Learning Models for Aspect-Based Sentiment Analysis: A Retrospective Overview and Bibliometric Analysis. *Cognitive Computation*, *16*(6), 3518–3556. https://doi.org/10.1007/s12559-024-10331-y

Chen, X., Xie, H., Tao, X., Wang, F. L., Zhang, D., & Dai, H.-N. (2025). A computational analysis of aspect-based sentiment analysis research through bibliometric mapping and topic modeling. *Journal of Big Data*, *12*(1), 40. https://doi.org/10.1186/s40537-025-01068-y

Cui, J., Wang, Z., Ho, S.-B., & Cambria, E. (2023). Survey on sentiment analysis: Evolution of research methods and topics. *Artificial Intelligence Review*, *56*(8), 8469–8510. https://doi.org/10.1007/s10462-022-10386-z

D'Aniello, G., Gaeta, M., & La Rocca, I. (2022). KnowMIS-ABSA: An overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis. *Artificial Intelligence Review*, *55*(7), 5543–5574. https://doi.org/10.1007/s10462-021-10134-9

Diekson, Z. A., Prakoso, M. R. B., Putra, M. S. Q., Syaputra, M. S. A. F., Achmad, S., & Sutoyo, R. (2023). Sentiment analysis for customer review: Case study of Traveloka. *Procedia Computer Science*, *216*, 682–690. https://doi.org/10.1016/j.procs.2022.12.184

Diwali, A., Saeedi, K., Dashtipour, K., Gogate, M., Cambria, E., & Hussain, A. (2024). Sentiment Analysis Meets Explainable Artificial Intelligence: A Survey on Explainable Sentiment Analysis. *IEEE Transactions on Affective Computing*, *15*(3), 837–846. https://doi.org/10.1109/TAFFC.2023.3296373

D'Orazio, M., Di Giuseppe, E., Mariotti, C., & Muccioli, M. F. (2025). Leveraging user-generated content to enhance heritage monitoring protocols and resources allocation through aspect-based sentiment analysis. *Journal of Information Technology in Construction*, *29*, 397–417. https://doi.org/10.36680/j.itcon.2025.017

Eid, Y., Zayed, H., & Medhat, W. (2024). A-MASA: Arabic Multi-Domain Aspect-Based Sentiment Analysis Datasets. *Procedia Computer Science*, *244*, 202–211. https://doi.org/10.1016/j.procs.2024.10.193

*name of corresponding author

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

Florindi, F., Fedele, P., & Dimitri, G. M. (2024). A novel solution for the development of a sentimental analysis chatbot integrating ChatGPT. *Personal and Ubiquitous Computing*, *28*(6), 947–960. https://doi.org/10.1007/s00779-024-01824-6

Gogineni, A. K., Reddy, S. K. S., Kakarala, H., Gavini, Y. C., Venkat, M. P., Hajarathaiah, K., & Enduri, M. K. (2023). A Hybrid Deep Learning Framework for Efficient Sentiment Analysis. *International Journal of Advanced Computer Science and Applications*, *14*(12). https://doi.org/10.14569/IJACSA.2023.01412105

Gu, Q., Wang, Z., Zhang, H., Sui, S., & Wang, R. (2024). Aspect-Level Sentiment Analysis Based on Syntax-Aware and Graph Convolutional Networks. *Applied Sciences*, *14*(2), 729. https://doi.org/10.3390/app14020729

Gupta, S., Singhal, N., Hundekari, S., Upreti, K., Gautam, A., Kumar, P., & Verma, R. (2024). Aspect Based Feature Extraction in Sentiment Analysis using Bi-GRU-LSTM Model. *Journal of Mobile Multimedia*, 935–960. https://doi.org/10.13052/jmm1550-4646.2048

Hasan, M., Ghani, Md. R., & Hasan, K. M. A. (2024). Aspect based sentiment analysis datasets for Bangla text. *Data in Brief*, *57*, 111107. https://doi.org/10.1016/j.dib.2024.111107

Hercig, T., Brychcín, T., Svoboda, L., & Konkol, M. (2016). UWB at SemEval-2016 Task 5: Aspect Based Sentiment Analysis. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 342–349. https://doi.org/10.18653/v1/S16-1055

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

Huang, J., Cui, Y., & Wang, S. (2023). Adaptive Local Context and Syntactic Feature Modeling for Aspect-Based Sentiment Analysis. *Applied Sciences*, *13*(1), 603. https://doi.org/10.3390/app13010603

Islam, Md. S., Kabir, M. N., Ghani, N. A., Zamli, K. Z., Zulkifli, N. S. A., Rahman, Md. M., & Moni, M. A. (2024). Challenges and future in deep learning for sentiment analysis: A comprehensive review and a proposed novel hybrid approach. *Artificial Intelligence Review*, *57*(3), 62. https://doi.org/10.1007/s10462-023-10651-9

Isnan, M., Elwirehardja, G. N., & Pardamean, B. (2023). Sentiment Analysis for TikTok Review Using VADER Sentiment and SVM Model. *Procedia Computer Science*, *227*, 168–175. https://doi.org/10.1016/j.procs.2023.10.514

Jiang, Q., Chen, L., Xu, R., Ao, X., & Yang, M. (2019). A Challenge Dataset and Effective Models for Aspect-Based Sentiment Analysis. *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 6279–6284. https://doi.org/10.18653/v1/D19-1654

Jim, J. R., Talukder, M. A. R., Malakar, P., Kabir, M. M., Nur, K., & Mridha, M. F. (2024). Recent advancements and challenges of NLP-based sentiment analysis: A state-of-the-art review. *Natural Language Processing Journal*, *6*, 100059. https://doi.org/10.1016/j.nlp.2024.100059

Liang, B., Su, H., Gui, L., Cambria, E., & Xu, R. (2022). Aspect-based sentiment analysis via affective knowledge enhanced graph convolutional networks. *Knowledge-Based Systems*, *235*, 107643. https://doi.org/10.1016/j.knosys.2021.107643

Mughal, N., Mujtaba, G., Shaikh, S., Kumar, A., & Daudpota, S. M. (2024). Comparative Analysis of Deep Natural Networks and Large Language Models for Aspect-Based Sentiment Analysis. *IEEE Access*, *12*, 60943–60959. https://doi.org/10.1109/ACCESS.2024.3386969

Obiedat, R., Al-Darras, D., Alzaghoul, E., & Harfoushi, O. (2021). Arabic Aspect-Based Sentiment Analysis: A Systematic Literature Review. *IEEE Access*, *9*, 152628–152645. https://doi.org/10.1109/ACCESS.2021.3127140

Perikos, I., & Diamantopoulos, A. (2024). Explainable Aspect-Based Sentiment Analysis Using Transformer Models. *Big Data and Cognitive Computing*, *8*(11), 141. https://doi.org/10.3390/bdcc8110141

Pontiki, M., Galanis, D., Papageorgiou, H., Androutsopoulos, I., Manandhar, S., Al-Smadi, M., ... & Eryiğit, G. (2016, June). Semeval-2016 task 5: Aspect based sentiment analysis. In *Proceedings of the 10th international workshop on semantic evaluation (SemEval-2016)* (pp. 19-30).

*name of corresponding author

Raghunathan, N., & Saravanakumar, K. (2023). Challenges and Issues in Sentiment Analysis: A Comprehensive Survey. *IEEE Access*, *11*, 69626–69642. https://doi.org/10.1109/ACCESS.2023.3293041

Shaik, T., Tao, X., Li, Y., Dann, C., McDonald, J., Redmond, P., & Galligan, L. (2022). A Review of the Trends and Challenges in Adopting Natural Language Processing Methods for Education Feedback Analysis. *IEEE Access*, *10*, 56720–56739. https://doi.org/10.1109/ACCESS.2022.3177752

Toh, Z., & Su, J. (2016). NLANGP at SemEval-2016 Task 5: Improving Aspect Based Sentiment Analysis using Neural Network Features. *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016)*, 282–288. https://doi.org/10.18653/v1/S16-1045

Toms, W. (n.d.). *Explainable Aspect-Based Sentiment Analysis Using Attention Mechanisms and SHAP*.

Wankhade, M., Rao, A. C. S., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, *55*(7), 5731–5780. https://doi.org/10.1007/s10462-022-10144-1

Ye, L., Johar, G., & Alkawaz, M. H. (n.d.). *Aspect-Based Sentiment Analysis Based on Knowledge Enhancement and Pre-trained Language Model*.

Zhang, H., Cheah, Y.-N., Alyasiri, O. M., & An, J. (2024). Exploring aspect-based sentiment quadruple extraction with implicit aspects, opinions, and ChatGPT: A comprehensive survey. *Artificial Intelligence Review*, *57*(2), 17. https://doi.org/10.1007/s10462-023-10633-x

Zhang, H., & Shafiq, M. O. (2024). Survey of transformers and towards ensemble learning using transformers for natural language processing. *Journal of Big Data*, *11*(1), 25. https://doi.org/10.1186/s40537-023-00842-0

Zhang, Y., Yang, Y., Liang, B., Chen, S., Qin, B., & Xu, R. (2023). An Empirical Study of Sentiment-Enhanced Pre-Training for Aspect-Based Sentiment Analysis. *Findings of the Association for Computational Linguistics: ACL 2023*, 9633–9651. https://doi.org/10.18653/v1/2023.findings-acl.612

*name of corresponding author