Aspect-based sentiment-analysis using topic modelling and machine-learning

Radhika Jinendra Dhanal[1,2], Vijay Ram Ghorpade[3]
[1]Department of Computer Science and Engineering, D. Y. Patil College of Engineering and Technology, Shivaji University, Kolhapur, India
[2]Department of Technology, Shivaji University, Kolhapur, India
[3]Department of Computer Science and Engineering, Bharati Vidyapeeth's College of Engineering, Shivaji University, Kolhapur, India

Article Info

Article history:

Received Apr 16, 2024
Revised Aug 6, 2024
Accepted Aug 14, 2024

Keywords:

BERT
Latent Dirichlet allocation
Machine learning
SemEval
Sentiment
TF-BERT

ABSTRACT

This study addresses the critical need for an accurate aspect-based sentiment-analysis (ABSA) model to understand sentiments effectively. The existing ABSA models often face challenges in accurately extracting aspects and determining sentiment polarity from textual data. Therefore, we propose a novel approach leveraging latent-Dirichlet-allocation (LDA) for aspect extraction and transformer-based bidirectional-encoder-representations from transformers (TF-BERT) for sentiment-polarity evaluation. The experiments were carried out on SemEval 2014 laptop and restaurant datasets. Also, a multi-domain dataset was generated by combining SemEval 2014, Amazon, and hospital reviews. The results demonstrate the superiority of the LDA-TF-BERT model, achieving 82.19% accuracy and 79.52% Macro-F1 score for the laptop task and 86.26% accuracy of 87.26% and 81.27% for Macro-F1 score for the restaurant task. This showcases the model's robustness and effectiveness in accurately analyzing textual data and extracting meaningful insights. The novelty of our work lies in combining LDA and TF-BERT, providing a comprehensive and accurate ABSA solution for various industries, thereby contributing significantly to the advancement of sentiment analysis techniques.

Corresponding Author:

Radhika Jinendra Dhanal
Department of Computer Science and Engineering, D. Y. Patil College of Engineering and Technology, Shivaji University
Kolhapur, India
Email: dhanallraddheka@gmail.com

1. INTRODUCTION

One subfield of natural-language-processing (NLP) known as aspect-based sentiment-analysis (ABSA) looks for patterns in how people feel about certain parts of textual content, including customer-feedbacks, social-media posts, or product-reviews [1]. ABSA differs from conventional sentiment assessment methods by focusing on identifying and evaluating the polarity of sentiment (neutral, negative, or positive) towards specific aspects along with entities discussed in a given text. Unlike conventional methods which offer a general sentiment rating to the entire text, ABSA attempts to offer a more granular evaluation by examining sentiments regarding individual elements within the text [2]. For instance, in a restaurant-review, ABSA might analyze sentiments towards aspects like food-quality, ambiance, service, and pricing separately [3]. ABSA finds applications across various industries where understanding customer opinions and sentiments towards specific features or products is crucial. In the retail and e-commerce sector, ABSA helps businesses analyze customer feedback on product features, pricing, shipping, and customer service to improve customer experience and make data-driven decisions [4]. In the hospitality industry, ABSA is used to analyze sentiments toward hotel amenities, room cleanliness, staff behavior, and dining options to enhance guest satisfaction [5]. It is also valuable in sectors like healthcare, automotive, finance, and more, where customer feedback and sentiment analysis play a vital role in business strategies and product development [6].

The need for ABSA arises from the limitations of traditional sentiment analysis [7], [8] which provides a generalized sentiment score without considering the importance related to specific aspects. ABSA enables a more granular understanding of customer opinions, allowing businesses to precisely identify strengths, weaknesses, and areas for improvement. However, traditional deep-learning (DL) [9], [10] and machine-learning (ML) [11], [12] models used in ABSA have certain disadvantages. Moreover, DL models in ABSA are often trained on specific datasets and domains, making them less adaptable to new contexts or industries [13]. They may also suffer from bias and lack of generalizability, as they focus on the given area of training and may not perform well on diverse datasets or emerging trends [14]. Additionally, many existing models in ABSA directly evaluate the overall polarity of reviews without considering aspect-level sentiments, leading to unreliable and biased results [15].

To address these challenges, this work presents a novel approach, i.e., latent-Dirichlet-allocation (LDA) for aspect extraction, which can identify topics or aspects within a corpus of reviews. LDA helps in identifying related words to these topics, which are considered as aspects for sentiment analysis. Additionally, advanced models like Transformer-based bidirectional-encoder-representations from transformers (TF-BERT) are used in this work for evaluating sentiment polarity, both at the overall review level and aspect level. By leveraging LDA for aspect extraction and TF-BERT for sentiment analysis, the proposed approach combines aspect-level sentiment analysis with overall sentiment evaluation, providing a more comprehensive understanding of customer sentiments. This approach aims to overcome the limitations of traditional ABSA models by considering diverse datasets, extracting aspects dynamically, and evaluating aspect-level sentiments to reduce bias and improve the reliability of sentiment analysis results in various industries. The contributions of this work can be summarized as follows:

– Introduced a novel approach termed LDA-TF-BERT, which combines LDA for aspect extraction and TF-BERT for sentiment analysis and polarity evaluation.
– The model's performance was rigorously evaluated utilizing the standard SemEval-2014 Task-4 dataset, a widely recognized benchmark for sentiment analysis tasks.
– In addition to the SemEval-2014 Task-4 dataset, this work has also been evaluated using diverse reviews from multiple domains, including hospital and Amazon reviews.
– The results obtained from the LDA-TF-BERT model showcase better performance in comparison with existing sentiment analysis models.

This work is structured as follows. Section 2 delves into the literature survey, providing an overview of existing research and methodologies in sentiment-analysis and ABSA. In section 3, the LDA-TF-BERT model is thoroughly discussed, detailing its architecture, methodology, and implementation for sentiment analysis and polarity evaluation. Following that, section 4 focuses on evaluating the LDA-TF-BERT model in terms of accuracy and macro-F-score, presenting a comparative analysis with current works for highlighting LDA-TF-BERT performance and advancements. Finally, section 5 encapsulates the conclusion of the work, summarizing key findings, contributions, and potential avenues for future research in the area of sentiment-analysis and ABSA.

2. LITERATURE SURVEY

The literature survey delves into various methodologies and findings from recent studies in ABSA and sentiment-analysis. This section covers a range of approaches aimed at predicting sentiment polarities and extracting aspect-sentiment pairs within sentences or documents. In [16], the objective of this study was to develop a predictive model for sentiment-polarities associated with various aspects within documents or sentences. This research was motivated by the observation that previous research had overlooked the crucial factors of aspect sentiment-polarity and local-context relationship. To overcome this issue, this work proposed a local-context-focus (LCF) approach from the multi-head self-attention (MHSA) module previously presented for ABSA. This work utilized two layers called context-features dynamic-weight (CDW) and context-features dynamic-mask (CDM) to prioritize local-context words. They achieved 75.08% and 74.45% accuracy with CDM and CDW, respectively, for the laptop dataset, and 80.98% and 80.89% for the restaurant dataset. The Twitter dataset yielded an accuracy of 71.24%. Further, Phan and Ogunbona [17], explored sentence grammatical-aspects and used a self-attention approach for syntactical-learning and proposed a model called LCF on syntax aspect-based-classification (LCFS-ASC). They combined contextualized embeddings (i.e., RoBERTa and BERT), dependency-based, and part-of-speech (PoS) embeddings for aspect extraction and enhancing performance. The LCFS-ASC achieved an accuracy of 80.52 with CDW and 80.34 with CDM for the laptop dataset. For the restaurant dataset, they achieved an accuracy of 80.31 with CDW and 80.10 with CDM.

The study in [18] was centered around the extraction of aspect-sentiment polarity-pairs. The approach involved utilizing sentence information, edge-labels, and dependency-trees to achieve this goal. For this, they proposed a model called phrase-dependency graph-attention-network (PD-RGAT) and conducted experiments with Global-Vectors (GloVE) and BERT pre-training models. PD-RGAT showed better effectiveness comparable to current works on the SemEval-2014 Task-4 dataset. He *et al.* [19] main focus was to understand the relationship between aspect sentiment-polarity and global-context. Hence, proposed a multilingual learning model using CDM and CDW called local and global-context-focus CDM-CDW (LGCF-CDM-CDW). The LGCF-CDM-CDW achieved 81.29% accuracy for the laptop dataset and 85.52% accuracy for the restaurant dataset. Yu and Zhang [20] focus was to detect aspect-based relevant words utilizing an attention-based approach. This work introduced an aspect-position information extraction approach for long-distance dependencies (sentences) and proposed the model called multi-weight graph-convolutional-network (MWGCN) and achieved accuracies of 78.8% and 83.82% for the laptop and restaurant datasets, respectively. Ma *et al.* [21] aimed to correct the errors faced by dependency-parsers and to understand the complex sentence structures, hence, presented a Multi-GCN model. The Multi-GCN model achieved accuracies of 798.8% for the laptop and 83.82% for the restaurant dataset. Lin and Joe [22] proposed a masked attention mechanism for noise reduction and feature enhancement in the global text called CDM-CDW. The CDM-CDW achieved an accuracy of 80% for the laptop and 80.59% for the restaurant datasets. Finally, Zhao *et al.* [23] considered relationships among the dependency-edges and nodes for aspect sentiment-polarity detection and proposed a model called structured-dependency tree-based GCN (SDTGCN) approach. The SDTGCN achieved accuracies of 78.64% for the laptop and 83.32% for the restaurant dataset.

While the above studies provide valuable insights and advancements in ABSA, it is essential to acknowledge certain limitations inherent in these methodologies and findings. Some models may perform well in specific domains like laptops or restaurants but may struggle when applied to different domains or industries. Adapting these models to diverse domains remains a challenge. The scalability of models, especially those incorporating attention mechanisms or graph-based approaches, may be limited when dealing with large-scale datasets or real-time applications. Moreover, the advanced models discussed above require substantial computational resources and training time, making them less accessible for researchers or organizations with limited computing capabilities. Also, like many artificial intelligent (AI) systems, aspect-based sentiment classifiers may exhibit biases based on the training data, potentially leading to unfair or biased predictions, especially concerning sensitive topics or underrepresented groups. Given the rapid evolution of language use and emerging trends, models developed based on older datasets or linguistic patterns may struggle to adapt to new linguistics or expressions. Hence, to solve all these problems, LDA-TF-BERT has been proposed. The LDA approach helps to recognize topics along with related words which helps to extract aspects efficiently, whereas the TF-BERT helps to detect the polarity of aspects and sentences efficiently. As many of the models discussed above rely heavily on the SemEval-2014 dataset for evaluation, hence the proposed LDA-TF-BERT has also been evaluated utilizing the SemEval-2014 Task-4 dataset. The LDA-TF-BERT has been discussed in more depth in the below section.

3. MODEL

In Figure 1, the LDA-TF-BERT model architecture is presented. The initial step of the LDA-TF-BERT model involves extracting review text from the corpus dataset. This text is then subjected to aspect extraction and polarity extraction using dedicated models, i.e., LDA and TF-BERT respectively. The aspect extraction model is responsible for identifying aspects within the text, while the polarity extraction model determines the polarities of aspects and sentences. Ultimately, by combining the extracted aspects, sentence polarities, and aspect polarities, the overall sentiment polarity of the sentence is determined.

3.1. LDA model for aspect extraction

LDA has emerged as a powerful method for uncovering hidden patterns and structures within review datasets (denoted as R). It leverages the capability to analyze data and generate topic-vectors by modeling likelihood distributions. In LDA, as depicted in Figure 2, the variables are defined as follows: n represents the position of a word within a specific review R, α, and η are parameters governing topics and proportions, respectively, while β evaluates the mixture of words across different topics. K signifies overall topics, θ_R denotes the distribution of topics for every review, $Z_{R,n}$ denotes the matrix encapsulating reviews and words, and $W_{R,n}$ represents the likelihood distribution of words. Further, R denotes the overall reviews, and N denotes the overall words present in reviews. In the LDA, it is considered that the prior-distribution of review topics follows a Dirichlet-distribution [24], which is a fundamental aspect of the model's underlying probabilistic framework.

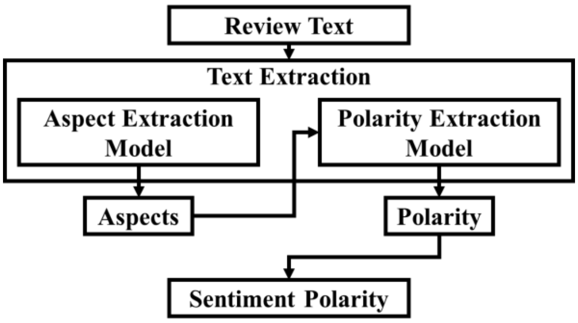

Figure 1. Architecture of LDA-TF-BERT model

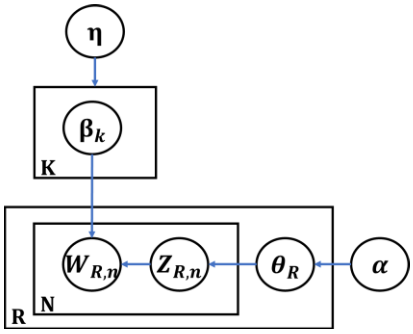

Figure 2. LDA model

In LDA, the topic distribution for a specific review R is represented as θ_R = $Dirichlet(\vec{\alpha})$, where $\vec{\alpha}$ is a hyper-parameter vector with K dimensions, corresponding to the number of topics. The assignment of a topic to the n^{th} word in review R, denoted as $z_{R,n}$, is calculated using the multinomial distribution as $z_{R,n} = multi(\theta_R)$. It is to be noted that LDA considers the prior distribution of words within every topic to be governed by Dirichlet-distribution as discussed above. Specifically, for any topic K, its word-distribution is expressed as β_k=$Dirichlet(\vec{\eta})$, where $\vec{\eta}$ represents another hyper-parameter vector with vocabulary size dimensions. Furthermore, the observed word probability distribution for each word $\mathcal{W}_{Rn}$ in review R is determined as $\mathcal{W}_{Rn} = multi(\beta_{zRn})$, where β_{zRn} refers to the word distribution associated with the topic assignment $z_{R,n}$ for the given word. In (1), the comprehensive distribution encompassing all visible and hidden variables within the LDA model can be efficiently approximated through Gibbs sampling, a technique commonly used in Bayesian inference and probabilistic modeling [25]. This process aids in inferring the latent topics and word distributions within a corpus of documents, enabling the extraction of meaningful insights and patterns.

$$p(\mathcal{W}_R, z_R, \theta_R, \beta_k|\alpha, \beta) = \prod_{n=1}^{N} p(\theta_R|\,\alpha)p(z_{Rn}|\theta_R)p(\beta_k|\beta)p(\mathcal{W}_{Rn}|\beta_k) \tag{1}$$

The coherence score for each topic across different clusters is determined by evaluating the (2) [26].

$$\phi S_i(\vec{u}, \vec{w}) = \frac{\sum_{i=1}^{|W|} u_i \cdot w_i}{||\vec{u}||_2 \cdot ||\vec{w}||_2} \tag{2}$$

ϕ signifies the confirmation measure, while S_i represents the pair of vectors u and w. Here, $\vec{u}$ and $\vec{w}$ denote the cosine vector similarity. Following the assessment of the coherence score using (2), the topics with the highest likelihood are chosen. This selection process is facilitated by utilizing the pyLDAvis library, which visualizes the distinct topics identified. From the different topics, its correlated words are considered as aspects.

3.2. TF-BERT model for polarity classification

In this section, a hybrid joint-extraction approach using the TF-BERT model is presented for extracting both the polarity of aspects and reviews. In this work, the BERT has been pre-trained and fine-tuned utilizing a transformer taken from the Hugging-face [27]. Consider the dataset X which represents a set of reviews, where x are the different sentences and N represents the tokens (words). This can be represented as (3). From X, the polarity is identified using a tuple set T, as presented in (4).

$$X = \{x^1, x^2, x^3, \dots, x^N\} \tag{3}$$

$$T = \{(a_i, o_i, s_i)\}\,|T| \tag{4}$$

where a_i represents the aspects, o_i represents the polarity of aspects, s_i represents the polarity of reviews, $i = 1, \dots, n$ and $|T|$ represents how many sets are present for the given corpus dataset. Moreover, the hybrid joint-extraction approach utilizes a position tagging technique for defining aspects which are present from the starting point to the ending point of a sentence represented as A_s and A_e. Similarly, the position tagging technique is also used for defining polarity present from the starting point to the ending point of a sentence represented as O_s and O_e. Also, as the review texts are embedded, this work utilizes a position-encoding process to get the position of aspects and their associated polarity. Using the BERT, the tokens are further contextualized and represented as (5).

$$CR_e = \{r_{e1}, r_{e2}, \dots, r_{el}\} \tag{5}$$

where, CR_e represents a set of overall tokens and r_{el} represents individual tokens. Further, for the classification of tokens, this work utilizes a fine-tuning process. During fine-tuning, BERT uses the output from its final layer (classify token) as input to a SoftMax classifier. The SoftMax function is applied to the logits (raw scores) produced by the classifier, which converts these scores into probabilities. These probabilities indicate the likelihood of each class (e.g., positive, negative, neutral) for sentiment analysis, and the highest probability class is selected as the predicted sentiment label. The probability is defined as (6).

$$Y = Softmax(CR_e W_e + b_e) \tag{6}$$

where, W_e is the weight matrix that is used to linearly transform the contextualized tokens into a format that is suitable for classification and b_e is a bias vector that is added to the linear transformation output to shift the result and enhance the flexibility of the BERT. After the process of classification of tokens, the outcomes are separated as sentence polarity and aspect polarity represented as $Y_O = \{o_1, o_2, \dots, o_m\}$, $Y_A = \{a_1, a_2, \dots., a_m\}$. The complete flow is presented in Figure 3. From sentence polarity $Y_O = \{o_1, o_2, \dots, o_m\}$ and aspect polarity $Y_A = \{a_1, a_2, \dots, a_m\}$, the overall sentiment polarity is evaluated using accuracy and macro F-score as (7)-(9).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{7}$$

$$F-Score = \frac{2\times(\frac{TP}{TP+FP})\times(\frac{TP}{TP+FN})}{(\frac{TP}{TP+FP})+(\frac{TP}{TP+FN})} \tag{8}$$

$$Macro\ F-Score = Average\ of\ F-score \tag{9}$$

where, TP represents $True\ Positive$, TN represents $True\ Negative$, FP represents $False\ Postive$ and FN represent $False\ Negative$. The results for the LDA-TF-BERT model are evaluated using (7) to (9). Further, in the next section, the complete results achieved by the LDA-TF-BERT are discussed in detail.

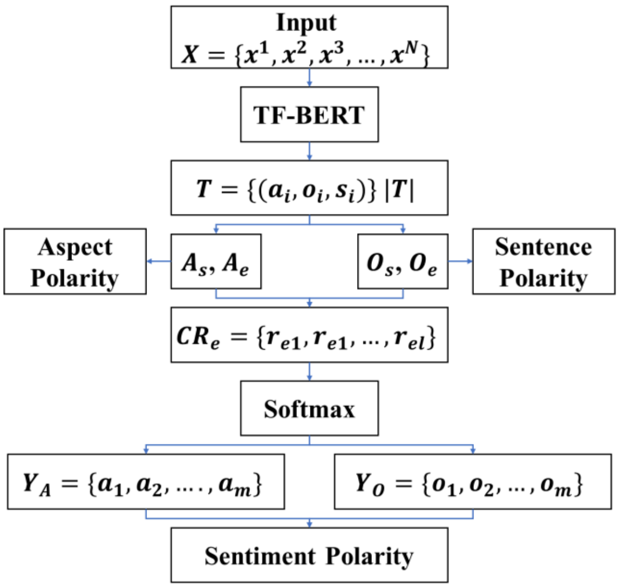

Figure 3. TF-BERT model

4. RESULTS AND DISCUSSION
4.1. System requirements
The evaluation of the LDA-TF-BERT model was performed utilizing an Intel i7 processor with 12 cores and 16 GB of RAM, operating on Windows 11. The coding for the model was implemented in Python, and the execution was carried out within the Anaconda environment. The parameter settings used in this work are presented in Table 1.

Table 1. Parameter settings

Parameter	Value
Dropout-rate	0.1
Optimizer	Adam
Learning rate	2.00E-05
Batch size	32
Max sequence length	128
Max epoch	10

4.2. Dataset
In this work, hospital datasets were collected from [28], [29]. Also, this work considered the SemEval 2014 Task 4 [30] for evaluation. The laptop 2014 and restaurant 2014 are two subtasks within the SemEval sentiment analysis dataset. These datasets are labeled as having different aspects and polarity. Polarity includes sentiments, i.e., neutral, negative, or positive associated with these aspects. Further, a multi-source dataset has been generated for this work. The dataset consists of information collected from diverse sources, such as Hospitals [28], [29], SemEval 2014 Task 4 [30], and Amazon [31]. For the hospital dataset, data extraction was conducted from Practo and Mouthshut using a web-scrapping application programming interface (API). The SemEval 2014 Task 4 dataset included exclusively laptop and restaurant data. Additionally, the Amazon dataset comprised information related to cell phones and electronics. These separate datasets were merged to form a comprehensive multi-source domain dataset.

4.3. Aspect polarity
This section delves into the analysis of aspect-based sentiment, showcasing how different aspects within sentences can be associated with varying polarity scores, and providing insights into sentiments expressed within text data. In this section, Table 2 is discussed, which presents a collection of sentences along with their corresponding aspects and polarities. Each sentence is analyzed to identify the specific aspect being discussed and its sentiment-polarity related to that aspect. For instance, the sentence “Boot time is super-fast, around anywhere from 35 seconds to 1 minute” highlights the aspect of speed, which is associated with a positive polarity of 0. Conversely, the sentence “tech support would not fix the problem unless I bought your plan for $150 plus” focuses on the aspect of the problem, with a negative polarity of -1. Similarly, other sentences such as “but in resume this computer rocks!” emphasize the aspect of the computer, conveying a positive polarity of 1, while “Setup was easy” highlights the aspect of setup with a positive polarity. Lastly, the sentence “Did not enjoy the new Windows 8 and touchscreen functions” does not specify a particular aspect and is labeled as “None” with a negative polarity of 0.

Table 2. Sentences and their related aspects and polarity

Sentence	Aspect	Polarity
Boot time is super-fast, around anywhere from 35 seconds to 1 minute.	Fast	0
tech support would not fix the problem unless I bought your plan for $150 plus.	Problem	-1
but in resume this computer rocks!	Computer	1
The setup was easy.	Setup	1
Did not enjoy the new Windows 8 and touchscreen functions.	None	0

4.4. Sentence polarity
This section provides a detailed breakdown of sentiment analysis, showcasing how probabilities in different sentiment categories contribute to the overall sentiment polarity of each sentence. In this section, Table 3 is discussed, which provides an analysis of sentiment polarity for a set of sentences. The table includes columns for negative, neutral, and positive probabilities, along with the overall sentiment polarity score. Each sentence is evaluated to determine the probability distribution across these sentiment categories, indicating the degree of negativity, neutrality, and positivity expressed in the text. For instance, the sentence

"Boot time is super-fast, around anywhere from 35 seconds to 1 minute" has a relatively balanced distribution among the three categories, resulting in a neutral sentiment polarity score of 0. Conversely, the sentence "tech support would not fix the problem unless I bought your plan for $150 plus" has a high probability in the negative category, leading to a negative sentiment polarity score of -1. Similarly, sentences like "but in resume this computer rocks!" and "Setup was easy" exhibit strong positive probabilities, resulting in positive sentiment polarity scores of 1 for both sentences. On the other hand, the sentence "Did not enjoy the new Windows 8 and touchscreen functions" shows a predominant probability in the negative category, contributing to a negative sentiment polarity score of -1.

Table 3. Sentences polarity scores

Sentence	Negative	Neutral	Positive	Polarity
Boot time is super fast, around anywhere from 35 seconds to 1 minute.	0.0313	0.5616	0.4069	0
tech support would not fix the problem unless I bought your plan for $150 plus.	0.78885	0.19955	0.0118	-1
but in resume this computer rocks!	0.005	0.0518	0.9427	1
Setup was easy.	0.021	0.3667	0.6119	1
Did not enjoy the new Windows 8 and touchscreen functions.	0.927	0.06666	0.0060	-1

4.5. Overall sentiment polarity

The results from the SemEval laptop task evaluation reveal interesting insights into the performance of various models in ABSA as presented in Figure 4. Starting with the RGAT model, it achieved an accuracy of 78.02% and a Macro-F1 score of 74, indicating a decent performance in identifying aspects and sentiments within laptop-related reviews. Moving on to MultiGCN, it showed a slight improvement with an accuracy of 78.8% and a Macro-F1 score of 74.97, showcasing its effectiveness in capturing syntactic and semantic information for sentiment analysis tasks. The MWGCN model demonstrated further improvement with an accuracy of 79.78% and a remarkable Macro-F1 score of 76.68, highlighting its capability to extract contextual features and aspect-related semantics accurately. The CDM-CDW model showcased competitive performance with an accuracy of 80% and an impressive Macro-F1 score of 79.81, indicating its effectiveness in leveraging context features and dynamic weighting for sentiment analysis tasks. Finally, the LDA-TF-BERT model emerged as the top performer with an accuracy of 82.19% and a Macro-F1 score of 79.52, showcasing its superiority in utilizing LDA for aspect extraction and TF-BERT for sentiment polarity evaluation.

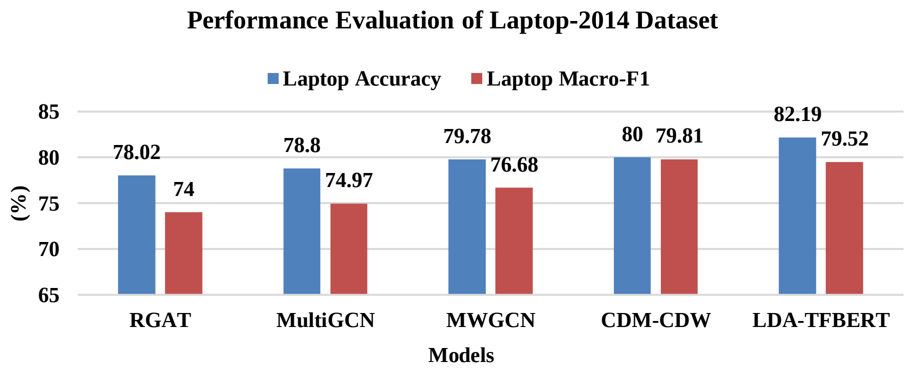

Figure 4. Performance evaluation using laptop-2014 dataset

The results from the SemEval restaurant task evaluation provide valuable insights into the performance of different models in aspect-based sentiment analysis within the restaurant domain as presented in Figure 5. Beginning with the RGAT model, it achieved a notable accuracy of 83.55% and a respectable Macro-F1 score of 75.99, showcasing its effectiveness in identifying aspects and sentiments within restaurant reviews. The MultiGCN model demonstrated a slight improvement with an accuracy of 83.82% and a Macro-F1 score of 77.1, indicating its capability to capture syntactic and semantic information for sentiment analysis tasks in the restaurant domain. Moving on to the MWGCN model, it showcased significant improvement with an impressive accuracy of 86.36% and an outstanding Macro-F1 score of 80.54, highlighting its ability to extract contextual features and aspect-related semantics accurately, thereby enhancing sentiment analysis performance substantially. The CDM-CDW model, although showing a competitive accuracy of 80.5%, fell slightly behind with a Macro-F1 score of 80.3, suggesting that while it effectively leveraged context features and dynamic weighting, there is room for improvement in capturing sentiments within restaurant reviews. Finally, the LDA-TF-BERT model emerged as the top performer with an exceptional accuracy of 87.26% and an impressive Macro-F1 score of 81.27, showcasing its superiority in LDA for aspect extraction and TF-BERT for sentiment polarity evaluation. This demonstrates the effectiveness of combining topic modeling and DL techniques for robust ABSA in the restaurant domain, leading to an enhanced understanding of customer sentiments and preferences.

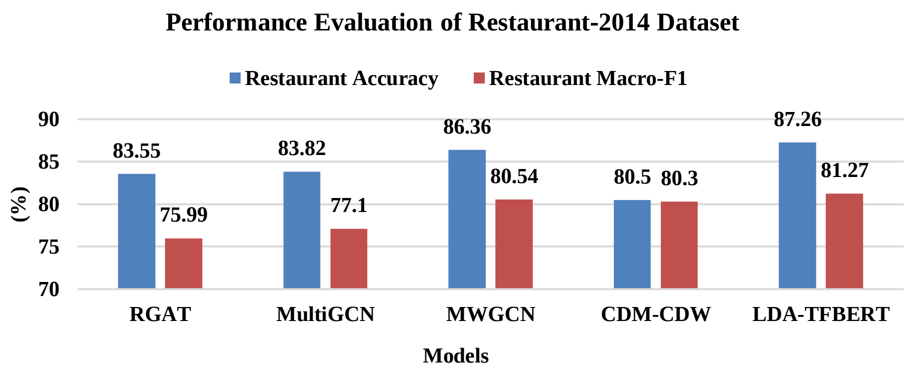

Figure 5. Performance evaluation using laptop-2014 dataset

Further, Table 4 presents a comparative study of various models in terms of accuracy and Macro-F1 scores for both the laptop and restaurant tasks, along with a multi-category evaluation. Each model's performance is analyzed across different years, providing insights into their effectiveness in ABSA. Starting with the LCF-BERT-CDM and LCF-BERT-CDW models from 2019, they showcase competitive performance with accuracy ranging from 79.91% to 80.05% and Macro-F1 scores ranging from 79.69% to 79.85% in the laptop task. Similarly, in the restaurant task, their accuracy ranges from 80.13% to 80.2%, with Macro-F1 scores between 79.97% and 80.03%. These models establish a solid foundation for ABSA. Moving on to the RGAT model from 2022, it exhibits decent accuracy in the laptop task (78.02%) but performs better in the Restaurant task with an accuracy of 83.55%. However, its Macro-F1 scores are comparatively lower, suggesting room for improvement in capturing sentiments. The LGCF-CDM-CDW model from 2019 shows promising results, especially in the restaurant task, with accuracy reaching 85.52% and a Macro-F1 score of 79.85%. This indicates its effectiveness in handling ABSA in restaurant-related reviews. The proposed LDA-TF-BERT model outperforms most of the previous models, showcasing superior accuracy and Macro-F1 scores across all tasks. In the Laptop task, it achieves an accuracy of 82.19% and a Macro-F1 score of 79.52, while in the restaurant task, it achieves an accuracy of 86.26% and a Macro-F1 score of 80.27. The multi-domain evaluation further highlights its robustness, with an accuracy of 87.14% and a Macro-F1 score of 85.15. Overall, the comparative study underscores the advancements in aspect-based sentiment analysis models, with newer models like LDA-TF-BERT demonstrating superior performance in accurately detecting aspects and sentiments in reviews, thus contributing significantly to market insights and customer understanding.

Table 4. Comparative study

Models	Laptop		Restaurant		Multi-Domain	
	Accuracy	Macro-F1	Accuracy	Macro-F1	Accuracy	Macro-F1
LCF-BERT-CDM, 2019, [16]	79.91	79.69	80.2	80.03	-	-
LCF-BERT-CDW, 2019, [16]	80.05	79.85	80.13	79.97	-	-
LCFS-BERT-CDM, 2020, [17]	79.88	79.56	80.34	80.19	-	-
LCFS-BERT-CDW, 2020, [17]	79.93	79.66	80.41	80.25	-	-
RGAT, 2022, [18]	78.02	74	83.55	75.99	-	-
LGCF-CDM-CDW [19]	81.29	78.86	85.52	79.85	-	-
MWGCN, 2023 [20]	79.78	76.68	86.36	80.54	-	-
Multi-GCN, 2023, [21]	78.8	74.97	83.82	77.1	-	-
CDM-CDW, 2023 [22]	80	79.81	80.59	80.39	-	-
SDTGCN, 2024 [23]	78.64	75.50	83.32	76.13	-	-
LDA-TF-BERT (Proposed)	82.19	79.52	87.26	81.27	87.14	85.15

5. CONCLUSION

In conclusion, this comprehensive work has delved deep into the realm of ABSA. Through a systematic approach, various models and methodologies were evaluated and compared across different domains, including the laptop and restaurant tasks, to discern their efficacy in extracting aspects and determining sentiment polarity from textual data. The study began with an exploration of existing literature, highlighting the importance of ABSA in understanding customer sentiments, market trends, and product evaluations. The need for accurate and efficient ABSA models was emphasized, especially in industries such as e-commerce, hospitality, and technology, where customer feedback plays a pivotal role. The methodology detailed the LDA-TF-BERT model and evaluation. Diverse datasets from sources like SemEval, Amazon, and hospital reviews were utilized to create a comprehensive multi-source domain dataset, ensuring the robustness and generalization of the models. The experimental results showcased the performance of various models, including RGAT, Multi-GCN, MWGCN, CDM-CDW, and the proposed LDA-TF-BERT model. The evaluation metrics such as accuracy and Macro-F1 scores were analyzed across different tasks and domains, providing valuable insights into each model's strengths and limitations. Notably, the LDA-TF-BERT model emerged as a top performer, demonstrating superior accuracy and Macro-F1 scores in both the Laptop and Restaurant tasks. Its ability to leverage LDA for aspect extraction and TF-BERT for sentiment polarity evaluation proved highly effective, showcasing its potential for real-world applications in market analysis and customer feedback interpretation. In future work, the model can be evaluated using different datasets to validate its performance.

REFERENCES

[1] G. Kontonatsios *et al.*, "FABSA: an aspect-based sentiment analysis dataset of user reviews," *Neurocomputing*, vol. 562, Dec. 2023, doi: 10.1016/j.neucom.2023.126867.

[2] S. Datta and S. Chakrabarti, "Aspect based sentiment analysis for demonetization tweets by optimized recurrent neural network using fire fly-oriented multi-verse optimizer," *Sādhanā*, vol. 46, no. 2, Jun. 2021, doi: 10.1007/s12046-021-01608-1.

[3] G. Kaur and A. Sharma, "HAS: hybrid analysis of sentiments for the perspective of customer review summarization," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 9, pp. 11971–11984, Sep. 2023, doi: 10.1007/s12652-022-03748-6.

[4] H. Taherdoost and M. Madanchian, "Artificial intelligence and sentiment analysis: a review in competitive research," *Computers*, vol. 12, no. 2, Feb. 2023, doi: 10.3390/computers12020037.

[5] N. Jeong and J. Lee, "An aspect-based review analysis using ChatGPT for the exploration of hotel service failures," *Sustainability*, vol. 16, no. 4, Feb. 2024, doi: 10.3390/su16041640.

[6] Y. C. Hua, P. Denny, K. Taskova, and J. Wicker, "A systematic review of aspect-based sentiment analysis (ABSA): domains, methods, and trends," *arXiv e-prints, arXiv-2311*, 2023.

[7] G. D'Aniello, M. Gaeta, and I. La Rocca, "KnowMIS-ABSA: an overview and a reference model for applications of sentiment analysis and aspect-based sentiment analysis," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5543–5574, Oct. 2022, doi: 10.1007/s10462-021-10134-9.

[8] R. K. Yadav, L. Jiao, M. Goodwin, and O.-C. Granmo, "Positionless aspect based sentiment analysis using attention mechanism," *Knowledge-Based Systems*, vol. 226, Aug. 2021, doi: 10.1016/j.knosys.2021.107136.

[9] S. Kalbhor and D. Goyal, "Survey on ABSA based on machine learning, deep learning and transfer learning approach," *AIP Conference Proceedings*, 2023, doi: 10.1063/5.0154549.

[10] B. Harjo and R. Abdullah, "Attention-based sentence extraction for aspect-based sentiment analysis with implicit aspect cases in hotel review using machine learning algorithm, semantic similarity, and BERT," *International Journal of Intelligent Engineering and Systems*, vol. 16, no. 3, pp. 189–200, Jun. 2023, doi: 10.22266/ijies2023.0630.15.

[11] K. W. Trisna and H. J. Jie, "Deep learning approach for aspect-based sentiment classification: a comparative review," *Applied Artificial Intelligence*, vol. 36, no. 1, Dec. 2022, doi: 10.1080/08839514.2021.2014186.

[12] H. H. Do, P. W. C. Prasad, A. Maag, and A. Alsadoon, "Deep learning for aspect-based sentiment analysis: a comparative review," *Expert Systems with Applications*, vol. 118, pp. 272–299, Mar. 2019, doi: 10.1016/j.eswa.2018.10.003.

[13] W. G. Hatcher and W. Yu, "A survey of deep learning: platforms, applications and emerging research trends," *IEEE Access*, vol. 6, pp. 24411–24432, 2018, doi: 10.1109/ACCESS.2018.2830661.

[14] S. Razavi, "Deep learning, explained: Fundamentals, explainability, and bridge ability to process-based modelling," *Environmental Modelling & Software*, vol. 144, Oct. 2021, doi: 10.1016/j.envsoft.2021.105159.

[15] J. Sun, P. Han, Z. Cheng, E. Wu, and W. Wang, "Transformer based multi-grained attention network for aspect-based sentiment analysis," *IEEE Access*, vol. 8, pp. 211152–211163, 2020, doi: 10.1109/ACCESS.2020.3039470.

[16] B. Zeng, H. Yang, R. Xu, W. Zhou, and X. Han, "LCF: a local context focus mechanism for aspect-based sentiment classification," *Applied Sciences*, vol. 9, no. 16, Aug. 2019, doi: 10.3390/app9163389.

[17] M. H. Phan and P. O. Ogunbona, "Modelling context and syntactical features for aspect-based sentiment analysis," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 3211–3220, doi: 10.18653/v1/2020.acl-main.293.

[18] H. Wu, Z. Zhang, S. Shi, Q. Wu, and H. Song, "Phrase dependency relational graph attention network for aspect-based sentiment analysis," *Knowledge-Based Systems*, vol. 236, Jan. 2022, doi: 10.1016/j.knosys.2021.107736.

[19] J. He, A. Wumaier, Z. Kadeer, W. Sun, X. Xin, and L. Zheng, "A local and global context focus multilingual learning model for aspect-based sentiment analysis," *IEEE Access*, vol. 10, pp. 84135–84146, 2022, doi: 10.1109/ACCESS.2022.3197218.

[20] B. Yu and S. Zhang, "A novel weight-oriented graph convolutional network for aspect-based sentiment analysis," *The Journal of Supercomputing*, vol. 79, no. 1, pp. 947–972, Jan. 2023, doi: 10.1007/s11227-022-04689-9.

[21] Y. Ma, R. Song, X. Gu, Q. Shen, and H. Xu, "Multiple graph convolutional networks for aspect-based sentiment analysis," *Applied Intelligence*, vol. 53, no. 10, pp. 12985–12998, May 2023, doi: 10.1007/s10489-022-04023-z.

[22] T. Lin and I. Joe, "An adaptive masked attention mechanism to act on the local text in a global context for aspect-based sentiment analysis," *IEEE Access*, vol. 11, pp. 43055–43066, 2023, doi: 10.1109/ACCESS.2023.3270927.

[23] Q. Zhao, F. Yang, D. An, and J. Lian, “Modeling structured dependency tree with graph convolutional networks for aspect-level sentiment classification,” *Sensors*, vol. 24, no. 2, Jan. 2024, doi: 10.3390/s24020418.

[24] A. Farkhod, A. Abdusalomov, F. Makhmudov, and Y. I. Cho, “LDA-based topic modeling sentiment analysis using topic/document/sentence (TDS) model,” *Applied Sciences*, vol. 11, no. 23, Nov. 2021, doi: 10.3390/app112311091.

[25] K. S. Arun and V. K. Govindan, “A hybrid deep learning architecture for latent topic-based image retrieval,” *Data Science and Engineering*, vol. 3, no. 2, pp. 166–195, Jun. 2018, doi: 10.1007/s41019-018-0063-7.

[26] S. Syed and M. Spruit, “Full-text or abstract? examining topic coherence scores using latent Dirichlet allocation,” in *2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, Oct. 2017, pp. 165–174, doi: 10.1109/DSAA.2017.61.

[27] “BERT,” *Hugging Face*. https://huggingface.co/docs/transformers/en/model_doc/bert (accessed Aug. 30, 2024).

[28] Practo, “Book doctor appointments online, order medicine, diagnostic tests, consult,” *practo.com*, 2017. https://www.practo.com/bangalore/doctors (accessed Aug. 30, 2024).

[29] “India’s largest review platform,” *mouthshut.com*, https://www.mouthshut.com/category/bangalore-hospitals-and-clinics-reviews-925044773 (accessed Feb. 13, 2024).

[30] M. Pontiki, D. Galanis, J. Pavlopoulos, H. Papageorgiou, I. Androutsopoulos, and S. Manandhar, “SemEval-2014 Task 4: aspect based sentiment analysis,” in *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, 2014, pp. 27–35, doi: 10.3115/v1/S14-2004.

[31] J. Ni, J. Li, and J. McAuley, “Justifying recommendations using distantly-labeled reviews and fine-grained aspects,” in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, 2019, pp. 188–197, doi: 10.18653/v1/D19-1018.

BIOGRAPHIES OF AUTHORS

Radhika Jinendra Dhanal is an academician from Kolhapur, Maharashtra, with a passion for computer science and information systems. She is currently affiliated with DY Patil College of Engineering and Technology, situated in Kasaba Bawada, Kolhapur, where she has been contributing to the field of education since 2003. Her academic journey led her to pursue a master's in engineering (M.E.) from Shivaji University in 2012, specializing in her area of interest. Her commitment to academic excellence and a deep understanding of computer science and information systems laid the foundation for her successful career. With a rich academic background, a wealth of experience, and a commitment to advancing the field of Computer Science, Miss Radhika Dhanal stands as a noteworthy figure in the academic landscape of DY Patil College of Engineering and Technology, Kolhapur. Her journey continues to inspire both students and fellow educators alike. She can be contacted at email: dhanallraddheka@gmail.com.

Vijay Ram Ghorpade is an academic researcher from Shivaji University. He is the Professor/Principal of Bharti Vidyapeeth’s College of Engineering, Kolhapur. He has contributed to research on topics: Wireless sensor networks and Quality of service. He has a H-index of 7, co-authored 24 publications receiving 146 citations. He can be contacted at email: vijayghorpade@hotmail.com.