engineering proceedings

MDPI

Proceeding Paper

Parameter-Efficient Adaptation of Qwen2.5 for Aspect-Based Sentiment Analysis Using Low-Rank Adaptation and Parameter-Efficient Fine-Tuning †

Pei Ying Lim *, Chuk Fong Ho and Chi Wee Tan

Faculty of Computing and Information Technology, Tunku Abdul Rahman University of Management and Technology, Kuala Lumpur 53300, Malaysia; cfho@tarc.edu.my (C.F.H.); chiwee@tarc.edu.my (C.W.T.)

* Correspondence: pylim-wp19@student.tarc.edu.my; Tel.: +60-14-6480830

† Presented at 2025 IEEE International Conference on Computation, Big-Data and Engineering (ICCBE), Penang, Malaysia, 27–29 June 2025.

Abstract

Aspect-based sentiment analysis (ABSA) plays a vital role in deriving fine-grained sentiment from textual content. As large language models (LLMs) are increasingly adopted for automated data annotation in natural language processing (NLP), concerns have emerged regarding the accuracy of their outputs. Despite their capacity to generate large volumes of labeled data, LLMs often suffer from overconfidence in predictions, high uncertainty in complex contexts, and difficulty capturing nuanced meanings, which compromise the quality of annotations and, in turn, the performance of downstream models. This underscores the need to enhance LLM adaptability while maintaining annotation accuracy. To address these limitations, we integrated low-rank adaptation (LoRA) with parameter-efficient fine-tuning (PEFT) for adapting Qwen2.5 to ABSA. LoRA reduces the number of trainable parameters by decomposing weight updates into low-rank matrices, while PEFT introduces modular adapter layers with scaled gradient updates and dynamic rank allocation. Using the standard SemEval 2014 Laptop dataset, Qwen2.5-3B fine-tuned with LoRA and PEFT achieves 64.50% accuracy, outperforming its baseline of 24.05%. Likewise, Qwen2.5-7B attains 77.50%, compared with a baseline of 34.63%. These results highlight the potential of parameter-efficient methods to improve the accuracy of LLMs in ABSA annotation tasks, especially under resource constraints. Such results lay the groundwork for scalable, reproducible LLM deployment and open avenues for future research in cross-domain adapter transferability and dynamic rank optimization.

Keywords: aspect-based sentiment analysis; large language models; Qwen2.5; parameter-efficient fine-tuning; low-rank adaptation

Academic Editors: Teen-Hang Meen, Wei Chien and Cheng-Fu Yang

Published: 9 March 2026

1. Introduction

Aspect-based sentiment analysis (ABSA) is an important subfield of sentiment analysis in natural language processing (NLP) that focuses on extracting fine-grained opinions from text by identifying sentiments expressed toward specific aspects of entities, such as "food" or "environment" in restaurant reviews (Figure 1) [1,2].

This granularity provides actionable information for customer feedback analysis [3], product improvement [3,4], and reputation management [4]. However, traditional ABSA approaches face significant challenges, particularly in handling context dependency [5,6] and aligning aspect terms with their corresponding opinion expressions [1,5,6]. Ambiguity in language makes words carry multiple meanings depending on context and complicates

accurate aspect–opinion alignment and sentiment polarity detection, often limiting the effectiveness of conventional algorithms [1,2]. To address these challenges, large language models (LLMs) are used to strengthen ABSA pipelines, automating data annotation [7–9], which remains a critical bottleneck in traditional approaches.

Result of ABSA:

Aspect	Sentiment
food	positive
environment	negative

Figure 1. Sample of ABSA.

The advantages of LLMs, such as Tongyi Qianwen (Qwen), enable data annotation by automating the labeling process for complex NLP, including ABSA [10]. While LLMs have shown promising capabilities in generating annotations at scale and reducing the dependence on manual labeling, their outputs sometimes lack consistency or contain errors, necessitating careful validation or human-in-the-loop refinement to ensure high-quality annotations [10–12]. Besides, LLM-based annotation systems exhibit notable limitations. They might display overconfidence in uncertain contexts, struggle to capture nuanced sentiment, and misalign with human judgments, especially when subtle interpretive cues are required [13,14]. These challenges highlight the need for robust validation and calibration mechanisms when leveraging LLMs for annotation tasks [13,14].

While LLMs exhibit strong zero-shot and few-shot capabilities [15], achieving optimal performance on specialized tasks such as ABSA often requires fine-tuning, also referred to as model calibration [12,16]. Fine-tuning involves updating the model's parameters on task-specific annotated data to better capture the nuances and domain-specific language patterns relevant to ABSA [16,17]. In the process, validation on a held-out dataset must be conducted to monitor the model's generalization performance and prevent overfitting, ensuring that improvements on training data translate to real-world effectiveness [18–20].

However, fine-tuning large-scale LLMs is computationally expensive due to their massive parameter counts, which pose significant resource and time constraints [21–24]. To overcome these constraints, researchers have proposed Parameter-Efficient Fine-Tuning (PEFT) methods. PEFT updates a small subset of parameters while keeping the majority fixed, significantly lowering the computational burden without compromising performance [25]. Among PEFT methods, one notable PEFT method is Low-Rank Adaptation (LoRA), which has gained attention due to its efficiency and effectiveness in adapting LLMs [26]. LoRA injects trainable low-rank matrices into existing model layers, leveraging the insight that task-specific knowledge can often be represented in a lower-dimensional subspace [26]. This approach enables efficient adaptation of LLMs with minimal memory and compute requirements, making fine-tuning more accessible for specialized NLP tasks like ABSA.

To address the limitations of LLMs in ABSA, including annotation inconsistency, overconfidence in ambiguous contexts, and difficulty in capturing nuanced sentiment, we investigated the effectiveness of fine-tuning LLMs using PEFT techniques. In particular, we explored how to integrate LoRA, a prominent PEFT method, to improve model adaptability and annotation quality without incurring the high computational costs of fine-tuning. By combining LoRA's low-rank decomposition with PEFT's efficient update mechanisms, the scalability and reliability of fine-tuned Qwen2.5 model variants are enhanced for ABSA. Beyond evaluating fine-tuning effectiveness, a dynamic rank allocation strategy

in LoRA was also proposed in this study to better balance computational overhead and task-specific performance. The results of this study can be used to optimize resource efficiency and improve model robustness and alignment with human judgment. A practical and reproducible approach was also proposed for advancing LLM-based ABSA systems through targeted fine-tuning and architectural optimization.

2. Related Work

2.1. ABSA Methodologies

Traditional ABSA approaches involve aspect extraction and sentiment classification. However, recent advances favor end-to-end learning models that jointly optimize these tasks, improving performance and reducing error propagation [27–29]. Transformer-based architectures, especially those built on Bidirectional Encoder Representations from Transformers (BERT) and its variants, have become dominant in ABSA due to their strong contextual representation capabilities [30–32]. BERT's bidirectional self-attention mechanism enables a nuanced understanding of sentiment and aspect terms by considering context from both left and right sides simultaneously, outperforming earlier unidirectional models [33]. WordTransABSA further enhances such capability by leveraging the entire transformer parameters and utilizing sentiment-related pivot tokens to predict affective tokens for target words, showing superior results in both full-data and few-shot learning scenarios [34]. Moreover, Large Language Model Meta AI (LLaMA) improves cross-domain generalization, reducing the need for extensive domain-specific fine-tuning [35]. These transformer-based end-to-end methods have thus set new benchmarks in ABSA tasks across various domains [35–37].

Despite the strong performance of transformer-based end-to-end models such as BERT and its variants, as well as LLaMA for improved cross-domain generalization, limitations still hinder their effectiveness in ABSA tasks. One major limitation is annotation inconsistency, particularly in datasets involving implicit aspects. Human annotation in these cases is cognitively demanding and often leads to lower inter-annotator agreement, as identifying abstract or implied sentiments is inherently subjective and error-prone [8]. Another limitation is overconfidence in ambiguous or context-dependent cases, where LLMs generate confident yet incorrect predictions, undermining the robustness and reliability of sentiment analysis systems [38]. Lastly, LLMs often struggle to capture nuanced or implicit sentiment, especially when trained on synthetic datasets that lack the contextual richness and lexical diversity of real-world text [8]. These limitations highlight the need for improved annotation strategies, better model calibration, and more sophisticated reasoning techniques, such as multi-step prompting or iterative refinement, to enhance the reliability and generalizability of LLM-based ABSA across diverse application domains.

2.2. PEFT in LLMs

While these limitations pose challenges to LLM-based ABSA, recent advancements in fine-tuning strategies, particularly PEFT, enable a promising methodology. LLMs significantly advance ABSA by enabling powerful contextual understanding and generalization [12,39]. However, despite their effectiveness, fine-tuning large-scale LLMs remains computationally expensive and resource-intensive due to their massive parameter sizes, limiting their accessibility and scalability in practical applications [21–24]. To overcome these challenges, PEFT techniques are introduced to full fine-tuning [25,40–43]. Early approaches involved updating all model parameters, but PEFT techniques such as adapters and LoRA have enabled efficient adaptation by tuning only a small subset of parameters or injecting lightweight modules within the frozen backbone [26,44]. This evolution allows for rapid and resource-efficient customization of LLMs for downstream tasks without com-

promising performance [33]. These strategies have made it feasible to apply LLMs across specialized tasks such as ABSA, even under resource constraints. In line with this trend, the Qwen model adopts PEFT methods to efficiently adapt to diverse NLP applications, balancing fine-tuning cost and inference efficiency [45–47]. Although detailed specifics on Qwen adaptations are emerging, the trend aligns with broader PEFT strategies that prioritize modular and scalable fine-tuning [33].

Nevertheless, the application of PEFT techniques, particularly LoRA, to ABSA using Qwen models remains unexplored. This highlights the need for a systematic investigation into how PEFT enhances Qwen's performance on ABSA tasks while simultaneously minimizing computational overhead. This study aims to fill this gap by developing and evaluating the effectiveness and efficiency of PEFT-based fine-tuning strategies tailored for Qwen in ABSA, thereby advancing efficient and scalable sentiment analysis solutions.

3. Methodology

3.1. Data Description

The Semantic Evaluation (SemEval) 2014 Laptop domain dataset: https://aclanthology.org/S14-2004/ (accessed on 20 May 2025) is widely used for ABSA [48–56]. It consists of over 3000 English sentences extracted from customer reviews of laptops. The dataset is divided into a training set containing 3045 records, a testing set with 800 records, and an evaluation set comprising 219 records. Each entry in the dataset includes fields such as SentenceID, Raw_text, AspectTerms, and Aspect Categories, enabling detailed and structured analysis of customer opinions. For this dataset, experienced annotators tag aspect terms within sentences (Subtask 1) and assign sentiment polarities to these aspects (Subtask 2) [48]. The dataset is provided in XML format, with detailed annotations specifying the exact aspect terms and their corresponding sentiment polarity labels, which include positive, negative, neutral, conflict, and none (Table 1) [48].

Table 1. Sample records from the SemEval 2014 Laptop domain dataset.

SentenceID	Raw_text	AspectTerms	AspectCategories
2339	I charge it at night and skip taking the cord with me because of the good battery life.	[{'term': 'cord', 'polarity': 'neutral'}, {'term': 'battery life', 'polarity': 'positive'}]	[{'category': 'noaspectcategory', 'polarity': 'none'}]
812	I bought an HP Pavilion DV4-1222nr laptop and have had so many problems with the computer.	[{'term': 'noaspectterm', 'polarity': 'none'}]	[{'category': 'noaspectcategory', 'polarity': 'none'}]
562	Did not enjoy the new Windows 8 and touchscreen functions.	[{'term': 'Windows 8', 'polarity': 'negative'}, {'term': 'touchscreen functions', 'polarity': 'negative'}]	[{'category': 'noaspectcategory', 'polarity': 'none'}]
912	The price is higher than most laptops out there; however, he/she will get what they paid for, which is a great computer.	[{'term': 'price', 'polarity': 'conflict'}]	[{'category': 'noaspectcategory', 'polarity': 'none'}]

This dataset was introduced as part of SemEval-2014 Task 4 to foster research in fine-grained sentiment analysis by focusing on identifying specific aspects of target entities and the sentiments expressed toward them [48]. Among the various domains, the laptop reviews domain is notable because it presents unique challenges due to the technical nature of the product and the variety of aspects, such as battery, screen, and performance, on which customers frequently comment. Its well-annotated and domain-specific nature has

made it a standard benchmark for evaluating ABSA models. It has been extensively used in research to train and assess traditional machine learning approaches and LLMs in aspect term extraction and sentiment classification [25].

3.2. Qwen Models

In the experiments, we utilized the Qwen2.5 series of LLMs, specifically the Qwen2.5-3B and Qwen2.5-7B variants. Qwen2.5 represents a significant advancement over its predecessors, trained on an expanded dataset of 18 trillion tokens, which enhances its common sense, expert knowledge, and reasoning capabilities [57]. The model includes base pretrained and instruction-tuned variants, demonstrating state-of-the-art performance across diverse benchmarks in language understanding, reasoning, coding, and human preference alignment [57]. Qwen2.5 models support quantization and mixed-precision training, facilitating efficient deployment while maintaining high accuracy and versatility in various NLP [57].

3.3. LoRA

Among various PEFT techniques, LoRA was chosen in this study due to its strong balance of efficiency, performance, and ease of integration. Unlike other PEFT techniques that either introduce additional latency (e.g., adapters) or require complex tuning (e.g., prefix tuning), LoRA injects trainable low-rank matrices directly into transformer attention layers while keeping the original pretrained weights frozen [26]. This approach drastically reduces the number of trainable parameters, often to less than 0.1% of the full model size, enabling fine-tuning on limited computational resources without sacrificing model expressiveness or inference speed [26]. Moreover, LoRA demonstrates superior or comparable performance to full fine-tuning and other PEFT methods across various LLM architectures, including GPT, LLaMA, and Qwen, making it a practical and scalable choice for adapting large models to downstream tasks [58–60]. Additionally, LoRA's compatibility with quantization techniques (e.g., QLoRA) further enhances memory efficiency, supporting deployment in resource-constrained environments. These advantages collectively motivate the choice of LoRA over alternative PEFT methods for efficient and effective fine-tuning.

The Qwen models are fine-tuned using LoRA with the parameters shown in Figure 2, for the following reasons.

- The choice of rank r = 16 strikes a balance between model capacity and computational efficiency, consistent with prior studies that show moderate ranks achieve strong performance without excessive parameter growth [61].
- A scaling factor $\alpha = 32$ is used to appropriately scale the LoRA updates, stabilizing training and ensuring effective adaptation [61].
- Targeting the query, key, value, and output projection layers aligns with LoRA's original design and has been empirically validated as effective for Qwen models [26,62,63].
- A dropout rate of 0.1 is applied to mitigate overfitting during fine-tuning, which is especially important when adapting large models on limited datasets [64].
- The bias parameter is set to "none" to reduce complexity and focus adaptation on the core projection weights [26,65].
- Specifying the task type as "CAUSAL_LM" ensures compatibility with Qwen's autoregressive architecture and training objectives [66].

```
lora_config = LoraConfig(
    r=16,
    lora_alpha=32,
    target_modules=["q_proj", "k_proj", "v_proj", "o_proj"],
    lora_dropout=0.1,
    bias="none",
    task_type="CAUSAL_LM"
)
```

Figure 2. Configuration of LoRA Parameters.

3.4. Model Evaluation

To evaluate the performance of the LLM-based ABSA model, accuracy was calculated. Accuracy measures the proportion of correctly predicted sentiment labels compared to the ground truth annotations in the SemEval 2014 Laptop domain dataset, which is a widely accepted benchmark for ABSA tasks.

In this dataset, each review sentence contained an aspectTerms field, which is a list of aspect–sentiment pairs. For evaluation, these records were split into term and polarity, effectively transforming each aspect–sentiment pair into an individual evaluation instance. This process increases the number of evaluation samples from the original 800 review records to 1032 aspect-level records, enabling a fine-grained and accurate assessment of the model's ability to predict sentiment polarity for each specific aspect.

Accuracy was computed as the ratio of correctly predicted polarity labels to the total number of aspect-level records, reflecting a clear and interpretable measure of the model's effectiveness in identifying the correct sentiment for each aspect term. Accuracy was used in previous studies on ions using the SemEval 2014 Laptop dataset, facilitating consistent comparison with existing ABSA models [20,48–50,67].

4. Results and Discussion

The accuracy and matched records are presented in Figure 3. Fine-tuning using PEFT and LoRA enhances the performance of both the Qwen2.5-3B and Qwen2.5-7B models. The fine-tuned models achieved higher accuracy and matched a larger number of records from the testing dataset compared to their non-fine-tuned counterparts. For instance, the accuracy of Qwen2.5-3B improved from 24.50% to 64.50%, while that of Qwen2.5-7B rose from 34.63% to 77.50%. This performance boost indicates that fine-tuning plays a crucial role in adapting LLMs to specific tasks, showing a substantial improvement in predictive accuracy. The results demonstrate the effectiveness of applying fine-tuning techniques over using pretrained models alone.

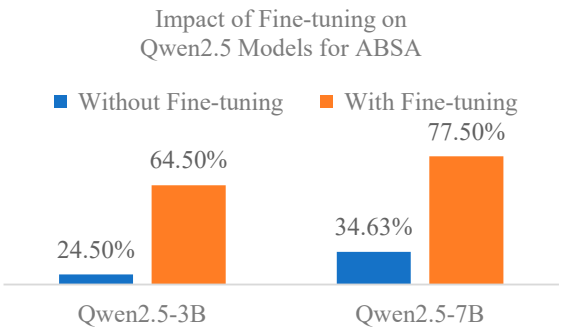

Figure 3. Accuracies of Qwen2.5 models with and without fine-tuning.

In addition to enhanced accuracy, the fine-tuned models produced more consistent annotations across similar input samples, which reduces annotation inconsistency, a known

limitation in LLM-based ABSA systems. Moreover, a review of model confidence scores showed fewer confidently incorrect predictions in ambiguous or borderline sentiment cases, suggesting a mitigation of overconfidence in uncertain contexts. Finally, qualitative inspection of outputs on sentences with implicit or subtle sentiment cues demonstrated that the fine-tuned models were better able to infer and align aspect–opinion pairs correctly, thereby addressing the challenge of capturing nuanced or implicit sentiment. These results collectively suggest that PEFT with LoRA enhances quantitative performance metrics and overcomes the qualitative limitations of LLMs in ABSA.

While fine-tuning the Qwen2.5-3B and Qwen2.5-7B models using PEFT with LoRA significantly improved accuracy on the ABSA task, it is important to consider the computational resources required. On a system equipped with a 13th Gen Intel Core i7-13700F processor and 64 GB RAM, the fine-tuning durations were approximately 15 h and 36 min for Qwen2.5-3B and about 74 h for Qwen2.5-7B. These lengths of time reflect a substantial but manageable resource investment, considering the model sizes and hardware constraints.

Full fine-tuning of Qwen2.5-14B and other dense variants requires multi-GPU setups with over 60 GB of GPU VRAM and training durations extending from several days to weeks (Table 2). In contrast, PEFT with LoRA reduces hardware requirements and training time. The ability to fine-tune sizable LLMs on a single commodity machine without sacrificing performance highlights the practical value of PEFT-LoRA techniques, broadening accessibility for researchers and practitioners with limited computational resources.

Table 2. Resource requirements and fine-tuning time for large Qwen models without parameter-efficient techniques.

LLM Model	Parameter Size (Billion)	Hardware Setup	Random Access Memory (RAM)/Graphics Processing Unit (GPU)	Fine-Tuning Time
Qwen2.5-14B [68]	14 B	Multi-GPU (A100 or equivalent)	Estimated >60GB GPU video RAM (VRAM)	Not explicitly reported
Qwen2 dense models [66]	0.5B to 72B	Large GPU clusters (A100 80GB GPUs)	Up to 80GB+ VRAM per GPU	Days to weeks (full pretraining and fine-tuning)
Qwen2.5-1M series [68]	7B and 14B	Multi-GPU setups	High GPU VRAM (>60GB typical)	Not explicitly stated
This study (Qwen2.5-3B & 7B + LoRA)	3B and 7B	Single-machine (Intel i7-13700F, 64GB RAM, RTX 4070 Ti)	64GB RAM, 12GB GPU VRAM	15.6 h (3B) 74 h (7B)

5. Conclusions

The effectiveness of parameter-efficient techniques in adapting LLMs to specialized NLP tasks, such as ABSA, was validated in this study. By integrating LoRA with PEFT, we fine-tuned Qwen2.5-3B and Qwen2.5-7B models and achieved substantial performance improvements, enhancing accuracy from 24.50% to 64.50% and from 34.63% to 77.50%, respectively. These results validate the capability of LoRA and PEFT to overcome the computational burdens of full fine-tuning while preserving model effectiveness.

The developed method is feasible on commodity hardware, reinforcing the practicality of deploying large-scale models in resource-constrained settings. Despite these promising outcomes, its reliance on a single-domain, English-only dataset with a relatively small size must be mitigated. Therefore, future studies should explore multilingual and cross-domain generalization, dynamic rank allocation strategies, and the transferability of adapters across

models and tasks. These directions can further advance the scalability, adaptability, and robustness of LLMs in fine-grained sentiment analysis and beyond.

Author Contributions: Conceptualization, P.Y.L.; methodology, P.Y.L.; software, P.Y.L.; validation, P.Y.L., C.F.H. and C.W.T.; formal analysis, P.Y.L.; investigation, P.Y.L.; data curation, P.Y.L.; writing—original draft preparation, P.Y.L.; writing—review and editing, P.Y.L., C.F.H. and C.W.T.; visualization, P.Y.L.; supervision, C.F.H. and C.W.T.; project administration, P.Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received financial support from Tunku Abdul Rahman University of Management and Technology (TAR UMT), Malaysia.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available in this manuscript.

Acknowledgments: We would like to express our gratitude to Tunku Abdul Rahman University of Management and Technology (TAR UMT), Malaysia, for the resources provided to carry out this work.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

ABSA	Aspect-based Sentiment Analysis
LLMs	Large Language Models
NLP	Natural Language Processing
LoRA	Low-Rank Adaptation
PEFT	Parameter-Efficient Fine-Tuning
Qwen	Tongyi Qianwen

References

1. Zhang, W.; Li, X.; Deng, Y.; Bing, L.; Lam, W. A Survey on Aspect-Based Sentiment Analysis: Tasks, Methods, and Challenges. *IEEE Trans. Knowl. Data Eng.* **2022**, *35*, 11019–11038. [CrossRef]
2. Chifu, A.G.; Fournier, S. Sentiment Difficulty in Aspect-Based Sentiment Analysis. *Mathematics* **2023**, *11*, 4647. [CrossRef]
3. Hua, Y.C.; Denny, P.; Taskova, K.; Wicker, J. A Systematic Review of Aspect-Based Sentiment Analysis: Domains, Methods, and Trends. *Artif. Intell. Rev.* **2023**, *57*, 296. [CrossRef]
4. Ismet, H.T.; Mustaqim, T.; Purwitasari, D. Aspect Based Sentiment Analysis of Product Review Using Memory Network. *Sci. J. Inform.* **2022**, *9*, 73–83. [CrossRef]
5. Xing, B.; Tsang, I.W. Out of Context: A New Clue for Context Modeling of Aspect-Based Sentiment Analysis. *J. Artif. Intell. Res.* **2022**, *74*, 627–659. [CrossRef]
6. Nazir, A.; Rao, Y.; Wu, L.; Sun, L. Issues and Challenges of Aspect-Based Sentiment Analysis: A Comprehensive Survey. *IEEE Trans. Affect. Comput.* **2022**, *13*, 845–863. [CrossRef]
7. Simmering, P.F.; Huoviala, P. Large Language Models for Aspect-Based Sentiment Analysis. *arXiv* **2023**, arXiv:2310.18025. [CrossRef]
8. Neveditsin, N.; Lingras, P.; Mago, V. From Annotation to Adaptation: Metrics, Synthetic Data, and Aspect Extraction for Aspect-Based Sentiment Analysis with Large Language Models. In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 4: Student Research Workshop), Albuquerque, NM, USA, 29 April–4 May 2025*; Association for Computational Linguistics: Kerrville, TX, USA, 2025.
9. Zhong, Q.; Li, H.; Zhuang, L.; Liu, J.; Du, B. Iterative Data Generation with Large Language Models for Aspect-Based Sentiment Analysis. *arXiv* **2024**, arXiv:2407.00341.
10. Tan, Z.; Li, D.; Wang, S.; Beigi, A.; Jiang, B.; Bhattacharjee, A.; Karami, M.; Li, J.; Cheng, L.; Liu, H. Large Language Models for Data Annotation and Synthesis: A Survey. *arXiv* **2024**, arXiv:2402.13446.
11. Yang, X.; Zhan, R.; Wong, D.F.; Wu, J.; Chao, L.S. Human-in-the-Loop Machine Translation with Large Language Model. *arXiv* **2023**, arXiv:2310.08908.

12. Zhou, C.; Song, D.; Tian, Y.; Wu, Z.; Wang, H.; Zhang, X.; Yang, J.; Yang, Z.; Zhang, S. A Comprehensive Evaluation of Large Language Models on Aspect-Based Sentiment Analysis. *arXiv* **2024**, arXiv:2412.02279. [CrossRef]
13. Pangakis, N.; Wolken, S.; Fasching, N. Automated Annotation with Generative AI Requires Validation. *arXiv* **2023**, arXiv:2306.00176. [CrossRef]
14. Gligorić, K.; Zrnic, T.; Lee, C.; Candès, E.J.; Jurafsky, D. Can Unconfident LLM Annotations Be Used for Confident Conclusions? In *Proceedings of the 2025 Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), Albuquerque, NM, USA, 29 April–4 May 2025*; Association for Computational Linguistics: Kerrville, TX, USA, 2024.
15. Huang, J.; Cui, Y.; Liu, J.; Liu, M. Supervised and Few-Shot Learning for Aspect-Based Sentiment Analysis of Instruction Prompt. *Electronics* **2024**, *13*, 1924. [CrossRef]
16. Parthasarathy, V.B.; Zafar, A.; Khan, A.; Shahid, A. The Ultimate Guide to Fine-Tuning LLMs from Basics to Breakthroughs: An Exhaustive Review of Technologies, Research, Best Practices, Applied Research Challenges and Opportunities. *arXiv* **2024**, arXiv:2408.13296. [CrossRef]
17. Ding, X.; Zhou, J.; Dou, L.; Chen, Q.; Wu, Y.; Chen, C.; He, L. Boosting Large Language Models with Continual Learning for Aspect-Based Sentiment Analysis. In *Findings of the Association for Computational Linguistics: EMNLP 2024*; Association for Computational Linguistics: Kerrville, TX, USA, 2024.
18. Šmíd, J.; Přibá, P.; Přibáň, P.; Král, P. LLaMA-Based Models for Aspect-Based Sentiment Analysis. In *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis, Bangkok, Thailand, 15 August 2024*; Association for Computational Linguistics: Kerrville, TX, USA, 2024.
19. Zhang, Y.; Zeng, J.; Hu, W.; Wang, Z.; Chen, S.; Xu, R. Self-Training with Pseudo-Label Scorer for Aspect Sentiment Quad Prediction. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Bangkok, Thailand, 11–16 August 2024*; Association for Computational Linguistics: Kerrville, TX, USA, 2024; Volume 1.
20. Scaria, K.; Gupta, H.; Goyal, S.; Sawant, S.A.; Mishra, S.; Baral, C. InstructABSA: Instruction Learning for Aspect Based Sentiment Analysis. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers), Mexico City, Mexico, 16–21 June 2024*; Short Papers; Association for Computational Linguistics: Kerrville, TX, USA, 2024; Volume 2.
21. Azizi, S.; Kundu, S.; Pedram, M. LaMDA: Large Model Fine-Tuning via Spectrally Decomposed Low-Dimensional Adaptation. *arXiv* **2024**, arXiv:2406.12832.
22. Zhang, Y.; Li, P.; Hong, J.; Li, J.; Zhang, Y.; Zheng, W.; Chen, P.-Y.; Lee, J.D.; Yin, W.; Hong, M.; et al. Revisiting Zeroth-Order Optimization for Memory-Efficient LLM Fine-Tuning: A Benchmark. *arXiv* **2024**, arXiv:2402.11592.
23. Wu, X.K.; Chen, M.; Li, W.; Wang, R.; Lu, L.; Liu, J.; Hwang, K.; Hao, Y.; Pan, Y.; Meng, Q.; et al. LLM Fine-Tuning: Concepts, Opportunities, and Challenges. *Big Data Cogn. Comput.* **2025**, *9*, 87. [CrossRef]
24. Zhang, B.; Liu, Z.; Cherry, C.; Firat, O. When Scaling Meets Llm Finetuning: The Effect of Data, Model And Finetuning Method. *arXiv* **2024**, arXiv:2402.17193. [CrossRef]
25. Balne, C.C.S.; Bhaduri, S.; Roy, T.; Jain, V.; Chadha, A. Parameter Efficient Fine Tuning: A Comprehensive Analysis Across Applications. *arXiv* **2024**, arXiv:2404.13506. [CrossRef]
26. Hu, E.; Shen, Y.; Wallis, P.; Allen-Zhu, Z.; Li, Y.; Wang, S.; Wang, L.; Chen, W. LORA: Low-rank adaptation of large language models. *ICLR* **2022**, *1*, 3.
27. Liu, S.; Zhou, J.; Zhu, Q.; Chen, Q.; Bai, Q.; Xiao, J.; He, L. Let's Rectify Step by Step: Improving Aspect-Based Sentiment Analysis with Diffusion Models. *arXiv* **2024**, arXiv:2402.15289.
28. Schmitt, M.; Steinheber, S.; Schreiber, K.; Roth, B. Joint Aspect and Polarity Classification for Aspect-Based Sentiment Analysis with End-to-End Neural Networks. *arXiv* **2018**, arXiv:1808.09238.
29. Mao, Y.; Shen, Y.; Yu, C.; Cai, L. A Joint Training Dual-MRC Framework for Aspect Based Sentiment Analysis. *Proc. AAAI Conf. Artif. Intell.* **2021**, *35*, 13543–13551. [CrossRef]
30. Hoang, M.; Alija Bihorac, O.; Rouces, J. Aspect-Based Sentiment Analysis Using BERT. In *Proceedings of the 22nd Nordic Conference on Computational Linguistics, Turku, Finland, 30 September–2 October 2019*; Linköping University Electronic Press: Linköping, Sweden, 2019.
31. Xu, H.; Shu, L.; Yu, P.S.; Liu, B. Understanding Pre-Trained BERT for Aspect-Based Sentiment Analysis. In *Proceedings of the 28th International Conference on Computational Linguistics, Barcelona, Spain, 8–13 December 2020*; International Committee on Computational Linguistics: New York, NY, USA, 2020.
32. Zhang, M.; Zhu, Y.; Liu, Z.; Bao, Z.; Wu, Y.; Sun, X.; Xu, L. Span-Level Aspect-Based Sentiment Analysis via Table Filling. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Toronto, ON, Canada, 9–14 July 2023*; Long Papers; Association for Computational Linguistics: Kerrville, TX, USA, 2023; Volume 1.
33. Ghosh, K.K.; Sur, C. Learning to Extract Cross-Domain Aspects and Understanding Sentiments Using Large Language Models. *arXiv* **2025**, arXiv:2501.08974. [CrossRef]

34. Jin, W.; Zhao, B.; Zhang, Y.; Huang, J.; Yu, H. WordTransABSA: Enhancing Aspect-Based Sentiment Analysis with Masked Language Modeling for Affective Token Prediction. *Expert Syst. Appl.* **2024**, *238*, 122289. [CrossRef]
35. Musa, A.; Adam, F.M.; Ibrahim, U.; Zandam, A.Y. HauBERT: A Transformer Model for Aspect-Based Sentiment Analysis of Hausa-Language Movie Reviews. *Eng. Proc.* **2025**, *87*, 43.
36. Chaudhry, H.N.; Kulsoom, F.; Ullah Khan, Z.; Aman, M.; Khan, S.U.; Albanyan, A. TASCI: Transformers for Aspect-Based Sentiment Analysis with Contextual Intent Integration. *PeerJ Comput. Sci.* **2025**, *11*, e2760. [CrossRef]
37. Taj, S.; Daudpota, S.M.; Imran, A.S.; Kastrati, Z. Aspect-Based Sentiment Analysis for Software Requirements Elicitation Using Fine-Tuned Bidirectional Encoder Representations from Transformers and Explainable Artificial Intelligence. *Eng. Appl. Artif. Intell.* **2025**, *151*, 110632. [CrossRef]
38. Wang, Q.; Ding, K.; Liang, B.; Yang, M.; Xu, R. Reducing Spurious Correlations in Aspect-Based Sentiment Analysis with Explanation from Large Language Models. In *Findings of the Association for Computational Linguistics: EMNLP 2023*; Association for Computational Linguistics: Kerrville, TX, USA, 2023; p. 2941.
39. Cao, J.; Li, J.; Yang, Z.; Zhou, R. Enhanced Multimodal Aspect-Based Sentiment Analysis by LLM-Generated Rationales. In *International Conference on Neural Information Processing*; Springer Nature: Singapore, 2025.
40. Xu, L.; Xie, H.; Qin, S.-Z.J.; Tao, X.; Wang, F.L. Parameter-Efficient Fine-Tuning Methods for Pretrained Language Models: A Critical Review and Assessment. *arXiv* **2023**, arXiv:2312.12148. [CrossRef]
41. Prottasha, N.J.; Chowdhury, U.R.; Mohanto, S.; Nuzhat, T.; Sami, A.A.; Ali, M.S.; Sobuj, M.S.I.; Raman, H.; Kowsher, M.; Garibay, O.O. PEFT A2Z: Parameter-Efficient Fine-Tuning Survey for Large Language and Vision Models. *arXiv* **2025**, arXiv:2504.14117.
42. Shankar Pandey, D.; Pyakurel, S.; Yu, Q. Be Confident in What You Know: Bayesian Parameter Efficient Fine-Tuning of Vision Foundation Models. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 44814–44844.
43. Liao, B.; Meng, Y.; Monz, C. Parameter-Efficient Fine-Tuning without Introducing New Latency. In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), Toronto, ON, Canada, 9–14 July 2023*; Association for Computational Linguistics: Kerrville, TX, USA, 2023; Long Papers; Volume 1.
44. Chen, K.; Pang, Y.; Yang, Z. Parameter-Efficient Fine-Tuning with Adapters. *arXiv* **2024**, arXiv:2405.05493.
45. Zhou, X.; He, J.; Ke, Y.; Zhu, G.; Gutiérrez-Basulto, V.; Pan, J.Z. An Empirical Study on Parameter-Efficient Fine-Tuning for MultiModal Large Language Models. In *Findings of the Association for Computational Linguistics: ACL 2024*; Association for Computational Linguistics: Kerrville, TX, USA, 2024.
46. Zhang, D.; Feng, T.; Xue, L.; Wang, Y.; Dong, Y.; Tang, J. Parameter-Efficient Fine-Tuning for Foundation Models. *arXiv* **2025**, arXiv:2501.13787.
47. Haque, S.; Eberhart, Z.; Bansal, A.; McMillan, C. Semantic Similarity Metrics for Evaluating Source Code Summarization. In *Proceedings of the IEEE International Conference on Program Comprehension, Virtual, 16–17 May 2022*; IEEE Computer Society: New York, NY, USA, 2022; Volume 2022, pp. 36–47.
48. Pontiki, M.; Papageorgiou, H.; Galanis, D.; Androutsopoulos, I.; Pavlopoulos, J.; Manandhar, S. SemEval-2014 Task 4: Aspect Based Sentiment Analysis. In *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval-2016), San Diego, CA, USA, 16–17 June 2016*; Association for Computational Linguistics: Kerrville, TX, USA, 2014.
49. Wang, B.; Liu, M. Deep Learning for Aspect-Based Sentiment Analysis. In *2021 International Conference on Machine Learning and Intelligent Systems Engineering (MLISE), Chongqing, China, 9–11 July 2021*; IEEE: New York, NY, USA, 2021.
50. Jayakody, D.; Isuranda, K.; Malkith, A.V.A.; de Silva, N.; Ponnamperuma, S.R.; Sandamali, G.G.N.; Sudheera, K.L.K. Aspect-Based Sentiment Analysis Techniques: A Comparative Study. In *2024 Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, 8–10 August 2024*; IEEE: New York, NY, USA, 2024. [CrossRef]
51. Li, X.; Bing, L.; Li, P.; Lam, W. A Unified Model for Opinion Target Extraction and Target Sentiment Prediction. *Proc. AAAI Conf. Artif. Intell.* **2018**, *33*, 6714–6721. [CrossRef]
52. Hu, M.; Peng, Y.; Huang, Z.; Li, D.; Lv, Y. Open-Domain Targeted Sentiment Analysis via Span-Based Extraction and Classification. *arXiv* **2019**, arXiv:1906.03820.
53. Chen, Z.; Qian, T. Relation-Aware Collaborative Learning for Unified Aspect-Based Sentiment Analysis. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, Online, 5–10 July 2020*; Association for Computational Linguistics: Kerrville, TX, USA, 2020.
54. Li, X.; Bing, L.; Zhang, W.; Lam, W. Exploiting BERT for End-to-End Aspect-Based Sentiment Analysis. *arXiv* **2019**, arXiv:1910.00883.
55. Luo, H.; Li, T.; Liu, B.; Zhang, J. DOER: Dual Cross-Shared RNN for Aspect Term-Polarity Co-Extraction. *arXiv* **2019**, arXiv:1906.01794.
56. He, R.; Lee, W.S.; Ng, H.T.; Dahlmeier, D. An Interactive Multi-Task Learning Network for End-to-End Aspect-Based Sentiment Analysis. *arXiv* **2019**, arXiv:1906.06906.
57. Bai, J.; Bai, S.; Chu, Y.; Cui, Z.; Dang, K.; Deng, X.; Fan, Y.; Ge, W.; Han, Y.; Huang, F.; et al. Qwen Technical Report. *arXiv* **2023**, arXiv:2309.16609. [CrossRef]

58. Albert, P.; Zhang, F.Z.; Saratchandran, H.; Rodriguez-Opazo, C.; van den Hengel, A.; Abbasnejad, E. RandLoRA: Full-Rank Parameter-Efficient Fine-Tuning of Large Models. *arXiv* **2025**, arXiv:2502.00987.
59. Li, Y.; Han, S.; Ji, S. VB-LoRA: Extreme Parameter Efficient Fine-Tuning with Vector Banks. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 16724–16751.
60. Tian, C.; Shi, Z.; Guo, Z.; Li, L.; Xu, C. HydraLoRA: An Asymmetric LoRA Architecture for Efficient Fine-Tuning. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 9565–9584.
61. Kim, D.; Lee, G.; Shim, K.; Shim, B. Preserving Pre-Trained Representation Space: On Effectiveness of Prefix-Tuning for Large Multi-Modal Models. *arXiv* **2024**, arXiv:2411.00029.
62. Hsu, C.-Y.; Tsai, Y.-L.; Lin, C.-H.; Chen, P.-Y.; Yu, C.-M.; Huang, C.-Y. Safe LoRA: The Silver Lining of Reducing Safety Risks When Fine-Tuning Large Language Models. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 65072–65094.
63. Qing, P.; Gao, C.; Zhou, Y.; Diao, X.; Yang, Y.; Vosoughi, S. AlphaLoRA: Assigning LoRA Experts Based on Layer Training Quality. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, Miami, FL, USA, 12–16 November 2024*; Association for Computational Linguistics: Kerrville, TX, USA, 2024.
64. Lin, Y.; Ma, X.; Chu, X.; Jin, Y.; Yang, Z.; Wang, Y.; Mei, H. LoRA Dropout as a Sparsity Regularizer for Overfitting Control. *arXiv* **2024**, arXiv:2404.09610. [CrossRef]
65. Prottasha, N.J.; Mahmud, A.; Sobuj, M.S.I.; Bhat, P.; Kowsher, M.; Yousefi, N.; Garibay, O.O. Parameter-Efficient Fine-Tuning of Large Language Models Using Semantic Knowledge Tuning. *Sci. Rep.* **2024**, *14*, 30667. [CrossRef]
66. Yang, A.; Yang, B.; Hui, B.; Zheng, B.; Yu, B.; Zhou, C.; Li, C.; Li, C.; Liu, D.; Huang, F.; et al. Qwen2 Technical Report. *arXiv* **2024**, arXiv:2407.10671.
67. Madhoushi, Z.; Razak, A.; Suhaila Zainudin, H. Aspect-Based Sentiment Analysis Methods In Recent Years. *Asia-Pacific J. Inf. Technol. Multimedia* **2019**, *7*, 79–96. [CrossRef]
68. Yang, A.; Yu, B.; Li, C.; Liu, D.; Huang, F.; Huang, H.; Jiang, J.; Tu, J.; Zhang, J.; Zhou, J.; et al. Qwen2.5-1M Technical Report. *arXiv* **2025**, arXiv:2501.15383. [CrossRef]