symmetry

MDPI

Article

Symmetry- and Asymmetry-Aware Domain Adaptation for Cross-Domain Sentiment Analysis

Chumsak Sibunruang [1], Jantima Polpinij [2,*], Manasawee Kaenampornpan [3], Thananchai Khamket [1], Jaturong Som-ard [4], Anirut Chottanom [1], Jatuphum Juanchaiyaphum [1], Vuttichai Vichianchai [1] and Bancha Luaphol [5]

[1] Department of Information Technology, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand; chumsak.s@msu.ac.th (C.S.); thananchai.k@msu.ac.th (T.K.); anirut@msu.ac.th (A.C.); jatuphum.j@msu.ac.th (J.J.); vuttichai.v@msu.ac.th (V.V.)
[2] Department of Computer Science, Faculty of Informatics, Mahasarakham University, Mahasarakham 44150, Thailand
[3] Department of Computer Engineering, Faculty of Engineering, Khon Kaen University, Khon Kaen 40002, Thailand; manaska@kku.ac.th
[4] Department of Geography, Faculty of Humanities and Social Sciences, Mahasarakham University, Mahasarakham 44150, Thailand; jaturong.s@msu.ac.th
[5] Department of Digital Technology, Faculty of Administrative Science, Kalasin University, Kalasin 46000, Thailand; bancha.lu@ksu.ac.th
* Correspondence: jantima.p@msu.ac.th; Tel.: +66-8-4326-0182

Abstract

Cross-domain sentiment analysis remains challenging due to distributional shifts and heterogeneous sentiment expressions across platforms. Existing domain adaptation approaches primarily focus on enforcing domain-invariant representations. However, such symmetry-preserving strategies often overlook directional and expression-level asymmetries. These asymmetries naturally arise in real-world sentiment data, particularly for context-inferred sentiment expressions. In this work, we propose a novel symmetry- and asymmetry-aware domain adaptation framework for cross-domain sentiment classification. The framework models symmetry through explicit multi-source distribution alignment, which captures transferable sentiment knowledge across domains. Additionally, aspect-level structural supervision organizes representations according to shared linguistic aspects. To address asymmetry, a directional divergence regularization is introduced. This component models expression-level and directional discrepancies between source and target domains. Importantly, the framework operates without requiring target-domain annotations. Experiments are conducted under a multi-source unsupervised domain adaptation setting using sentence-level hotel review datasets collected from multiple online platforms. Empirical results demonstrate strong performance for the proposed framework. It achieves an average Accuracy of 82.0% and Macro-F1 of 80.6%. The framework consistently and statistically significantly outperforms source-only, multi-source, and transformer-based adversarial adaptation baselines across all evaluated target domains ($p < 0.05$). Additional analyses confirm improved robustness to implicit sentiment expressions and platform-induced asymmetries. These findings highlight the importance of jointly modeling symmetry and asymmetry for robust cross-domain sentiment adaptation and provide a unified and deployable solution for sentiment analysis under realistic platform shifts.

Keywords: cross-domain sentiment analysis; domain adaptation; symmetry and asymmetry modeling; representation learning; implicit sentiment expression; multi-source learning

Check for updates

Academic Editors: Zhixun Su and Jie Yang

Received: 3 January 2026
Revised: 2 February 2026
Accepted: 11 February 2026
Published: 14 February 2026
Corrected: 9 September 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Sentiment analysis (SA) is a fundamental problem in natural language processing (NLP) and data mining. It is essential for various applications, including customer feedback analysis, social media monitoring, and decision support systems [1,2]. Over the past decade, sentiment classification has achieved enormous progress, driven by increasingly effective text representation methods. These methods have evolved from traditional feature engineering to distributed word embeddings and contextualized language models [3–5]. However, sentiment analysis in real-world settings faces the common issue of domain dependency. Models trained on one domain often fail to generalize to another domain due to disparities in vocabulary, linguistic style, contextual meaning, and sentiment expression patterns.

Consequently, domain adaptation has been one of the primary research topics in this field to address this challenge. Early work by Glorot et al. [6] showed that it is possible to learn domain-invariant representations with deep networks, which limits performance degradation across domains. Later studies investigated several approaches. These include robust adaptation with limited resources [7], multi-source domain adaptation [8], and multi-level contextual sentiment knowledge [9]. These works demonstrate the necessity of sharing transferable information across domains. Other researchers have explored mining domain-shared knowledge [10], projecting embeddings into shared latent spaces [11], and integrating syntactic structures into cross-domain sentiment models [12]. Despite these advances, domain shift still remains a significant challenge. This is particularly true when sentiment polarity changes, context meaning varies, or labeled data in the target domain are scarce [13].

From a machine learning perspective, cross-domain sentiment analysis naturally reflects the interplay between symmetric and asymmetric structures. In this context, symmetry refers to shared or invariant structures across domains. Examples include identical emotional expressions, universal sentiment patterns, and latent semantic representations that remain stable under domain transformations. In contrast, asymmetry represents domain-specific and/or directional gaps, such as vocabulary differences, contextual polarity shift, imbalanced class distributions, and non-reciprocal relationships between source and target domains. While many existing domain adaptation methods implicitly aim to achieve symmetry by minimizing domain discrepancy or enforcing representation alignment [6,11], they often fail to explicitly model expression-level and directional asymmetries. This limitation can reduce performance in realistic cross-domain sentiment analysis scenarios.

In this study, symmetry refers to domain-invariant distributional alignment. Under this approach, sentiment representations from different domains are encouraged to follow similar feature distributions. This assumes that sentiment semantics are shared across domains. This property is commonly operationalized through shared encoders and distribution alignment objectives, such as Maximum Mean Discrepancy (MMD) and related discrepancy minimization techniques.

In contrast, asymmetry denotes expression-level and directional discrepancies. These arise when sentiment is conveyed differently across domains, even when underlying sentiment polarity is preserved. Such asymmetries may originate from implicit sentiment expressions, domain-specific aspects, or contextual cues. These factors violate strict invariance assumptions. In this work, asymmetry is operationalized through dedicated learning components and loss terms that explicitly model non-invariant properties rather than treating them as noise.

Recent work on sentiment analysis and text classification also highlights this imbalance. Hybrid models and embedding-based sentiment classifiers have achieved better perfor-

mance by integrating different representations or leveraging contextual semantics [14–16]. Contrastive learning has also been proposed to capture implicit sentiment and aspect-level information [17,18]. Meanwhile, studies on text representation modeling and sentiment analysis report that there is no universal model or transfer adaptation method that performs well on various domains, tasks, or different data situations [2–4,19–21]. Our results indicate that successful cross-domain sentiment analysis not only relies on exploiting shared semantics but also requires modeling and adapting to domain-specific asymmetries.

Despite the rich literature on domain adaptation for sentiment analysis, existing approaches predominantly emphasize learning domain-invariant representations through distribution alignment or shared feature spaces. Symmetry-preserving strategies effectively transfer shared sentiment semantics, but they fail to explicitly model expression-level and directional discrepancies that naturally occur across domains. In particular, implicit sentiment expressions, domain-specific aspects, and contextual polarity shifts introduce asymmetries that cannot be fully addressed by invariance-based alignment alone. Several studies have improved adaptability through architectural or algorithmic enhancements. However, the interplay between symmetry-preserving alignment and asymmetry-aware modeling remains insufficiently explored in a unified framework.

To summarize, the remaining research gaps can be identified as follows:

- Most existing domain adaptation methods focus on symmetry-preserving domain-invariant alignment. Expression-level and directional asymmetries are not explicitly modeled.
- Implicit sentiment expressions and contextual polarity shifts introduce non-invariant behaviors. These cannot be fully captured by invariance-based alignment alone.
- The interaction between symmetry-aware alignment mechanisms and asymmetry-aware modeling components has not been systematically investigated, especially within a unified multi-source domain adaptation framework.

Consequently, it remains unclear how jointly modeling symmetry and asymmetry affects cross-domain performance. This is particularly relevant for robustness, generalization, and context-inferred (implicit) sentiment expressions.

Motivated by these gaps, this study addresses the following research questions:

1. How can symmetry-preserving domain-invariant alignment and asymmetry-aware modeling be jointly integrated into a unified multi-source domain adaptation framework for sentiment analysis?
2. To what extent does explicitly modeling expression-level and directional asymmetries improve cross-domain sentiment classification performance compared to symmetry-only adaptation approaches?
3. How do symmetry-aware and asymmetry-aware components contribute to cross-domain generalization, especially for context-inferred (implicit) sentiment expressions?

The aim of this work is to contribute to a more principled and interpretable understanding of domain adaptation in sentiment analysis through the explicit modeling of symmetry and asymmetry within the framework. Our contributions in this paper can be summarized as follows: (i) a conceptual contribution that reconceptualizes cross-domain sentiment analysis from the perspective of symmetry and asymmetry in data representations, (ii) a methodological contribution that integrates symmetry-aware and asymmetry-aware principles into domain adaptation; and (iii) an empirical contribution that demonstrates enhanced robustness and generalization across domains. Situated within the context of symmetry and asymmetry in machine learning and data mining, this study also serves as a

tentative exploration that combines theoretical analysis with practical approaches guided by empirical experiments.

The rest of the paper is structured as follows. Related works are discussed in Section 2. The proposed symmetry- and asymmetry-aware domain adaptation approach is described in Section 3. Experimental configurations and detailed analyses of results are presented in Section 4. Section 5 concludes the work with directions for future research.

2. Dataset Annotations

2.1. Data Sources and Collection

The data used in this work complements and consolidates data sources from prior studies on sentiment analysis. The dataset is designed to serve symmetry- and asymmetry-informed domain adaptation. It comprises customer review text from various online platforms. This provides diverse sentiments spanning different domains and writing conventions.

There are two dominant classes of data sources. The first consists of expert-annotated sentence-level hotel reviews, primarily gathered from TripAdvisor. Complete reviews are tokenized into separate sentences to facilitate fine-grained sentiment analysis. The second category includes multi-platform customer reviews derived from TripAdvisor, Booking.com, and Agoda. These platforms vary significantly in review length, writing style, and context emphasis. This variation offers a more realistic cross-domain adaptation setting.

Data are collected using the same pipeline. Raw reviews are filtered by removing duplicate and non-informative content. Only English texts are retained. Reviews are split into sentences, which serve as basic analysis units. This multi-source approach deliberately preserves platform variety in order to analyze both common sentiment trends and domain-specific biases.

2.2. Annotation Schema

To ensure consistency across all data sources, a unified annotation schema is applied to all datasets. Each sentence is annotated along three dimensions. The first dimension corresponds to the aspect category, such as staff service, cleanliness, value for money, and location convenience, representing shared and transferable aspects across customer reviews. The second dimension captures the expression type, including whether sentiment is expressed explicitly or implicitly. The third dimension denotes sentiment polarity (positive or negative); sentences with neutral polarity are excluded to maintain a clear separation between sentiment classes.

By applying a uniform annotation schema across platforms, sentences from different sources share the same label space. This design enables structural symmetry across domains. Simultaneously, it allows data-driven asymmetry to emerge through linguistic and contextual variations.

This study adopts binary sentiment classification to reduce task complexity and avoid confounding effects from fine-grained polarity distinctions. This choice allows the evaluation to focus specifically on domain shift and adaptation behavior rather than multi-class sentiment categorization difficulty. Such a setting is commonly used in domain adaptation studies to isolate representation transferability across domains [6,8,11].

Across the collected datasets, a substantial proportion of sentences express sentiment implicitly rather than through explicit sentiment-bearing words. According to expert annotations, approximately 35% of sentences require context-inferred sentiment interpretation, while the remaining 65% contain explicit sentiment expressions. This distribution aligns with prior findings in sentiment analysis literature. These studies report that implicit

sentiment is prevalent in real-world user-generated content, particularly in review and opinion domains [2,17,18].

2.3. Label Construction and Quality Control

The sentence-level annotations are conducted by two or more sentiment analysis experts. Each sentence is individually annotated, and inconsistencies are resolved through expert discussion to reach a consensus. This process is especially relevant for implicit sentiment expressions, which often lack explicit lexical indicators.

For multi-platform data, we also pay attention to inter-platform consistency. Annotators are trained to adhere to the meaning expressed rather than the surface form, and regular calibration is performed on the annotation guidelines. Label distributions (histograms) are observed to minimize the potential for extreme class imbalance, while still keeping the data appropriate for comparison between supervised learning and domain adaptation.

Inter-annotator agreement was calculated to evaluate annotation stability. This analysis used an independent, randomly selected portion of data before consensus. Cohen's κ was calculated for sentiment polarity and expression-type annotations. The agreement scores reflect high inter-annotator reliability. The k value was 0.78 for polarity and 0.71 for explicit versus implicit expression classification. These findings indicate that despite the subjectivity inherent in implicit sentiment inference, the annotation process produces reliable and consistent labels.

Detailed annotation guidelines, including aspect definitions and criteria for identifying implicit sentiment expressions, are available from the corresponding author upon reasonable request.

2.4. Multi-Source Mapping for Domain Adaptation

The annotated data was then split into source and target domains, which vary based on the platform of origin, quantity of the data, and style. Platforms with sufficient labeled data serve as source domains. In contrast, platforms with a lower portion of labeled training data, a higher proportion of implicit expressions, or totally different writing styles are regarded as target platforms.

These factors create a situation where label-level symmetry is maintained through the common annotation schema, but distributional and representational asymmetry occur due to stylistic differences across platforms, review length, and expression patterns. This setting is more realistic for cross-domain sentiment analysis scenarios and serves as an effective benchmark for symmetry- versus asymmetry-based adaptation approaches.

In this study, we follow a symmetric-asymmetric approach. We enforce symmetry by learning domain-invariant representations, while leveraging the expressive power of asymmetry by explicitly modeling expression-specific differences in emotion distributions between domains. This dual approach cannot be aligned through invariance only.

Table 1 provides an overview of the two primary classes of data sources that we utilized in this study. It illustrates how label-level symmetry is preserved across a unified annotation schema, while distributional and stylistic asymmetry emerge between platforms.

The datasets are structured into two main classes of sources. The first consists of expert-annotated, sentence-level hotel reviews, which provide high-quality and reliable sentiment labels. The second comprises multi-platform customer reviews collected from different online platforms, which introduce stylistic variations and distributional discrepancies across sources. Within this framework, sentiment expressions are further categorized into two types: Direct (explicit aspect mentions) and Context-inferred types,

where sentiment polarity is inferred from contextual information rather than explicit sentiment-bearing words.

Table 1. Overview of the two dominant classes of data sources used in this study.

Data Source Class	Expert-annotated hotel reviews	Multi-platform customer reviews	Total (All sources)
Platform/Origin	TripAdvisor	TripAdvisor, Booking.com, Agoda	Multi-source
Analysis Unit	Sentence	Sentence	Sentence
No. of Sentences	2000	6000	8000
Aspect Categories	Staff service, Cleanliness, Value for money, Location convenience	Staff service, Cleanliness, Value for money, Location convenience	Unified schema (4 aspects)
Sentiment Polarity	Positive/Negative	Positive/Negative	Binary
Sentiment Expression	Direct/Context-inferred	Direct/Context-inferred	Direct/Context-inferred

Table 1 summarizes the dataset construction, where label-level symmetry is preserved through a unified annotation schema while realistic distributional and stylistic asymmetries naturally arise across platforms. This section proposes a symmetry- and asymmetry-aware domain adaptation approach for cross-domain sentiment classification.

All datasets are drawn from the hotel review domain in order to control topical variability and isolate the effects of platform-induced domain shift. By fixing the topical domain, differences across datasets primarily stem from platform-specific characteristics, including writing styles, review length, contextual emphasis, and sentiment expression patterns. This approach eliminates semantic differences between unrelated domains. This design enables more controlled evaluation of symmetry-preserving alignment and asymmetry-aware modeling under realistic platform shifts.

In order to provide finer granularity than the high-level source categorization in Table 1, Table 2 reports the distribution of explicit and implicit sentiment expressions for each platform. This platform-based breakdown characterizes how sentiment expression patterns vary across sources in greater detail. This analysis is particularly relevant because the present study focuses on implicit sentiment and platform-induced asymmetries.

Table 2. Distribution of Explicit and Implicit Sentiment Expressions per Platform.

Platform	Number of Sentences	Explicit (%)	Implicit (%)
TripAdvisor	3200	66.4	33.6
Booking.com	2850	70.2	29.8
Agoda	2450	61.7	38.3

Overall, the dataset design described in Table 1 and platform-level statistics reported in Table 2 illustrate a constrained yet realistic cross-domain setting. In this setting, label-level symmetry is maintained, even as platform-induced asymmetries naturally surface. This dataset setting directly motivates our symmetry-and asymmetry-aware domain adaptation method.

2.5. Recent Advances Beyond Symmetry-Only Domain Adaptation

Recent progress in sentiment analysis and representation learning has focused on robustness under complex and heterogeneous data conditions. Contemporary studies explore multimodal fusion, ensemble learning, and structured representation modeling to better capture implicit affective cues and contextual variations. For example, FPA-GCN [22] enhances aspect sentiment triplet extraction (ASTE) through prediction-aware graph convolution. It emphasizes fine-grained structural modeling at the aspect level.

However, despite their effectiveness for within-domain sentiment understanding, these approaches do not explicitly address cross-domain distribution shift.

Ensemble models like MLM-EOE [23] provide enhanced robustness by using multi-expert networks to address noisy and emotion-dependent cues. Wang et al. [24] combine multimodal spatio-temporal features for automatic depression recognition. Recent multimodal sentiment models (e.g., 3WD-DRT [25]) adopt asymmetric decision strategies and cost-sensitive objectives to balance diverse modalities and uneven error risks. These approaches illustrate that asymmetric modeling is useful when sentiment signals are skewed or vague. However, they are usually designed for a single-domain setting or multimodal project, but not for domain adaptation.

In addition to sentiment analysis, related work in affective computing and representation disentanglement has emphasized separating invariant factors from variant ones. For example, TalkingStyle [26] disentangles identity-invariant and style-specific factors by learning a latent style space. Wu et al. [27] proposed Happiness Prediction with Domain Knowledge Integration, which incorporates domain constraints to ensure consistency in explanations across predictive factors. Although these studies are not intended for sentiment domain adaptation, they support the view that rigid invariance constraints alone are insufficient for modeling real-world affective data.

In general, although much research has succeeded in robustness [8,11], structure modeling [18], and task-specific performance–especially using multimodal or ensemble-based architectures [27] or transformer-based dynamic context modeling [28], existing methods primarily focus on either representation learning or downstream task optimization. However, the coexistence of domain-invariant (symmetric) and domain-specific (asymmetric) properties is rarely modeled explicitly in cross-domain sentiment adaptation. Few works consider the interaction between symmetry-preserving alignment and asymmetry-aware modeling. This gap motivates our investigation of the symmetry- and asymmetry-adaptation model.

To clarify the positioning of recent studies and their relationship to our work, Table 3 provides an analytical categorization of representative methods recommended by the reviewers. The table offers an at-a-glance comparison of how these studies address sentiment representation robustness, asymmetry consideration, and contextual variability. It also shows where each approach falls short for cross-domain sentiment adaptation. This comparison demonstrates that, although recent studies have incorporated certain aspects of robustness, ensemble modeling, or asymmetric decision mechanisms, none of them explicitly and jointly model both symmetry and asymmetry within a multi-source cross-domain sentiment adaptation framework.

Table 3. Analytical grouping of recent studies with respect to cross-domain adaptation, symmetry modeling, and asymmetry handling.

Study	Main Relevance	Key Limitation
FPA-GCN [22]	Aspect-level structural modeling for sentiment representation	Not designed for the domain adaptation
MLM-EOE [23]	Ensemble robustness and handling noisy or implicit affective signals	Not cross-domain; no symmetry modeling
3WD-DRT [25]	Asymmetric decision strategies and cost-sensitive sentiment modeling	Multimodal focus; not textual domain adaptation
TalkingStyle [26]	Explicit disentanglement of invariant and variant factors	Not related to NLP or sentiment analysis
Happiness Prediction [27]	Integration of domain knowledge and explanation consistency	Not a sentiment classification task
DPA-MVSNet [28]	Dynamic context modeling under distribution shift	Computer vision domain

Table 3 provides a concise overview of recent representative studies suggested by the reviewers, focusing on their modeling schemes and limitations regarding cross-domain sentiment analysis. Rather than performance metrics, this table illustrates the specific problems addressed by each approach (e.g., aspect structure, ensemble robustness, or asymmetric decision modeling) and highlights limitations in the context of cross-domain adaptation. By clearly stating these boundaries, Table 2 provides a guideline for distinguishing robustness-driven or task-specific modeling from cross-domain sentiment adaptation requirements.

A broader conceptual pattern emerges from the analytical grouping presented in Table 3. Although recent research has increasingly emphasized the need for robust, generalizable, and asymmetric modeling, these contributions are typically developed in isolation. They are largely focused on single-domain or multimodal settings. Prior cross-domain sentiment studies have not considered the joint optimization of domain-invariant (symmetric) alignment and domain- or expression-specific (asymmetric) discrepancies. Table 4 provides a qualitative comparison of recent sentiment models across key dimensions, including cross-domain adaptation, symmetry modeling, and asymmetry handling. This comparison clarifies the distinction and situates the proposed method within the context of prior work.

Table 4. Conceptual Positioning of Recent Sentiment Models.

Method	Cross-Domain	Symmetry Modeling	Asymmetry Modeling	Multi-Source	Primary Focus
FPA-GCN	×	×	×	×	Aspect-level sentiment extraction
MLM-EOE	×	×	✓ (implicit)	×	Ensemble robustness
3WD-DRT	×	×	✓ (decision-level)	×	Cost-sensitive multimodal sentiment
TalkingStyle	×	✓ (implicit)	✓ (implicit)	×	Disentanglement of invariant vs. variant factors
Happiness Prediction	×	×	✓ (knowledge-driven)	×	Explanation consistency
DPA-MVSNet	×	×	✓ (context-driven)	×	Dynamic context modeling under distribution shift
Proposed	✓	✓ (explicit)	✓ (explicit)	✓	Cross-domain robustness

A check mark (✓) denotes that the method supports or incorporates the specified capability, while a cross mark (×) indicates that it does not.

Table 4 positions recent sentiment-related approaches according to their modeling scope and adaptation capabilities. While prior studies emphasize aspect structure, ensemble robustness, or asymmetric decision mechanisms, they are predominantly designed for within-domain or multimodal scenarios. In contrast, our proposed framework explicitly integrates symmetry-preserving domain alignment and asymmetry-aware modeling within a multi-source cross-domain sentiment adaptation setting.

The conceptual positioning summarized in Table 4 demonstrates that recent sentiment-related studies address robustness, contextual variability, or asymmetric modeling from different perspectives, yet largely in isolation. While these approaches contribute valuable insights, they do not provide a unified solution for cross-domain sentiment adaptation. In particular, the explicit and joint treatment of symmetry (domain-invariant alignment) and asymmetry (domain- or expression-specific discrepancies) remains absent from existing approaches.

To clarify how the proposed framework advances beyond these closely related approaches, Table 5 presents a focused comparison of its novelty along key dimensions relevant to symmetry- and asymmetry-aware cross-domain sentiment adaptation.

Table 5. Novelty Comparison with Recent Representative Works.

Dimension	FPA-GCN	MLM-EOE	3WD-DRT	Proposed
Cross-domain adaptation	×	×	×	✓
Explicit symmetry modeling	×	×	×	✓
Explicit asymmetry modeling	×	✓ (implicit)	✓ (decision-level)	✓ (representation-level)
Multi-source learning	×	×	×	✓
Implicit sentiment robustness	✓	✓	✓	✓
Unified symmetry–asymmetry framework	×	×	×	✓

A check mark (✓) denotes that the method supports or incorporates the specified capability, while a cross mark (×) indicates that it does not.

Table 5 focuses on recent sentiment modeling approaches that are most closely related to cross-domain sentiment classification. Unlike Table 3, which provides a conceptual overview, this table highlights the specific novelty of the proposed framework in jointly modeling symmetry and asymmetry within a multi-source adaptation setting.

3. The Proposed Method

Our novel framework addresses cross-domain sentiment classification under real-world platform shifts. It does this by formulating the problem as multi-source domain adaptation. In this scenario, we have sentences with known authorship from numerous source domains $\left\{D_s^{(k)}\right\}_{k=1}^{K}$. The test target domain D_t exhibits different language distributions due to platform characteristics (such as writing style, review length, and context emphasis). Such distributional discrepancy is a common source of performance loss when a model learned on source domains is directly deployed to new target domains [6,7,13,21]. To address this problem, the proposed approach is divided into four main components: (i) sentence-level representation learning, (ii) aspect-level structure induction (to capture common semantics), (iii) explicit multi-source alignment to obtain domain-invariant representations, and (iv) asymmetry-aware learning to handle directionality and expression-level disagreement. The core idea is to exploit shared, invariant sentiment structure (symmetry). At the same time, we explicitly model domain- and expression-specific irregularities caused by noise (asymmetry). This leads to more robust performance in cross-domain adaptation [8–11,19–21].

3.1. Preprocessing and Sentence-Level Representations

Given a review document from any outlet, the text undergoes standard preprocessing, including tokenization, lowercasing, punctuation normalization, and stop-word removal. Sentence-level modeling is adopted because sentiment cues are usually localized. Sentence granularity also helps alleviate noise from long and heterogeneous reviews, which aligns with the common practice of sentiment analysis [1,2]. A sentence is represented as $x = [w_1, w_2, \ldots, w_n]$.

The framework builds representations that are comparable across domains using embedding-based sentence encodings. These encodings are highly effective for text classification and transfer because they capture semantic similarity beyond simple lexical overlap [3,4]. A combinatory approach is used to incorporate lexical salience and semantic generalization. Firstly, we calculate TF–IDF weights to emphasize the informative tokens:

$$tf(w, x) = \log\left(1 + freq(w \in x)\right) \quad (1)$$

$$idf(w) = log\left(1 + \frac{N}{df(w)}\right) \tag{2}$$

$$TF - IDF(w, x) = tf(w, x) \times idf(w) \tag{3}$$

where N is the number of sentences and $df(w)$ is the number of occurrences of token w. The sentence presentation is then computed as its TF–IDF weighted embedding as follows.

$$z(x) = \frac{\sum_{w \in x} TF - IDF(w, s) \cdot e(w)}{\sum_{w \in x} TF - IDF(w, x)} \tag{4}$$

where $e(w) \in \mathbb{R}^d$ is a pre-trained embedding vector. The hybrid feature representation retains semantic closeness between domains and mitigates platform-specific verbosity [1,3,4].

When stronger contextual modeling is desired—for example, in context-inferred sentiment where polarity must be expressed without explicit aspect terms—the same pipeline can be instantiated using a contextual encoder. In that case, the sentence representation is computed from a pooled Transformer representation:

$$z(x) = f_\theta(x)_{[CLS]} \tag{5}$$

where $x = [w_1, w_2, \ldots, w_n]$ denotes an input sentence obtained after tokenization and standard preprocessing. Sentences are encoded by using a pre-trained contextual text encoder $f_\theta(.)$ (e.g., BERT, RoBERTa or XLNet) parameterized by θ. The input sequence is mapped into contextualized hidden states by the encoder, which models both local and global dependencies between tokens. The parameter set θ, which comprises all learnable components in the encoder (e.g., embedding layers and Transformer layers), is shared across the source and target domains to guide domain-invariant representation learning.

In Transformer architectures, a special token [CLS] is added to the start of the input sequence. The hidden vector corresponding to this token in the last encoder layer, denoted as f_θ (x)$_{[CLS]}$, is commonly used as the hidden representation of the entire sentence. Thus, the sentence-level representation $z(x)$ is derived from this [CLS] vector, which is then fed into subsequent modules, including the sentiment classifier, the aspect prediction module, and the domain alignment mechanism.

Contextual encoders consistently outperform their non-contextual versions in terms of text representation quality and classification performance. They also outperform classical baselines, particularly on tasks involving implicit or distributed sentiment cues [5,16,29–33]

3.2. Aspect-Level Structure Induction for Symmetric Semantic Organization

Although sentence embeddings are designed to offer a generic semantic space, cross-domain sentiment classification benefits from maintaining a common aspect-level structure. User reviews across domains often focus on the same criteria, even if the mode of expression differs. Aspect-based sentiment analysis has demonstrated that modeling aspect structure improves interpretability and helps avoid confounding signals [19,20]. Thus, the proposed structure is symmetric in that it organizes representations based on the four shared topic types (staff service, cleanliness, value for money and location convenience).

When the annotated data contain aspect labels, an auxiliary aspect prediction head is introduced:

$$\hat{P}(a|x) \ = softmax(W_a z(x) + b_a), \mathcal{L}_{asp} = CE(a, \ \hat{P}Za|x) \tag{6}$$

Here, a is the aspect label of the input sentence x. Given the sentence representation $z(x)$, the auxiliary aspect prediction head infers the distribution $\hat{P}(a|x)$ over possible aspects

via a softmax function with W_a and b_a as parameters. The aspect prediction loss $\mathcal{L}_{asp}$ is the cross-entropy between ground-truth and the predicted distributions. This auxiliary goal encourages the learned representations to retain aspect-discriminative semantics, which helps maintain a consistent aspect-level organization across domains [9,10,19].

For data with rare aspect supervision, or to improve the aspect structure of context-inferred expressions, we further perform clustering directly over the embedding space using K-means with $K = 4$.

$$\min_{\{\mu_k\}_{k=1}^{4}} \sum_i \min_k \|z(x_i) - \mu_k\|_2^2,\ \hat{a}_i = arg\ \min_k \|z(x_i) - \mu_k\|_2^2 \tag{7}$$

The goal of the above formulation is clustering-based aspect induction. Here, $z(x_i)$ demotes the sentence-level representation of i-th sentence, and $\{\mu_k\}_{k=1}^{4}$ is the set of cluster centers, each representing one aspect category. The optimization objective minimizes the sum of squared Euclidean distances from each sentence representation to its nearest centroid, thereby encouraging semantic grouping. The pseudo-aspect label $\hat{a}_i$ assigned to sentence x_i is obtained by choosing the cluster whose centroid μ_k is nearest to $z(x_i)$. This clustering procedure allows us to aggregate sentence representations into a semantically coherent aspect-level structure, particularly for sentence where aspect information is not explicitly expressed.

3.3. Multi-Source Domain Alignment (Core Domain Adaptation Step)

Explicit domain adaptation occurs during this stage. Even with powerful encoders, representation distributions across different platforms are not well aligned due to domain shift, which can affect generalization. Classic and recent cross-domain sentiment classification studies have indicated the importance of mitigating such mismatch through adaptation approaches, especially when target labels are limited or unavailable [6–8,11,13,21]. We aim to learn domain-invariant representations and simultaneously align the latent distributions of each source domain with the target data.

Let $D_s^{(k)}$ represent the labeled source-domain data k, and let D_t denote the target domain. Let P_k and Q be the distributions of sentence representations $z(x)$ induced by $D_s^{(k)}$ and D_t, respectively. In order to promote representation-level symmetry across domains, we minimizes a multi-source Maximum Mean Discrepancy (MMD) loss:

$$MMD^2(P_k, Q) = \left\|\mathbb{E}_{z\sim P_k}[\phi(z)] - \mathbb{E}_{z\sim Q}[\phi(z)]\right\|_2^2, \tag{8}$$

$$\mathcal{L}_{align} = \sum_{k=1}^{K} MMD^2(P_k, Q) \tag{9}$$

where P_k is the distribution of sentence representations z in k-th source domain, and Q is the distribution of sentence representations from the target domain. Here $\phi(.)$ is an implicit feature mapping determined by a kernel function, enabling distribution comparison in high-dimensional space. The expectation $\mathbb{E}[\cdot]$ is taken over each distribution. The MMD measures the distance between the source and target distributions in the latent space, with smaller values indicating better alignment. The global alignment loss $\mathcal{L}_{align}$ is obtained by calculating the summation of distances between all K source domains and the target domain using MMD. This further encourages the encoder to learn domain-invariant representations that are equally shared across multiple sources and the target.

The global alignment loss $\mathcal{L}_{align}$ is computed as the sum of MMD distances between all K source domains and the target domain. This forces the encoder to acquire domain-invariant features that are learned across multiple sources and the target domain [9,10,19].

3.4. Asymmetry-Aware Learning for Domain-Specific and Expression-Level Discrepancies

Alignment alone does not eliminate all cross-domain errors because real data contain asymmetries beyond simple distributional differences. In this study, two specific types of asymmetries are especially relevant: (i) directionality of the source mixture versus the target shift, and (ii) variations in sentiment expression modality between direct and context-inferred sentiments. Previous work has focused on the resistance of cross-domain sentiment analysis in such settings, especially when little or no target supervision is available [2,7,21]. Our new framework therefore adds asymmetry-aware components to the training objective.

A sentiment classifier predicts polarity from $z(x)$:

$$P(y|x) = softmax\left(W_y z(x) + b_y\right),\ y\ \in \{0,\ 1\} \tag{10}$$

Let y be the sentiment label of the input sentence x, where $y \in \{0, 1\}$ denotes negative and positive sentiment class labels. Given a sentence-level representation $z(x)$, the sentiment classification layer calculates the conditional probability distribution $P(y|x)$ by using softmax with the parameterized weight matrix W_y and bias vector b_y. The parameters W_y and b_y map the latent representation into the sentiment label space with class probabilities, which are employed for predicting the sentiment.

The supervised learning signal for labeled sources is obtained by cross-entropy. In order to mitigate the bias toward simpler direct expressions and improve performance on the more difficult context-inferred cases, which are the cases where "everyone agrees this is difficult". This is the key challenge of sentiment analysis and aspects-based settings. We reweight our samples by expression type:

$$w_i \begin{cases} \alpha, ti = \text{context-inferred} \\ 1, ti = \text{direct} \end{cases} (\alpha > 1)\ \mathcal{L}_{cls} = \sum_{(x_i, y_i) \in D_s} w_i CE(y_i, P(y|x_i)) \tag{11}$$

Here, $\mathcal{L}_{cls}$ is the weighted cross-entropy loss over the labeled source dataset D_s, where each (x_i, y_i) contributes with weight w_i. This loss is the main supervisory signal for training the sentiment classifier.

This architecture explicitly considers expression-level asymmetry, which is closely related to implicit emotion comprehension [2,19,20]. Recent methods enhance the implicit sentiment learning with supervised contrastive pre-training argue for separation between representation for explicit and implicit cues. Such a term can be added if necessary:

$$\mathcal{L}_{supcon} = \sum_i \frac{-1}{|P(i)|} \sum_{p \in P(i)} log \frac{\exp\left(sim(z_i, z_p)/\tau\right)}{\sum_{a \in A(i)} \exp\left(sim(z_i, z_a)/\tau\right)} \tag{12}$$

Here, $P(i)$ denotes the set of positive sample sharing the same label, and $A(i)$ is the set of all samples in the current batch excluding i [17,18]. τ is a temperature parameter that controls the sharpness of the distribution. The supervised contrastive loss $\mathcal{L}_{supcon}$ encourages representations with the same label to be closer together while pushing apart those with different labels in the latent space.

To model discrepancies of direction (i.e., asymmetry from source to target), we introduce a regularization term based on the Kullback–Leibler (KL) divergence between predictive marginals induced by the source and the target. Let

$$\overline{P}_t = \frac{1}{|D_t|} \sum_{x \in D_t} P(y|x),\ \overline{P}_s = \frac{1}{|D_s|} \sum_{x \in D_t} P(y|x) \tag{13}$$

A directional divergence is then minimized:

$$\mathcal{L}_{dir} = KL\left(\overline{P}_t \middle|\middle| \overline{P}_s\right), \tag{14}$$

This formulation does not assume reciprocity in the source mixture. This kind of directional robustness is intuitively aligned with focus on cross-domain robustness under distribution shift [2,33].

Finally, all components are trained jointly in an end-to-end optimization:

$$\mathcal{L} = \mathcal{L}_{cls} + \beta\mathcal{L}_{asp} + \eta\mathcal{L}_{supcon} + \lambda\mathcal{L}_{align} + \gamma\mathcal{L}_{dir} \quad (15)$$

Here $\mathcal{L}$ is the total training objective with multiple loss terms. $\mathcal{L}_{cls}$ is the sentiment classification loss, and $\mathcal{L}_{asp}$ enforces a spectral aspect-aware constraint. The supervised contrastive loss $\mathcal{L}_{supcon}$ promotes discriminative clustering of representations by sentiment labels, while $\mathcal{L}_{align}$ maintains domain-invariant information by minimizing the discrepancy between source and target distributions. The directional loss $\mathcal{L}_{dir}$ explicitly models asymmetric relationships between domains or expression types. The scalar coefficients β, η, λ, and γ are hyperparameters that balance the relative importance of aspect supervision, contrastive learning, domain alignment, and asymmetry modeling during training [6–13].

3.5. Training Protocol

The presented framework is trained in an end-to-end manner by jointly exploiting labeled data from multiple source domains and unlabeled data from the target domain. At each training iteration, mini-batches are sampled from all source domains in a balanced manner to prevent any single platform from dominating, and these are combined with a mini-batch from the target domain. All samples pass through a common sentence encoder. Parameters are shared across domains, which facilitates the learning of cross-domain transferable representations.

Supervised sentiment learning is conducted on labeled source data. It uses an expression-sensitive weighting approach to promote context-inferred sentiment expressions over direct ones. This architecture minimizes the bias toward salient sentiment-related clues and improves robustness against polarity inference from contexts [34,35]. Furthermore, we employ an auxiliary aspect-prediction objective to retain consistent aspect-level semantics across the domains when aspect annotations are available.

Domain adaptation is explicitly pursued via a multi-source alignment objective, which enhances the correspondence between latent source representations and the target distribution. This alignment is directly optimized together with the sentiment classification task, allowing the encoder to simultaneously capture domain-invariant information under the sentiment's supervision [36]. Moreover, in order to tackle residual directional domain shift, an extra regularization term penalizes the discrepancy between source-induced and target-induced predictive distributions.

The global training objective is updated with a gradient-based optimizer, such as Adam. Hyperparameters that control the relative importance of sentiment supervision, aspect structure, domain alignment, and asymmetry-aware regularization are chosen based on validation using source data or held-out development data. This protocol ensures robust convergence and balances symmetry preservation with asymmetry modeling in the cross-domain setting [37,38].

3.6. Theoretical Discussion on Directional Divergence and Adaptation Risk

Although the proposed directional divergence term is introduced as an auxiliary regularization objective, it is closely related to divergence-based domain adaptation theory. Classical adaptation bounds suggest that the expected target-domain risk can be upper-bounded by the source-domain risk. The bounds also include a discrepancy measure between source and target distributions, as well as a shared hypothesis complexity term.

In this context, minimizing a suitable divergence between source and target distributions serves as a surrogate for reducing the target-domain error under distribution shift.

Unlike symmetric discrepancy measures commonly adopted in domain-invariant representation learning, cross-domain sentiment transfer inherently exhibits directional characteristics, where knowledge is transferred from labeled source domains to an unlabeled target domain [39,40]. As a result, symmetric alignment alone may be insufficient to capture non-reciprocal discrepancies arising from domain-specific expressions or platform-induced linguistic variations [41]. The Kullback–Leibler (KL) divergence provides a natural asymmetric measure for modeling such directional discrepancies, as it penalizes mismatches between predictive distributions in a non-symmetric manner, which aligns with the source-to-target adaptation setting considered in this study [42].

From a risk minimization perspective, the proposed KL-based directional divergence can be interpreted as a regularizer that constrains the discrepancy between source-induced and target-induced predictive marginals [43,44]. By encouraging consistency between these predictive distributions, the model implicitly reduces the uncertainty of target-domain predictions, thereby making the divergence term a surrogate objective for controlling target-domain risk. Similar divergence-based regularization strategies have been widely employed in prior domain adaptation and representation learning frameworks, in which measures such as MMD, Jensen–Shannon divergence, or Wasserstein distance are optimized without requiring explicit guarantees on exact risk minimization.

Regarding optimization stability, the directional divergence term is formulated as a differentiable and bounded regularization loss that is jointly optimized with the primary classification objective. Since the KL divergence is applied to normalized predictive distributions and weighted as an auxiliary term, it does not dominate the training dynamics but instead complements the symmetry-preserving alignment objective. In practice, this joint optimization strategy follows common design principles in multi-objective domain adaptation frameworks and does not introduce additional instability beyond that of standard divergence-based regularization methods.

4. Evaluation

This section examines the proposed framework in more realistic cross-domain sentiment classification scenarios. The benchmark evaluates performance not only in predictiveness but also regarding the ability to handle domain shift and the effectiveness of multi-source adaptation, as well as the contribution of symmetry- and asymmetry-aware factors. In order to make the comparison fair and stable, various evaluation metrics and complementary experimental analysis are adopted.

4.1. Experimental Setup and Metrics

In order to evaluate the proposed symmetry- and asymmetry-aware framework, we perform experiments under a cross-domain sentiment classification scenario that has been used in related works of sentiment domain adaptation [1–3]. The training data are labeled reviews from the target domain identified in multiple preferred domains, and the target domain has only the unlabeled data for training, which will be used solely for evaluation. We note that this setting is representative of common practical domain adaptation scenarios, where annotations exist for an established data source (e.g., an old platform) but are scarce or absent for a newer or emerging domain [6,8,12].

The models are trained with a shared sentence encoder to enable parameter sharing across domains. This approach is commonly used to facilitate transferable representations in cross-domain sentiment analysis [6,7,11]. During training, mini-batches are formed by sampling data from all sources in a balanced manner, combined with samples from the

target domain. This strategy aligns with multi-source adaptation trends [3,4]. We select the model and tune hyperparameters using a held-out validation set derived solely from the source domains. Specifically, we do not access target-domain labels to prevent information leakage and maintain a fair evaluation protocol [6,13].

For performance measurement, we evaluate the sentiment classification task using Accuracy and Macro-averaged F1-score (Macro-F1). These are well-established metrics for sentiment analysis and domain adaptation work [1,7,11]. Accuracy is a straightforward measure of general correctness. However, we prioritize Macro-F1 because it is less sensitive to class imbalance and better reflects performance stability across sentiment classes, which is especially valuable in cross-domain scenarios [1,2]. Moreover, to evaluate LPRS's robustness to expression-level asymmetry, we report separate results for sentences dominated by direct (explicit) sentiment expressions and for those dominated by context-dependent (implicit) sentiment expressions. Recent works in sentiment analysis have emphasized the importance of distinguishing these expression types [2,17,18,20].

To ensure a fair comparison, all results are averaged over multiple runs initialized with different random seeds. Our proposed method is compared with several state-of-the-art models, including: (i) source-only training without adaptation and (ii) multi-source training without explicit alignment. Source-only training without adaptation and multi-source training without explicit alignment are widely used baseline models in cross-domain sentiment classification [6,8,11,13]. This comparison demonstrates the role of domain adaptation and our asymmetry-aware components in enhancing cross-domain generalization.

To assess the robustness of the observed performance gains, all experiments were repeated five times using different random seeds. We conducted paired two-tailed t-tests comparing the proposed method and the strongest baseline for both Accuracy and Macro-F1 scores. Improvements achieved by the proposed framework are statistically significant ($p < 0.05$) across all evaluated target domains.

4.2. Cross-Domain Results

In the following section, we describe our empirical results and analyze them from different angles to assess the effectiveness of our approach. The evaluation comprises four complementary experiments, each focusing on specific properties of cross-domain sentiment analysis. These properties include general classification performance, domain-generalization ability, robustness against expression asymmetry, and an analysis of individual model components.

4.2.1. Overall Sentiment Classification Performance

The first set of experiments measures the overall sentiment classification performance in cross-domain cases. The objective is to determine whether this framework can leverage multiple source domains to build strong predictive models that can be transferred to a held-out target domain.

Table 6 shows the classification performance in terms of Accuracy and Macro-F1. We compare the proposed symmetry- and asymmetry-aware framework against source-only training, single-source adaptation, and multi-source baselines.

Table 6 presents several training configurations that progressively incorporate symmetry and asymmetry modeling. These configurations range from source-only learning without domain adaptation to the proposed framework, which jointly models domain-invariant symmetry and domain-specific asymmetry. As shown in the table, source-only training exhibits notable performance degradation when applied to an out-of-domain target, highlighting the impact of domain shift.

Introducing multiple source domains improves generalization performance by reinforcing shared (symmetric) sentiment patterns across domains. Further performance gains are observed when explicit domain alignment is applied, indicating the benefit of enforcing domain-invariant representations. However, these improvements remain limited when only symmetry-preserving mechanisms are considered.

Table 6. Cross-Domain Sentiment Classification Results.

Model Configuration	Adaptation Mechanism	Accuracy (%)	Macro-F1 (%)
Source-only Training	No symmetry modeling	74.2 ± 0.6	71.8 ± 0.7
Single-Source Adaptation	Limited symmetry across domains	76.5 ± 0.5	74.1 ± 0.6
Multi-Source Training	Implicit symmetry via data aggregation	78.3 ± 0.5	76.0 ± 0.6
Multi-Source + Domain Alignment	Explicit symmetry enforcement	80.1 ± 0.4	78.4 ± 0.5
Proposed Framework	Symmetry + Asymmetry modeling	82.0 ± 0.3	80.6 ± 0.4

In contrast, the proposed framework, which integrates both symmetry and asymmetry modeling, achieves the best performance across both Accuracy and Macro-F1 metrics. Importantly, the reported improvements are consistent across multiple runs with relatively low variance. This suggests stable and robust behavior under different random initializations. These results support the hypothesis that modeling expression-level and domain-specific asymmetries complements symmetric alignment. Consequently, this approach leads to more reliable cross-domain sentiment classification performance.

4.2.2. Cross-Domain Generalization Under Domain Shift

This experiment investigates cross-domain generalization under domain shift by comparing model performance across various target domains. All models are trained on the same source domains, and evaluation is performed independently for each unseen target domain.

As presented in Table 7, the source-only trained system exhibits noticeable performance variability across different target domains. This variability indicates limited robustness to domain shift. Introducing multiple source domains consistently improves performance. This supports the intuition that exposure to more diverse training data facilitates the learning of more generalizable sentiment patterns.

Table 7. Cross-Domain Generalization across Target Domains.

Model Configuration	Macro-F1 for Target A (%)	Macro-F1 for Target B (%)	Macro-F1 for Target C (%)	Avg. (%)
Source-only Training	70.5 ± 0.7	72.1 ± 0.6	73.0 ± 0.6	71.9 ± 0.6
Multi-Source Training	74.2 ± 0.6	75.8 ± 0.5	76.1 ± 0.5	75.4 ± 0.5
Multi-Source + Domain Alignment	76.8 ± 0.5	78.2 ± 0.4	78.5 ± 0.4	77.8 ± 0.4
Proposed Framework	79.1 ± 0.4	80.4 ± 0.3	80.9 ± 0.3	80.1 ± 0.3

The addition of explicit domain alignment further stabilizes performance across target domains, thereby reducing sensitivity to domain-specific distributional differences. However, these symmetry-preserving strategies alone remain insufficient to fully address cross-domain discrepancies.

Importantly, the proposed framework achieves the highest Macro-F1 scores across all target domains and demonstrates consistently low variance. These results indicate stable

generalization behavior. This suggests that jointly enforcing domain-invariant symmetry while retaining domain-specific asymmetry enables the model to capture transferable representations. Consequently, the representations remain effective under varying domain shifts. These results further confirm that the proposed approach does not overfit to any specific source–target pair, but instead causes the model to generalize reliably across multiple target domains.

4.2.3. Analysis of Robustness Under Expression Asymmetry

The third experiment examines robustness to expression-level asymmetry. This phenomenon occurs when sentiment polarity must be derived from context rather than being directly stated. This problem is particularly difficult in cross-domain settings, where writing styles and context clues differ widely across platforms. Therefore, we also report the performance on sentences that express explicit sentiment versus those requiring implicit inference based on context.

As shown in Table 8, both models achieve strong performance on sentences containing explicit sentiment expressions, where sentiment polarity is directly stated. However, the baseline model exhibits a pronounced performance drop on implicit sentiment expressions. This results in a large performance gap between explicit and implicit cases. This behavior indicates limited robustness when sentiment must be inferred primarily from context rather than explicit sentiment-bearing words.

Table 8. Robustness to Expression Asymmetry.

Model	Explicit F1 (%)	Implicit F1 (%)	Gap
Multi-Source + Alignment	83.2 ± 0.5	71.5 ± 0.6	−11.7
Proposed Method	84.0 ± 0.4	76.8 ± 0.5	−7.2

In contrast, the proposed framework demonstrates substantially improved performance on implicit sentiment expressions while maintaining competitive performance on explicit cases. More importantly, it consistently reduces the gap between explicit and implicit performance with relatively low variance across runs. This reduction suggests that expression-level asymmetry exerts less influence on the proposed framework. Specifically, the framework is better able to capture contextual sentiment cues beyond explicit lexical indicators. These results further validate the effectiveness of the expression-aware learning strategy in enhancing robustness to implicit sentiment expressions.

4.2.4. Component Contribution Analysis

We conduct an ablation study on the target domain sentiment classification task to investigate how each component of the proposed framework contributes to cross-domain sentiment classification. Table 9 compiles the results.

Table 9. Ablation Study on Target Domain.

Variant	Accuracy (%)	Macro-F1 (%)
Full model	82.0 ± 0.3	80.6 ± 0.4
− Aspect loss	80.4 ± 0.4	78.9 ± 0.5
− Alignment loss	79.1 ± 0.5	77.3 ± 0.6
− SupCon loss	80.0 ± 0.4	78.2 ± 0.5
− Asymmetry-aware loss	79.5 ± 0.4	77.8 ± 0.5

The dash (–) denotes that the corresponding component is removed from the full model for ablation analysis.

Table 9 presents an ablation study that examines the contribution of each major component in the proposed framework. The full model integrates sentiment supervision, aspect-level structural modeling, multi-source domain alignment, supervised contrastive learning, and asymmetry-aware modeling. Each ablated variant removes one component while keeping all others unchanged, enabling an analysis of the individual effect of a specific module.

Removing any single component consistently leads to performance degradation in both Accuracy and Macro-F1. This indicates that all modules contribute meaningfully to the framework's overall effectiveness. In particular, removing the alignment loss results in the largest performance drop, highlighting the critical role of enforcing domain-invariant representations in cross-domain sentiment classification.

Performance also decreases when the asymmetry-aware loss is removed. This suggests that modeling domain-specific and context-inferred sentiment expressions is essential, even when symmetric alignment is present. Similarly, excluding aspect-level supervision or supervised contrastive learning leads to noticeable but smaller degradations, confirming their complementary roles in stabilizing representation learning.

Overall, the ablation results demonstrate that strong cross-domain sentiment performance arises from the synergistic interaction between symmetry-preserving components and explicit asymmetry-aware modeling, rather than from any single mechanism in isolation.

4.3. Comparison Scope and Baseline Justification

Although the proposed framework is evaluated against multiple representative baselines, recent advances in transformer-based domain adaptation motivate a more direct empirical comparison with strong neural baselines [45]. To address this, we include an additional adversarial domain adaptation baseline, DANN-BERT [46], which represents a canonical and widely adopted transformer-based adversarial adaptation approach.

The comparison is conducted under the same multi-source unsupervised domain adaptation setting. In this setting, labeled data are available only from multiple source domains, and no target-domain annotations are used during training. DANN-BERT integrates a gradient reversal mechanism into the BERT encoder to encourage domain-invariant representations, acting as a representative adversarial alignment strategy in the transformer setting.

This comparison is designed to assess whether adversarial domain alignment alone is sufficient to handle the complex symmetry–asymmetry interactions present in cross-domain sentiment analysis, particularly under platform-induced distribution shifts. All models are evaluated using the same experimental protocol and metrics to ensure a fair and controlled comparison.

As shown in Table 10, incorporating adversarial domain alignment via DANN-BERT yields clear improvements over source-only and naive multi-source baselines. This confirms the effectiveness of adversarial strategies in reducing domain discrepancy. However, the proposed symmetry–asymmetry-aware framework consistently outperforms DANN-BERT across both Accuracy and Macro-F1 metrics, with stable performance observed across multiple runs.

This performance gap suggests that while adversarial alignment is effective in promoting domain invariance, it may not sufficiently capture directional and expression-level asymmetries inherent in real-world cross-domain sentiment data. By explicitly modeling both symmetry-preserving alignment and asymmetry-aware discrepancies, the proposed framework achieves more robust and consistent adaptation under a multi-source unsupervised setting.

Table 10. Comparison with Transformer-Based Adaptation Baselines.

Method	Setting	Accuracy (%)	Macro-F1 (%)
Source-only BERT	-	75.8 ± 0.6	73.2 ± 0.7
Multi-source BERT	-	78.4 ± 0.5	76.1 ± 0.6
DANN-BERT	Adversarial DA	80.1 ± 0.4	78.3 ± 0.5
Proposed (Sym–Asym)	Multi-source UDA	82.0 ± 0.3	80.6 ± 0.4

Importantly, these results do not imply a limitation of adversarial domain adaptation itself. Instead, they highlight the complementary role of asymmetry-aware modeling in scenarios where sentiment expressions and contextual polarity vary across platforms. Overall, this empirical comparison demonstrates that explicitly addressing symmetry–asymmetry interactions provides measurable and reliable benefits beyond adversarial alignment alone.

5. Conclusions

This study presents a novel symmetry- and asymmetry-aware domain adaptation framework for cross-domain sentiment analysis. This approach addresses the limitations of approaches that rely solely on domain-invariant representation learning. By explicitly integrating symmetry-preserving alignment with asymmetry-aware modeling, the proposed framework captures both shared sentiment structures and domain-specific or expression-level variations. This is particularly beneficial for context-inferred sentiment expressions that are difficult to handle under strict invariance assumptions.

In contrast to prior domain adaptation methods that emphasize representation alignment alone, this work demonstrates the practical necessity of accounting for residual asymmetries that naturally arise across platforms. Experimental results show that modeling directional and expression-level discrepancies effectively complements symmetric alignment, leading to more stable and robust performance under domain shift. Notably, the proposed framework consistently outperforms strong transformer-based adversarial adaptation baselines, with statistically significant improvements observed across multiple runs and all evaluated target domains.

Further analyses reveal that symmetry-aware and asymmetry-aware components contribute complementary effects rather than functioning as isolated mechanisms. Ablation and robustness studies confirm that explicit asymmetry modeling is particularly important for handling implicit sentiment expressions, while multi-source symmetry enforcement remains essential for transferable representation learning.

Despite the encouraging results, this study has several limitations. The experiments focus on hotel review datasets to control topical variability and isolate platform-induced domain shift. Furthermore, the framework is evaluated on binary sentiment classification with English-language data only. Moreover, aspect-level structure is based on a predefined schema, and the adaptation setting assumes unlabeled target-domain data. Future work will extend this framework to multilingual and fine-grained sentiment settings, incorporate limited target-domain supervision, and explore multimodal sentiment adaptation involving visual and acoustic cues.

Overall, this work demonstrates that jointly modeling symmetry and asymmetry provides a robust and principled solution for cross-domain sentiment analysis under realistic platform shifts. This offers both theoretical insight and practical value for real-world deployment.

Author Contributions: Conceptualization, C.S., J.P. and M.K.; methodology, C.S., J.P. and M.K.; formal analysis, C.S., J.P. and M.K.; investigation, C.S., J.P. and M.K.; data curation, T.K., J.S.-a., A.C., J.J., V.V. and B.L.; validation, T.K., J.S.-a., A.C., J.J., V.V. and B.L.; writing—original draft preparation, C.S., J.P. and M.K.; writing—review and editing, C.S., J.P. and M.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research received internal funding provided by Mahasarakham University, Thailand. Funding number is IT1-2568.

Data Availability Statement: The data used in this study were collected from publicly accessible online review platforms, including TripAdvisor, Booking.com, and Agoda. The raw review content is publicly available through these platforms; however, the annotated and processed dataset generated in this study cannot be publicly shared due to platform usage policies and the involvement of expert annotations. The processed data may be made available from the corresponding author upon reasonable request for academic research purposes.

Acknowledgments: This work was financially supported by Mahasarakham University.

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

References

1. Medhat, W.; Hassan, A.; Korashy, H. Sentiment analysis algorithms and applications: A survey. *Ain Shams Eng. J.* **2014**, *5*, 1093–1113. [CrossRef]
2. Poria, S.; Hazarika, D.; Majumder, N.; Mihalcea, R. Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research. *IEEE Trans. Affect. Comput.* **2023**, *14*, 108–132. [CrossRef]
3. Patil, R.; Boit, S.; Gudivada, V.; Nandigam, J. A survey of text representation and embedding techniques in NLP. *IEEE Access* **2023**, *11*, 36120–36146. [CrossRef]
4. Da Costa, L.S.; Oliveira, I.L.; Fileto, R. Text classification using embeddings: A survey. *Knowl. Inf. Syst.* **2023**, *65*, 2761–2803. [CrossRef]
5. Subakti, A.; Murfi, H.; Hariadi, N. The performance of BERT as data representation of text clustering. *J. Big Data* **2022**, *9*, 15. [CrossRef]
6. Glorot, X.; Bordes, A.; Bengio, Y. Domain adaptation for large-scale sentiment classification: A deep learning approach. In Proceedings of the International Conference on Machine Learning, Bellevue, WA, USA, 28 June–2 July 2011.
7. Elming, J.; Plank, B.; Hovy, D. Robust cross-domain sentiment analysis for low-resource languages. In Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, Gothenburg, Sweden, 27 April 2014.
8. Wu, F.; Huang, Y. Sentiment domain adaptation with multiple sources. In Proceedings of the Annual Meeting of the Association for Computational Linguistics, Berlin, Germany, 7–12 August 2016.
9. Wu, F.; Wu, S.; Huang, Y.; Huang, S.; Qin, Y. Sentiment domain adaptation with multi-level contextual sentiment knowledge. In Proceedings of the ACM International Conference on Information and Knowledge Management, Indianapolis, IN, USA, 24–28 October 2016.
10. Zhou, G.; Huang, X. Modeling and mining domain-shared knowledge for sentiment analysis. *ACM Trans. Inf. Syst.* **2017**, *36*, 1–36. [CrossRef]
11. Barnes, J.; Klinger, R.; Schulte im Walde, S. Projecting embeddings for domain adaptation: Joint modeling of sentiment analysis in diverse domains. In Proceedings of the International Conference on Computational Linguistics, Santa Fe, NM, USA, 20–26 August 2018.
12. Cheeti, S.; Stanescu, A.; Caragea, D. Cross-domain sentiment classification using an adapted naïve Bayes approach and features derived from syntax trees. In Proceedings of the International Conference on Knowledge Discovery and Information Retrieval and the International Conference on Knowledge Management and Information Sharing, Paris, France, 29–31 October 2013; pp. 169–176.
13. Yu, H.; Pan, Y.; Zhou, C. Domain adaptation approach for sentiment analysis. In Proceedings of the International Conference on Compute and Data Analysis, Lakeland, FL, USA, 5–7 April 2019.
14. Zainuddin, N.; Selamat, A.; Ibrahim, R. Hybrid sentiment classification on Twitter aspect-based sentiment analysis. *Appl. Intell.* **2018**, *48*, 1218–1232. [CrossRef]
15. Giménez, M.; Palanca, J.; Botti, V. Semantic-based padding in convolutional neural networks for improving NLP performance: A case study in sentiment analysis. *Neurocomputing* **2020**, *378*, 315–323. [CrossRef]

16. Li, X.; Fu, X.; Xu, G.; Yang, Y.; Wang, J.; Jin, L. Enhancing BERT representation with context-aware embedding for aspect-based sentiment analysis. *IEEE Access* **2020**, *8*, 46868–46876. [CrossRef]
17. Li, Z.; Zou, Y.; Zhang, C.; Zhang, Q.; Wei, Z. Learning implicit sentiment in aspect-based sentiment analysis with supervised contrastive pre-training. *arXiv* **2021**. [CrossRef]
18. Fu, J.; Li, X.; Zhu, Y.; Du, Y.; Fan, Y.; Chen, X.; Huang, D.; Wang, S. An implicit aspect-based sentiment analysis method using supervised contrastive learning and knowledge embedding. *Appl. Soft Comput.* **2024**, *167*, 112233. [CrossRef]
19. Brauwers, G.; Frasincar, F. A survey on aspect-based sentiment classification. *ACM Comput. Surv.* **2022**, *55*, 1–37. [CrossRef]
20. 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]
21. Chan, J.Y.-L.; Bea, K.T.; Leow, S.M.H.; Phoong, S.W.; Cheng, W.K. A review of sentiment analysis based on sequential transfer learning. *Artif. Intell. Rev.* **2023**, *56*, 749–780. [CrossRef]
22. Jiang, H.; Chen, X.; Miao, D.; Zhang, H.; Qin, X.; Gu, X.; Lu, P.; Li, X. Fpa-GCN: Enhancing aspect sentiment triplet extraction with feature-rich prediction-aware graph convolutional networks. *Appl. Intell.* **2025**, *55*, 740. [CrossRef]
23. Lin, Z.; Wang, Y.; Zhou, Y.; Du, F.; Yang, Y. MLM-EOE: Automatic depression detection via sentimental annotation and multi-expert ensemble. *IEEE Trans. Affect. Comput.* **2025**, *16*, 2842–2858. [CrossRef]
24. Wang, Y.; Lin, Z.; Yang, C.; Zhou, Y.; Yang, Y. Automatic depression recognition with an ensemble of multimodal spatio-temporal routing features. *IEEE Trans. Affect. Comput.* **2025**, *16*, 1855–1872. [CrossRef]
25. Jiang, H.; Chen, X.; Miao, D.; Zhang, H.; Qin, X.; Du, S.; Lu, P. 3WD-DRT: A three-way decision enhanced dynamic routing transformer for cost-sensitive multimodal sentiment analysis. *Inf. Sci.* **2026**, *725*, 122704. [CrossRef]
26. Chai, Y.; Weng, Y.; Shao, T.; Zhou, K. A semantic talking style space for speech-driven facial animation. *IEEE Trans. Vis. Comput. Graph.* **2025**, *31*, 10801–10814. [CrossRef]
27. Wu, X.; Li, L.; Tao, X.; Xing, F.; Yuan, J. Happiness prediction with domain knowledge integration and explanation consistency. *IEEE Trans. Comput. Soc. Syst.* **2025**, *12*, 2949–2962. [CrossRef]
28. Ji, J.; Han, B.; Liu, X.; Xu, X.; Aysa, A.; Ubul, K. DPA-MVSNet: Dynamic context perception multi-view stereo with transformers and data augmentation. *Knowl.-Based Syst.* **2025**, *325*, 113852. [CrossRef]
29. Arabadzhieva-Kalcheva, N.; Kovachev, I. Comparison of BERT and XLNet accuracy with classical methods in text classification. In *Proceedings of the International Conference on Biomedical Innovations and Applications*; IEEE: New York, NY, USA, 2021.
30. Shahriar, A.; Pandit, D.; Rahman, S. XLNet-CNN: Combining global and local context for multi-label text classification. In *Proceedings of the International Conference on Networking, Systems, and Security*; Association for Computing Machinery: New York, NY, USA, 2024; pp. 24–31.
31. Jose, A.; Harikumar, S. Predicting IMDb movie ratings using RoBERTa embeddings and neural networks. In *Responsible Data Science*; Springer: Singapore, 2022; pp. 181–189.
32. Mollah, A.R.; Kabir, J.; Kabir, M.; Reza, S. Detection of fake news with RoBERTa-based embedding and modified deep neural network architecture. In *Proceedings of the International Conference on Computer and Information Technology*; IEEE: New York, NY, USA, 2023; pp. 1–6.
33. Ghanem, R.; Erbay, H.; Bakour, K. Content-based spam detection on social networks using RoBERTa embedding and stacked BLSTM. *SN Comput. Sci.* **2023**, *4*, 380. [CrossRef]
34. Alizadeh, M.; Seilsepour, A. A novel self-supervised sentiment classification approach using semantic labeling based on contextual embeddings. *Multimed. Tools Appl.* **2024**, *84*, 10195–10220. [CrossRef]
35. He, K.; Mao, R.; Gong, T.; Li, C.; Cambria, E. Meta-based self-training and re-weighting for aspect-based sentiment analysis. *IEEE Trans. Affect. Comput.* **2023**, *14*, 1731–1742. [CrossRef]
36. Fu, Y.; Liu, Y. Contrastive transformer-based domain adaptation for multi-source cross-domain sentiment classification. *Knowl.-Based Syst.* **2022**, *245*, 108649. [CrossRef]
37. Zhao, C.; Wu, M.; Yang, X.; Sun, X.; Wang, S.; Li, D. Cross-domain aspect-based sentiment classification with a pre-training and fine-tuning strategy for low-resource domains. *ACM Trans. Asian Low-Resour. Lang. Inf. Process.* **2024**, *23*, 1–22. [CrossRef]
38. Deng, Y.; Zhang, W.; Pan, S.J.; Bing, L. Bidirectional generative framework for cross-domain aspect-based sentiment analysis. In *Proceedings of the Annual Meeting of the Association for Computational Linguistics*; Association for Computational Linguistics: Toronto, ON, Canada, 2023.
39. Luo, Y.; Guo, F.; Liu, Z.; Zhang, Y. Mere contrastive learning for cross-domain sentiment analysis. In *Proceedings of the International Conference on Computational Linguistics*; International Committee on Computational Linguistics: Gyeongju, Republic of Korea, 2022.
40. Wang, Q.; Wen, Z.; Ding, K.; Liang, B.; Xu, R. Cross-domain sentiment analysis via disentangled representation and prototypical learning. *IEEE Trans. Affect. Comput.* **2025**, *16*, 264–276. [CrossRef]
41. Wang, K.; Wan, X. Counterfactual representation augmentation for cross-domain sentiment analysis. *IEEE Trans. Affect. Comput.* **2023**, *14*, 1979–1990. [CrossRef]

42. Gopalakrishnan, S.; Chen, V.Z.; Dou, W.; Zadrozny, W. On the relation between K–L divergence and transfer learning performance on causality extraction tasks. *Nat. Lang. Process. J.* **2024**, *6*, 100055. [CrossRef]
43. Nguyen, A.; Tran, T.; Gal, Y.; Torr, P.H.S.; Baydin, A.G. KL-guided domain adaptation. In Proceedings of the International Conference on Learning Representations, Virtual, 3–7 May 2021.
44. Yu, D.; Yao, K.; Su, H.; Li, G.; Seide, F. KL-divergence regularized deep neural network adaptation for improved large vocabulary speech recognition. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*; IEEE: New York, NY, USA, 2013.
45. Wright, D.; Augenstein, I. Transformer based multi-source domain adaptation. *arXiv* **2020**. [CrossRef]
46. Xiao, X.; Chen, X.; Li, W. Improved domain-adversarial neural networks with mixup and pseudo-labeling. In *Proceedings of the International Conference on Cyber-Physical Social Intelligence (ICCSI)*; IEEE: New York, NY, USA, 2024.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.