Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2025 SCPE. Volume 26, Issues 4, pp. 1671–1683, DOI 10.12694/scpe.v26i4.4566

ACDPSNET: ADAPTIVE CROSS DOMAIN POLARITY ASPECT LEVEL LEARNING SCALABLE COMPUTING MODEL FOR SENTIMENT CLASSIFICATION AND QUANTIFICATION

JHANSI RANI T[*], SWAPNA NEERUMALLA[†], AKUNDI SAI HANUMAN [‡], B. VEERASEKHAR REDDY [§], AND KAYAM SAIKUMAR[¶]

Abstract. Automatic sentiment classification, identifying opinions as positive, negative, or neutral, is essential across diverse applications. However, applying a sentiment classifier trained on labeled data from one domain to a different domain often leads to degraded performance, as domain-specific language terms common in the source domain may not appear in the target domain. This research proposes an Adaptive Cross-Domain Polarity-Specific Network (ACDPSNet) for sentiment classification and quantification across domains. The model leverages labeled data from the source domain alongside labeled and unlabeled data from the target domain to build a robust, adaptable domain adaptation framework. Sensitivity to sentiment is enhanced by embedding polarity-specific sentiment annotations into semantic vectors, enabling accurate computation of distributional similarities between terms. The framework integrates a classifier that is both domain-specific and domain-invariant to ensure accurate analysis and classification. ACDPSNet achieves notable performance improvements, with an accuracy of 98.76%, recall of 97.85%, throughput of 96.94%, and a positive learning expression rate of 97.76%, demonstrating significant advancements over existing approaches. These metrics underscore ACDPSNet's effectiveness in adapting to new domains, achieving high sentiment quantification accuracy, and enhancing cross-domain polarity detection.

Key words: Polarity, Domain Transfer, Scalable computing Pivot Model, Domain Adaptation and Cross-Domain Sentiment Classification

1. Introduction. The increasing importance of sentiment analysis across various applications has led to significant research interest in this field. Traditional studies often emphasize predicting the sentiment of entire texts, ranging from paragraphs to individual phrases [1]. However, accurately discerning sentiment towards specific aspects within a text is essential, as it requires an in-depth understanding of the contextual language surrounding those aspects. This challenge, known as automatic sentence-level sentiment classification [2], is crucial for applications such as market analysis, opinion mining, and contextual advertising.

Cross-domain sentiment classification is particularly challenging because it involves applying classifiers trained on one domain (source domain) to a different domain (target domain). This challenge entails two significant issues: identifying the common characteristics between the source and target domains and developing a learning framework that incorporates this relatedness. In our research, we propose a hybrid approach for emotion classification across domains to address these challenges effectively.

In sentiment analysis, accurately determining the polarity of opinions at the aspect level is critical, especially when dealing with the complexities of language, context, and topic variations across different domains. Traditional sentiment classification methods often struggle to maintain accuracy in cross-domain scenarios, where these variations can significantly impact performance. The challenge is even greater when the objective extends beyond classification to quantification—estimating the prevalence of each sentiment class within a dataset.

[*]Department of CSE, GITAM School of Technology, GITAM Deemed-to-be-University, Hyderabad, India (jhansi.rani.t@gmail.com)

[†]Vignana Bharathi Institute of Technology, Hyderabad, India (swapnakiran29@gmail.com)

[‡]Gokaraju Lailavathi womens engineering college, Hyderabad, Telangana, India (a_saihanuman@hotmail.com)

[§]Department of information technology,MLR Institute of Technology, Hyderabad-500043, Telangana, India (veerasekhar@mlrinstitutions.ac.in)

[¶]Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Hyderabad-500075, Telangana, India (saikumar.kayam@klh.edu.in)

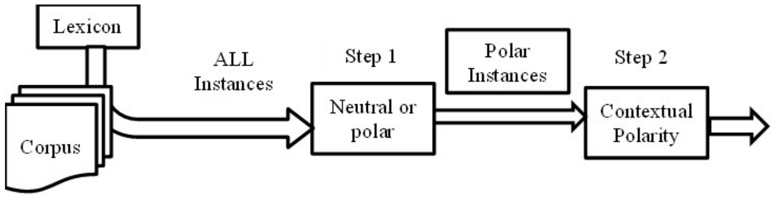

Fig. 1.1: Positive, Negative and both neural

Adaptive Cross-Domain Polarity Aspect-Level Sentiment Classification and Quantification represents a pivotal area of research, aimed at overcoming these challenges. This approach focuses on developing models that can adapt to new domains with minimal labeled data while providing robust sentiment quantification. By leveraging the potential of cross-domain transfer learning, these models can be effectively applied to new domains, thereby enhancing both classification accuracy and the reliability of sentiment quantification.

Polarity classification. Figure 1.1 illustrates the outcomes of our proposed research, which focuses on developing and evaluating adaptive methods that not only classify sentiment at the aspect level but also quantify the distribution of sentiment across different domains. This dual approach is particularly important in scenarios where understanding both the intensity and distribution of sentiment is as crucial as identifying the sentiment itself. By addressing these challenges, our research contributes to the broader objective of creating more flexible and accurate sentiment analysis systems capable of operating effectively across a wide range of domains and contexts.

Aspect-level sentiment classification (ASC) is a nuanced task in sentiment analysis that aims to determine the polarity—positive, neutral, or negative—towards specific opinion targets within a sentence. For example, in the sentence "Average to good Thai food, but terrible delivery," the opinion targets are "Thai food" and "delivery," with corresponding sentiments being positive and negative, respectively. The advent of deep learning in natural language processing (NLP) has significantly advanced ASC tasks, with neural network models often outperforming traditional machine learning techniques. These deep learning models have demonstrated superior performance in ASC due to their ability to capture complex patterns in the data. Given the specific requirements of ASC, particularly the need to distinguish between sentiments expressed towards different targets within the same context, recent studies have increasingly incorporated attention mechanisms into deep learning models. These attention-based approaches improve sentiment prediction by focusing on sentiment-laden words that are relevant to specific targets.

However, the effectiveness of deep learning models heavily depends on the availability of sufficient training data. In practical applications, generating aspect-level training data requires extensive manual annotation, which limits the size of available public datasets and, consequently, the performance of neural network models. On the other hand, large volumes of document-level sentiment classification (DSC) labeled data are available from many online review platforms, offering rich emotional insights and semantic patterns. This presents an important research question: how can the valuable knowledge contained in DSC data be harnessed to enhance ASC tasks, especially when aspect-level resources are limited. The core challenge in cross-domain sentiment classification is training a classifier on one or more source domains and effectively applying it to a different target domain. Our proposed framework, named Adaptive Cross-Domain Polarity Aspect-Level Sentiment Classification and Quantification (ACDPASCQ), addresses this challenge by integrating various methodologies. It employs co-training on target unlabeled data to achieve invariant classification and analysis, extracts both domain-invariant and domain-specific aspects from the target domain data, and identifies significant polarity words that are consistent across domains.

The objectives the work follows:

1. Transferable Information Across Domains
2. Identification of Target Features from Source and Target Domains
3. Classification and Analysis of Cross-Domain Sentiment

By integrating these approaches within the ACDPASCQ framework, we aim to provide a comprehensive solution to the challenges posed by cross-domain sentiment classification.

2. Related work. The field of sentiment classification systems is generally divided into two main categories: single-domain classifiers [2] and cross-domain classifiers [3]. Our research centers on document-level cross-domain sentiment classification [4]. In this context, a classifier is initially trained for single-domain sentiment categorization using labeled data specific to the application domain [4]. For example, Turney [5] introduced an approach using point-wise mutual information to assess the sentiment of a word by analyzing its co-occurrence with manually selected positive (e.g., fine, wonderful, fantastic) and negative (e.g., dreadful, unpleasant, weak) words. While single-domain sentiment classification has been extensively explored [6], recent advancements in domain adaptation techniques have brought cross-domain sentiment classification into focus.

The objective of sentiment classification is to assign an emotional polarity to a given text, typically categorized as positive, negative, neutral, or into more nuanced categories. This area has garnered significant attention in recent years [7], particularly because real-world texts often involve multiple target entities or specific aspects of a topic. Customer reviews serve a dual purpose: they help other customers make informed decisions and assist online retailers in predicting sales [8] and understanding customer preferences [9]. This understanding enables retailers to craft effective marketing strategies to boost revenue. However, the sheer volume of reviews across diverse domains on these platforms presents a challenge in efficiently extracting the most relevant information. Consequently, researchers have increasingly focused on developing automated methods for cross-domain aspect-based sentiment classification [10].

The primary challenge in cross-domain aspect-based sentiment classification lies in the discrepancy between the training and testing data, which originate from different domains and thus exhibit distinct characteristics [11]. Previous studies have explored two main approaches to address this issue: data-based and feature-based. The data-based approach focuses on constructing a training dataset that closely resembles the target data. Typically, this involves generating pseudo-labels for the target data and incorporating these labeled target data into the training set [12]. However, the effectiveness of this approach heavily depends on the quality of the generated pseudo-labels, which directly impacts the model's performance. A feature-based approach has been proposed [13] to overcome the limitations of the data-based approach. Rather than generating pseudo-labels, this approach seeks to identify domain-independent features shared between the source and target domains. These features often include syntactic dependency relations and domain-independent words [14]. Researchers have employed models such as Conditional Random Fields [15] and Recurrent Neural Networks [16] to encode syntactic dependency relations. For domain-independent words, higher weights are assigned to these words than domain-specific ones [17].

By leveraging domain-independent information, connections between the source and target domains are established, enabling the trained model to perform well across both domains. Aspect-level sentiment classification (ASC) has seen significant advancements, particularly with the adoption of deep learning methods that enhance the precision and accuracy of sentiment analysis at a granular level [18]. Traditional approaches to ASC often relied on supervised learning techniques that required substantial amounts of annotated data, posing limitations, especially when adapting models to new domains with scarce labeled data.

Cross-Domain Sentiment Analysis. This area of research has emerged to address the challenge of transferring knowledge from a source domain, rich in labeled data, to a target domain with limited or no labeled data. Early works, such as those by [19], utilized domain adaptation techniques to reduce discrepancies between domains, enabling more effective sentiment classification across different contexts. These foundational approaches paved the way for more sophisticated methods that incorporate deep learning.

Deep Learning and Attention Mechanisms. The advent of deep learning has introduced neural networks as the backbone for ASC. Models such as Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been widely adopted for their ability to learn features from data automatically. Attention mechanisms have further enhanced ASC by allowing models to focus on relevant text parts when making sentiment predictions. Attention-based models, such as the Hierarchical Attention Network [20] and the Aspect-Based Sentiment Classification model [21], have significantly improved capturing sentiment at the aspect level. Domain Adaptation Techniques: Recent studies have explored various domain adaptation techniques in conjunction with deep learning to improve cross-domain ASC. Models like Transfer Learning with Fine-Tuning and Adversarial Training for Domain Adaptation have shown promise in transferring sentiment knowledge from one domain to another, improving classification accuracy in the target domain. These methods are particularly effective when domain-specific language and context differ significantly. Sentiment Quantification: Beyond classification, sentiment quantification has become an important research area. The goal is not just to classify sentiment but to quantify sentiment distribution across different categories within a dataset. Traditional sentiment quantification methods, such as those proposed by [22], relied on aggregated document-level predictions. However, recent advancements have integrated quantification techniques with deep learning models, allowing for a more accurate estimation of sentiment prevalence in diverse domains. Challenges and Future Directions: Despite these advancements, challenges remain in achieving robust cross-domain ASC and sentiment quantification. One of the main difficulties is the heterogeneity of language and sentiment expression across domains, which can lead to reduced model performance. Future research will likely focus on developing more sophisticated domain adaptation techniques, possibly leveraging unsupervised or semi-supervised learning to reduce reliance on labeled data [23]. Additionally, integrating quantification with ASC in a unified framework remains an open research question with the potential to significantly enhance the practical applicability of sentiment analysis models in real-world scenarios [24]. The existing models are facing issues with misclasses balancing and batch normalization. The open-source dataset used by earlier models can drop its MAP (Mean Average Precision) and performance [25]. The limitations of existing models can be crossed over through custom deep-learning models with benchmark dataset training.

3. Proposed Methodology. This paper first evaluates the adaptive cross-domain polarity aspect level sentiment classification and quantification (ACDPASCQ) framework for domain adaptation. We design a polarity of aspect-level sentiment analysis by deriving the polarity of words for labeled and unlabelled data with a pivot model. Classification of labeled and unable data across domains using domain invariant and specific classifier.

Aspect-Level Sentiment Classification. In the early research on aspect-level sentiment classification (ASC), the primary methodologies were heavily dependent on feature engineering. For example, Kiritchenko et al. [9] utilized n-gram features and developed new lexical resources, integrating these features into classification models using Support Vector Machines (SVM). Similarly, Yi and Zhou [10] designed sentiment feature vectors by calculating sentiment values and employed models such as Naive Bayes and SVM for training. Although these methods achieved notable results, their performance was largely contingent on the quality of manually crafted features, which demanded significant time and effort in feature design. The advent of deep learning addressed these limitations by enabling neural network models to automatically learn crucial sentiment features from text based on sentence vectors, eliminating the need for manual feature construction. Dong et al. [4] were among the first to apply Recurrent Neural Networks (RNNs) to ASC, improving sentiment classification accuracy by using RNNs to extract sentiment polarity from text and integrating syntactic structure information to support the model. Xue et al. [3] introduced a convolutional neural network (CNN) model with a gating mechanism that selectively outputs emotional features through convolutional layers. To mitigate the issue of gradient explosion associated with RNNs, Tang [2] proposed the use of Long Short-Term Memory (LSTM) networks, which model the left and right contexts of a given opinion target to extract sentiment information. However, due to the fine-grained nature of ASC, these models often struggled to effectively capture the relationship between the context and the specified opinion target. To address this challenge, subsequent research focused on incorporating attention mechanisms to capture target-dependent sentiment contexts. Wang [5] proposed an attention-based LSTM that enhances relevance by concatenating aspect word vectors with context vectors and then applying self-attention to extract sentiment knowledge specific to the aspect word and its context. Xu [11] introduced a dual attention module, combining global and local attention to capture different granularities of interaction information between aspects and contexts. Lin [12] utilized multi-head target-specific self-attention to better capture global dependencies and introduced target-sensitive transformations to address target-specific sentiment. Liu [13] further refined sentiment feature extraction by employing multi-head attention to capture semantic information between words related to the specified aspect and replaced the softmax function in the classification layer with SVM to improve feature representation in high-dimensional space. Li [6] integrated syntactic dependency information with semantic information, facilitating interaction between aspect words and sentences through attention-based graph convolutional networks (GCNs). Huang [14] developed a contextual location weighting function that considers the positional information of aspect words within the context, thereby reducing the influence of surrounding words on sentiment polarity. Despite the significant improvements in ASC performance brought about by these deep learning-based methods, they remain highly dependent on the availability of data. The limited size of existing ASC datasets constrains the ability of these supervised models to realize their full potential. To overcome this limitation, this paper proposes leveraging the Document-Level Sentiment Classification (DSC) task to transfer a large amount of sentiment knowledge, thereby mitigating the impact of insufficient data on ASC performance.

Transfer Learning. Transfer learning aims to utilize knowledge from one or more related tasks (source tasks) and apply it to a different but related target task. Transfer learning techniques in natural language processing (NLP) are generally categorized into three main types: instance transfer, model transfer, and domain adaptation. These methods have been successfully implemented in various NLP subtasks, including machine translation [14], question-answering systems [16], and speech recognition [17]. Model transfer has become a well-established method for leveraging knowledge from document-level tasks to support aspect-level tasks.

Model transfer is often employed in multi-task learning, where data from multiple related subtasks are used, and shared modules are applied to learn the relationships between these tasks, thereby extracting additional useful information. For instance, Xu [18] introduced a pre-training plus multi-task learning model. This approach first involves training on a document-level dataset using a shared BiLSTM module to obtain pre-trained weights. These weights are then retained as initialization parameters for the shared part of the model. Subsequently, aspect-level data are fed into the pre-trained model to train both tasks simultaneously, allowing for fine-tuning of the weights.

To enable the flexible application of document-level knowledge, Chen [32] proposed a model based on transfer capsules. Unlike the method used by Xu [18], Chen's approach employs multi-task learning with heterogeneous datasets [20], where both document-level and aspect-level datasets are fed into the model simultaneously. The shared parameters are dynamically optimized, and at the upper layer of the model, semantic capsules and dynamic routing are combined with the transferred knowledge.

While these methods are highly effective, they are limited by the hard parameter sharing inherent in vanilla model transfer (VMT) [21] used within the shared modules. This hard sharing can lead to a scenario where the shared module negatively impacts learning in the target task due to differences in tasks and data. In contrast, our auto-adaptive model transfer method addresses these discrepancies, allowing the model to learn more refined and relevant information from the auxiliary task, thereby enhancing the accuracy of ASC.

Domain Adoption. A set of ns fully named Ds = (xs 1, ys 1) and (xs ns, ys ns) = Rd = Y chosen from the Ps(X, Y) array make up the root domain. Also divided into nl (nl ns) categorized points is the aim data. Dl t = (xt 1, yt 1), xt nl, yt nl Nu (nu nl) unlabelled points and Ru Y from the distribution Pt(X, Y) Du t is equal to (xt nl+1,yt nl+1), The objective is to build a classifier for target data using the source domain data and a few selected target domain data.

In this section we present CMD parameters used to calculate the variance between two Random variables in the distribution of probabilities. Here, we extract the domain-specific and Domain invariant representation from the target domain instances. Finally, demonstrate how these two interpretations can be mixed using the co-training framework shown in figure 3.1.

3.1. Central Moment Discrepancy (CMD). Zellinger et al. (2017) suggested the CMD parameter to calculate the difference between two (high-dimensional) random variables in the probability distributions. Let X and Y be random samples bounded at the interval [a,b]N with p and q, the corresponding probability distributions. The CMD regulator CMDK is specified as

$$CMD(X,Y) = \frac{1}{|b-a|}\left\|E(X)-E(Y)\right\|_2 + \frac{1}{|b-a|^k}\sum_{k=2}^{k}\left\|C_k(X))-C_k(Y)\right\|_2 \tag{3.1}$$

The vector of analytical expectations based on the X sample is denoted by

$$E(X) = 1|X|Px\epsilon Xx. \tag{3.2}$$

$$C_k(X) = (E(\prod_{i=1}^{n})X_i - E(X_i))^{ri})r_{i\geq 0}, \sum_{i}^{n} r_i = k \tag{3.3}$$

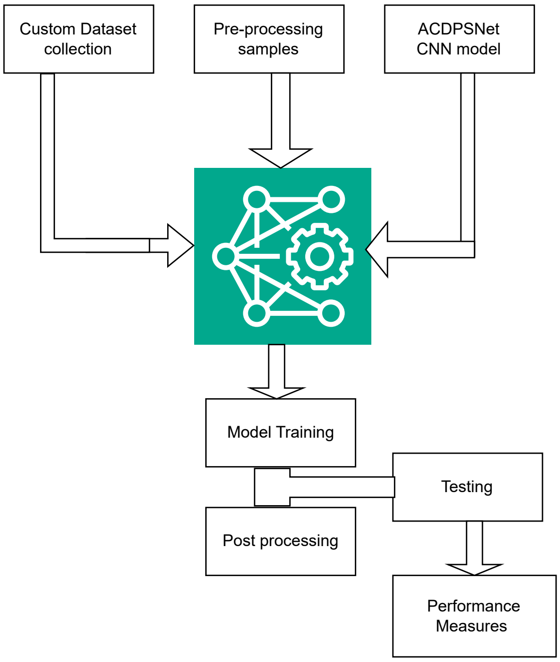

Fig. 3.1: Proposed model Block Diagram

Is the matrix of all kth order sequence main co-ordinate times X. The implicit interpretation of this equation is that whenever two distributions of probabilities are close, There will be a closer core moment of greater order.

3.2. Retrieve Domain Specific and Domain Invariant Representations. Our goal in this work is to obtain both a domain-specific counterpart and a domain-invariant representative for every target instance. Using two separate mappers, Et and Ec, respectively, Data is converted into a unique hidden space for the desired domain and a domain-invariant hidden space.

$$H^s_{spe} = E_t(X_s, \Theta^t_e) \tag{3.4}$$

$$H^t_{spe} = E_t(X_t, \Theta^t_e) \tag{3.5}$$

$$H^t_{spe} = E_t(X_t, \Theta^t_e) \tag{3.6}$$

$$H^s_{inv} = E_t(X_s, \Theta^c_e) \tag{3.7}$$

$$H^t_{inv} = E_t(X_t, \Theta^c_e) \tag{3.8}$$

The goal domain-specific mapper in this case is Et stands for the mapper that is domain invariant, and Ec. The related parameters are defined by $\Theta^{\ c}_{e}$ and $\Theta^{\ t}_{e}$ Encode signifies the subscript e. Depending on the hidden $H^t_{inv} and H^t_{spe}$ representations, we are creating an autoencoder for the examples of target domains:

$$X_t = D_t(H^t_{inv}, H^t_{spe}, \Theta^t_d) \tag{3.9}$$

As far as parameters are concerned,Θ_d^t the d subscript denotes decoding. The resulting reconstruction failure is the mean square error described as:

$$L_{recon} = \frac{1}{n_t}\sum_{i}^{n_t}\frac{1}{k}\left\|X_t^i - X_t^i\right\|^2 \tag{3.10}$$

where Xi t is the target domain data, for instance, ith, and k is the input function vector component. Remember that in this task, the auto-encoder receives only target instances, as our goal is to obtain accurate information about the target domain. The hidden representation of the $(H - inv^s)$ source data and the $(H - inv^t)$ target data have the CMD regularizer added. The corresponding loss will be explained as follows:

$$L_{sim} = CMD_k(H_{inv}^s, H_{inv}^t) \tag{3.11}$$

Ec is encouraged to encode domain-specific invariant features when this loss is minimized since it will force the H_{inv}^s, H_{inv}^t distributions to be identical. The loss in question will be explained as follows:

$$L_{diff} = -CMD_k(H_{spe}^s, H_{spe}^t) \tag{3.12}$$

Minimizing the error allows H_{spe}^s propagation to vary from H_{spe}^t which in effect enables Et to encode different domain features.

3.3. Polarity of Words in the Labelled Source Domain. The statistically significant correlation between a word and a class label is supported by the chi-square test. We give each word in the domain a polarity orientation based on this relationship. Since the target domain data is unlabelled, a χ^2 check cannot be utilized to determine the words' meaning. However, we exploit the fact that only some terms in the target domain that are known to be significant in the source domain need to be characterized as meaning in order to acquire SCP terms across domains. It is presumed that a term that meets the χ^2 test criteria for relevance in the source domain and appears frequently (nearly) in the target domain is also significant. According to the χ^2 test, the source domain is also significant in the target domain when it appears more frequently than a particular threshold (nearly).

$$count_t(significant_s(w)) > \Theta \Rightarrow significant_t(w) \tag{3.13}$$

The labelled source (s) domain's significance of the word w is guaranteed by [significant]-s, whereas [count]-s provides the normalized count of the w in t. Using this supposition as a foundation, we fix the value of Θ.

Word Polarity in the Target Unlabelled Domain. Positive words typically exist in polar corpuses with other positive words, whereas negative words typically occur in conjunction with other negative words (Sharma et al., 2015). Mikolov et al. (2013) discovered that nearby words, such "go" and "to," had a higher degree of similarity in their meaning vectors than far-off words or words that are not nearby. Using the publicly available word2 and the skip-gram model toolbox, we measured the context vector (conVec) of a word (w) (Mikolov et al., 2013).7. This model predicts words within a given window by using the Huffman code of each word as an input to a log-linear classifier with a persistent projection layer. Equation 3's decision-making process outlines how to assign polarity to an unknown term in the target domain.

$$If(cosine(conVec(w), conVec(PosPivot)) > cosine(conVec(w), conVec(NegPivot))) \Rightarrow Positive \tag{3.14}$$

$$If(cosine(conVec(w), conVec(PosPivot)) < cosine(conVec(w), conVec(NegPivot))) \Rightarrow Negative \tag{3.15}$$

3.4. Pivot Selection Method. We found empirically that a polar term that has the largest percentage in the corpus offers more coverage by using context vector to estimate the polarity orientation of other terms. Furthermore, it is discovered that a polar term with the highest frequency in the target domain works better as the pivot for input term polarity detection. A few phrases in the electronics domain are shown in Table 3.1 whose polarity orientation is determined by similarity scores acquired using PosPivot and NegPivot. The inferred polarity orientation of the words along with the cosine-similarity scores using PosPivot (excellent) and NegPivot (bad) is shown in table 3.1.

Table 3.1: The inferred polarity orientation of the words

word	Great	Poor	Polarity
Noisy	0.03	0.24	Neg
Crap	0.04	0.28	Neg
weak	0.05	0.21	Neg
Defective	0.21	0.70	Neg
Sturdy	0.43	0.04	Pos
Durable	0.44	0.00	Pos
perfect	0.48	0.20	Pos
Handy	0.60	0.21	Pos

3.5. Domain Transferable Knowledge. To determine the importance and polarity of terms in the labelled source data and the unlabelled target data, the suggested algorithm makes use of the previously discussed techniques. Significant, consistent polarity (SCP) characteristics are a group of phrases that are relevant in both fields and have the same polarity orientation. These features are used to classify attitudes that are cross-domain. The weights learned by the classification algorithm for the SCP features in the labeled source domain can be reused in the unlabelled target domain for sentiment classification because SCP features have clear impacts in both domains. Based on the cosine correlation function, it differentiates the positive or negative polarity.

Algorithm 1 Aspect Level Domain Adaptation

Input:

- L_s: Instances labeled in the source domain
- L_t: Instances labeled in the target domain
- U_t: Unlabeled instances in the target domain

Representations:

- H^s_{inv}: Invariant representation for L_s
- H^t_{inv}: Invariant representation for L_t
- H^t_{spec}: Specific representation for L_t

Steps:

1. **Train Classifier:** Train classifier F_c on labeled instances L_s and L_t using invariant representations H^s_{inv} and H^t_{inv}.
2. **Classify Unlabeled Data:** Apply classifier F_c to predict labels for instances in U_t.
3. **Select High-Confidence Predictions:**
   - Identify instances in U_t with the highest confidence scores.
   - Select positive instances p and negative instances n to form the subset U^t_c containing these high-confidence predictions.
4. **Train Specific Classifier:** Train a separate classifier F_t on L_t using the specific representation H^t_{spec}.
5. **Refine Predictions on U_t:**
   - Apply classifier F_t to predict labels for instances in U_t.
   - Select positive instances p and negative instances n with the highest confidence scores, resulting in a subset U^t_t.
6. **Update Unlabeled Set and Target Labels:**
   - Remove instances in U^t_c and U^t_t from U_t.
   - Add the selected high-confidence instances U^t_c to L_t and assign their predicted labels.
7. **Repeat Steps 2–6** until the desired performance is achieved on the development dataset.

3.6. Domain invariant and specific classifier. This model's training is divided into two parts, with one for the domain invariant classifier, Fc, and one for the domain-specific classifier, The training objective for Fc is to minimize the following failure concerning parameters

$$\Theta = \Theta_e^c, \Theta_e^c, \Theta_d^t, \Theta_c^t \tag{3.16}$$

$$L = L_{recon}(\Theta_e^c, \Theta_e^t, \Theta_d^t) + \alpha L_c(\Theta_e^c, \Theta_c^t) + \gamma L_{sim}(\Theta_e^c) + \lambda L_{diff}(\Theta_e^t) \tag{3.17}$$

where the weights α, γ, and λ control how the words of loss are connected.$L(\Theta)$ denotes that throughout training on the parameters, failure, L, is balanced. Moreover, Lc indicates that the domain's invariant representation is not classified. Described by the ground-truth class 's negative log-likelihood for instances of both source domain and target domain

$$L_c = \frac{1}{n_s + l_t} \sum_{i=1}^{n_s} -Y_s^i log F_e(Y_s^i \Bigg| E_c(L_s^i)) + \frac{1}{n_s + l_t} \sum_{i=1}^{L_t} -Y_t^i log F_e(Y_t^i \Bigg| E_c(L_t^i)) \tag{3.18}$$

The dynamic number of target data tagged in each iteration is shown by Y I t, which is the one-hot encoding of the class label for the source in the example. The training challenge for Ft is to minimize the subsequent parameter failure.

$$\Theta = \Theta_e^c, \Theta_e^c, \Theta_d^t, \Theta_c^t \tag{3.19}$$

$$L = L_{recon}(\Theta_e^c, \Theta_e^t, \Theta_d^t) + \beta L_c(\Theta_e^c, \Theta_c^t) + \gamma L_{sim}(\Theta_e^c) + \lambda L_{diff}(\Theta_e^t) \tag{3.20}$$

where the weights $\gamma and \lambda$ correspond to the classifier's weights, the weight β and Fc control the classification failure portion. In contrast, Lt is the domain-specific representation based on the target domain's negative log-likelihood of the ground-truth class.

$$L_t = \frac{1}{l_t} \sum_{i=1}^{l_t} -Y_t^i log F_t(Y_t^i | E_t(L_t^i)) \tag{3.21}$$

The cosine of the angle formed by the two vectors that represent the lexical items u and v is what this represents.

$$\tau(v, u) = \frac{\sum w\epsilon\Gamma(v) f(u, w)}{\|u\| \, \|v\|} \tag{3.22}$$

$$\|v\| = \sqrt{\sum w\epsilon\Gamma(u)(f(v, w))^2}, \|u\| = \sqrt{\sum w\epsilon\Gamma(u) > 0(f(u, w))^2} \tag{3.23}$$

Here, $\Gamma(v) = \{x | f(v, x) > 0\}$ is the collection of features (x) in the feature vector for element v that have positive pmi values. Cosine similarity is a commonly utilized relatedness metric in many natural language processing tasks. We cluster related terms using Lin's proposed similitude measure. For word clustering tasks, this metric has been demonstrated to perform better than various other comparisons. Computed in the manner shown below:

$$\tau(v, u) = \frac{\sum w\epsilon\Gamma(v) \cap \Gamma(u)(f(v, w) + f(u, w))}{\sum w\epsilon\Gamma(v)(f(v, w) + \sum w\epsilon\Gamma(u) f(u, w)} \tag{3.24}$$

Last equation defines this measure of relatedness, which is the one put out in this study. Similar to Lin's estimate of similarity and Cosine similarity, this relatedness estimate is asymmetric.

We build a baseline relatedness measure in the last equation by swapping the two arguments, u and v, to show the asymmetric existence of the connection measure presented in previous equations. The following formula accurately determines the inverted baseline:

$$\tau(v, u) = \frac{\sum w\epsilon \{x | f(u, x) > 0\} f(u, w)}{\sum w\epsilon \{x | f(v, x) > 0\} f(u, w)} \tag{3.25}$$

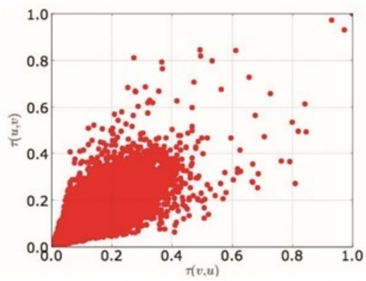

Fig. 3.2: Correlation between relatedness scores

Table 4.1: Equally balanced between positive and negative.

Review Type	Description
Positive	Reviews expressing positive sentiment
Negative	Reviews expressing negative sentiment
Unlabelled	Reviews without assigned sentiment
Processed	Reviews that have undergone preprocessing
Balanced	Preprocessed reviews with balanced labels

Remember that this baseline gives higher relative scores to expansion candidates, commonly included in user feedback, since the denominator consists of the sum of point-wise reciprocal knowledge values for terms co-occurring with correlation sources shown in figure 3.2.

Using the relatedness measure, we create a sentiment-sensitive thesaurus by listing lexical items v that co-occur with u 4 ($i.e., f(u,v) > 0$) for each lexical element u in descending order of the relatedness values $\tau(v,u)$.

4. Experiment model.

4.1. Dataset. This work evaluates the suggested approach with alternative sentiment classification models using the multi-domain sentiment dataset from Amazon product reviews. The dataset comprises product reviews of mobiles, kitchen sets, books, and electronics. Each review has a rating of (0 to 5 stars). The reviewers are transferred into positive, negative, and moderate labels. The product review dataset contains labeled and unlabelled data; the proposed framework extracts the features to determine the aspect levels using polarity and pivot models. In this experiment, one or more other domains serve as sources, and we choose each domain as the target domain. Reviews of the source domain and destination domain differ according to the data records they pertain to. To implement proposed framework, we have designed a framework using Python based packages, we adopted the machine learning packages to determine the results we divide the data into the three models, train (60%), validation (20%) and test (20%). The dataset samples and experiment were conducted at the KL University Hyderabad data center.

4.2. Results and Discussion. Figure 4.1 depicts the sentiment classifier for different electronic source domains. The accuracy is high for the kitchen domain for a single source. The accuracy is high for mobile and kitchen domains when two sources are combined, but when all three source domains are integrated, the best accuracy is obtained.

Figure 4.2 depicts the classification accuracy of target-labeled data against many source domains.

Figure 4.3 depicts that SU+ and TU+ denote both source and target domains for unlabelled data.SU- and TU+ denote only a Target domain exists. The higher accuracy is achieved when source and target unlabelled data is used, and poor performance is achieved when source and target unlabelled data is not used.

The sentiment classification for various target domains using the Adaptive Cross Domain Polarity aspect

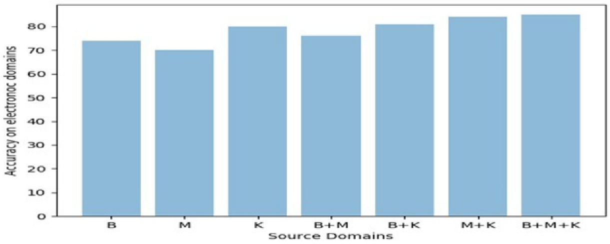

Fig. 4.1: Accuracy on the effect of multiple source domains

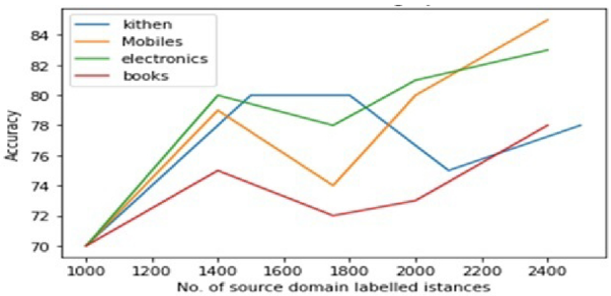

Fig. 4.2: Effect on source domain labeled data

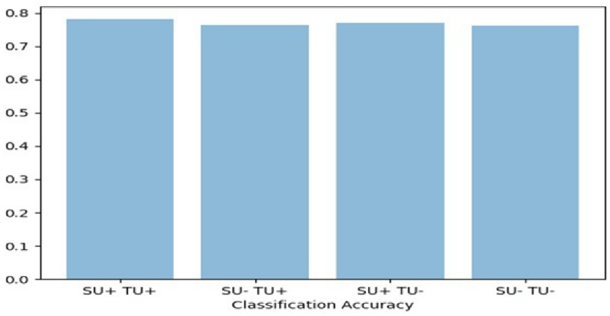

Fig. 4.3: Impact of Unlabelled Data's Source and Target Domains

Table 4.2: Sentiment classification for various target domains

Author	Technique	+ve Learning Expression Rate	Accuracy	Recall	F-score
Murugappan et.al [13]	DWT and KNN	45.87	82.32	80	81
Taran et.al [19]	CIF and MC-LS-SVM (MH)	67.85	86	76	78.1
Krishna et.al [20]	TQWT and ELM (MH)	78.65	87.1	80.1	82.5
Bajaj et.al [21]	FAWT and KNN	87.43	86.1	85.9	83.1
Proposed	(ACDPSNet)	97.76	98.76	97.85	96.94

level is shown in Figure 4.4 above and table 4.2.

Figure 4.5 briefly explains the performance measures of the proposed model compared with other models.

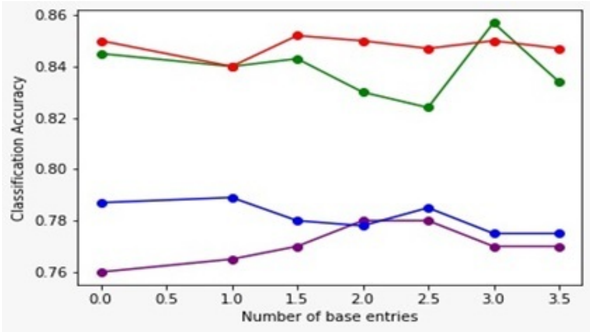

Fig. 4.4: Adaptive Cross-domain Polarity aspect-level sentiment quantification and categorization performance.

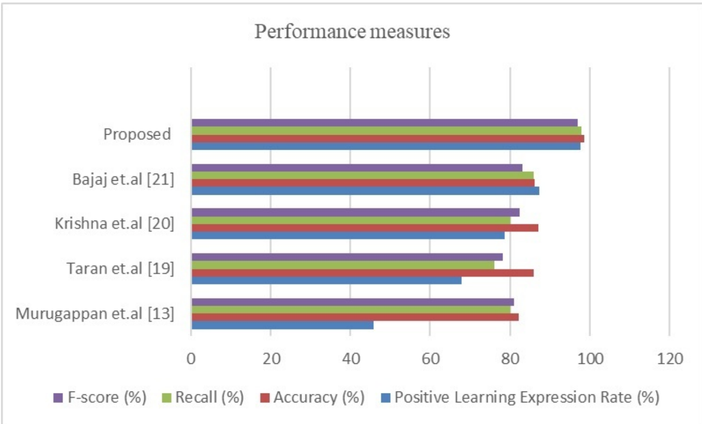

Fig. 4.5: Performance measures

5. Conclusion. Our research highlights the importance of incorporating domain-specific knowledge into domain adaptation tasks for sentiment classification and quantification. The proposed Adaptive Cross-Domain Polarity-Specific Network (ACDPSNet) demonstrates that by integrating polarity-specific sentiment annotations into semantic vectors and utilizing both labeled and unlabeled data from multiple domains, we can significantly enhance the model's ability to adapt to new domains. Our approach effectively addresses the common challenges of feature mismatch in cross-domain sentiment analysis, achieving notable improvements in accuracy, recall, throughput, and positive learning expression rate compared to existing methods. The findings of this study suggest that domain-specific information, often overlooked in favor of domain-invariant techniques, can be crucial for improving performance in domain adaptation scenarios, even when in-domain labeled data is sparse. This represents a significant shift from traditional methods that rely heavily on domain-invariant features. By successfully leveraging domain-specific information, our approach not only improves the robustness of sentiment classification models but also offers a scalable solution for various domain adaptation challenges. Future work could explore the application of this framework to other complex cross-domain tasks, further refining the adaptive mechanisms and exploring the integration of additional contextual factors. The promising results of this research indicate that ACDPSNet has the potential to be a valuable tool in the development of more adaptable and accurate sentiment analysis systems across diverse domains.

REFERENCES

[1] B. Liu, *"Sentiment analysis and opinion mining,"* , Synthesis Lectures on Human Language Technologies, pp. 1–167, 2012.

[2] P. D. Turney, *"Thumbs up or thumbs down? semantic orientation applied to unsupervised classification of reviews,"*, in ACL 2002, 2002, pp. 417–424.

[3] B. Pang and L. Lee, , *"Opinion mining and sentiment analysis,"* Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, pp. 1–135, 2008.

[4] Y. Lu, C. Zhai, and N. Sundaresan, , *"Rated aspect summarization of short comments,"* in WWW 2009, 2009, pp. 131–140.

[5] T.-K. Fan and C.-H. Chang, , *"Sentiment-oriented contextual advertising,"* Knowledge and Information Systems, vol. 23, no. 3, pp. 321–344, 2010.

[6] M. Hu and B. Liu, , *"Mining and summarizing customer reviews,"* in KDD 2004, 2004, pp. 168–177.

[7] J. Blitzer, M. Dredze, and F. Pereira, , *"Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification,"* in ACL 2007, 2007, pp. 440–447.

[8] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, , , *"Cross-domain sentiment classification via spectral feature alignment,"* in WWW 2010, 2010.

[9] H. Fang,, , *"A re-examination of query expansion using lexical resources,"* in ACL 2008, 2008, pp. 139–147.

[10] G. Salton and C. Buckley, , *"Introduction to Modern Information Retreival. McGraw-Hill Book Company, 1983.*

[11] D. Shen, J. Wu, B. Cao, J.-T. Sun, Q. Yang, Z. Chen, and Y. Li, , , *"Exploiting term relationship to boost text classification,"* in CIKM'09, 2009, pp. 1637 – 1640.

[12] T. Briscoe, J. Carroll, and R. Watson, , *"The second release of the rasp system,"* in COLING/ACL 2006 Interactive Presentation Sessions, 2006. https://www.cs.jhu.edu/ mdredze/datasets/sentiment/

[13] K. Saikumar and V. Rajesh, *A machine intelligence technique for predicting cardiovascular disease (CVD) using Radiology Dataset*, International Journal of System Assurance Engineering and Management, vol. 15, no. 1, pp. 135–151, 2024.

[14] D. S. Reddy, K. S. Charan, S. K. Kayam, and P. Rajalakshmi, *Robust Obstacle Detection and Collision Warning for Autonomous Vehicles Using Autoware Universe*, in 2024 16th International Conference on Computer and Automation Engineering (ICCAE), IEEE, March 2024, pp. 378–384.

[15] S. S. Pillalamarri, K. Saikumar, S. A. Al-Ameedee, S. R. Abdul Kadeem, and M. J. Hussein, *An Open CV-Based Social Distance Violation Identification Using a Deep Learning Technique*, in Micro-Electronics and Telecommunication Engineering: Proceedings of 6th ICMETE 2022, Singapore: Springer Nature Singapore, 2023, pp. 611–620.

[16] M. Wankhade, A. C. S. Rao, and C. Kulkarni, *A survey on sentiment analysis methods, applications, and challenges*, Artificial Intelligence Review, vol. 55, no. 7, pp. 5731–5780, 2022.

[17] D. M. E. D. M. Hussein, *A survey on sentiment analysis challenges*, Journal of King Saud University-Engineering Sciences, vol. 30, no. 4, pp. 330–338, 2018.

[18] A. Yadav and D. K. Vishwakarma, *Sentiment analysis using deep learning architectures: a review*, Artificial Intelligence Review, vol. 53, no. 6, pp. 4335–4385, 2020.

[19] A. M. Ramadhani and H. S. Goo, *Twitter sentiment analysis using deep learning methods*, in 2017 7th International Annual Engineering Seminar (InAES), IEEE, August 2017, pp. 1–4.

[20] W. Souma, I. Vodenska, and H. Aoyama, *Enhanced news sentiment analysis using deep learning methods*, Journal of Computational Social Science, vol. 2, no. 1, pp. 33–46, 2019.

[21] P. Vateekul and T. Koomsubha, *A study of sentiment analysis using deep learning techniques on Thai Twitter data*, in 2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), IEEE, July 2016, pp. 1–6.

[22] J. Duchi, E. Hazan, and Y. Singer, *Adaptive subgradient methods for online learning and stochastic optimization*, The Journal of Machine Learning Research, vol. 12, pp. 2121–2159, 2011.

[23] Y. Jin, H. Zhang, and D. Du, *Improving deep belief networks via delta rule for sentiment classification*, in Proceedings of 2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI), 2017, pp. 410–414.

[24] K. Srinivas Rao, D. V. Divakara Rao, I. Patel, K. Saikumar, and D. Vijendra Babu, *Automatic prediction and identification of smart women safety wearable device using Dc-RFO-IoT*, Journal of Information Technology Management, vol. 15, Special Issue, pp. 34–51, 2023.

[25] N. J. Bommagani, A. Venkataramana, R. Vemulapalli, T. R. Singasani, A. K. Pani, M. B. Challageri, and S. Kayam, *Artificial Butterfly Optimizer Based Two-Layer Convolutional Neural Network with Polarized Attention Mechanism for Human Activity Recognition*, Mathematical Modelling of Engineering Problems, vol. 11, no. 3, 2024.

Edited by: Dhilip Kumar V
Special issue on: Unleashing the power of Edge AI for Scalable Image and Video Processing
Received: Aug 13, 2024
Accepted: Nov 19, 2024