ITC 1/52
Information Technology and Control
Vol. 52 / No. 1 / 2023
pp. 100-110
DOI 10.5755/j01.itc.52.1.32119

Improved Feature Representation Using Collaborative Network for Cross-Domain Sentiment Analysis

Received 2022/08/22 | Accepted after revision 2022/11/07

https://doi.org/10.5755/j01.itc.52.1.32119

HOW TO CITE: Gunasekar, M., Thilagamani, S. (2023). Improved Feature Representation Using Collaborative Network for Cross-Domain Sentiment Analysis. *Information Technology and Control*, 52(1), 100-110. https://doi.org/10.5755/j01.itc.52.1.32119

Improved Feature Representation Using Collaborative Network for Cross-Domain Sentiment Analysis

M. Gunasekar
Department of Information Technology, Kongu Engineering College, Erode-638060, India

S. Thilagamani
Department of Computer Science and Engineering, M. Kumarasamy College of Engineering, Karur-639113, India

Corresponding author: mgunasekar1@outlook.com

Sentiment Analysis task helps us to estimate the opinion of a person from his reviews or comments about a product, person, politics, etc., Cross-Domain Sentiment Analysis (CDSA) empowers the Sentiment models with the ability to forecast the opinion of a review coming from a different domain other than the domain where the model is trained. The challenge of the CDSA model relies on bridging the relationship between words in the source and target domain. Several types of research in CDSA focus on determining the domain invariant features to adapt the model to the target domain, such model shows less focus on aspect terms of the sentence. We propose CWAN (Collaborative Word Attention Network), which integrates aspects and domain invariant features of the sentences to calculate the sentiment. CWAN uses attention networks to capture the domain-independent features and aspects of the sentences. The sentence and aspect attention models are executed collaboratively to determine the sentiment of the sentence. Amazon product review dataset is used in this experiment. The performance of the CWAN model is compared with other baseline CDSA models. The results show that CWAN outperforms other baseline models.

KEYWORDS: Cross Domain Sentiment Analysis, Domain invariant features, Attention Network, Bi-LSTM, Aspect-based sentiment analysis.

1. Introduction

Sentiment analysis task is carried out by companies to measure the opinion of their products in the market using customer reviews. It allows people to make significant decisions on their business products. Social media and the customer feedback corner of a website capture a massive volume of data each day. Synthesizing those data manually is challenging; that is where sentiment analysis comes into light. Sentiment analysis automates the opinion mining process and eases the task. A fair amount of opinioned data is required to train a sentiment analysis model. Most of the works in sentiment analysis intend to identify opinions for data on the domain in which the model is trained. The sentiment classifier trained on a particular domain shows high accuracy in sentiment prediction on the same domain than a different domain dataset. Cross-Domain sentiment suits the situation where the labeled data in the target domain is less or not available.

The cross-domain classifier has to learn and understand the aspect terms in the reviews for better sentiment prediction. Many research works in the literature focuses on learning the shared features across the domains. Yang [20] introduced transfer learning for domain adaptation. The similar words in source and target domain are utilized in transfer learning to make the model adapt to the target domain. Blitzer [5] proposed structural correspondence learning to learn domain specific features from the sentences. Pan[19] used mutual information score to distinguish between domain specific and domain invariant features. In the literature [7], [21] and [25] many methods are formulated to address the domain adaptation using attention networks. Attention transfer networks [14], [29]are applied to learn common features between the source and target domain using attention networks.

Concentrating only on the shared features might work well for source domain label prediction but not on the target domain. The proposed model intends to enhance the accuracy of the sentiment classifier by improving the feature representation for classification tasks. We propose CWAN (Collaborative word attention network), which learns the sentence and aspect representation using a collaborative word attention mechanism. We hypothesize that the collaborative learning between the sentence and aspect vector provides a better sentiment feature representation. The proposed model addresses the problem of domain adaptation by learning the domain independent feature representation to bridge the gap between the domains on which model is trained and tested. The experiments are executed by training the model using the source dataset and making it adapt to the target dataset.

The previous studies on cross-domain sentiment analysis proclaim that domain invariant features could be extracted with the help of domain variant features from the source and target domains. However, the previous works did not consider the aspect terms which indirectly signify the sentiment in the sentence. Consider the example review from the Electronics domain and Kitchen domain in Table 1. The reviews have one common aspect, “design” the sentiment word expressed on this aspect is positive, but the overall sentiment of reviews 1 and 2 is Negative and Positive, respectively. This shows that in review 1, aspect “Design” has a low weight compared to another aspect, “Memory”. In review 2 the weight of aspect “design” is high compared to another aspect, “rack space”. From this example, it is evident that the importance of the aspect may differ in the same domain based on the weights attained by each aspect. The aspects shared between the domains are highly utilized to identify the domain invariant features and help us in the transfer learning process.

The proposed method consists of two attention networks. One is to identify the common features between the domains. The second one utilizes the common features across the domains to extract the

Table 1

Example of customer reviews with aspect words

	Review 1	Review 2
Reviews	“The design of the mobile is good, but the Memory is too low”	“The design of the Utensils holder is good and the rack space is not bad”
Domain	Electronics	Kitchen
Aspects	Design, Memory	Design, Rack Space
Sentiment	Negative	Positive

information from aspects. The aspects from both source and target domain are extracted. We Propose CWAN, which absorbs the hidden state representation of the sentence and aspect vector and jointly learns the new vector representation of the sentence. The new vector representation is prepared by using mean pooling on sentence and aspect vector to create the new sentence and aspect representation. This representation is rich in sentiment awareness and helps in predicting the sentiment of the sentence.

2. Review of Literature

Sentiment analysis may usually be split into classes i.e. positive and negative based on the sentiment classification. The approaches for analyzing textual sentiment are largely divided into machine-based and lexicon based learning. The approach based on Lexicons uses primarily the lexicon to retrieve key terms in the corpus of emotional language.

Mohammad et al. [18] utilizing crowdsourcing to develop emotional features for sentiment analysis to forecast emotions. Xing et al. [27] proposed a lexicon-based method that trains the model initially using raw sentiment classifier and the model is further trained cognitively for life-long learning from wrongly predicted sentences. The approach traces the wrongly expected sentences and imitates the lifetime dictionary learning as supervision. In recent years, the machine learning approach has become a common study rules. This approach for text-sentiment analysis is developed by modeling training and the text-sentiment analysis is then conducted. Guna et al. [9] use Naïve Bayes (NB), and Support Vector Machine (SVM) to identify and investigate the sentiment analysis using ensemble feature selection methods. Desai et al. [15] investigate approaches for sentiment analysis using machine learning methods.

The word vector referred is a low-dimensional vector and a similar text search field, which corresponds to the word distribution. The term vector represents the point in an N-dimensional space and it provides the relationship between the data points or semantics of the terms. In addition, the capacity of the neural network to interpret structural knowledge also opens the door for word vectors to develop [4], [6], [13] and applications.

Analysis of text sentiments entails the issue of information sharing between various realms. Meng et al. [16] proposes a cross-domain textual sentiment using a multi-layer CNN that does share the weights of the features from both the domains. In the cross-domain text-like analysis, Huang et al. [13] compare NB, SVM and expectation maximization (EM). The combination features of Xia et al. [26] were used with the NB, EM and SVM approaches as part of the speech and word relationship and improved study outcomes than the standard one-channel learning methods are used. Tang et al. [24] have previously studied uses of profound research techniques for sentimental interpretation which are superior to conventional approaches for classifying sentiment, views and emotional dictionaries.

Ashraf et al. [2] created an ensemble model using LSTM-GRU to analyse the sentiment of the cryptocurrency tweets. Using a deep learning process for modeling sentences Yu et al. [16] prove that the deep learning is superior to the conventional learning model. As the semantic representation in various fields can best be disclosed and obtained by machine learning, it usually is superior to the graph model and the statistical learning models.

Aspect based models provide more intuitive word representation such that the sentence vectors are associated learned jointly with the aspect learning. Tang et al. [23] generated a deep memory network with multiple convolutional layer to estimate the context words while categorizing aspect terms from the sentence. Attention networks are utilized to capture the aspect terms. Huang et al. [10] proposed a novel Attention-over-Attention (AOA) model that interactively learns the aspect terms and sentence representation. AOA also focus on contextual learning to improve the vector representation. Rietzler et al. [22] created an Aspect Based Sentiment Analysis (ABSA) model involving detection and classification task. First step is to detect aspect category and aspect target. The second step is to classify the sentiment label on aspect category and aspect target.

Li et al. [12] proposed a novel SAL (Selective Adversarial Learning) model which learns the word correlation automatically. The weights of the word are perceived dynamically such that word with highest weight is considered to be the significant word in the sentence. Gong et al. [8] created a model that jointly learns an instance based adaptation and feature vector representation. Pre-trained BERT model is adopted in his work to capture the domain independent features. Zhang et al. [28] proposed a transfer learning model that comprises of two tasks, sentiment classification and domain invariant feature identification. The two tasks are executed simultaneously to infer the context aware domain representation. Ashraf et al. [3] developed a model to identify the best calling apps from google play store using app reviews.

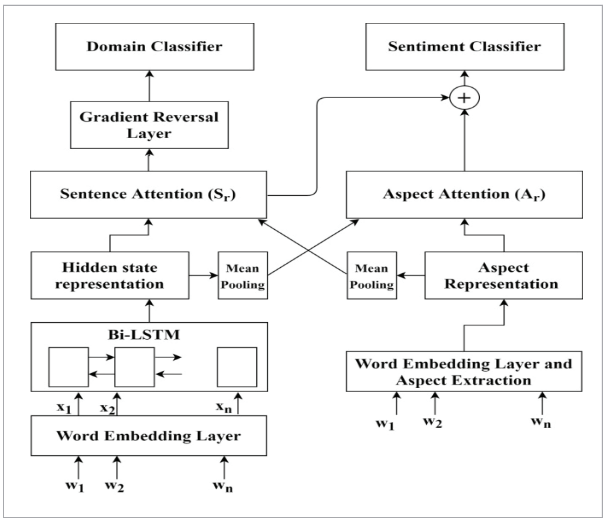
However, machine learning methods allow instruction and data testing in distributed manner. The conventional research has shown that for sentimental analysis, the supervised classification approach is successful. The texts are typically classified into three groups according to various granularities. The established classified data does not in most cases fall into the same domain as the data to be assessed, which obviously reduces the efficiency of the supervised classification algorithm, leading to cross-domain analysis.

3. Methodology

The proposed work focuses on generating aspect aware feature vector by jointly learning the domain invariant features and aspect features using attention networks. The architecture of proposed work is depicted in the Figure 1. Cross-domain sentiment analysis gains its attention from its ability to predict sentiment with less knowledge on the target domain.

Figure 1

Architecture of CWAN



Aspect attention is an important factor to understand the real context of a review. In the proposed work the hidden state representation of aspect words are captured using Bi-GRU. The hidden state representation for the sentence is learned using Bi-LSTM. Then the sentence and aspect vectors are learned collaboratively using mutual learning based on the mean pooling value of the sentence and aspect vectors. Gradient reversal layer is used to learn the common features from both domains. As a result of GRL the training process is reversed and the domain classifier can separate the domain invariant features from the sentence.

3.1. Word Encoding

The input representation for a deep learning model has to be numerical vectors. To convert text to a numerical vector representation, we use the Embedding layer. The Encoding process converts each word from a review into a numerical vector. Given a review S, containing words $\{w_1, w_2, w_3, \ldots w_n\}$ where w is word drawn from vocabulary V and n is the length of the review. The value of n is made constant for all reviews in the corpus.

3.2. Hidden State Representation Using Bi-LSTM

The words are transferred into embeddings $E = \{x_1, x_2, x_3, \ldots x_n\}$. Since RNN is prone to gradient vanishing problem, we prefer Bi-LSTM over RNN to learn hidden state representation for sentence embeddings E. Bi-LSTM uses cell state to store the data for longer time to make the hidden state learning deep and intuitive.

Bi-LSTM converts the output vector of word embedding layer E into a hidden state representation $H=\{h_1, h_2, h_3, \ldots h_t\}$ where t ranges from 1 to n. The hidden state representation at h_t contains the information of h_{t-1} and its antecedents $h_{t-2}, h_{t-3}, \ldots h_1$.

$$i_t = \delta(W_{ip}x_t + W_{hi}h_{t-1} + b_{ip}) \tag{1}$$

$$f_t = \delta(W_{fi}x_t + W_{hfi}h_{t-1} + b_{fi}) \tag{2}$$

$$c_t = f_t c_{t-1} + i_t \tau(W_{ct}h_{t-1} + b_{ct}) \tag{3}$$

$$o_t = \delta(W_{op}x_t + W_{op}h_{t-1} + b_{ot}) \tag{4}$$

$$h_t = o_t \tanh(c_t) . \quad (5)$$

W_{ip} and b_{ip} are weight and bias for input gate i_t, W_{ft} and b_{ft} are weight and bias of forget gate f_t and W_{op} and b_{ot} are weight and bias of output gate o_t. c_t maintains the cell information. h_t contains the hidden state representation for each reviews in the corpus.

3.3. Aspect Extraction Using Bi-GRU

Utilizing the Glove (pre-trained model) we convert each review from our corpus into an Embedding vector with fixed dimensions. Aspect extraction is a sequence tagging method; we adopt Bi-GRU to learn relation among the words in the reviews. The aspect learning method utilizes transfer learning to capture the relationship between a word and its adjoining words. The word sequence from the reviews is fed into Bi-GRU_{AE} to create a relation among word vector representation.

$S=\{ w_1, w_2, w_3,....w_n\}$

$S_E=\{x_1,x_2,x_3,....x_n\}$

$S_a=$ Bi-$GRU_{AE}(S_E)$

S_a is a relation aware vector generated using Sentence embedding S_E using Glove model with Bi-GRU. We use conditional random field (CRF) to model the dependence between the consecutive words and aspect-labels

3.4. Collaborative Learning

The interactive learning of sentence and aspect representation guides the classifier in sentiment prediction. To reduce the size of features and uphold only the significant features we adopt a non-linear down sampling method. We use mean pooling to do down sampling the hidden state and aspect representations. The mean pooling mechanism constructs H_{sp} sentence pooling and H_{ap} aspect pooling vector using following equations classify the sentiment and domain

$$H_{sp} = \sum_{i=1}^{n} \frac{H_i}{n} \quad (6)$$

$$H_{ap} = \sum_{i=1}^{m} \frac{s_a^i}{m} . \quad (7)$$

Attention mechanism is adopted to highlight the important words in the input sentence. The input words are assigned with weights at each step. To perform interactive learning, the pooling vectors H_{sp} and H_{ap} are added to S_a (relation aware Aspect vector) and H (hidden state representation of review) respectively. The hidden state sentence vector is updated as,

$H=\{h_1,h_2,h_3,...h_t, H_{ap}\}$

$S_a=\{h_1,h_2,h_3,...h_t, H_{sp}\}$

3.5. Sentence Representation

To obtain the sentence attention vector α_i we adopt the following equation,

$$\alpha_i = \frac{\exp(\gamma(H_i))}{\sum_{j=1}^{n+1} \exp(\gamma(H_j))} . \quad (8)$$

Score function for sentence attention vector is calculated using the following equation,

$$\gamma(H_i) = \tanh(H_i W_s + b_s) , \quad (9)$$

where W_s is weight matrix and b_s is bias matrix. α_i improves the explanatory ability of the model and it helps us in drawing out the high sentiment score word [17]. Attention vector also eases the cross domain transfer learning. The computation of final sentence representation is done using the sentence attention vector α_i as follows,

$$H_r = \sum_{i=1}^{n+1} \alpha_i H_i . \quad (10)$$

3.6. Aspect Representation

The same method is followed to obtain the aspect attention vector β_i,

$$\beta_i = \frac{\exp(\lambda(H_{ap}, H_{sp}))}{\sum_{j=1}^{n+1} \exp(\lambda(H_{ap}, H_{sp}))} . \quad (11)$$

Score function for aspect attention vector is calculated using the following equation,

$$\gamma(H_{ap}, H_{sp}) = \tanh(H_{ap} H_{sp} W_s + b_s) , \quad (12)$$

where W_s is weight matrix and b_s is bias matrix. β_i identifies the sentiment and domain invariant features by combining the sentence and aspect pooling vector [11]. The final aspect representation is calculated using the following equation,

$$A_r = \sum_{i=1}^{m} \beta_i S_a^i \quad (13)$$

3.7. Domain Classifier

Domain classification and sentiment classification task are executed parallel to reduce the error in classification. The domain classifier utilizes GRL to extract the domain invariant features from source and target domain [1]. Gradient Reversal layer (GRL) to reduce the divergence between two domains. The domain classifier discriminates between the source and target domain. It makes use of the entire corpus from source domain E_s and target domain S_t to carryout domain invariant and domain specific feature representation task. The feature representation vector from GRL is assigned to softmax layer for domain classification

$$y_d^{'} = soft\max(W_d H_r + b_a)\,. \quad (14)$$

3.8. Sentiment Classifier

The feature representation generated using the aspect and sentence awareness performs classification more intuitively. The weights of the aspects are learned to improve the significance of domain specific words in the sentence. The sentiment analysis is performed after combining the sentence and aspect representation as follows,

$$y_s^{'} = soft\max(W_s[H_r \oplus A_r] + b_s)\,. \quad (15)$$

3.9. Dataset Pre-processing

To make a comparative study of our model with previous baseline models we adopt the standard Amazon product review dataset. Since there are different domains in the Amazon product review dataset we choose Books (B), DVD (D), Electronics (E) and Kitchen (K) datasets. For our experiment we consider 2000 positive and 2000 negative labeled reviews from each domain. Also we consider 2000 unlabeled data from each domain. The dataset is further divided as train and test set with 80 and 20 percent respectively.

We consider 2000 positive, negative and unlabeled reviews from Amazon product review dataset. The notion of cross domain sentiment analysis is to predict target domain by learning from source domain. Our model utilizes the labeled source domain data and unlabeled target domain data for training the model. The prediction accuracy of the model is checked with the labeled target domain data. Data pre-processing involves a series of steps to convert raw data into a useful data for our experiment. This process eradicates the absurd data from the dataset. The data pre-processing involves

- Removal of stop words from the corpus.
- Removal of punctuations, expressions, numerical tokens and special characters.
- Lemmatization - converting the words to its dictionary form.

Table 2
Amazon product review dataset

Dataset	Type	Positive Reviews	Negative Reviews	Unlabeled
Books	Amazon Product Reviews	2000	2000	2000
DVD		2000	2000	2000
Electronics		2000	2000	2000
Kitchen		2000	2000	2000

3.10. Implementation

In the proposed work the embedding layer take the reviews from the source and target domain dataset and creates a vector representation for the reviews. The sentence vector is converted into feature vector using bi-directional LSTM. The proposed approach extracts the aspects using the Bi-GRU aspect extraction. The attention mechanism is applied on the hidden state representation of the sentence and aspect vectors. Domain classifier is trained using both source and target domain dataset. Sentiment classifier is trained using source labeled dataset. Our work makes use of two separate learning tasks domain and sentiment classification. The Cross entropy loss function is utilized to train domain and sentiment classifier.

3.11. Algorithm: Aspect Feature Extraction and Classification

Input: Data D_s source and D_t target domain

// Data Pre-Processing:

Step 1: For data in D_s and D_t

Step 2: Remove absurd data

Step 3: End

// Word Embedding

Step 4: For each epoch=1 to max-epoch do

Step 5: For each review in D_s, D_t do

Step 6: Create Word Vector (E_i) using embedding layer

Step 7: Compute sentence Hidden state representation (H_i) using Bi-LSTM

Step 8: Compute Aspect Extraction and representation (AE_i) using Bi-GRU

Step 9: Calculate sentence attention vector α_i

Step 10: Calculate aspect attention vector β_i

Step 11: Compute Sentence representation $(H_r) = \sum_{i=1}^{n+1} \alpha_i H_i$

Step 12: Compute Aspect representation $(A_r) = \sum_{i=1}^{m} \beta_i S_a^i$

// Domain Classification:

Step 13: For each row in H_r do

Step 14: $y_d^{'} = soft\max(W_d H_r + b_a)$

Step 15: End

// Sentiment Classification:

Step 16: For each row in H_r and A_r do

Step 17: $y_s^{'} = soft\max(W_s[H_r \oplus A_r] + b_s)$

Step 18: End

Step 19: (Training: cross entropy loss function)

Step 20: End

Step 21: End

3.12. Hyperparameter Setting

The length of the sentences is fixed to 200 words. We use word2vec to create word embeddings and it is a 200-dimension vector. The dimensions of the Attention, LSTM vectors are 64 and 64 respectively. The Uniform distribution μ(-0.01,0.01) is adopted to initialize all weight matrices. The weight matrices are randomly initialized. Relu activation function is used. The learning rate is considered to be 0.001 and the l2 normalization is 0.0001. The number of iterations is set to 40 epochs. The progress of the training phase is monitored continuously to stop training when there is no significant improvement in the validation test sets. This strategy stops the training phase when there are no significant changes in the validation test sets after 15 epochs.

3.13. Baseline Models Compared

SCL-MI (Blitzer, Dredze, and Pereira 2007): In SCL-MI pivot features are identified using mutual information of word with the source label.

MSDA (Chen et al. 2012): This model learns its feature representation from a large corrupted input dataset. Domain adaptation is made through unsupervised learning on the union of the source and target dataset.

DANN (Ganin) has focused on the domain adaptation of the model using labeled source domain and unlabeled target domain datasets. Domain adapted feature alignment is possible for all neural networks which uses back propagation.

CNN-aux: provides two auxiliary tasks to learn hidden feature representation and the sentiment classification interactively.

AMN (Zheng) provides a joint learning method to dig out the domain independent features from the source and target domain. This approach reduces the discriminating characteristics between the features of the two domains.

4. Results and Discussion

The performance of CWAN model is assessed using the accuracy index of the confusion matrix. The general parameter to assess the efficiency of a model is the Accuracy of that model. Accuracy is calculated as the ratio of the correctly predicted samples to the total samples. Parameters of the confusion matrix are,

TP: Computes the total amount of positive samples predicted correctly

TN: Computes the total amount of negative samples predicted correctly

FP: Computes the total amount of negative samples predicted as positive

FN: Computes the total amount of positive samples predicted as negative

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (16)$$

$$Precision = \frac{TP}{TP + FP} \quad (17)$$

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

$$F1score = 2 * \frac{Precision * Recall}{Precision + Recall}. \quad (19)$$

In our experiment, training and testing of the model is carried out using source domain and target domain dataset respectively. Figure 2 shows the results of the model that is trained using DVD, Electronics and Kitchen dataset and is tested on Books dataset. The results illustrate that the model performs better on DVD domain dataset compared to other domain dataset.

Figure 2

Accuracy of adaptation between DVD, Electronics, Kitchen domains with Book domain

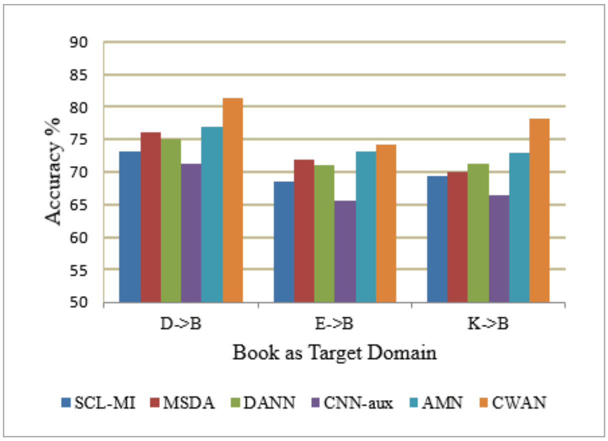


Figure 3

Accuracy of adaptation between Book, Electronics, Kitchen domains with DVD domain

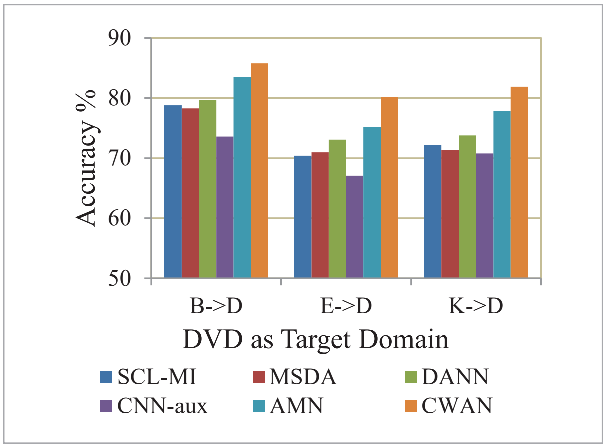
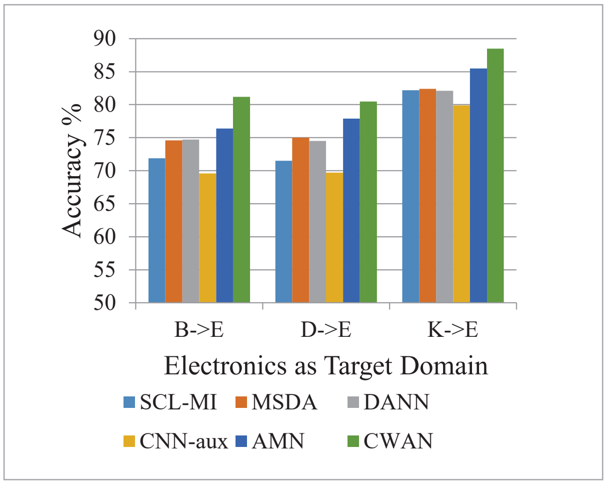


Figure 4

Accuracy of adaptation between Book, DVD, Kitchen domains with Electronics domain



Figures show the comparison of CWAN with other conventional models on Amazon product review dataset.

Figure 3 depicts the performance of the model that is trained using Books, Electronics and Kitchen dataset and is tested on DVD dataset. The model shows better accuracy when trained on Book source domain dataset compared to other domain dataset.

Figure 4 depicts the performance of the model that is trained using Books, DVD and Kitchen dataset and is tested on Electronics dataset. The model shows better accuracy when trained on Kitchen source domain dataset compared to other domain dataset.

Figure 5 depicts the performance of the model that is trained using Books, DVD and Electronics dataset and is tested on Kitchen dataset. The model shows better accuracy when trained on Electronics source domain dataset compared to other domain dataset.

From the results it is evident that the dataset combinations Book – DVD and Electronics – Kitchen shares more common features. While observing the performance of CWAN on the pairs of dataset B->D,

Figure 5

Accuracy of adaptation between Book, DVD, Electronics domains with Kitchen domain

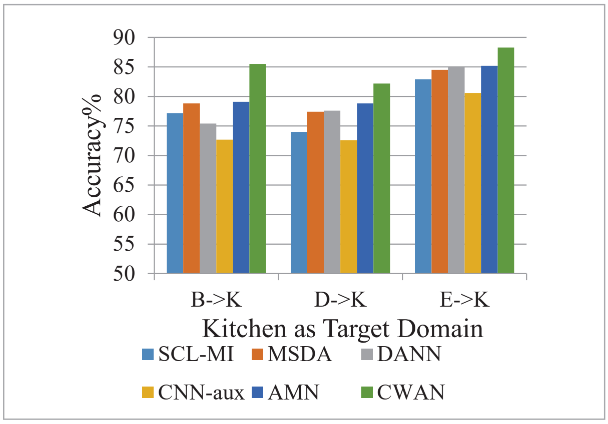


Figure 6

Average Accuracy of the models

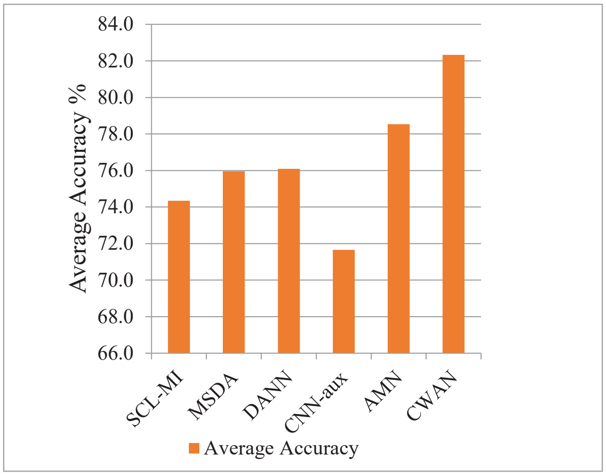
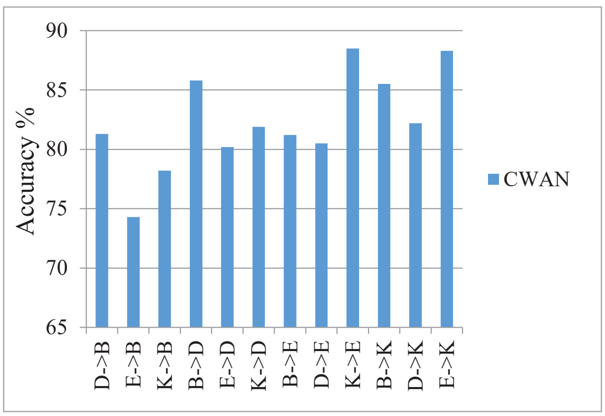


Figure 7

Accuracy of different combination of dataset



D->B, E->K and K->E the accuracy of the model is above 80%.

From Figures 8-10 it is evident that the model trained on source domain kitchen and tested on target domain Electronics and vice versa has shown better Precision, Recall and F1 measure.

When comparing the performance of CWAN with other baseline models, the accuracy of CWAN model out-

Figure 8

Precision of CWAN across 4 datasets

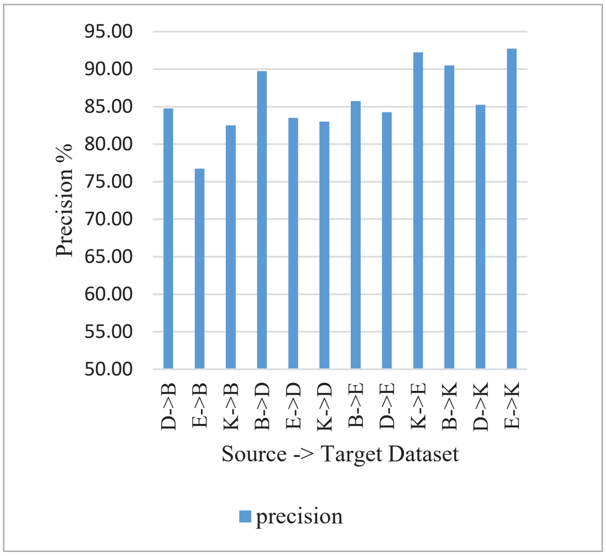


Figure 9

Recall of CWAN across 4 datasets

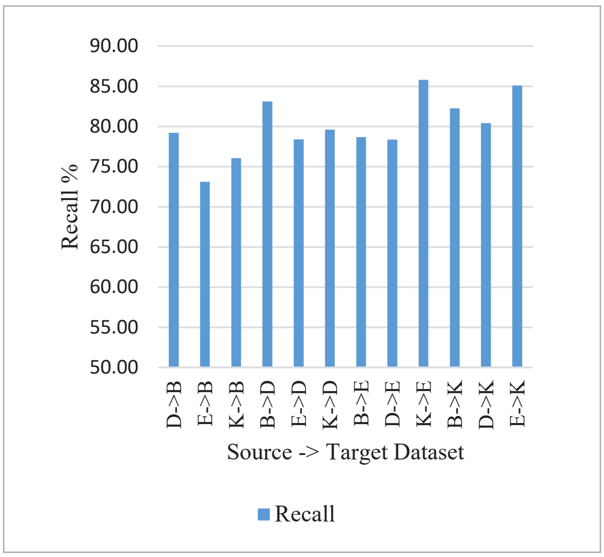
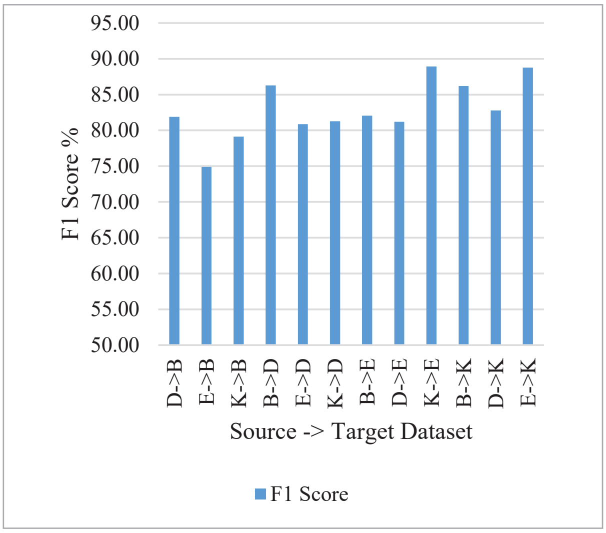


Figure 10
F1 Score of CWAN across 4 datasets



perform other models. The accuracy of all the baseline models and CWAN model are averaged and represented in Figure 6. The graph shows that the average accuracy of CWAN is higher compared to other baseline models average accuracy. The CWAN model is tested on different combinations of source and target domain datasets. From Figure 7 it is evident that the dataset combination Kitchen and electronics shows higher accuracy than any other dataset combinations.

5. Conclusion

The proposed CWAN model uses two attention networks that are collaboratively executed to learn representation for domain and sentiment classifier. The sentence attention network generates new representation by utilizing the hidden state representation of sentence and the mean pooling on the aspect attention vector. Aspect attention network creates aspect representation from the hidden state representation of the aspects and the mean pooling on the sentence attention vector. The new sentence and aspect vectors have sentiment and domain influential information respectively. The experiment illustrates that the combination of different datasets produce different results. It is also shown that the proposed model accuracy is higher comparing to other baseline models considered. Though there are different combinations of source and target domain datasets are tested, Books->DVD and Electronics->Kitchen combination shows better results than other dataset combinations.

References

1. Abu-Shareha, Q. Y., Abualhaj, M. M. A Hotel Recommender System Based on Multi-criteria Collaborative Filtering. Information Technology and Control, 2022, 51(2), 390- 402. https://doi.org/10.5755/j01.itc.51.2.30701
2. Aslam, N., Rustam, F., Lee, E., Washington, P. B., Ashraf, I. Sentiment Analysis and Emotion Detection on Cryptocurrency Related Tweets Using Ensemble LSTM-GRU Model. IEEE Access, 2022 Apr 7, 10, 39313-39324. https://doi.org/10.1109/ACCESS.2022.3165621
3. Aslam, N., Xia, K., Rustam, F., Hameed, A., Ashraf, I. Using Aspect-Level Sentiments for Calling App Recommendation with Hybrid Deep-Learning Models. Applied Sciences, 2022 Aug 26, 12(17), 8522. https://doi.org/10.3390/app12178522
4. Bengio, Y., Ducharme, R., Vincent, P., A Neural Probabilistic Language Model, Advances in neural information processing systems, 2000, 13.
5. Blitzer, J., McDonald, R., Pereira, F. Domain Adaptation with Structural Correspondence Learning. COLING/ACL 2006 - EMNLP 2006 2006 Conf. Empir. Methods Nat. Lang. Process. Proc. Conf., 2006, 120-128. https://doi.org/10.3115/1610075.1610094
6. Demeester, T., Rocktäschel, T., Riedel, S., Lifted Rule Injection for Relation Embeddings, EMNLP 2016 - Conf. Empir. Methods Nat. Lang. Process. Proc., 2016, 1389-1399. https://doi.org/10.18653/v1/D16-1146
7. Gan, C., Wang, L., Zhang, Z., Wang, Z., Sparse Attention Based Separable Dilated Convolutional Neural Network for Targeted Sentiment Analysis, Knowledge-Based Syst., 2020, 188, 104827. https://doi.org/10.1016/j.knosys.2019.06.035
8. Gong, C., Yu, J. Unified Feature and Instance Based Domain Adaptation for Aspect-Based Sentiment Analysis, Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), 2020, 7035-7045. https://doi.org/10.18653/v1/2020.emnlp-main.572
9. Gunasekar, M., Thilagamani, S. Performance Analysis of Ensemble Feature Selection Method Under SVM and BMNB Classifiers for Sentiment Analysis. International Journal of Scientific & Technology Research, 2020, 9(2), 1536-1540.

10. Huang, B., Ou, Y., Carley, K. M. Aspect Level Sentiment Classification with Attention-over-Attention Neural Networks. International Conference on Social Computing, Behavioral-cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation, 2018, 197-206. https://doi.org/10.1007/978-3-319-93372-6_22

11. Kapočiūtė-Dzikienė, J., Damaševičius, R., Woźniak, M. Sentiment Analysis of Lithuanian Texts Using Traditional and Deep Learning Approaches. Computers, 2019, 8(1), 4. https://doi.org/10.3390/computers8010004

12. Li, Zheng, Xin Li, Ying Wei, Lidong Bing, Yu Zhang, and Qiang Yang. Transferable End-to-End Aspect-based Sentiment Analysis with Selective Adversarial Learning. International Joint Conference on Natural Language Processing, 2019, 4590-4600. https://doi.org/10.18653/v1/D19-1466

13. Li, J., Li, J., Fu, X., Masud, M. A., Huang, J. Z., Learning Distributed Word Representation with Multi-Contextual Mixed Embedding, Knowledge-Based Systems. 2016, 106, 220-230. https://doi.org/10.1016/j.knosys.2016.05.045

14. Li, Z., Zhang, Y., Wei, Y., Wu, Y., Yang, Q. End-to-End Adversarial Memory Network for Cross-Domain Sentiment Classification. International Joint Conferences on Artificial Intelligence Organization, 2017, 2237-2243. https://doi.org/10.24963/ijcai.2017/311

15. Mdesai, M., Mehta, M. A. Techniques for Sentiment Analysis of Twitter Data: A Comprehensive Survey. Proceeding - IEEE International Conference on Computing, Communication & Automation, 2016, 149-154. https://doi.org/10.1109/CCAA.2016.7813707

16. Meng, J., Long, Y., Yu, Y., Zhao, D., Liu, S. Cross-Domain Text Sentiment Analysis Based on CNN_FT Method. Information, 2019, 10(5). https://doi.org/10.3390/info10050162

17. Mir, J., Mahmood, A. Movie Aspects Identification Model (MAIM) for Aspect Based Sentiment Analysis. Information Technology and Control, 2020, 49(4), 564-582. https://doi.org/10.5755/j01.itc.49.4.25350

18. Mohammad, S. M., Zhu, X., Kiritchenko, S., Martin. J. Sentiment, Emotion, Purpose, and Style in Electoral Tweets. Information Processing & Management, 2015, 51(4), 480-499. https://doi.org/10.1016/j.ipm.2014.09.003

19. Pan, S. J., Ni, X., Sun, J. T., Yang, Q., Chen, Z. Cross-Domain Sentiment Classification via Spectral Feature Alignment. Proceedings of 19th International Conference on World Wide Web, WWW'10 2010, 751-760. https://doi.org/10.1145/1772690.1772767

20. Pan, S. J., Yang, Q., A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering, 2009, 22(10), 1345-1359. https://doi.org/10.1109/TKDE.2009.191

21. Qu, Z., Wang, Y., Wang, X., Zheng, S. A Transfer Learning Based Hierarchical Attention Neural Network for Sentiment Classification. International Conference on Data Mining and Big Data, 2018, 10943, 383-392. https://doi.org/10.1007/978-3-319-93803-5_36

22. Rietzler, A., Stabinger, S., Opitz, P., Engl, S. Adapt or Get Left Behind: Domain Adaptation Through BERT Language Model Finetuning for Aspect-Target Sentiment Classification. International Conference on Language Resources and Evaluation, 2019, 4933-4941.

23. Tang, D., Qin, B., Liu, T. Aspect Level Sentiment Classification with Deep Memory Network. 2014, arXiv preprint arXiv:1605.08900.

24. Tang, D., Qin, B., Liu, T. Deep Learning for Sentiment Analysis: Successful Approaches and Future Challenges. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2015, 5(6), 292-303. https://doi.org/10.1002/widm.1171

25. Wang, Y., Huang, M., Zhao, L., Zhu, X. Attention-Based LSTM for Aspect-Level Sentiment Classification, EMNLP 2016 - Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, 2016, 606-615. https://doi.org/10.18653/v1/D16-1058

26. Xia, R., Zong, C., Hu, X., Cambria, E. Feature Ensemble Plus Sample Selection: Domain Adaptation for Sentiment Classification. IEEE Intelligent Systems, 2013, 28(3), 10-18. https://doi.org/10.1109/MIS.2013.27

27. Xing, F. Z., Pallucchini, F., Cambria, E. Cognitive-Inspired Domain Adaptation of Sentiment Lexicons, Information Processing & Management, 2019, 56(3), 554-564. https://doi.org/10.1016/j.ipm.2018.11.002

28. Zhang, K., Liu, Q., Qian, H., Xiang, B., Cui, Q., Zhou, J. EATN : An Efficient Adaptive Transfer Network for Aspect-level Sentiment Analysis. IEEE Transactions on Knowledge and Data Engineering, 2021, 4347. https://doi.org/10.1109/TKDE.2021.3075238

29. Zhang, K., Zhang, H., Liu, Q., Zhao, H., Zhu, H., Chen, E. Interactive Attention Transfer Network for Cross-Domain Sentiment Classification. Proceedings of the AAAI Conference on Artificial Intelligence 2019, 33(1), 5773-5780. https://doi.org/10.1609/aaai.v33i01.33015773