IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

SENTIMENT CLASSIFICATION ACROSS MULTIPLE SOURCES USING AN INTEGRATED POLARITY-SCORE BASED EMBEDDING IN DEEP LEARNING MODEL.

1Parvati Kadli, 2Vidyavathi B M
1Associate Professor and HOD, 2Professor and HOD
1Computer Science and Engineering,
1PDIT, Hosapete, Karnataka, India

Abstract: The increasing prevalence of the Internet and the advent of Web 2.0 have led to a heightened focus on sentiment analysis of freely expressed opinions in different social media platforms. Sentiment analysis plays significant role in various applications such as review-based product recommendations and opinion mining. This study presents cross-domain-labeled Web sources (Amazon and Tripadvisor) in a unique cross-source cross-domain sentiment categorization approach. We propose a novel architecture named the Deep Learned Model with Integrated binary embedding. This model combines the strengths of Bidirectional Long-Short Term Memory (Bi-LSTM), Bi-Gated Recurrent Unit (GRU), and Convolutional Neural Network (CNN).. The proposed approach achieves an accuracy of over 80% in sentiment analysis of Facebook and Twitter datasets.

Index Terms **- Convolutional neural network; cross-domain data; sentiment analysis; social media; Facebook; Twitter; Amazon; Tripadvisor.**

1.Introduction

Technological advancements, like the proliferation of the Internet, the Web 2.0 phenomena, and widespread usage of mobile phones, have enhanced the accessibility of liberally prejudiced script (e.g., blog assessments, social network remarks). The utility of sentiment assessment, also known as opinion mining, that employs computer methods to systematically assess individual views, feelings, and assessments regarding objects (e.g., goods, facilities, companies) [1], is enhanced by this big data source of unorganized texts. Numerous researches [2-4] have looked at the patterns of opinion in social media and their possible influence on judgement. As a result, sentiment assessment is a critical component of current decision support systems, assisting with judgments in a variety of real scenarios, including hotels, financial markets, [7], and road mishaps [8].

Recognizing the prominence of social media sites (like Facebook, Twitter), supervised machine learning algorithms for sentiment assessment of online network messages have indeed been developed [9]. Developing an effective machine learning classifier for certain sentiment domain and data source, on the other hand, necessitates a significant amount of work from data analysts and the implementation of computation trials. Furthermore, as compared to other areas, few contain fewer dataset (such as maximum Amazon feedback are regarding electronics). These 2 challenges are addressed utilizing cross-domain sentiment assessment,[10], [11], current transfer learning research area which tries utilizing sentiment types that have already been adapted to certain domains (e.g., electronics) to forecast sentiment of messages by different domains (e.g., volumes).User-labeled contributions are frequently requested on current Online sites. Amazon and Tripadvisor, for instance, encourage people to write feedback using a 5-star review system. Other social networks, on the other hand, have a scarcity of sentiment-labeled information. Facebook, for instance, is prominent networking sites with over two billion active users, but a tiny portion of Facebook sites enable tagged evaluations. Furthermore, with 330 million active monthly users, Twitter is a relevant networking site which has been utilized for communicating thoughts regarding a broad variety of subjects, such as goods [12] or financial sectors [13]. However, Twitter-labeled information is significantly harder to obtain, frequently necessitating arduous human labor. There may also be disparities in the sorts of writings published on various Web sites. Twitter, for instance, limits the maximum length of text words, although Facebook will not. The bulk of cross-domain research uses a solitary Web data, as indicated in Sec. 2. (e.g., Amazon assessments). As seen in Refs. [14] And [15], combining datasets can be beneficial in terms of improving data accuracy and reducing bias. As a result, analyzing that which we call "cross-source cross-domain" sentiment categorization, wherein cross-domain information through 1 or additional labelled source materials is utilized for generating sentiment assessment methods which are then utilized to categorize non-labeled cross-domain texts by another sources, has probable improvement and study profit. We suggest such a method in this work, with subsequent primary offerings:

1) We use several datasets and domains for developing and evaluating the systems in cross-source cross-domain sentiment categorization. To train the sentiment classification systems, we use cross-domain big data labelled resources from several Web sites (Amazon and Tripadvisor). The learnt models are then utilized to forecast sentiment on 2 unlabeled social media resources, Facebook and Twitter datasets.
(2) This paper offers the Deep learned model, which is a system that relies on Bidirectional Long-Short Term Memory (Bi-LSTM), Bi-Gated Recurrent Unit (GRU), and Convolutional Neural Network (CNN) using a 3-phase innovative embedding architecture and attention network that is compared to prior research.

2.LITERATURE REVIEW

Dipak Patel and Kiran [16] Amin developed a novel multi-source sentiment analysis approach in 2021, which contains 6 steps for domain customization. Higher order statistics dependent on text characteristics are used to examine for similarity between assertions. The modified cross entropy statistic is used in the higher order statistics dependent characteristics collection. The residual characteristics are then supplied to suggested classifier, which estimates polarity of particular domain. Neural Network (NN) is used to classify the data. The parameters of NN are optimized utilizing the Improved Grey Wolf Optimization (IGWO) method that is a more advanced variant of the GWO procedure. Lastly, the model's efficiency is evaluated to that of many other domain modification methods, and it is found to be superior given the growth is 28% over 6% precision.

Miguel et al. [17] published a paper in 2019 that describes a new strategy for improving a collection of current Sentiment Analysis (SA) algorithms in ensemble classifier centered on input text domain. The domain adaptation problem has been alleviated using a specific set of SA approaches. Finally, the results demonstrated the usefulness of the selected approach by demonstrating improved sentiment analysis performance.
Xing et al. [18] proposed new technique in 2019 for concurrently training word polarity and a vanilla sentiment classifier on objective area. The incorrectly anticipated phrases, in particular, were logged consecutively and used as inspection. The selected model improved with regards to sentiment classification for multiple domains employing emotive lexicons, according to research findings from array of well-known datasets.

Yin et al. [19] proposed new "CITK approach for cross-domain sentiment categorization" in 2019. The selected strategy used a capsule network to encode domain invariant information, which helped to overcome the data barrier between the destination and origin domains. Furthermore, BERT was used to convert phrases to similar sizes, a process known as pre-training that resulted in a better standard model embedding model. In the conclusion, the investigative results showed that the chosen model outperformed the old models.

Bollegala et al. [20] showed "embedding learning" and built 3 processes which collected: (a) pivots characteristics (i.e., typical features which arise in both the objective and origin domains), (b) label specifications in source documents, and (c) geometric characteristics in both the destination and source domains. As a result of the selected "joint optimization" method, embeddings that were amenable to SA classification were found. Furthermore, the results of this technique have shown that the provided model performs better in regards of SA cataloguing.

Manshu and Bing [21] designed a HANP framework for completing CDSC activities in 2019. By taking into account prior information, the offered technique was able to get both domain particular and domain individual features. Furthermore, the HANP incorporates an attention-based method, allowing for the recording of relevant phrases and words connected with moods. Finally, tests on "Amazon review datasets" revealed that chosen HANP accomplished significantly compared to current systems.

Manshu and Xuemin [22] suggested an end-to-end model in 2019. There were two subcategories in this strategy: one was a "CTN," and the other was a "CHAN." Based on the attention focused approach, the later one caught relevant phrases relating sentiments. Finally, "Amazon review datasets" were utilized for validating improvement of projected strategy.

3. Dataset and Pre-Processing

3.1. Sentiment Analysis Data

Here, we look at 2 sentiment output label collections, each having two categories ("negative and positive"). The datasets (except Twitter dataset) used in the study are open source and may be found at https://github.com/paolazola/Cross-source-crossdomain-sentiment-analysis. Texts are derived by 4 primary data sources:

This repository contains two Python pickle dictionaries with labelled data for cross-source and cross-domain sentiment assessment. 2 folders are connected to English texts. The following items make up the Dataset:

1. Amazon: It comprises a collection of 75,000 ratings published in English from January to February 2018 on various Amazon goods (such as electrical gadgets, kitchen items, clothing, and homewares). Every comment includes the date, a brief title, and the sentiment (stated as a 5-star rating) expressed by the individual who wrote it. The user’s name has been removed for reasons of confidentiality.
2. Tripadvisor: It comprises a collection of 75,000 English comments on hotels, restaurants, and cities that were obtained between January and February 2018 from Tripadvisor.com. Every comment includes the date, a brief title, and the sentiment (conveyed as a 5-star grade) expressed by the individual who wrote it. The user’s name has been removed for reasons of confidentiality.
3. Facebook: There are 5,782 English Facebook postings in total. Only particular public sites with a 5-star rating scheme are mentioned in the posts. Colleges, gatherings, renowned individuals, residents, celebrations, businesses, and towns are among the themes included in the sampling evaluations conducted from January to February 2018. Every piece in the group is associated with the user's determined sentiment (represented as a 5-star grade). The user’s name has been removed for reasons of confidentiality.

4. The twitter dataset is made up of multimodal user postings published on Twitter between 2014 and 2017 [23] [24]. Every one of the entities fall into one of four categories: Person, Place, Organization, and Other Information

3.2 Pre-Processing

Facebook, Trip Advisor, and Amazon datasets have been updated to operate with the Deep Learning Architecture that has been designed. Pickle is the file format for the datasets that are accessible. As a result, they're transformed to text form. The information is then preprocessed in the following way:

- The rating of every comment is regarded for categorizing them as positive and negative classes.
- If rating > =4, then review is ‘positive’,
- If rating <4, then review is ‘negative.
- Stopwords are detached

The Twitter dataset is labelled with three values: 1, 0, and -1. Because we are working with two courses, the following reviews are taken into account:

- If the label is 1, the review is favourable.
- If the label is -1, the review is not favourable(negative)

4. Embedding Patterns

4.1. Embedding Schemes

Because it is simple to design, the bag-of-words model was employed to represent text for sentiment categorization. However, because the data is so large, the sizes of features are enormous. This will result in a shallow representation of the word vector matrix, perhaps resulting in overfitting. This prompted Mikolov et al. [25] to propose the word embedding model, which aids in the learning of vector representations of words. Word embedding [26] is a technique for representing sentences that involves mapping words into low-dimensional vectors. The vectors are denoted in constant space, and words are charted onto nearby locations as well. It is accomplished with the aid of neural network and language model which has been pre-trained with corpus data. This will aid in the extraction of words' semantic and syntactic interpretations from unstructured information. Collobert et al [27, 28] presented architecture for natural language analysis that focuses on the similar embedding for multiple purposes. Other study [29] suggested utilizing neural word embedding to predict sentiments.

For embedding, predictive approaches such as Probabilistic Language Models (NNLM) [30], N-grams, and Word2Vec [31] are utilized. A feed forward network is used by Word2Vec. It contains stages for input, projection, and output. Continuous Skip-gram Model (Skip-gram) and [Continuous Bag-of-Words Model] are 2 methods utilized for forecasting (CBOW). In order to maximize the categorization of words, skip gramme considers the texts in the phrase.

Caiand Wan [32] presented a multi-domain neural sentiment categorization system that can be used to grasp and learn domain-aware word embedding entirely. It is accomplished from exchanging word embeddings from multiple domains in order to obtain domain-specific data. For NLP purposes, a novel word format called GloVe [33] has been proposed. The benefits of local context window approaches and global matrix factorization are combined in this method. FastText and WordEmbed are two more word embedding models.

4.1.1. Proposed Polarity Score based Binary Pattern Vector (PSBPV)

The power of rule-based sentiment lexicons, particularly VADER (Valence Aware Dictionary for Sentiment Reasoning), is used for building bit binary pattern for word embedding in this research. VADER[34] is particularly responsive to microblogging situations. It converts lexical aspects of a text into sentiment ratings, which are measures of emotional intensity. This application calculates the sentiment value of a text by adding intensity of every word in text. It gives a grade for each of the four forms: positive, negative, neutral, and compound. It gives a rating oscillating between -1 to 1, with -1 as the most negative and 1 as the most positive. Compound rating is used to construct the embedding bit pattern in our technique.The phrase is broken into words after it has been pre-processed, like stop word removal, stemming, and so on. The polarity rating for every word is calculated using VADER.

In this approach, a compound rating is computed for every word, and a sequence is formed dependent on whether or not this number falls inside a certain limit. The preceding table 1 shows how to create a Polarity Score dependent Binary Pattern.

© 2023 IJNRD | Volume 8, Issue 5 May 2023 | ISSN: 2456-4184 | IJNRD.ORG

Pseudo code

```
for each word w in sentence:
    [positive_score, negative_score, neural_score, compound_score] = Vader Polarity Score (w)
    if compound_score>=0.6
        PBP_v=[1, 0, 0, 0, 0]
    else if compound_score>=0.3
        PBP_v=[0, 1, 0, 0, 0]
    else if compound_score>=0
        PBP_v=[0, 0, 1, 0, 0]
    else if compound_score>=-0.5
        PBP_v=[0, 0, 0, 1, 0]
    else
        PBP_v=[0, 0, 0, 0, 0]
    end
end
```

For Instance:
Let us contemplate the input sentence is like this

I	love	this	beautiful	Place

After subsequently eliminating stop words, it will be

Stopwords	love	beautiful	Place
Vader's Compound score	0.3612	0.413	0.0
Polarity score based Pattern	{0,1,0,0,0}	{0,1,0,0,0}	{0,0,1,0,0}

Aside from state-of-the-art embeddings like Glove, FastText, and Word2vec, the suggested embedding assigns coefficients to every word to differentiate its sentiment lexicon. This sentiment score-based word layer embedding aids model in cross-domain sentiment forecasting. Every word has their own rating and significance, but depending on the domain and adjacent words, it might reflect a variety of sentiments. As a result, this embedding, alongside the BiLSTM model, is included in this suggested Tri model network to increase sentiment analysis effectiveness across domains.

4.1.2 GloVe Model

By producing the relationship among these, this model uses the frequency of incidences of words, that is key source of accessible data. Utilizing this data, a latest method for encoding words known as GloVe was developed (Global Vectors). Contemplate the letters I j, and k. Number of times these terms appear together is N ijandN jk. The co-occurrence of I and j is N iji, while co-occurrence of j and k is N jk. In GloVe, word embedding is made possible by combining this information with its probability.

4.1.3 FastText Model

In the year 2016, Facebook AI suggested FastText. It's a natural language processing (NLP) package for word embedding. With a hierarchical system classification, FastText employs the Continuous Bag of Words (CBOW) model to represent words. By anticipating category, CBOW may anticipate the word. Softmax layer is replaced by hierarchical softmax in this case.
When the emphasis is on words and phrases, sentiment categorization produces greater results. The researchers employed the attention mechanism [35-38] to achieve this aim through collecting the distance dependence. For boosting the value of words and sentences, hierarchical attention networks [39] have been proposed at both the sentence and word levels. For multitasking, another attention model [40] was presented that shared the feature extractor. Shareable sentence representation learning [41] generates alternative representations via an attention mechanism and task-independent query vectors. Cai and Wan [42] developed a domain-aware attention mechanism-based multi-domain neural sentiment cataloguing system. Zhang et al. [43] presented an collaborative attention transfer network and investigated impact of aspect network and sentence network attention.

5. Proposed Architecture –Deep Learned Model with Polarity Score Based Embedding

Long short-term memory (LSTM) is kind of Recurrent Neural Network (RNN) which can understand how order affects sequence prediction. The outcome of text categorization utilizing LSTM was significant. When using LSTM, the sparsity of the text data leads to a high level of difficulty. Bidirectional LSTM (BiLSTM) was offered as a solution to the algorithm's complication when employed for text categorization. This paper presents an integrated architecture that includes a BiLSTM [44], BiGRU [45], and CNN attention mechanisms, as well as novel word embedding to improve pre-trained word embedding. Figure 2 depicts model architecture that has been suggested.

There are 3 types of models in this suggested network. The Glove embedding method is utilized in the first model to produce embedding patterns, that are subsequently injected into a BiLSTM tier having 256 nodes, succeeded with Average pooling tier having pool capacity of 2 and a Context Attention tier having loss of 0.5. Eventually, from this initial model, 100 dense features are retrieved. The next model is built by aid of BiGRU, which has 128 nodes and uses Fast Text embedding. The successive layers are deployed in the same way as the first model, and 100 dense features are extracted at the end. Three types of convolutional levels with variable number filters and kernel sizes make up the third integrated model. The first convolutional layer is made up of 100 filters having a kernel size of 3 and relu activation on glove embedding, while 2nd and 3rd convolutional tiers are made up of 6 and 300 filters, correspondingly, with kernel sizes of 3 and 5. On suggested polarity score-based embedding and FastText embedding, Relu and softplus activations are implemented after 2nd and 3rd convolutional tiers, respectively. Every convolutional tier was input into a 2-pool average pool tier. Using a concatenation layer, the mean pool from the 2nd convolutional tier on PSBP is merged with the average pools from the 1st and 3rd convolutional regions, resulting in Integrated Polarity Score Dependent Binary Pattern Embedding (IPSBPE) with Glove and Fasttext CNN model.

Attention layer, trailed by Dropout layer, and dense level with 100 pieces, effectively processes the output from those 2 concatenation tiers. With the aid of the Softmax tier, every one of the characteristics from the 4 types of dense tiers is fused to generate an efficient deep feature to describe the emotion polarity of a phrase.

The suggested PSBP in this model is a five-bit word embedding that efficiently retrieves the spatial data of a phrase. Both the previous and following context descriptions are accessed utilizing BiLSTM. The data provided by the underlying tiers of BiLSTM, BiGRU, and the integrated model of CNN is given varied emphasis utilizing the attention mechanism. Lastly, the dense layer employs the softmax classifier to categorize the collected context data. This architecture can capture both the spatial information of phrases and the universal semantics of sentences.

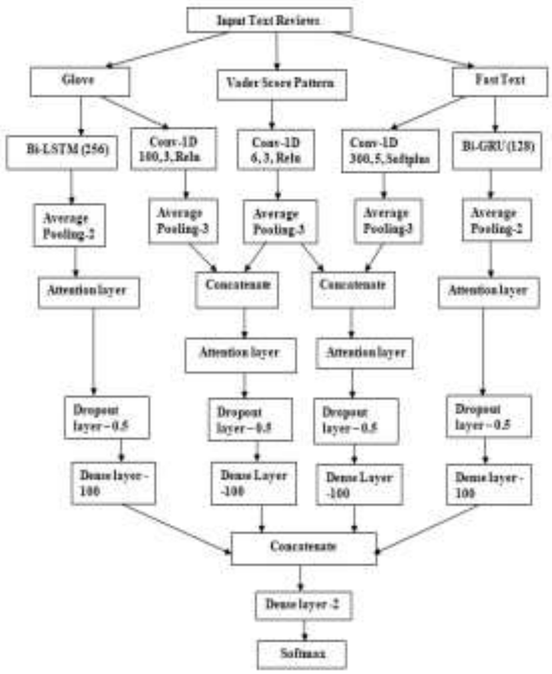

Figure 1: Architecture of Deep learned model

5.1. BiLSTM

For categorization, the suggested technique uses the BiLSTM [44] model. A BiLSTM is a sequence computing type that consists of 2 LSTMs. The input is first collected in a forward direction by the LSTM. The 2nd LSTM works in the opposite direction. Glove is used to construct word embedding, and the retrieved features are input into the BiLSTM tier, which extracts sentence level features to reflect each phrase's emotion polarity. The Domain Exact Attention Method, in conjunction with the BiLSTM layer, is used to choose significant characteristics and increase the approach's performance.

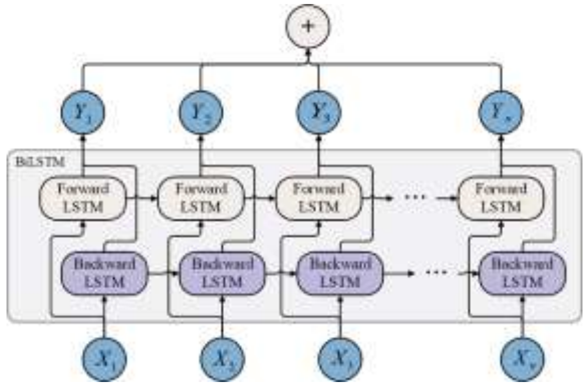

Figure 2: BiLSTM Architecture

5.2. BiGRU

We utilize a bidirectional GRU [45] to generate word annotations by combining input by both directions for each word and including contextual data into annotation.

Bidirectional GRU contains forward GRU $\vec{f}$ and backward GRU $\overleftarrow{f}$. The forward GRU $\vec{f}$ inspects sentence s_i by w_{i1} to w_{iK} as revealed in Eq. (2). Backward GRU $\overleftarrow{f}$ inspects sentence s_i by w_{iT} to w_{i1} as revealed in Eq. (3).

$$x_{ik} = W_e w_{ik}, k \in [1, K], \qquad [1]$$

$$\vec{h}_{ik} = \overrightarrow{GRU}(x_{ik}), k \in [1, K], \qquad [2]$$

$$\overleftarrow{h}_{it} = \overrightarrow{GRU}(x_{ik}), k \in [K, 1]. \qquad [3]$$

An explanation for a specified word $w_{ik}$is gotten via concatenating the forward hidden state $\vec{h}_{ik}$ and backward hidden state$\overleftarrow{h}_{ik}$, i.e., h_{ik} = $\vec{h}_{ik}$,$\overleftarrow{h}_{ik}$]. This type reviews the information of whole sentence w_{ik} as well.

GRU stands for Gated Recurrent Unit and is a form of Recurrent Neural Network (RNN). The forget gate and the input gate are combined into a solitary update gate in GRU. At the very same time, GRU integrates the hidden state and the cell state. Figure 2 depicts the unit topology of the system.

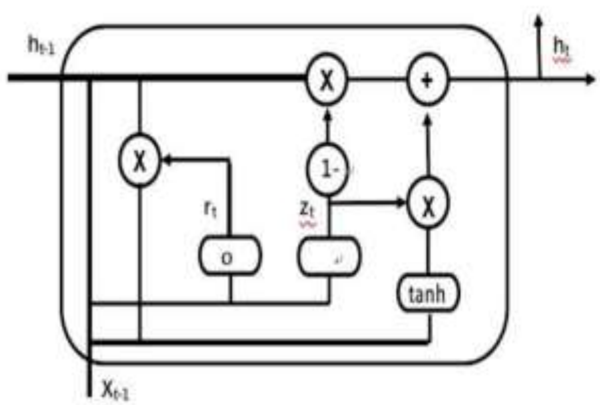

Figure 3: GRU Coding Unit

Precise approximation process is achieved following Eqs (4) – Eqs (7).

$$Z_t = \sigma(W_i * [h_{t-1}, x_t]) \qquad [4]$$

$$r_t = \sigma(W_i * [h_{t-1}, x_t]) \qquad [5]$$

$$h_t = tanh(W_c * [r_t . h_{t-1}, x_t]) \qquad [6]$$

$$h_t = (1 - z_t). c_{t-1} + z_t . \widetilde{h_t} \qquad [7]$$

In which σ is signified like sigmoid function.

· is signified like dot product.

x_t is signified like input vector at time t.

h_t is signified like hidden state at time t.

z_t is signified like update gate.

r_t is signified like reset gate.

The data streams into the following phase are controlled by this update gate. As a result, the control data is compromised. The concealed state's output is determined by the reset gate.

5.3. Convolution layer

Every neuron function as a kernel in a recurrent layer, which is made up of convolutional kernels. When a kernel is symmetric, convolution process develops correlation process [46,47]. Eq. (8) shows how to express the convolution process

$$f_l^k(p,q) = \sum_c \sum_{x,y} i_c(x,y).\, e_l^k\,(u,v) \qquad [8]$$

In which i_c (x, y) is signified as component of input tensor I_c.

This input is component wise increased from value of e_l^k(u, v). e_l^k(u, v) is counted as index location of k^{th} convolutional kernel l k of first tier. Output feature-map of k^{th} convolutional process is signified as Equation 9

$$F_l^k = [f_l^k, \dots, f_l^k(p,q), \dots, f_l^k(P,Q)]. \qquad [9]$$

Usually, the sliding kernel idea is used to extract several sets of characteristics from an input. As a result, convolutional operations can be used as a weight-sharing system. The CNN parameter in completely linked networks offers efficient results with similar collection of values on input, which can be image, audio, text, or signal. The kind of padding, number of filters, and orientation of convolution are all used to classify convolution processes [48]. 3 kinds of one-dimensional Convolution layers are employed in this study to remove cross-domain features with sufficient strength to accurately depict every domain.2 Convolutional tiers get the embedding characteristics via the conventional embedding tier, and one Convolutional layer gets the characteristics via the polarity dependent embedding and conventional embedding levels in the shape of correlation mode.

5.4. Average Pooling

To decrease measurement of feature of depiction while keeping significant properties, the pooling approach [49] is applied. It's also referred to as a non-linear down-sampling technique. As a result, after obtaining the hidden state interpretations of sentences, the pooling technique is used in this approach. Pooling may be done using a variety of non-linear functions. Pooling is the most effective of the functions. Lastly, using the following Eq., the sentence pooling vector (i.e.,h_s^p) is calculated throughout Eq. (9).

$$h_s^p = \sum_{i=1}^{n} h_a^i / n \qquad [10]$$

5.5. Context Attention Layer

The contributions to the portrayal of the essential to the sentence interpretation are not distributed evenly throughout the works. To represent the phrase, we use the Attention mechanism approach to determine the most dominating characteristic vector of words. A sentence vector is created using the representation of those informative words [50].

$$u_{it} = \tanh(W_w h_{it} + b_w) \qquad [11]$$

$$\propto_{it} = \frac{\exp(u_{it}^T u_w)}{\sum_t \exp(u_{it}^T u_w)} \qquad [12]$$

$$s_i = \sum_t \propto_{it} h_{it} \qquad [13]$$

The term annotation h_{it} is supplied constantly via a single tier MLP to produce u_{it}, which is a concealed representation of h_{it}. Word level context vector u_w is used for determining significance of word resemblance of u_{it}. This approach uses a softmax method to produce a normalized significance weight. According to the weights, the sentence vector α_{it} is computed utilizing a weighted amount of the word annotations. At elevated description of set query, context vector u_w is provided. "What is the informative term," for instance, may be the inquiry. This approach may be found in almost all memory networks [51], [52]. During the training method, the word context vector u_w is mistakenly introduced and jointly learnt.

5.6. Dropout Layer

In this suggested work, dropout is used to avoid the problem of overfitting. After the attention method, the dropout procedure is carried out by assistance of a dropout layer, having 50% dropout degree. As an alternative to avoidable variables, this dropout level is used to focus on distinguishing forthcoming variables [53].

5.7. Dense Layer

A thick tier is made up of network layer neurons. The input from every level is collected in the dense layer, which is made up of neurons. Weight matrix w, a bias vector b, and authorizations from proceeding level an are all elements for every dense layer. Eq. (14) represents the dense layer.

$$y = (a(x.w) + b) \qquad [14]$$

In which a is signified as element-wise argument, w is signified as weights matrix and bias is signified as bias vector created from layer [54].

5.8. Softmax Layer

The Softmax layer represents output layer of every Neural Network (NN). Nature of input text and functionalities of hidden layer are anticipated based on deep score. Sentiment polarity of reviews in cross domain [55] is predicted in the Softmax layer using the deep score. The Softmax classifiers are given the review sentence illustration z_s to imagine the probability distribution $\hat{y}_s \in R^c$ of sentiment groups as indicated in Eq (15).

$$\hat{y}_s = softmax(W_s z_s + b_s) \qquad [15]$$

In which c is signified as amount of sentiment groups, and allocated c = 2. In this planned model, $W_s \in R^{c \times d_h}$ and $b_s \in R^c$ is signified as weight matrix and bias correspondingly.

6. Performance Evaluation

F-measure, Accuracy, Retention, and Efficiency are the quality measures used to assess outcomes. The True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) values for provided classes are used for constructing these functioning measures.

True Positive (TP): These are the positive numbers that have been calculated accurately. Value of the real class is denoted as yes, and value of the computed class is likewise displayed as yes, according to TP.

True Negatives (TN): These are the negative numbers that were computed accurately. The value of a genuine class is denoted as no, and the value of a computed category is similarly displayed as no, according to TN.

False Positives (FP): Value of a genuine category is presented as no, whereas number of a computed category is presented as yes, according to FP.

False Negatives (FN): Value of a genuine class is presented as yes, whereas result of a computed category is presented as no, according to FN.

6.1. Precision

It is called as proportion of precisely computed positive annotations to whole computed positive annotations. This is specified from Eq.16

$$Precision = \frac{TP}{TP+FP} \qquad [16]$$

6.2. Recall (Sensitivity)

It is called as proportion of precisely calculated positive annotations to every annotation in actual class. This is known as Sensitivity as well. This is specified from Eq. [17]

$$Recall = \frac{TP}{TP+FN} \qquad [17]$$

6.3. F1 Score

It is called as weighted mean of Accuracy and Recall. It is bench mark metric. Therefore, this score contemplates both False Positives (FP) and False Negatives (FN) into relation. This is specified from Eq. 18

$$F1Score = \frac{2*(Recall*Precision)}{(Recall+Precision)} \qquad [18]$$

6.4. Accuracy

It is called as utmost insightful functioning amount. This is described as proportion of precisely computed annotation to whole annotations. This is specified from Eq. [19]

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \qquad [19]$$

6.5 AUC

The Area Under the Curve (AUC) is brief of ROC curve which evaluates classifier's capability to discriminate among classes. As a consequence, better classifier's AUC score is, better it is at differentiating amongst positive and negative categories.

AUC is among the most essential assessment measures for assessing a categorization model's performance. It is a metric for evaluating the performance of a classification issue at various thresholds.

7. Results

The experimental outcomes are revealed below in Table1 and Table 2. The Notation Dataset1→Dataset 2 represents that the model is trained on Dataset1 and tested on Datset2.

Table 1 displays outcomes of proposed technique related to other method using CNN [45], with regards to AUC, Accuracy, and F1 score. The CNN model is trained on both Amazon and Trip advisor (TA) to test on facebook or Twitter reviews. We trained our Deep learned model with integrated pattern (DLMIBP) model on **Amazon samples** only and validated it on Facebook and Twitter. From the table below, we can see that our technique outperforms other technique based on CNN model [45] in all the three cases in terms of AUC, Accuracy, and F1 Score.

Table 1: Comparison of overall functioning Performance among DLMIBP and CNN[45]

Case 1		AUC	Accuracy	F1 Score
Training data	Testing data			
Amazon→ Facebook(DLMIBP)		81%	82%	73%
Amazon TA→ Facebook(CNN)		81%	81%	72%
Case 2		AUC	Accuracy	F1 Score
Amazon→ Trip Advisor(DLMIBP)		82%	77%	73%
Amazon→ Trip Advisor(CNN)		78%	76%	66%
Case 3		AUC	Accuracy	F1 Score
Amazon→ Twitter(DLMIBP)		71%	65%	65%
Amazon / TA→ Twitter(CNN)		68%	61%	60%

The proposed Deep learned model with integrated pattern (DLMIBP) model's AUC is compared to the other method, CNN[45]. Our strategy outperforms CNN with Amazon as source reviews and target as other social media reviews. Better results of AUC are observed . The difference in AUC between the presented technique and CNN is shown in Figure 4.

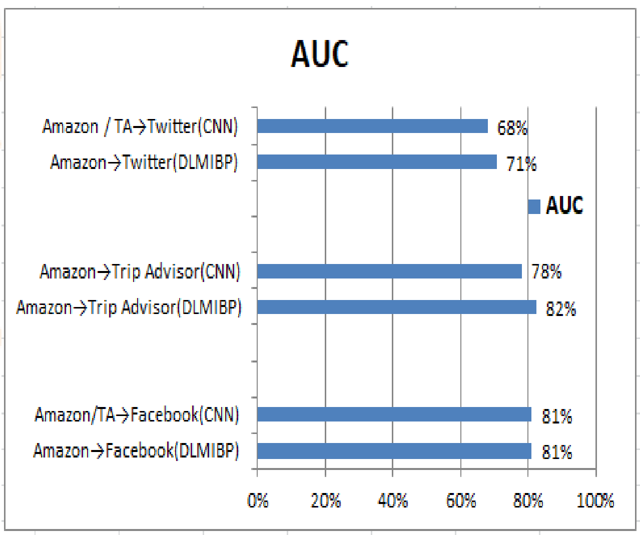

Figure 4: Comparison of AUC among DLMIBP and CNN[45].

The proposed DLMIBP model's Accuracy is compared to the other method, CNN[45]. Our strategy outperforms CNN with Amazon as source reviews and target as other social media reviews. Better results of Accuracy are observed. The difference in Accuracy between the presented technique and CNN is shown in Figure 5.

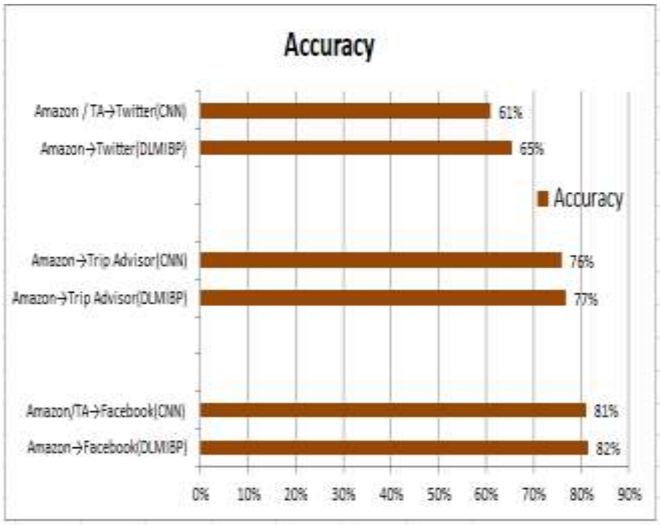

Figure 5: Comparison of Accuracy among DLMIBP technique and CNN[45]

The proposed DLMIBP model's F1_Score is compared to the other method, CNN[45]. Our strategy outperforms CNN with Amazon as source reviews and target as other social media reviews. Better results of F1_Score are observed. The difference in F1_Score between the presented technique and CNN is shown in Figure 6.

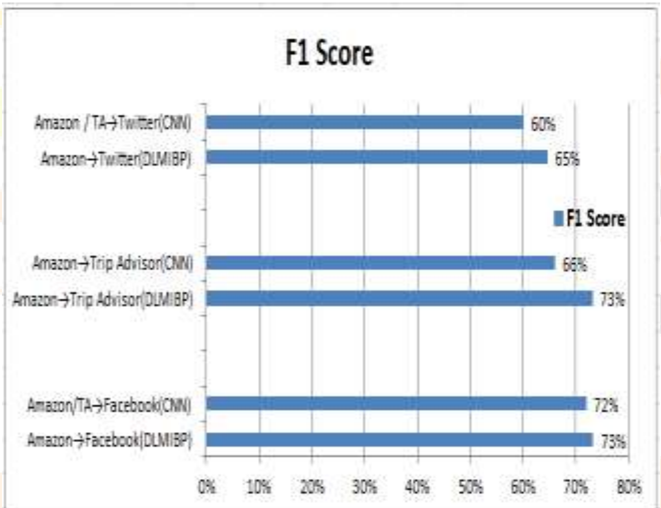

Figure 6: Comparison of F1 Score among DLMIBP technique and CNN.

Table 2 displays outcomes of DLMIBP in comparison to another technique, CNN [45], in terms of AUC, Accuracy, and F1 score. Here, The CNN model is trained on both Amazon and Trip advisor reviews to test on facebook and Twitter reviews. We trained our model on **Trip advisor** samples only and validated it on Facebook and Twitter reviews. From the table 2, we understand that our technique outdoes former system CNN[45] with regards to AUC, Accuracy, and F1 Score in case 3.

Table 2: Comparison of overall functioning performance among DLMIBP and CNN[45]

Case 1		AUC	Accuracy	F1 Score
Training data	Testing data			
Tripadvisor→Amazon (DLMIBP)		82%	76%	70%
Tripadvisor→Amazon (CNN)		81%	75%	78%
Case 2		AUC	Accuracy	F1 Score
Tripadvisor→Facebook(DLMIBP)		80%	79%	72%
Tripadvisor/Amazon→Facebook(CNN)		81%	81%	72%
Case 3		AUC	Accuracy	F1 Score
Tripadvisor→Twitter(DLMIBP)		70%	62%	62%
Tripadvisor/Amazon→Twitter(CNN)		68%	61%	60%

The proposed DLMIBP model's AUC is compared to the other method, CNN[45]. Our strategy outperforms CNN with Tripadvisor as source reviews and target as other social media reviews. Better results of AUC are observed . The difference in AUC between the presented technique and CNN is shown in Figure 7.

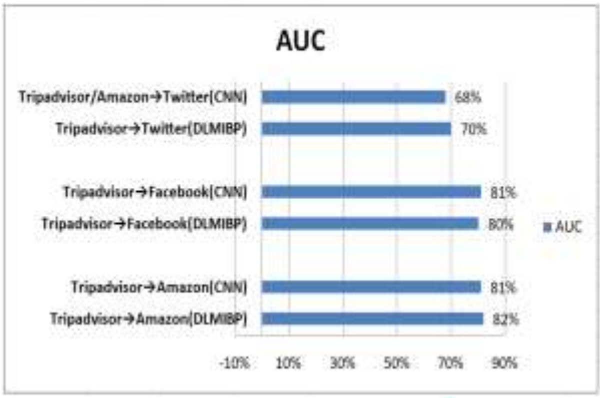

Figure 7: Comparison of AUC among DLMIBP technique and CNN[45]

The proposed DLMIBP model's Accuracy is compared to the other method, CNN[45]. Our strategy outperforms CNN with Tripadvisor as source reviews and target as Twitter and Amazon. The difference in Accuracy between the presented technique and CNN is shown in Figure 8.

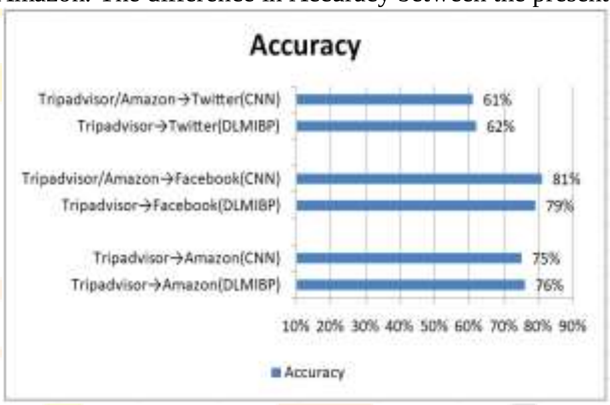

Figure 8: Comparison of Accuracy among DLMIBP technique and CNN[45]

The proposed DLMIBP model's F1_Score is compared to the other method, CNN[45]. Our strategy outperforms CNN with Tripadvisor as source reviews and target as other social media reviews. Better results of F1_Score are observed with one exception on target Amazon reviews. The difference in F1_Score between the presented technique and CNN is shown in Figure 9.

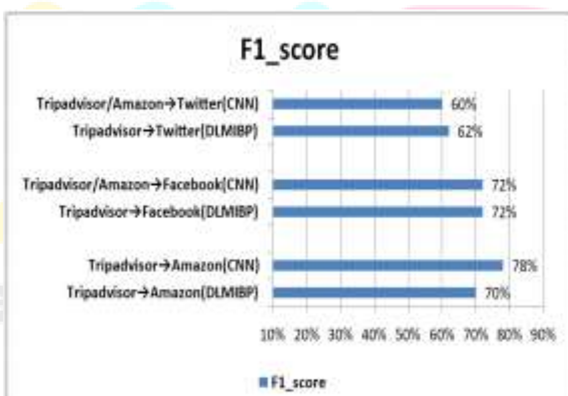

Figure 9: Comparison of F1_Score among DLMIBP technique and CNN[45]

8. Conclusion and Discussion

In this study, we conducted an investigation into a unique approach for cross-source cross-domain sentiment analysis. Our objective was to efficiently categorize sentiment for different components, such as restaurants, hotels, books, and music available on various platforms. To achieve this, we developed a sentiment classifier using labeled Web sources like Amazon or Tripadvisor, and then adapted the model to predict sentiment of reviews in unlabeled social media feedback from platforms like Facebook and Twitter. The proposed model, called DLMIBP, incorporates Bidirectional Long-Short Term Memory (Bi-LSTM), Bi-Gated Recurrent Unit (GRU), and Convolutional Neural Network (CNN) with a three-phase embedding architecture and attention network. This approach eliminates the need to create separate sentiment models for each data source and simplifies the categorization of unlabeled texts. Finally, we compared the performance of our DLMIBP model with CNN in terms of AUC, Accuracy, and F1 Score, and observed that our technique outperformed CNN, particularly when trained on Amazon reviews.

References

[1] B. Liu, Sentiment analysis and opinion mining, Synthesis Lectures on Human Language Technologies 5(1) (2012) 1–167.

[2] Y. Dong, Q. Zha, H. Zhang, G. Kou, H. Fujita, F. Chiclana and E. Herrera-Viedma, Consensus reaching in social network group decision making: Research paradigms and challenges, Knowledge-Based Systems 162 (2018) 3–13.

[3] Y. Dong, M. Zhan, G. Kou, Z. Ding and H. Liang, A survey on the fusion process in opinion dynamics, Information Fusion 43 (2018) 57–65.

[4] R. Ureña, G. Kou, Y. Dong, F. Chiclana and E. Herrera-Viedma, A review on trust propagation and opinion dynamics in social networks and group decision making frameworks, Information Sciences 478 (2019) 461–475.

[5] H. Shi and X. Li, A sentiment analysis model for hotel reviews based on supervised learning, in 2011 International Conference on Machine Learning and Cybernetics (ICMLC), Vol. 3 (IEEE, 2011), pp. 950–954.

[6] N. Oliveira, P. Cortez and N. Areal, Stock market sentiment lexicon acquisition using microblogging data and statistical measures, Decision Support Systems 85 (2016) 62–73.

[7] N. Wang, S. Ke, Y. Chen, T. Yan and A. Lim, Textual sentiment of Chinese microblog toward the stock market, International Journal of Information Technology and Decision Making 18(2) (2019) 649–671.

[8] X. Fu, J. Lee, C. Yan and L. Gao, Mining newsworthy events in the tra±c accident domain from Chinese microblog, International Journal of Information Technology and Decision Making 18(2) (2019) 717–742.

[9] B. Liu, Sentiment Analysis: Mining Opinions, Sentiments, and Emotions (Cambridge University Press, 2015).

[10] J. Blitzer, M. Dredze and F. Pereira, Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification, in Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics (Prague, Czech Republic, 2007), pp. 440–447.

[11] A. Wallin, Sentiment analysis of Amazon reviews and perception of product features, PhD thesis, Master's thesis, Lund University (2014).

[12] A. Go, R. Bhayani and L. Huang, Twitter sentiment classification using distant supervision, CS224N Project Report, Stanford 1(12) (2009).

[13] N. Oliveira, P. Cortez and N. Areal, The impact of microblogging data for stock market prediction: Using twitter to predict returns, volatility, trading volume and survey sentiment indices, Expert Systems with Applications 73 (2017) 125–144.

[14] L. Dalla Valle and R. S. Kenett, Official statistics data integration for enhanced informationquality, QualityandReliabilityEngineering International 31(7)(2015) 1281–1300.

[15] L. Dalla Valle and R. Kenett, Social media big data integration: A new approach based on calibration, Expert Systems with Applications 111 (2018) 76–90.

[16] Dipak Patel and Kiran Amin, "Multi-Source Domain Adaptation in Sentiment Analysis using Optimized Neural Network and Cross-Domain Semantic Library", International Journal Of Intelligent Engineering And Systems. Revised May 2021

[17] M. López, A. Valdivia, E. M. Cámara, M. V. Luzón, and F. Herrera, "E2SAM: EVolutionary ensemble of sentiment analysis methods for domain adaptation", Information Sciences, Vol. 480, pp. 273-286, 2019.

[18] F. Z. Xing, F. Pallucchini, and E. Cambria, "Cognitive-inspired domain adaptation of sentiment lexicons", Information Processing & Management, Vol. 56, No. 3, pp. 554-564, 2019.

[19] H. Yin, P. Liu, Z. Zhu, W. Li, and Q. Wang, "Capsule Network With Identifying Transferable Knowledge for Cross-Domain Sentiment Classification", IEEE Access, Vol. 7, pp. 153171-153182, 2019.

[20] D. Bollegala, T. Mu, and J. Y. Goulermas, "Cross-Domain Sentiment Classification Using Sentiment Sensitive Embeddings", IEEE Transactions on Knowledge and Data Engineering, Vol. 28, No. 2, pp. 398-410, 2016.

[21] T. Manshu and W. Bing, "Adding Prior Knowledge in Hierarchical Attention Neural Network for Cross Domain Sentiment Classification", IEEE Access, Vol. 7, pp. 32578-32588, 2019.

[22] T. Manshu and Z. Xuemin, "CCHAN: An End to End Model for Cross Domain Sentiment Classification", IEEE Access, Vol. 7, pp. 50232-50239, 2019.

[23] Di Lu, Leonardo Neves, Vitor Carvalho, Ning Zhang, and Heng Ji. Visual attention model for name tagging in multimodal social media. In Proceedings of ACL, 2018. ;

[24] Qi Zhang, Jinlan Fu, Xiaoyu Liu, and Xuanjing Huang. Adaptive coattention network for named entity recognition in tweets. In Proceedings of AAAI, 2018.

[25] T. Mikolov, K. Chen, G. Corrado and J. Dean, E±cient estimation of word representations in vector space, arXiv:1301.3781

[26] Bengio, Y.; Ducharme, R.; Vincent, P.; and Jauvin, C. 2003. A neural probabilistic language model. Journal of machine learning research 3(Feb):1137–1155.

[27] R. Collobert, J. Weston, A unified architecture for natural language processing: deep neural networks with multitask learning, in: Machine Learning, Proceedings of the Twenty-Fifth International Conference (ICML 2008), Helsinki, Finland, June 5-9, 2008, 2008, pp. 160–167. doi:10.1145/1390156.1390177. URL http://doi.acm.org/10.1145/1390156.1390177

[28] R. Collobert, J. Weston, L. Bottou, M. Karlen, K. Kavukcuoglu, P. P. Kuksa, Natural language processing (almost) from scratch, Journal of Machine Learning Research 12 (2011) 2493–2537. URL http://dl.acm.org/citation.cfm?id=2078186

[29] M. Dragoni, G. Petrucci, A neural word embeddings approach for multi-domain sentiment analysis, IEEE Trans. Affective Computing 8 (4) (2017) 457–470. doi:10.1109/TAFFC.2017.2717879. URL https://doi.org/10.1109/TAFFC.2017.2717879

[30] Y. Bengio, R. Ducharme, P. Vincent and C. Jauvin, A neural probabilistic language model, Journal of Machine Learning Research 3 (2003) 1137–1155.

[31] T. Mikolov, K. Chen, G. Corrado and J. Dean, E±cient estimation of word representations in vector space, arXiv:1301.3781

[32] YitaoCai and Xiaojun Wan, "Multi-Domain Sentiment Classification Based on Domain-Aware Embedding and Attention," Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19), pp.4904 – 4911.

[33] J. Pennington, R. Socher, and C. Manning. Glove: Global vectors for word representation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 1532–1543, 2014.

[34] Bojanowski, P., Grave, E., Joulin, A., and Mikolov, T. (2017). Enriching word vectors with subword information. Transactions of the Association for Computational Linguistics, 5:135–146.

[35] [64] Peters, M. E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., and Zettlemoyer, L. (2018). Deep contextualized word representations. arXiv preprint arXiv:1802.05365.

[36] Gilbert, CJ Hutto Eric (2014). "VADER: A parsimonious rule-based model for sentiment analysis of social media text." In Eighth International Conference on Weblogs and Social Media (ICWSM-14).Available at (20/04/16) http://comp. social. gatech. edu/papers/icwsm14. VADER.hutto. pdf. 2014.

[37] DzmitryBahdanau, Kyunghyun Cho, and YoshuaBengio. Neural machine translation by jointly learning to align and translate.arXiv preprint arXiv:1409.0473, 2014.

[38] Minh-ThangLuong, Hieu Pham, and Christopher D Manning. Effective approaches to attention-based neural machine translation. arXiv preprint arXiv:1508.04025, 2015.

[39] AshishVaswani, Noam Shazeer, NikiParmar, JakobUszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and IlliaPolosukhin. Attention is all you need. In Advances in Neural Information Processing Systems, pages 5998–6008, 2017.

[40] Zichao Yang, Diyi Yang, Chris Dyer, Xiaodong He, Alex Smola, and Eduard Hovy. Hierarchical attention networks for document classification. In Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 1480–1489, 2016.

[41] Zhigang Yuan, Sixing Wu, Fangzhao Wu, Junxin Liu, and Yongfeng Huang.Domain attention model for multi-domain sentiment classification.KnowledgeBased Systems, 155:1–10, 2018.

[42] RenjieZheng, Junkun Chen, and XipengQiu. Same representation, different attentions: Shareable sentence representation learning from multiple tasks. arXiv preprint arXiv:1804.08139, 2018.

[43] YitaoCai and Xiaojun Wan, "Multi-Domain Sentiment Classification Based on Domain-Aware Embedding and Attention," Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19), pp.4904 – 4911.

[44] K. Zhang, H. Zhang, Q. Liu, H. Zhao, H. Zhu and E. Chen, Interactive attention transfer network for cross-domain sentiment classification, in The 33rd AAAI Conference on Artificial Intelligence (AAAI-2019) (Honolulu, Hawaii, USA, 2019), pp. 5773–5780.

[45] YitaoCai and Xiaojun Wan, Multi-Domain Sentiment Classification Based on Domain-Aware Embedding and Attention, Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence (IJCAI-19), 2019.

[46] DzmitryBahdanau, Kyunghyun Cho, and YoshuaBengio. 2014. Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv:1409.0473.

[47] Ian Goodfellow, Bengio Y, Courville A (2017) Deep learning. Nat Methods 13:35. doi: 10.1038/nmeth.3707

[48] Bouvrie J (2006) 1 Introduction Notes on Convolutional Neural Networks. doi: http://dx.doi.org/10.1016/j.protcy.2014.09.007

[49] LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444. doi: 10.1038/nature14539

[50] Zhang, K., Zhang, H., Liu, Q., Zhao, H., Zhu, H., & Chen, E. (2019). Interactive Attention Transfer Network for Cross-Domain Sentiment Classification. Proceedings of the AAAI Conference on Artificial Intelligence, 33(01), 5773-5780.

[51] Yang, Zichao& Yang, Diyi& Dyer, Chris & He, Xiaodong&Smola, Alex &Hovy, Eduard. (2016). Hierarchical Attention Networks for Document Classification. 1480-1489. 10.18653/v1/N16-1174.

[52] SainbayarSukhbaatar, Arthur Szlam, Jason Weston, and Rob Fergus. 2015. End-to-end memory networks. arXiv preprint arXiv:1503.08895.

[53] Ankit Kumar, OzanIrsoy, Jonathan Su, James Bradbury, Robert English, Brian Pierce, Peter Ondruska, IshaanGulrajani, and Richard Socher. 2015. Ask me anything: Dynamic memory networks for natural language processing. arXiv preprint arXiv:1506.07285.

[54] Sangeetha, K., Prabha, D. Sentiment analysis of student feedback using multi-head attention fusion model of word and context embedding for LSTM. J Ambient Intell Human Comput 12, 4117–4126 (2021). https://doi.org/10.1007/s12652-020-01791-9

[55] Y. Han, M. Liu and W. Jing, "Aspect-Level Drug Reviews Sentiment Analysis Based on Double BiGRU and Knowledge Transfer," in IEEE Access, vol. 8, pp. 21314-21325, 2020, doi: 10.1109/ACCESS.2020.2969473.