Knowledge-Based Systems 213 (2021) 106423

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Automatic construction of domain-specific sentiment lexicon for unsupervised domain adaptation and sentiment classification

Omid Mohamad Beigi, Mohammad H. Moattar *

Computer Engineering Department, Mashhad Branch, Islamic Azad University, Mashhad, Iran

ARTICLE INFO

Article history:
Received 10 January 2020
Received in revised form 15 August 2020
Accepted 1 September 2020
Available online 7 November 2020

Keywords:
Sentiment analysis
Domain adaptation
Domain independent lexicon
Multilayer perceptron

ABSTRACT

Sentiment analysis has long been suffering from inaccuracies using either machine learning methods that mostly benefit from text features or sentiment lexicon-based methods that are prone to domain-dependent problems. Furthermore, since labeling is a time-consuming and an expensive task, supervised machine learning methods suffer from the drawback of insufficient labeled samples. To tackle the mentioned issues, this paper proposes a novel approach with a hybrid of a neural network and a sentiment lexicon. This combination can simultaneously adapt word polarities to the target domain and leverage the polarity of whole document in order to alleviate the need for large labeled corpora in an unsupervised manner. In this respect, a sentiment lexicon is constructed from the source domain in the preprocessing phase using the labeled data. In the Next phase, having a Multilayer Perceptron (MLP), the weights of the first hidden layer are set to the corresponding polarity of each word from the retrieved sentiment lexicon and the network is trained. Finally, a Domain-Independent Lexicon (DIL) is introduced which contains words (mostly adjectives) with static positive or negative scores independent from a specific domain. After feeding the target domain to the pre-trained model, the total accuracy of the framework is enhanced by estimating the sentiment polarity of each sentence using the summation of the scores of the constitutive domain independent words. The experiments on Amazon multi-domain sentiment dataset illustrate that our approach significantly outperforms several alternative previous approaches of unsupervised domain adaptation.

1. Introduction

The phenomenal growth of online platforms such as Amazon (amazon.com), eBay (ebay.com), Tmall (tmall.com), IMDB (imdb.com) and Jingdong (jd.com) have accumulated massive user generated contents (UGCs) such as movie reviews, product reviews, blogs, microblogs and so on [1,2]. Sentiment analysis is defined as the process of determining the inclination of people's opinion towards a subject through Natural Language Processing. Nowadays, since sentiment classification is useful for both consumers to make decision and sellers to understand the public thinking that leads to the better sale, it has been applied in numerous fields. Consequently, the interest in developing sentiment analysis methods has increased in both the academic and the business world [3].

Sentiment classification is known as either 3-class classification problem with positive, negative and neutral opinions of a user concerning a topic (academically called "domain"), or a 2-class classification problem known as "bipolar classification" which does not include neutral class. Since sentiment lexicon only contains words with their positive/negative score, in this paper, bipolar sentiment classification (i.e. polarity detection) is studied.

There have been a variety of approaches for constructing the sentiment lexicon manually [4] or automatically [5], however the polarity of an individual word might be interchangeable from domain to domain. For instance, in Electronics domain, the word "*easy*" in "*This is an easy way to charge your phone*" conveys a positive sentiment, however in DVD domain, "*easy*" may convey negative sentiment, e.g. "*The plot is easy to guess*". Hence, using one sentiment lexicon for multiple domains often performs unsatisfactory and therefore, constructing a domain adaptation method has become a hot research topic.

To tackle the above issue, some approaches have used traditional classification models such as Naïve Bayes (NB) or Support Vector Machine (SVM) with unigram, bigram, or POS features [6]. These approaches are mostly focusing on extracting predefined features since their training models are based on high dimensional and sparse data structure. However, it is more desirable to make learning algorithms less dependent on feature engineering [7]. Another solution is to create a unique classifier for each domain [8,9]. For example, the approach in [8] uses machine learning methods to classify the sentiments of movie reviews,

* Corresponding author.
E-mail addresses: omidbeigi23@gmail.com (O. Mohamad Beigi), moattar@mshdiau.ac.ir (M.H. Moattar).

https://doi.org/10.1016/j.knosys.2020.106423
0950-7051/© 2020 Published by Elsevier B.V.

however it requires a large amount of labeled data in each domain which is time consuming and expensive and without enough labeled data, it is quite impractical to get an accurate result. Supervised learning methods are defined as the approaches that are domain specific, require labeled data for training and the classification of unseen objects is done based on the training set. These approaches benefit from higher accuracy in most cases. In some cases, most of the data is unlabeled and the target is to build the classifier based on the existing labeled data which is called "semi-supervised classification". On the other hand, unsupervised learning which only includes unlabeled data, can be applied faster for the sake of excluding the annotating time.

Motivated by the above observations, an unsupervised domain adaptation model is proposed to adapt a sentiment lexicon from a source domain to a target domain without any supervision. For this purpose, using labeled samples from a source domain, initial polarities of each individual word is obtained. Then a Multilayer Perceptron (MLP) is trained in a way that the extracted word polarities from the source domain are used as the weights of its first hidden layer. Following, the target domain is applied to the MLP model when the weights of each layer are fully trained from the data in the source domain. In this regard, to overcome the issue of domain-specific word polarity fluctuation, which may show inaccurate results in the target domain, the polarities of a source domain are adapted to the polarity of the target domain by taking advantage of the prior knowledge of Domain Independent Lexicon (DIL) words. Domain-Independent Lexicon is a sentiment lexicon which only contains words with static polarities, independent from the domain they are used in. In this case, we can make sure that the polarity of some specific words such as "*great*", "*excellent*" and "*fantastic*" are positive in all domains and the polarity of "*hate*", "*awful*" and "*worst*" are negative regardless of their domains. Extensive experiments are conducted on 13 amazon sentiment domains to evaluate the performance of the proposed model. The experimental results demonstrate the effectiveness of the proposed approach which effectively outperforms baseline methods. The main contributions of our work are listed as follows:

1. Generating a Sentiment Lexicon for an unsupervised dataset.

2. An effective method is proposed for adapting a pre-trained model to a target domain, using Domain Independent Lexicon and the generated sentiment lexicon. Then the framework calculates an estimation of each sentence polarity to assign a tag to each unlabeled sample. In this case, we increase or decrease polarity of words in each sentence according to the predicted polarity of the sentence (i.e. it is positive or negative according to its DIL sentiment score).

1. A new Multilayer Perceptron model is proposed as the final classifier, which utilizes the generated sentiment lexicon as the weights of its first hidden layer, and is trained to output an accurate sentiment score.

The rest of the paper is organized as follows. Section 2 introduces some representative related works. The proposed model is described in details in Section 3. Section 4 explains the experimental results and analysis. Section 5 concludes and summarizes the paper as well as its limitations.

2. Related work

Sentiment Analysis has been an important problem in the NLP community for a long time [10,11]. It is defined as the computational treatment of opinions, sentiments and subjectivity in a text [12]. In this section several works related to the sentiment lexicon construction and domain adaption is introduced.

2.1. Sentiment lexicon construction

Sentiment Lexicon plays a crucial role in sentiment analysis. It is defined as a set of words or/and phrases with either numeric scores or positive/negative labels [1], in a way that shows the sentiment or emotional orientation of words, and has attracted a lot of research interests recently [13–15]. Sentiment Lexicons are mostly known as either domain-specific lexicons or general lexicons. An abundance of studies of developing a sentiment lexicon depend on the concept of general sentiment. In this regard, a general sentiment lexicon provides sentiment polarity of a word in general contexts which may not be adaptable to all possible domains. On the other hand, a domain-specific lexicon is adaptable to a specific lexicon and contains sentiment polarity of a word in that specific context. Following this, [16] classified the methods of generating domain-specific lexicons into three categories of linguistic rules-based approaches, corpora-based statistical learning approaches, and dictionary-based approaches.

2.1.1. Linguistic rules-based

Linguistic rules-based approaches refer to specific language phenomena, dependency relationships, conjunction words and so on, which are extracted from language contextual information or linguistic rules. Taking advantage of context-dependent opinion words, [17] proposed a method based on intra-sentence and inter-sentence conjunctive rules that assign sentiment polarities to opinion words according to specific opinion targets. This method determines sentiment polarities of opinion targets instead of generating a sentiment lexicon. Authors of [18] generated a topic-specific method by using a bootstrapping method and formed a general-purpose sentiment lexicon. This method is applied to opinionated blog post retrieval tasks to test its validity. Linguistic rules-based approaches suffer from corpora like blogs or comments in which people do not follow grammatical rules, and this adds to the difficulties in creating domain-specific sentiment lexicon [19].

2.1.2. Corpora-based statistical learning approaches

Corpora-based approaches retrieve words sentiment by using the mutual reinforcement between documents and words and by exploring practical usage text data which requires high quality corpora. This information can also be extracted using the co-occurrence pattern with the general sentiment lexicon. Ref. [20] uses the co-occurrence pattern of negations and adverbials to generate an automatic word lexicon. To extract the semantic orientation of a word, [21] proposed a method to infer it from the statistical association with a set of positive and negative words. In this approach, the semantic orientation of a given word is calculated from the strength of its association with a set of positive words, minus the strength of its association with a set of negative words. Proposing a clause-level sentiment analysis, [22] identified domain-specific sentiment atoms using the parse tree of each sentence and the general sentiment lexicon information.

2.1.3. Dictionary-based approaches

Dictionary-based approaches make use of previous sentiment lexicon resources in unsupervised way to generate a domain-specific sentiment lexicon. A semi-automatic approach is proposed by [23] for generating a sentiment lexicon which assigns sentiment values to sentiment terms via crowd-sourcing. In this regard, the authors introduced a bootstrapping process operating on unlabeled documents to extend the created lexicons, and to customize them according to the particular case. The in [24] extracts sentiments at the segment level by proposing a semi-supervised framework called ReNew to generate a domain-specific sentiment lexicon. ReNew adopts dependency relation pairs as basic elements in the generated sentiment lexicon in order to capture the sentiment of words.

2.1.4. Generating a sentiment lexicon

Researches have contributed two main approaches for creating sentiment lexicons, namely manual and automatic creations. Manual approaches are the most reliable and straightforward approaches since each sentiment value of a term is determined by experts, but in the same way they are the most labor intensive and expensive. MPQA [4], LIWC [25], Hu and Liu's Lexicon [26] and VADER [27] are some instances of these approaches. To create a Sentiment Lexicon using automatic approaches which require fewer human efforts and consequently results in faster inclusion of lexicon items, a substantial amount of approaches is proposed. Most of these approaches use a small group of seed words (e.g. great, hate, love, etc.) which are collected manually and make use of heterogeneous sources of information in corpus to propagate sentiment information from seed words. Syntactic rules [28], morphological relations [29] and co-occurrences [30] are some instances of these approaches.

2.2. Supervisory characteristics

2.2.1. Supervised and semi-supervised algorithms

Most of the studies regarding opinion mining and sentiment analysis are classified under supervised or semi-supervised categories. Such algorithms require labeled data which are formed as texts or words with their sentiment/opinion label (e.g. *love, joy, surprise, fear, sadness and etc.* in case of opinion or *positive, neutral and negative* in case of sentiment).

Authors of [31] proposed a supervised feature-based sentiment classification method to determine the sentiment of a document based on the features, such as emoticons, exclamation and question mark symbol and unigrams. To increase the efficiency of their sentiment classification approach, the authors compared the performance of different supervised classification methods and found discriminative multinomial Bayes and sequential minimal optimization approaches as the best considering the overall result.

Also [32] introduced a supervised machine learning approach based on Conditional Random Field (CRF) to extracts features and their corresponding opinions. Two datasets from shopping websites are considered. Each word in the datasets is tagged with its parts of speech (POS) after preprocessing. Preprocessing phase includes removing abnormal characters and HTML tags. Then each document is split into sentences using OpenNLP. At last, the proposed CRF model is trained on the datasets. To improve the performance of the aspect-level sentiment classification, [33] takes advantage of the Topical Word Embedding (TWE) model to learn aspect-specific word embedding with corresponding sentiment vector in semi-supervised way. Also [34] creates a semi-supervised learning framework that uses unsupervised information to improve SVM-based classification for tweet sentiment analysis.

2.2.2. Unsupervised algorithms

In contrast to the supervised algorithms which are time-consuming and expensive, unsupervised algorithms are proposed to determine the polarity score of a document without previous labeled data. The proposed approach in [35] fuses the information from wide coverage Urdu Sentiment Lexicon and the efficient Urdu Sentiment Analyzer to outperform supervised machine learning approaches (i.e. Support Vector Machine, Decision Tree and K Nearest Neighbor). An unsupervised polarity detection technique is introduced by [36] which employs a Word2Vec tool to label each word with a polarity score. The proposed approach computes the semantic word vectors of the preprocessed datasets. A prior knowledge of seed words is considered to calculate the polarities of each new word by determining the similarity between the word and its relative seed words.

In another research proposed by [37], a word vector representation and a K-means clustering is used to summarize the features of the products without any supervision. To do so, after removing stop words from the collected reviews, each unlabeled word is represented with a Word2Vec vector. Then the dataset is divided into different clusters based on the features of the products. The experiments are conducted on reviews collected from a shopping mall in China, using three typical feature mining approaches.

By taking advantage of the hierarchical relationship between different features of a product, [38] introduced an unsupervised approach to aggregate the sentiment about various features of a product and compute the overall polarity of a review. The importance of each feature in the product's tree is considered to leverage the performance. To generate the product-specific ontology tree, the ConceptNet is used [38]. Using frequent nouns used in a review of a customer and the extracted relationships from ConceptNet, the initial product tree is constructed. A dependency parser is then employed to detect opinion words describing each feature of product reviews. At last, a polarity aggregation formula is proposed to calculate the sentiment score by considering the depth of a feature node in the ontology tree (the closest node to the root has the highest importance).

2.3. Sentiment domain adaptation

In order to adapt a sentiment lexicon of the source domain to a target domain, transfer learning is usually used in most of the adaptation methods. In these group of approaches, sentiment classifiers are trained for one or more source domains with sufficient labeled samples, and then applied to the target domain where there are insufficient or no labeled data. In [39], the general sentiment information in sentiment lexicons is actively adapted to the target domain which does not rely completely on the labeled data of the source domain. Some studies have introduced methods to pass the information from person to person which is called as Word of Mouth (WOM). The approach proposed in [40] increases the WOM quality and builds the sentiment lexicons from the contextual information, which are adaptable to other domains. Finally, these contextual sentiment lexicons are integrated with preference vector modeling for WOM quality classification.

Another study introduced by [41] uses unannotated corpus and a dictionary to adapt sentiment lexicons for domain-specific classification. Two large corpora consisting of tweets related to stock market with 743,069 samples and one million political tweets are used for the experiments. Also, five lexicons are used for seeds and baselines. For the purpose of adaptive lexicon learning, [42] proposed a genetic algorithm for sentiment analysis in the microblogs. In [42] a combination of corpora-based and lexicon-based approaches is used. Then the proposed genetic algorithm, which generates adaptive sentiment lexicon, is used for sentiment classification which is an optimization problem.

In transfer learning based approaches, most domain adaptation models focus on homogeneous unsupervised domain adaptation (HoUDA). It means that the source and target domains have similar, same-dimensionality feature spaces and there are no labeled instances in the target domain [43]. However, heterogeneous unsupervised domain adaptation (HeUDA) models are proposed to handle the situation where the target domain is heterogeneous and unlabeled. In the field of domain adaptation, "heterogeneity" often refers to either the difference in the dimensionality of source and target domains or differences in the features of the two domains. The authors of [44] propose the Grassmann-Linear monotonic maps-geodesic flow kernel (GLG) model which presents an unsupervised knowledge transfer theorem that transfers knowledge from source domain to a heterogeneous and unlabeled target domain. Although GLG avoids

Table 1
Description of the notations used in the paper.

Notation	Definition
x	Represents each single word in a lexicon or a corpus. **X** refers to all the words inside them.
T	Represents the preprocessed input sentences as a *D-dimensional* vector. Each index of the record ***T*** represents the unique location of the word ***x*** in it.
pl(x)	Polarity score of each word ***x*** in the Source Sentiment Lexicon.
sc(x)	Polarity score of each seed word ***x*** retrieved from the general sentiment lexicons.
DIL	Domain Independent Lexicon containing seed words as well as their sentiment scores.
TPL	Primary Target Sentiment Lexicon generated in each epoch using ESL.
ESL	Evaluation Sentiment Lexicon containing ***pl(x)*** for each seed word ***x*** and ***sc(x)*** for the remaining ***X*** in ***DIL***.
TSL	Adapted Target Sentiment Lexicon.

negative transfer using Linear monotonic maps (LMM) mapping function, it does not provide sufficient theoretical guarantees to directly address multiclass classification problems. LMMs, key mapping functions in GLG, can only guarantee that the probability of positive labels (denoted by P1), will not change to 1 - P1 after mapping this instance set to its homogeneous representation.

A method called Geometric Knowledge Embedding (GKE) is proposed in [45] which exploits the geometric information of data in order to learn discriminative and transferable representations. It first applies a Convolutional Neural Network (CNN) to extract high-level features from the input data, and then constructs a graph called Graph Convolutional Network (GCN) on the extracted features to introduce the similarity relationship between both the source and the target domains to the model. Although the authors proposed a method to alleviate the issue of domain discrepancy, GCN works best on homogeneous domains where the source and the target domains have features in common.

Considering Online Heterogeneous Transfer learning problem (OHT), [46] developed a method to label online data (as target domain) using source data and co-occurrence data from the offline sources. In this regard, an offline decision function is built based on a heterogeneous similarity by taking advantage of the labeled source domain and the unlabeled co-occurrence data. Then, an online decision function is built on the target data sequence. Finally, to combine these two kinds of decision functions, a weighting strategy is applied in order to boost performance. The OHT approach is also limited when the number of training samples is limited; however our proposed approach mostly takes advantage of the DIL and the target domain features to build a sentiment lexicon for the sake of classification.

The proposed approach of [31] is the most similar study to this paper which introduced a sentiment lexicon adaptation method. In this respect, a fully labeled lexicon dataset is considered. If the predicted label of a sentence does not match the actual label of that sentence, the algorithm changes the polarity score of each word during the learning phase. The polarity of every single word of an incorrectly predicted sentence is changed by a constant small value (namely an epsilon value) to get the correct final result which makes an unrealistic amount of computation and in some cases, it reaches the threshold before reaching the expected result. Adding or subtracting this constant is a bottleneck to this approach since each individual word requires different amount of value to be added to, which does not guarantee the convergence of the polarity scores changing process. For example, in "The great plot of movie was unpredictable" if the polarity of the word "unpredictable" is taken from electronics domain which we assumed it to be negative, the value of epsilon which has to be added to "unpredictable" must differ from the epsilon value that should be added to word "great" which is positive in both domains.

On the other hand, the proposed approach of this paper adapts the sentiment lexicon of a source domain to an unsupervised target domain using information from both source domain and extracted Domain Independent Lexicon polarities. On the other hand, the proposed approach transfers the probability of each single word in source domain to an unlabeled heterogeneous target domain and optimizes itself according to the target domain solely, without experiencing negative transfer learning in the experiences. Consequently, it leads to a better convergence according to the results.

3. Proposed framework

The main objective of this research is to perform two main tasks of Sentiment Analysis simultaneously. The first task is to adapt a source sentiment lexicon to a target domain by generating a new domain-specific sentiment lexicon from the source domain. Furthermore, the proposed framework trains a neural network to assign positive or negative label to each document using extracted information from the generated Target Sentiment Lexicon (**TSL**) without any supervision. In order to leverage the performance of labeling the target domain, an Evaluation Sentiment Lexicon (**ESL**) is created by taking advantage of a prior knowledge consisting of *initial words* (*seed words)* extracted from three general sentiment lexicons which are intrinsically independent from the context Fig. 3. The proposed approach is based on three main phases namely, preprocessing, source-domain sentiment extraction and modeling, and target domain adaptation. The notations used in the remaining parts of the paper are presented in Table 1. From these steps, we emphasize on the second stage which is the main contribution of the framework and uses a Multilayer Perceptron model for effective polarity modeling and adaptation.

3.1. Preprocessing

Preprocessing plays a crucial role in analyzing and constructing a sentiment lexicon due to the importance of sentiment information in each word. There are plenty of challenges for NLP which are posed by the language used in social media; sites, blogs and etc. letter or word duplications and misspellings are examples of the problems with typographical characteristics of words.

For instance:

sweeeeeeeet → *sweet (replicated characters)*

w8 → *wait (abbreviations)*

wat r u doin? → *what are you doing? (misspellings)*

Additionally, English sentences are filled with words like "is, are, a, an, I, you, etc." which do not carry any particular polarity scores. Such words that are known as "stop words", have high frequency and by eliminating them the sentiment of a sentence stays the same. To remove stop words from the sentences, a file consisting of a list of stop words is applied. Finally, in order to reduce the sparsity caused by different forms of a same word, lemmatization is considered.

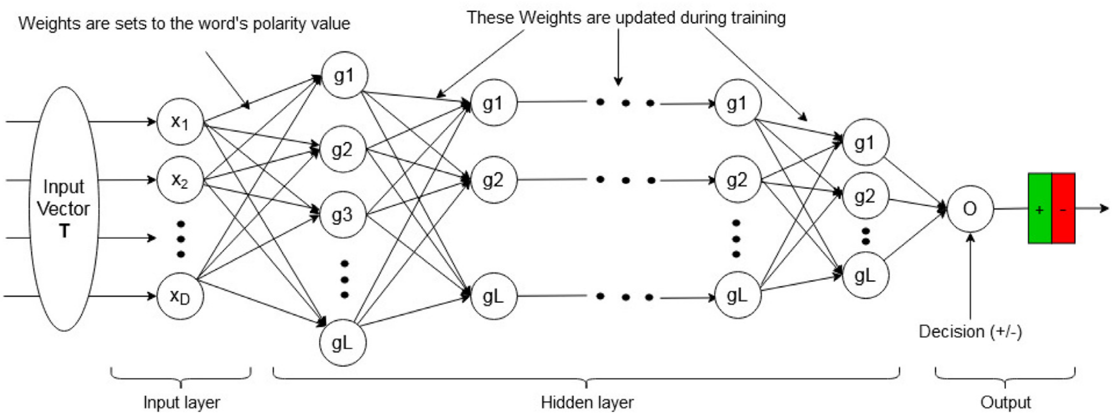

Fig. 1. Simplified illustration of the Multilayer Perceptron Model.

3.2. Source-domain sentiment extraction and modeling

3.2.1. Sentiment lexicon creation

A sentiment lexicon, also called a *polarity lexicon* or *an opinion lexicon* consists of a list of words with associated values representing their sentiment polarities. In this paper, sentiment value of a word $\boldsymbol{x}$, $\boldsymbol{pl}\,(\boldsymbol{x})$, is considered as a real-valued number. Apart from sentiment lexicon and term frequency, which this paper takes advantage of, there are plenty of other features proposed for sentiment classification such as POS tags [47], negators [48] and syntactic features [49]. However, this research focuses on adapting a sentiment lexicon rather than training the best classifier. Hence, only sentiment information is used. To generate the domain-specific sentiment lexicon, polarity score of each word, $\boldsymbol{pl}\,(\boldsymbol{x})$, is calculated. Firstly, the frequency of each word $\boldsymbol{x_i}$ in the positive corpus is subtracted from its frequency in the negative corpus, where $\boldsymbol{i}$ is the word index in the lexicon. Then, all extracted values are divided by a constant (vocabulary size of the corpus) to be normalized (Eq. (1)) without distorting the differences in the ranges of values or losing any information.

$$pl\,(x_i) = \frac{\#(x_i \text{ } in\ positive\ corpus) - \#(x_i \text{ } in\ negative\ corpus)}{\#(all\ words\ in\ corpus)} \tag{1}$$

Due to various connotations of each word in different domains, this approach is advantageous to obtain domain-specific sentiment of each word according to its context.

3.2.2. Constructing the multilayer perceptron

The main contribution of this paper and the proposed approach is to use Multilayer Perceptron to learn and modify the sentiment polarities of the words. There are plenty of studies using neural networks as an effective way of sentiment classification specifically Recurrent Neural Networks (RNNs), which are one of the earliest pioneer models in supervised domains [50–52]. However, training and testing these methods are time consuming, and without having sufficient supervised samples (with true label for each sample, which rarely happens in unsupervised domain adaptation), the obtained results from these classifiers may perform poorly. It is also worth mentioning that the amount of resources such as GPU, CPU and RAM used for the sake of processing in these classifiers are higher than an MLP model. On the other hand, we want both the general knowledge of the language and the embedded knowledge in a specific domain.

Hence, this study proposes a Multilayer Perceptron (MLP) with a new architecture in its first hidden layer; that outperforms many other RNNs. In this respect, the input layer of the MLP model is fed with record $\boldsymbol{T}$. The output is the predicted sentiment as either a positive or a negative real value. The proposed approach uses corresponding sentiment value $\boldsymbol{pl}(\boldsymbol{x})$ of each word $\boldsymbol{x}$ which is extracted from the source sentiment lexicon (as mentioned in Section 3.2.1) as the weights of the first hidden layer (see Fig. 1).

3.3. Target domain adaptation

3.3.1. Domain-independent feature selection

Words contained in a corpus can be divided into two groups: domain-specific words and domain-independent words [53]. Domain-specific words can have fluctuating sentiment polarities among different contexts and domains (e.g. "*high*" in "*high price*" that indicates a negative sentiment and "*high definition television*" which indicates a positive sentiment). On the other hand, *domain-independent* words are used to convey the same sentiment orientation in different domains such as "good, bad, hate and etc.". In this step, our work relies on domain-independent words extracted from three general sentiment lexicons, i.e., Bing Liu's sentiment lexicon[1] [26], SentiWordNet[2] [54] and MPQA [4][3] which are publicly available. SentiWordNet includes 117,686 terms with a positive and a negative score assigned to each term. We assign each term with a positive (negative) value if the corresponding positive score is higher (lower) than the negative score. Also, all words with equal positive and negative scores are removed as they are considered to be neutral and they barely carry any sentiment (Eq. (2)). Furthermore, since in SentiWordNet, "*polysomic words*" can have different sentiment scores if a word has more than one semantic meaning, we also remove them from the process to provide higher consistency.

$$polarity\,(x \in SentiWordNet) = \begin{cases} 1, & if\ (positive\_score\,(x) - negative\_score\,(x)) > 0 \\ -1, & if\ (positive\_score\,(x) - negative\_score\,(x)) < 0 \\ 0, & otherwise \end{cases} \tag{2}$$

Polysomic **Xs** with different sentiments are ignored. Bing Liu's sentiment lexicon [26] consists of 2006 positive terms and 4783 negative terms. These words are labeled with either "positive" or "negative" to express the sentiment inclination. In the MPQA, 6886 words are provided with their sentiment scores as well as their POS tags. Following the applied rule on SentiWordNet, all words with different POS tagging that have different polarity values are removed.

Since the construction of these general sentiment lexicons is based on various semantic thesaurus, corpora and methods,

1 https://www.cs.uic.edu/liub/FBS/sentiment-analysis.html#datasets.

2 http://sentiwordnet.isti.cnr.it/.

3 http://mpqa.cs.pitt.edu/.

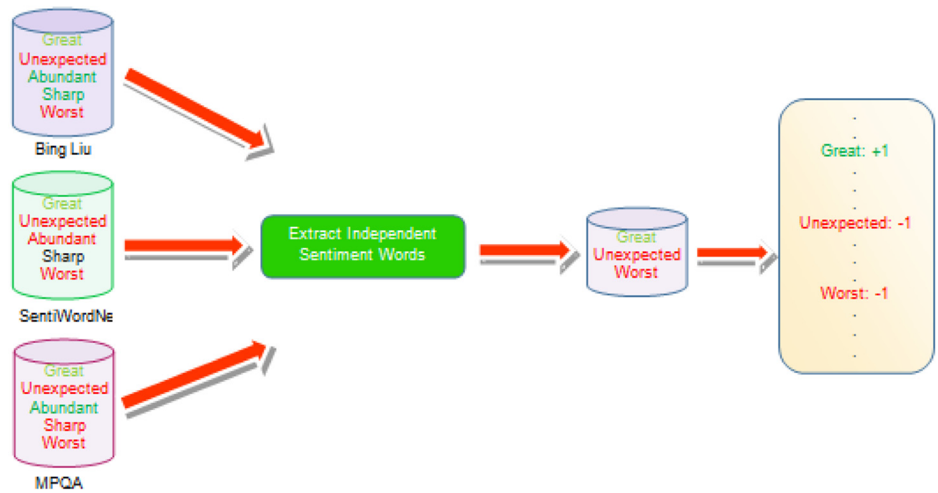

Fig. 2. An illustrative model of generating Domain-Independent Lexicon. The words in green, red and black colors in the sentiment lexicons represent positive, negative and neutral words respectively. . (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

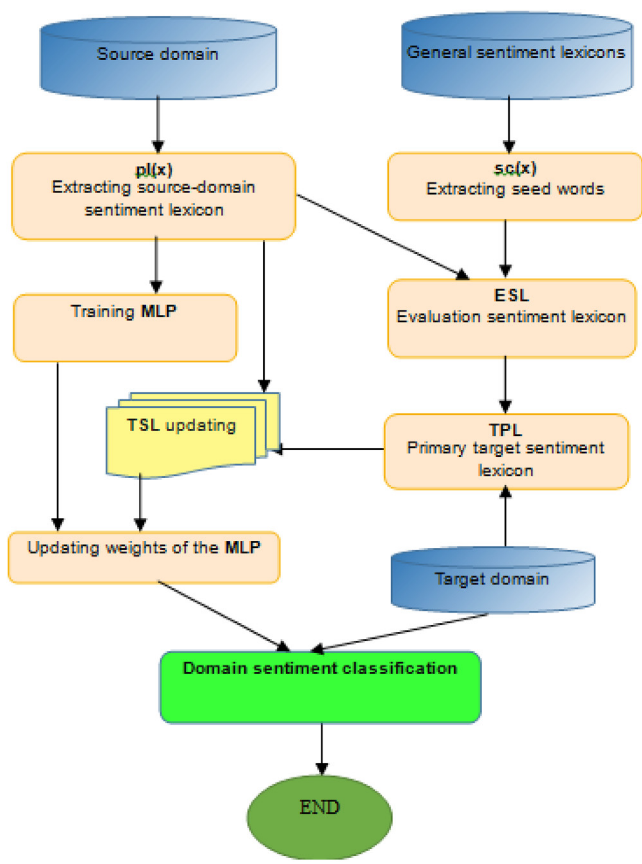

Fig. 3. The workflow of the proposed framework.

domain-specific words in these lexicons usually do not have the same polarity score. For instance, in Bing Liu's sentiment lexicon [26] "*abundant*" expresses a positive sentiment in contrast to SentiWordNet sentiment lexicon [54] in which it has negative polarity. For this purpose, these three general sentiment lexicons are fused so their sentiment information is extracted using the method described below. By this approach, we overcome the inconsistencies to build the primary Domain-Independent Lexicon (DIL). An illustrated preview of the approach is shown in Fig. 2.

The retrieved sentiment information for each word $\boldsymbol{x}$ in these three lexicons are used to construct the DIL as follows:

$$sc\,(x) = \begin{cases} 1, & \textit{if } x \textit{ is positive in all three sentiment lexicons} \\ -1, & \textit{if } x \textit{ is negative in all three sentiment lexicons} \\ 0, & \textit{otherwise} \end{cases} \tag{3}$$

In this end, a total number of 2213 domain-independent words $\boldsymbol{x}$ as well as their polarity value of $\boldsymbol{sc}(\boldsymbol{x})$ is obtained. Since the purpose of this paper is to apply sentiment polarities of a source domain to an unseen target domain, it is essential to only take advantage of seed words that each of them has a unique polarity in all sentiment lexicons, which reduces the chances of uncertainty in unseen domains.

3.3.2. Target lexicon adaptation

The purpose of this subsection is to adapt the source domain sentiment lexicon to a target domain in an unsupervised manner using the constructed DIL as prior knowledge. After applying preprocessing on the target corpus, all keywords are extracted for generating a new primary Target Sentiment Lexicon (***TSL***).

3.3.2.1. Evaluating sentiment lexicon creation. At this point, to leverage consistency and to increase efficiency of overall classification performance, all polarity values $\boldsymbol{sc}(\boldsymbol{x})$ in DIL are substituted with sentiment polarity $\boldsymbol{pl}(\boldsymbol{x})$ of its corresponding word in the source sentiment lexicon. If a word does not exist in source sentiment lexicon, the value in DIL remains unchanged (Eq. (4)). This new generated sentiment lexicon which is called Evaluating Sentiment Lexicon (**ESL**) is created to evaluate each sentence in the target corpus. The usage of ESL is precisely described in Section 3.3.2.2.

$$ESL_x = \begin{cases} x\colon pl\,(x) & \textit{if } x \in \textit{source sentiment lexicon} \\ x\colon sc\,(x) & \textit{if } x \notin \textit{source sentiment lexicon} \end{cases} \tag{4}$$

where $\boldsymbol{ESL_X}$ contains a list of ordered pairs of the word $\boldsymbol{x}$, and its corresponding polarity.

3.3.2.2. Primary target-domain classification. In this end, we assign a "positive" or "negative" label to each input sentence using **ESL** according to the following approach. Although in an unsupervised manner the samples are not associated with sentiment labels explicitly, many of them contain strong sentiment orientations and can provide rich sentiment information. For instance, by only taking advantage of *seed words* (e.g. *great*, *mind-blowing*, *fabulous* as positive words), we can deduce that the sentiment orientation of the phrase "*I enjoyed the Great acting with mind-blowing setting and fabulous plot*" is positive.

Following the above hypothesis, an approach is proposed to predict the label of each record ***T*** using the ESL sentiment information. To classify each record ***T*** into a positive or negative category, we substitute each ***w*** in ***T*** with its corresponding polarity score from **ESL**. The polarity label of record ***T*** is determined by accumulating the polarities of the containing words. The positive (negative) value obtained from this calculation categorizes the record ***T*** in positive (negative) class. Following is an example of the approach:

Example 1.

$$ESL = \{(great, 0.271), (hate, -0.307), (excellent, 0.198), (unexpected, -0.002)\}$$

Original input sentence of:

T = "Great plot of movie with excellent acting of scarlett johansson, I'm wondering how people hate it"

Which is converted to:

$$T = 0.271 + 0 + 0 + 0 + 0 + 0.198 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + 0 + (-0.307) = 0.162$$

Therefore, the above example is classified in the *positive* category.

Note: In the conducted experiments, the preprocessed sentences are considered but for the sake of readability, the original sentences are used here.

3.3.2.3. Target-domain sentiment lexicon creation. Consequently, all records from the target domain are assigned with a label tag. Using (Eq. (1)) to calculate the polarity ***TPL(x)*** of each word ***x***, sufficient sentiment information to construct Target-Domain Sentiment Lexicon (***TSL***) according to the (Eq. (5)) is obtained. The proposed approach constructs the TSL using sentiment information from both source sentiment lexicon and the generated primary target-domain lexicon (TPL) in order to overcome their limitations regarding the wrong sentiment inclination of domain-dependent adjectives.

$$TSL_x = \{x: pl(x) + TPL(x), \forall x \in TPL\} \tag{5}$$

For a better consistency, these values are updated in multiple iterations:

$$TSL'_x = \{x: TSL(x) + TPL(x), \forall x \in TPL\} \tag{6}$$

where $\boldsymbol{TSL_x}$ is the polarity of the word ***x*** in the previously generated Target Sentiment Lexicon, while the $\boldsymbol{TSL'_x}$ is the ordered pair of the word ***x*** and its corresponding polarity in the new adapted lexicon.

3.3.2.4. Lexicon expansion. In order to improve the robustness of the classification, we propose an approach to expand and update **ESL** using sentiment information from three sources, namely tagged corpus of target domain which contains the sentiment polarities of TPL words with "adjective", "adverb" or "noun" POS tag, source sentiment lexicon which corresponds to the polarity extracted from the Source Sentiment Lexicon and the previously computed polarity from ESL. In this regard, the polarities retrieved from the mentioned three sources are summed together and used to update the polarity of the words in the ESL.

$$ESL'_x = \{x: pl(x) + TPL(x) + ESL(x) \,|\, x \in POS\,(adjective, adverb, noun), \forall x \in TSL\} \tag{7}$$

In which ***ESL(x)*** is the polarity of the word ***x*** in the ESL. If a word ***x*** does not exist in neither ***pl*** (***x***) nor ***TPL(x)***, the default value of 0 is assigned to it.

3.3.3. Target-domain sentiment classification

Another important step of the proposed approach which takes benefit from the sentiment model is to classify the sentences from the target domain using the updated MLP model. The generated Target Sentiment Lexicon (TSL) is used to update the previously trained Multilayer Perceptron (MLP) to get the best result. In this regard, the weights of MLP model except for the first layer is updated until it reaches a consistency and outputs the sentiment score of a document, accordingly. Both the accuracies from utilizing just the sentiment information and the neural network are computed and compared. The complete workflow of the proposed approach is shown in Fig. 3.

4. Experiments

In this section, the proposed approach is validated against different datasets and lexicons which are described in Section 4.1. The experiments are conducted on two datasets and three general sentiment lexicons in order to evaluate the performance of model. Then, the proposed unsupervised target domain adaptation model is compared with several state-of-the-art methods and the effectiveness of these approaches on multiple domains is discussed.

This study uses *Accuracy*, *F1 measure*, *Precision* and *Recall* as the evaluation metrics. Accuracy is defined as proportion of correctly classified documents to the total number of documents (Eq. (8)); in this way, error rate refers to incorrectly classified documents to correctly classified documents.

$$acc(T_x) = \frac{\sum_{j \in T_x} \#\ of\ \hat{y}_j = y_j}{\#\ of\ T_x} \tag{8}$$

The Precision and Recall are computed using (Eq. (9)) and (Eq. (10)), respectively. In this respect, Precision is defined as the ratio of correctly predicted positive observation to the total predicted positive observations, and Recall is defined as the ratio of correctly predicted positive observations to the all observations in actual class.

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

where TP (True Positive) refers to the number of positive tuples classified correctly as positive by the classifier; FP (False Positive) refers to the number of negative tuples wrongly labeled as positive; and FN (False Negative) refers to the number of positive tuples wrongly labeled as negative.

Following the above, the F measure is defined as the weighted harmonic mean of the Precision and Recall and is computed according to (Eq. (11)).

$$F1 - measure = \frac{2 * Precision * Recall}{Precision + Recall} \tag{11}$$

4.1. Datasets and lexicons

The proposed approach is evaluated on Multi-Domain Sentiment Dataset v2.0[4] collected by [55]. It contains product reviews of Amazon.com from many product domains. Each balanced domain in Amazon dataset has 1000 positive reviews as well as 1000 negative reviews in XML format. The second dataset containing Movie reviews is Sentence Polarity dataset v1.0[5] [56] which contains 5331 positive and 5331 negative sentences. Our

[4] http://www.cs.jhu.edu/~mdredze/datasets/sentiment/.

[5] http://www.cs.cornell.edu/pabo/movie-review-data/.

Table 2
Statistics for balanced datasets.

Domain	Apparel	Electronics	Kitchen	Health	Movie	Book	DVD	Average
Positive	1000	1000	1000	1000	5331	1000	1000	1618.71
Negative	1000	1000	1000	1000	5331	1000	1000	1618.71
Unlabeled	7252	21,009	17,856	5225	0	973,194	122,438	163,853.4

Table 3
Statistics for imbalanced datasets.

Domain	Automotive	Musical	Tools	Gourmet food	Grocery	Jewelry	Baby	Average
Positive	584	284	98	1000	1000	1000	1000	709.42
Negative	152	48	14	208	352	292	900	295.14
Unlabeled	0	0	0	367	1280	689	2356	670.28

Table 4
Statistics for general sentiment lexicons.

	SentiWordNet	MPQA	Bing Liu's
Positive words	13,128	2718	2036
Negative words	14,726	4912	4814
All words (including neutrals)	117,659	8222	6850

experiments are based on seven balanced domains of *Apparel, Electronics, Kitchen, Healthcare, Book, DVD and Movie* (Table 2) and seven imbalanced domains of *Automotive, Jewelry, Musical, Tools, Gourmet food, Grocery and Baby* (Table 3) to demonstrate the effectiveness of the proposed approach.

The General Sentiment Lexicons used as prior knowledge are described in the following paragraphs. These lexicons contain words as well as their sentiment polarity (positive, negative and neutral in some cases) which are publicly available. In the proposed model, we only take advantage of positive and negative words, and discard the neutral ones.

SentiWordNet v.3[6] [54]: this sentiment lexicon is specifically designed for sentiment and opinion mining with 117,686 terms. Each term in SentiWordNet is assigned with a positive and negative value within [0, 1] range. For the experiments, by subtracting the negative value from positive value of a word, the aggregated polarity value is computed (Eq. (2)). For instance, the word "*dangerous*" with positive value of *0.0* and negative value of 0.75 is assigned with the polarity of −0.75. Also, all terms with equal values of positive and negative polarity are removed from lexicon since they are known as *neutral words* and does not convey any sentiment information.

Bing Liu's [26]: Including 2006 positive words and 4783 negative words. This lexicon consists of sentimental words, slangs, misspelled words and common variants. The assigned polarity for each word is either positive (+1) or negative (−1).

MPQA [4]: Consists of 6886 words as well as their sentiment polarities, each word is also assigned with one or more Part of Speech (POS) tag(s). In the experiments, we only take advantage of the words with the same polarity score for all POS tags in order to leverage the consistency.

As explained in Section 3, a selection set of seed words is extracted from all of the abovementioned lexicons to improve the performance of final accuracies. Consequently, a total number of 2213 domain-independent words are obtained. Table 4 provides statistical summaries of these lexicons.

4.2. Performance evaluation

In this section, the effectiveness of the proposed approaches is examined. Several experiments are conducted to compare our method with some baseline methods. To do so, the accuracy of different baseline methods is compared with the proposed method using the best fit domains as starting point. Next, the second experiments, the results from some more baseline methods are represented in more details. In this part, the source domains and the target domains of methods are considered to be the same. Finally, the effectiveness of the proposed approaches is evaluated on multiple imbalanced heterogeneous domains.

4.2.1. Experiment I

The first experiment is performed on two sets of domains. First set of domains includes *Apparel, Electronics, Kitchen, Healthcare and Movie*. The results are illustrated in Table 5. Next set of domains involves *Book, DVD, Electronics and Kitchen* that the results are presented in Table 6. Applied methods on these domains are as follow:

(1) Cognitive-inspired Domain Adaptation with Higher-level Supervision (CDAHS) [31] which develops heuristic rules like exploration–exploitation trade-off and some cognitive constraints in its model in order to create a sentiment lexicon and expects to change the polarity of all words from source domain to adapt them to a target domain. Although it uses some labeled data from target domain which is considered to be a semi-supervised learning algorithm, this is the closest proposed method to our approach to the best of our knowledge.

(2) Delta TFIDF [57]: *Delta Term Frequency-Inverse Document Frequency* uses no prior knowledge for word sentiment. It is an intuitive general-purpose technique to efficiently weight word scores before classification.

(3) *SVM with* Naïve Bayes features (NBSVM) [3]: NBSVM is used to increase the performance of SVM *(Support-vector Machine)*. For this purpose, if *SVM* model is not confident enough, it uses Naïve Bayes log-count ratios as features for the modification.

(4) LS, SVM, LR [58]: *Least Squares method, Support Vector Machine and Logistic Regression* are three known supervised classification methods. These three methods are trained on labeled corpora of target domains, so they only take advantage of the domain-specific sentiment information of target domain.

(5) ASL [59] uses integer linear programming to adapt an existing sentiment lexicon. It requires previous labeled samples of the target domain and therefore it is categorized under the domain-specific sentiment classification approaches.

(6) Automatic Induction of lexicon Polarity Scores (AIPS) [60] is one of the most noticeable methods which are compared in this paper.

(7) DIL classification: Sentiment Classification using the sentiment lexicon (ESL) introduced in Section 3.3.2.2. An example of this classification is provided in Example 1.

(8) MLP (Multilayer Perceptron) neural network; to demonstrate the effectiveness of the proposed method, and to ensure that the generated sentiment lexicon is an important key to the method, all layers of the MLP model are trained.

[6] Implemented with NLTK.

Table 5
Sentiment classification accuracies (in percent) for different balanced domains.

	Apparel	Electronics	Kitchen	Healthcare	Movie	Average
NBSVM	75.4	65.8	67.1	65.2	77.9	70.2
TFIDF	74.2	66.0	65.0	65.0	75.4	69.1
AIPS	54.5	53.3	51.3	49.0	53.1	52.2
CDAHS	74.7	69.2	69.7	66.5	77.9	71.6
DIL	72.9	72.8	73.8	71.9	61.6	70.6
MLP	50.0	50.0	49.3	49.6	50.1	50.0
The proposed approach	**75.6**	**77.6**	**76.3**	**76.5**	**64.5**	**74.1**

Table 6
Sentiment classification accuracies (in percent) for different imbalanced domains.

	Book	DVD	Electronics	Kitchen	Average
LS	69.1	70.6	75.5	77.8	73.2
SVM	69.7	70.4	75.1	77.2	73.1
LR	70.3	71.2	76.5	**78.5**	74.1
ASL	62.7	64.7	66.1	68.3	64.4
DIL	71.4	66.7	72.8	73.8	71.1
MLP	50.1	49.1	51.2	50.6	49.9
The proposed approach	**73.0**	**72.4**	**77.6**	76.3	**74.8**

Table 7
Sentiment classification accuracies (in percent) for different domains with respect to their source domains.

	B → D	B → E	B → K	D → B	D → E	D → K	E → B	E → D	E → K	K → B	K → D	K → E
Naïve Bayes	50.35	48.55	51.95	50.1	48.85	50.3	49	51.45	49.4	50.55	50.45	51.4
SVM	51.6	50	51.05	51.05	48.55	50.2	50.55	49.95	52.2	52.2	51.6	49.2
DIL	**67.88**	**65.48**	**66.48**	71.48	66.88	66.78	**63.28**	64.13	73.88	62.23	62.98	74.83
MLP	49.13	49.64	50.56	50.34	50.00	49.67	49.47	50.01	50.13	50.25	49.03	49.18
SWN MLP	59.53	57.07	58.13	**74.09**	67.88	**69.18**	**64.03**	63.73	76.30	65.88	66.23	76.24
MSWN	61.88	58.38	58.33	73.64	67.93	68.13	63.68	63.68	75.04	65.73	**66.98**	77.03
The proposed approach	67.73	60.38	61.63	73.08	**68.68**	68.13	62.53	**64.63**	**76.33**	**66.48**	65.53	**77.43**

It is also worthy to note that in case of target domains, above methods either use semi-supervised or supervised learning algorithms, while we only take advantage of target documents without any supervision. Represented accuracies of CDAHS [31] shown in Table 5 is considered to be the best results from using either Opinion Lexicon, SentiWordNet, L&M or SenticNet as starting point.

In the same way, the domains used as the source-domain for the proposed approach and DIL method in experiments of Tables 5 and 6 are considered to be the best results among all other domains due to the majority of domain-dependent words with the same sentiment orientation in both source and target domain. Also as seen in the above tables, using solely the MLP model does not have satisfying results at all. These domains are as follows:

Kitchen → Apparel; Apparel → Electronics; Electronics → Kitchen; Apparel → Healthcare; DVD → Movie; DVD → Book; Music → DVD, where the word before arrow indicates the source domain and the word after arrow indicates the target domain.

4.2.2. Experiment II

In this subsection, the comparison between the proposed method and other baseline methods is discussed in more detail. In this regard, we demonstrate the effectiveness of the proposed method by comparing it with *Naïve Bayes* and *SVM classification* methods. For the conducted experiments, Rapid Miner 5.3.015 software is used. For further analysis, a text mining plugin from the software is used to convert non-structured textual data into structured format.

As the only configuration, linear kernel is used for *SVM*. Also, to demonstrate the effectiveness of fusing three general sentiment lexicons, more experiments are conducted fusing only two general sentiment lexicons (MPQA and SentiWordNet, denoted as MSWN), and solely SentiWordNet (denoted as SWN-MLP). Domains used for these experiments are *Book (B), DVD (D), Electronics (E) and Kitchen (K)* as source and target domains, including 12 cross-domains of B → D; B → E; B → K; D → B; D → E; D → K; E → B; E → D; E → K; K → B; K → D; K → E. Accuracy results are represented in Table 7.

According to Tables 5–7, it is evident that the proposed model either outperforms the DIL classification or has almost the same accuracy. This is the result of training the MLP by initializing the best first hidden layer's weights with the generated sentiment lexicon (**TSL**).

As the experimental results of Table 7 denotes, although taking advantage of only one or two general sentiment lexicons to build the DIL, results in almost the same outputs as fusing all three general sentiment lexicons, in unseen domains, it is more convenient to fuse all three general sentiment lexicons to ensure that a seed word has the same sentiment in all contexts and domains.

4.2.3. Qualitative analysis

For evaluating the model, four evaluation metrics, i.e., accuracy, F1 measure, precision and recall are used to show the effectiveness of the model's performance.

In terms of accuracy, two state-of-the-art deep networks are considered, and the results are compared and illustrated in Tables 8 and 9. These deep networks are:

1. Bidirectional Long Short-Term Memory (Bi-LSTM) network [61], which is defined as an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Having access to both past and future input features of a sequence for a given time, Bi-LSTM network can effectively utilize past features (via forward states), and future features (via backward states) for a specific time frame.
2. Bidirectional Encoder Representations from Transformers (BERT) model [62] is designed to pre-train deep bidirectional representations from unlabeled text by jointly conditioning on both left and right context in all layers. It introduces two self-supervised tasks for learning contextual word representations: masked language modeling and next sentence prediction.

Table 8
Sentiment classification accuracies (in percent) for multiple target domains using Baby, Automotive and Musical as source domain.

	B → A	B → M	B → T	B → G	B → Gr	B → J	A → G	A → M	A → T	A → J	M → G	M → A	M → T	M → J
BiLSTM	76.49	65.05	64.36	75.13	52.05	50.32	47.93	41.96	61.56	52.09	47.62	52.12	49.56	50.21
BERT	79.45	85.80	88.28	82.85	74.01	77.45	**82.85**	**85.80**	81.08	77.45	82.85	79.45	**88.28**	77.45
DIL	84.21	89.72	87.38	**84.92**	79.71	82.33	**82.85**	85.49	87.38	77.45	**82.93**	**80.00**	87.38	77.38
The proposed approach	**85.44**	**89.74**	**87.39**	82.93	**80.53**	**83.42**	82.77	85.50	**87.39**	**77.53**	82.77	79.46	87.39	**77.46**

Table 9
Sentiment classification accuracies (in percent) for multiple target domains using Tools, Gourmet Food, Grocery and Jewelry as source domain.

	T → M	T → G	T → A	T → J	G → A	G → M	G → T	G → J	Gr → A	Gr → G	Gr → M	Gr → T	J → A	J → G	J → M	J → T
BiLSTM	52.74	49.39	53.80	57.59	48.97	57.10	49.08	49.71	47.85	49.27	56.86	42.86	48.22	47.78	53.14	48.23
BERT	**85.80**	**82.85**	79.45	**77.45**	79.45	**85.80**	**86.48**	77.45	79.45	82.85	**85.80**	**88.28**	79.45	82.85	85.80	88.28
DIL	85.49	83.18	**79.59**	**77.45**	79.45	85.49	87.38	77.38	79.49	82.93	**85.80**	87.38	79.59	**83.09**	85.49	87.38
The proposed approach	85.50	82.60	79.32	77.23	**79.46**	85.50	87.39	**77.61**	**79.59**	**83.09**	85.50	87.39	**79.72**	**83.09**	**86.10**	**87.39**

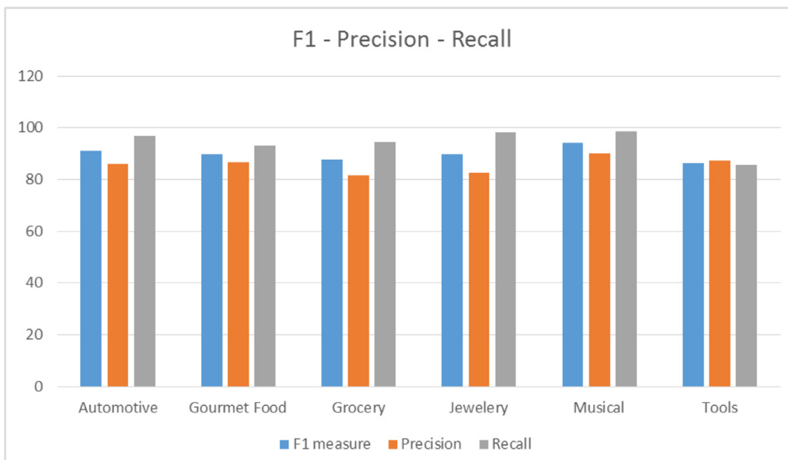

Fig. 4. F1 measure, Precision and Recall resulting from various domains.

The two deep network models are used as a classifier, after labeling the unsupervised dataset using the proposed DIL method. These experiments are conducted on 7 heterogeneous imbalanced domains, that are Automotive (A), Musical(M), Tools (T), Gourmet food (G), Grocery (Gr), Jewelry (J) and Baby (B). Tables 8 and 9 demonstrate the accuracies obtained from these domains, which fluctuates between **77.23** and **89.7**. Reviews from *Baby* sub-directory as source domain and *Musical* as target domain has the best accuracy, and most compatibility.

It is evident that the Bi-LSTM model has proved a poor performance in terms of accuracy, comparing to the other models. Furthermore, the retrieved accuracies from BERT model are almost the same as the proposed MLP model, and even outperformed in ten cross-domains (out of 30 experiments) by a slight percentage. However, it is worth mentioning that the MLP model uses much less resources (CPU, RAM and GPU), as well as having a faster training time that makes it more beneficial, in comparison with the BERT model. These experiments are conducted on the cross-domains used in Tables 8 and 9. Also in terms of F1-measure, Precision and Recall, Fig. 4 reports the results of *Automotive, Gourmet food, Grocery, Jewelry, Musical and Tools* domains as the target domains and *Baby* domain as the source, which have the most compatibility in case of accuracy (Table 8).

In contrast to the domain-independent words, experiments show that the domain-dependent words convey different sentiment orientation among various domains after adaptation (Fig. 5). In Tables 8 and 9, it is evident that using cross-domains of T → j, M → j and A → j which have the least domain-dependent words with the same sentiment orientation lead to a lower accuracy compared to B → M, A → T and J → T.

Fig. 5 shows the domain-dependent words that most intensively changed their polarity scores according to the proposed approach. However, in the experiments, many word polarities changed from negative(positive) to positive(negative), neutral to positive(negative) or even enhanced their sentiment score towards their sentiment orientation.

5. Conclusion and future works

Sentiment Analysis is a popular research area and has attracted lots of attentions recently. Yet, domain adaptation with almost no supervision is becoming a hot topic in this field. This paper proposed a novel sentiment classification model to tackle two main tasks regarding sentiment analysis i.e., generating a domain-specific sentiment lexicon from a source domain and a method for classifying documents into two classes of "positive" and "negative" using Neural Network. The proposed method for the first task extracts sentiment information from a source domain and adapts it to the target domain by taking advantage of a Domain-Independent Lexicon (DIL). The DIL, which is consisted of words as well as their polarities, is constructed from three general sentiment lexicons, i.e., SentiWordNet, MPQA and Bing Liu's sentiment lexicon. For this purpose, words with the same polarities in three lexicons are considered. Then, a Multi-Layer Perceptron is trained on source domain and target domain, respectively. The proposed approach uses the generated sentiment lexicon of both source and target domains as the weights of first hidden layer.

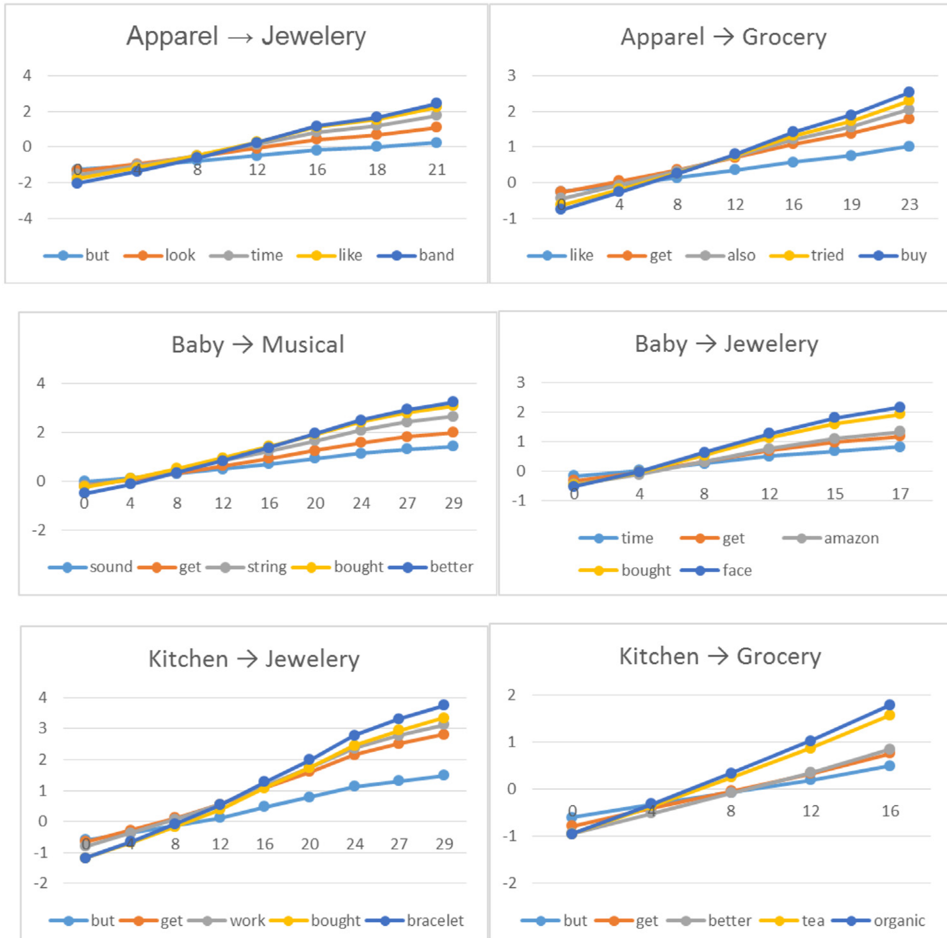

Fig. 5. Sentiment shifts of 30 sample words in different domains.

The main contribution of this paper is that it is unsupervised on the target domain. To the best of our knowledge, few studies have focused on adapting a sentiment lexicon without any supervision. Despite of this, the performance of the proposed method outperforms several supervised and semi-supervised popular methods. Moreover, in terms of accuracy, this work has no negative learning in case of using sufficient source domain. The experiments are conducted on two datasets: 13 domains from the Amazon dataset and the Movie dataset. The experimental results confirm the effectiveness of the model. According to the results, the proposed approach is competitive with the state-of-the-art methods. However, this work has its own limitation. That is, each domain uses different range of vocabularies from one another. In this case, the chances that source domain has the same vocabularies as the target domain does, may decrease. In future works, we aim to investigate into (a) overcoming above-mentioned limitations by fusing multiple source domains and using their sentiment information to generate a target sentiment lexicon; (b) extending the proposed approach to concept-level sentiment adaptation.

CRediT authorship contribution statement

Omid Mohamad Beigi: Developing the idea, Software development, Experimentation, Writing, Original draft preparation, Visualization. **Mohammad H. Moattar:** The main Idea, Methodology, Experimental design, Supervision, Writing, Reviewing and editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] B. Liu, Sentiment analysis and opinion mining, in: Synthesis Lectures on Human Language Technologies, 2012, pp. 1–167.
[2] B. Pang, L. Lee, Opinion mining and sentiment analysis, Found. Trends Inf. Retr. 2 (1–2) (2008) 1–135.
[3] F. Ribeiro, M. Araújo, P. Gonçalves, M. Gonçalves, F. Benevenuto, A benchmark comparison of state-of-the-practice sentiment analysis methods, EPJ Data Sci. 5 (2016) 1–29.
[4] T. Wilson, J. Wiebe, P. Hoffmann, Recognizing contextual polarity in phrase-level sentiment analysis, in: In Proceedings of the Conference on Human Language Technology and Empirical Methods in Natural Language Processing, 2005.
[5] F. Wu, Y. Huang, Y. Song, S. Liu, Towards building a high-quality microblog-specific chinese sentiment lexicon, Decis. Support 87 (2016) 39–49.
[6] X. Ouyang, P. Zhou, C.H. Li, L. Liu, Sentiment analysis using convolutional neural network, in: IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications;Pervasive Intelligence and Computing, 2015.
[7] Y. Bengio, R. Ducharme, P. Vincent, C. Jauvin, A neural probabilistic language model, J. Mach. Learn. Res. (2003) 1137–1155.
[8] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up? sentiment classification using machine learning techniques, 2002, cs.CL/0205070.
[9] B. Chen, L. Zhu, D. Kifer, D. Lee, What is an opinion about? exploring political standpoints using opinion scoring model, in: Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, Atlanta, Georgia, USA, 2010.

[10] A.L. Maas, R.E. Daly, P.T. Pham, D. Huang, A.Y. Ng, C. Potts, Learning word vectors for sentiment analysis, in: NAACL HLT 2011, Association for Computational Linguistics, Portland, Oregon, USA, 2011.
[11] S. Wang, C.D. Manning, Baselines and bigrams: Simple, good sentiment and topic classification, in: Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics.
[12] B. Pang, L. Lee, Opinion mining and sentiment analysis, Found. Trends Inf. Retr. 2 (2008) 1–135.
[13] G. Qiu, B. Liu, J. Bu, C. Chen, Expanding domain sentiment lexicon through double propagation, in: In Proceedings of the 21st International Joint Conference on Artificial Intelligence, 2009.
[14] D. Tang, F. Wei, B. Qin, M. Zhou, T. Liu, Building large-scale twitter-specific sentiment lexicon: A representation learning approach, in: In Proceedings of COLING 2014, the 25th International Conference on Computational Linguistics:Technical Papers, 2014.
[15] A. Neviarouskaya, H. Prendinger, M. Ishizuka, Sentiful: Generating a reliable lexicon for sentiment analysis, in: In 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, 2009.
[16] C. Zhao, S. Wang, D. Li, Exploiting social and local contexts propagation for inducing chinese microblog-specific sentiment lexicons, Comput. Speech Lang. 57–81 (2019) 57–81.
[17] X. Ding, B. Liu, P.S. Yu, A holistic lexicon-based approach to opinion mining, in: In Proceedings of the 2008 International Conference on Web Search and Data Mining, 2008.
[18] V. Jijkoun, M. de Rijke, W. Weerkamp, Generating focused topic-specific sentiment lexicons, in: Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics. ACL, 2010.
[19] H. Cho, S. Kim, J. Lee, J.-S. Lee, Data-driven integration of multiple sentiment dictionaries for lexicon-based sentiment classification of product reviews, Knowl. Based Syst. (2014) 61–71.
[20] T. Zagibalov, J. Carroll, Automatic seed word selection for unsupervised sentiment classification of Chinese text, in: Proceedings of the 22nd International Conference on Computational Linguistics, 2008.
[21] P. Turney, M. Littman, Measuring praise and criticism: Inference of semantic orientation from association, ACM Trans. Inf. Syst. 4 (2003) 315–346.
[22] H. Kanayama, T. Nasukawa, Fully automatic lexicon expansion for domainoriented sentiment analysis, in: Proceedings of the 2006 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2006.
[23] A. Weichselbraun, S. Gindl, A. Scharl, Using games with a purpose and bootstrapping to create domain-specific sentiment lexicons, in: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, ACM, 2011.
[24] Z. Zhang, M. Singh, Renew: A semi-supervised framework for generating domain-specific lexicons and sentiment analysis, in: Proceedings of the 52nd Annual Meeting of Association for Computational Linguistics. ACL, 2014.
[25] J.W. Pennebaker, M.E. Francis, R.J. Booth, Linguistic Inquiry and Word Count:LIWC, Erlbaum, Mahwah, NJ, 2001.
[26] M. Hu, B. Liu, Mining and summarizing customer reviews, in: In Proceedings of the tenth ACM SIGKDD international conference on Knowledge Discovery and Data Mining, 2004.
[27] C.H.E. Gilbert, Vader: A parsimonious rule-based model for sentiment analysis of social media text, in: In Eighth International Conference on Weblogs and Social Media (ICWSM-14), 2014.
[28] G. Qiu, B. Liu, J. Bu, C. Chen, Opinion word expansion and target extraction through double propagation, Comput. Linguist. 37 (2011) 9–27.
[29] V. Hatzivassiloglou, K.R. McKeown, Predicting the semantic orientation of adjectives, in: Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics, Association for Computational Linguistics, 1997, pp. 174–181.
[30] S. Tan, Q. Wu, A random walk algorithm for automatic construction of domain-oriented sentiment lexicon, Expert Syst. Appl. (2011) 12094–12100.
[31] F.Z. Xing, F. Pallucchini, E. Cambria, Cognitive-inspired domain adaptation of sentiment lexicons, Inf. Process. Manage. 56 (2019) 554–564.
[32] A.K. Samha, Y. Li, J. Zhang, Aspect - based opinion mining from product reviews using conditional random fields, in: In AusDM : The 13th Australasian Data Mining Conference, University of Technology, Sydney, Australia, 2015.
[33] X. Fu, Y. Wei, F. Xu, T. Wang, Y. Lu, J. Li, J.Z. Huang, Semi-supervised aspect-level sentiment classification model based on variational autoencoder, Knowl.-Based Syst. 171 (2019) 81–92.
[34] N.F.F. Da Silva, L.F.S. Coletta, E.R. Hruschka, E.R. Hruschka Jr., Using unsupervised information to improve semi-supervised tweet sentiment classification, Inform. Sci. 355356 (2016) 348–365.
[35] N. Mukhtar, M.A. Khan, N. Chiragh, Lexicon-based approach outperforms Supervised Machine Learning approach for Urdu Sentiment Analysis in multiple domains, Telemat. Inform. 35 (2018) 2173–2183.
[36] A. García-Pablos, M. Cuadros, G. Rigau, Unsupervised word polarity tagging by exploiting continuous word representations, Proces. Leng. Nat. 55 (2015) 127–134.
[37] K. Ye, L. Li, M. Guo, Y. Qian, H.Y. B, Summarizing product aspects from massive online review with word representation, Knowl. Sci. Eng. Manage. 9403 (71572029) (2015) 318–323.
[38] S. Mukherjee, S. Joshi, Sentiment aggregation using conceptnet ontology, in: Proceedings of the Sixth International Joint Conference on Natural Language Processing, (October), 570578., 2013.
[39] W. Fangzhao, H. Yongfeng, Y. Jun, Active sentiment domain adaptation, in: Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, Department of Electronic Engineering, Tsinghua University.
[40] C. Hung, Word of mouth quality classification based on contextual sentiment lexicons, Inf. Process. Manage. 53 (4) (2017) 751–763.
[41] S. Deng, A. P.Sinha, H. Zhao, Adapting sentiment lexicons to domain-specific social media texts, Decis. Support Syst. 94 (2017) 65–76.
[42] H. Keshavarz, M. Saniee Abadeh, ALGA: Adaptive lexicon learning using genetic algorithm for sentiment analysis of microblogs, Knowl.-Based Syst. 122 (2017) 1–16.
[43] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, J.W. Vaughan, A theory of learning from different domains, Mach. Learn. 79 (2010) 151–175.
[44] F. Liu, G. Zhang, J. Lu, Heterogeneous domain adaptation: An unsupervised approach, IEEE Trans. Neural Netw. Learn. Syst. (2020) 1–15.
[45] H. Wu, Y. Yan, Y. Ye, M.K. Ng, Q. Wu, Geometric Knowledge Embedding for unsupervised domain adaptation, Knowl.-Based Syst. 191 (5) (2020) 105155.
[46] Y. Yan, Q. Wu, M. Tan, M.K. Ng, H. Min, I.W. Tsang, Online heterogeneous transfer by hedge ensemble of offline and online decisions, IEEE Trans. Neural Netw. Learn. Syst. 29 (7) (2017) 3252–3263.
[47] J. S.Deshmukh, A. Kumar Tripathy, Entropy based classifier for cross-domain opinion mining, Appl. Comput. Inform. 14 (2018) 55–64.
[48] N. Pröllochs, S. Feuerriegel, D. Neumann, Negation scope detection in sentiment analysis, Decis. Support Syst. 88 (2016) 67–75.
[49] A.S. Manek, P.D. Shenoy, M.C. Mohan, K.R. Venugopal, Aspect term extraction for sentiment analysis in large movie reviews using Gini Index feature selection method and SVM classifier.
[50] C. Chen, R. Zhuo, R. Jiangtao, Gated recurrent neural network with sentimental relations for sentiment classification, Inform. Sci. 502 (2019) 268–278.
[51] R. Socher, J. Pennington, E. Huang, A. Ng, C. Manning, Semi-supervised recursive autoencoders for predicting sentiment distributions, in: Proceedings of EMNLP, 2011.
[52] R. Socher, A.W.J. Perelygin, J. Chuang, C. Manning, A. Ng, C. Potts, Recursive deep models for semantic compositionality over a sentiment treebank, in: Proceedings of EMNLP, 2013.
[53] S.J. Pan, X. Ni, J.-T. Sun, Q. Yang, Z. Chen, Cross-domain sentiment classification via spectral feature alignment, in: Proceedings of the 19th International Conference on World Wide Web, ACM, 2010.
[54] S. Baccianella, A. Esuli, F. Sebastiani, 3.0: A publicly available lexical resource for opinion mining An enhanced lexical resource for sentiment analysis and opinion mining, in: 7th Language Resources and Evaluation Conference, 2010.
[55] J. Blitzer, M. Dredze, F. Pereira, [53] Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification, 2007, pp. 440–447.
[56] B. Pang, L. Lee, Seeing stars: Exploiting class relationships for sentiment categorization with respect to rating scales, 2005, pp. 115–124.
[57] J. Martineau, T. Finin, Delta TFIDF: An improved feature space for sentiment analysis, in: ICWSM, 2009, pp. 258–261.
[58] W.F. zhao, H. Yongfeng, Y. Zhigang, Domain-specific sentiment classification via fusing sentiment knowledge from multiple sources, Inf. Fusion 35 (2017) 26–37.
[59] Y. Choi, C. Cardie, Adapting a polarity lexicon using integer linear programming for domain-specific sentiment classification, in: In:Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, Association for Computational Linguistics, 2009.
[60] S.M. Mohammad, S. Kiritchenko, X. Zhu, NRC-Canada: Building the state-of-the-art in sentiment analysis of tweets, in: The Seventh International Workshop on Semantic Evaluation, 2013.
[61] Z. Huang, W. Xu, K. Yu, Bidirectional LSTM-CRF Models for Sequence Tagging, Comput. Lang. (2015).
[62] J. Devlin, M.-W. Chang, K. Lee, K. Toutanova, BERT: Pre-training of deep bidirectional Transformers for Language understanding, Comput. Lang. (2019).