Improving SFA Algorithm by Employing Multi-Source Data

Saeed Dehghani Ashkezari
Computer Dept.
Yazd University
Yazd, Iran
sdehghani@stu.yazd.ac.ir

Vali Derhami
Computer Dept.
Yazd University
Yazd, Iran
vderhami@yazd.ac.ir

Samira MashtiZadeh
Computer Dept.
Science and Art University
Yazd, Iran
sm.ms20270@gmail.com

Received:2018/11/15 Revised:2019/03/04 Accepted:2019/04/27

***Abstract*— In recent years, discriminative learning methods have widely been used in various areas of Natural Language Processing (NLP). These methods achieve the best performance, when the set of training and testing samples have the same distribution. However, in many applications of NLP, the lack of labeled datasets for some domains is a serious challenge. In such conditions, we need to develop a model based on domains with rich labeled instances and apply it to the domain with no labeled instances. In this research, a method for sentiment classification of opinions into positive and negative groups, which represent the users' feelings, is offered based on multi-source transfer learning. The proposed method here employs Spectral Feature Alignment algorithm to adapt different domains. Furthermore, according to the Majority Voting, accuracy is assigned to classifications trained on different domains based on the Majority Voting Error. Ultimately, decisions are made for each classification based on the calculated error. The Amazon datasets for four different categories, each of which contains 1000 positive and 1000 negative samples, are exploited to train the proposed model. Meanwhile, each category includes unlabeled samples that are used to select pivot features. The accuracy values of 85.5%, 86.4%, 83.5% and 90.1% obtained for Electronics, DVD, Books and Kitchen domains respectively, show the effectiveness of the proposed method compared with similar methods.**

Keywords*— Sentiment Classification of Opinions Transfer Learning, Spectral Feature Alignment, Majority Voting, Majority Voting Error.

1. INTRODUCTION

Today online shopping websites play an important role in e-commerce. Most of these sites persuade customers to express their opinions about the purchased goods and services, by providing an appropriate environment and interface. The opinions of customers who have purchased a product or a service have become one of the main references to guide other customers about the quality of that product or service [1]. Park et al. [2] confirmed that opinions have a direct impact on products' sales, and customers decide to buy products that have gotten positive reviews more probably.

Different opinions about a product will help potential buyers take advantage of the experiences of other people who have already purchased it. When there are, however, so many opinions, it's impossible to review all of them, making customers confused and making their final decisions harder. Hence, in the last decade, a new research field called "Opinion Mining" has emerged and various studies have been conducted in this scope.

Opinion Mining is one of the relatively new fields which aims to study methods of analyzing individuals' opinions, expressed in their unstructured writings on the Web by the help of Artificial Intelligence, Information Retrieval and Natural Language Processing. These opinions range from business issues (opinions about a product or a service [3]) to political opinions (opinions of individuals about the candidates [4]) and social networks (such as twitter) [5]. Despite knowing a particular person's opinion about a product or a service alone may not help so much, knowing the opinions of all individuals and groups may result in a wise crowd-based conclusion.

The main purpose of this research is to provide a new method for the sentiment classification of opinions. To classify opinions about a product such as a camera into positive and negative categories, we need to collect and label a large number of opinions. As the process of labeling may be time-consuming and costly, transfer learning is used to reduce the efforts needed to label opinions about a new product or service.

A large number of machine learning methods work only with a common assumption: the train and test data act at the same feature space and distribution. When the distribution of data is changed, most statistical models need to be rebuilt with the help of new training data [6]. Thus, inter-domain sentiment classification is applied in order to cover the difference of these domains. Although this method can identify a more stable representation of inter-domain data compared with previous approaches, it suffers from weaknesses such as using only one source domain for training. Hence, taking advantage of several sources may lead to better results, an approach based on Multi-source Spectral Feature Alignment method is offered here.

Then this Multi-source method that exploits several sources, is used for classifying the opinions of a different domain. There are a large number of unlabeled opinions in the source and target domains, so the similarity between two domains can be efficiently mapped. This idea is based on the assumption that distribution of similar words is identical and the repeated words behave the same way[7]. Spectral clustering has been used to adapt domain-independent and domain-specific features. Resulting clusters are used to reduce the mismatch between domain-specific features. These clusters are eventually used as a feature space for representing data samples. As stated previously, another major goal of this research is concurrent exploitation of different domains for training.

Briefly, in this paper a new and efficient method for classifying opinions into positive and negative labels, based on Multi-source learning is introduced. The proposed model is based on Spectral Feature Alignment (SFA) approach, employing several source domains for training and exploiting the resulting model to classify opinions of a different domain.

The remaining article is organized as follows: Section 2 reviews works already done. Section 3 outlines the proposed method for the sentiment classification of opinions. In Section 4, results of experiments are described, and section 5 is dedicated to summing up and highlighting future works.

2. LITERATURE REVIEW

In many classification methods, it is assumed that the train and test data have the same distribution. This assumption can be analyzed from two aspects: (a) documents belonging to both source and target domains can be displayed based on the same feature space (identical words) and (b) words of both domains have the same distributions. According to the first aspect, words are repeated in two domains. The second aspect considers the likelihood of words' occurrence in the two domains. A lot of research has been done to solve this problem, all of which use labeled data of other domains. The basic idea behind these methods is to map features of the target domain with features of the source domain based on domain- independent words.

To determine sentimental polarity of opinions, we attempt to identify positive or negative label of each review. This is typically done based on other labeled opinions that already exist. But sometimes such data is not available for some domains. On the other hand, other domains may include information that can be used to construct models in other domains. For example, opinions contained in the domain of "electronic devices" may not have labels, while the "book" and "kitchen" domains have opinions whose polarity label is available. In this situation, domain adaptation techniques would be used.

In domain adaptation, we attempt to classify opinions by transferring knowledge from a labeled domain (source) to another unlabeled domain (target). So far, various studies have been done in the field of domain adaptation [8-10].

Here basic idea is to transform the data representations of source domains into target domains so that they present the same joint distribution of observations and labels. Blitzer et al. [10] proposed the structural correspondence learning (SCL) algorithm to exploit domain adaptation techniques for sentiment classification. The main idea of SCL is to achieve feature alignment in different domains by choosing a set of pivot features and modeling the oorrll iii ons bwween 'pvvot faauur'' nnd ohher faauurss (ceeed 'non-pvvot feuuur'''). Pnn tt ll. proposed a domnnn adaptation approach called spectral feature alignment (SFA) [11] that tries to find an alignment between domain-specific and domain-independent features by performing spectral clustering based on a bipartite graph. This graph is constructed based on co-occurring relationship between domain-specific and domain-independent features.

In some studies, attempts have been made to fill the gap between source and target domains by means of foreign knowledge sources. For example, Wang et al. [12] proposed a common clustering approach to transfer labels between two domains with the help of Wikipedia in order to represent documents based on concepts. Xiang et al. [13] presented another similar approach for exploiting Wikipedia and other resources for domain adaptation. In sentiment classification, some methods use sentimental dictionaries as external sources. The JSTM model [14] also uses a sentimental dictionary as a reservoir to prioritize the sentimental meaning of words. In [15] SentiWordNet is used for domain adaptation. The main drawback of such methods is the need for the availability of external resources, and that their efficiency depends on the quality of these resources.

Bollegala et al. [16] presented a method for sentiment classification of opinions. In their method, a polarity-based thesaurus is created for several labeled source domains and a target domain, and then the ultimate feature vector is expanded to learn a binary classification.

Franco Salvador et al. [17] proposed a meta-learning-based inter-domain approach in the effort of building a model for opinion classification. Using BabelNet's sentimental network, they exploit multilingual features to resolve the ambiguity and to expand the dictionary used in their method.

Fang et al. [18] presented a hybrid method that integrates sentimental information of a source domain with information from a set of pre-selected sentimental words.

In [19], a novel approach called words alignment based on association rules (WAAR) for cross-domain sentiment classification is proposed which can establish an indirect mapping relationship between domain-specific words in different domains by learning the strong association rules between domain-shared words and domain-specific words in the same domain. In this way, the differences between the source domain and target domain can be reduced to some extent, and a more accurate cross-domain classifier can be trained.

Semi-supervised learning is an effective method for coping with the insufficiency of labeled data in machine learning. In [20], a cooperative semi-supervised learning method based on the hybrid mechanism of active learning and self-learning is proposed for text sentiment classification.

Liu et al. [21] proposed a semi-supervised sentiment classification based on auxiliary task learning, namely Aux-

LSTM, which is used to assist learning the sentiment classification task with a small amount of human-annotated samples by training auto-annotated samples. Their method first annotates the unlabeled samples automatically with IG algorithm to obtain the auto-annotated samples. Then, it assists in sentiment classification of the human-annotated samples (main task) through the sentiment classification of the auto-annotated samples (auxiliary task). Finally, joint learning the loss function of the two task to improve the performance of the main task.

Xia et al. [22] proposed a dual-view co-training algorithm based on dual-view bag-of-words representation for semi-supervised sentiment classification. In dual-view BOW, antonymous reviews are constructed automatically and a review text is modeled by a pair of BOWs with opposite views. Then they make use of the original and antonymous views in pairs, in the training, bootstrapping and testing process, all based on a joint observation of two views.

In [23], Peng et al. propose a method to simultaneously extract domain specific and invariant representations and train a classifier on each of the representation, respectively. A few target domain labeled data is also introduced for learning domain-specific information. To effectively utilize the target domain labeled data, they train the domain-invariant representation based classifier with both the source and target domain labeled data and train the domain-specific representation based classifier with only the target domain labeled data. These two classifiers then boost each other in a co-training style.

3. SFA ALGORITHM

Based on SFA algorithm, words from source domains and target domains are aligned to bridge the gaps between them. Here the way this algorithm works is explained by the original example.

Let's consider some of sample opinions expressed in Table 1 from electronics and video games domains [11]. In this table, words such as compact, sharp and hooked are domain-specific that have been used in just a specific domain. Words such as good and excited are domain-independent as they behave equally in source and target domains. By representing each opinion by means of bag of words, Table 2 is obtained. It's evident that domain-specific features prevent to build an efficient training model based on electronics domain and apply it in video games domain. By the way, there are some methods of bridging between source and target domains.

One of the most popular methods of domain adaptation is Spectral Feature Alignment (SFA) method. Based on this algorithm, domain-specific words are clustered by their behavior with domain-independent words. It means that we consider their occurrences with domain-independent words (consider Table 3).

There are three ways to select domain-independent features:

1. Feature selection based on their frequencies in both domains: More specifically, given the number l of domain-independent features to be selected, we choose features that occur more than k times in both of the source and target domains. k is set to be the largest number such that we can get at least l such features.
2. Exploiting mutual dependence between features and labels in the source domain data: In information theory, mutual information is used to measure the mutual dependence between two random variables. Feature selection using mutual information can help identify features relevant to source domain labels. But there is no guarantee that the selected features act similarly in both domains.
3. Selecting features by the modified mutual information criterion (1):

TABLE 2. BAG-OF-WORDS REPRESENTATIONS OF ELECTRONICS (E) AND VIDEO GAMES (V) REVIEWS

		…	compact	realistic	sharp	hooked	blurry	boring
E	+	…	1	1	0	0	0	0
	+	…	0	1	0	0	0	0
	-	…	0	0	1	0	0	0
V	+	…	0	0	0	1	0	0
	+	…	0	0	0	1	1	0
	-	…	0	0	0	0	0	1

TABLE 3. A CO-OCCURRENCE MATRIX OF DOMAIN-SPECIFIC AND DOMAIN-INDEPENDENT WORDS

	compact	realistic	sharp	hooked	blurry	boring
good	1	1	1	1	0	0
exciting	0	0	1	1	0	0
never_buy	0	0	0	0	1	1

TABLE 1. CROSS-DOMAIN SENTIMENT CLASSIFICATION EXAMPLES: REVIEWS OF ELECTRONICS AND VIDEO GAMES PRODUCTS

	electronics	*video games*
+	**Compact**; easy to operate; very *good* picture quality; looks sharp!	A very *good* game! It is action packed and full of *excitement*. I am very much **hooked** on this game.
+	I purchased this unit from Circuit City and I was very *excited* about the quality of the picture. It is really *nice* and **sharp**.	Very **realistic** shooting action and *good* plots. We played this and were **hooked**.
-	It is also quite **blurry** in very dark settings. I will *never buy* HP again.	The game is so **boring**. I am extremely unhappy and will probably *never buy* UbiSoft again.

$$I(X^i; D) = \sum_{d \in D} \sum_{x \in X^i, x \neq 0} p(x, d) \log \left(\frac{p(x, d)}{p(x)\, p(d)} \right) \quad (1)$$

where D is a domain variable and we only sum over non-zero values of a specific feature X^i. The smaller $I(X^i;D)$ is, the more likely that X^i can be treated as a domain-independent feature [11].

For our example, a bipartite (domain-independent and domain-specific features) graph is created as Figure 1.

Obviously, clustering methods such as k-means will lead to the Table 4 with which we can train a convenient model. It's easy to see that we have a good situation for training our model now. In fact, this algorithm adapts domain-specific words in identical clusters and based on domain-independent words, the relationship between different domains is obtained. In other words, a map between source and target domains is established.

In graph spectral theory, it's assumed that if two nodes (e.g. domain-specific words) in a graph were connected to many common nodes (e.g. domain-independent words), then they would be very similar (or quite related). Therefor SFA algorithm exploits spectral clustering that has some advantages like:

1. It makes no assumption on the form of the data clusters by transforming the data clustering to graph partitioning problem.
2. It provides good clustering results, specially it is invariant to cluster shapes and densities.
3. It's reasonably faster for sparse data sets of several thousand elements.

Table 4. Ideal representations of domain-specific words

		…	Compact_realistic	Sharp_hooked	Blurry_boring
E	+	…	1	0	0
	+	…	1	0	0
	-	…	0	1	0
V	+	…	0	1	0
	+	…	0	1	1
	-	…	0	0	1

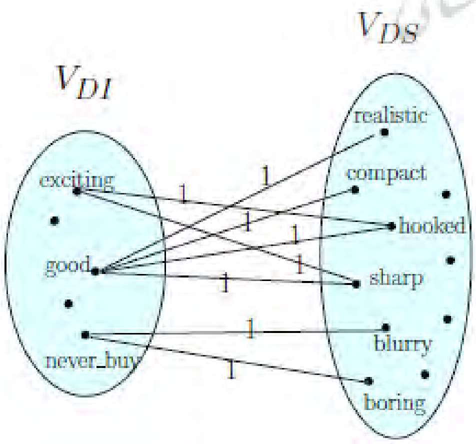


Fig.1. A bipartite graph example of domain-specific and domain-independent features

In other words, unlike K-Means and classic algorithms, spectral clustering can group objects belonging to irregular form groups based on connectivity, see Figure 2 [24]:

Given a set of points $V = \{v_1, v_2, ..., v_n\}$ and their corresponding weighted graph G, the goal is to cluster the points into k clusters, where k is an input parameter.

1. Form an affinity matrix for V : $A \in R\ n\times n$, where $A_{ij} = m_{ij}$, if $i \neq j$; $A_{ii} = 0$.
2. Form a diagonal matrix D, where $D_{ii} = H_j A_{ij}$, and construct the matrix $L = D^{-222}AD^{-222}$.
3. Find the k largest eigenvectors of L, $u_1, u_2, ..., u_k$, and form the matrix $U = [u_1\ u_2 ... u_k] \in R\ n\times k$.
4. Normalize U, such that $U_{ij} = U_{ij} / (\sum j\ U^2{}_{ij})^{1/2}$.
5. Apply the k-means algorithm on U to cluster the n points into k clusters.

Based on the above spectral clustering steps, given labeled source domain data Dsrc and unlabeled target domain data Dtar, the number of clusters K and the number of domain-independent features m, the SFA algorithm may be described as below:

1: Select l domain-independent and m-l domain-specific features.

2: Calculate (DI-word)-(DS-word) co-occurrence matrix $M \in R^{mm\times\times k}$.

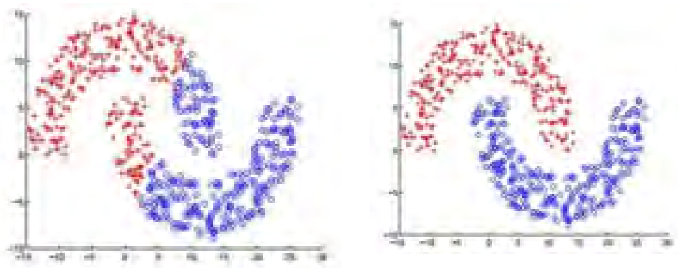
3: Construct matrix $L = D^{-122}AD^{-222}$, where:

$$A = \begin{bmatrix} 0 & M \\ M^T & 0 \end{bmatrix}$$

4: Find the K largest eigenvectors of L, $u_1, u_2, ..., u_K$, and form the matrix $U = [u_1\ u_2 ... u_K] \in R^{m\times K}$.

5: Train classifiers on the source using augmented features (original features + new features).

Based on the above description, the standard spectral clustering algorithm clusters n points to k discrete nndoooos, whcdh aan be rffrrrdd oo ss “dssrreee uuusrrring.. ii ng and ee [25] proved that the k principal components of a term-document co-occurrence matrix, which are referred to as the k largest eigenvectors u1, u2, ..., uk in step 4, are actually the continuous solution of the cluster membership indicators of documents in the k-means clustering method.



k-means clustring spectral clustring

Fig.2. Comparison between K-means and spectral clustering

4. THE PROPOSED METHOD

In this research, a method based on Spectral Feature Alignment algorithm, trained with several sources is used to classify opinions of another different domain. There are a large number of unlabeled items in the source and target domains, so the similarity map of these words from both domains can be used. This idea is based on the assumption that the distribution of synonym words and the words that are repeated with them are the same. Spectral clustering is used to adapt domain-independent and domain-dependent features. Resulting clusters are then used to reduce the mismatch between domain-specific features. These clusters are eventually used as a feature space to represent data samples.

But as said previously, there are sometimes more than one labeled source to train a transferring model, and if a model were built by the aid of more than one source, it would lead to better results. Hence, the method presented here is an attempt to use more than one source domain to train the classifier model.

By considering several sources, there would be several classifiers to be used simultaneously in order to achieve the most effective results. The easiest way is to compare the results of each resource and select the best classifier, which will not certainly produce the best performance [26]. To achieve maximum efficiency, a solution may be achieved by combining the results. Hence, in this paper, the Majority Voting method is used to combine the results of all classifiers, as it is one of the most common ensemble methods in classification tasks [27] as well as it's a reasonable choice for a balanced data set [28]. As shown in Figure. 3, there are more than one classifier (depending on the source domains) and the majority of results are combined based on voting. More exactly, based on the voting criteria, the majority of positive or negative labels of opinions are determined.

More specifically, this combination works as follows:

(1) take the result of each method applied to a single message;

(2) check the most frequent polarity given by all methods;

(3) assign the most frequent polarity as the final polarity of this message [29].

In order to choose the best combination of classifiers, we use a criterion that measures the error rate of the combined classifiers. The criterion used here is the Majority Voting Error (MVE). If the majority voting for each sample were defined as (2):

$$y_i^{MV} = \begin{cases} 1, & if \sum_{j=1}^{M} y_{ij} \geq \frac{M}{2} \\ 0, & if \sum_{j=1}^{M} y_{ij} < \frac{M}{2} \end{cases} \tag{2}$$

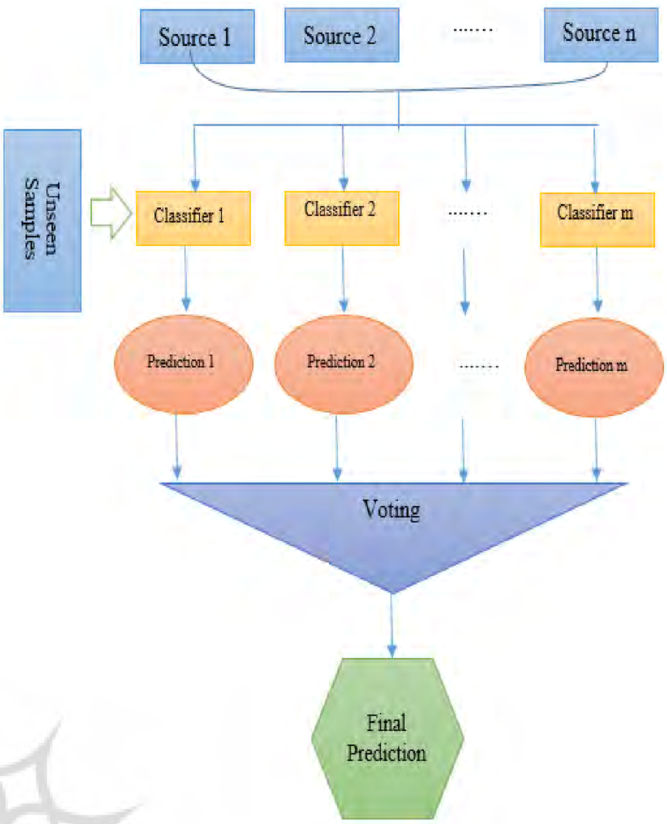


Fig. 3. Combining the results of more than one classifier for final prediction

Then the majority voting error might be calculated according to (3):

$$MVE = \frac{1}{N}\sum_{i=1}^{N} y_i^{MV} \tag{3}$$

Furthermore let m(xi) denote the number of classifiers producing error for the input sample x_i. It can be expressed by (4):

$$m(x_i) = \sum_{i=1}^{M} y_{ij} \tag{4}$$

where y_{ij} is the binary output from the j^{th} classifier for the i^{th} input sample. Finally let $e_j = \frac{1}{N}\sum_{i=1}^{N} y_{ij}$ denote the error rate of j^{th} classifier and accordingly the ensemble mean error rate be defined by (5) [30]:

$$\bar{e} = \frac{1}{M}\sum_{j=1}^{M} e_j \tag{5}$$

Due to the fact that any source domain may have a different consistency and correspondence with the target domain, accuracy is used as the coefficient factor for the target adaptation weight with the source domains.

We also for better combining the advantages of classifier, use a dynamic selection voting method. In this method the algorithms that are initially used for building the ensemble are tested in a small subset of the training set and if they have statistically worse accuracy than the most accurate algorithm, they do not participate to the final decision of the ensemble. The presented methodology for combining classifiers is a six-step strategy:

1. The dataset is sampled at random about 20% of the initial set
2. The new dataset is divided at random into three equal parts
3. Two of three parts are used for training of algorithms and the remaining data is the testing set
4. The results of three tests are averaged
5. The algorithms that have statistically worse accuracy than the most accurate are not used by the ensemble
6. The remaining algorithms are then executed on the full training set in order to produce the prediction model with simple voting.

In other words, the classification process includes two phases (Figure 4): (1) learning phase, and (2) application phase. During the learning phase, a set of base classifiers is generated and each base classifier in the ensemble (classifiers h_1 ... h_n) is trained. For the ensemble classification, the corresponding classifications of the base classifiers are combined with selective voting $h^* = F(h_1, h_2, ..., h_n)$ to produce the final classification of the ensemble. At the application phase, a new instance (x,?) is given with the unknown value y to be classified by the ensemble. As a result, the class value y^* is then predicted as $y^* = h^*(x)$.

5. RESULTS

The Multi-Domain Sentiment Dataset (MDS) was first introduced by Blitzer et al. [10]. This dataset was provided based on opinions on the Amazon website for four domains: Book, DVD, Electronics and Kitchen. In order to evaluate the proposed method and the possibility of

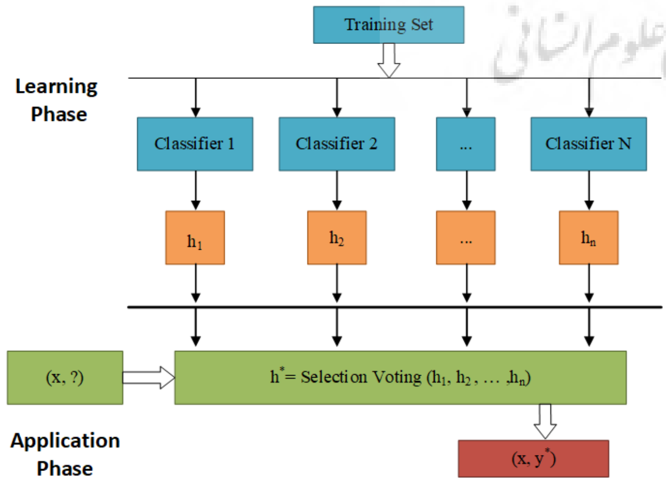


Fig. 4. Dynamic combining the results of more than one classifier for final prediction

comparing results with similar methods, the above dataset that has been used in many studies was used.

The Amazon website gives points from 1 to 5 in each opinion. In this dataset, opinions with scores larger than 3 are labeled as positive, and smaller scores receive a negative label. The rest of opinions, as they may not be polarized, are obscure and ignored. Finally, for each domain, 1000 positive and 1,000 negative samples are considered. In addition to labeled data for each domain, there are also unlabeled samples that range from 3586 to 5945 opinions for different domains.

Table 5 below shows some of the most frequent features of each domain. As mentioned earlier, domain-independent features are those features that belong and behave equally to both of the source and target domains. For example, 'excellent' is a domain-independent feature belonging to the domains 'Book' and 'Electronics'. We used the frequency criterion (FQ) for feature selection to show the efficiency of our proposed model more obviously. Results of the proposed method are so remarkable and attractive: better results by means of weaker criterion of feature selection (Table 9)!

Table 6 shows the number of labeled/unlabeled documents, the number of unique terms, and the total number of terms for each domain. As with similar methods presented previously, we randomly divided each dataset into two parts; 1600 samples for training and 400 samples for testing.

Based on most studies about sentiment classification of opinions, we use Accuracy criterion to evaluate our proposed

TABLE 5. Some of the most frequent features of 4 domains

Domain	Polarity	Features
Independent	+	excellent, great, best, perfect, love, wonderful, loved, enjoy
	-	bad, waste, boring, disappointed, worst, poor, disappointing, disappointment, terrible, poorly, off, broken
Book	+	excellent, easy, loved, enjoyable, fun, favorite, must_read, important, novel
	-	boring, disappointing, bad, instead, waste, little, poorly, unfortunately
DVD	+	enjoy, hope, loved, better than, best, first, classic, back
	-	worst, boring, bad, the_worst, terrible, waste, awful, horrible, dull, lame, hard
Kitchen	+	easy, great, perfect, love, easy_to, best, little, well, good, nice, long, durable, clean
	-	disappointed, back, poor, broken, return, off, returned, broke, waste, tried
Electronics	+	excellent, great, perfect, best, love, easy_to, easy, little, the_best, works, good, nice, wonderful
	-	disappointed, poor, waste, bad, worst, back, broken, return, horrible, off, tried, poorly

TABLE 6. Number of labeled/unlabeled samples

Domain	Labeled	Unlabeled	Terms	Occurrence
Book	2000	4465	195887	445793
DVD	2000	3586	188778	370844
Electronics	2000	5681	111407	392699
Kitchen	2000	5945	93474	351162

method and compare the results with other works. The Accuracy criterion, as seen in the relationship (6), calculates the correct classified documents divided by all outputs of the model.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

In this relation, TP, TN, FP, and FN, represent the actual number of positives, actual negatives, false positives and false negatives, respectively.

It should be noted that the reason for choosing this criterion is that there is a balanced number of positive and negative samples in each dataset.

Tables 7 and 8 display the best results of basic SFA algorithm for each target domain by using FQ as the feature selection method and all of FQ, MI and DI methods, respectively. The best source domain for each target domain has also been determined.

In order to evaluate efficiency of the proposed method, the results are compared with the methods presented earlier. Table IX shows results of the proposed method (static and dynamic) for different target domains.

Table 10 contains results of different algorithms described earlier, as well as the proposed (static and dynamic) methods.

In the Figures 5 to 8, the results of the proposed multisource SFA algorithms are shown in comparison with the single-source SFA algorithm and other methods above.

TABLE 7. THE BEST RESULTS OF THE BASIC SFA ALGORITHM (FQ)

No	Target	Source	Accuracy
1	Electronics	Kitchen	84.9
2	DVD	Book	81.25
3	Book	DVD	78.25
4	Kitchen	Electronics	85.8

TABLE 8. THE BEST RESULTS OF THE BASIC SFA ALGORITHM (FQ, MI, DI))

No	Target	Source	Selection Method	Accuracy
1	Electronics	Kitchen	DI	85.05
2	DVD	Book	DI	81.35
3	Book	DVD	MI	79.8
4	Kitchen	Electronics	DI, MI	86.75

TABLE 9. RESULTS OF THE PROPOSED METHOD (THREE SOURCES)

No	Source	Target	Accuracy (Static)	Accuracy (Dynamic
1	{DVD, Book, Kitchen}	Electronics	85.13	85.5
2	{Book, Kitchen, Electronics}	DVD	86.36	86.4
3	{DVD, Kitchen, Electronics}	Book	82.68	83.5
4	{Book, DVD, Electronics}	Kitchen	89.72	90.1

TABLE 10. SUMMARY OF THE VARIOUS ALGORITHMS' RESULTS

	Electronics	DVD	Books	Kitchen
SFA (FQ)	84.9	81.25	78.25	85.8
SFA (Best)	85.05	81.35	79.8	86.75
Proposed (Static)	85.13	86.36	82.68	89.72
Proposed (Dynamic)	85.5	86.4	83.5	90.1
SCL-MI	86.8	77.2	79.7	85.9
[18]	84.18	79.13	78.29	86.29
[19]	85.05	81.25	79.6	85.03
[20]	77.8	77.1	74.4	80
[21]	74.8	73	73.8	77.8
[22]	76.9	73.8	72.1	78
[23]	87.2	83.1	81.8	87.3

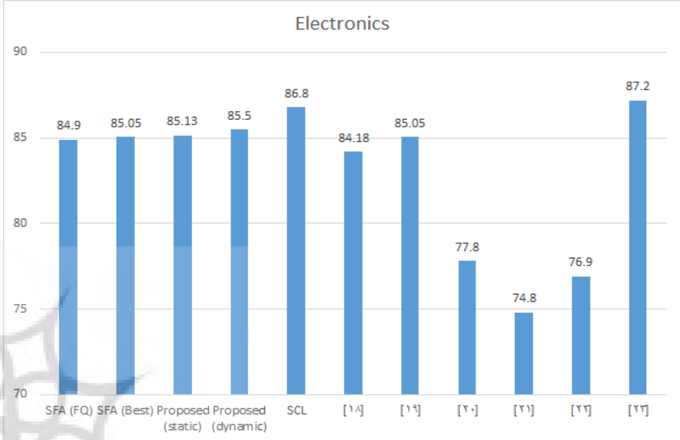


Fig. 5: Electronics domain

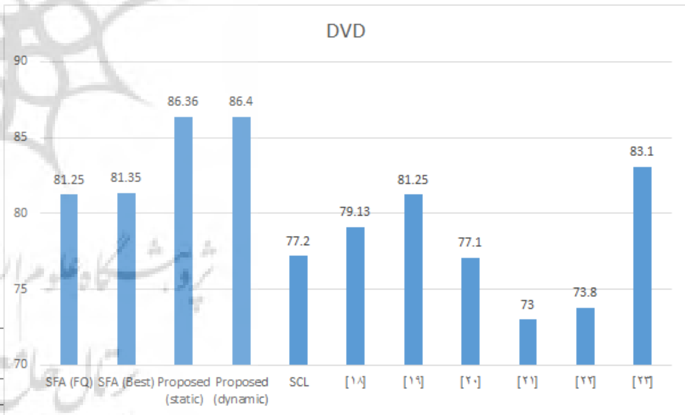


Fig. 6: DVD domain

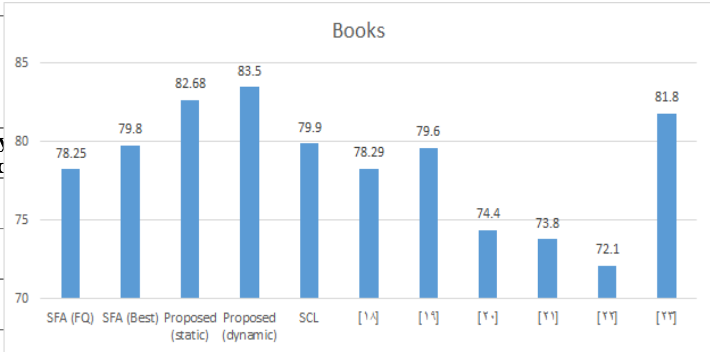


Fig. 7: Books domain

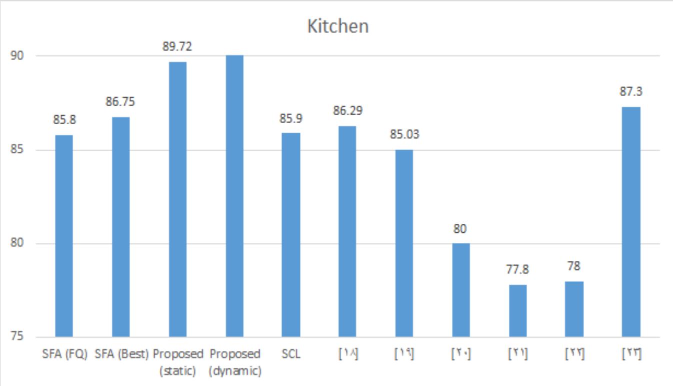


Fig. 8: Kitchen domain

6. CONCLUSIONS AND FUTURE WORKS

The Spectral Feature Alignment (SFA) method is an efficient method for multi-source sentiment classification of texts. This approach uses unlabeled documents to find the corresponding pair of words between different domains. In this paper, multiple domains are used simultaneously as the source, and after training the hybrid classification model, it is used for the sentiment classification of the target test domain. As seen above, the proposed method offers better performance in comparison with the base method of single-source SFA. For future studies, additional domains from a dataset other than Amazon can be used to improve the performance of the method. Moreover, the previous multisource methods that have had a good result can be employed in order to improve results.

REFERENCES

[1] Hu, N., Liu, L., & Zhang, J. J. (2013), "Do online reviews affect product sales? The role of reviewer characteristics and temporal effects", *Information Technology and Management*, 9(3), 201-214.

[2] Park, D. H., Lee, J., & Han, I. (2007), "The effect of online consumer reviews on consumer purchasing intention: The moderating role of involvement", *International Journal of Electronic Commerce*, 11(4), 125-148.

[3] Umar, S., Maryam, Azhar, F., Malik, S., and Samdani, G. "Sentiment Analysis approaches and applications: A survey", *International Journal of Computer Applications*, 181(1), 1-9.

[4] Bharathi S., & Geetha, A. (2017), "Sentiment analysis for effective stock market prediction", *International Journal of Intelligent Engineering & Systems*, 10(3), 146-154.

[5] Dattu, B. S., & Gore, D. V. (2015), "A survey on sentiment analysis on twitter data using different techniques", *International Journal of Computer Science and Information Technologies (IJCSIT)*, 6(6), 5358-5362.

[6] Chen, J., & Liu, X. (2016), "Transfer learning with one-class data", *Pattern Recognition Letters*, 37, 32-40.

[7] Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016), "A survey of transfer learning", *Journal of Big Data*, 3(1), 1-40.

[8] Banea, C., Mihalcea, R., Wiebe, J., & Hassan, S. (2008, October), "Multilingual subjectivity analysis using machine translation", In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, (pp. 127-135), Association for Computational Linguistics.

[9] Prettenhofer, P., & Stein, B. (2011), "Cross-lingual adaptation using structural correspondence learning", *ACM Transactions on. Intelligent Systems and Technology*, 3(1), 13:1-13:22.

[10] Blitzer, J. McDonald, R. & Pereira, F. (2006), "Domain adaptation with structural correspondence learning", In *Proceedings of the 2006 conference on empirical methods in natural language processing*, (pp. 120-128), Association for Computational Linguistics.

[11] Pan, S. J., Ni, X., Sun, J. T., Yang, Q., & Chen, Z. (2010, April), "Cross-domain sentiment classification via spectral feature alignment", In *Proceedings of the 19th international conference on World wide web*, (pp. 751-760), ACM.

[12] Wang, P., Domeniconi, C., & Hu, J. (2008, December), "Using wikipedia for co-clustering based cross-domain text classification", In *2008 Eighth IEEE international conference on Data Mining*, (pp. 1085-1090). IEEE.[13] Xiang, E. W., Cao, B., Hu, D. H., & Yang, Q. (2010). "Bridging domains using world-wide knowledge for transfer learning", *IEEE Trans. Knowl. Data Eng.*, 22(6), 770-783.

[14] He, Y., Lin, C., & Alani, H. (2011), "Automatically extracting polarity-bearing topics for cross-domain sentiment classification", In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, (pp. 123-131), Association for Computational Linguistics.

[15] Denecke, K. (2009, November). "Are SentiWordNet scores suited for multi-domain sentiment classification?", In *2009 Fourth International Conference on Digital Information Management*, (pp. 1-6). IEEE.

[16] Bollegala, D., Weir, D., & Carroll, J. (2011). "Using multiple sources to construct a sentiment sensitive thesaurus for cross-domain sentiment classification", In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies-Volume 1*, (pp. 132-141), Association for Computational Linguistics.

[17] Franco-Salvador, M., Cruz, F. L., Troyano, J. A., & Rosso, P. (2015). "Cross-domain polarity classification using a knowledge-enhanced meta-classifier", *Knowledge-Based Systems*., 86, 46–56.

[18] Fang, F., Dutta, K., & Datta, A. (2014). "Domain adaptation for sentiment classification in light of multiple sources", *INFORMS Journal of Computing*, 26(3), 586–598.

[19] Jia, X., Jin, Y., Li, N., Su, X., Cardiff, B., & Bhanu, B. (2018). "Words alignment based on association rules for cross-domain sentiment classification", *Frontiers of Information Technology & Electronic Engineering*, 19(2), 260-272.

[20] Li, Y., Lv, Y., Wang, S., Liang, J., Li, J., & Li, X. (2019). "Cooperative Hybrid Semi-Supervised Learning for Text Sentiment Classification", *Symmetry*, 11(2):133, 260-272.

[21] Liu, H., Wang, J., Li, S., Li, J., & Zhou, G. (2018). "Semi-supervised Sentiment Classification Based on Auxiliary Task Learning", In *CCF International Conference on Natural Language Processing and Chinese Computing (NLPCC)*, (pp. 372-382). Springer, Cham.

[22] Xia, R., Wang, C., Dai, X., & Li, T. (2015). "Co-training for semi-supervised sentiment classification based on dual-view bags-of-words representation", In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing*, (pp. 1054-1063), Association for Computational Linguistics.

[23] Peng, M., Zhang, Q., Jiang Y., & Huang, X. (2018). "Cross-Domain Sentiment Classification with Target Domain Specific Information", In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, (pp. 2505-2513), Association for Computational Linguistics.

[24] Ben Ayed, A., Ben Halima, M., & Alimi, A. M., (2017), "Adaptive fuzzy exponent cluster ensemble system based feature selection and spectral clustering", *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, (pp. 1-6). IEEE.

[25] Ding, C. & He, X. (2004). "K-means clustering via principal component analysis", In *Proceedings of the 21st international conference on Machine learning*, (pp. 225–232), ICML 2004.

[26] Smith, M. R., & Martinez, T., (2018). "The robustness of majority voting compared to filtering misclassified instances in supervised classification tasks", *Artificial Intelligence Review*, 49(1), 105-130.

[27] Zhou, Z. H. (2012), *"Ensemble Methods: Foundations and Algorithms*", Chapman and Hall/CRC.

[28] Stepanov, E. A., & Riccardi, G. (2011), "Detecting General Opinions from Customer Surveys", *11th IEEE International Conference on Data Mining Workshops*, (pp. 115-122}, IEEE.

[29] Goncalves, P., Dalip, D. H., Costa, H., Goncalves, M. A., & Benevenuto, F. (2016). "On the Cobb ination of “Off-The-Shelf” Sentiment Analysis Methods" In *proceedings of the 31st ACM Symposium on Applied Computing (SAC)* (pp. 1158-1165), ACM

[30] Ruta, D., & Gabrys, B. (2005). "Classifier selection for Majority Voting", *Information Fusion*, 6(1), 63–81.

Saeed Dehghani Ashkezari is a PhD candidate of Computer Engineering in Yazd University. His work focuses mainly on the opinion mining, advanced algorithms and knowledge engineering.

Vali Derhami received the B.Sc. degree from Isfahan University of Technology, Iran, in 1996. He received M.S. and Ph.D. degrees from Tarbiat Modares University, Iran, in 1998 and 2007, respectively. Currently, he is Associate professor in computer engineering department in Yazd University. His research interests are neural fuzzy systems, intelligent control, reinforcement learning, and robotics.

Samira MashtiZadeh received the M.Sc. degree from Science and Art University.in 2018. Her main research areas are database management, opinion mining and machine learning

.