Cross Domain Sentiment Classification Techniques: A Review

Parvati Kadli
Associate Professor
PDIT, Hosapete
Karnataka, India

Vidyavathi B. M.
Professor
BITM, Ballari
Karnataka, India

ABSTRACT
With the explosive growth in the availability of online resources, sentiment analysis has become an interesting topic for researchers working in the field of natural language processing and text mining. The social media corpus can span many different domains. It is difficult to get annotated data of all domains that can be used to train a learning model. Hence continuous efforts are made to tackle the issue and many techniques have been designed to improve cross domain sentiment analysis. In this paper we present literature review of methods and techniques employed for cross domain sentiment analysis. The aim of the review is to present an overview of techniques and approaches, datasets used to solve cross domain sentiment classification problem in the research work carried out in the recent years.

General Terms
Sentiment Analysis, Classifier, Dataset, Features.

Keywords
Cross Domain Sentiment Classification (CDSC), Source Domain, Target Domain.

1. INTRODUCTION
Users express their opinions about products and services they consume in social media like reviews, blog spots, shopping sites, twitters etc. Sentiment analysis is a computational study of people's attitude, appraisals and opinions about individuals, issues, entities, topics, events and products [1]-[5]. Sentiment analysis includes the concepts of natural language processing, machine learning and computation linguistics. It aims at classifying sentiment data into polarity categories. Users do not specify sentiment polarity explicitly. Hence, we need to predict it from text data generated by users.

One of the main requirements for accurate performance is annotated data in various domains. This would imply huge cost for large numbers of domains and prevent us from exploiting the information shared across domains. Also, it is not feasible to develop different models for different domains for classification. Research work is taken up to solve this issue. One feasible solution is to develop a single system for sentiment classification using labeled and unlabeled data from different domains and apply it for any target domain. This is Cross Domain Sentiment Analysis. This study aims to present recent works on such cross-domain sentiment classification. Organization of the paper is, section 2 explains the challenges in CDSC and section 3 briefs the early research and baseline methods. Section 4 explains the key techniques for CDSC. The last sections present general discussion and conclusion.

2. CHALLENGES IN CDSC
The most critical challenge is that sentiment analysis is highly dependent on the domain i.e. a technique performing well on one domain might perform poorly on another. It is challenging as machine learning techniques used for cross domain classification perform well with labeled documents and hence are highly domain sensitive. A mismatch between review ratings and review text also affects performance [19].

We get inconsistent results because of poor target domain compared to rich labeled source domain, using which the classifier is trained. Some of the main challenges are as follows:

- Sparsity: When the target corpora contains words or phrases that do not appear or rarely appear in source domain.
- Polysemy: The meaning of the same word appearing in source and target domain changes based on the context of the respective domain.
- Feature Divergence: If the classifier is trained on source specific features and these may mismatch with domain specific features on which the classifier is applied. Feature divergence refers to the mismatch in source domain specific features and target domain specific features [6]-[7]
- Polarity Divergence: Same word may have difference polarity in different domains. Example cheap may be positive in one domain and may have negative meaning in some other domain.

3. EARLY RESEARCH AND BASELINE METHODS
In the early days, classifiers were trained and tested on a same domain. This is single domain classification. The first results of polarity classification using machine learning techniques were reported by Pang et al. [8]. Movie reviews were extracted from IMDB. First results on CDSC were given by Blitzer et al [20] Reviews on Books, Electronics, DVDs and kitchen domain were used. In other approaches groups of classifiers were trained on source domains [9]. For example in TPLSA (Topic-Bridged Probabilistic Latent Semantic Analysis) developed by [21] joint Probabilistic model is used to bridge the test and training domains. Identification of prime topic is obtained as a concurrent decomposition of contingency tables which are based on occurrence of terms in both test and training domain documents. Later collaborative dual PLSA was developed by [22] which exploited commonality and domain distinction among multiple domains. Document class and word concept are two latent concepts of this model. For Evaluation of new approaches developed baseline methods like SCL, SFA, SCL-MI techniques are used.

4. KEY TECHNIQUES FOR CDSC

Some of the key techniques developed are briefed as below:

4.1 Spectral Feature Alignment (SFA)

The algorithm [6] tries to find a new data representation which reduces the gap between source and target domain. Using the words which are domain independent a bipartite graph is constructed to model co-occurrence relationship between domain specific and domain independent words. It represents the probability of alignment of domain specific words to more common domain independent words. Feature clusters are formed by using spectral clustering algorithm on bipartite graph. Cluster thus reduces mismatch between domain specific words of different domains. This was used to train the classifier for sentiment classification. Experiments in the real-world domains have shown promising performance compared to other base line classifiers.

4.2 Structured Correspondence Learning

This algorithm was proposed by [11] to learn features from variety of domains. Unlabeled data from both source and target domains are used. The frequently occurring features in both domains called pivot features are estimated which are considered correspondences among features. Then a discriminative learner is used in training a classifier. An extension of SCL, SCL-Mutual Information (MI) model was developed by [7] as SCL depends on the choice of pivot features. If the choice is not good the performance is adversely affected. Here Using the mutual information between features and a domain label top pivot features are selected. Later the binary classifier is trained by the SCL algorithm & evaluated on test domain.

4.3 Joint sentiment topic (JST) model

The JST model [13], based on Latent Dirichlet Allocation (LDA) model [14], is a probabilistic modeling framework. JST is completely unsupervised. JST model is extension of LDA model [14]. This was developed to detect a topic and sentiment simultaneously from the text. Discriminative classifier marks a decision boundary that maximizes separation measure between classes in JST model. Clusters of different terms exhibiting a similar sentiment are formed by JST. Information gain criteria are used to select better features for CDSC. Later Dynamic JST was developed [15]. This identifies & tracks interests & changes the topic & sentiment with time. Dynamically both sentiment and topic are captured assuming the dependency of current sentiment-topic-specific word distributions on earlier distributions.

4.4 Active Learning and Deep Learning

Under the category of semi supervised machine learning, Active learning is considered as a special case. Here the learning algorithm interactively queries the user to get desired results at new data points [16]. i.e. It gets additional labeled target data from source domain information. [17] Proposed CDSC using an active learning approach. For sample selection a method called Query by committee (QBC) is incorporated and for classification combination of two classifiers is used. One classifier is trained on labeled source domain data & another on target domain labeled data. Later both are trained by unlabeled data of target domain with label propagation algorithm. These two classifiers select informative data by QBC and take combined classification decision. This approach was found to produce good results after addition of 1000 labeled sentiments from new domain to the existing data. The results thus attained accuracy approximately same as accuracy when trained on 10,000 annotated sentences.

The Deep learning technique is unsupervised and discovers intermediate concepts common to both target and source domains. These features are used to train the classifiers. In [18] first high-level features are extracted using stacked denoising Auto encoder with rectifier units. Second transformed labeled data from source domain are used by classifier for learning.

4.5 Topic Modeling

Here approaches are based on LSI. The aim is to get term-document matrix of low dimensions denoted on latent topics. Clustering techniques used in topic modeling do not require label information. There are four main techniques in this category:

4.5.1 Topical Correspondence Transfer (TCT)

In [24], the domain specific information is learnt from several domains and unified topics are created with the help of knowledge about shared topics. Documents are represented as term document matrix. By applying least squares penalty based on specific model, a document's sentiment labels are obtained. Thus, the differences among source & target domains are bridged by the hidden correspondence between the shared topics.

4.5.2 Bridged Topic Model (BTM)

In [25], Direct and Indirect co-citation relationships are found using an auxiliary link network. These relationships are then used to bridge the gap between source and target domains. Latest topic module is framed combining content information and link structure. In [17] senti-rank algorithm is used to get sentiment scores for target domain documents. Later Intrinsic structures of target domain are represented using the small numbers of labeled documents identified by source. Next the structure of target domain, manifold ranking scores, resulting from application of manifold ranking algorithm, labels the target domain data.

4.5.3 Latent Direct Analysis (LDA)

In [26], Real time transfer learning framework based on LDA is proposed. Here topic space is learnt from social streams in real time via online streaming LDA. Transfer learning framework is created by incorporating topic models learnt from social streams. This leads to real time CD graph spectra analysis.

4.5.4 Probability Latent Semantic Analysis (PLSA)

In [27] PLSA supervised adaptive transfer algorithm for CD text classification was proposed. PLSA modified using the latent variable made it a supervised learning algorithm. The class conditional probability of specific word conditioned on a class is estimated directly during initialization & is then fixed in the model fitting step to train the algorithm on source domain documents. For testing documents in target domain, the word category probabilities are assigned read only and learned. So, word category problems serve as bridge between two domains.

In [28] latent sentiment factorization algorithm based on probabilistic matrix factorization is developed. Sentiment correlations between domain shared and domain specific words in two dimensional spaces are exploited to bridge the gap between domains.

4.6 Approaches based on Thesaurus

CDSA can be done using thesaurus. In [29] feature mismatches are avoided by automatic classifier which is based on a sentiment sensitive thesaurus. The relatedness of characteristics is calculated from labeled data of several source data and unlabeled data from source & target data to conceptualize the sources. This conceptualized thesaurus is used to extend feature vectors, which are applied as training & test data on binary classifier. This approach gave comparatively better results than many baselines.

In [30] CDSC problem is modeled as embedded learning. A joint optimization method is developed to learn embeddings sensitive to classification. Optimizing three objective functions based on Distributional properties of pivot, Label constrains in source domain documents and Geometric properties in source and target domain unlabeled documents jointly revealed better performance in some experiments. With respect to only individual optimization, objective function based on geometric function has performed the best. In [31] vocabulary mismatches between source and target domains are addressed using word embeddings and canonical correlation analysis corresponding to feature learning and feature subspace mapping. It presents a generic method to solve the problem which is simpler yet produces competitive results in comparison with more complicated methods.

4.7 Case Based Reasoning (CBR) Techniques

Drawing knowledge from similar past examples and applying that knowledge to predict the outcome of new unseen case is the idea in case based reasoning approaches. In [32] CBR is used for handling CDSC problem. Here case base is developed from learning set of labeled out of domain opinion documents. Case base has two important portions.

Case Description:

It is feature vector based on a document's statistics which is used as a documents signature for retrieval purposes.

Case Solution:

This is information about successful predictions made during training. It contains all lexicons that made positive forecasts during training. The CBR technique was tested on user created reviews in six domains. It was compared to single lexicon classifier and the performance was found competitive.

In [33] domain explicit dictionary is built by combining large data from a specific domain and information from many preexisting dictionaries. Stochastically sentiment score was formulated and assigned to handle domain explicit variations.

4.8 Feature Based Techniques

In Features representation and transfer method [27], the main task is feature representation. Feature ensemble plus sample solution (SS-FE) is a comprehensive approach proposed in [41]. Here both labeling adaption & instance adaption are considered for domain adaption. FE model learns new labeling function in a feature reweighting manner. For instance, adaption PCA based sample selection is proposed. Domain pairs where distributions vary to larger extent improvement is due to instance adaption. In [34] different representations namely text based, features based, lexicon based and combined representations are desired to tackle domain dependence issue. An Ensemble algorithm consisting of several classifiers is created and each one is trained by one of the distinct feature representations. In [35] an approach which addresses both feature divergence and polarity divergence is proposed. A set of high polarity features are created using high polarity independent features of both domains and polarity of source domain features is transferred to the target domain.

4.9 Graph Based Approaches

Weighted graphs can be used to represent the data where data instances are vertices and weights on edges between vertices indicate the similarity between instances. If instances are strongly connected then they belong to same class. Label Propagation(LP) [36] is one of the first graph based algorithm developed. For SC documents are nodes and iterative process transfer information from labeled to unlabeled nodes. Iterations continue until convergence is achieved. In sentiment classification scenario the closeness of documents is denoted by edge weights. In [42] some modification is done on graph structures and parameters are varied to compare various graph based algorithms. In [43] effectiveness of graph based algorithm is compared. Here various sentiment similarity measures are investigated to assess better performance. In [38] Emotion keywords are employed to automatically extract labeled samples from target domain with high precision.

4.10 Domain Complexity and Similarity Approaches

One of the methods for domain adaption is by considering domain similarity. Samples from training data belonging to source domain that are similar to those in target domain are selected. The amount of domain similarity between source & target domains and degree of complexity of source & target domain helps to determine the reduction factor of training data set size. In [39] training data found are similar to test domain data as more similarity leads to more accurate performance. In document level polarity classification, it was found that rare words proportions correlate best with in-domain accuracy. Also, it was shown that performance loss was influenced by domain complexity represented as independent vector. In [40] divergence in term distribution and unigram distribution is domain similarity and domain complexity is assessed by homogeneity.

4.11 Knowledge Enhanced Meta Classifier

KE-Meta (Knowledge Enhanced Meta learning) [12] adds knowledge features to bag of words, n- grams or lexical resource based classifiers. Semantic network is used for word sense disambiguation. A vocabulary expansion based classifier is developed using the disambiguated terms.

4.12 Distance Based Model

In [10] review documents are classified using distance based predictive model. The distance metric and the training corpus are defined. A new review classified becomes part of training dataset and the distance metric is used to identify it. Majority rule strategy is used to classify the unlabelled reviews.

The table below gives the summary of selected research studies. They are sorted from early years to the most recent times.

Table 1. Summary of Methods, Datasets, Classifiers and Findings.

Publication	Dataset	Classifier	Findings
CDSC via SFA [6].	Product reviews from a) Amazon b) Yelp c) City search.	Simple Vector Machine (SVM)	Frame work is proved applicable for both entrance level and document level classification activities. The results are effective.
Domain adaption for large scale SC: A deep learning Approach [18].	Product reviews from Amazon	SVM	Domain adaption is successfully performed on an industrial scale dataset of 22 domains.
Automatically extracting polarity-bearing topics for CDSC [13].	Movie reviews from IMDB	SVM, Naive's Bayesian(NB), Maximum Entropy(ME)	Augmented features representation used to train in-domain supervised classifiers achieve state of art performance. It is simple & does not require parameter training.
A two stage frame work for CDSC [23]	Reviews of note books, books and hotels	Expectation Maximization (EM)	This can be used as high performance sentiment transfer technique as results shown high precision enhancement.
Social transfer Cross Domain transfer learning from social streams for media application [26]	Tweets and YouTube videos	SVM	Performance is better than in traditional learners, creates an interoperable connection across social domains & video leading to many CD applications.
Do neighbors help? An exploration of graph based algorithms for CDSC [37]	Product reviews of Amazon.	SVM (LIB SVM)	The best of the parameters are analyzed on two graph based algorithms. Results show that no optimal values for all domain pairs exist & that the values are influenced by the domain characteristics. Dominant regularity among number of source and target domain neighbors is not found.
Bibliographies or blenders which resource is best for Cross Domain Analysis [40]	Multi-domain Dataset from Amazon reviews	Linear Regression Model	Measures of domain similarity are found. Accuracy loss is modeled by a linear regression and tested. Accuracy loss is predicted with an average error of 15% & maximum error of 3.4%
Domain adaption using domain similarity & domain complexity based instance selection for CDSA [39]	10 product reviews of Amazon	SVM	Variance in domain complexity & similarity can be used for estimating parameter settings. Achieved better performance compared to natural baselines and also competitive results with state of art CDSC approaches.
A case based approach to CDSC [32]	1. IMDB dataset of film review. 2. Hotel Reviews 3. Product reviews from Amazon.com	K-Nearest Neighbor(KNN)	Demonstrates that preselection of lexicon corresponding to domain is not required & performs better than a baseline single lexicon classifier.
Semi supervised Vs. Cross domain graphs for SA [42]	Product reviews from Amazon	Graph based LP algorithm LIB SVM	Demonstrate that Graph based semi supervised method is suitable if there is large difference in source & target domains and GB-CDL is a competitive alternative to fully supervised technique.
Active learning for CDSC [17]	Multi-domain emotional comments	Maximum Entropy with LP based classifier	In this approach QBC based samples selected and combination based classifier have achieved comparable performance over in domain classifiers

			& some strong baselines.
Dynamic Joint Sentiment Topic model [15]	Movie review dataset	EM	The proposed approach sequentially updates the model with newly arrived data and show the effectiveness of model on the add on reviews entered between 2011to 2017
Employing emotion keywords to improve CDSC [38]	Multi domain emotional comments corpus	ME	Effectiveness is demonstrated by empirical results. Performance is superior to methods using only unlabeled target domain
Feature ensemble Plus sample selection: Domain adaption for sentiment classification [41]	Multi domain data set by Daume III	NB	As both labeling adaption and instance adaption are considered experimental results show significant improvements compared to individual FE and PCA-SS
CDSC using sentiment sensitive thesaurus [29]	product reviews of Amazon	L1 regularized logistic regression based binary classifier	Proposed method outperforms several baselines. The created SST groups words accurately expressing similar sentiments in comparison with Sentiwordnet.
A link bridged topic model for CDSC [25]	Scientific data from Cora data set	SVM	This model achieved effective knowledge transformation between domains. Prediction accuracy is significantly improved compared to state of art algorithms.
Data intensive review mining for SC across heterogeneous domains [10]	Trip Advisor dataset (hotel reviews from(tripadvisor.com)	KNN	Experimental results show satisfactory performances with respect to both accuracy and computational efficiency
An ensemble model for cross domain polarity classification on Twitter [34]	• Stanford twitter data set. • Obama healthcare reforms • Obama health care debate	SVM MNB	High accuracy of 81.81% on training set is obtained by combining algorithms trained on different features of generic training set. Better results are obtained compared to all out of domain approaches
Cross Domain opinion word identification with QBC active learning [43]	Review sentences on restaurant, movies & hotels. Tsai et.al. (2014)	SVM	The method shows that by adding only 1000 labeled sentences from the new domain to the existing labeled data systems achieves same level as if model trained with the 10000 labels.
Exploring ensemble models in taxonomy based CDSC [44]	Product reviews from three different domain trees in Amazon.	SVM with SFA	Experiments results show ensemble algorithms consisting of a SVM & SFA algorithm is able to comprehend the effect of different model algorithms.
Supervised PLSA for CDSC [27]	News Group posts on 20 sub categories. six datasets from twenty news groups,3 datasets from Reuters 21578.	EM algorithm	Efficient performance on nine cross domain text classification bench mark datasets is proved by Supervised adaptive transfer (SATPLSA) algorithms.
CDSC via topical correspondence transfer [24].	Reviews on Amazon.	SVM	Experiment conducted on reviews show TCT significantly outperforms the baseline methods & achieves accuracy which is competitive with state of art CDSC techniques.
Building domain specific sentiment lexicons combining information from many sentiment lexicons and a domain specific corpus. [33]	Product reviews from www.komplett.no. and mpx.no	Using sentiment lexicon & score.	Demonstration that combining information from both source sentiment lexicons and the domain specific corpus to build a lexicon, results in better performance than lexicon that depends

			only on source lexicons' information.
Cross domain polarity classification using knowledge enhanced meta classifier. [12]	Product reviews of Amazon.	SVM	The generic characteristic of KE meta is because meta classifier does not perform domain adaption. Additional information is provided by word sense disambiguation and vocabulary expansion which is not in bag of words and n- gram based classification.
CDSC feature divergence polarity divergence or both. [35]	Amazon product reviews	Linear classifier.	Results shows that TPF is superior to P only (polarity diverse) and F only (features divergence) and in comparison, with state of art algorithms TPF outperforms in 6 tasks.
CDSC with word embeddings and canonical correlation analysis. [31]	Product reviews of Amazon	SVM	Experiments shown that feature subspace mapping technique used makes this approach a generic one. It has achieved competitive results on 12 target source domain pairs.
Leveraging latent sentiment constraint in probabilistic matrix for CDSC. [28].	Product reviews of Amazon.com (by Blitzer et.al)	LIBSVM	Comparative study on LSF, SCL, SFA & TCT is made using Amazon datasets. LSF's performance is better and also achieves accuracy level comparable to TCT for CDSC.
CDSC using sentiment sensitive embeddings. [30]	Product reviews of Amazon.com (by Blitzer et.al)	Logistic regression classifier	The objective function which considers geometric properties in target and source domain document has resulted in best performance. Also, better performance is achieved by optimizing all objective functions rather than individual optimization.

5. DISCUSSION

Majorly studies in the area of sentiment classification aims at reduction in distribution difference among the domains. This is a trivial task as most of the techniques are domain dependent and distribution discrepancy in feature space reduces the efficiency. Performance of many techniques is dependent on the availability of labeled data. Larger the difference between test data and training data poorer is the performance. In sentiment classification studies generally deal with binary classification and unfortunately feasible results are not provided. On the same lines in cross domain learning, even though no human interference is required, one main factor is dependency on similarity between the domains under consideration. Therefore, accuracy can be improved by designing and applying novel methods for feature representation, extensive testing and realization of potential of different ensemble methods or combined methods. Also, polarity divergence and feature divergence should be given due importance in the methods. Results needs be stabilized across a wide range of domains. There are few more difficult challenges that need attention in the field of CDSC. Real world datasets of industries containing numerous domains pose similar challenges. Factors based on cultural diversities, linguistic variations, contextual differences and noises embedded in dataset affecting the CDSC techniques make it very difficult to gain high level of accuracy.

6. CONCLUSION

Training the learning models with annotated data for sentiment analysis aids for higher accuracy. But as there is lack of annotated data studies are focusing on developing techniques which are domain independent or deriving features that can bridge the gap across different domains considered for sentiment classification. We can conclude that to find better solutions for learning problems in CDSC researchers can work to develop learning methods considering the nature and structure of the data/reviews belonging to different domains and distributional similarity among the domains. Feature expansion and pivot features representation can be focused upon to develop better learning models.

7. REFERENCES

[1] Immonen, P. Pääkkönen, and E. Ovaska, ``Evaluating the Quality of Social Media Data in Big Data Architecture,'' *IEEE Access*, vol. 3,pp. 2028–2043,015.

[2] D. Jiang, X. Luo, J. Xuan, and Z. Xu, ``Sentiment computing for the news event based on the social media big data, "*IEEE Access*, vol. 5, pp. 2373–2382, 2016.

[3] M. N. Injadat, F. Salo, and A. B. Nassif, ``Data mining techniques in social media: A survey,'' *Neurocomputing*, vol. 214, pp. 654–670, Nov. 2016.

[4] T. A. A. Al-Moslmi, *Machine Learning and Lexicon-Based Approach forArabic Sentiment Analysis*. Bangi, Malaysia: Fakulti Teknologi & Sains Maklumat/Institut, 2014.

[5] N. Omar, M. Albared, T. Al-Moslmi, and A. Al-Shabi, ``A comparative study of feature selection and machine learning algorithms for Arabic sentiment classification,'' in *Information Retrieval Technology*. Springer, 2014, pp. 429–443.

[6] S. J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, ``Cross-domain sentiment classification via spectral feature alignment,'' in *Proc. 19th Int. Conf. Worldwide Web*, vol. 10. 2010, p. 751.

[7] J. Blitzer, M. Dredze, and F. Pereira, ``Biographies, bollywood, boomboxes and blenders: Domain adaptation for sentiment classification,'' in *Proc. ACL*, vol. 7. 2007, pp. 440–447.

[8] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up? Sentiment classification using machine learning techniques. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 79–86, 2002.

[9] S. Li and C. Zong, ``Multi-domain adaptation for sentiment classification: Using multiple classifier combining methods,'' in *Proc. Int. Conf. Natural Lang. Process. Knowl. Eng. (NLP-KE)*, 2008, pp. 1–8.

[10] F Bisio, P. Gastaldo, C. Peretti, R. Zunino, and E. Cambria, ``Data intensive review mining for sentiment classification across heterogeneous domains,'' in *Proc. IEEE/ACM Int. Conf. Adv. Social Netw. Anal. Mining*, vol. 13.Aug. 2013, pp. 1061–1067.

[11] J. Blitzer, R. McDonald, and F. Pereira, ``Domain adaptation with structural correspondence learning,'' in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2006, pp. 120–128.

[12] M. Franco-Salvador, F. L. Cruz, J. A. Troyano, and P. Rosso, ``Cross-domain polarity classification using a knowledge-enhanced meta classifier,'' *Knowl.-Based Syst.*, vol. 86, pp. 46–56, Jun. 2015.

[13] Y. He, C. Lin, and He, H. Alani, ``Automatically extracting polarity-bearing topics for cross-domain sentiment classification," in Proc.49th Annual meeting, Jun 2011,pp. 123–131.

[14] D. M. Blei and M. I. Jordan, ``Modeling annotated data,'' in *Proc. 26th Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2003, pp. 127–134.

[15] Y. He, C. Lin, W. Gao, and K.-F. Wong, ``Dynamic joint sentiment-topic model,'' *ACM Trans. Intell. Syst. Technol.*, vol. 5, no. 1, p. 6, 2013.

[16] B. Settles, *Active Learning Literature Survey*. Madison, WI, USA: Univ. Wisconsin, 2010.

[17] S. Li, Y. Xue, Z. Wang, and G. Zhou, ``Active learning for cross-domain sentiment classification,'' in *Proc. 23rd Int. Joint Conf. Artif. Intell.*, 2013,pp. 2127–2133.

[18] X. Glorot, A. Bordes, and Y. Bengio, ``Domain adaptation for large-scale sentiment classification: A deep learning approach,'' in *Proc. 28th Int. Conf. Mach. Learn.*, 2011, pp. 513–520.

[19] J.Carrillo de Albornoz, L. Plaza, and P. Gervás, ``A hybrid approach to emotional sentence polarity and intensity classification,'' in *Proc. 14th Conf. Comput. Natural Lang. Learn.*, 2010, pp. 153–161.

[20] J. Blitzer, M. Dredze, F. Pereira, Biographies, bollywood, boomboxes and blenders: Domain adaptation for sentiment classification, in: Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, 2007, pp. 187–205.

[21] G.-R. Xue, W. Dai, Q. Yang, and Y. Yu, ``Topic-bridged PLSA for crossdomain text classification,'' in *Proc. 31st Annu. Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2008, pp. 627–634.

[22] F. Zhuang *et al.*, ``Collaborative dual-PLSA: Mining distinction and commonality across multiple domains for text classification," in *Proc. 19th ACM Int. Conf. Inf. Knowl. Manage.*, 2010, pp. 359–368.

[23] Q. Wu and S. Tan, ``A two-stage framework for cross-domain sentiment classification," *Expert Syst. Appl.*, vol. 38, no. 11, pp. 14269–14275, 2011.

[24] G. Zhou, Y. Zhou, X. Guo, X. Tu, and T. He, ``Cross-domain sentiment classification via topical correspondence transfer," *Neurocomputing*, vol. 159, pp. 298–305, Dec. 2015.

[25] P. Yang, W. Gao, Q. Tan, and K.-F. Wong, ``A link-bridged topic model for cross-domain document classification," *Inf. Process. Manage.*, vol. 49, no. 6, pp. 1181–1193.Nov.2013.

[26] S. D. Roy, T. Mei, W. Zeng, and S. Li, ``Social Transfer: Cross-domain transfer learning from social streams for media applications," in *Proc. 20th ACM Int. Conf. Multimedia*, 2012, pp. 649–658.

[27] R. Zhao and K. Mao, ``Supervised adaptive-transfer PLSA for cross domain text classification," in *Proc. IEEE Int. Conf. Data Mining Work shop*, Jan. 2014, pp. 259–266.

[28] J. Liang, K. Zhang, X. Zhou, Y. Hu, J. Tan, and S. Bai, ``Leveraging latent sentiment constraint in probabilistic matrix factorization for cross domain sentiment classification," *Proc. Comput. Sci.*, vol. 80, pp. 366–375, Mar. 2016.

[29] D. Bollegala, D. Weir, and J. Carroll, ``Cross-domain sentiment classification using a sentiment sensitive thesaurus," *IEEE Trans. Knowl. DataEng.*, vol. 5, no. 8, pp. 1719–1731, Aug. 2013.

[30] D. Bollegala, T. Mu, and J. Y. Goulermas, ``Cross-domain sentiment classification using sentiment sensitive embeddings," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 2, pp. 398–410, Feb. 2016.

[31] N. X. Bach, V. T. Hai, and T. M. Phuong, ``Cross-domain sentiment classification with word embeddings and canonical correlation analysis," in *Proc. 7th Symp. Inf. Commun. Technol.*, 2016, pp. 159–166.

[32] B. Ohana, S. J. Delany, and B. Tierney, ``A case-based approach to cross domain sentiment classification," in *Lecture Notes in Computer Science*, vol. 7466. Cham, Switzerland: Springer, 2012, pp. 284–296.

[33] H. Hammer, A.Yazidi, A. Bai, and P. Engelstad, ``Building domain specific sentiment lexicons combining information from many sentiment lexicons and a domain specific corpus," *Comput. Sci. Appl.*, vol. 456, pp. 205–216, Dec. 2015.

[34] A Tsakalidis, ``An ensemble model for cross-domain polarity classification on twitter," in *Web Information Systems Engineering*. Cham, Switzerland: Springer, 2014, pp. 168–177.

[35] Y. Zhang, X. Hu, P. Li, L. Li, and X. Wu, ``Cross-domain sentiment classification-feature divergence, polarity divergence or both?" *Pattern Recognit. Lett.*, vol. 65, pp. 44–50, Jun. 2015.

[36] X. Zhu and Z. Ghahramani, *Learning From Labeled and Unlabeled Data With Label Propagation*. Seattle, WA, USA: Semantic Scholar, 2002.

[37] N. Ponomareva and M. Thelwall, ``Do neighbours help?: An exploration of graph-based algorithms for cross-domain sentiment classification," in *Proc. Joint Conf. Empirical Methods Natural Lang. Process. Comput Natural Lang. Learn.*, Jul. 2012, pp. 655–665.

[38] Z. Zhu, D. Dai, Y. Ding, J. Qian, and S. Li, ``Employing emotion keywords to improve cross- domain sentiment classification," in *Proc. Chin. Lexical Semantics*, 2013, pp. 64–71.

[39] R. Remus, ``Domain adaptation using domain similarity- and domain complexity-based instance selection for cross-domain sentiment analysis," in *Proc.-12th IEEE Int. Conf. Data Mining Workshops*, Jun. 2012, pp. 717–723.

[40] N. Ponomareva and M. Thelwall, ``Biographies or blenders: Which resource is best for cross-domain sentiment analysis?" in *Lecture Notes in Computer Science*, vol. 7181. Cham, Switzerland: Springer, 2012, pp. 488–499.

[41] R. Xia, C. Zong, X. Hu, and E. Cambria, ``Feature ensemble plus sample selection: Domain adaptation for sentiment classi_cation," *IEEE Intell. Syst.*, vol. 28, no. 3, pp. 10–18, May 2013.

[42] N. Ponomareva and M. Thelwall, ``Semi-supervised vs. cross-domain graphs for sentiment analysis," in *Proc. RANLP*, 2013, pp. 571–578.

[43] Y. Tsai, R. T. Tsai, C. Chueh, and S. Chang, ``Cross-domain opinion word identification with query-by-committee active learning," in *Technologies and Applications of Artificial Intelligence*. Cham, Switzerland: Springer, 2014, pp. 334–343.

[44] C. Lin, Y. Lee, C. Yu, and H. Chen, ``Exploring ensemble of models in taxonomy-based cross-domain sentiment classification," in *Proc. 23rd ACM Int. Conf. Conf. Inf. Knowl. Manage.-(CIKM)*, 2014, pp. 1279–1288.