International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682
Vol.3, No.8, August 2014 DOI:10.15693/ijaist/2014.v3i8.114-119

Survey Paper: Domain Adaptation for Sentiment Classification

P.Manivannan
Assistant Professor/Department of CSE
Arulmurugan College of Engineering, Karur, India

Dr.C.S.Kanimoli Selvi
Associate Professor/Department of CSE
Kongu Engineering College, Erode, India

***Abstract*— Online reviews make sentiment classification an interesting topic in industrial research. Given a review about a product – the goal is to classify whether it is positive or negative. Reviews are in different domain and it is difficult to collect data and train them for the entire domain. Domain adaptation is a fundamental problem in natural language processing (NLP). Transfer learning or domain adaptations are tools for sentiment analysis applications. In Sentiment classification the labeled training data comes from one distribution (source domain) and test data from other distribution (target domain). Such mismatches are considered different and it is usually very hard to measure and formulate these distribution differences. In order to build a sentiment classifier using text categorization methods it is necessary to take into account between these two distributions. The need for domain adaptation is widespread in many real world applications especially in computer vision this is the main challenge. Domain adaptation or Transfer learning is a key feature that requires a useful framework for sentiment classification. This paper provides an overview of literature dealing with challenges and techniques for domain adaptation with special focus on natural language processing.**

***Index terms* - Sentiment Classification, Domain Adaptation, Natural Language Processing, Transfer Learning, Classifier.**

I. INTRODUCTION

In Natural language processing (NLP), opinion mining is one of the applications in recent years with the biggest growth in industry applications. Web is the best way of expressing opinion about products. Opinion is huge in number of websites and is growing rapidly. They refer opinion mining as text classification task. Sentiment analysis is a subfield of text mining. In many aspects it is often referred to as opinion mining, sentiment classification, sentiment analysis and subjectivity analysis. It analyzes opinions of peoples expressed in text which could be emotions, judgments, evaluations and wishes. People not only express opinion about products and service but also various topics and issues from social domain. The basic task of sentiment analysis is classifying the polarity of given text as positive or negative.

Sentiment analysis is done at 3 different levels. They are Document level, Sentence level and Entity or Aspect level [1]. In Document level, analysis is performed for the whole document, in Sentence level it is related to find sentiments from sentences, whether each sentences express a positive, negative or neutral opinion and in Entity or Aspect level they perform fine-grained analysis. In document level sentiment classification most of the work has been focused using machine learning techniques [2] [3]. Major challenges involved in sentiment analysis are Named Entity Extraction, Information Extraction, Sentiment Determination, Co-reference Resolution, Relation Extraction and Domain Dependency. Content in languages are ambiguous and it is hard to analysis. New challenges in sentiment analysis are addressed in [13].

There are two types of sentiment classification technique for text used in natural language processing. They are supervised and un-supervised classification techniques. Supervised classification techniques [2] are used in machine learning approach to classify the review. It is most widely used in movie reviews to perform sentiment analysis. Corpus is formed to represent data in the document, and then the corpus is trained using the classifier. Naïve Bayes , Maximum Entropy and Support Vector Machines are machine learning approaches trained on different features sets. SVM outperforms all the other approaches presented in the work which is trained on unigram bag-of-words features set. Machine learning approaches has drawback on opinion polarity detection for both data set topic and time period for collected data set [4].

Un-supervised classification techniques it does not require any training data. K-means clustering, Hierarchical clustering are clustering algorithms used to classify data into categories. Turney et al. [5] used Point-wise Mutual Information unsupervised learning techniques for classifying reviews. For phrase the semantic orientation is calculated by comparing its similarity. Natural language processing approaches in un-supervised classification are domain independent opinion features [7][8][9]. Domain dependent is also considered for phrase sentiments using contextual Information. Zhang et al. [10] used word dependency structure to classify the sentiment using rule based semantic analysis. Maas et al. [11] used both supervised and un-supervised techniques by capturing semantic term document information to learn word vector.

In addition to the above two classification technique we have another one which is called as Semi-supervised classification approach. Several NLP task has been applied. In [46] Clark et al. achieve improvements by applying self-training to POS-tagging. In [47] Steedman et al. used parsing, he report that self-training improves the accuracy. In domain adaptation self training is applied on several NLP tasks such as cross language text classification (shi et al. [48]), POS tagging (Jiang and Zhai [49]). The performance improves on the target domain by modeling annotated

source domain data and unannotated target domain data in the training process. In this review the literature focus on challenges on cross domain adaptation and the number of methods proposed for cross domain adaptations.

II. Domain adaptation

In supervised learning techniques labeled data are widely used for training and it has been successfully applied to build a classifier for a particular domain in sentiment classification. The performance of supervised learning algorithm depends on labeled training data and also expensive to build an accurate sentiment classifier. Hence these approaches are domain dependent. When we apply the trained classifier for different domain their performance will low due to mismatch of data in training and testing. In order to handle the mismatch between the training and test data from different distribution we need to build a classifier using the labeled data in one domain to perform well in other domain. This problem which we call it is as "Transfer Learning" or "Domain Adaptation".

In sentiment classification problem we require large amount of labeled data for training. These data can be time-consuming and expensive for acquisition. Training a classifier in particular domains will lack in large number of training examples. These approaches are investigated and compared in the paper [30]. Challenges in domain adaptation are

1. How much domain are related to define measure allowing one to quantify
2. To determine the attitude of the review
3. Some review will not express any sentiments
4. Overall sentiment polarity of the text
5. Natural Language Processing techniques to preprocess

Domain adaptation is an interesting topic which has constant attracting attenuation. The problem of domain adaptation is to generalize the training labeled data in source domain to an unlabeled data in the target domain or labeled data in target domain. In one domain the same word will be positive, but in other domain it may refer to negative. For example "This Camera Suck" usually indicates negative sentiments for camera domain but "This vacuum cleaner sucks" it is positive sentiments for vacuum cleaner domain. It is important for machine learning, computer vision and natural language processing [12]. This idea is proposed by Aue & Gamon [21] to learn a polarity classifier. They proposed opinionated data from one or more domain that are different from target domain. Significant difference in data the classifier learned one domain not performed well in other domain. Here we are going to discuss most domain adaptation algorithm.

III. Algorithm

A. Structured Correspondence Learning Algorithm

Structured Correspondence Learning (SCL) proposed by John Blitzer et al. [14] to learn the features from different domain. It is a discriminative model for adapting source domain to target domain. This domain adaptation is a binary classification which is defined as "A domain is a pair consisting of a distribution D on X and a labeling function f : X [0,1]". They consider two domains, a source domain and a target domain. They measure the distance between two distribution using hypothesis distance measures using KL divergence. They identify features from different domain by modeling their correlations with pivot features which are most useful for semi-supervised learning. Non-pivot features are correlated with the same pivot features which are assumed to be similar, and treat them in discriminative learner. POS tag is used for end-to-end parsing to improve accuracy. In this case no labeled training data is used for test domain hence it is useful to model and utilize the correlation between features in different domains.

B. Spectral Feature Alignment (SFA) Algorithm

Spectral Feature Alignment (SFA) proposed by Sinno Jialin Pan et al. [15] to align words which are specific to that particular domain from different domains into unified cluster, with the help of domain independent words as a connection. It reduces the gap between domain-specific words of the two domains, and then it is used to train sentiment classifier in the target domain accurately. To construct domain independent words they used bipartite graph to model co-occurrence relationship between domain-specific and domain-independent words into a set of clusters, which reduce the mismatch between domain specific words of both domains. With the help of this cluster they train classifier for sentiment classification. They applied this algorithm for both document level and sentence level sentiment classification task. This is an improved method compared to SCL for all most all pairs of domain.

C. Joint sentiment-topic model (JST)

Joint sentiment topic model was proposed by He et al. [22]. It is extended form Latent Dirichlet Allocation (LDA) model proposed by Blie et al. to detect sentiment and topic simultaneously from text. Discriminative classifiers seek to find a decision boundary that maximizes certain measure of separation between classes. JST learning is domain independent polarity word prior information. It is an In-domain supervised classifier. Gibbs sampling was used to estimate the posterior distribution by sequentially sampling each variable. It allows clustering different terms which share similar sentiment. Feature argumentation and selection according to the information gain criteria for cross domain classification.

D. Active Learning

Active learning is a supervised learning techniques used to control data in which the leaner is used for learning data. . It is applied in information extraction [17], named entity recognition [20], parts of speech tagging [18], text categorization [16] and word sense disambiguation [19]. It takes input from a set of labeled data and produce classifier without using target label relatively to set of new labeled data. This is the method used to leverage the source domain information for acquiring some extra labeled target data. The key is to choose the data from which it learns to perform well with less training. The challenge in active learning is size of labeled data both in source domain and target domain. In [6]

he proposed a approach for cross domain sentiment classification by leveraging QBC based sample selection.

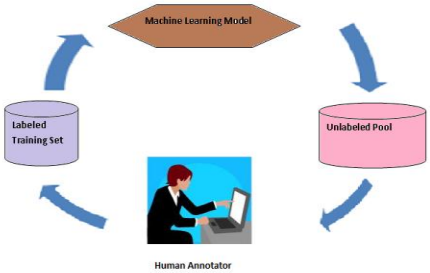

Fig. 1. Active Learning

E. Graph Based Algorithm

It is a group technique to model data as a graph for documents [23] [24] [26]. It represents relationship among each pair in the labeled and unlabeled data. The problem is choosing the best set of parameter to demonstrate the characteristic of a corresponding domain pair. [24] Proposed a graph-based method which is a semi-supervised learning by applying label propagation to constructs a graph with labeled and unlabeled examples as nodes and their similarity relationships as edges. It makes the assumption of label smoothness over the graph. [28] In this frame work to learn predictive structures on hypothesis spaces using unlabeled data, and then use these structures to enhance learning. In [29], they explore graph-based algorithms which refer to a group of techniques that model data as a graph of documents. This data representation takes into account not only document contents but also document connectivity which is modeled as document sentiment similarity rather than content similarity.

F. Sentiment sensitive thesaurus

Sentiment sensitive thesaurus method proposed by Bollegala et al. [27] relies on word co-occurrences. The method for automatic construction of a sentiment-sensitive thesaurus using labeled and unlabeled review using POS tagging to filter out function words. In order to find the word co-occurrence they used point wise mutual information where each lexical element is either unigram or bigram connected to a list of related lexical elements which appear most frequently in the sentence. To expand the feature vector they use this thesaurus with related element on the training step to defeat the feature mismatch problem. Relatedness measure is used to measure the neighbors for particular lexical element from the list of different lexical element which is proposed by the author. The relatedness measure is compared with various measures and the proposed method achieved the overall performance followed by Reversed baseline, Lin's Similarity Measure, and the Cosine Similarity. Binary classifier is used for training and testing.

IV. Feature Extraction Techniques

A. POS – Parts of Speech

Classifications of words according to their relations t o each other and to the things they represent.Different parts of speech name actions, name the performers of actions, describe the performers oractions, and so on. The common parts of spe ech are adjectives, adverbs, articles, conjunctions,interjections, nouns, prepositions, pronouns, and verbs. The work on subjectivity detection [31] between the presence of adjectives and sentence subjectivity meant for POS tag. [2][32] Using adjectives as the only features results in worse performance than using the same number of most frequent unigrams.

B. CRF – Conditional Random Field

It is a class of statistical modeling method often applied in pattern recognition and machine learning, where they used for structured prediction. It is an extension of logistic regression to sequential data, which are used in natural language processing. Conditional Random Fields (CRF) in [33] takes input as a sequence of tokens, to calculate the probabilities of the various possible labeling and chooses the one with the maximum probability. In [34] used to extract opinion targets with the help of noun combination which is modeled as sequence segmentation and labeling task.

C. Hidden Markov Model

Hidden Markov Model is popular in building system for tag disambiguation. In parts-of-speech tagging it is one the best known stochastic algorithm. [43] Russsel and Norvig gives a more general introduction to Hidden Marko Models. According to joint distribution P(X, Y), Hidden Markov Model belongs to group of generative models as they assume the data. Hidden Markov Models (HMMs) on unlabeled sentences from two domains [44] Huang & Yates proposed to induce hidden states as latent features for training. They empirically demonstrated the efficacy of their approach on out-of-domain part-of-speech tagging and syntactic chunking tasks. Their learning technique is also further exploited in [45] (Huang & Yates), which aims to learn a multidimensional feature representation by simultaneously training multiple HMMs with different initializations.

D. TF & IDF – Term Frequency & Inverse Document Frequency

Term frequency and Inverse document frequency are based on traditional term weighting functions used to identify the index term and also its sequence using the vectors to be processed. In sentiment classification some these techniques are adaptated [37][38]. In [39] TF-IDF weighting method is applied to obtain clustering for opinion. It is also used in discovering the bridging terms which relate to different domains [40]. In [41] they used to transform the original feature space into another low or high dimensional feature space that reduces the domain distance.

V. Machine learning techniques

A. Point wise Mutual Information (PMI)

One of the approaches that can determine which words are mostly used in natural language to express positive

or negative sentiment is calculating the PMI of each word in a document with respect to the label of that document in the training phase. For instance the PMI of the word "Excellent" is positively higher in "Positive" labeled documents than in "Negative" labeled documents where the PMI is negatively lower. Intuitively, by selecting those words that have higher PMI in positive and negative documents, which means they are mostly used to express positive and negative sentiments and consequently can lead to better classifiers. On the other hand, as we can select the best features (words) based on their PMI rank, we can manage the dimensionality of the feature vector for each document and therefore construct a more reasonable matrix in terms of dimensionality per each dataset, which implies faster classification.

B. Mutual Information

Mutual information is used to measure the mutual dependence between two random variables in information theory. Feature selection using mutual information can help in identifying features relevant to source domain labels. In [33], mutual information is applied on source domain labeled data to select features as "pivots". Feature selection using mutual information can help identify features relevant to source domain labels. In [34] "pivots" are referred as domain independent features to measure the dependence between features and domain. It is domain specific if the feature has high mutual information otherwise it is domain dependent.

C. Chi-Square

Chi Square is used to measure the dependence between the word feature and its class categories in text classification. It is a statistical test based on feature occurrence which is an independent class value [35][36]. It is also used measure term similarity between the source domain (sd) which is called as term and target domain (td) which is called as category. In order to measure the dependence between sd and td, where A is number of sentences in which sd and td co-occur, B is the number of sentences in which sd co-occur without td, C is the number of sentences in which td co-occur without sd, D is the number of sentences in which neither td not sd occur where A,B,C,D (or N) is the total number of sentences occurs in the documents..

$$x^2(sd,td) = \frac{N*(AD-CB)^2}{(A+C)*(B+D)*(A+B)*(C+D)}$$

If chi-square score is zero, then the sd and td are independent.

VI. Conclusion

In this paper we studied a supervised and un-supervised learning method for cross domain sentiment classification. Sentiment classification is a domain specific in which classifier trained in one domain will not perform in other domain. However the classifier has to customize the new domain. Text categorization is a challenge task to build a high quality sentiment classifier due to lack of labeled data in a target domain. Identification of features and opinion is important in feature based opinion mining techniques. Different algorithm for domain adaptation we discussed in this paper is applied to provide a better classification. In order to improve accuracy in the classification more work has to carry out with various feature extraction and machine learning techniques. This review gives a better understanding for researchers in cross domain sentiment classification.

TABLE I. SUMMARY OF THE SURVEY

Paper	ML/NLP	Algorithms	Dataset	Classifier
[25]	POS	SCL	Amazon	SVM
[14]	POS	SCL-MI	Amazon	SVM
[37]	TF and IDF	PLSA	UseNet news articles, SRAA & Newswire articles	SVM and NBC
[42]	N-gram method	feature-level fusion	Amazon	LIBSVM
[20]	LaSA and LDA	MI	Wikipedia, Chinese newspapers	RRM classifier
[41]	TF and IDF	MI	Newsgroup & SRAA	NBC
[15]	Bipartite Grpah	SFA	Amazon, Yelp and Citysearch websites	SVM
[22]	Gibbs sampling	JST	MDS	SVM,NBC, ME
[29]	POS	RANK and OPTIM alg.	Amazon	SVM
[6]	Bipartite Graph	LP	Amazon	ME
[27]	POS	PMI	Amazon	Binary Classifier
[38]	TF and IDF	SMART and BM25 tf.idf variants	Amazon, movie dataset from new groups	SVM and NBC
[34]	POS	CRF Algorithm	Amazon	SVM
[40]	TF and IDF	CrossBee system	Gold Standard Dataset	Centroid Similarity, Random Forest & SVM.
[45]	CRF	HHM based Model	Wall Street Journal	ME

REFERENCES

[1]. Bing Liu, Sentiment Analysis and Opinion Mining, Morgan and Claypool Publishers, May 2012.p.18-19,27-24,44-45,47,90-101.
[2]. B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs Up? Sentiment Classification Using Machine Learning Techniques," Proc. ACL-02 Conf. Empirical Methods in Natural Language Processing, pp. 79-86, 2002.
[3]. B. Pang and L. Lee, "Opinion Mining and Sentiment Analysis," Foundations and Trends in Information Retrieval, vol. 2, nos. 1/2, pp. 1-135, 2008.
[4]. Read, J., Using emoticons to reduce dependency in machine learning techniques for sentiment classification. In Proc of the ACL Student Research Workshop, 2005.
[5]. Turney, P. D. Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews. In Proc of the 40th Ann meeting of the ACL, 417–42, 2002.
[6]. Shoushan Li, Yunxia Xue, Zhongqing Wang, Guodong Zhou. "Active Learning for Cross- Domain Sentiment Classification". In Proc of the 23rd Int'l Joint Conf on Artificial Intelligence, 2127-2133, 2013.
[7]. S.-M. Kim and E. Hovy, "Extracting Opinions, Opinion Holders, and Topics Expressed in Online News Media Text," Proc. ACL/COLING Workshop Sentiment and Subjectivity in Text, 2006.
[8]. G. Qiu, C. Wang, J. Bu, K. Liu, and C. Chen, "Incorporate the Syntactic Knowledge in Opinion Mining in User-Generated Content," Proc. WWW 2008 Workshop NLP Challenges in the Information Explosion Era, 2008.
[9]. G. Qiu, B. Liu, J. Bu, and C. Chen, "Opinion Word Expansion and Target Extraction through Double Propagation," Computational Linguistics, pp. 9-27, 2011.
[10]. C. Zhang, D. Zeng, J. Li, F.-Y. Wang, and W. Zuo, "Sentiment Analysis of Chinese Documents: From Sentence to Document Level," J. Am. Soc. Information Science and Technology, vol. 60, no. 12, pp. 2474-2487, Dec. 2009.
[11]. A.L. Maas, R.E. Daly, P.T. Pham, D. Huang, A.Y. Ng, and C. Potts "Learning Word Vectors for Sentiment Analysis," Proc. 49th Ann. Meeting of the ACL: Human Language Technologies, pp. 142-150, 2011.
[12]. H. Daume III. Frustratingly easy domain adaptation. In ACL, 2007.
[13]. Pang.B and Lee.L (2012), 'Opinion Mining and Sentiment Analysis', Foundations and Trends in Information Retrieval, Vol. 4, No. 10 pp. 1-168
[14]. J. Blitzer, M. Dredze, and F. Pereira, "Biographies, Bollywood, Boom-Boxes and Blenders: Domain Adaptation for Sentiment Classification," Proc. 45th Ann. Meeting of the ACL '07, pp. 440-447, 2007
[15]. S.J. Pan, X. Ni, J.-T. Sun, Q. Yang, and Z. Chen, "Cross-Domain Sentiment Classification via Spectral Feature Alignment," Proc. 19th Int'l Conf. World Wide Web, 2010.
[16]. Hoi, Steven C. H., Rong Jin and Michael R. Lyu 2006. Large-scale text categorization by batch mode active learning. Proc. 15th Int'l World Wide Web Conf 2006.
[17]. Culotta, Aron, Trausti Kristjansson, Andrew McCallum and Paul Viola 2006. Corrective feedback and persistent learning for information extraction. Journal of Artificial Intelligence 170 (14): 1101–1122.
[18]. Ringger, Eric, Peter McClanahan, Robbie Haertel, George Busby, Marc Carmen, James Carroll, Kevin Seppi and Deryle Lonsdale 2007. Active learning for part-of-speech tagging: Accelerating corpus annotation. Proc of the Linguistic Annotation Workshop, 101–108. ACL, Prague, Czech Republic.
[19]. Zhu, Jingbo, Huizhen Wang and Eduard Hovy. Learning a stopping criterion for active learning for word sense disambiguation and text classification. Proc of the 3rd Int'l Joint Conf on Natural Language Processing (IJCNLP 2008), 366–372. Hyderabad, India.
[20]. H. Guo, H. Zhu, Z. Guo, X. Zhang, X. Wu, and Z. Su, "Domain Adaptation with Latent Semantic Association for Named Entity Recognition," Proc. Ann. Conf. North Am. Ch. ACL (NAACL '09), pp. 281-289, 2009..
[21]. Aue, A. & Gamon, M. Customizing sentiment classifiers to new domains: a case study. In Proc of Int'l Conf on Recent Advances in Natural Language Processing 2005, 207--218.
[22]. He Y., C. Lin and H. Alani. 2011. Automatically Extracting Polarity-Bearing Topics for Cross-Domain Sentiment Classification. In Proceeding of ACL-11, pp.123-131.
[23]. Bo Pang and Lillian Lee. 2005. Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales. In Proc of the 43rd Ann Meeting on ACL '05, pages 115–124
[24]. Xiaojin Zhu, Zoubin Ghahramani, and John Lafferty. 2003. Semi-supervised learning using gaussian fields and harmonic functions. In Proc of the 20th Int'l Conf on Machine Learning (ICML '03), pages 912–919.
[25]. John Blitzer, Ryan McDonald, and Fernando Pereira 2006. Domain adaptation with structural correspondence learning. EMNLP.
[26]. Qiong Wu, Songbo Tan, and Xueqi Cheng. Graph ranking for sentiment transfer. In Proc of the ACL-IJCNLP 2009 Conf Short Papers, pages 317–320.
[27]. Danushka Bollegala, David Weir, and John Carroll. Cross domain Sentiment using sentiment sensitive thesaurus. In IEEE Transaction on Knowledge and Data Engineering, Vol . 25, No. 8, August 2013 pages 1719–1731.
[28]. Ando, R. K. and Zhang, T. A Framework for Learning Predictive Structures from Multiple Tasks and Unlabeled Data. J. Mach. Learn. Res. 6, pp. 1817-1853, 2005.
[29]. Natalia Ponomareva, Mike Thelwall. Do Neighbours Help? An Exploration of Graph-based Algorithms for Cross-domain Sentiment Classification. Proc of the 2012 Joint Conf on Empirical Methods in NLP and Computational Natural Language Learning, pages 655–665, ACL
[30]. A. Aue and M. Gamon, "Customizing Sentiment Classifiers to New Domains: A Case Study," technical report, Microsoft Research, 2005.
[31]. Vasileios Hatzivassiloglou and Janyce Wiebe, 2000. Effects of Adjective Orientation and Gradability on Sentence Subjectivity. In Proc of the Int'l Conf on Computational Linguistics (COLING), pages 299-305.
[32]. Farah Benamara, Carmine Cesarano, Antonio Picariello, Diego Reforgiato and V. S. Subrahmanian, 2007.

Sentiment Analysis: Adjectives and Adverbs are Better than Adjectives Alone. In Proc of the Int'l Conf on Weblogs and Social Media (ICWSM).

[33]. John Lafferty, Andrew McCallum, and Fernando Pereira. 2001. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In Proc of the 18th Int'l Conf on Machine Learning, ICML'01.

[34]. Niklas Jakob, Iryna Gurevych, Extracting Opinion Targets in a Single- and Cross-Domain Setting with Conditional Random Fields Proc of the 2010 Conf on Empirical Methods in NLP, pages 1035–1045, MIT, Massachusetts, USA, 9-11 October 2010.ACL

[35]. Tan, S. and Zhang, J. An empirical study of sentiment analysis for chinese documents. Expert Systems with Applications, 34, 2008, 2622-2629.

[36]. Yang, Y., & Pedersen, Jan O. (1997). A comparative study on feature selection in text categorization. ICML, 412–420.

[37]. Xue, G. R., Dai, W. Y., Yang, Q., & Yu, Y. Topic-bridged PLSA for cross-domain text classification. In SIGIR-2008 (pp. 627–634).

[38]. G. Paltoglou and M. Thelwall, "A study of Information Retrieval weighting schemes for sentiment analysis," In Proc. of 48th Ann Meeting of the ACL, pp. 1386-1395, 2010.

[39]. G. Li and F. Liu, "A Clustering-Based Approach on Sentiment Analysis," In Proc. of Int'l Conf. on Intelligent Systems and Knowledge Engineering, pp 331-337, 2011.

[40]. Matjaz Jursic, Bojan Cestnik, Tanja Urbancic, Nada Lavrac. Cross-domain Literature Mining: Finding Bridging Concepts with CrossBee in Int'l Conf on Computational Creativity 2012.

[41]. Wang, P., Domeniconi, C., & Hu, J. Using wikipedia for co-clustering based cross-domain text classification. In: ICDM-2008 (pp. 1085–1090).

[42]. Shoushan Li and Chengqing Zong. 2008. Multi-domain sentiment classification. In Proc of ACL-08: HLT, Short Papers, pp 257–260.

[43]. Russell, S. J. and Norvig, P. (2010). Artificial Intelligence - A Modern Approach. Pearson Education, 3rd edition.

[44]. Huang, F. and Yates, A. Distributional representations for handling sparsity in supervised sequence labeling. In Proc. of the Annual Meeting of the ACL, 2009.

[45]. Huang, F. and Yates, A. Exploring representation learning approaches to domain adaptation. In Proc. of the Annual Meeting of the ACL, 2010.

[46]. Clark, S., Curran, J., and Osborne, M. (2003). Bootstrapping POS-taggers using unlabelled data. In Proc of CoNLL-2003, pages 49–55.

[47]. Steedman, M., Osborne, M., Sarkar, A., Clark, S., Hwa, R., Hockenmaier, J., Ruhlen, P., Baker, S., and Crim, J. (2003). Bootstrapping statistical parsers from small datasets. In Proc of EACL, Budapest, Hungary.

[48]. Shi, L., Mihalcea, R., and Tian, M. (2010). Cross language text classification by model translation and semi-supervised learning. In Proc of the 2010 Conference on Empirical Methods in NLP, pages 1057–1067, Cambridge, MA. ACL.

[49]. Jiang, J and Zhai C. (2007). Instance weighting for domain adaptation in nlp. In Proc of the 45th Annual Meeting of the ACL, pages 264-271, Prague, Czech Republic.

Authors Profile

P.Manivannan is a faculty member in the Department of Computer Science and Engineering of Arulmurugan College of Engineering, Karur. He received Bachelor degree B.Sc., in Computer Science from Government Arts and Science College, Karur and M.Sc., in Computer Science from Bishop Heber College, Tiruchirappalli in 2003. Then he received M.E., degree in Computer Science and Engineering from the Thiagarajar College of Engineering, Madurai, Anna University, Chennai in 2008. Currently he is doing **Ph.D.** in Anna University, Chennai, India. His research interest includes Opinion Mining.

Dr.C.S.Kanimoli Selvi is a faculty member in the Department of Computer Science and Engineering of Kongu Engineering College, Perundurai, Erode, Tamil Nadu, India. She received her Bachelor's degree in Computer Science in1994 from Bharathiar University and a Master's degree in Computer Applications from Bharathiar University in 1998. Then, she received her Master's degree in Computer Science and Engineering from Anna University, Chennai in 2004. She has completed her Ph.D in 2011. She has been in the teaching profession for the past 15 years. Her areas of academic interest include data mining, database management systems and cloud computing. She has published 14 articles in international journals and more than 20 papers international and national conferences.