WORLD WIDE WEB KNOWLEDGE FREE DATA TRANSFER BY USING EXTENSION DOMAINS

Bhanu Nareesha Vanamali[1], Vesalapu Suneetha[2]

1. M.Tech Scholar, Department of Information Technology, Avanthi Institute of Engineering and Technology (A.I.E.T) Narisipatnam, Andhra Pradesh, India.
2. Assistant Professor, Department of Information Technology, Avanthi Institute of Engineering and Technology (A.I.E.T), Narsipatnam, Andhra Pradesh, India.

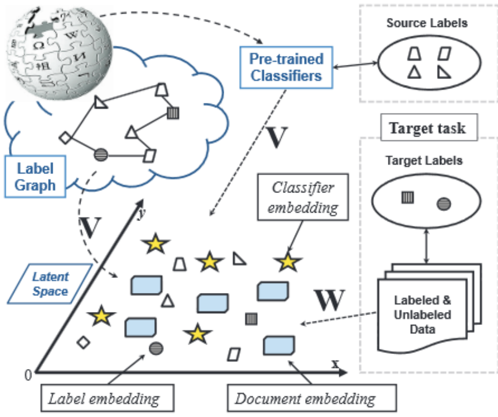
Abstract: The Problems are occurring due to the lack of ground-truth labeled knowledge. It is usually expensive to label new information instances for coaching a model. To solve this problem, domain adaptation in transfer learning has been proposed to classify target domain information by using another source domain information, even when the info may have completely different distributions. However, domain adaptation could not work well when the variations between the source and target domains are massive. In this paper, we have a tendency to design a unique transfer learning approach, referred to as BIG (Bridging Info Gap), to effectively extract useful data during a worldwide knowledge base, which is then used to link the supply and target domains for improving the classification performance. BIG works when the supply and target domains share the same feature space however totally different underlying knowledge distributions. Using the auxiliary source data, we tend to will extract bridge that permits cross-domain text classification problems to be solved using standard semi supervised learning algorithms. A major contribution of our work is that with BIG, a large amount of worldwide data will be simply custom-made and used for learning in the target domain. We conduct experiments on several real-world cross-domain text classification tasks and demonstrate that our proposed approach will outperform many existing domain adaptation approaches considerably.

Introduction

Transfer learning aims to improve the learning performance in a target domain using knowledge extracted from related source domains or tasks. What distinguishes transfer learning from other traditional learning is that either the source and target domains, or the target and source tasks, or both, are different. Transfer learning is particularly useful when we only have limited labeled data in a target domain, which requires that we consult one or more auxiliary tasks or domains to gain insight on how to solve the target problem [Pan and Yang, 2010] . Many transfer learning approaches have been developed over the years. For example, [Rainaet al., 2006;Daiet al., 2009] proposed to learn text classifiers by transferring knowledge from other text or image domains. [Panet al., 2010] and [Prettenhofer and Stein, 2011] proposed feature-based transfer learning methods for solving natural language processing tasks. [Dinget al., 2011] proposed to adopt a boosting based approach to select the most discriminative feature for knowledge transfer in target domain. In many typical transfer learning settings, a major assumption is that source data are provided by the problem designers. This places a big burden on the designer of the learning problem, since to improve the performance of learning, the “right” source data must be provided as well for effective transfer learning. However, it is very difficult to identify a proper set of source data. We often meet with the situation where we have a target task to solve, but we are at a loss at identifying from an extremely large number of choices of potential sources to use. For example, we may be given some text data to classify with limited labels, but we are only told to make use of the data on the World Wide Web! In such a situation, not only are we missing the source data, we also lack a scalable transfer learning method. This problem makes it difficult to benefit from many of the advantages of transfer learning. In this paper, we propose a novel framework tap into the online information sources without asking the user for a specific source data set for a given target classification problem. For simplicity, we assume that the online information source and the target task share the same feature space, but the label spaces may be different or even disjoint. Our problem can be informally stated as follows. We are given a target text classification problem with categorical classes YT. Besides the class label for the target classification problem, we optionally have a labeled training data that have a small number of class labels however, we assume that the labels are not sufficient to build an effective classification model using traditional machine learning algorithms. We also have an entire information source K available online, such as the Wikipedia which is used in this paper. K consists of a collection of labeled text data, while the label space YS in K may be different from YT in the target data. To help bridge between YS and YT, we also assume that we have a collection of social media that have been labeled via tags Ytag , such that Ytag. has overlap with YS and is a superset of YT . In other words,the target labels are covered by the tags in Ytag, which in turn have overlaps of those labels in the online information source. Our goal is to build a "bridge" and select a subset of the K as the source data to transfer the knowledge for the target task. To build an effective bridge between YS and YT, we propose to embed all labels into a latent Euclidean space using a graph representation. As a result, the relationship between labels can be represented by the distance between the corresponding prototypes of the labels in the latent space. Further more, we show that predictions made by each source classifier can also be mapped into the latent space, which makes the knowledge transfer from source classifiers possible. Finally, we apply a regularization framework to learn an effective classifier for the target text classification task. In this manner, our transfer learning framework does not depend on the specification of a precise source data set by the problem designer, and for this reason we call it "source-selection-free transfer learning" (SSFTL for short). There are several advantages associated with the SSFTL framework. First, since the online information source K is available ahead of the classifier training time for the target task. We can thus "compile" a large number of potential classifiers ahead of time for efficient classification, because they can be reused for different target tasks. Second, because we use a graph Laplacian to represent the label-translation process, the mapping between the target and online information source labels can be done very efficiently, resulting in a highly scalable architecture for transfer learning. Third, the class labels for the target learning task can vary from task to task, as long as they can be covered by the social media that serve as a bridge. This adds a lot to the flexibility of the learning process. Finally, our framework can be easily scaled up when the size of the online information source increases.

Source-Selection-Free Transfer Learning:

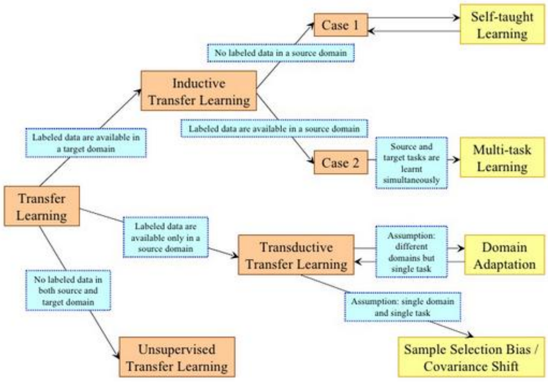
Transfer learning has been proposed to address the machine learning problems when there is an insufficient amount of labeled training data. Typical transfer learning approaches require one or more source datasets be given by the designers of the learning problem. However, how to identify the right source data to enable effective knowledge transfer has been an unsolved problem, which limits the applicability of many transfer learning approaches. In this paper, we propose a novel transfer learning approach that requires no specific source data be given; instead, for a given target learning task, the system can find the right subset to use as the auxiliary data from an extremely large collection on the world wide web. In our approach, which is known as source-free transfer learning (SFTL), we are given a target data set to classify, where the training data may have an insufficient amount of labeled data only. To build a classifier, SFTL turns to some very large knowledge sources such as the Wikipedia for help, by identifying a portion of the knowledge base as the potential source data. Since the labels provided by the world-wide knowledge may not match exactly with those in the target task, a translation must be done through a translator, to achieve knowledge transfer. This is done by consulting the social media which we compile into a graph Palladian based representation. One advantage of our approach is its source-free nature; learning task users no longer need to find the necessary source data to start learning. Another advantage is scalability; unlike many previous transfer learning approaches, which are difficult to scale up to the WWW scale, our approach is highly scalable and can offset much of the training work to offline stage. We demonstrate these advantages through extensive experiments on several real-world datasets, both in terms of efficiency of learning and classification performance. There are two challenges we need to address for the proposed problem, (1) since the label spaces of the auxiliary and target tasks may be different, a crucial research issue is how to build a bridge between these tasks via exploring relationships between the auxiliary and target labels, and (2) another challenge is how to make use of the pre-trained source classifiers to train a target classifier with the learned relationship



We address the above two challenges separately in our two step transfer learning framework as shown in Figure. In the first step, we construct a label graph by utilizing the social tagging data, and then learn a latent low-dimensional space of the labels via graph embedding. In the second step, we propose a principled regularization framework to transfer the knowledge encoded in the source classifiers

Source-Selection-Free Transfer Learning Algorithm:

Transfer learning addresses the problems that labeled training data are insufficient to produce a high-performance model. Typically, given a target learning task, most transfer learning approaches require to select one or more auxiliary tasks as sources by the designers. However, how to select the right source data to enable effective knowledge transfer automatically is still an unsolved problem, which limits the applicability of transfer learning. In this paper, we take one step ahead and propose a novel transfer learning framework, known as source-selection-free transfer learning (SSFTL), to free users from the need to select source domains. Instead of asking the users for source and target data pairs, as traditional transfer learning does, SSFTL turns to some online information sources such as World Wide Web or the Wikipedia for help. The source data for transfer learning can be hidden somewhere within this large online information source, but the users do not know where they are. Based on the online information sources, we train a large number of classifiers. Then, given a target task, a bridge is built for labels of the potential source candidates and the target domain data in SSFTL via some large online social media with tag cloud as a label translator. An added advantage of SSFTL is that, unlike many previous transfer learning approaches, which are difficult to scale up to the Web scale, SSFTL is highly scalable and can offset much of the training work to offline stage. We demonstrate the effectiveness and efficiency of SSFTL through extensive experiments on several real-world datasets in text classification.



In this section, we show that the structure of the social tagging data can be exploited to extract the relationship between the target and auxiliary labels. Since the label names are usually short and sparse, it is very hard for us to identify their correspondence based on some similarity measure using their word feature space alone. In order to uncover the intrinsic relationships between the target and source labels, we turn to some social media such as Delicious, which can help to bridge different label sets together. Delicious can be viewed as a Tag Cloud, where different users may use different tags to label one Web page. Each tag can be treated as a label, and the tag co-occurrence relationship carries rich label correspondence information. In order to exploit the underlying structure of the graph in the social media data, we apply the graph spectral techniques [Chung, 1997] on the graph to map each node in the graph to a low-dimensional latent space. In this way, each label will have a coordinate on this latent space, and we call it the prototype of this class. Because the dimension of such latent space can be much lower than the original word feature space, the mismatch problem caused by the label sparseness can be alleviated. Then the relationships between labels, e.g., similar or dissimilar, can be represented by the distance between their corresponding prototypes in the latent space, e.g., close to or far away from each other. Recall that Y is a union set of all source and target labels. For each label $y \in Y$, we can find its corresponding category in the social media data K. We can further extract a subgraph G that contains all target labels YT and the auxiliary labels YS from K. For each label y, we aim to recover its low-dimensional representation $v_y \in R^{m\times 1}$. In this paper, we propose to use LaplacianEigenmap [Belkin and Niyogi,2003] to recover the latent matrix $V= [v1, . . .vq]' \in R^{q\times m}$. Given a label graph G with its corresponding weight or neighborhood matrix $M \in R^{q\times q}$, LaplacianEigenmap aims to learn V by solving the following optimization problem,

$$\mathrm{Min}_v \mathrm{tr}(V'LV) \mathrm{s.t.} V'DV=I_m,$$

Where D is a diagonal matrix with $D_{ii}= \sum_j M_{ij}$, and $L=D-M$ is the Laplacianmatrix[Belkin and iyogi, 2003]. Im is an identity matrix of dimensionalitym. Note that if the distance between two label prototypes across the auxiliary and target domains in the latent space is small, then it implies that these two labels are semantically similar. Thus, based on the distance between label prototypes in the latent space, we are able to transfer knowledge across domains

Knowledge Transfer:

Based on the discovery of the relationships between labels, we propose a principled regularization framework for source selection-free transfer learning. In particular, with the matrix V estimated by using Laplacian graph embedding, for each label y∈Y, we have its M dimensional representation as vy. Therefore, for the target classification task which tries to learn a classifierg :x→y with y∈YT, we can transform it to a regression problem which aims at earning a regression model g:x→vy with vy=V′φ(y)∈Rm×1. In this paper, we assume is a linear model which can be written as g(x) =W′x with W∈Rd×m Recall that we are given a few target labeled data DℓT and some target unlabeled data DuT in the target domain. In transfer learning, the labeled data are too few to learn a prediction model. We thus show how to use the unlabeled data in our framework for transfer learning. We can make predictions on the unlabeled data DuT by using a combined mapping of all auxiliary classifiers as,

$$\mathbf{V}'\mathbf{F}_S^u = \mathbf{V}'\sum_{i=1}^{k}\varepsilon_i\mathbf{F}_{S_i}^u,$$

Where FuSi = [fSil+1, . . . ,fSin]∈Rq×(n−ℓ) is the predictions of auxiliary classifier fSi on DuT, and {εi}'s are weights for the source classifiers{fSi}'s. We will introduce an effective approach to estimate εi at the end of this section. The prototypes of the target labels may be enveloped and separated by those of the auxiliary labels in the latent space, which implies that the auxiliary classifiers may be helpful for the target egression problem. Therefore, the combined mapping of the auxiliary classifiers can be used to regularize the target classifier on the unlabeled target data as follows ΩDuT(W) =1n−ℓ‖W′Xu−V′FuS‖2F, where Xu∈Rd×(n−ℓ)is the unlabeled data matrix. Finally, we obtain the following optimization problem, min MINWΩDℓT(W) +λ1||W||2F+λ2ΩDuT(W) Where ΩD ℓT=1ℓW′Xℓ−V′Tφ(Yℓ))2Fis the loss function on the labeled data matrix Xℓ∈Rd×ℓ and φ (Yℓ)∈Rq×ℓ is the corresponding label matrix. Such loss function can be replaced with some other large margin losses proposed in [Quadriantoet al., 2010] or [Weinberger and Chapelle,2009] which can penalize misclassification errors. The additional regularization term ||W||2Fis used to avoid over fitting. Note that the knowledge from the auxiliary classifiers is encoded in ΩDuT(W), and the relationships between the target and auxiliary labels are encoded in V. We show that the model parameter W in can be solved analytically in the following proposition. Proposition 1 The optimization problem in Eq.(3) has an optimal solution in a closed form as

W= (A+λ1Id)−1XF′V,

Where A ⩾0is entirely independent of W,X= [XℓXu] ∈Rd×n and F = [φ(Yℓ)FuS]∈Rq×n Due to space limit, we omit the proof of the proposition. If there is no labeled data available in the target task, i.e., ℓ=0, SSFTL reduces to an unsupervised learning model which only minimizes the second and the third terms in Eq.(3). It can be proved that in this case, the model parameter W still has a closed form solution, which is similar to Eq.(4). So far, we have presented our SSFTL framework for transfer learning. We now introduce how to estimate the weights {εi}'s in Eq.(2). When there is no labeled data in he target domain, we can set εi= 1/k, which is a uniform weighting approach. However, if we have a few labeled data in the target domain, we can use the following simple yet effective approach to estimate {εi}'s. First, we can use each source classifier fSi to make predictions on the labeled target data DℓT by using the following rule,

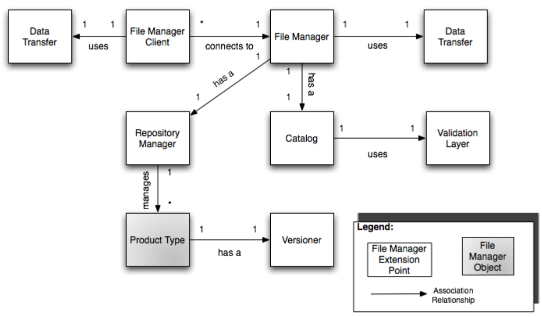
$$\mathrm{argmax}_y P(y|xlj) \propto -V' fSi(xlj) - V' \varphi(y)^2{}_2$$

We then calculate the classification accuracy hi of each fSi, setεi=hi and normalize εi, s.t.,∑iεi=1. Such weights can help select the most useful knowledge to transfer.

With the parameter matrix W, we can make prediction on any incoming test data x using the following rule

$$y^* = \arg\max_y P(y|\boldsymbol{x}) = \frac{e^{-||\mathbf{W}'\boldsymbol{x} - \boldsymbol{v}_y||_2^2}}{\sum_{y \in \mathcal{Y}_T} e^{-||\mathbf{W}'\boldsymbol{x} - \boldsymbol{v}_y||_2^2}},$$

Where the denominator is a normalization term, which ensures that ∀y∈Y T,0≤P(y|x)≤1 and ∑yP(y|x) = 1. In practice, in learning the parameter matrix W in Eq.(4), we can apply linear conjugate gradient for estimating each column of W independently, without computing the inverse of a d ×d matrix, which can be solved efficiently. Further-more, the complexity of computing FuSisO((n−l)dk). Since each source classifier is independent, we can further parallelize the computing process of FuS, which can become much more efficient. In making the prediction, the time complexity of SSFTL is O (md|YT|), which is independent of the number of auxiliary classifiers, and m is usually very small, e.g. m = 100 in our experiments, thus very efficient



An Experimental Protocol Data Transfer by Using Extension Domains:

As part of the initiative, we are experimenting with alternative protocols to help reduce the latency of web pages. One of these experiments is Specific Domine (pronounced "SPeeDY"), an application-layer protocol for transporting content over the web, designed specifically for minimal latency. In addition to a specification of the protocol, we have developed a SPECIFIC DOMINE -enabled Google Chrome browser and open-source web server. In lab tests, we have compared the performance of these applications over HTTP and Specific Domine, and have observed up to 64% reductions in page load times in Specific Domine. We hope to engage the open source community to contribute ideas, feedback, code, and test results, to make SPECIFIC DOMINE the next-generation application protocol for a faster web.

Background Web Protocols and Web Latency:

Today, HTTP and TCP are the protocols of the web. TCP is the generic, reliable transport protocol, providing guaranteed delivery, duplicate suppression, in-order delivery, flow control, congestion avoidance and other transport features. HTTP is the application level protocol providing basic request/response semantics. While we believe that there may be opportunities to improve latency at the transport layer, our initial investigations have focussed on the application layer, HTTP.

Unfortunately, HTTP was not particularly designed for latency. Furthermore, the web pages transmitted today are significantly different from web pages 10 years ago and demand improvements to HTTP that could not have been anticipated when HTTP was developed. The following are some of the features of HTTP that inhibit optimal performance:

- Single request per connection. Because HTTP can only fetch one resource at a time (HTTP pipelining helps, but still enforces only a FIFO queue), a server delay of 500 ms prevents reuse of the TCP channel for additional requests. Browsers work around this problem by using multiple connections. Since 2008, most browsers have finally moved from 2 connections per domain to 6.
- Exclusively client-initiated requests. In HTTP, only the client can initiate a request. Even if the server knows the client needs a resource, it has no mechanism to inform the client and must instead wait to receive a request for the resource from the client.
- Uncompressed request and response headers. Request headers today vary in size from ~200 bytes to over 2KB. As applications use more cookies and user agents expand features, typical header sizes of 700-800 bytes is common. For modems or ADSL connections, in which the uplink bandwidth is fairly low, this latency can be significant. Reducing the data in headers could directly improve the serialization latency to send requests.
- Redundant headers. In addition, several headers are repeatedly sent across requests on the same channel. However, headers such as the User-Agent, Host, and Accept* are generally static and do not need to be resent.
- Optional data compression. HTTP uses optional compression encodings for data. Content should always be sent in a compressed format.

Previous approaches:

SPECIFIC DOMINE is not the only research to make HTTP faster. There have been other proposed solutions to web latency, mostly at the level of the transport or session layer:

- Stream Control Transmission Protocol (SCTP) -- a transport-layer protocol to replace TCP, which provides multiplexed streams and stream-aware congestion control.
- HTTP over SCTP -- a proposal for running HTTP over SCTP. Comparison of HTTP Over SCTP and TCP in High Delay Networks describes a research study comparing the performance over both transport protocols.
- Structured Stream Transport (SST) -- a protocol which invents "structured streams": lightweight, independent streams to be carried over a common transport. It replaces TCP or runs on top of UDP.
- MUX and SMUX -- intermediate-layer protocols (in between the transport and application layers) that provide multiplexing of streams. They were proposed years ago at the same time as HTTP/1.1.

These proposals offer solutions to some of the web's latency problems, but not all. The problems inherent in HTTP (compression, prioritization, etc.) should still be fixed, regardless of the underlying transport protocol. In any case, in practical terms, changing the transport is very difficult to deploy. Instead, we believe that there is much low-hanging fruit to be gotten by addressing the shortcomings at the application layer. Such an approach requires minimal changes to existing infrastructure, and (we think) can yield significant performance gains.

Goals for Specific Domine :

The Specific Domine project defines and implements an application-layer protocol for the web which greatly reduces latency. The high-level goals for Specific Domine are:

- To target a 50% reduction in page load time. Our preliminary results have come close to this target (see below).
- To minimize deployment complexity. SPECIFIC DOMINE uses TCP as the underlying transport layer, so requires no changes to existing networking infrastructure.
- To avoid the need for any changes to content by website authors. The only changes required to support SPECIFIC DOMINE are in the client user agent and web server applications.
- To bring together like-minded parties interested in exploring protocols as a way of solving the latency problem. We hope to develop this new protocol in partnership with the open-source community and industry specialists.

Basic Features of Multiplexed Streams

IT allows for unlimited concurrent streams over a single TCP connection. Because requests are interleaved on a single channel, the efficiency of TCP is much higher: fewer network connections need to be made, and fewer, but more densely packed, packets are issued.

Request prioritization:

Although unlimited parallel streams solve the serialization problem, they introduce another one: if bandwidth on the channel is constrained, the client may block requests for fear of clogging the channel. To overcome this problem, SPECIFIC DOMINE implements request priorities: the client can request as many items as it wants from the server, and assign a priority to each request. This prevents the network channel from being congested with non-critical resources when a high priority request is pending.

Framework for Work:

The check pointing extension is laid out as follows:

(1) the name of the SMTP service extension defined here is checkpointing;

(2) the EHLO keyword value associated with the extension is CHECKPOINT;

(3) no parameter is used with the CHECKPOINT EHLO keyword;

(4) one optional parameter using the keyword TRANSID is added to the MAIL FROM command. The value associated with this parameter, coupled with the name of the client taken from EHLO command, forms a globally unique value that identifies this particular transaction and serves to distinguish it from all others. This value iscase-sensitive. The syntax of the value is as follows, using the ABNF notation of

Our initial results are promising, but we don't know how well they represent the real world. In addition, there are still areas in which SPECIFIC DOMINE could improve. In particular:

- Bandwidth efficiency is still low. Although dialup bandwidth efficiency rate is close to 90%, for high-speed connections efficiency is only about ~32%.
- SSL poses other latency and deployment challenges. Among these are: the additional RTTs for the SSL handshake; encryption; difficulty of caching for some proxies. We need to do more SSL tuning.
- Our packet loss results are not conclusive. Although much research on packet-loss has been done, we don't have enough data to build a realistic model model for packet loss on the Web. We need to gather this data to be able to provide more accurate packet loss simulations.
- SPECIFIC DOMINE single connection loss recovery sometimes underperforms multiple connections. That is, opening multiple connections is still faster than losing a single connection when the RTT is very high. We need to figure out when it is appropriate for the SPECIFIC DOMINE client to make a new connection or close an old connection and what effect this may have on servers.
- The server can implement more intelligence than we have built in so far. We need more research in the areas of server-initiated streams, obtaining client network information for prefetching suggestions.

Related Works:

Most previous works of transfer learning methods in text classification require the label spaces between the source and target tasks to be the same, and assume the difference between domains is only caused by the mismatch of data distributions [Pan and Yang, 2010]. In order to guarantee the scalability of large scale transfer learning, Duanet al., proposed a domain adaptation machine (DAM)[Duanet al.,2009] for transfer learning. Similar to SSFTL, in DAM, the knowledge carried by the source domains are encoded in the compact model parameters instead of the reuse of the raw data. Gaoet al. [2008] proposed a locally weighted ensemble framework (LWE) to combine multiple models for transfer learning. However, either DAM or LWE needs to assume the label spaces for the source and target tasks be the same, which cannot be applied to solve the label mismatch problem. More recently, the label mismatch problem in instance based transfer learning has attracted more and more attention. Some research works, such as risk-sensitive spectral partition (RSP) [Shiet al., 2009], EigenTransfer [Daietal., 2009] and multi-task learning with mutual information (MTL-MI) [Quadriantoet al., 2010], introduced some transfer learning methods for learning the label correspondence. However, their learning processes require maintaining all the training data from the auxiliary domain, which is ineffective for large scale setting if not impossible. As far as we know, SSFTL is the first work to address the above two challenges of heterogeneous label spaces and scalability due to large auxiliary domain. There are also some recent works on label embedding [Weinberger and Chapelle, 2009; Bengioet al., 2010] to discover a compressed space for large-scale multiple classes, such that a multi-class problem can be transformed to a regression problem. Our work is focused on exploring the relationships between the source and target labels to bridge two domains to enable knowledge transfer

Conclusions and Future Work:

In this paper, we proposed a novel transfer learning frame work, known as source selection-free transfer learning (SSFTL), to solve transfer learning problems when the potentialauxiliary data is embedded in very large online information sources. In our SSFTL framework, the label sets across domains can be different. We compile the label sets into a graph Laplacian for automatic label bridging, such that model designers no longer need to select task-specific source-domain data. SSFTL is highly scalable because the processing of the online information source can be done offline and reused for different tasks. Extensive experiments have been conducted to verify that SSFTL is efficient and effective for transfer learning. In the future, we will extend SSFTL along the following directions: (1) extend SSFTL to achieve knowledge transfer on heterogeneous feature spaces; (2) generalize SSFTL to truly “source-free” via transferring knowledge with different forms from the World Wide Web

References:

1. R. Raina, A. Battle, H. Lee, B. Packer, and A.Y. Ng, "Self-Taught Learning: Transfer Learning from Unlabeled Data," Proc. 24th Ann. Int'l Conf. Machine Learning (ICML '07), pp. 759-766, June 2007.
2. W. Dai, G.-R. Xue, Q. Yang, and Y. Yu, "Co-Clustering Based Classification for Out-of-Domain Documents," Proc. 13th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (KDD '07), pp. 210-219, Aug. 2007.
3. D.M. Blei, A.Y. Ng, and M.I. Jordan, "Latent Dirichlet Allocation," Proc. 14st Ann. Conf. Neural Information Processing Systems (NIPS '01), pp. 601-608, Dec. 2001.
4. Beitzelel., 2005 Steven M. Beitzel, Eric C. Jensen, OphirFrieder, David D. Lewis, AbdurChowdhury, and AleksanderKolcz. Improving automatic query classification via semiupervised learning. In ICDM, 2005.
5. Belkin and Niyogi, 2003 MikhailBelkin and ParthaNiyogi. Laplacianeigenmapsfor mensionality reduction and data representation. Neural Computation,15(6):1373–1396, 2003.
6. Bengioetal., 2010 SamyBengio, Jason Weston, andDavidGrangier. Label embedding trees for large multiclass tasks. InNIPS 23. 2010.
7. Chung, 1997 Fan R. K. Chung.Spectral Graph Theory.American Mathematical Society, 1997.
8. Daietal., 2009 Wenyuan Dai, Ou Jin, Gui-RongXue,Qiang Yang, and Yong Yu. Eigentransfer: a unified framework for transfer learning. In ICML, 2009.
9. Dingetal., 2011 Wei Ding, Tomasz F. Stepinski, YangMu, Lourenco Bandeira, Ricardo icardo, YouxiWu,Zhenyu Lu, Tianyu Cao, and Xindong Wu. Sub-kilometercrater discovery with boosting and transfer learning.ACMTIST, To appear 2011.
10. Duanetal., 2009 LixinDuan, Ivor W. Tsang, Dong Xu, and Tat-Seng Chua.

Domain adaptation from multiple sources via auxiliary classifiers. In ICML , 2009. \
11. Gaoetal., 2008 Jing Gao, Wei Fan, Jing Jiang, and Jiawei Han. Knowledge transfer via multiple model local structure mapping. InKDD, 2008.
12. Pan and Yang, 2010 SinnoJialin Pan and Qiang Yang. A survey on transfer learning.IEEE TKDE, 22(10):1345–1359, October 2010.
13. Panetal., 2010 SinnoJialin Pan, Xiaochuan Ni, Jian-TaoSun, Qiang Yang, and Zheng Chen. Cross-domain sentiment classification via spectral feauture alignment. In WWW, 2010.
14. Phanetal., 2008 XuanHieuPhan, Minh Le Nguyen, and Susumu Horiguchi. Learning to classify short and sparse text & web with hidden topics from large-scale data collections. In WWW, 2008.
15. Prettenhofer and Stein, 2011Peter Prettenhofer and Benno Stein. Cross-lingual adaptation using structural correspondence learning. ACM TIST , To appear 2011.
16. Quadriantoetal., 2010 Novi Quadrianto, Alex J. Smola,Tiberio S. Caetano, S.V.N. Vishwanathan, and James Petterson. Multitask learning without label correspondences. In NIPS 23, 2010.
17. Rainaetal., 2006 RajatRaina, Andrew Y. Ng, and Daphne Koller. Constructing informative priors using transfer learning. In ICML , 2006.
18. Shietal., 2009 Xiaoxiao Shi, Wei Fan, Qiang Yang, and JiangtaoRen. Relaxed transfer of different classes via spectral partition. In ECML/PKDD , 2009.
19. Weinberger and Chapelle, 2009 Kilian Q Weinberger and Olivier Chapelle. Large margin taxonomy embedding for document categorization. In NIPS 21. 2009

Biographies:

Bhanu Nareesha Vanamali pursuing her M.Tech in Department of Information Technology at Avanthi Institute of Engineering And Technology (AIET), Narsipatnam, Andhra Pradesh, India. Her research interest includes World Wide Web Knowledge Free Data Transfer by Using Extension Domains.

Vesalapu Suneetha, Assistant Professor, M.Tech in Department of Information Technology at Avanthi Institute of Engineering and Technology(A.I.E.T), Narsipatnam, Andhra Pradesh, India.. She Guided for the project and involved in all technical aspects.