Journal of Fundamental and Applied Sciences

ISSN 1112-9867

Available online at http://www.jfas.info

Research Article

IMPROVE A LEARNER: A SURVEY OF TRANSFER LEARNING

A. S. Bais

Northeastern University, 3-11 Heping District, Shenyang, China

Received: 04 March 2018 / Accepted: 24 April 2018 / Published online: 01 May 2018

ABSTRACT

Transfer learning is a subclass of machine learning, which uses training data (source), drawn from a diverse domain than that of the testing data (target). The real world is messy and contains an infinite range of novel eventualities. Transfer learning across different feature spaces is usually a tougher problem than Transfer Learning within the common feature space. This survey paper defines transfer learning to feature representation that maps the target domain to the source domains exploiting a set of meticulously manufactured features and applications to transfer learning. This paper systematically examines a feature based transfer techniques and reviews some current analysis on the topic of negative transfer.

Keywords: Data mining, survey, Transfer learning, Traditional machine learning.

Author Correspondence, e-mail: adityas.bais@gmail.com

doi: http://dx.doi.org/10.4314/jfas.v10i2.10

1. INTRODUCTION

Transfer learning has developed in the computer science writing as a method for transferring knowledge from a source domain to a target domain. Unlike traditional machine learning and semi-supervised algorithms, transfer learning consider that the domains of the training data and the test data can be different.

We can entice from practically non-technical experiences comprehend why transfer learning is probable, cogitate an example 4-people who want to learn how to play tennis, three people have experience how to play table tennis, while one person has no previous experience through the playing tennis. The person with sports background will be able to learn the tennis in a more efficient manner by transferring previously learned sport knowledge to the task learning to play the tennis. When the new task is haggard from a distinct population than the previous, this is often considered to transfer learning.

In transfer learning, we've got the following three main research issues: What to transfer, how to transfer, and when to transfer. “What to transfer” queries, which amount of knowledge can be transferred through domains or tasks. Some knowledge is specifically for individual domains or tasks, and a few data could also be common between different domains such, they'll facilitate improved performance for the target domain or task. Once discovering that data is transferred, learning algorithms must be developed to transfer the data, that corresponds to the “how to transfer” issue.

“When to transfer” asks during which things, transferring skills ought to be performed. Also, we have a tendency to have an interest in knowing during which things, and data mustn't be transferred. In some things, once the supply domain and a target domain aren't associated with one another, the brute-force transfer could also be unsuccessful. In the worst case, it should even hurt the performance of learning within the target domain, a scenario that is commonly remarked as negative transfer. However, a way to avoid negative transfer is a crucial open issue that's attracting a lot of and a lot of attention in the future.

This survey article, we give an overview of the field of the transfer learning as it pertains to data mining tasks for classification, regression and clustering developed in machine learning. This survey paper provides transfer learning feature representation followed by heterogeneous transfer learning and homogeneous transfer learning.

The rest of the survey paper is structured as follows. In the next section, first give the background and related work followed by Transfer learning technique and their different setting in a given table in details. After that in Section 3 feature representation transfers, we review some research on the Cross-Domain Knowledge transfer, Homogeneous transfer learning, Heterogeneous transfer learning and negative transfer. For each setting, we review different approaches that are given in tables in details. In the section4 provides some examples of transfer learning application. Finally, Conclusion and discussion are summarized in section 5.

2. BACKGROUND AND RELATED WORK

Research on transfer learning has fascinated more and more since 1995 in different titles: learning to learn, lifelong learning, knowledge transfer, inductive knowledge-based inductive bias, Meta-learning, and incremental/cumulative learning. Table 1 shows a simple comparison of the machine learning technologies from a different perspective.

Table 1. Comparisons of the machine learning technologies

Learning types	Data processing tasks	Distinction norm	Learning algorithms	References
Supervised learning	Classification/ Regression/ Estimation	Computational classifiers	Support vector machine	[1]
		Statistical classifiers	Naïve Bayes	[2]
			Hidden Markov model	
			Bayesian networks	[3]
		Connectionist classifiers	Neural networks	[4]
Unsupervised learning	Clustering /Prediction	Parametric	K-means	[5]
			Gaussian mixture model	[6]
		Nonparametric	Dirichlet process mixture model	[6]
			X-means	[5]
Transfer learning	Decision-mechanism /mapping	Model based	Regression, classification	
Reinforcement learning	Decision-making	Model-free	Q-learning	
			R-learning	
		Model-based	TD learning	
			Sarsa learning	

2.1. Transfer learning technique

In transfer learning, both the training data and the testing data can contribute to two types of domains:

1) Target domain and

2) Source domain.

The target domain contains the testing instances, that are the task of the categorization system, and the source domain contains training instances, which are underneath a unique distribution with the target domain data. Transfer learning allows for the training and test data have different domains and tasks.

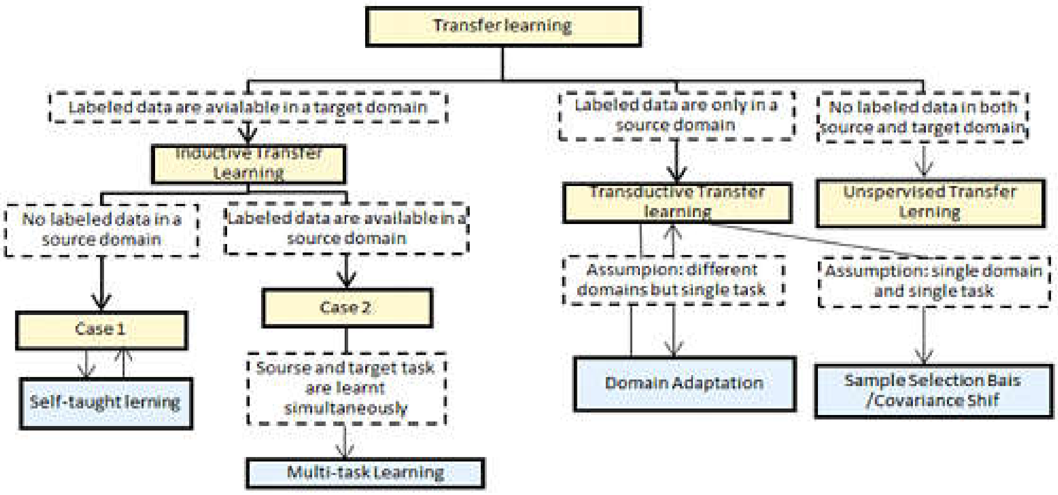

Fig 1. An overview of different settings of transfer learning

3. FEATURE REPRESENTATION TRANSFER

Feature representation level knowledge transfer is a fashionable transfer learning class that maps the target domain to the source domains exploiting a set of meticulously manufactured features. Table 2 shows the most characteristics of the feature representation level knowledge transfer approaches that obtainable. Thus, we have induce a tendency to review the feature level knowledge transfer techniques according to two knowledge types. Later we discuss two categories of transfer learning, that observed as homogeneous transfer learning and heterogeneous transfer learning.

Cross-Domain Knowledge transfer: In the configuration of cross domain learning to rank drawback, two elements of labeled data are obtainable. In different words, this can be a supervised transfer learning problem. Wherever the source domain and target domain data share a typical document feature space $X \in R^d$; whereas the data distributions for them are different.

Cross-View Knowledge Transfer: Cross-view knowledge transfer is consider as a peculiar instance of cross-domain knowledge transfer, wherever the divergences across domains are caused by viewpoint changes. The task is to recognize action categories within the target view victimization training samples from one or a lot of different views. Generating view-invariant features to deal with the cross-view visual pattern recognition issues attract vital attention within the computer vision field, particularly for cross-view action recognition.

Table 2. Feature representation level knowledge transfer approaches

Adaptation methods	Target domain Label	Source domain labe	Adaptation type	Application	Ref.
Sparse Approximation	A small set available	Unavailable	Cross-domain	Image classification	[7]
Kernel Based	Unavailable/a small set	Available	Cross-domain	Object recognition	
Sparse Coding	Available	Unavailable	Cross-domain	Digit recognition	[8]
Gaussian Mixture Model	Unavailable	Available	Cross-domain	Action detection	[9]
Geometric Reasoning	Available	Available	Cross-view	Gesture detection, tracking, action recognition	
Similarity Mining	Unavailable	Available	Cross-domain	Activity recognition	
Visual Pattern Properties	Available	Available	Cross-view	Action recognition	
Self-Similarity Matrix	Available	Available	Cross-view	Action recognition	
Bipartite Graph	Unavailable	Available	Cross-view	Action recognition	[10]
Linear Transformation	Unavailable/a small set	Available	Cross-view	Action recognition	
Dictionary Learning	Unavailable/a small set	Available	Cross-view	Action recognition	[11]
Quantize Aspect	Unavailable	Available	Cross-view	Action recognition	

Homogeneous transfer learning: Homogeneous transfer learning is that the case where $\chi_S = \chi_T$. The methodology of homogeneous transfer learning is directly applied to a big data environment. An example of homogeneous multi-source transfer learning task is a picture classification in computer vision area wherever the acquisition of the labeled pictures of the target domain is pricey. Table 3 shows homogeneous transfer learning approaches surveyed in homogeneous transfer with data available in the source and target.

Table 3. Section listing different characteristics of each approach in Homogeneous transfer learning

Approach	**Transfer category**	**Source data**	**Target data**	**Multiple sources**	**Generic solution**	**Negative transfer**	**Ref**
CP-MDA	Parameter	Labeled	Limited labels	✓	✓		[12]
2SW-MDA	Asymmetric feature	Labeled	Unlabeled	✓	✓		[12]
FAM	Asymmetric feature	Labeled	Limited labels	✓	✓		[13]
DTMKL	Asymmetric feature	Labeled	Unlabeled		✓		[14]
JDA	Asymmetric feature	Labeled	Unlabeled		✓		[15]
ARTL	Asymmetric feature	Labeled	Unlabeled		✓		[16]
TCA	Symmetric feature	Labeled	Unlabeled		✓		[33]
SFA	Symmetric feature	Labeled	Limited labels	✓	✓		[34]
SDA	Symmetric feature	Labeled	Unlabeled		✓		[17]
GFK	Symmetric feature	Labeled	Unlabeled		✓	✓	[18]
DCP	Symmetric feature	Labeled	Unlabeled		✓		[19]
TCNN	Symmetric feature	Labeled	Limited labels		✓		[20]
MMKT	Parameter	Labeled	Limited labels	✓	✓	✓	[35]
DSM	Parameter	Labeled	Unlabeled	✓	✓	✓	[21]
MsTrAdaBoost	Instance	Labeled	Limited labels	✓	✓	✓	[22]
TaskTrAdaBoost	Parameter	Labeled	Limited labels	✓	✓	✓	[22]
RAP	Relational	Labeled	Unlabeled				[23]
SSFE	Hybrid (instance and feature)	Labeled	Limited labels				[24]

Heterogeneous transfer learning: In this instance, $X_s \neq X_t$ and $Y_s \neq Y_t$ as the source and target domains can share no features and labels, whereas the dimensions of the feature spaces may differ as well. This method, therefore desires a feature and label space transformations to bridge the gap for knowledge transfer, in addition, as handling the cross-domain data distribution variations. Heterogeneous transfer learning applications that are covered include image recognition [25, 26, 27, 28, 31, and 40], multi-language text classification [25, 29, 31, 36, and 39] single language text classification [32], drug efficacy classification [27], human activity classification [37], and software defect classification [40]. Table 4 shows a surveyed in Heterogeneous transfer learning division list several features of each approach.

Table4. A surveyed in Heterogeneous transfer learning section of each approach

Approach	Transfer category	Source data	Target data	Multiple sources	Generic solution	Negative transfer	Ref
CLSCL	Symmetric feature	Labeled	Unlabeled				[36]
HeMap	Symmetric feature	Labeled	Limited labels		✓		[27]
DAMA	Symmetric feature	Labeled	Limited labels	✓	✓		[32]
HTLIC	Symmetric feature	Unlabeled	Abundant labels				[40]
TTI	Symmetric feature	Labeled	Limited labels				[28]
HFA	Symmetric feature	Labeled	Limited labels		✓		[31]
SHFA	Symmetric feature	Labeled	Limited labels		✓		[25]
ARC-t	Asymmetric feature	Labeled	Limited labels		✓		[26]
MOMAP	Asymmetric feature	Labeled	Limited labels	✓			[37]
SHFR	Asymmetric feature	Labeled	Limited labels		✓		[29]
HDP	Asymmetric feature	Labeled	Unlabeled		✓		[38]
HHTL	Asymmetric feature	Labeled	Unlabeled		✓		[39]

Negative transfer

Despite the fact that the way to avoid negative transfer is a vital issue. Some works have been exploited to research relatedness among tasks and task clustering techniques, which may facilitate giving guidance on a way to avoid negative transfer automatically.

Cases of negative transfer area unit incontestable by Rosenstein in experiments employing a hierarchical Naive Bayes classifier. The author additionally demonstrates the prospect of negative transfer goes down because the variety of labeled target training samples goes up. The paper by Eaton [30] proposes to create a target learner supported a transferability live from multiple related source domains. The strategy by Eaton is tested in conjunction with a wherever the source domains area unit manually selected to be associated with the target, the main method that uses all sources accessible, and a baseline technique that doesn't use transfer learning. Classification accuracy is that the performance metric measured within the input space. Overall, the Eaton [30] approach performs the most effective. The subject of negative transfer could be a fertile space for additional analysis.

4. TRANSFER LEARNING APPLICATIONS

The transfer learning has been applied to many real-world applications. There are a number of application examples concerning adapts to new domains. For text connected analysis, i.e., exploring the helpful information from a given document, the learning drawback of the way to label the text documents across totally different distributions is addressed. Reliable cross lingual adaptation strategies would enable us to leverage the large amounts of labeled data we have in English and apply them to any language, significantly underserved and actually low-resource language.

Within the literature, there are varieties of transfer learning solutions that are specific to the applying of advice systems. Recommendation systems give users with recommendations or ratings for a selected domain (e.g. Shopping for merchandise, comments on the merchandise, historical places, etc.), that is supported historical information. To better predict economic condition mapping, Xie[41] proposes associate approach kind of like Oquab[42] that uses a convolution neural network model. The primary prediction model is trained to predict getting dark intensity level from the source image data. The ultimate target prediction model predicts the economic condition mapping from the source getting dark intensity level data. Transfer learning is applies to the sector of biology as well as medical issues, biological modeling designs, ecological problems, and so on. In applications associated with medical problems, Caruana recommended victimization multi-task learning in artificial neural networks, associated projected an inductive transfer learning approach for respiratory disorder risk prediction. Another application area of transfer learning is finance, like within the area of automobile insurance risk estimations and monetary early warning systems. Niculescu-Mizil associated Caruana conferred an inductive transfer learning approach that collectively learns multiple Bayesian network structures rather than adaptation probabilistic networks from multiple connected data sets. Transfer learning victimization procedure intelligence has been applied in business management. For example, Roy associated Kaelbling projected an economical Bayesian task-level transfer learning to tackle the user's behavior within the meeting domain. Learning from simulations has the advantage of creating data gathering simple as objects are simply delimited and analyzed, whereas at the same time sanctionative quick coaching, as learning is parallelized across multiple instances.

5. CONCLUSION

In this survey paper, we have reviewed the several current trends of transfer learning techniques on visual categorization tasks. Many real-world applications that transfer learning is applied to are discussed in this survey paper. In addition, how to avoid negative transfer learning is an open problem in classical transfer learning, but also in computational intelligence-based transfer learning. In this survey homogeneous transfer learning is measured, surveyed and additionally present solution for the topic of negative transfer. Homogeneous transfer learning solutions varied necessities and unlabeled data that also are conferred as a key attribute. Numerous of real-life applications that transfer learning is functional to are enumerated and state. In some cases, the proposed transfer learning solutions are found a specific to the underlying application and can't be generically used for an alternative application.

6. REFERENCES

1. J Zheng, F Shen, H Fan, J Zhao, An online incremental learning support vector machine for large-scale data. Neural Computer Appl 22 (5), 1023–1035 (2013).
2. P. Rashidi, D.J. Cook, L.B. Holder, and M. Schmitter-Edgecombe. Discovering activities to recognize and track in a smart environment. IEEE Transactions on Knowledge and Data Engineering, 23(4):527 –539, April 2011.
3. K Yue, Q Fang, X Wang, J Li, W Weiy, A parallel and incremental approach for data-intensive learning of Bayesian networks. IEEE Trans Cybern 99, 1–15 (2015).

4. X Dong, Y Li, C Wu, Y Cai, A learner based on neural network for cognitive radio, in Proceedings of the 12th IEEE International Conference on Communication Technology (ICCT) (Nanjing, 2010), pp. 893–896.

5. AM Bkassiny, SK Jayaweera, Y Li, Multidimensional dirichlet process-based non-parametric signal classification for autonomous self-learning cognitive radios. IEEE Trans Wirel Commun 12(11), 5413–5423 (2013).

6. M Bkassiny, SK Jayaweera, Y Li, Multidimensional dirichlet process-based non-parametric signal classification for autonomous self-learning cognitive radios. IEEE Trans Wirel Commun 12(11), 5413 5423 (2013).

7. R. Li and T. Zickler, “Discriminative virtual views for cross-view action recognition,” in Proc. 25th IEEE Conf. Comput. Vis. Pattern Recognit. Providence, RI, USA, Jun. 2012, pp. 2855–2862.

8. R. Raina, A. Battle, H. Lee, B. Packer, and A. Y. Ng, “Self-taught learning: Transfer learning from unlabeled data,” in Proc. 24th Int. Conf. Mach. Learn.,.

9. L. Cao, Z. Liu, and T. S. Huang, “Cross-dataset action detection,” in Proc. 23rd IEEE Conf. Comput. Vis. Pattern Recognit., San Francisco, CA, USA, Jun. 2010, pp. 1998–2005.

10. J. Liu, M. Shah, B. Kuipers, and S. Savarese, “Cross-view action recognition via view knowledge transfer,” in Proc. 24th IEEE Conf. Comput. Vis. Pattern Recognit., Colorado Springs, CO, USA, Jun. 2011, pp. 3209–3216.

11. F. Zhu and L. Shao, “Correspondence-free dictionary learning for cross view action recognition,” in Proc. 22nd Int. Conf. Pattern Recognit., Stockholm, Sweden, Aug. 2014.

12. Chattopadhyay R, Ye J, Panchanathan S, Fan W, Davidson I. Multi-source domain adaptation and its application to early detection of fatigue. ACM Trans Knowl Dis Data (Best of SIGKDD 2011 TKDD Homepage archive) 2011; 6(4) (Article 18).

13. Daumé H III. Frustratingly easy domain adaptation. In: Proceedings of ACL. 2007. p. 256–63

14. Duan L, Tsang IW, Xu D. Domain transfer multiple kernel learning. IEEE Trans Pattern Anal Mach Intell. 2012;34(3):465–79

15. Long M, Wang J, Ding G, Sun J, Yu PS. Transfer feature learning with joint distribution adaptation. In: Proceedings of the 2013 IEEE international conference on computer vision. 2013. p. 2200–07

16. Long M, Wang J, Ding G, Pan SJ, Yu PS. Adaptation regularization: a general framework for transfer learning. IEEE Trans Knowl Data Eng. 2014;26(5):1076–89.

17. Glorot X, Bordes A, Bengio Y. Domain adaptation for large-scale sentiment classification: A deep learning approach. In: Proceedings of the twenty-eight international conference on machine learning, vol. 27. 2011. p.97–110.

18. Gong B, Shi Y, Sha F, Grauman K. Geodesic flow kernel for unsupervised domain adaptation. In: Proceedings of the 2012 IEEE conference on computer vision and pattern recognition. 2012. p. 2066–73.

19. Shi Y, Sha F. Information-theoretical learning of discriminative clusters for unsupervised domain adaptation. In: Proceedings of the 29th international conference on machine learning. 2012. p. 1–8.

20 Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the 2014 IEEE conference on computer vision and pattern recognition. 2013. p. 1717–24.

21. Duan L, Xu D, Chang SF. Exploiting web images for event recognition in consumer videos: a multiple source domain adaptation approach. In: IEEE 2012 conference on computer vision and pattern recognition. 2012. p. 1338–45.

22. Yao Y, Doretto G. Boosting for transfer learning with multiple sources. In: Proceedings of the IEEE computer society conference on computer vision and pattern recognition. 2010. p. 1855–62.

23. Li F, Pan SJ, Jin O, Yang Q, Zhu X. Cross-domain co-extraction of sentiment and topic lexicons. In: Proceedings of the 50th annual meeting of the association for computational linguistics long papers, vol. 1. 2012. p. 410–19.

24. Xia R, Zong C, Hu X, Cambria E. Feature ensemble plus sample selection: domain adaptation for sentiment classification. IEEE Intell Syst. 2013;28(3):10–8.

25. Li W, Duan L, Xu D, Tsang IW. Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation. IEEE Trans Pattern Anal Mach Intell. 2014;36(6):1134–48.

26. Kulis B, Saenko K, Darrell T. What you saw is not what you get: domain adaptation using asymmetric kernel transforms. In: IEEE 2011 conference on computer vision and pattern recognition. 2011. p. 1785–92.

27. Shi X, Liu Q, Fan W, Yu PS, Zhu R. Transfer learning on heterogeneous feature spaces via spectral transformation. In: 2010 IEEE international conference on data mining. 2010. p. 1049–1054.

28. Qi GJ, Aggarwal C, Huang T. Towards semantic knowledge propagation from text corpus to Web images. In: Proceedings of the 20th international conference on world wide web. p. 297–306.

29. Zhou JT, Tsang IW, Pan SJ Tan M. Heterogeneous domain adaptation for multiple classes. In: International conference on artificial intelligence and statistics. 2014. p. 1095–103.

30. Eaton E, des Jardins M, Lane T. Modeling transfer relationships between learning tasks for improved inductive transfer. Proc Mach Learn Knowl Disc Database. 2008;5211:317–32..

31. Duan L, Xu D, Tsang IW. Learning with augmented features for heterogeneous domain adaptation. IEEE Trans Pattern Anal Mach Intell. 2012;36(6):1134–48.

32. Wang C, Mahadevan S. Heterogeneous domain adaptation using manifold alignment. In: Proceedings of the 22nd international joint conference on artificial intelligence, vol. 2. 2011. p. 541–46.

33. Pan SJ, Tsang IW, Kwok JT, Yang Q. Domain adaptation via transfer component analysis. IEEE Trans Neural Netw.2009;22(2):199–210.

34. Pan SJ, Ni X, Sun JT, Yang Q, Chen Z. Cross-domain sentiment classification via spectral feature alignment. In: Proceedings of the 19th international conference on world wide web. 2010. p. 751–60.

35. Tommasi T, Orabona F, Caputo B. Safety in numbers: learning categories from few examples with multi model knowledge transfer. IEEE Conf Comput Vision Pattern Recog. 2010;2010:3081–8.

36. Prettenhofer P, Stein B. (2010) Cross-language text classification using structural correspondence learning. In: Proceedings of the 48th annual meeting of the association for computational linguistics. 2010. p. 1118–27.

37. Harel M, Mannor S. Learning from multiple outlooks. In: Proceedings of the 28th international conference on machine learning. 2011. p. 401–8.

38. Nam J, Kim S (2015) Heterogeneous defect prediction. In: Proceedings of the 2015 10th joint meeting on foundations of software engineering. 2015. p. 508–19.

39. Zhou JT, Pan S, Tsang IW, Yan Y. Hybrid heterogeneous transfer learning through deep learning. In: Proceedings of the national conference on artificial intelligence, vol. 3. 2014. p. 2213–20.

40. Zhu Y, Chen Y, Lu Z, Pan S, Xue G, Yu Y, Yang Q. Heterogeneous transfer learning for image classification. In: Proceedings of the national conference on artificial intelligence, vol. 2. 2011. p. 1304–9.

41. Xie M, Jean N, Burke M, Lobell D, Ermon S. Transfer learning from deep features for remote sensing and poverty mapping. In: Proceedings 30th AAAI conference on artificial intelligence. 2015. p. 1–10.

42. Oquab M, Bottou L, Laptev I, Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In: Proceedings of the 2014 IEEE conference on computer vision and pattern recognition. 2013. p. 1717–24.

How to cite this article:
Bais AS. Improve a learner: a survey of transfer learning. J. Fundam. Appl. Sci., 2018, *10(2), 136-147.*