Active Learning for Domain Adaptation

V. Sreeramdass*, V. Piratla, S. Sarawagi, S. Chakrabarti
Department of Computer Science and Engineering
Indian Institute of Technology Bombay
Powai, India
varshiths@cse.iitb.ac.in

Abstract

This work surveys and categorizes methods for adapting models to domains that are different from the source in an active learning setting.

1 Introduction

Domain Adaptation is a crucial task in modern machine learning. While models trained to solve general tasks are readily available, for the sake of real life deployments, the performance of the model in a narrow domain that is relevant to the application is more important than the generality of the algorithm. Domain Adaptation takes multiple forms. While some aim to build domain invariant representations, others learn correspondences between examples from the source and the target domains.

Active Learning is a method that attempts to learn a particular task by actively choosing data points for which labels are queried. This is to be data efficient by ensuring that only those examples' labels are queried from which the model can learn something it did not know before.

In the effort to adapt models to new domains data efficiently, active learning can be quite beneficial. It can also help in understanding what aspects domains differ in by observing the distribution of query signals over the set of candidate examples.

2 Classical Active Learning Methods

The core component of Classical Active Learning algorithms is to design acquisition functions that determine whether a particular example is to be considered as a candidate for which the label is to be queried. The four factors that are crucial when picking a candidate for labeling are:

- Representative of unseen data: The example picked should be characteristic of what part of the distribution of the target domain the model is yet to understand.
- Redundancy: If the model is fairly confident about the candidate's label, the particular example could be redundant to the learning process.
- Trade off between learning to be confident vs learning a wider distribution.
- Mis-calibration: Deep neural networks could assign confidences that are not reflective of the distribution learned by the model and so confidence measures could be unreliable. A back-off strategy could be important.

A few designs of acquisition functions for classifiers are laid out in Gal et al. (2017). Besides from baseline indicators like Entropy, Max-Confidence, Variation Ratios, Gal et al. (2017) advocates the use of dropout layers and broadly ensembles to model uncertainty towards labeling a particular sample. One particular signal is the number of disagreements between various predictions of the models in the ensemble, or in this case, for various random dropout values. This serves as a robust measure of certainty of the model towards the current example and indicate its candidacy for label acquisition.

Xiao & Guo (2013) proposed the use of multi-view methods to train different models on different views of the input. These models trained on multiple views serve as an ensemble which again serve as a source of uncertainty estimates.

Because uncertainty measures do not capture distributional information of the target domain, Li & Guo (2013) augments entropy measure with an information density measure which is the mutual information between the candidate and the remaining set of unlabeled examples in a Gaussian Process framework. The information density measure takes the form:

$$d(x_i) = I(x_i, X_{U_i}) = H(x_i) - H(x_i|X_{U_i})$$

where H is entropy, x_i the candidate, X_{U_i} the remaining set of unlabeled examples excluding x_i.

Among other techniques like query-by-committee, Li et al. (2013) uses label propagation by training a model on a small subset of the target domain labeled set and generating high confidence labels for unlabeled data.

Rai et al. (2010) trained a classifier/separator to assign domain to a particular example. Examples which the separator assigns high confidence to for not belonging to the source domain are queried.

Settles et al. (2008) proposed the use of Expected Gradient Length (EGL). EGL is a measure of sensitivity of the network to one particular example; the expectation of the length of the gradient update for a particular label, over the predicted label distribution.

3 Active Learning for Sequences

While many of the acquisition functions are designed for classifiers, they can be extended to sequence labeling tasks. Settles & Craven (2008) analyses the use of the following methods:

- Total Token Entropy: entropy at each position of the input sentence averaged
- EGL: Expected Gradient Length but the set of weights are limited to embeddings
- Information Density Measure with various sequence kernels defined over state traces

Zhang et al. (2017) applies EGL, ensemble methods on sentence and document classification tasks and achieves state of the art while Shen et al. (2017) uses normalized likelihood, ensemble methods to achieve the same for named entity recognition.

4 Learnable Active Learning

While most of the work in the field focused on designing acquisition functions, Fang et al. (2017) pioneered in the realm of learning an acquisition function through a reinforcement learning approach. In an episode, the policy network which is the acquisition function, goes over all the examples in the unlabeled set and decides which examples to include one by one with the model being trained after each step. The reward associated with this process is the increase in the performance of the classifier at the exhaustion of the budget. The approach is quite straightforward but computationally heavy.

Liu et al. (2018a) designed a similar approach where the classifier is trained after a batch of examples is included in the labeled set. However, the policy network here is learned by imitation learning by a strategic brute force search for the most rewarding batch among the set of unlabeled examples. Liu et al. (2018b) applied this approach to learn to actively learn machine translation.

5 Conclusion

This survey attempts to broadly categorize the work on Active Learning and study the variations in methods and their applicability to various tasks.

References

Omar Abdelwahab and Adel Elmaghraby. Deep learning based vs. markov chain based text generation for cross domain adaptation for sentiment classification. In *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, pp. 252–255. IEEE, 2018.

Meng Fang, Yuan Li, and Trevor Cohn. Learning how to active learn: A deep reinforcement learning approach. *arXiv preprint arXiv:1708.02383*, 2017.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 1126–1135. JMLR. org, 2017.

Yarin Gal, Riashat Islam, and Zoubin Ghahramani. Deep bayesian active learning with image data. In *Proceedings of the 34th International Conference on Machine Learning-Volume 70*, pp. 1183–1192. JMLR. org, 2017.

Behnam Gholami, Pritish Sahu, Ognjen Rudovic, Konstantinos Bousmalis, and Vladimir Pavlovic. Unsupervised multi-target domain adaptation: An information theoretic approach. *arXiv preprint arXiv:1810.11547*, 2018.

Da Li, Yongxin Yang, Yi-Zhe Song, and Timothy M Hospedales. Learning to generalize: Meta-learning for domain generalization. In *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

Shoushan Li, Yunxia Xue, Zhongqing Wang, and Guodong Zhou. Active learning for cross-domain sentiment classification. In *Twenty-Third International Joint Conference on Artificial Intelligence*, 2013.

Xin Li and Yuhong Guo. Adaptive active learning for image classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 859–866, 2013.

Ming Liu, Wray Buntine, and Gholamreza Haffari. Learning how to actively learn: A deep imitation learning approach. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pp. 1874–1883, 2018a.

Ming Liu, Wray Buntine, and Gholamreza Haffari. Learning to actively learn neural machine translation. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pp. 334–344, 2018b.

Piyush Rai, Avishek Saha, Hal Daumé III, and Suresh Venkatasubramanian. Domain adaptation meets active learning. In *Proceedings of the NAACL HLT 2010 Workshop on Active Learning for Natural Language Processing*, pp. 27–32. Association for Computational Linguistics, 2010.

Avishek Saha, Piyush Rai, Hal Daumé, Suresh Venkatasubramanian, and Scott L DuVall. Active supervised domain adaptation. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, pp. 97–112. Springer, 2011.

Burr Settles and Mark Craven. An analysis of active learning strategies for sequence labeling tasks. In *Proceedings of the conference on empirical methods in natural language processing*, pp. 1070–1079. Association for Computational Linguistics, 2008.

Burr Settles, Mark Craven, and Soumya Ray. Multiple-instance active learning. In *Advances in neural information processing systems*, pp. 1289–1296, 2008.

Yanyao Shen, Hyokun Yun, Zachary C Lipton, Yakov Kronrod, and Animashree Anandkumar. Deep active learning for named entity recognition. *arXiv preprint arXiv:1707.05928*, 2017.

Eric Tzeng, Judy Hoffman, Kate Saenko, and Trevor Darrell. Adversarial discriminative domain adaptation, 2017.

Min Xiao and Yuhong Guo. Online active learning for cost sensitive domain adaptation. In *Proceedings of the Seventeenth Conference on Computational Natural Language Learning*, pp. 1–9, 2013.

Ye Zhang, Matthew Lease, and Byron C Wallace. Active discriminative text representation learning. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.

Guangyou Zhou, Zhiwen Xie, Jimmy Xiangji Huang, and Tingting He. Bi-transferring deep neural networks for domain adaptation. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, volume 1, pp. 322–332, 2016a.

Joey Tianyi Zhou, Sinno Jialin Pan, Ivor W Tsang, and Shen-Shyang Ho. Transfer learning for cross-language text categorization through active correspondences construction. In *Thirtieth AAAI Conference on Artificial Intelligence*, 2016b.