Learning From Multiple Domains*
(Extended Abstract)

Zahra Taghiyarrenani, Sławomir Nowaczyk, Sepideh Pashami and Mohamed-Rafik Bouguelia

Abstract— **Domain adaptation (DA) transfers knowledge between domains by adapting them. The most well-known DA scenario in the literature is adapting two source and target domains using the available labeled source samples to construct a model generalizable to the target domain. Although the primary purpose of DA is to compensate for the target domain's labeled data shortage, the concept of adaptation can be utilized to solve other problems [1].**

One issue that may occur during adaptation is the problem of class misalignment, which would result in a negative transfer. Therefore, preventing negative transfer should be considered while designing DA methods. In addition, the sample availability in domains is another matter that should also be taken into account.

Considering the two mentioned matters, we aim to develop DA techniques to solve primarily predictive maintenance problems.

We consider a spectrum of cases with different amounts of available target data. One endpoint is the case in which we have access to enough labeled target samples for all classes. In this case, we use the concept of DA for 1) Analyzing two different physical properties, i.e., vibration and current, to measure their robustness for fault identification [2] and 2) Developing a denoising method to construct a robust model for a noisy test environment [3].

Next, we consider the case where we have access to unlabeled and a few labeled target samples. We aim to prevent negative transfer using the few labeled samples available while adapting source and target domains. We construct a unified feature representation using a few-shot and an adaptation learning technique to achieve this [4].

In the next considered setting, we assume we only have access to very few labeled target samples, which are insufficient to train a domain-specific model. Furthermore, for the first time in the literature, we solve the DA for regression in a setting in which it adapts multiple domains with any arbitrary shift [5].

Sometimes, due to the dynamic nature of the environment, we need to update a model to reflect the changes continuously. An example is in the field of computer network security. There is always the possibility of intrusion into a computer network, which makes each Intrusion Detection System (IDS) subject to concept shifts. In addition, different types of intrusions may occur in different networks. This thesis presents a framework for handling concept shift in one single network through incremental learning and simultaneously adapting samples from different networks to transfer knowledge about various intrusions. In addition, we employ active learning to use expert knowledge to label the samples for adaptation purposes [6].

All cases mentioned so far have the same label space for the source and target domains during adaptation. Occasionally, this is not the case, and we do not have access to samples for specific classes, either in the source or target.

One case is when we cannot access some classes in the source domain. This setting is called Partial Domain Adaptation (PDA). This setting is beneficial to network traffic classification systems because, in general, every network has different types of applications and, therefore, different types of traffic. We develop a method for transferring knowledge from a source network to a target network even if the source network does not contain all types of traffic [7].

Another case is when we have access to unlabeled target samples but not for all classes. We call this Limited Domain Adaptation (LDA) setting and propose a DA method for fault identification. The motivation behind this setting is that for developing a fault identification model for a system, we don't want to wait until the occurrence of all faults to collect even unlabeled samples; instead, we aim to use the knowledge about those faults from other domains [8].

Results on synthetic and real-world datasets for the scenarios mentioned above indicate that the proposed methods outperform the state-of-art and are effective and practical in solving real-world problems.

All of the above cases, however, have one thing in common: for the adaptation of domains, we gather all the samples (either labeled or unlabeled from either the source or target domains) and then apply DA techniques; in fact, the cases mentioned were all categorized as Centralized DA. There is, however, another case we call Decentralized DA, which in the literature is referred to as Heterogeneous Federated Learning.

A Federated Learning (FL) platform allows clients to collaborate in order to train Machine Learning models in a privacy-preserving manner, i.e., without being required to share any data; Clients collaborate through a server and, by sharing only their models, construct a generalizable model. In FL, one of the challenges is the statistical heterogeneity among clients; clients' data follow different distributions, so heterogeneous FL or decentralized DA solutions are needed.

One example is when we aim to build a model for a specific type of equipment. Different companies may use that equipment under different working conditions and aim to develop a model specific to that type of equipment. Nevertheless, companies refuse to share their data, instead sharing only their models. Therefore, on the one hand, data from different companies should be adapted through DA techniques. On the other hand, all the data is not centralized for DA to be applied.

In the context of fault identification, we examined the impact of statistical data heterogeneity on Federated Learning (FL). Based on our analysis of four types of heterogeneity among the clients, we discuss considerations that should be taken into consideration when designing FL solutions.

Finally, we proved theoretically that the distribution distance between clients affects the performance of an FL model. Following that, we proposed a method to remove heterogeneity between clients where the server constructs the client-specific generator. Data is generated by each generator so that its distribution follows that of other clients. As a result, the distribution of data among clients will be identical, and therefore, the heterogeneity between clients will no longer exist. As removing heterogeneity means adapting the clients' data, conventional

*This work is funded in part by KK Foundation and Vinnova.

Authors are with Center for Applied Intelligence Systems Research (CAISR), Halmstad University, Halmstad, Sweden. {firstname.lastname}@hh.se

FL algorithms such as FedAvg might be sufficient to train a well-generalizable model.

REFERENCES

[1] Taghiyarrenani Z. Learning from Multiple Domains (Licentiate dissertation, Halmstad University Press).

[2] Taghiyarrenani Z, Berenji A. An Analysis of Vibrations and Currents for Broken Rotor Bar Detection in Three-phase Induction Motors. InPHM Society European Conference 2022 Jun 29 (Vol. 7, No. 1, pp. 43-48).

[3] Taghiyarrenani Z, Berenji A. Noise-robust representation for fault identification with limited data via data augmentation. InPHM Society European Conference 2022 Jun 29 (Vol. 7, No. 1, pp. 473-479).

[4] Taghiyarrenani Z. Learning from Multiple Domains (Doctoral dissertation, Halmstad University Press).

[5] Taghiyarrenani Z, Nowaczyk S, Pashami S, Bouguelia MR. Multi-domain adaptation for regression under conditional distribution shift. Expert Systems with Applications. 2023 Mar 20:119907.

[6] Mahdavi E, Fanian A, Mirzaei A, Taghiyarrenani Z. ITL-IDS: Incremental Transfer Learning for Intrusion Detection Systems. Knowledge-Based Systems. 2022 Oct 11;253:109542.

[7] Taghiyarrenani, Z. and Farsi, H., 2023, January. Domain Adaptation with Maximum Margin Criterion with application to network traffic classification. In Machine Learning and Principles and Practice of Knowledge Discovery in Databases: International Workshops of ECML PKDD 2022, Grenoble, France, September 19–23, 2022, Proceedings, Part II (pp. 159-169). Cham: Springer Nature Switzerland.

[8] Taghiyarrenani Z, Nowaczyk S, Pashami S, Bouguelia MR. Towards Geometry-Preserving Domain Adaptation for Fault Identification. InMachine Learning and Principles and Practice of Knowledge Discovery in Databases: International Workshops of ECML PKDD 2022, Grenoble, France, September 19–23, 2022, Proceedings, Part II 2023 Jan 31 (pp. 451-460). Cham: Springer Nature Switzerland.