S3AI - www.S3AI.at
Security and Safety for Shared Artificial Intelligence by Deep Model Design

Host: SCCH, www.scch.at

Programme: COMET – Competence Centers for Excellent Technologies

Programme line: COMET-Module

Type of project: Strategic research project, 5 years, Start: Jan 2020

TACKLING DATA SHIFT IN MACHINE LEARNING (ML)

NOTABLE-TOP-5% DISTINCTION BY TOP ML CONFERENCE (ICLR'23)

Data shift is a wide-spread phenomenon in machine learning applications [1,2]. It is of particular interest in connection with personalized AI, such as that used in personalized medicine. The phenomenon however is more general. It always occurs when a machine learning model is transferred from a laboratory environment based on previously collected training data to a real application. Often there is a selection bias in the training data, or there are some data augmentation methods [5] at play to keep data labeling costs low. More often simulations are used to generate data, but then the real distribution will also be different from the training conditions.

Therefore, it would be desirable to be able to **learn** a model **on unlabeled data** from a target input distribution B using labeled data from a different source distribution A. See Fig. 1. for an illustration.

The related sub-field machine learning is referred to as *Unsupervised Domain Adaptation*. A basic question is whether and to which accuracy a learned baseline model can be transferred. In general, the Domain Adaptation problem is unsolvable, as a training set distribution $p_A(x,y)$ and an arbitrary distribution of interest $p_B(x,y)$ could be arbitrarily far

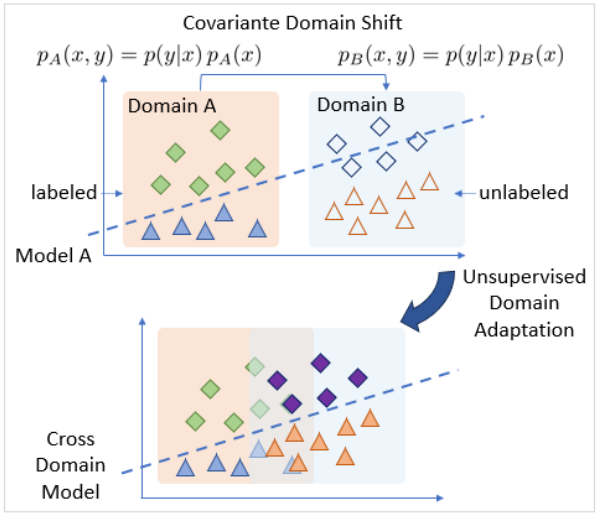

Fig. 1: Illustration of Unsupervised Domain Adaptation

apart. However, in many applications it seems to be reasonable to rely on the so-called covariate shift assumption, where the conditional probability distribution $p(y|x)$ is the same, i.e. $p_A(x,y) = p(y|x)\, p_A(x)$ and $p_B(x,y) = p(y|x)\, p_B(x)$ with different (personalized) marginal probabilities $p_A(x)$ and $p_B(x)$. While the conditional probability distribution comes from the causal relationship between between input and output, e.g. due some physiological mechanisms, the marginal distribution is very patient-dependent.

S3AI Methodology

Our approach proposes a theoretical framework for choosing hyper-parameters in unsupervised domain adaptation. The main strategy is to compute an **aggregation of models with target error bound**, which theoretically relies on the extension of importance weighted least squares to linear aggregation of vector-valued functions.

Citation from a reviewer (see open review [4]):

> "The paper conducts large scale comparative experiments on language, images, time-series classification tasks. The proposed method outperforms IWV and DEV and **sets a new state-of-the-art performance**."

Relevance for Personalized Medicine and Outlook

For example, imagine the context of diabetes and the problem to control the insulin supply a machine learning model for predicting the blood glucose concentration for some prediction time horizon (10, 20 or 30min). Such models are part of edge AI devices for diabetes patients for controlling the insulin supply and collect measurements from the past. Imagine a machine learning model, for example, a prediction model for the blood glucose concentration for some prediction time horizon (10, 20 or 30min). Such models are part of edge AI devices for diabetes patients for controlling the insulin supply. As patients have different lifestyles and physical conditions, it is more realistic to assume a data shift to a baseline model from the beginning. Any type 1 diabetes patient is recommended to keep a blood glucose diary, where daytime measurements are recorded. Then, this past data from a patient dairy together with unlabeled data from a training set can be used to estimate the Radon-Nikodym derivative [3] between these distributions, to determine an optimal aggregated prediction [4] based on a pool of predictors of choice.

Related S3AI Publications

[1] W. Zellinger, B. A. Moser, and S. Saminger-Platz, On generalization in moment-based domain adaptation. *Annals of Mathematics and Artificial Intelligence* 89, pp. 333-369, 2021, doi.org/10.1007/s10472-020-09719-x

[2] W. Zellinger, N. Shepeleva, M.-C. Dinu, H. Eghbal-zadeh, H.D. Nguyen, B. Nessler, S. Pereverzyev, B.A. Moser, The balancing principle for parameter choice in distance-regularized domain adaptation. *Advances in Neural Information Processing Systems* (NeurIPS 2021),

[3] E. R. Gizewski, L. Mayer, B. A. Moser, D. H. Nguyen, S. Pereverzyev Jr, S. V. Pereverzyev, N. Shepeleva, and W. Zellinger, On a regularization of unsupervised domain adaptation in RKHS, Applied and Computational Harmonic Analysis, 57:201–227, 2022, doi.org/10.1016/j.ins.2023.119838

[4] M.-C. Dinu, M. Holzleitner, M. Beck, H. D. Nguyen, A. Huber, H. Eghbal-zadeh, B.A. Moser, S. Pereverzyev, S. Hochreiter, W. Zellinger, Addressing Parameter Choice Issues in Unsupervised Domain Adaptation by Aggregation, in International Conference on Learning Representation, ICLR 2023 (notable-top-5%), openreview.net/forum?id=M95oDwJXayG

[5] H. Eghbal-zadeh, W. Zellinger, M. Pintor, K. Grosse, K. Koutini, B. A. Moser, B. Biggio, G. Widmer, Rethinking data augmentation for adversarial robustness, Information Sciences, Volume 654, 2024, 119838, ISSN 0020-0255, doi.org/10.1016/j.ins.2023.119838 .

Consortium:

JKU Institut für Machine Learning · KU LEUVEN · University of Cagliari, Italy · COSIC · TISSUEGNOSTICS PRECISION THAT INSPIRES · RICAM JOHANN-RADON-INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS · RM GROUP · EYES we make machines see · KTM · PKE · ventopay CUSTOMIZED PAYMENT SYSTEMS

Contact:

Bernhard A. Moser, S3AI Coordinator, SCCH, T +43 50 343 833, bernhard.moser@scch.at

Federal Ministry Republic of Austria Climate Action, Environment, Energy, Mobility, Innovation and Technology

Federal Ministry Republic of Austria Digital and Economic Affairs

FFG Promoting Innovation.

Austrian Research Promotion Agency
Sensengasse 1, A-1090 Vienna
P +43 (0) 5 77 55 - 0
office@ffg.at
www.ffg.at