CSC one-pager round 2023

Project title: Federated Learning using Privileged Information

Principal Investigators:

Dr. Evgueni Smirnov, Assistant Professor, Department of Advanced Computing Sciences, Maastricht University

Dr. Mirela Popa, Assistant Professor, Department of Advanced Computing Sciences, Maastricht University

Promotor: Prof. Dr. Anna Wilbik, Department of Advanced Computing Sciences, Maastricht University

Introduction: This project proposes a new paradigm of learning using privileged information (LUPI) in federated setting. Assume that we have N learning agents A_n s.t.:

(1) the data representation *is shared*; i.e. all the agents use common set (view) X of input variables and output variable Y; and

(2) each training dataset D_n *is not shared* among the agents.

The test instances, however, are represented in a smaller view $X' \subseteq X$. t. Therefore, the privileged view $X \setminus X'$ is only available during the training phase. The task of federated LUPI is to train a predictor h: $X' \to Y$ from training data D_n in view X when no data exchange is allowed. Federated LUPI has many appealing applications when the test instances are systematically incomplete and the datasets are not shared; e.g. monitoring patients, when a half of medical tests are performed, that is realized by a federated predictor trained on the data of different hospitals that are not allowed to exchange the data.

Objectives: This proposal aims at developing approaches to federated learning using privileged information.

Methods: The project will be organized in two stages: model-specific and model-agnostic. In the model-specific stage we will combine LUPI models of SVM, nearest neighbor, decision trees/rules, etc. with their federated counterparts. In the model- agnostic stage we will employ meta modeling of learning agents A_n using to-be-developed predictors from the previous stage.

Impact: The proposed approach will be applied in an ICU readmission prediction system developed by a medical data-science company.

Team: The supervisors' team has experience in theory and practice of machine learning as well as in guiding PhD students on this topic.

Requirements candidate: Solid background in math, computer science, and machine learning on a Master level. Good English language skills.

Keywords: machine learning, federated learning, learning using privileged information

Top 5 relevant selected publications:

1. S. Zhou, E. Smirnov, G. Schoenmakers, R. Peeters, X. Wu, *Conformal Feature-Selection Wrappers and ensembles for negative-transfer avoidance*, Neurocomputing 397, 309-319, 2020
2. S. Zhou, E. Smirnov, G. Schoenmakers, R. Peeters. *Conformity-based Source Subset Selection for Instance Transfer*. Neurocomputing 258: 41-51, 2017
3. F. Ismailoglu, R. Cavill, E. Smirnov, S. Zhou, P. Collins, R. Peeters, *Heterogeneous Domain Adaptation for IHC Classification of Breast Cancer Subtypes,* IEEE ACM Trans. Comput. Biol. Bioinform. 17(1), 347-353, 2020
4. A. Wilbik and P. Grefen, *Towards a Federated Fuzzy Learning System*, In Proc of the *2021 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 2021, pp. 1-6, doi: 10.1109/FUZZ45933.2021.9494392.
5. I. Belderbos, T. de Jong and M. Popa, *GANs Based Conditional Aerial Images Generation for Imbalanced Learning*, in Proceedings of the Third International Conference on Pattern Recognition and AI, Lecture Notes in Computer Science, volume 13364, 330—342, 2022