Neurocomputing 397 (2020) 264–265

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

Editorial

Special Issue on Conformal and Probabilistic Prediction with Applications

Alex Gammerman[a,*], Vladimir Vovk[a], Zhiyuan Luo[a], Evgueni Smirnov[b], Ralf Peeters[b]

[a] *Royal Holloway, University of London, Egham, Surrey, UK*
[b] *Maastricht University, the Netherlands*

ARTICLE INFO

Article history:
Received 20 October 2019
Accepted 8 November 2019
Available online 30 November 2019

Communicated by Prof. Zidong Wang

Keywords:
Competitive online prediction
Conformal clustering
Conformal prediction
Domain adaptation
Predictive distributions
Regression
Transfer learning

This Special Issue of Neurocomputing is devoted to conformal prediction, a relatively recent method of machine learning. In its basic form the method was developed in the late 1990s and early 2000s, and the main results of that period were summarized in the 2005 monograph [1]. Since then the method have been developed in various directions, and some later work is described in [2]. Since 2012 there have been eight annual symposia on Conformal and Probabilistic Prediction (COPA), held in various places of Europe. A sizable fraction of papers in this Special Issue are expanded and updated versions of papers published in the proceedings of the seventh COPA [3], which took place at the University of Maastricht in the summer of 2018. The Special Issue also includes several papers prepared specifically for it.

An advantage of conformal prediction is that the predictions it makes (usually in the form of prediction sets or predictive distributions) are well-calibrated under weak assumptions. For example, if we use a conformal predictor at a certain significance level ϵ in the online mode of prediction with IID observations, then errors at different trials will be made independently with probability ϵ. Therefore, we can control the number of errors by setting up a required significance level. This property is useful in many applications, and indeed, there have been applications of conformal prediction in medicine, environment, industry, and other areas.

The papers in this Special Issue can be roughly divided into three groups. First we have three papers treating the basic setting of conformal prediction, in which we have a homogeneous training set and the test observations are coming from the same distribution; in other words, all observations are assumed to be IID. In “Efficient conformal predictor ensembles”, Henrik Linusson, Ulf Johansson, and Henrik Boström construct new computationally efficient versions of conformal predictors, namely, ensemble conformal predictors. They introduce a novel technique, out-of-bag calibration, and study a new important property, *p*-value stability. “Multi-level conformal clustering: A distribution-free technique for clustering and anomaly detection” by Ilia Nouretdinov, James Gammerman, Matteo Fontana, and Daljit Rehal applies conformal prediction to clustering, presenting and studying the technique of multi-level conformal clustering. It builds on previous work by Lei et al. [4]. “Computationally efficient versions of conformal predictive distributions” by Vladimir Vovk, Ivan Petej, Ilia Nouretdinov, Valery Manokhin, and Alex Gammerman applies conformal prediction to probabilistic regression, following a line of work started at COPA 2017. For computational efficiency, this paper concentrates on split and cross conformal predictive distributions.

The following four papers treat various non-IID settings. The area of transfer learning is represented by “Conformal feature-selection wrappers and ensembles for negative-transfer avoidance”

* Corresponding author
E-mail addresses: a.gammerman@rhul.ac.uk (A. Gammerman), v.vovk@rhul.ac.uk (V. Vovk), zhiyuan.luo@cs.rhul.ac.uk (Z. Luo), smirnov@maastrichtuniversity.nl (E. Smirnov), ralf.peeters@maastrichtuniversity.nl (R. Peeters).

https://doi.org/10.1016/j.neucom.2019.11.025

(Shuang Zhou, Evgueni Smirnov, Gijs Schoenmakers, Ralf Peeters, and Xi Wu), which explains how conformal prediction can be used to avoid negative (i.e., harmful) transfer, and "Transfer learning extensions for the probabilistic classification vector machine" (Christoph Raab and Frank-Michael Schleif), which emphasizes sparsity and interpretability. The closely related area of domain adaptation is represented by "Audio-visual domain adaptation using conditional semi-supervised Generative Adversarial Networks" (Christos Athanasiadis, Enrique Hortal, and Stylianos Asteriadis), which applies inductive conformal prediction to the knowledge transfer for visual and audio data. Finally, "Gaussian process classification for variable fidelity data" (Nikita Klyuchnikov and Evgeny Burnaev) considers the setting in which the training set consists of two subsets, high- and low-quality, and we would like to take both into account fully when developing our classifier. Conformal prediction in non-IID settings has also been tackled recently outside the COPA series of workshops and special issues: see, e.g., 2019 papers [5,6]. This direction of research is important for applications, and we hope it will develop and grow further.

The final two papers are devoted to competitive online prediction, which establishes performance guarantees without making any statistical assumptions (not even IID, as in conformal prediction). "Adaptive hedging under delayed feedback" (Alexander Korotin, Vladimir V'yugin, and Evgeny Burnaev) studies the case of delayed feedback: the tth true label is disclosed not straight away, but in time D_t. This field is plagued by the existence of what the authors call "replicated algorithms": running a bunch of no-delay algorithms in the delayed framework, which on one hand is obviously wasteful but on the other hand achieves asymptotically optimal or nearly optimal results. "Universal algorithms for multinomial logistic regression under Kullback–Leibler game" by Raisa Dzhamtyrova and Yuri Kalnishkan develops new powerful algorithms for probabilistic classification (for extra generality, the outcomes are also probability measures, which leads to the "Kullback–Leibler game"). It includes extensions such as discounted loss (the undiscounted case is included as special case), tracking the best expert (multinomial logistic regression function), and kernelization.

Declaration of Competing Interest

There are no conflicting interests.

References

[1] V. Vovk, A. Gammerman, G. Shafer, Algorithmic Learning in a Random World, Springer, New York, 2005.
[2] V.N. Balasubramanian, S.-S. Ho, V. Vovk, Conformal Prediction for Reliable Machine Learning: Theory, Adaptations, and Applications, Elsevier, Amsterdam, 2014.
[3] A. Gammerman, V. Vovk, Z. Luo, E. Smirnov, R. Peeters, Conformal and probabilistic prediction and applications (proceedings of the seventh symposium, COPA 2018), in: Proceedings of Machine Learning Research, 91, 2018.
[4] J. Lei, A. Rinaldo, L. Wasserman, A conformal prediction approach to explore functional data, Ann. Math. Artif. Intell. 74 (2015) 29–43.
[5] R.F. Barber, E.J. Candès, A. Ramdas, R.J. Tibshirani, Conformal prediction under covariate shift, in: Advances in Neural Information Processing Systems 32 (NIPS 2019) pre-proceedings, 2019.
[6] L. Guan, R. Tibshirani, Prediction and Outlier Detection in Classification Problems, Tech. Rep. arXiv:1905.04396 [stat.ME], arXiv.org e-Print archive, 2019.