Abstract Viktor Bengs:

Title: The Fundamental Issues of Evidential Deep Learning

Abstract:

The representation and quantification of uncertainty in machine learning, most notably of predictive uncertainty in the setting of supervised learning, has recently attracted increasing attention [1]. Going beyond standard probabilistic prediction, various methods have been proposed that seek to distinguish between so-called aleatoric and epistemic uncertainty [2,3]. While the former refers to variability due to inherently random effects, the latter is uncertainty caused by a lack of knowledge and hence relates to the epistemic state of an agent.

One popular way for distinguishing these uncertainties in a prediction task is by means of the Evidential Deep Learning (EDL) paradigm [4]. Here, one essentially learns a model (usually a deep neural network) by empirical risk minimization, whose output for a query instance are the parameters of a parameterized family of a second-order distribution (e.g. Dirichlet or Normal-Inverse Gamma). The idea is that the learner uses the learned second-order distributional to assign probabilities to all possible label distributions for prediction purposes, rather than making a single point prediction, thus expressing aleatoric and epistemic uncertainty simultaneously.

Even if these approaches lead to good and plausible results in practice, the theoretical foundation of these approaches is not flawless. On the contrary, we will see in this talk that the whole underlying design idea of the EDL paradigm can be viewed extremely critical in the sense that the loss functions proposed to induce such second-order distributions provide no incentive for the learner to faithfully represent its epistemic uncertainty. These problems are made explicit in two ways: first, by the theoretical convergence analysis of the underlying empirical risk minimizer, and second, by the analysis of the proposed loss functions in terms of a (second-order) scoring rule.

[1] Hüllermeier, E. and Waegeman, W. Aleatoric and epistemic uncertainty in machine learning: An introduction to concepts and methods. Machine Learning, 110(3):457–506, 2021. doi: 10.1007/s10994-021-05946-3.

[2] Senge, R., Bösner, S., Dembczynski, K., Haasenritter, J., Hirsch, O., Donner-Banzhoff, N., and Hüllermeier, E. Reliable classification: Learning classifiers that distinguish aleatoric and epistemic uncertainty. Information Sciences, 255:16–29, 2014

[3] Kendall, A. and Gal, Y. What uncertainties do we need in Bayesian deep learning for computer vision? In Proc. NIPS, 30th Advances in Neural Information Processing Systems, pp. 5574–5584, 2017.

[4] Ulmer, D., Hardmeier, C., and Frellsen, J. (2023). Prior and posterior networks: A survey on evidential deep learning methods for uncertainty estimation. Transaction of Machine Learning Research.