

Efficient Predictive Uncertainty Estimators for Deep Probabilistic Models

Julissa Villanueva Llerena, Denis Deratani Mauá

Institute of Mathematics and Statistics

University of São Paulo, Brazil

{jgville, ddm}@ime.usp.br

Abstract

Deep Probabilistic Models (DPM) based on arithmetic circuits representation, such as Sum-Product Networks (SPN) and Probabilistic Sentential Decision Diagrams (PSDD), have shown competitive performance in several machine learning tasks with interesting properties (Poon and Domingos 2011; Kisa et al. 2014). Due to the high number of parameters and scarce data, DPMs can produce unreliable and overconfident inference. This research aims at increasing the robustness of predictive inference with DPMs by obtaining new estimators of the predictive uncertainty. This problem is not new and the literature on deep models contains many solutions. However the probabilistic nature of DPMs offer new possibilities to achieve accurate estimates at low computational costs, but also new challenges, as the range of different types of predictions is much larger than with traditional deep models. To cope with such issues, we plan on investigating two different approaches. The first approach is to perform a global sensitivity analysis on the parameters, measuring the variability of the output to perturbations of the model weights. The second approach is to capture the variability of the prediction with respect to changes in the model architecture. Our approaches shall be evaluated on challenging tasks such as image completion, multilabel classification.

Problem Statement

Deep models are now the state-of-the-art technique in several machine learning tasks. This enormous success has been followed by a large number of criticisms regarding their robustness, reliability and interpretability (Caruana et al. 2015). In fact, examples of overconfident and wrong predictions made by deep models are on the rise in the literature (Nguyen, Yosinski, and Clune 2015). This has renewed interest in developing new techniques for estimating the uncertainty about a model’s predictions, that is, for measuring how much the model “knows what it knows” (Lakshminarayanan, Pritzel, and Blundell 2017). One seemingly growing consensus is that the model “may simply not be the best judge of its trustworthiness” (Jiang et al. 2018). SPNs are a special type of compositional deep neural networks whose architecture represents context-specific probabilistic

independences and mixture models (Poon and Domingos 2011). Formally, a SPN is an acyclic computation graph with an initial layer that computes univariate probability distributions and internal layers that compute either products or weighted sums. Importantly, each node/neuron in a SPN computes a joint probability distribution of the variables that are connected (directly or indirectly) to it, and the network outputs a joint probability distribution over the variables in the input. A PSDD is also a tractable representation of probability distributions that are subject to logical constraints (Kisa et al. 2014).

The probabilistic semantics of SPNs and PSDDs allows for greater flexibility. Another gain of the probabilistic semantics is to allow for efficient learning of the architecture from data, a task that still lacks good algorithms for general deep neural networks. DPMs generalize poorly on regions with insufficient statistical support or conflicting data, leading to unreliable, overconfident and prior-dependent conclusions. It is thus important to ascertain the predictive uncertainty of inferences with the model, that is, to provide a confidence or trust measure for each prediction. This project will investigate new approaches to computing predictive uncertainty estimators for DPMs. To meet the needs of many applications, we focus on estimators that can be obtained in time comparable to the computation of a prediction (often linear in the size of the model). This excludes many of the current approaches (e.g. Bayesian inference), making the problem non trivial.

Contributions

We will follow two different approaches towards increasing the robustness of DPMs via predictive uncertainty estimators. The first approach estimates the uncertainty introduced by the parameter learning algorithm, and consists in analyzing the sensitivity of predictions to small perturbations of the weights. As a first attempt we will obtain uncertainty estimators by allowing each weight (or set of related weights) to vary independently within a pre-defined convex set. This allows an interval around the prediction to be estimated efficiently by propagating upper and lower bounds through the network. However, allowing weights to vary independently can lead to over conservative estimates, since many of the weights are highly correlated. To counter that, we plan on developing methods that allow weights to vary

given a certain budget, so that the variability of one weight connects to another. We expect this to lead to harder computational problems, which will probably require approximate solutions to ensure efficiency. The second approach consists in estimating the uncertainty introduced by the architecture used. To measure that, we propose using an ensemble of diverse SPNs (not necessarily the highest-scoring ones), so as to obtain estimates of the variability of the prediction as the network structure is changed. We plan on obtaining such diverse set by sampling structures. Here two different strategies can be adopted: either sampling uniformly, or sampling towards high-scoring networks while ensuring diversity. The algorithms developed will be evaluated by their performance on real-world tasks such as image completion, missing data imputation and multilabel classifications.

As our first contribution, we develop efficient algorithms for global quantitative analysis of Maximum-A-posteriori (MAP) inference in selective SPNs, a subclass of SPNs that allows tractable MAP inference with only a small decrease in accuracy (Peharz, Gens, and Domingos 2014). In particular, we devise a polynomial-time procedure to decide whether a given MAP configuration is *robust* with respect to changes in the model parameters. I evaluated our algorithms in two different tasks: completing missing values and performing multilabel classification. The results showed that our approach can discriminate easy- and hard-to-classify instances, often more accurately than criteria based on the probabilities induced by the model (Villanueva and Mauá 2019). We are currently carrying out a more extensive empirical evaluation, and adapting some of the algorithms to the formalism of PSDDs.

Related Work

We have already started pursuing the first approach, that obtains predictive estimators by varying weights independently. We showed that certain queries do not allow for polynomial-time estimators, while others admit very efficient routines. Empirical results show that this strategy outperforms predictive uncertainty estimators that rely on the model's prediction (Mauá et al. 2018; Villanueva and Mauá 2019; Antonucci, Facchini, and Mattei 2019). There are however still many variants of architectures and queries that remain unexplored. There has been a number of proposals for performing Bayesian inference with SPNs (Zhao et al. 2016), although not with the aim of estimating predictive uncertainty. These proposals either convert the problem into a numerical optimization or resort to sampling. In either case, the computation of the predictive uncertainty estimate is very sensitive to the choice of priors, and takes considerably more time than the computation of the prediction. The situation is similar to Bayesian inference with general neural networks – although recent works have advocated using drop-out as a form of efficient and approximate Bayesian inference estimate, this violates the assumptions of DPMs. More recently, there has been many proposals that try to overcome such issues, often restricted to classification (Jiang et al. 2018). Another approach consist in use ensemble of diverse networks to obtain estimates (Lakshminarayanan, Pritzel, and Blundell 2017) proposed, similar to

what we propose here. Unlike our proposed approach, they propose using a fixed architecture and different initializations of parameters. We have very recently started investigating approaches for sampling network structures of a certain type uniformly (Mattei et al. 2019).

Acknowledgments

This study was financed by the Brazilian Agency *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES) - Finance Code 001, CNPq grants no. 303920/2016-5 and 420669/2016-7.

References

- Antonucci, A.; Facchini, A.; and Mattei, L. 2019. Credal sentential decision diagrams. In *International Symposium on Imprecise Probabilities: Theories and Applications*.
- Caruana, R.; Lou, Y.; Gehrke, J.; Koch, P.; Sturm, M.; and Elhadad, N. 2015. Intelligible models for healthcare: Predicting pneumonia risk and hospital 30-day readmission. In *ACM International Conference on Knowledge Discovery and Data Mining*.
- Jiang, H.; Kim, B.; Guan, M.; and Gupta, M. 2018. To trust or not to trust a classifier. In *Advances in Neural Information Processing Systems*.
- Kisa, D.; Van den Broeck, G.; Choi, A.; and Darwiche, A. 2014. Probabilistic sentential decision diagrams. In *International Conference on the Principles of Knowledge Representation and Reasoning*.
- Lakshminarayanan, B.; Pritzel, A.; and Blundell, C. 2017. Simple and scalable predictive uncertainty estimation using deep ensembles. In *Advances in Neural Information Processing Systems*.
- Mattei, L.; Soares, D.; Antonucci, A.; Mauá, D.; and Facchini, A. 2019. Exploring the space of probabilistic sentential decision diagrams. In *Workshop of Tractable Probabilistic Modeling*.
- Mauá, D. D.; Conaty, D.; Cozman, F. G.; Poppenhaeger, K.; and de Campos, C. P. 2018. Robustifying sum-product networks. *International Journal of Approximate Reasoning*.
- Nguyen, A.; Yosinski, J.; and Clune, J. 2015. Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. In *IEEE conference on computer vision and pattern recognition*.
- Peharz, R.; Gens, R.; and Domingos, P. 2014. Learning selective sum-product networks. In *Workshop of Tractable Probabilistic Modeling*.
- Poon, H., and Domingos, P. 2011. Sum-product networks: A new deep architecture. In *Conference on Uncertainty in Artificial Intelligence*.
- Villanueva, J., and Mauá, D. 2019. Robust analysis of map inference in selective sum-product networks. In *International Symposium on Imprecise Probabilities: Theories and Applications*.
- Zhao, H.; Adel, T.; Gordon, G.; and Amos, B. 2016. Collapsed variational inference for sum-product networks. In *International Conference on Machine Learning*.