Nonparametric Bayes for structure learning of SPNs

MSc project

June, 2022

Project description

Sum-product-networks (SPNs, Poon and Domingos [3]) are probabilistic models that allow tractable computation of a large class of inferential problems, such as marginalization or conditioning of random variables. SPNs share limited similarity to Bayesian networks [7] as their computational graph is a DAG and because they require similar procedures to infer their structure (see Fig. 1). While the methodology to learn Bayesian networks is well developed, SPN structure learning algorithms are so far somewhat ad-hoc and poorly theoretically motivated. SPNs can grow infinitely in size since leaf nodes do not directly correspond to random variables as in Bayesian networks. This suggests the application of nonparametric Bayesian priors for SPN structure learning which are (possibly) infinite dimensional priors that can grow in dimensionality appropriately with the amount of available data[2].

Here, we propose learning the structure of an SPN using a nested Chinese restaurant process [1],a Bayesian nonparametric method, to allow the SPN to grow infinitely in depth and width. The procedure first draws a structure from the nCRP and then assigns scope functions and parameters following Trapp et al. [4] in an MCMC framework. Our method is the first that frames structure and parameter inference in a single probabilistic model and hence should not only serve well as future baseline to evaluate structure learning methods but also open up future directions of research building on nonparametric Bayesian methodology in probabilistic circuit inference. In addition, the method should be easily accommodated to variational approaches via tree-based stick-breaking [6] which could speed up the inference significantly.

For this project, the student will read the recent literature on probabilistic circuits and nonparametric Bayesian methods, develop a generative model for SPNs based on the nCRP, and implement it in a software package. A focus of the project will be the efficient implementation of the methodology to be actually able to use it in practice hence a good understanding of scientific computing is helpful.

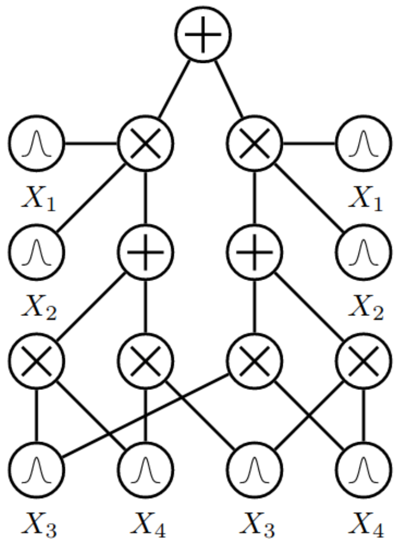

Figure 1: In contrast to Bayesian networks where every node encodes a random variable, an SPN consists of three types of nodes: sum-, product- and leaf-nodes. Sum-nodes induce mixture distributions, product-nodes factorizations, while leaf-nodes encode a random variable. Note that the same random variable can be represented by multiple leaf nodes. The root node encode the joint probability distribution (figure adopted from Vergari et al. [5]).

Additional Information

- **What will you learn?**
  - Good understanding of tractable deep learning, Bayesian statistics, Bayesian nonparametric methods
  - Python packaging and tools of the Python ecosystem (alternatively Julia possible)
- **Requirements**
  - Student of MSc program in computer science, data science, or statistics
  - Experience in scientific computing with Python (JAX) or Julia (!)
  - Basics of statistical machine learning and deep learning
  - Interest in probability theory, stochastic processes, Bayesian statistics
- **Supervisor** Dr Simon Dirmeier (simon.dirmeier@sdsc.ethz.ch) and Prof Dr Fernando Perez Cruz

References

[1] David M Blei, Thomas L Griffiths, and Michael I Jordan. "The nested chinese restaurant process and bayesian nonparametric inference of topic hierarchies". In: *Journal of the ACM (JACM)* 57.2 (2010), pp. 1–30.

[2] Subhashis Ghosal and Aad Van der Vaart. *Fundamentals of nonparametric Bayesian inference*. Vol. 44. Cambridge University Press, 2017.

[3] Hoifung Poon and Pedro Domingos. "Sum-product networks: A new deep architecture". In: *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)*. IEEE. 2011, pp. 689–690.

[4] Martin Trapp et al. "Bayesian learning of sum-product networks". In: *Advances in neural information processing systems* 32 (2019).

[5] Antonio Vergari et al. *Tutorial on "Probabilistic Circuits: Representations, Inference, Learning and Applications"*. 2020. URL: http://web.cs.ucla.edu/~guyvdb/talks/ECAI20-tutorial/.

[6] Chong Wang and David Blei. "Variational inference for the nested Chinese restaurant process". In: *Advances in Neural Information Processing Systems* 22 (2009).

[7] Han Zhao, Mazen Melibari, and Pascal Poupart. "On the relationship between sum-product networks and Bayesian networks". In: *International Conference on Machine Learning*. PMLR. 2015, pp. 116–124.