Robust Machine Learning by Credal Sentential Decision Diagrams

Abstract for the MLDM Workshop of the AI*IA 2019 Conference
November 2019, Cosenza, Italy

Lilith Mattei, Alessandro Antonucci, and Alessandro Facchini
Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA)
Galleria 2 - Manno (Switzerland)
`{lilith,alessandro,alessandro.facchini}@idsia.ch`

Abstract

Probabilistic sentential decision diagrams [2] are logical circuits annotated by probability mass functions on the disjunctive gates. This allows for a compact representation of joint mass functions consistent with logical constraints. We present a recent generalisation of the probabilistic quantification of these models, that allows to replace the local probabilities with *credal* sets, i.e., sets of mass functions specified by linear constraints [1]. Such a relaxed quantification induces a joint credal set, that sharply assigns probability zero to states inconsistent with the constraints. Algorithms to compute lower and upper bounds of marginal and conditional queries have been already derived in [1]. The first task can be achieved in linear time with respect to the diagram size for marginal queries. The same can be done for conditional queries if the topology of the circuit is singly connected. Here we also derive a polynomial-time algorithm to evaluate the robustness of a most probable explanation, i.e., to check whether all the most probable explanations in the probabilistic sentential decision diagrams consistent with the credal sentential decision diagram coincide. For a first validation of this new class of models, we focus on two machine learning tasks involving logical constraints: (i) classification under the closed-world assumption, i.e., only the observed instances are possible; (ii) preference learning by modelling each label in each rank as a separate Boolean variable. A set-valued quantification of these models is provided by the imprecise Dirichlet model, that offers cautious estimates of the local parameters when only small amounts of training data are available. The algorithm for conditional queries is adopted for the classification task, while preference learning is reduced to the robustness of the most probable explanation. Preliminary tests on benchmark data are promising: the credal models can successfully separate the easy instances (corresponding to the robust ones), i.e, those on which a standard model would provide more accurate results, from the difficult ones.

References

1. Antonucci, A., Facchini, A., Mattei, L.: Credal sentential decision diagrams. In: Proceedings of the Eleventh International Symposium on Imprecise Probabilities: Theories and Applications. Proceedings of Machine Learning Research, vol. 103, pp. 14–22. PMLR (2019)
2. Kisa, D., Van den Broeck, G., Choi, A., Darwiche, A.: Probabilistic sentential decision diagrams. In: Fourteenth International Conference on the Principles of Knowledge Representation and Reasoning (2014)