On Bridging the Gap Between Machine Learning and Knowledge Representation and Reasoning: The Case of Abstract Argumentation

Isabelle Kuhlmann

Artificial Intelligence Group, University of Hagen, Germany
isabelle.kuhlmann@fernuni-hagen.de

Motivation Two major fields in Artificial Intelligence are *Machine Learning* (ML) and *Knowledge Representation and Reasoning* (KRR). In ML, algorithms typically have the advantage that, once they are trained, they can yield solutions rather fast, but the disadvantage that the results are not guaranteed to be correct, and that solutions usually have no rationale or justification comprehensible to humans. By contrast, in KRR, algorithms are typically sound and complete. Hence, they yield correct, and usually "explainable" results. Consequently, a combination of ML and KRR approaches is highly promising for a plethora of problems originating from both fields. On the one hand, KRR methods can be used, e.g., to make ML approaches explainable; on the other hand, ML algorithms can be used to speed up KRR approaches. For instance, ML can be used for algorithm selection or for approximation in KRR. In the following, we will illustrate an example of the latter in order to outline both opportunities and challenges of combining ML and KRR.

Deciding Acceptability in Abstract Argumentation Using Deep Learning An *abstract argumentation framework* (AF) [2] is a tuple $F = (\mathsf{Args}, R)$, with Args being a set of *arguments* and $R \subseteq \mathsf{Args} \times \mathsf{Args}$ representing an *attack relation* between such arguments. AFs can be viewed as directed graphs, where the nodes represent Args and the edges R. In Abstract Argumentation (AA), one is usually interested in identifying *extensions*, i.e., sets of arguments that are jointly acceptable. Which constraints a set has to satisfy in order to be considered an extension, depends on the *semantics* used in the reasoning process. Typical problems in AA are deciding whether an argument is credulously accepted (DC) or skeptically accepted (DS), i.e., whether it is contained in at least one, or in all extensions of a given semantics, respectively. Most algorithms for solving such problems are sound and complete (see [6] for a recent overview). However, these problems exhibit a high complexity—e.g., deciding DS for *preferred* semantics is Π_2^P-complete [3]—which hinders a scalable behavior. Therefore, some authors suggest to use deep learning approaches [5, 1, 7] to compute solutions which are only approximate on the one hand, but fast to obtain on the other hand.

The objective of an ML system is to "learn" from given data (*training set*) and to use the acquired "knowledge" to make predictions about previously unknown data. In a feasibility study, Kuhlmann and Thimm [5] trained a *Graph*

Convolutional Network (GCN) [4] on a set of AFs. Each argument in the training set had a label marking it as either "accepted" (a) or "not accepted" (na). The authors demonstrated that a GCN is able to correctly classify arguments as a or na to a certain degree, but also pointed out some issues. The classes are often unevenly distributed, as there are usually more unaccepted than accepted arguments in an AF. This issue is generally not new in ML; there exist, e.g., a number of *augmentation* techniques which counteract this problem. However, augmentation methods for other graph data (e.g., citation or social networks), cannot simply be adopted for AFs, as they mostly consist of adding/deleting arguments or entire sub-graphs, or of manipulating node features. Since arguments do not possess any features, and modifying the graph topology could change the arguments' acceptability status, we cannot use such methods. Malmqvist et al. [7] address this problem by introducing a scheme to dynamically balance the training data, as well as a randomized training regime. Further, Craandijk and Bex [1] propose an *Argumentation Graph Neural Network* (AGNN) which learns a message-passing algorithm.

Even though some of the results (in particular those in [1]) are quite promising, the existing works on this topic are difficult to compare, since they all use different datasets—which is most likely due to the lack of a standard dataset for such purposes. Hence, data selection poses an additional challenge in combining ML and argumentation. Another problem of applying existing techniques, such as GCNs, to AA, lies in the fact that they are based on the assumption that closely connected data points are similar and are thus likely to belong to the same class. Although this is true for most graph-structured data (again, such as citation or social networks), it is explicitly not true for AFs: If one argument attacks another, they are direct neighbors, but they cannot both be accepted.

Conclusion and Future Work In the scope of this work, we discussed an example of combining ML (specifically, neural networks) with KRR (specifically, AFs). We explained that there are already quite promising results regarding this problem, but we also identified challenges. Examples of future work include the need for an appropriate (standard) dataset, as well as some learning approaches which are specifically designed with the "adversarial" nature of AFs in mind.

References

1. Craandijk, D., Bex, F.: Deep learning for abstract argumentation semantics. In: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. p. 1667–1673 (2020)
2. Dung, P.M.: On the acceptability of arguments and its fundamental role in non-monotonic reasoning, logic programming and n-person games. Artificial Intelligence **77**(2), 321–358 (1995)
3. Dvorák, W., Dunne, P.E.: Computational problems in formal argumentation and their complexity. Handbook of formal argumentation pp. 631–687 (2018)
4. Kipf, T.N., Welling, M.: Semi-supervised classification with graph convolutional networks. In: International Conference on Learning Representations (ICLR) (2017)

5. Kuhlmann, I., Thimm, M.: Using graph convolutional networks for approximate reasoning with abstract argumentation frameworks: A feasibility study. In: International Conference on Scalable Uncertainty Management. pp. 24–37. Springer (2019)
6. Lagniez, J.M., Lonca, E., Mailly, J.G., Rossit, J.: Design and results of ICCMA 2021. arXiv preprint arXiv:2109.08884 (2021)
7. Malmqvist, L., Yuan, T., Nightingale, P., Manandhar, S.: Determining the acceptability of abstract arguments with graph convolutional networks. In: SAFA@ COMMA. pp. 47–56 (2020)