Logics in Machine Learning and Data Mining: Achievements and Open Issues

Francesca A. Lisi

Dipartimento di Informatica &
Centro Interdipartimentale di Logica e Applicazioni (CILA)
Università degli Studi di Bari "Aldo Moro", Italy
`FrancescaAlessandra.Lisi@uniba.it`

Abstract. This short paper overviews 20 years of work done in logic-based Machine Learning and Data Mining along three different directions of research. The aim is to discuss the achievements and the open issues with reference to some challenging applications which involve representation and reasoning.

Keywords: Inductive Logic Programming · Ontology Reasoning · Description Logics · Fuzzy Logic · Metamodeling

1 Introduction

The current hype about AI is mainly due to a number of successful applications of Machine Learning (ML) and Data Mining (DM) algorithms in challenging domains such as vision. Most of these algorithms belong to the new generation of neural networks known under the name of "deep learning". Deep learning follows a function-based approach, *i.e.*, it formulates ML tasks as function-fitting problems. Neural networks are therefore considered as examples of *model-free AI* according to the definition given by Hector Geffner in his keynote talk at IJCAI 2018 [12]. While analyzing the limitations of this approach, several works (see, *e.g.*, [4]) have placed the emphasis on the need to construct and use models, as required by *model-based AI*, for the sake of interpretability and explainability. The model-based approach - a distinguishing feature of what is currently referred to as "good old-fashioned AI" - requires one to represent knowledge about entities of a domain of interest and involves reasoning with such knowledge. Logics and probability are among the main tools of this approach today.

As a drawback of the popularity of deep learning, the emerging trend is to have ML streamlined into neural network research. Yet, the variety of ML and DM algorithms is wide enough to have a significant overlap with model-based AI. *Inductive Logic Programming* (ILP) [35] is considered as the major logic-based (thus, model-based) approach to learning and mining rules from structured data. Originally focused on the induction of logic programs, due to the common roots with *Logic Programming* (LP) [32], ILP has significantly widened its scope over the years to cover all aspects of learning and mining in logic [34]. Notable is the exploration of the intersections to statistical learning and other probabilistic approaches (see, *e.g.*, [38] for a survey).

In the following section I will overview the work done in ILP over the past 20 years along three different directions of research. The aim is to discuss the achievements and the open issues with reference to some challenging applications which involve representation and reasoning. In Section 3 I will conclude the paper with final remarks.

2 Three Cases for Logics in ML and DM

2.1 Combining rules and ontologies

Rules are widely used in Knowledge Engineering (KE) as a powerful way of modeling knowledge. However, the acquisition of rules for very large Knowledge Bases (KBs) still remains a very demanding KE activity. A partial automation of the rule authoring task can be of help even though the automatically produced rules are not guaranteed to be correct. A viable solution to this KE bottleneck is just applying ILP algorithms. ILP has been historically concerned with learning rules from examples and background knowledge with the aim of prediction (see, *e.g.*, the system FOIL [37]). However, ILP has also been applied to tasks - such as association rule mining - other than classification where the scope of induction is description rathen than prediction. A notable example of this kind of ILP systems is WARMR [7] which mines frequent DATALOG queries.

With the advent of the Semantic Web new challenges and opportunities have been presented to ILP. In particular, ontologies and their logical foundations in the family of *Description Logics* (DLs) [2] raised several issues for the direct application of existing ILP systems, thus urging the extension and/or adaptation of the ILP methodological apparatus to the novel context. The reason for this is the following: LP and DLs are both based on fragments of First Order Logic (FOL), yet characterized by different semantic assumptions [33]. Though a partial overlap exists between LP and DLs, even more interesting is a combination of the two via several integration schemes that are aimed at designing very expressive FOL languages and ultimately overcoming the aforementioned semantic mismatch (see, *e.g.*, [9] for a survey). A representative example of this class of hybrid KR formalisms is $\mathcal{AL}$-LOG [8] which tightly integrates DATALOG and the DL $\mathcal{ALC}$.

Starting from the seminal work by Rouveirol and Ventos [39], several proposals in ILP testify the great potential of these formalisms also from the ML&DM perspective [15,25,19,20,21,23]. Originally motivated by a spatial data mining application [1] and inspired by WARMR, $\mathcal{AL}$-QUIN [25,21] is an ILP system for mining association rules at multiple levels of granularity within the KR framework of $\mathcal{AL}$-LOG. Here, reasoning in $\mathcal{AL}$-LOG allows for the actual exploitation of taxonomies possibly made available as background knowledge, such as the classification of spatial objects in geographic information systems (see [1,25] for examples of application in this context).

2.2 Dealing with imprecision and granularity

Imprecision is a weak form of vagueness, not to be mistaken for uncertainty, which is often formalized with fuzzy set theory. For instance, spatial notions such as the distance between two sites can be naturally represented with fuzzy sets (modeling the degrees of distance, *e.g.*, high, medium and low) if one is interested in their human perception rather than in precise measurements. In order to deal with imprecision in Ontology Reasoning several fuzzy extensions of DLs have been proposed (see, *e.g.*, [40] for an overview).

The problem of automatically managing the evolution of fuzzy DL ontologies has attracted some interest in the ILP community [16,14,27,28]. Iglesias and Lehmann [14] extend DL-Learner [18] (the state-of-the-art ILP system for learning in DLs) with some of the most up-to-date fuzzy ontology tools. Notably, the resulting system can learn fuzzy OWL DL equivalence axioms from FuzzyOWL 2 [1] ontologies by interfacing the *fuzzyDL* [2] reasoner. Lisi and Straccia [27] propose *Soft*FOIL, a FOIL-like method for learning fuzzy $\mathcal{EL}$ GCI axioms from fuzzy DL assertions. In [31], the same authors present FOIL-$\mathcal{DL}$, another FOIL-like method which, conversely, is designed for learning fuzzy $\mathcal{EL}(\mathbf{D})$ GCI axioms from crisp DL assertions. As opposite to *Soft*FOIL, FOIL-$\mathcal{DL}$ has been implemented and tested [28], notably in a real-world tourism application where fuzzy DLs come into play for modeling imprecise knowledge such as the hotel price ranges.

Imprecision dealt with fuzzy sets is strongly related to the notion of information granule. In [26], Lisi and Mencar propose a granular computing method for OWL 2 ontologies with the ultimate goal of optimizing the learning process when dealing with a huge number of relations, *e.g.*, those concerning the distance between places in the abovementioned tourism application. Here, information granulation encompasses the use of fuzzy quantifiers such as "most" and "a few" in OWL 2 ontologies as detailed in [30]. Soft quantification has been also explored in statistical relational learning [10].

2.3 Modeling and metamodeling

Research in ML and DM has traditionally focussed on designing effective algorithms for solving particular tasks, most of which can be seen as Constraint Satisfaction Problems (CSPs) or Optimization Problems (OPs). However, there is an increasing interest in providing the user with a means for specifying what the ML/DM problem in hand actually is, rather than letting him struggle to outline how the solution to that problem needs to be computed (see the recent note by De Raedt [6]). This corresponds to a *model+solver* approach to ML and DM, in which the user specifies the problem in a *declarative modeling language* and the system automatically transforms such models into a format that can be used by a *solver* to efficiently generate a solution. For instance, constraint programming has been successfully applied to itemset mining problems (see, *e.g.*, [13]

[1] `http://www.straccia.info/software/FuzzyOWL/`
[2] `http://www.straccia.info/software/fuzzyDL/intro.html`

for a comprehensive account). The *model+solver* approach is also at the basis of Meta-Interpretive Learning (MIL) [36], a novel and promising ILP framework. MIL uses descriptions in the form of meta-rules (expressed in a higher-order dyadic Datalog fragment) with procedural constraints incorporated within a meta-interpreter, which could be eventually implemented by relying on *Answer Set Programming* (ASP) solvers (see [11] for an updated overview).

The importance of metamodeling in several applications has been recently recognized in the KR community, with an increasing interest in higher-order DLs. In particular, De Giacomo *et al.* [5] augment a DL with variables that may be interpreted - in a Henkin semantics - as individuals, concepts, and roles at the same time, obtaining a new logic $Hi(\mathcal{DL})$. Colucci *et al.* [3] introduce second-order features in DLs under Henkin semantics for modeling several forms of non-standard reasoning. Lisi [22] extends [3] to some variants of Concept Learning, thus being the first to propose higher-order DLs as a means for metamodeling in ML and DM. In [29], the proposed *model+solver* approach combines the efficacy of higher-order DLs in metamodeling (as shown in [22]) with the efficiency of ASP solvers in dealing with CSPs and OPs. More recently, higher-order DLs have been considered as a starting point for the definition of a metaquerying language for mining the Web of Data [24].

3 Final remarks

Initiatives such as the workshop series promoted by the Association for Neuro-Symbolic Integration (NeSy)[3] since 2005 testify the need to address a fundamental open issue in AI: How to come up with a computational model capable of learning and reasoning both at the symbolic and the sub-symbolic level?

One such issue is also addressed by the Angry Birds AI[4] competition, built around what is currently considered a challenging problem for AI: to build an intelligent agent that can play new levels of the Angry Birds game better than the best human players. This is a very difficult problem as it requires agents to predict the outcome of physical actions without having complete knowledge of the world, and then to select a good action out of infinitely many possible actions. A distinguishing feature of future AI systems is just this capability of interacting with the physical world. The Angry Birds AI competition provides a simplified and controlled environment for developing and testing this capability.

The ILP works overviewed in this short paper testify an effort towards the integration between learning and reasoning, mostly at the symbolic level. However, the use of fuzzy logic could be considered as an attempt at dealing with the sub-symbolic level. Also, as opposed to neural networks, fuzzy systems have the potential of being interpretable and explainable.

A notorious drawback for ILP is the cost of computation. One of the advantages of the model+solver approach should be just to choose the most efficient solver to improve the performance of the learning process while preserving the

[3] `http://www.neural-symbolic.org/`
[4] `https://aibirds.org/`

declarativity of the model. In this respect, Geffner's vision [12] of true AI based on the integration between model-free learners and model-based solvers is a great source of inspiration.

Acknowledgments This work was partially funded by the INdAM - GNCS Project 2019 "Metodi per il trattamento di incertezza ed imprecisione nella rappresentazione e revisione di conoscenza", and by the Università degli Studi di Bari "Aldo Moro" under the IDEA Giovani Ricercatori 2011 grant "Dealing with Vague Knowledge in Ontology Refinement".

References

1. Appice, A., Ceci, M., Lanza, A., Lisi, F.A., Malerba, D.: Discovery of spatial association rules in geo-referenced census data: A relational mining approach. Intelligent Data Analysis **7**(6), 541–566 (2003), `http://content.iospress.com/articles/intelligent-data-analysis/ida00146`
2. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P. (eds.): The Description Logic Handbook: Theory, Implementation and Applications (2nd ed.). Cambridge University Press (2007)
3. Colucci, S., Di Noia, T., Di Sciascio, E., Donini, F.M., Ragone, A.: A unified framework for non-standard reasoning services in description logics. In: Coelho, H., Studer, R., Wooldridge, M. (eds.) ECAI 2010 - 19th European Conference on Artificial Intelligence, Lisbon, Portugal, August 16-20, 2010, Proceedings. Frontiers in Artificial Intelligence and Applications, vol. 215, pp. 479–484. IOS Press (2010)
4. Darwiche, A.: Human-level intelligence or animal-like abilities? Commun. ACM **61**(10), 56–67 (2018). `https://doi.org/10.1145/3271625`
5. De Giacomo, G., Lenzerini, M., Rosati, R.: Higher-order description logics for domain metamodeling. In: Burgard, W., Roth, D. (eds.) Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2011, San Francisco, California, USA, August 7-11, 2011 (2011)
6. De Raedt, L.: Languages for learning and mining. In: Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, January 25-30, 2015, Austin, Texas, USA. pp. 4107–4111 (2015), `http://www.aaai.org/ocs/index.php/AAAI/AAAI15/paper/view/9934`
7. Dehaspe, L., Toivonen, H.: Discovery of frequent Datalog patterns. Data Mining and Knowledge Discovery **3**, 7–36 (1999)
8. Donini, F.M., Lenzerini, M., Nardi, D., Schaerf, A.: $\mathcal{AL}$-log: Integrating Datalog and Description Logics. Journal of Intelligent Information Systems **10**(3), 227–252 (1998)
9. Drabent, W., Eiter, T., Ianni, G., Krennwallner, T., Lukasiewicz, T., Maluszynski, J.: Hybrid Reasoning with Rules and Ontologies. In: Bry, F., Maluszynski, J. (eds.) Semantic Techniques for the Web, The REWERSE Perspective, Lecture Notes in Computer Science, vol. 5500, pp. 1–49. Springer (2009)
10. Farnadi, G., Bach, S.H., Moens, M., Getoor, L., Cock, M.D.: Soft quantification in statistical relational learning. Machine Learning **106**(12), 1971–1991 (2017). https://doi.org/10.1007/s10994-017-5647-3, `https://doi.org/10.1007/s10994-017-5647-3`

11. Gebser, M., Leone, N., Maratea, M., Perri, S., Ricca, F., Schaub, T.: Evaluation techniques and systems for answer set programming: a survey. In: Lang [17], pp. 5450–5456. https://doi.org/10.24963/ijcai.2018/769
12. Geffner, H.: Model-free, model-based, and general intelligence. In: Lang [17], pp. 10–17. https://doi.org/10.24963/ijcai.2018/2
13. Guns, T., Nijssen, S., De Raedt, L.: Itemset mining: A constraint programming perspective. Artificial Intelligence **175**(12-13), 1951–1983 (2011)
14. Iglesias, J., Lehmann, J.: Towards integrating fuzzy logic capabilities into an ontology-based inductive logic programming framework. In: Proc. of the 11th Int. Conf. on Intelligent Systems Design and Applications. IEEE Press (2011)
15. Kietz, J.U.: Learnability of description logic programs. In: Matwin, S., Sammut, C. (eds.) Inductive Logic Programming, 12th International Conference, ILP 2002, Sydney, Australia, July 9-11, 2002. Revised Papers. Lecture Notes in Computer Science, vol. 2583, pp. 117–132. Springer (2003)
16. Konstantopoulos, S., Charalambidis, A.: Formulating description logic learning as an inductive logic programming task. In: Proc. of the 19th IEEE Int. Conf. on Fuzzy Systems. pp. 1–7. IEEE Press (2010)
17. Lang, J. (ed.): Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence, IJCAI 2018, July 13-19, 2018, Stockholm, Sweden. ijcai.org (2018), http://www.ijcai.org/proceedings/2018/
18. Lehmann, J.: DL-Learner: Learning Concepts in Description Logics. Journal of Machine Learning Research **10**, 2639–2642 (2009)
19. Lisi, F.A.: Building Rules on Top of Ontologies for the Semantic Web with Inductive Logic Programming. Theory and Practice of Logic Programming **8**(03), 271–300 (2008)
20. Lisi, F.A.: Inductive Logic Programming in Databases: From Datalog to $\mathcal{DL}$+log. Theory and Practice of Logic Programming **10**(3), 331–359 (2010)
21. Lisi, F.A.: $\mathcal{AL}$-QuIn: An Onto-Relational Learning System for Semantic Web Mining. International Journal on Semantic Web and Information Systems **7**(3), 1–22 (2011)
22. Lisi, F.A.: A declarative modeling language for concept learning in description logics. In: Riguzzi, F., Zelezny, F. (eds.) Inductive Logic Programming, 22nd International Conference, ILP 2012, Dubrovnik, Croatia, September 17-19, 2012, Revised Selected Papers. Lecture Notes in Computer Science, vol. 7842. Springer Berlin Heidelberg (2013)
23. Lisi, F.A.: Learning onto-relational rules with inductive logic programming. In: Lehmann, J., Völker, J. (eds.) Perspectives on Ontology Learning, Studies on the Semantic Web, vol. 18, pp. 93–111. IOS Press/AKA (2014)
24. Lisi, F.A.: Mining the web of data with metaqueries. In: Riguzzi, F., Bellodi, E., Zese, R. (eds.) Up-and-Coming and Short Papers of the 28th International Conference on Inductive Logic Programming (ILP 2018), Ferrara, Italy, September 2-4, 2018. CEUR Workshop Proceedings, vol. 2206, pp. 92–99. CEUR-WS.org (2018), http://ceur-ws.org/Vol-2206/paper8.pdf
25. Lisi, F.A., Malerba, D.: Inducing Multi-Level Association Rules from Multiple Relations. Machine Learning **55**, 175–210 (2004)
26. Lisi, F.A., Mencar, C.: A granular computing method for OWL ontologies. Fundamenta Informaticae **159**(1–2), 147–174 (2018). https://doi.org/10.3233/FI-2018-1661
27. Lisi, F.A., Straccia, U.: A logic-based computational method for the automated induction of fuzzy ontology axioms. Fundamenta Informaticae **124**(4), 503–519 (2013)

28. Lisi, F.A., Straccia, U.: Learning in description logics with fuzzy concrete domains. Fundamenta Informaticae **140**(3-4), 373–391 (2015). http://dx.doi.org/10.3233/FI-2015-1259
29. Lisi, F.A.: A model+solver approach to concept learning. In: Adorni, G., Cagnoni, S., Gori, M., Maratea, M. (eds.) AI*IA 2016: Advances in Artificial Intelligence - XVth International Conference of the Italian Association for Artificial Intelligence, Genova, Italy, November 29 - December 1, 2016, Proceedings. Lecture Notes in Computer Science, vol. 10037, pp. 266–279. Springer (2016), http://dx.doi.org/10.1007/978-3-319-49130-1_20
30. Lisi, F.A., Mencar, C.: Introducing fuzzy quantification in OWL 2 ontologies. In: Monica, D.D., Murano, A., Rubin, S., Sauro, L. (eds.) Joint Proceedings of the 18th Italian Conference on Theoretical Computer Science and the 32nd Italian Conference on Computational Logic co-located with the 2017 IEEE International Workshop on Measurements and Networking (2017 IEEE M&N), Naples, Italy, September 26-28, 2017. CEUR Workshop Proceedings, vol. 1949, pp. 321–325. CEUR-WS.org (2017), http://ceur-ws.org/Vol-1949/CILCpaper08.pdf
31. Lisi, F.A., Straccia, U.: A FOIL-like Method for Learning under Incompleteness and Vagueness. In: Zaverucha, G., Santos Costa, V., Paes, A. (eds.) Inductive Logic Programming - 23rd International Conference, ILP 2013, Rio de Janeiro, Brazil, August 28-30, 2013, Revised Selected Papers. Lecture Notes in Computer Science, vol. 8812, pp. 123–139. Springer (2014). https://doi.org/10.1007/978-3-662-44923-3_9
32. Lloyd, J.W.: Foundations of Logic Programming. Springer, 2nd edn. (1987)
33. Motik, B., Rosati, R.: Reconciling description logics and rules. J. ACM **57**(5) (2010)
34. Muggleton, S., De Raedt, L., Poole, D., Bratko, I., Flach, P.A., Inoue, K., Srinivasan, A.: ILP turns 20 - Biography and future challenges. Machine Learning **86**(1), 3–23 (2012)
35. Muggleton, S.H.: Inductive logic programming. In: Arikawa, S., Goto, S., Ohsuga, S., Yokomori, T. (eds.) Proceedings of the 1st Conference on Algorithmic Learning Theory. Springer/Ohmsma (1990)
36. Muggleton, S.H., Lin, D.: Meta-interpretive learning of higher-order dyadic datalog: Predicate invention revisited. In: Rossi, F. (ed.) IJCAI 2013, Proceedings of the 23rd International Joint Conference on Artificial Intelligence, Beijing, China, August 3-9, 2013. IJCAI/AAAI (2013)
37. Quinlan, J.R.: Learning logical definitions from relations. Machine Learning **5**, 239–266 (1990)
38. Riguzzi, F., Bellodi, E., Zese, R.: A history of probabilistic inductive logic programming. Frontiers in Robotics and AI **2014** (2014). http://dx.doi.org/10.3389/frobt.2014.00006
39. Rouveirol, C., Ventos, V.: Towards Learning in CARIN-$\mathcal{ALN}$. In: Cussens, J., Frisch, A.M. (eds.) Inductive Logic Programming, 10th International Conference, ILP 2000, London, UK, July 24-27, 2000, Proceedings. Lecture Notes in Artificial Intelligence, vol. 1866, pp. 191–208. Springer (2000)
40. Straccia, U.: All about fuzzy description logics and applications. In: Faber, W., Paschke, A. (eds.) Reasoning Web. Web Logic Rules - 11th International Summer School 2015, Berlin, Germany, July 31 - August 4, 2015, Tutorial Lectures. Lecture Notes in Computer Science, vol. 9203, pp. 1–31. Springer (2015). http://dx.doi.org/10.1007/978-3-319-21768-0