Addressing the Meta-Learning problem with Metaheuristics

(Extended Abstract)

William Raynaut
Chantal Soule-Dupuy
Nathalie Valles-Parlangeau
Cedric Dray
Philippe Valet

IRIT UMR 5505, UT1, UT3, INSERM
Universite de Toulouse, France
firstname.lastname@irit.fr

February 23, 2017

Abstract

The many machine learning and data mining techniques produced over the last decades can prove invaluable assets in diverse fields, but choosing the most appropriate for a given application may be very difficult for a non-expert. Our objective is thus to provide modelling assistance using a meta-learning approach based on an evolutionary metaheuristic. We present the intended workflow of such modelling assistant and the expected challenges along our line of work.

1 Motivation

The machine learning community has produced, over the last decades, an important variety of techniques and algorithms addressing the learning problem. But among different datasets, the performance of such techniques can vary a lot, leaving useless on a particular domain an algorithm that excelled on another. It has also been shown that no individual algorithm could perform better than all the other on every possible dataset [12]. These issues led to extensive study of the Meta-Learning problem : "*Which learning algorithm will perform best on a given learning problem ?*".

This problem has notably been addressed via rule-generation in [1][4], where the meta-knowledge describes the conditions under which the significant performance difference between algorithms holds. Despite many other successful applications over a limited range of learning tasks, and the insight that it could be solved by an infinite recursion of adaptive learners [11], the Meta-Learning problem still carries many open perspectives.

A computer program qualifies as a learning machine if its performance improves with experience. In most approaches of the Meta-Learning problem, this is achieved via the generation of meta-knowledge, which is then used to train a meta-learner. Our own perspective view on the matter is that the meta-knowledge can be viewed as a population of meta-instances, each describing the evaluated application of a learning task to a given dataset, and that a good solution to a given meta-learning task can be obtained via the evolutionary exploration of this population. Such approach is giving interesting results among other classes of problems [13], but, to our knowledge, has not yet been explored regarding Meta-Learning. We will detail its intended workflow in section 2, then the expected challenges and our inceptive approach tackling them in section 3.

2 Intended workflow

Our objective is to provide modelling assistance using a meta-learning approach. This section will detail a modelling assistant intended workflow illustrated by figure 1.

First, from the end-user point of view, we got a dataset describing features of a number of instances, and a specific need toward the modelling of this dataset. For instance, a medical practitioner trying to predict a pathology will give priority to the model predictive sensitivity, while a researcher modelling a misunderstood phenomena will care more about informativity and explainability...

Then, from the modelling assistant point of view, we face two main issues : find a good solution to the given problem, and generate meta-data to find better solutions in the future.

Let us consider the meta-data as a population of meta-instances. Each of those meta-instances represents the application of a modelling treatment to a specific dataset, which is then evaluated on different criteria. For example, one of our meta-instances could denote that the dataset submitted by our medical practitioner was once modelled using a Naïve-Bayes classification algorithm, which achieved a given predictive accuracy, sensitivity and specificity. Those meta-instances are thus described by meta-features, characterising respectively the dataset, the diverse applied treatments, and the evaluation of the resulting model.

Finally, to find a good solution to the problem as close as possible to the user's need, we intend to apply an evolutionary metaheuristic over this meta-instance population. The function evaluating the meta-instances fitness should take into account the user specified objectives and preferences to reflect the fact

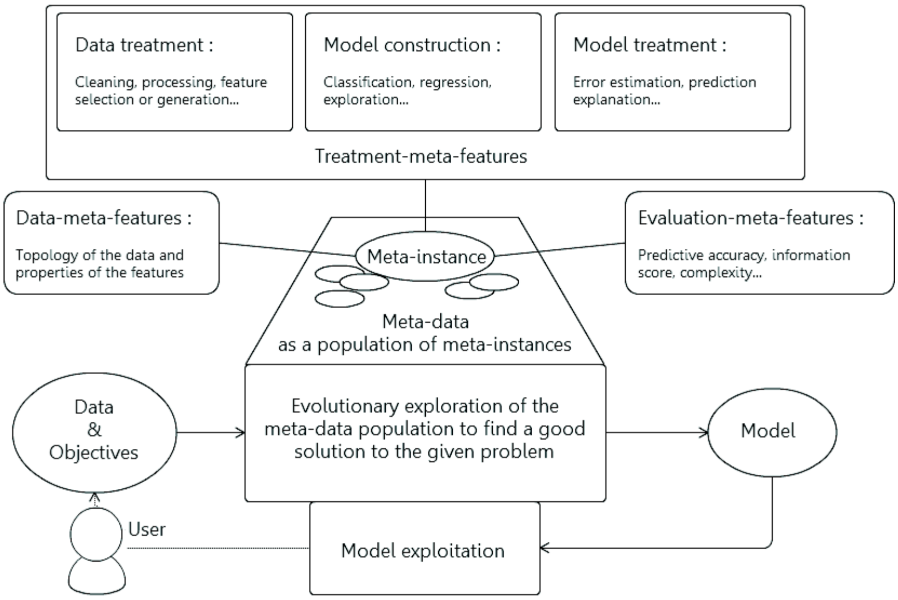

Figure 1: Modeling assistant

that the best fit meta-instance (so the best modelling way) on two different problems has no reason to be the same. The process is thus to use the best available meta-instances to *evolve* solutions to the problem, which are new meta-instances enriching the meta-data, and repeat until a solution deemed good enough is obtained.

3 Expected challenges

One of the first challenges will be to define the meta-features that will describe our meta-instances. Those sets of meta-features should be large enough to characterise well any modelling task, but a balance must be found to avoid the abundance of indecisive features and limit computational complexity. Furthermore, in order to discriminate between meta-features or meta-instances according to the user's need, the comparison of meta-features of a particular meta-instance - or of a given meta-feature over several meta-instances - should be possible and make sense.

Meta-features describing the data will mostly consist of descriptions of the feature space topology, while those describing the modelling treatments should consider all potential treatments producing a model of the given dataset. The characterisation of those treatments may rely on algorithm profiles presented in [5]. The meta-features describing the evaluation of the resulting model should consider a wide range of criteria and allow some flexibility in its comparison to the user's need. Among many usual criteria, we are giving a particular attention to meaningful information-based criteria such as described in [6] , and wish to investigate the definition of some explainability criteria following [9, 8] prediction explanations. We consider the use of some kind of meta-level feature selection to restrain those sets, and plan to rely on the experiments of [10].

Finally two of the important challenges to address will be the definition of the evolutionary metaheuristic employed, and the creation of predatory mechanisms limiting the population of meta-instances. We expect to find good candidates among genetic algorithms [2], multiobjective evolutionary algorithms [14] and memetic algorithms [7, 3].

References

[1] David W Aha. Generalizing from case studies: A case study. In *9th International Conference on Machine Learning*, pages 1–10, 1992.

[2] Jaume Bacardit and Xavier Llorà. Large-scale data mining using genetics-based machine learning. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 3(1):37–61, 2013.

[3] Fabio Caraffini, Ferrante Neri, and Lorenzo Picinali. An analysis on separability for memetic computing automatic design. *Information Sciences*, 265:1–22, 2014.

[4] Joao Gama and Pavel Brazdil. Characterization of classification algorithms. In *Progress in Artificial Intelligence*, pages 189–200. Springer, 1995.

[5] Alexandros Kalousis and Maelanie Hilario. Building algorithm profiles for prior model selection in knowledge discovery systems. *International journal of engineering intelligent systems for electrical engineering and communications*, 8(2):77–88, 2000.

[6] Igor Kononenko and Ivan Bratko. Information-based evaluation criterion for classifier's performance. *Machine Learning*, 6(1):67–80, January 1991.

[7] Jim E Smith. Coevolving memetic algorithms: a review and progress report. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 37(1):6–17, 2007.

[8] Erik Strumbelj, Zoran Bosnic, Igor Kononenko, Branko Zakotnik, and Cvetka Grasic; Kuhar. Explanation and reliability of prediction models: The case of breast cancer recurrence. *Knowledge and Information Systems*, 24(2):305–324, August 2010.

[9] Erik Strumbelj and Igor Kononenko. Towards a model independent method for explaining classification for individual instances. In *Proceedings of the*

10th International Conference on Data Warehousing and Knowledge Discovery, DaWaK '08, pages 273–282, Berlin, Heidelberg, 2008. Springer-Verlag.

[10] Ljupco Todorovski, Pavel Brazdil, and Carlos Soares. Report on the experiments with feature selection in meta-level learning. In *Proceedings of the PKDD-00 workshop on data mining, decision support, meta-learning and ILP: forum for practical problem presentation and prospective solutions*, pages 27–39. Citeseer, 2000.

[11] Ricardo Vilalta and Youssef Drissi. A perspective view and survey of meta-learning. *Artificial Intelligence Review*, 18(2):77–95, October 2002.

[12] David H Wolpert. The lack of a priori distinctions between learning algorithms. *Neural computation*, 8(7):1341–1390, 1996.

[13] Lin Xu, Frank Hutter, Jonathan Shen, Holger H Hoos, and Kevin Leyton-Brown. Satzilla2012: improved algorithm selection based on cost-sensitive classification models. *SAT Challenge 2012: Solver and Benchmark Descriptions*, pages 57–58, 2012.

[14] Aimin Zhou, Bo-Yang Qu, Hui Li, Shi-Zheng Zhao, Ponnuthurai Nagaratnam Suganthan, and Qingfu Zhang. Multiobjective evolutionary algorithms: A survey of the state of the art. *Swarm and Evolutionary Computation*, 1(1):32–49, 2011.