Preference-Based Meta-Learning using Dyad Ranking: Recommending Algorithms in Cold-Start Situations (Extended Abstract)

Dirk Schäfer[1] and Eyke Hüllermeier[2]

[1] University of Marburg, Germany
[2] Department of Computer Science, University of Paderborn, Germany
dirk.schaefer@uni-marburg.de, eyke@upb.de

Preference learning in general and label ranking in particular have been applied successfully for meta-learning problems in the past [1, 4, 3]. The benefits of incorporating additional feature descriptions of alternatives in the context of preference learning have recently been shown for the *dyad ranking* framework [6]. Additional descriptions in the form of feature vectors are known in the recommender systems domain, too, where they are typically called *side-information* and used for tackling *cold-start problems*. These problems refer to situations where preference indicators (e.g., ratings) for new users or new items are not yet available (see Figure 1). In these situations, side-information helps by putting existing and new entities into relation. In this work, we make use

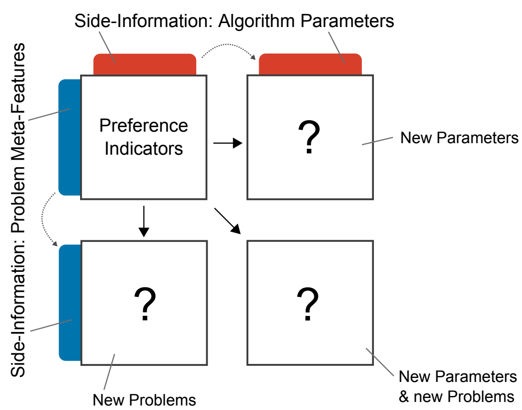

Fig. 1. Three kinds of cold-start problems are shown. They are characterized in that no preference indicators are available for algorithms or problems. Side-information can help in these situations for inferring preferences and thus recommendations.

of dyad ranking to predict a good ranking of candidate algorithms contextualized by problem instances, assuming that algorithms exhibit a representation in terms of a feature description. By generalizing over both, attributes of problems as well as algorithms, it becomes possible to tackle cold-start scenarios in which predictions are sought for algorithms that never occurred in the training data. A similar viewpoint towards meta-learning has been taken in [7, 5], where algorithm recommendation is tackled by means of collaborative filtering (CF) techniques. However, in contrast to the description of users and items in standard CF, side-information describing problems in meta-learning is usually carefully crafted [2]. As testbed, we present experimental results on the task of genetic algorithm (GA) recommendation in the cold-start situation corresponding to the lower right box in Figure 1. The (preference) meta-learning data set[3] for this experiment consists of rankings over 72 different parameterized GAs applied on the traveling salesman problem. The following leave-one-out cross validation (LOOCV) procedure over a total number of 246 examples (problems) and 72 GAs (referred to as labels) is applied: for a label A_j ($1 \leq j \leq 72$) the bilinear Plackett-Luce model [6] is trained on 245 examples and is then used to predict the ranking over all 72 labels for the left out example in two variants.

In the first variant (the "reference" situation), a method is trained on data where the label A_j *is part* of the *training set*, whereas in the second variant (the "cold start" situation) the same method is trained on data where A_j is completely *omitted*. In addition to the Kendall τ value that is used to quantify the quality of a predicted ranking in relation to a ground truth ranking, the deviation between the predicted rank of A_j and the true rank is recorded.

In the reference and the cold start situation, the Kendall τ values are almost identical. Moreover, the average deviation from the true rank in the reference case is 5.653 and in the cold-start scenario 5.712. These are first encouraging results. Future work could comprise experiments on further meta data sets and address the development of further approaches for cold-start problems.

References

1. Artur Aiguzhinov, Carlos Soares, and Ana Paula Serra. A Similarity-based Adaption of Naive Bayes for Label Ranking: Application to Metalearning for Algorithm Selection. *Planning to Learn Workshop (PlanLearn10) at ECAI*, pages 75–78, 2010.
2. Pavel Brazdil, Christophe Giraud-Carrier, Carlos Soares, and Ricardo Vilalta. *Metalearning: Applications to Data Mining.* Springer Publishing Company, Incorporated, 1st edition, 2008.
3. Johannes Fürnkranz and Eyke Hüllermeier. *Preference Learning.* Springer-Verlag New York, Inc., New York, NY, USA, 1st edition, 2010.
4. Jorge Kanda, Carlos Soares, Eduardo Hruschka, and Andre De Carvalho. A Meta-Learning Approach to Select Meta-Heuristics for the Traveling Salesman Problem Using MLP-Based Label Ranking. *19th International Conference on Neural Information Processing (ICONIP 2012)*, 7665 LNCS:488–495, 2012.
5. Mustafa Misir and Michèle Sebag. Algorithm Selection as a Collaborative Filtering Problem. Research report, INRIA, December 2013.
6. Dirk Schäfer and Eyke Hüllermeier. Dyad Ranking Using a Bilinear Plackett-Luce Model. In *Proceedings of the European Conference on Machine Learning and Principles and Practices of Knowledge Discovery in Databases.* Springer-Verlag, 2015.
7. David Stern, Horst Samulowitz, Luca Pulina, and Universita Genova. Collaborative Expert Portfolio Management. *Artificial Intelligence*, 116(3):179–184, 2010.

[3] Available at https://www.cs.uni-paderborn.de/fachgebiete/intelligente-systeme/