On Structured Output Training: Hard Cases and an Efficient Alternative*

Thomas Gärtner and Shankar Vembu

Fraunhofer IAIS, Schloß Birlinghoven, 53754 Sankt Augustin, Germany
{thomas.gaertner,shankar.vembu}@iais.fraunhofer.de

State-of-the-art structured prediction algorithms can be applied using off-the-shelf tools by implementing a joint kernel for inputs and outputs, and an algorithm for inference. The kernel is used for mapping the data to an appropriate feature space, while the inference algorithm is used for successively adding violated constraints to the optimisation problem. While this approach leads to efficient learning algorithms for many important real world problems, there are also many cases in which successively adding violated constraints is infeasible. As a simple yet relevant problem, we consider the prediction of routes (cyclic permutations) over a given set of points of interest. Solving this problem has many potential applications. For car drivers, prediction of individual routes can be used for intelligent car sharing applications or help optimise a hybrid vehicle's charge/discharge schedule. We show that state-of-the-art structured prediction algorithms cannot guarantee polynomial runtime for this output set of cyclic permutations.

Despite these hardness results, we show that efficient formulae for 'super-structure' counting can be derived and propose an alternative structured output training algorithm based on these counting formulae. By deriving 'super-structure' counting formulae for various combinatorial structures, we show that our approach subsumes many machine learning problems including multi-class, multi-label and hierarchical classification. Furthermore, our approach can be used for training complex combinatorial output sets for which the assumptions made in the literature do not hold. We empirically compare our algorithm with state-of-the-art general and special purpose algorithms on different structures. For multi-label and hierarchical classification, inference is trivial and our experiments demonstrate that our approach is competitive or better than the state-of-the-art. For route prediction, inference is hard and we focus on the training part of the algorithms, i.e., on synthetic data we compare the policy estimated by our approach to the policy estimated by SVM-Struct using approximate inference.

Reference

1. Gärtner, T., Vembu, S.: On Structured Output Training: Hard Cases and an Efficient Alternative. Machine Learning (2009) DOI: 10.1007/s10994-009-5129-3

* This is an extended abstract of an article published in the machine learning journal [1].

W. Buntine et al. (Eds.): ECML PKDD 2009, Part I, LNAI 5781, p. 7, 2009.