Similarity and kernels in machine learning

Zalán Bodó

Faculty of Mathematics and Computer Science, Babeş–Bolyai University

`Kogălniceanu 1., RO-400084, Cluj-Napoca`

Machine learning has become a very important subfield of computer science in the past decades, because in many situations, not necessarily in safety-critical scenarios, the human decision-making can be effectively substituted with it. Sometimes it is even a must to use such techniques to solve otherwise untreatable problems, consider for example content-based spam filtering.

This talk is about a special kind of machine learning algorithms, called *kernel methods*. From similarity, which is fundamental to learning [4], we arrive to kernel methods, that cover *"all geometric constructions that can be formulated in terms of angles, lengths and distances"* [3]. A kernel function, a peculiar similarity function, returns the dot product between two points in a so-called *feature space*, where data possibly gets a better representation, and allows to solve non-linear problems by means of linear problems.

The *kernelization* of an existing method, popular activity in the nineties, is presented through a simple prototype learning algorithm. The representer theorem, central to kernel methods, is only tangentially discussed to show the importance of data similarity. Then the *curse* and *blessings of dimensionality* are covered shortly: why a higher-dimensional space is sometimes better, and how kernel methods can be used for dimensionality reduction [2].

In the last part of the talk *semi-supervised* learning is introduced and along with it *data-dependent* kernels [5]. Using two cluster kernels [1], it is demonstrated that the more effective semi-supervised learning can be achieved by simply changing the (supervised) kernel function.

References

[1] Z. Bodó, L. Csató. Hierarchical and Reweighting Cluster Kernels for Semi-Supervised Learning. Int. J. of Computers, Communications & Control, Vol. V, No. 4, pp. 469-476, 2010.

[2] W. B. Johnson, J. Lindenstrauss. Extensions of Lipschitz mappings into a Hilbert space. Contemporary Mathematics, 26, pp. 189–206, 1984.

[3] B. Schölkopf, A. J. Smola. *Learning with Kernels*. MIT Press, Cambridge, Mass., 2002.

[4] R. Shahbazi, R. Raizada, S. Edelman. Similarity, kernels, and the fundamental constraints on cognition. Journal of Mathematical Psychology, vol. 70, pp. 21–34, 2016.

[5] X. Zhu, T. Rogers, R. Qian, C. Kalish. Humans perform semi-supervised classification too. AAAI, pp. 864–870, 2007.