Distance Metric Learning: Foundation, Methods and Applications

Hung Son Nguyen¹

Abstract: Distance Metric Learning (DML) is a machine learning approach that aims to learn a new distance metric from data, which improves the quality of the distance-based methods including classification, clustering, dimension reduction, kernel-based techniques, information (e.g. image) retrieval and computer vision. In case of supervised learning problems, the output of DML are embeddings, where the input data are mapped to improve a crisp or fuzzy classification process. In other words, DML focuses on searching for a transformation of the original space into a new Euclidean space, in which the distance of objects belonging to the same class becomes smaller than the distance between objects belonging to different classes.

This talk summarizes some of the recent developments in research and application of the distance metric learning method. Each DML algorithm consists of two main parts: (1) modelling the distance-based constrains in form of an optimisation problem from a given data set, and (2) proposing an efficient algorithm for solving the constructed in step (1) optimisation problem. In this talk, an overview of modelling techniques for classification, clustering and information retrieval tasks as well as the optimisation techniques including convex optimisation, different variants of gradient descent (stochastic, projected, ...), Frank-Wolfe algorithms or Bregman projection will be presented.

The lecture will also cover one of the latest research topics, which is Deep Distance Metric Learning. The goal of this research is to design a deep neural network that would be able to tell which objects in different images are visually similar and which are not? Therefore, the deep distance metric learning approach is actually a DLM method implemented by using deep neural networks. The advantage of such models is the discovery of highly representative non-linear embeddings.

Finally, the perspectives of Distance Metric Learning topic and the challenges in the near future for research and development will be presented.

Keywords: Distance Metric Learning, Deep DML

¹ Faculty of Mathematics, Informatics and Applied Mathematics University of Warsaw, Poland hungson@gmail.com, son@mimuw.edu.pl