

Visualizing non-metric similarities in multiple maps

Laurens van der Maaten · Geoffrey Hinton

Received: 25 October 2010 / Accepted: 17 November 2011 / Published online: 17 December 2011
© The Author(s) 2011. This article is published with open access at Springerlink.com

Abstract Techniques for multidimensional scaling visualize objects as points in a low-dimensional metric map. As a result, the visualizations are subject to the fundamental limitations of metric spaces. These limitations prevent multidimensional scaling from faithfully representing non-metric similarity data such as word associations or event co-occurrences. In particular, multidimensional scaling cannot faithfully represent intransitive pairwise similarities in a visualization, and it cannot faithfully visualize “central” objects. In this paper, we present an extension of a recently proposed multidimensional scaling technique called t-SNE. The extension aims to address the problems of traditional multidimensional scaling techniques when these techniques are used to visualize non-metric similarities. The new technique, called multiple maps t-SNE, alleviates these problems by constructing a collection of maps that reveal complementary structure in the similarity data. We apply multiple maps t-SNE to a large data set of word association data and to a data set of NIPS co-authorships, demonstrating its ability to successfully visualize non-metric similarities.

Keywords Multidimensional scaling · Embedding · Data visualization · Non-metric similarities

Editor: Paolo Frasconi.

L. van der Maaten (✉)
Pattern Recognition and Bioinformatics Laboratory, Delft University of Technology, Mekelweg 4,
Delft 2628 CD, The Netherlands
e-mail: lvdmaaten@gmail.com

G. Hinton
Department of Computer Science, University of Toronto, 6 King's College Road, M5S 3G4 Toronto,
ON, Canada
e-mail: hinton@cs.toronto.edu