Brandeis University
Department of Computer Science

Computer Science Seminar

Handong Zhao

Adobe Research

"Multi-modal Representation Learning, with an Application in Image/Text Understanding"

Multi-modal data are extensively accessible nowadays thanks to various types of features, sensors, and different channels. For example, the most popular commercial depth sensor Kinect uses both visible light and near-infrared sensors for depth estimation; visual question answering takes both visual images and natural language text as inputs; Marketing analyzer leverages user behavior logs as well as auxiliary text description for customer segmentation. All of them tend to facilitate better data representation in different applications. Recently there are a number of approaches proposed to deal with the representation learning in a multi-modal scenario. However, many existing approaches are either vulnerable to noises and outliers in real-world contaminated data, or inefficient to explore the rich information provided by the multi-modal data. In this talk, I will introduce some examples of my work in advancing the representation learning for multi-modal data analytics in both methodology and application, including 1) robust unsupervised learning for the contaminated multi-modal data; 2) application of multi-modal learning on image/text understanding. I will conclude this talk by describing the future research plan in the interdisciplinary field of data science.

Friday

September 13, 2019

10:00 AM

Volen 101

Hosted by: Hongfu Liu