

Supervised Feature Learning via Dependency Maximization

Chetan Tonde
Rutgers University

3/2/2016 at 02:00 pm
CBIM 22

Abstract

A key challenge in machine learning is to automatically extract relevant feature representations of data for a given task. This becomes especially formidable task for structured data like images, which are often highly structured and complex. In this thesis, we propose frameworks for supervised feature learning for structured and unstructured data, via dependency maximization.

In the first part of this dissertation we look at the problem of learning kernels for structured prediction. We present a novel framework called Twin Kernel Learning which proposes the idea of polynomial expansions of kernels, to learn kernels over structured data so as to maximize a dependency criterion called Hilbert-Schmidt Independence criterion (HSIC). We also give an efficient, matrix-decomposition based algorithm for learning these expansions and use it to learn covariance kernels of Twin Gaussian Processes. We demonstrate state-of-the-art empirical results on several synthetic and real-world datasets.

In the second part of this work, we present a novel framework for supervised dimensionality reduction based on a dependency criterion called Distance Correlation. Our framework is based on learning low-dimensional features which maximize squared sum of Distance Correlations of low dimensional features, with both, the response, and the covariates. We propose a novel algorithm to maximize our proposed objective, and also show superior empirical results over state-of-the-art on multiple datasets.