

# Abstract

Aug 16, 2006

An important problem in statistical machine learning is how to effectively model the predictions of multiple related tasks, which is known as *multi-task learning*. Different from *single-task learning* where tasks are learned separately, multi-task learning aims to jointly model those tasks. The main benefit of multi-task learning is that it can more effectively use training resources from all tasks and achieve better generalization performance when tasks are related. To be more specific, successfully addressing multi-task learning can not only allay the data paucity problem given many tasks, but also generalize to future tasks by transferring knowledge learned from existing tasks. Multiple tasks naturally exist in many applications, such as text classification, email anti-spam filtering, image classification, etc.

We present a novel probabilistic framework for multi-task learning where task relatedness is modeled using a shared structure through latent variables. Within such a framework, we study a series of important multi-task learning scenarios and propose suitable models accordingly, and show that the flexibility of the framework is achieved by allowing different assumptions about latent variables and the shared structure. In particular, we present sparsity models which are parsimonious and suitable for high-dimensional tasks; we propose the  $l_1 \circ l_p$  regularization method which is suitable for joint feature selection; we propose to use mixture models as the solution of the clusters of tasks scenario; we also extend our framework to unsupervised learning and show its connection to existing topic models. Furthermore, model selection techniques for multi-task learning are investigated since they play important roles in choosing the best joint model and generalizing to future tasks. Experiments are conducted to support our methods using both simulated datasets and real datasets from text classification, anti-spam filtering, handwritten letter recognition and collaborative filtering.