ABSTRACT

Larry Carin

Life-Long Learning with Application to Adaptive Sensing

Lawrence Carin
Duke University

In this set of talks we consider the problem of multi-task or life-long learning, for regression and classification problems. In this setting it is assumed that we have data from M learning tasks. If the objective is to simultaneously learn models for all M data sets, this is referred to as multi-task learning; if the goal is to learn a model for task M while exploiting data from M-1 previous tasks, this is referred to as life-long learning. The challenge to this problem is learning the inter-relationships between the different data sets, since this is typically unknown a priori. Employing a Bayesian setting, the learning of task inter-relations and the learning of the M models is addressed simultaneously.

The solution framework employs a hierarchical Bayesian solution. Specifically, the models associated with the M data sets share a hierarchical prior, with this manifested in two distinct ways. We first consider a formulation based on the Dirichlet process (DP), a nonparametric framework that naturally encourages sharing among the multiple tasks. The DP formulation is used in three related problems:

(i) classification based on a logistic-regression classifier,
(ii) analysis of sequential data via a DP-based MTL analysis of hidden Markov models, and
(iii) in the context of predictor variables employed within a linear regression formulation.

We also introduce a new hierarchical prior, termed the matrix stick breaking process (MSBP), which generalizes previous research on DP. Specifically, the MSBP represents a generalization of the stick-breaking view of DP, yielding a formulation that encourages predictor variables to manifest similar sharing properties across multiple tasks, without requiring that all predictors (or feature components) obey exactly the same type of sharing. The MSBP framework is particularly relevant in multi-modality sensing problems, for which one may expect that particular sensors (and their associated features) will have similar sharing properties across tasks, but the different sensors may have distinct sharing properties. In addition to demonstrating the MSBP on synthetic and real data, we examine several theoretical properties of this new modelling tool.

The Bayesian inference is implemented in two ways. We consider use of a Markov Chain Monte Carlo (MCMC) formulation, this providing a benchmark against which we may compare more computationally efficient approaches. With regard to the latter, we consider inference based on a variational Bayesian technique, which has proven to yield excellent approximate posterior estimates for many of the techniques considered here. The algorithms are examined using measured data from several sensing examples, and directions for future research are also outlined.

In addition to addressing detection, classification and regression tasks, we also address adaptive sensing with a partially observable Markov decision process (POMDP). The POMDP is formulated from a reinforcement learning perspective, and we demonstrate how the life-long learning techniques outlined above may be used within the POMDP, to place policy design within the context of all experience a given sensor has had previously.