When faced with learning a set of inter-related tasks from a limited amount of data, learning each task independently may lead to poor generalization performance. Multi-Task Learning (MTL) exploits the latent relations between tasks and overcomes data scarcity limitations by co-learning all these tasks simultaneously to offer improved performance. Although MTL has been actively investigated by the machine learning community, there are only a few studies examining the theoretical justification of this learning framework. These studies provide learning guarantees in the form of generalization error bounds which are considered as important problems in machine learning and statistical learning theory. This importance is twofold: (1) generalization bounds provide an upper-tail confidence interval for the true risk of a learning algorithm the latter of which cannot be precisely calculated due to its dependency to some unknown distribution \( P \) from which the data are drawn, (2) this type of bounds can also be employed as model selection tools, which lead to identifying more accurate learning models.

The generalization error bounds are typically expressed in terms of the empirical risk of the learning hypothesis along with a complexity measure of that hypothesis. Although different complexity measures can be used in deriving error bounds, Rademacher complexity has received considerable attention in recent years, as these complexity measures can potentially lead to tighter error bounds compared to the ones obtained by other complexity measures. However, one shortcoming of the general notion of Rademacher complexity is that it provides a global complexity estimate of the learning hypothesis space, which does not take into consideration the fact that learning algorithms, by design, pick functions belonging to a more favorable subset of this space, and they therefore yield better performing models than the worst case. To overcome the limitation of global Rademacher complexity, a more efficient notion of Rademacher complexity, the so-called local Rademacher complexity, has been considered, which leads to sharper learning bounds, and as such, compared to its global counterpart, guarantees a faster rate of convergence. Also, considering the fact that local bounds are expected to be tighter than the global ones, they can motivate better (more accurate) model selection algorithms.

While the previous MTL studies provide generalization bounds based on some other complexity measures, in this dissertation, we derive generalization error bounds for some popular kernel-based MTL hypothesis spaces based on the Local Rademacher Complexity (LRC) of those hypotheses. We show that these local bounds have faster convergence rate compared to the previous Global Rademacher Complexity (GRC)-based bounds. We then use our LRC-based MTL bounds to design a new kernel-based MTL model which benefits from strong learning guarantees. An optimization algorithm will be proposed to solve our new MTL problem. Finally, we run simulations on experimental data that compare our MTL model to some classical Multi-Task Multiple Kernel Learning (MTâ€œMKL) models designed based on the GRCs. Since the local Rademacher complexities are expected to be tighter than the global ones, our new model is also expected to show better performance compared to the GRCâ€œbased models.
Qipeng Phil Zheng, Industrial Engineering and Management Systems
Georgios, Anagnostopoulos, Electrical and Computer Engineering, Florida Institute of Technology
Petros, Xanthopoulos, Decision and Information Sciences, Stetson University

Approved for distribution by Mansooreh Mollaghasemi, Committee Chair, on June 20, 2017.

The public is welcome to attend.