

Challenges in Statistical Machine Learning

John Lafferty
School of Computer Science
Carnegie Mellon University

Abstract

A surge of research in machine learning during the past decade has led to powerful learning methods that are successfully being applied to a wide range of application domains, from search engines to computational biology and robotics. These advances have in part been achieved by refining the art and engineering practice of machine learning, paralleled by a confluence of machine learning and statistics. But an understanding of the scientific foundations and fundamental limits to learning from data can also be effectively leveraged in practice. In this overview of recent work we present some of the current technical challenges in the field of machine learning, focusing on high dimensional data and minimax rates of convergence. These challenges include understanding the role of sparsity in statistical learning, semi-supervised learning, the tradeoff between computation and risk, and structured prediction problems.