Introduction to Theory of Active Machine Learning

Robert Nowak

University of Wisconsin-Madison

Abstract

The field of Machine Learning (ML) has advanced considerably in recent years, but mostly in well-defined domains using huge amounts of human-labeled training data. Machines can recognize objects in images and translate text, but they must be trained with more images and text than a person can see in nearly a lifetime. Generating the necessary training data sets can require an enormous human effort. Active ML aims to address this issue by designing learning algorithms that automatically and adaptively select the most informative data for labeling so that human time is not wasted labeling irrelevant or trivial examples. This lecture will cover basic theory and algorithms for active machine learning, focusing on binary classification problems. I will highlight interesting issues that arise in actively learning linear classifiers in high dimensions and nonparametric active learning.