

Master's Thesis Defense Presentation
Spectral Methods for Outlier Detection in Machine Learning

by
Göker Erdoğan

Time: 31.05.2012 (Thursday) 14:00
Place: Ali Vahit Şahiner Seminar Room (Computer Engineering Building)

All interested are cordially invited.

ABSTRACT

Outliers are those instances that deviate significantly from the others in a sample. Their identification bears much importance since they carry valuable and actionable information in many real life scenarios. Spectral methods are unsupervised learning techniques that reveal low dimensional structure from high dimensional data. We analyze the spectral methods such as PCA, Laplacian eigenmaps (LEM), Kernel PCA (KPCA), Multidimensional scaling (MDS) and present a unified view. We argue that the ability of such methods to reduce dimensionality is valuable for outlier detection. Hence, we propose spectral outlier detection algorithms that combine outlier detection with spectral decomposition. The three outlier detection methods we use are Active-Outlier, Local outlier factor and Parzen windows. We combine these methods with the spectral methods of LEM and MDS to form our algorithm. We evaluate the performance of the our approach on various data sets and compare it with the performance of the outlier detection methods with no spectral transformations and with PCA. We observe that combining outlier detection methods with LEM increases the outlier detection accuracy. We discuss how the unique characteristics of LEM make it a valuable spectral method for outlier detection. We also confirm the merits of our approach on a face detection problem. Additionally, we provide an outlier detection toolbox that will be useful for researchers in this field containing the implementations of the outlier detection algorithms and the spectral methods.