Abstract

We propose a statistical model for the purpose of identifying a subgroup that has an enhanced treatment effect as well as the variables that are predictive of the subgroup membership. The need for such subgroup identification arises in clinical trials and in market segmentation analysis. By using a structured logistic-normal mixture model, our proposed framework enables us to perform a confirmatory statistical test for the existence of subgroups, and at the same time, to construct predictive scores for the subgroup membership. The inferential procedure proposed in the paper is built on the recent literature on hypothesis testing for Gaussian mixtures, but the structured logistic-normal mixture model enjoys some distinctive properties that are unavailable to the simpler Gaussian mixture models. In this talk we will present the asymptotic theory for the proposed test and provide some empirical results to demonstrate how the test can be used to reduce false positive reporting in subgroup identification. The talk is based on joint work with Dr. Juan Shen.