Extension of AUC Based on U-function

Osamu Komori
Institute of Statistical Mathematics, Japan

Abstract

We consider a two-group classification problem, where we can not assume any specific probability distributions to the observations, indicating heterogeneous property for both groups. This situation is often observed in the classification of patients who have similar but different types of diseases, such as Alzheimer's disease and other types of dementia. In this case AUC type objective function is more suitable to deal with the situation than t-statistics type objective function in which some specific distribution is assumed such as normality. We propose a class of generalized AUC based on U-function, where U is selected so that it minimizes the asymptotic variance of the coefficient of the discrimination function. We also try to extend this framework into Reproducing kernel Hilbert space so that it improves the classification accuracy.