Greedy Active Learning Algorithm for Logistic Regression Models

Hsiang-Ling Hsu
Institute of Statistics, National University of Kaohsiung

Abstract

We study a logistic model-based active learning procedure for binary classification problems, in which we adopt a batch subject selection strategy with a modified sequential experimental design method. Meanwhile, accompanying the proposed subject selection scheme, we conduct a greedy variable selection procedure such that we can update the classification model with all labeled training subjects efficiently. The proposed algorithm repeatedly performs both subject and variable selection steps until a prefixed stopping criterion is reached. Our numerical results show that the proposed procedure has competitive performances, with smaller training size and a more compact model compared with that of the classifier trained with all variables and a full data set. We also apply the proposed procedure to two data sets, wave and MAGIC gamma telescope data, to confirm the performance of our method. (This is a joint work with Yuan-Chin Ivan Chang and Ray-Bing Chen.)