

國立高雄大學統計學研究所

109 學年度書報討論題目暨摘要登記表

Extension of Greedy Active Learning Algorithm in Multi-class  
Logistic Regression Models

YI-MIN LU (盧宜敏)

摘要

Traditionally, a classification system facilitates us to predict the data, but we need large amounts of labeled training samples to construct such kind of detection system. But to obtain the labeled information might need an amount of costly resources. The active learning technology proposed by researchers comes up against this issue. In other words, we can use active learning to recruit the crucial data from a massive unlabeled data set into the labeled data set, which is utilized to construct the classifier. For binary data analysis, Hsu et al. (2019) proposed a greedy active learning algorithm, named as GATE algorithm, for the logistic regression models. The GATE algorithm not only considers the subject selection scheme, but also integrates the variable extraction step to build the classifier more efficiently. On the other hand, Li et al. (2020) proposed an active learning procedure of multi-class logistic model via individualized binary models. According to the previous literatures, we propose an active learning procedure that can determine the initial sample and also select subjects and variables simultaneously for apply multiple-class classification problems. In this work, we evaluate the classification performance of the final classification model through a simulation study.

Keywords : Active learning algorithm, Categorical data, Individualized binary estimation, Logistic regression model

指導教授簽名： 