Accepted Manuscript

Model free feature screening for ultrahigh dimensional data with responses missing at random

Peng Lai, Yiming Liu, Zhi Liu, Yi Wan

PII: S0167-9473(16)30189-X
DOI: http://dx.doi.org/10.1016/j.csda.2016.08.008
Reference: COMSTA 6330

To appear in: *Computational Statistics and Data Analysis*

Received date: 23 November 2015
Revised date: 31 May 2016
Accepted date: 10 August 2016

Please cite this article as: Lai, P., Liu, Y., Liu, Z., Wan, Y., Model free feature screening for ultrahigh dimensional data with responses missing at random. *Computational Statistics and Data Analysis* (2016), http://dx.doi.org/10.1016/j.csda.2016.08.008

Model free feature screening for ultrahigh dimensional data with responses missing at random

Peng LAI[a], Yiming LIU[b], Zhi LIU[c]*, and Yi WAN[c]

[a] *School of Mathematics and Statistics, Nanjing University of Information Science & Technology*

[b] *Devision of Mathematical Science, Nanyang Technological University*

[b] *University of Macau and UMacau Research Institute, Zhuhai*

Abstract

The paper concerns the feature screening for the ultrahigh dimensional data with responses missing at random. A model free feature screening procedure based on the inverse probability weighted methods has been proposed, where the Kolmogorov filter method is used to screen the important features under an unknown propensity score function. The suggested screening procedure has several desirable advantages. First, it has property of robust to heavy-tailed distributions of predictors and the presence of potential outliers. Second, it is a model free procedure with mild model assumptions. Third, it can deal with the missing data problem with responses missing at random. Monte Carlo simulation studies are conducted to examine the performance of the proposed procedure and a real data application is also conducted to evaluate and illustrate the proposed methods.

Keywords: ultrahigh dimensional data; missing at random; feature screening; sure screening property.

*Corresponding author: Tel: +853-88224494, Fax: +853-88222426, Email address: liuzhi@umac.mo