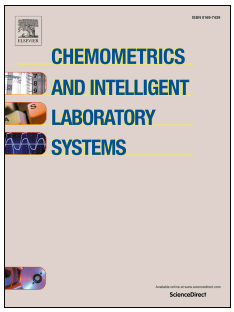
Accepted Manuscript



Survival forest with partial least squares for high dimensional censored data

Lifeng Zhou, Hong Wang, Qingsong Xu

PII: S0169-7439(18)30004-2

DOI: 10.1016/j.chemolab.2018.05.005

Reference: CHEMOM 3634

To appear in: *Chemometrics and Intelligent Laboratory Systems*

Received Date: 3 January 2018

Revised Date: 21 May 2018

Accepted Date: 22 May 2018

Please cite this article as: L. Zhou, H. Wang, Q. Xu, Survival forest with partial least squares for high dimensional censored data, *Chemometrics and Intelligent Laboratory Systems* (2018), doi: 10.1016/j.chemolab.2018.05.005.

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Survival Forest with Partial Least Squares For High Dimensional Censored Data

Lifeng Zhou[a], Hong Wang[b,*], Qingsong Xu[b]

[a]*School of Economics and Management, Changsha University, China*
[b]*School of Mathematics and Statistics, Central South University, China*

Abstract

Random forest and partial least squares have proved wide applicability in numerous contexts. However, the combination of these versatile tools has seldom been studied. Inspired by a relatively new decision tree ensemble called rotation forest, we introduce a new survival ensemble algorithm using partial least squares regression and the Buckley-James estimator within the framework of random forest. First, the approach taken to cope with the high dimensionality is to reduce the dimension by a random subspace method. Then, censored survival times are imputed by the Buckley-James estimator. After dimension reduction and time imputation, partial least squares regression is applied to extract the features. Similar to rotation forest, all extracted components are used as covariates in a bagged survival tree to predict the survival probabilities. Experimental results on a variety of simulation and real datasets demonstrate that the proposed approach is a strong competitor to other popular survival prediction models under high or ultra-high dimensional setting.

Keywords: survival ensemble, partial least squares, random survival forest, censored data, Buckley-James transformation

*Corresponding Author. School of Mathematics and Statistics, Central South University, China

Email addresses: lfzhou@csu.edu.cn (Lifeng Zhou), wh@csu.edu.cn (Hong Wang), qsxu@csu.edu.cn (Qingsong Xu)

Download English Version:

https://daneshyari.com/en/article/7561801

Download Persian Version:

https://daneshyari.com/article/7561801

Daneshyari.com