



國立高雄大學統計學研究所 巨量資料研究中心 聯合演講

【本演講將以 Google Meet 線上直播: <https://meet.google.com/rdi-zntt-bqj>】

Forward Stepwise Random Forest Analysis for Experimental Designs

講者 | 林長鋆教授 (國立中興大學統計學研究所)

時間 | 2022年6月1日(星期三)下午3:30~4:30

摘要

In experimental designs, it is usually assumed that the data follow normal distributions and the models have linear structures. In practice, experimenters may encounter different types of responses and be uncertain about model structures. If this is the case, traditional methods, such as the ANOVA and regression, are not suitable for data analysis and model selection. We introduce the random forest analysis, which is a powerful machine learning method capable of analyzing numerical and categorical data with complicated model structures. To perform model selection and factor identification with the random forest method, we propose a forward stepwise algorithm and develop Python and R codes based on minimizing the OOB error. Six examples including simulation and case studies are provided. We compare the performance of the proposed method and some frequently used analysis methods. Results show that the forward stepwise random forest analysis, in general, has a high power for identifying active factors and selects models that have high prediction accuracy.

近期演講內容: <http://www.stat.nuk.edu.tw/>

高大交通資訊: <http://www.stat.nuk.edu.tw/us.asp>

敬請公告

歡迎參加



(高大統計所)