European Network for Business and Industrial Statistics

Contribution ID: **42**   Type: **not specified**

Statistical and Machine Learning Methods in Monitoring: A Case Study on Student Performance

Monday, 16 September 2024 13:50 (20 minutes)

Despite the success of machine learning in the past several years, there has been an ongoing debate regarding the superiority of machine learning algorithms over simpler methods, particularly when working with structured, tabular data. To highlight this issue, we outline a concrete example by revisiting a case study on predictive monitoring in an educational context. In their work, the authors contrasted the performance of a simple regression model, a Bayesian multilevel regression approach, and an LSTM neural network to predict the probability of exceptional academic performance in a group of students. In this comparison, they found that the Bayesian multilevel model mostly outperformed the more complex neural network. In this work in progress, we focus on the case attributes that might have influenced the observed differences in outcomes among the models tested, and supplement the previous case study with supplementary analyses. We elaborate on characteristics that lead to comparable results derived from simpler, interpretable models in accordance with the data-generating mechanism and theoretical knowledge. We use these findings to discuss the general implications for other case studies.

Type of presentation

Talk

Classification

Mainly application

Keywords

Student monitoring, Predictive monitoring, Model complexity

Primary author: VON STACKELBERG, Paulina (University of Amsterdam)

Co-authors: Dr HUBERTS, Leo C.E. (University of New South Wales);   GOEDHART, Rob (University of Amsterdam);   Prof. DOES, Ronald J.M.M. (IBIS UvA and University of Amsterdam)

Presenter: VON STACKELBERG, Paulina (University of Amsterdam)

Session Classification: Process Monitoring and Control

Track Classification: Predictive Analytics