Statistics Seminar

Friday, March 22, 12:00 pm - 1:00 pm CT
Zoom

Machine Learning for Time Series.
Part II: Alternative approaches

Juana Sanchez, UCLA

Abstract: The Supervised Machine Learning pipeline was introduced in Part I of this seminar where we saw how traditional time series data analysis helped us determine the features that capture autocorrelations in a time series of a given length. These features create a new data set with a larger variable dimension but similar row dimensions (length depending on the lag features introduced). This data set is similar in shape to the rectangular datasets that we traditionally use for multiple regression of uncorrelated observations. This in turn allows this data set to be analyzed not only with multiple regression, but also with machine learning techniques such as random forests, gradient boosting and neural networks as if the data contained independent observations. In this Part II of the seminar, we focus on such methods and illustrate their use in forecasting applying them all to the same data set. As practitioners do, the accuracy of these forecasts is compared to benchmarks of traditional methods such as exponential smoothing and ARIMA models. The material presented will be easier to follow after viewing par I of the seminar. If time allows, we will discuss stakeholders concerned with multiple lines of products over time approach the same forecasting problem.

Bio: Dr. Juana Sanchez is a Senior Lecturer Emerita at UCLA Department of Statistics and Data Science. She is the author of "Time Series for Data Scientists" published by Cambridge University Press in 2023. (https://www.cambridge.org/gb/universitypress/subjects/statistics-probability/statistical-theory-and-methods/time-series-data-scientists-data-management-description-modeling-and-forecasting?format=HB) The book is accompanied by a web site, https://timeseriestime.org , where solutions to exercises and R programs used throughout the chapters and in the exercises, as well as errata sheet and supplements can be found. Juana has taught annually, since 1999, the course "Introduction to Time Series" to undergraduates at UCLA's Department of Statistics and Data Science. She has also published other books and numerous papers, is currently Editor of the Datasets and Stories section of the Journal of Statistics and Data Science Education, and has served in sections, chapters and committees of the ASA in the past. Throughout her career she has been actively engaged in the promotion of statistical literacy via national and international venues.
Webpage: http://statistics.ucla.edu/index.php/people1/all-faculty/7809-2/?smid=8829