School of Economics

Academic Year 2022-23
Term 1

DSA211 Statistical Learning with *R*

Instructor Name : Asst. Prof. Jing Rong GOH (Practice)
Email : jrgoh@smu.edu.sg
Office : SOE 5015.4
Tel : +65 68264850

1. Course Description

This course aims at introducing the concepts of statistics methodologies for searching analytical solutions to problems related in business with the practical use of big data. The use of statistical software R is integrated into most topics of the course. Topics include R-programming, Statistical Modelling, Linear Regression, Resampling Methods, Linear Model Selection and Regularization, Tree-based methods, and Unsupervised Learning.

2. Learning Objectives

Determining the appropriate and effective methods to analyze data is always considered as one of the most challenging decisions you need to make when you apply statistics approaches to solve real-life problems. This course equips students with appropriate statistical concepts for data analysis and R-programming techniques for implementing those cutting-edge statistical methods in solving practical issues.

Upon completion of this course, students will be able to:

- Appreciate statistical modelling
- Identify types of data
- Be familiar with statistical software R to deal with statistical issues
- Employ basic methods of exploratory data analysis
- Describe the differences between least squares and maximum likelihood parameter estimation
- Assess model accuracy
- Interpret diagnostic tests for model fit and assumption checking
- Apply Generalized Linear Regression approaches
- Use cross-validation and the Bootstrap
- Select the best linear model and apply regularization
- Use Tree-Based methods
- Understand unsupervised learning

This course is designed for students with basic knowledge of probability and mathematical statistics that is usually covered in the first introductory statistics course.

3. Pre-requisite/Co-requisite/Mutually Exclusive Course(s)
Please use the class search function at OASIS > Study > BOSS > BOSS Bidding (Plan & Bid page > Add to Cart and Perform Course Search) or the course catalogue in OASIS to check the most updated attributes of this course.

4. Assessment Methods
The assessment components for this course are:

Class Participation: 10%
Group Project: 15%
Assignments: 25%
Final Exam: 50%
Total: 100%

Class participation: Points will be lost for absences without valid reasons.

5. Course Assessment Details
Students will be assessed based on their class participation, assignments/group project, and final examination.

6. Recommended Text and Readings

- An Introduction to Statistical Learning, with Applications in R, James, Witten, Hastie, Tibshirani, 2013, New York: Springer. A PDF first and second version of the text can be downloaded at
    - http://www-bcf.usc.edu/~gareth/ISL/
    - An Introduction to Statistical Learning (statlearning.com)
- Statistical Regression and Classification, from Linear Models to Machine Learning, Matloff, 2017, CRC Press

7. University Policies

Academic Integrity
All acts of academic dishonesty (including, but not limited to, plagiarism, cheating, fabrication, facilitation of acts of academic dishonesty by others, unauthorized possession of exam questions, or tampering with the academic work of other students) are serious offences.

All work (whether oral or written) submitted for purposes of assessment must be the student's own work. Penalties for violation of the policy range from zero marks for the component assessment to expulsion, depending on the nature of the offense.

When in doubt, students should consult the instructors of the course. Details on the SMU Code of Academic Integrity may be accessed at http://www.smuscd.org/resources.html.

Copyright Notice
Please note that the course materials are meant for personal use only, namely, for the purposes of teaching, studying and research. You are strictly not permitted to make copies of or print additional copies or distribute such copies of the course materials or any parts thereof, for commercial gain or exchange. For the full copyright notice, please visit: https://smu.sg/Copyright-notice or OASIS > Campus Life & Exchange > Conduct & Discipline > University Council of Student Conduct.

Accessibility

SMU strives to make learning experiences accessible for all. If students anticipate or experience physical or academic barriers due to disability, please let the instructor know immediately. Students are also welcome to contact the university's disability services team if they have questions or concerns about academic provisions: included@smu.edu.sg.

Please be aware that the accessible tables in the seminar room should remain available for students who require them.

Emergency Preparedness for Teaching and Learning (EPTL)

Where there is an emergency that makes it infeasible to have classes on campus, classes will be conducted online via WebEx, with no disruption to the schedule. To familiarise students with the WebEx platform, part of this course may be conducted online. The instructor will inform students of which classes, if any, will be conducted as part of this EPTL initiative.

Updated on 6 June 2022