Math 4440/5440 Statistical Learning

Prerequisite: MATH 4410

Catalog Description: A survey of statistical methods for supervised and unsupervised learning tasks including linear regression, classification, nonparametric regression, tree-based methods, classification methods, and principal component analysis. Resampling methods for assessing goodness of fit and subset selection procedures. Prerequisite: MATH 4410.

Textbooks and Readings

An Introduction to Statistical Learning with Applications in R (2015). Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani, Springer, 978-1-4614-7138-7 (ebook) or 978-1-4614-7238-0.

The authors have permission from the publisher to distribute the pdf of the full text. It is available to download at this link:

http://www-bcf.usc.edu/~gareth/ISL/

On this web page you will also find the R code, and link to the R package ISLR that contains the data sets, so you may want to bookmark the web page.

The course content will closely follow the textbook.

Chapter 1: Introduction

Chapter 2: Statistical Learning

Chapter 3: Linear Regression

Chapter 4: Classification

Chapter 5: Resampling Methods

Chapter 6: Linear Model Selection and Regularization

Chapter 7: Moving Beyond Linearity

Chapter 8: Tree-Based Methods

Chapter 9: Support Vector Machines

Chapter 10: Unsupervised Learning

Software and IDE

- We will use the R statistical software, which is free to download from cran.r-project.org.
- We use the RStudio IDE, which is also free to download. https://www.rstudio.com/
- Assignments will usually be submitted online and prepared using the tools available in RStudio for integrated report writing (install the knitr package and its dependencies).
- Install the R package ISLR, which contains the data sets used in the textbook.

Assignments, Grading, and Policies

Assignments

- <u>Assignments</u> are due each week. See below for details about the format. Students should do the assignments independently, unless assigned as group work.
- Assignments in the form of online quizzes may be given occasionally.
- <u>Final Project</u>: Due at the schedule final examination.

Grading

Each assignment, quiz, and project is assigned a point value, and the grade recorded is the number of points earned. The course grade will be determined on the overall percentage score: 90% A, 80% B, 70% C, 60% D, 50% or less F.

All grades will be posted in Canvas.

Schedule and Canvas Notifications

- The details and due dates for assignments will be posted approximately one week in advance of the due date, sometimes several weeks in advance. Expect one assignment per week. The due dates will vary depending on the academic calendar (holidays, spring break, etc.). Information about the final project will be given after the spring break.
- Set your Canvas notifications to be notified immediately for any new or updated assignments, and check Canvas regularly for any announcements or discussions.

Format for submission of Assignments

- Most assignments will be submitted online in Canvas in html, Word, or pdf format. These reports are to be prepared using <u>R Markdown</u> in RStudio.
- Each student will need to install knitr package and its dependencies in order to create these reports.
- This format allows seamless integration of notes and textual material with R code and output, including graphics.
- For each report, create a markdown file (File > New File > R Markdown), edit the file, then click "Knit" to compile the report. The markdown file is for example, "myfile.Rmd" and the report file is for example "myfile.html". You submit the html report (not the Rmd source file) in Canvas.
- See the Help menu in RStudio (especially the Cheat Sheets).
- To create a Word format report instead of html, use the menu under "Knit".

Late Assignment Policy:

Unexcused late work will be assessed a penalty. The first instance is -10%. The second instance is -20%. The third and subsequent instances are penalized 50%.

Classroom Policies:

- During class, any personal devices are only to be used for MATH 4440 and not for personal web surfing or other courses. This includes tablets, laptops, cell phones, smart watches, etc.
- Acceptable uses of devices during class are reading the textbook, taking notes, or looking at MATH 4440 materials on Canvas.
- Unless otherwise instructed, class time is <u>not</u> to be used for completing assignments or quizzes.

- Cell phones are to be kept out of sight and not used during class. In case there is an urgent message expected, put your phone/notification on vibrate and please step out of the room in order to handle it. Texting or other types of messaging during class is not permitted.

General Information and Resources for Students

The Learning Commons

The **Learning Commons** provides "one-stop-shop" academic support within the Jerome Library in the areas of Academic Coaching, Supplemental Instruction, Writing Consultations, Math/Stats Tutoring, subject groups and individual assistance.

The Learning Commons is a collaborative environment designed to foster independent learning to meet the needs of any student in any course at any time in the learning process. For more information, or to make an appointment: tlc@bgsu.edu ; 419-372-2823; www.bgsu.edu/learning-commons.html .

University Libraries

The University Libraries supports the teaching, learning and research mission of BGSU by advancing scholarship and creativity through collections and user-centered services that connect faculty and students to high quality information resources. For more information, to reserve a study space or to make an appointment: http://www.bgsu.edu/library.html ;http://www.bgsu.edu/library/ask-us.html ; 419-372-6943; libhelp@bgsu.edu.

Academic Honesty Policy/Codes of Conduct

The instructor and students in this course will adhere to the University's general Codes of Conduct defined in the BGSU Student Handbook. The Code of Academic Conduct (Academic Honesty Policy) requires that students do not engage in academic dishonesty. For details, refer to the BGSU Codes of Conduct site at https://www.bgsu.edu/student-handbook/code-of-conduct.html.

The instructor and students will adhere to the general Code of Academic Conduct as outlined of the BGSU Student Handbook. Specifically, students will not cheat, fabricate, plagiarize or facilitate academic dishonesty. Students who passively engage in cheating (i.e. allowing others to cheat off of them) may receive the same consequences as the person copying. In group work, if your partner or teammates do all the work on an assignment, you should not be listed as a contributor and should receive no credit for that work. If you allow an assignment to be submitted listing you as a contributor, but you did not contribute, this is equivalent to plagiarism.

Classroom Expectations/Inclusion

Students are expected to display tolerance and respect in all communication. Comments and language should be respectful and appropriate for a college community. All comments should also follow acceptable grammar and spelling.

Disability Services

If you have a disability that I should be aware of, please notify me so that I can make arrangements to accommodate your learning needs. To get more information about your rights, contact the Office of Disability Services for Students located in 38 College Park, 419-372-8495. (http://www.bgsu.edu/disability-services.html)

Religious Holidays

It is the policy of the University to make every reasonable effort to allow students to observe their religious holidays without academic penalty. In such cases, it is the obligation of the student to provide the instructor with reasonable notice of the dates of religious holidays on which he or she will be absent. Absence from classes or examinations for religious reasons does not relieve the student of responsibility for completing required work missed. Following the necessary notification, the student should consult with the instructor to determine what appropriate alternative opportunity will be provided, allowing the student to fully complete his or her academic responsibilities. (As stated in The Academic Charter, B-II.G-4.b at:http://www.bgsu.edu/downloads/file919.pdf.

Technology Support

Provides a central point of contact for faculty, staff and students for questions, problem reports, service requests and inquiries for University computer systems and communications technologies at BGSU. Email: tsc@bgsu.edu Phone: (419) 372-0999.

Veterans

BGSU educators recognize student veterans' rights when entering and exiting the university system. If you are a student veteran, please communicate with your instructor so reasonable accommodations can be made for absence when drilling or being called to active duty. See (http://www.bgsu.edu/veteran/) for more information.