

FSRM 588: FINANCIAL DATA MINING 16:958:588:01

FALL 2021, MONDAY 6:00 PM - 9:00 PM, PH 115

1. COURSE INFORMATION

- Instructor: Zijian Guo
- Email: zijguo@stat.rutgers.edu
- **In-person class**
 - Follow the Rutgers classroom policy: wear the mask and keep social distance.
 - Course location: PH 115.
 - **Course time: 6:00pm - 9:00pm.**
- **Zoom office hour**
 - Meeting ID: 916 1368 8317 (If needed, Password: 716575)
 - Friday 1:00 pm - 2:00 pm or by appointment
- Teaching Assistant: Wei Yuan
 - Email: wy204@scarletmail.rutgers.edu
 - **R lab tutorial: 9:30am - 10:30 am, every Wednesday (via Zoom)**
 - Office hour: 10:30am - 11:30 am, every Wednesday (via Zoom)
 - Zoom link: <https://zoom.us/j/4605152013?pwd=RWdkK3gzZjRTY0dsZlBqdHBxaWQzUT09>
- *An Introduction to Statistical Learning with Applications in R*. This book is an easier version of *The Elements of Statistical Learning*. You can visit the website of the book: <http://www-bcf.usc.edu/~gareth/ISL/>.
- Recommended Reference: Textbook: *The Elements of Statistical Learning*, by Hastie, Tibshirani and Friedman. Springer, 2009, 2ed. Full text available from Springer <http://dx.doi.org/10.1007/978-0-387-84858-7> Access from campus or login via Rutgers account. You can visit the website of the book: <http://www-stat.stanford.edu/~tibs/ElemStatLearn/>.
- Software: R. Free software available at <http://www.r-project.org/>. If you go to **Manuals** on the left panel of the website, you will find a good introduction *An Introduction to R*.
- Course work: reading, assignments, quiz and a final project.
- Grades:
 - (1) Homework (40%): There will be 4-5 homework due in class.
 - (2) In-class quiz (10%): There will be 1 in-class quiz.
 - (3) Final project (50%):
 - formulate a group: three members each group
 - proposal 1 (10%) : formulate the problem
 - proposal 2 (10%): obtain some preliminary results
 - presentation (15%): present your results
 - final paper (15%): finalize your results
- Rules of the class

- (1) No late homework will be accepted.
 - (2) Students are encouraged to discuss the homework with classmates, the TA and the instructor. But each student needs to hand in an independent homework by himself/herself.
 - (3) When you send an email about this course, please use the title “FSRM 588: ”. This allows the instructor and the TA to reply to these emails with priority.
 - (4) When there is any problem with grading, please talk to the TA who grades the homework. If the problem is not resolved, submit a written request explaining the grading problem and also include a copy of your solution.
- Comments
 - (1) All students are required to read either the textbook or the recommended reference.
 - (2) The lectures will be based on the combination of the textbook, recommended reference and additional materials prepared by the instructor.

2. SYLLABUS (TENTATIVE)

- (1) **Introduction to Statistical Learning** (Chapter 1 in ESL)
- (2) **Supervised Learning: Linear Methods**
 - *Introduction to Supervised Learning* (Chapter 2 in ESL; Chapter 2 in ISL.)
 - *Linear Regression* (Chapter 3 in ESL; Chapter 3 in ISL.)
 - *Model Assessment and Selection* (Chapter 7 in ESL; Chapter 5 in ISL.)
 - *Linear Models for Classification* (Chapter 4 in ESL; Chapter 4 in ISL.)
 - *Penalized Regression Methods* (Chapter 3 in ESL; Chapter 6 in ISL.)
 - *Model Inference and Averaging* (Chapter 8 in ESL; Chapter 5 in ISL.)
- (3) **Supervised Learning: Nonparametric Methods**
 - *Basis Expansions and Regularization* (Chapter 5 in ESL; Chapter 6 in ISL.)
 - *Kernel Smoothing Methods* (Chapter 6 in ESL; Chapter 6 in ISL.)
- (4) **Supervised Learning: Advanced Methods**
 - *Additive Models, Trees* (Chapter 9 in ESL; Chapter 8 in ISL.)
 - *Random Forest* (Chapter 15 in ESL; Chapter 8 in ISL.)
 - *Support Vector Machine* (Chapter 12 in ESL; Chapter 9 in ISL.)
- (5) **Unsupervised Learning**
 - *Unsupervised Learning* (Chapter 14 in ESL; Chapter 10 in ISL.)
- (6) **Neural Network** (Chapter 11 in ESL.)
- (7) **Project Presentation**