EXST7142: Statistical Data Mining, Fall 2016

Instructor: Dr. Bin Li
Office: Office: Rm 173 Woodin Hall, LSU
Contact info: 578-1343; bli@lsu.edu
An Introduction to Statistical Learning with Applications in R, James, G., Witten, D., Hastie, T., and Tibshirani, R., (1st Ed.), Springer.

Handouts: The handouts can be purchased in Co-Op Bookstore (3960 Burbank Dr. 383-9870.)

Prerequisites: EXST 7013/7014/7015 and EXST 7034 or equivalent or consent of department head. Familiarity with matrix algebra and linear regression analysis, and working knowledge of one or more standard software packages such as R, S-plus, SAS, C/C++ or Matlab.


Computing: The lecture examples will be illustrated mainly in R. No restrictions on softwares used in project.

Final Grade: (55%) Homework, team project and presentation
(45%) Final project and exam (?)

List of topics:

- Introduction to R
- Introduction to statistical learning
- Model assessment
- Nonparametric regression and smoother (with case studies)
- Generalized additive models (with case studies)
- Classification and regression tree (CART) (with case studies)
- Ensemble methods: bagging and random forest (with case studies)
- Neural networks (with case studies)
- Clustering
- Mining association rules and market basket analysis (with case studies)
- Principal component analysis
- Case studies in data mining

Important dates:
Aug. 30 Final date for dropping without ‘W’
Aug. 31 Final date for adding course
Nov. 4 Final date for dropping courses