ISOM5535 - High Dimensional Statistics with Business Applications
Spring 2022

Instructor   Lilun DU, Assistant Professor of ISOM

Room: LSK4071

Phone: 2358-7730

Email: dulilun@ust.hk

Office Hour: by appointment

Instructional Assistant Elvis, LEE, Room LSK4065, Email: imelvis@ust.hk, Office Hour: by appointment

Prerequisites
ISOM5510 Data Analysis

Course Description
This course presents classical and modern approaches for analyzing multivariate and high dimensional data, including principal components, factor analysis, discriminant analysis, clustering. New developments on high dimensional features selection, nonparametric regression will also be covered. All these approaches will be analyzed in the context of Marketing, Finance, and other important business domains. Computational issues for both traditional and new methodologies will also be discussed.

Course Objectives
On successful completion of the course, students will be able to:

-   Use the concepts underlying the quantitative tools which deal with multivariate and high dimensional data, and put those tools together with data to draw meaningful conclusions.

-   Employ classical and advanced approaches used in multivariate and high dimensional data to different areas of business, including terminology and issues specific to finance, marketing, and E-commerce.

Syllabus
1. Preview
    -   Matrix Algebra, Multivariate Normal Distribution
2. Regression and Classification
    -   Lasso and Ridge regression
    -   Coordinate descent
    -   Regularized Logistic Regression
    -   Linear Discriminant Analysis, Quadrative Discriminant Analysis

3. Nonparametric Regression
    - Regression Splines
    - Smoothing Splines
    - Generalized Additive Model
4. Dimension Reduction
    - Principal Component Analysis: data example, basic concepts, interpretation
    - Factor Analysis: factor extraction, factor rotations, factor scores
5. Clustering
    - K-means Clustering, Hierarchical Clustering, EM Algorithm-based Clustering
6. Statistical Methods for Recommender Systems
    - Matrix Factorization
    - Alternating Least Squares Method
7. Trustworthy Online Controlled Experiments
    - A/A test, A/B test
    - Sensitivity Analysis

Course Materials

Reference Books:
1. Richar A. Johnson and Dean W.Wichern, Applied Multivariate Statistical Analysis, 6th Edition, Prentice Hall Englewood Cliffs, NJ, 2007.
2. James, G., Witten, D., Hastie, T., and Tibshirani, R., An Introduction to Statistical Learning, 1st Edition, Springer, 2013.
3. Agarwal, D. and Chen B., Statistical Methods for Recommender Systems, Cambridge University Press, 2016.
4. Kohavi, R., Tang, D., and Xu, Y., Trustworthy Online Controlled Experiments, Cambridge University Press, 2020.

Evaluation

Assignment 1 (6 students/group)   10%
Assignment 2 (6 students/group)   15%
Assignment 3 (6 students/group)   15%
Final Examination   50%
Class Participation (You are required to turn on camera for online section)   10%