IEDA 3560
Syllabus (Spring 23-24)

Ruohan Zhan
Jan 23, 2024

Predictive Analytics

Schedule:

Lectures: 1:00pm – 2:50pm, Friday. Rm 2504, Lift 25-26 (84)

Labs: 3:30pm – 5:20pm, Monday. Rm 3207, Lift 21

Instructor:

Ruohan Zhan. Email: rhzhan@ust.hk, Office: Room 5559E, Phone: 2358-7110.

TAs:

- Zezhen Ding. Email: zdingah@connect.ust.hk
- Yueying Hu. Email: yhucn@connect.ust.hk

Office hours:

- By appointment.

Course webpage: http://canvas.ust.hk
Lecture notes, tutorials, assignments, solutions, and other announcements will be posted on this site. Please check this site regularly.

Course Description: In this course, you will learn how to identify, evaluate, and capture business analytic opportunities that create value. Toward this end, you will learn basic analytic methods and analyze case studies on organizations that successfully deployed these techniques. We focus on how to use data to develop insights and predictive capabilities using machine learning, data mining, and forecasting techniques. Throughout the course, we explore the challenges that can arise in implementing analytical approaches within an organization. The course emphasizes that business analytics is a practical discipline which requires mastery of both methodology and business applications. The concepts learned in this class should help you identify opportunities in which business analytics can be used to improve performance and support important decisions. It will teach you important tools that can be used to transform data into high-impact business decisions. Lastly, it should make you alert to the ways that analytics can be used - and misused - within an organization.
Prerequisites: IEDA 2520, IEDA 2540.

Course Learning Outcomes: We have three goals in this course. The first is to help you think critically about data and the analyses based on those data - whether conducted by you or someone else. The second is to enable you to identify opportunities for creating value using predictive analytics. The third is to teach you essential tools and theory so that you can apply business analytics methods yourself. Upon the completion of this course, you will be able to:

- Master a range of tools to analyze data

- Think critically about data
- Identify opportunities for creating value using predictive analytics
- Apply these tools on real data in real applications
- Implement your methods using Python

Textbook: There is no required textbook for this course. All required class materials will be available on Canvas. However, if you want to pursue more advanced topics, here are two reference books that may be helpful:

- Machine Learning by Zhi-Hua Zhou.
- The Elements of Statistical Learning by Trevor Hastie, Robert Tibshirani, and Jerome Friedman.

Software and Coding:

- The class will use Python, a powerful and easy-to-use programming language.
- The class does assume prior programming knowledge at the level of what's been introduced in IEDA 2520 and 2540. We will provide an effective tutorial so that those without such prior knowledge will be able to catch up by self-learning.
- The class is not a programming class. Programming is a tool to implement your idea and method. Like any programming language, Python is best learned through practice. Despite the steep learning curve, students can become proficient in a few weeks with reasonable effort.

Tentative Course Outline:

1. Introduction of the course
2. Linear models and regression
3. Quality of predictions / financial analytics
4. Resampling methods: Cross-validation and the bootstrap
5. Subset selection, Ridge and lasso regressions
6. Logistic regression and linear discriminant analysis
7. Predicting outcomes / recommendation analytics
8. k-nearest neighbors and tree based methods
9. Support Vector Machines
10. Clustering
11. Principal component analysis and partial least squares

Grading: project (30%), final (40%), homework and lab (30%), participation (3%, bonus).

- The final exam will be scheduled by the university. During the exam, you will not be allowed to use books or notes. However, a specified cheat sheet will be allowed. Only specified models of calculators will be allowed. (Smart devices like iPhone or calculators with memory functions are strictly prohibited.)
- There will be project assigned toward the end of semester. The written report should consist of 1 page text presentation plus supporting analysis, charts and graphs. You also need to submit your code for the project. Details about the project will be posted on canvas right after easter break. Project presentation is scheduled during the last lecture time. Precisely, on 10 May.
- You may discuss your homework with instructor, TA, fellow students, and others. However, you are expected to write up your solutions on your own.