

The Analytics Edge

Fall 2018

Course Description

The increasing availability of data is changing the way organizations are thinking about themselves and the way they interact with the world. Data is helping improve the profit of businesses, the quality of life of individuals, the performance of sports teams and the way we network with each other. In this course you will learn how to use analytics to give you an edge. The course will expose students to real world examples of how analytics has and is being used. Through these examples you will learn how to use tools of predictive and prescriptive analytics such as linear regression, logistic regression, discrete choice models, classification and regression trees, random forests, clustering and optimization in practice. We will use the statistical software R (for predictive analytics) and the numerical computing software Julia with optimization packages (for prescriptive analytics) in the course.

Learning Objectives

1. Identify the link between data and models to help make better decisions that add value to individuals, companies and institutions
2. Describe data effectively, predict future outcomes and prescribe decisions using the tools of analytics

Measurable Outcomes

1. Develop a mathematical model from a given dataset
2. Solve the mathematical model using the tools of analytics such as linear regression, logistic regression, discrete choice models, classification and regression trees, random forests, clustering and optimization
3. Justify the decisions from the mathematical model are typically better than simple heuristics and expert opinions and how these models can leverage on human expertise

Textbook

- The Analytics Edge by Dimitris Bertsimas, Allison K. O'Hair and William R. Pulleyblank Dynamic Ideas, Belmont, Massachusetts, 2016

Notes and supplementary material will be provided each week. The students can also refer to the online course on edX offered by MIT - The Analytics Edge 15.071x. A related book for statistical learning is:

- An Introduction to Statistical Learning with Applications in R by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, Springer, 2014.

Software

We will make use of the free statistical software R for the course. Download this at <http://www.r-project.org/>.

We will also make use of the numerical computing software Julia for the course. Download this at <https://julialang.org/>.

Prerequisites

Statistics and optimization

Course Assessment

Mid-Term Test (Week 6)	30%
Competition (Week 13)	40%
Final Test (Week 14)	30%

Instructors

Karthik Natarajan. Email: karthik_natarajan@sutd.edu.sg

Stefano Galleli. Email: stefano_galleli@sutd.edu.sg

Benjamin Tan. Email: benjamin_tanwj@mymail.sutd.edu.sg

Tan Yi Long. Email: yilong_tan@mymail.sutd.edu.sg

Topics to be covered (Tentative)

Week 1	Introduction to Analytics and the Software R with Visualization
Week 2	Method: Linear Regression Predicting the quality and prices of wine (Wine analytics) Moneyball (Sports analytics)
Week 3	Method: Logistic Regression Predicting the failure of space shuttles (Challenger), Predicting the risk of coronary heart disease (Framingham Heart Study)
Week 4	Method: Multinomial Logit and Mixed Logit in Discrete Choice Predicting the Academy Award winners (Oscars) Estimating the preference for safety features in cars
Week 5	Methods: Sparsity and Model Selection Baseball (Sports) Cross-country growth regressions (Economics)
Week 6	Review and Test
Week 7	Break
Week 8	Method: Classification and Regression Trees (CART), Random Forests Forecasting Supreme Court Decisions (Law)
Week 9	Method: Logistic Regression, CART, Random Forests in Text Analytics Twitter (Social media), Enron (Email)
Week 10	Method: Clustering, Collaborative and Content Filtering Netflix, MovieLens (Recommendation systems)
Week 11	Method: Singular Value Decomposition, Censored Data Photos, Netflix, Redbook (Marital), Stanford Heart Transplant (Healthcare)
Week 12	Method: Optimization Revenue Management, Capstone Allocation
Week 13	Review and Competition
Week 14	Test