NUS SOC Summer Workshop 2021
Analytics & IoT
Web Mining
Course Information

Learning content and Teaching

😈 What will be covered during "trial" lectures?

Lecture 1: Introduction (3 hours):

- Introduction to Analytics
- Types of Decision Problem
- Statistical Learning
- Applications of Analytics
- Data Mining Process
- Introduction to Web Mining

Lecture 2: Mining Web Content I (3 hours):

- Introduction to Web Content Mining
- Web Basics
- Techniques for performing Web Content Mining
- Extracting Content from HTML Source

😈 What will be covered during the "advanced" seminars?

Lecture 3: Regression
- Simple Linear Regression
- Multi Linear Regression
- Coding Scheme for Categorical Variables
- Problems with Linear Regression

Lecture 4: Classification and Clustering
- Introduction to Classification
- Various Classification Techniques
- Assessing Model Accuracy
- Resampling Methods

Lecture 6: Mining Web Content II

- Document Object Model (DOM)

- XPath

- CSS Selectors

- Extracting Content using HTML Parser

- GUI Web Scrapers

Lecture 6: Mining Web Content III

- Web Application Design

- Mining Web Data using APIs

- Scraping using an actual browser/headless browser

Lecture 7: Recommender Systems

- Introduction to Recommender Systems

- Content-based Recommender Systems

- Collaborative Filtering Recommender System