



ISOM 5270
Big Data Analytics
Spring 2023 I

Professor	Rong Zheng, ISOM
Contact	Office: LSK 4042, Tel: 2358 7642, Email: rzheng@ust.hk, Note: begin subject: [ISOM5270]
Office Hours	By appt.
Teaching Assistant	TBC

1. Course Overview

"For every leader in the company, not just for me, there are decisions that can be made by analysis. These are the best kinds of decisions. They're fact-based decisions."

--- Amazon's Chairman, Jeff Bezos

Businesses, governments, and individuals create massive collections of data as a byproduct of their activities. Increasingly, decision-makers rely on intelligent technology to analyze data to improve decision-making. In many cases, automating the decision-making processes is required because of the sheer volume of data and the speed with which new data are generated.

The course will explain with real-world examples the uses and some details of various data analytics techniques. The emphasis is placed on understanding both the methods and their business applications. We will delve into the mechanics of how the methods work when necessary. You will establish analytical thinking to the problems and understand that the proper application of analytics is as much an art as it is a science.

After the completion of this course, the students are expected to achieve the following objectives.

1. Approach business problems data-analytically (intelligently). Think carefully & systematically about whether & how data can improve business performance. Be able to envision Big Data opportunities.
2. Know the basics of Big Data processes, techniques, & systems well enough to interact with data scientists, business analysts, marketers, and clients.
3. Gain hands-on experience in using Python to implement Big Data projects.

2. Course materials and resources

Lecture notes

Lecture notes will be the primary material for the class.

Text book:

"*Data Science for Business: What you need to know about data mining and data-analytic thinking*", by Foster Provost, Tom Fawcett, 2013 ISBN: 1449361323

Software

Orange (<https://orangedatamining.com/>): Free graphical interface platform for doing machine learning
Anaconda: Python development environment (included in Orange PC/Mac version)

3. Grading

The grade breakdown is as follows:

1. Class participation: 10%
2. Assignment (4): 60%
3. Final exam: 30%

CLASS PARTICIPATION:

This is a lecture-based course, but the participation is an essential part of the learning process in the form of active discussion. Points would depend on the *quantity* and *quality* of contribution to the discussion. Also, you are expected to follow basic classroom etiquette, including having all electronic devices turned off and refraining from chatting or doing other work or reading during class.

ASSIGNMENTS:

There are **Four** group assignments. They are hands-on data analysis mini projects. Each group can have up to **3** students. Completed assignments must be submitted in *prior* to the start of the class on the due date. The submission will be handled electronically on Canvas. Grading will include a peer evaluation.

Your team can use either Orange or Python to do assignment. Partial code will be provided if you choose to use Python.

FINAL EXAM:

The final exam will be conducted at the class time in the class time in Week 8. The exam will have a limit of 2 hrs. It is open notes. More information about the exam will be provided later.

5. Course Schedule

Date	Topic	Textbook chapter (Data Science for Business)	Due day
Class 1	Introduction to Big Data analytics. Data preprocessing.	Ch. 1,2	
Class 2	Model evaluation for supervised learning. Regression analysis. Case: Telco	Ch. 4,7,11	
Class 3	Building prediction model using decision tree learning method.	Ch. 3,5	Assignment 1 due
Class 4	Naïve bayes classification method and text analytics. Case: Trans-American Airlines	Ch. 9,10	Assignment 2 due
Class 5	Unsupervised learning method: clustering and association rule learning. Case: Dureas	Ch. 6,12	Assignment 3 due
Class 6	Prediction with nearest neighbor method. Online recommender systems. Case: Pedal Bikeshare	Ch. 6	
Class 7	Neural network and deep learning		Assignment 4 due
Class 8	Final exam		