269 Business Intelligence Technologies – Data Mining
Spring 2020

(A more technical title: Introduction to Machine Learning with Python)

University of California, Davis
Graduate School of Management

Professor Yinghui (Catherine) Yang
Room 3418, Gallagher Hall, UC Davis
530-754-5967
yiyang@ucdavis.edu; zoom link: https://zoom.us/j/3085360932

Davis (MGT&MPAc): Mon. 12:10pm-3:00pm, GH-2310, 3/30, 4/6, 4/13, 4/20, 4/27, 5/4, 5/11, 5/18, 6/1, 6/8
Sacramento: Friday 2pm-5pm, 6pm-9pm, MC-2205, 3/27, 4/10, 4/24, 5/8, 5/22
San Ramon: Saturday 9am-12pm, 1pm-4pm, BR-1502, 4/4, 4/18, 5/2, 5/16, 5/30

Course Description

Data mining is the process of converting "big data" into useful knowledge required to support decision-making. It automates the process of knowledge discovery, making us orders of magnitude more productive in our search for useful information than we would be otherwise. Virtually every business organization these days is in the process of exploring and implementing business intelligence or big data solutions to core business problems.

The course is recommended for students interested in understanding the techniques and applications of data mining and acquiring hands-on skills for making intelligent business decisions in data-rich organizations.

After taking this class, the students should be able to understand the data mining process, comprehend several popular data mining methods, use Python to address typical data mining problems with data mining and machines learning methods, and also formulate problems for a given data set.
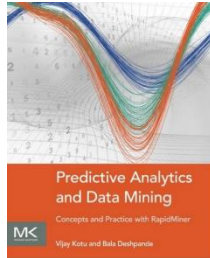
Software:

This is a very hands-on class. You will be learning Python to process data, perform data mining and machine learning tasks during the course of this class. You do not need to have programming background, but need to be comfortable with learning to program in Python. You should have access to a laptop which you can bring to class.

Prerequisite:

203B or equivalent is recommended. If you think you have adequate statistics background, please talk to me to get permission.

1

Required Textbook:

Predictive Analytics and Data Mining:
Concepts and Practice with RapidMiner

Vijay Kotu and Bala Deshpande,
2014, Morgan Kaufmann,
ISBN-10: 0128014601
ISBN-13: 978-0128014608

Grading:

Components	Grades
Class Participation	8%
9 Homework Assignments	67%
Final exam	25%

Class Participation: Class participation is evaluated by in-class exercises to see how well you have paid attention in class. In Lectures 2-9, you will work on an exercise which counts 1 points in total. If your answer is correct, you earn 1 point, if not correct, you earn 0.5, and you earn 0 if you are absent. You can't make up for a missing class unless you attend it at a different location.

Homework Assignments: There are 9 homework assignments. Because the solution to the homework will be discussed in class or posted on Canvas on the due date, late homework will not be accepted at all.

Canvas: All materials I need to hand out to you will be distributed via Canvas. All deliverables need to be submitted via Canvas.

Groups: I will post a sign-up sheet for you to self-select your group (in Homework 1). A group will work together on some homework assignment questions if any. You may also work together in class.

Final exam: We will have a take-home final exam which will be due one week after the last class.

Class Schedule: (Subject to Change)

No.	Davis	Sac	Bay	
1	3/30	3/27	4/4	Data Mining Problems, BI Process, Clustering 1
2	4/6	3/27	4/4	Clustering 2
3	4/13	4/10	4/18	Association Rules 1
4	4/20	4/10	4/18	Association Rules 2
5	4/27	4/24	5/2	Numeric Prediction 1
6	5/4	4/24	5/2	Numeric Prediction 2
7	5/11	5/8	5/16	Classification 1
8	5/18	5/8	5/16	Classification 2
9	6/1	5/22	5/30	Problem structuring, Connect 4 problems
10	No class	5/22	5/30	Analyze a data set

Notice of the Code of Academic Conduct: http://sja.ucdavis.edu/files/cac.pdf