

## **BSAN 750 Data Mining and Machine Learning (3credits)**

### **Fall 2021**

**Class Time:** T, H 11am– 12:15pm, CAPF 3056

**Instructor:** Shaobo Li, Ph.D. ([shaobo.li@ku.edu](mailto:shaobo.li@ku.edu))

Office Hours: T/H 12:15-1pm or by appointment

Office: CAPF 3166

**Teaching Assistant:** Lijun Chen ([lijun.chen@ku.edu](mailto:lijun.chen@ku.edu))

Office hours: F 9:30-10:30am or by appointment

<https://kansas.zoom.us/j/92319135308>

Meeting ID: 923 1913 5308

Passcode: 440440

**Prerequisites:** College level courses on

- Mathematics
- Probability and statistics
- Programming

**Course description:** The course introduces different machine learning techniques and how they are applied to real world problems. The course heavily relies on statistical programming R. Students are also encouraged to explore related materials and different programming languages by their own.

**Course Outcomes:**

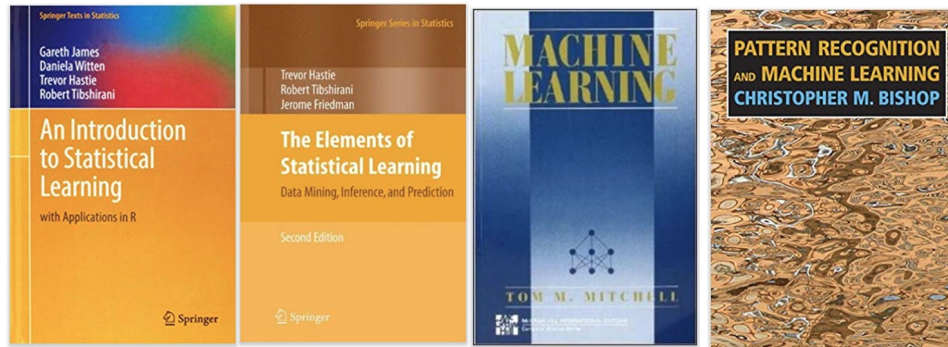
- Master level of computer programming skills
- Know popular machine learning algorithms
- Know fundamentals of machine learning theories
- Know how and why a specific algorithm works
- Know when a method works and when it does not

**Grading:**

- Group assignment 40% (Later submission will NOT be accepted)
- Online individual quizzes 40%
- Final group project 20%

**Course Format:** in person

- Delivery: mixture of lecture and lab
- Assessment: group assignment, individual quizzes and final group project
- Lecture videos also include offline videos as supplementary study materials
- You should expect heavy workload especially if you do not have experiences of coding.

**Recommended Textbook:**

**Academic Integrity:** The University of Kansas School of Business requires that all students conduct themselves in an ethical manner as outlined in the School of Business Honor Code. This dedication to ethical behavior is an integral part of the student experience, and academic integrity will be taken very seriously in this class. A breach of the honor code could potentially result in punitive punishments including expulsion from the University. More information about these expectations is available at: <http://business.ku.edu/honor-code>. If cheating is found, anyone involved will be given 0 grade, and case will be reported to university.

**Class Communication:**

We will use Blackboard to communicate. To ask questions, you may go to discussion board first to see if the same question has already been asked by someone, and if the answer is there. If not, you can send me email. To make sure that your message is not accidentally deleted as junk, please include 'BSAN750' in the **email subject** line.

**Other policies:** You need to send me early notice if you would miss a class or have possible late submission due to health-related issue, travel, or other emergencies. The instructor reserves the right to change the syllabus.

**Handling depression**

[KU Counseling and Psychological Services \(CAPS\)](#)

**More details on protecting our community from COVID-19:**

<https://protect.ku.edu/community-standards-pledge>

**Tentative Schedule**

<b>Week</b>	<b>Tuesday</b>	<b>Thursday</b>
1 Week of 8/23	Syllabus and Introduction	Basic statistics and database
2 Week of 8/30	R programming	R programming
3 Week of 9/6	<b>HW1 due by Monday</b> Exercise and review	<b>Quiz 1</b>
4 Week of 9/13	Overview of unsupervised learning Clustering	Clustering
5 Week of 9/20	Predictive analysis Nearest neighbor	Nearest neighbor
6 Week of 9/27	<b>HW2 due by Monday</b> Exercise and review	<b>Quiz 2</b>
7 Week of 10/4	Linear regression	Linear regression
8 Week of 10/11	<b>Fall break</b>	Feature selection
9 Week of 10/18	Feature selection	Feature selection
10 Week of 10/25	<b>HW3 due by Monday</b> Exercise and review	<b>Quiz 3</b>
11 Week of 11/1	Logistic regression	Logistic regression
12 Week of 11/8	Logistic regression	Logistic regression
13 Week of 11/15	<b>HW4 due by Monday</b> Exercise and review	<b>Quiz 4</b>
14 Week of 11/22	Decision tree (CART)	<b>Thanksgiving break</b>
15 Week of 11/29	Bagging and random forest	Boosting
16 Week of 12/6	Neural networks	Other topics Review
Final	<b>Final project presentation on Monday (12/13) 2-4pm</b>	

This tentative schedule may well be subject to change. A conversion to completely online format is still possible.