

BSAN 745

Advanced Machine Learning

Spring 2023



Time: Tuesday and Thursday, 11:00am-12:15pm

Location: 3056 CAPF

Instructor: Michael T. Lash, PhD

Office: 4166 CAPF

Office hours: 11:00am-1:00pm M and by appointment

Tel: (785) 864-0796

Email: Michael.Lash@ku.edu

Teaching Assistant: Shun Dong

Email: shun.dong@ku.edu

COURSE PREREQUISITES

None.

COURSE DESCRIPTION

This course takes a data-centric approach to machine learning, focusing on non-traditional data types, particularly text, networks/graphs, and images. These data types have become ubiquitous, invaluable sources of information and insight spanning the spectrum of industries and enterprise. It is therefore imperative that business analysts, data scientists, and machine learning practitioners have competency in not only dealing with these sources of data but in gleaning the treasure-trove of insights such data provide. This course equips students with the skills necessary to acquire, process, and learn from text, networks, and images, building on the foundations learned in the introduction to machine learning course (recommended, but not required). The primary emphasis in this course is text data; select lectures will be devoted to working with networks and images.

This is a 3-credit-hour semester-long course, with two 75-minute in-person class sessions on Tuesdays and Thursdays at 11:00am-12:15pm in CAPF 3056. For every credit hour students can expect to spend at least two hours per week in additional study and preparation.

LEARNING OUTCOMES

Students will learn the following:

- (1) How to acquire relevant, freely available data from the web (web scraping)
- (2) How to process non-typical data, such as text, image, and network data for use in downstream analytics tasks
- (3) Machine learning and analytics methods for text, images, and networks

TEXTS & MATERIALS

There is no textbook required for this course. All material will be made available by the instructor on Canvas, with the primary source being lectures and slides.

COURSE REQUIREMENTS AND GRADED ACTIVITIES

Your performance in this course will be evaluated in four areas: class participation, one exam, five individual homework assignments, and a group project. Class deliverables and the proportion of students earning each letter grade are provided below:

Deliverable	Percentage of grade	Proportion	Grade
Midterm Exam	15	<= 60%	A ±
Individual Homework Assignments x 5	5 x 11	<= 40%	B ±
Group Presentation	10	As need for lack of mastery	C, D, F
Group project (report)	20		

Class Participation

While class attendance and participation are not overtly graded, students should remain actively involved in the class if they expect to succeed. Please notify the instructor as soon as you are aware that you will not be able to attend one of the class sessions.

Midterm Exam

There will be a single midterm exam in this course, covering approximately two thirds (2/3rds) of the material – i.e., covering the text portion of the class. The midterm will be administered in-person and will be completed individually. Students will be allowed to use their notes and classroom materials, but not the internet.

Homework

Concepts introduced during lecture and practiced during lab sessions will be reinforced through regular homework assignments. There will be a total of five assignments, each worth 11% of student's final grade, throughout the course of the semester. The majority of these assignments will focus on the text-based portion of class (which is also the main topical thrust), with the final assignment focusing on learning from images.

Group Project

It is highly likely that you will work as a team member in your future vocation and it is therefore beneficial to learn effective teamwork practices. Also, in your future organization, you will be required to work with peers with complimentary skills for each project, whom you may or may not know beforehand. The group project will provide you with an opportunity to collectively apply what we cover during the lectures and labs to generate a comprehensive deliverable that will be presented and reported on at the

conclusion of the semester. The project will focus on text analytics, mining, and learning.

You will be assigned to a group by the instructor at the outset of the semester. As a group you will choose from among the topics (datasets) provided by the instructor. The topics will be given to groups on a first-come first-served basis (i.e., the group that selects a particular topic first will be given sole use of that topic).

After selecting a topic, groups will perform the necessary text-processing steps, basic analytics, and various text learning techniques discussed throughout the class to comprehensively assess and analyze the topic in question. At the conclusion of the semester groups will deliver a presentation on their topic, describing their processes and findings. The presentation will be worth 10% of the overall grade. Groups will also prepare a final written report describing the same. The report is worth 20% of each individual's final grade. Students will also complete and submit peer evaluations describing the contribution of each team member. Each student's grade may be adjusted based on these peer evaluations; discretion is ultimately left up to the instructor, however.

Additional details on the group project will be provided in class at the time the project is assigned.

COURSE POLICIES

Communication. Students can expect to receive weekly communications from the instructor (via the "Announcements" section on Canvas) and are responsible for the information communicated therein. Students are also responsible for any official correspondence sent through their University of Kansas email address. Students are expected to use such an email address (i.e. University address) to communicate with the instructor and can expect a response within 24-48 hours. Privacy considerations, such as federal law, may apply when using a non-University address for communication.

Timeliness. Late work will be penalized. If you have obligations that conflict with exam/assignment/report due dates, you should contact the instructor as soon as possible and ahead of the due date. Only in extremely rare circumstances will make-up exams be permitted after-the-fact. Equipment failure or conflicting time availability of team members is not an acceptable reason for turning in an assignment late. Late submissions will be penalized 25% of the assignment total each 24 hour period (i.e., after 72 hours, a late assignment will be assigned a 0).

Score/Grade appeals. If you feel that an assigned grade inaccurately reflects your performance, feel free to discuss the matter with the instructor. Timeliness is the responsibility of the student, however. As such, grades that have been assigned for more than one month are not open for appeal.

Software support. The instructor will provide software support for the software that is explicitly being used for course-work. Students will be responsible for troubleshooting virtually all potential software issues encountered when using non course-based software, including operating system issues.

ACADEMIC CODE OF HONOR

The KU School of Business seeks to develop future leaders with the highest ethical standards. It is through a strong code of conduct that we maintain mutual trust and respect among students, faculty, and staff. This code of conduct was developed by the students, faculty, and staff to articulate the School's core values and provide guidance on academic integrity. This code applies to the conduct of students, faculty, and staff at any function or academic activity conducted by the School of Business at the University of Kansas. See <https://business.ku.edu/services/student-academic-services/honor-code>.

Academic misconduct in this class will not be tolerated. Academic misconduct is not only an unethical behavior; it also deprives you of educational opportunities. There are many forms of cheating, including but not limited to:

- Possession of course material from previous sessions of this course.
- Plagiarism. All written work will be monitored by plagiarism detection software subscribed to by KU.
- Giving, receiving or unauthorized uses of aids during tests or giving test questions to another student who has not yet taken a given test.
- Following KU policy, posting content found on Blackboard/Canvas or any other class materials to a public website is a violation of intellectual property rights.
- Disrupting class is a violation of the honor code and is classified as behavioral misconduct.

ACCOMMODATIONS FOR STUDENTS WITH DISABILITIES

KU's Academic Achievement & Access Center (<http://www.achievement.ku.edu/>; 22 Strong Hall; 785-864-4064) coordinates services and accommodations for KU students with disabilities. Any students with disabilities for which they may request accommodation in KU classes should contact AAAC as soon as possible. Students should also contact the instructor privately regarding any requested accommodations.

POSITIVE CODE OF CONDUCT (R-E-S-P-E-C-T)

The School of Business promotes a Positive Code of Conduct based on seven principles:

Responsibility: We will accept and be accountable for the outcomes of our actions, work to exceed expectations, meet established deadlines for assignments, and represent the School of Business in a positive manner.

Enthusiasm: We will exhibit leadership, drive and passion in all of our pursuits both inside and outside of classroom by actively engaging in classroom discussions, events, and student organizations.

Self-Development: We will challenge ourselves to seek out opportunities for self-development, for cultivating belief in ourselves will enable us to foster respect for others.

Professional Integrity: We will be honest, trustworthy and genuine in all interactions.

Equity: We will contribute to the learning community at least as much as we receive, support equal opportunity for all students to learn, and act with respect toward the community and the environment.

Compassion: We will practice care, empathy, and understanding toward others by providing them with support or feedback when it is needed.

Teamwork: We will create an open atmosphere where everyone can contribute ideas and engage in respectful discussion without undue criticism, recognizing the value of a team in working toward a common goal.

COMMERCIAL NOTETAKING

Pursuant to the University of Kansas' Policy on Commercial Note-Taking Ventures, commercial notetaking is not permitted in this class. Lecture notes and course materials may be taken for personal use, for the purpose of mastering the course material, and may not be sold to any person or entity in any form.

Any student engaged in or contributing to the commercial exchange of notes or course materials will be subject to discipline, including academic misconduct charges, in accordance with University policy. Notetaking provided by a student volunteer for a student with a disability, as a reasonable accommodation under the ADA, is not the same as commercial notetaking and is not covered under this policy.

EQUAL OPPORTUNITY

The University of Kansas is proud of its goal to help all individuals realize their potential. To this end, the university is committed to providing an equal opportunity for all qualified individuals to be considered for employment, benefits and conditions of employment, educational programs and activities, regardless of race, religion, color, ethnicity, sex, disability, national origin, ancestry, age, status as a veteran, sexual orientation, marital status, parental status, gender identity, gender expression, or genetic information.

DIVERSITY & INCLUSION

As a premier international research university, the University of Kansas is committed to an open, diverse, and inclusive learning and working environment that nurtures the growth and development of all. KU holds steadfast in the belief that an array of values, interests, experiences, and intellectual and cultural viewpoints enrich learning and our workplace. The promotion of and support for a diverse and inclusive community of mutual respect require the engagement of the entire university.

CONCEALED CARRY LAW

Individuals who choose to carry concealed handguns are solely responsible to do so in a safe and secure manner in strict conformity with state and federal laws and KU weapons policy (refer to <http://concealedcarry.ku.edu/>). Safety measures outlined in the KU weapons policy specify that a concealed handgun:

- Must be under the constant control of the carrier.
- Must be out of view, concealed either on the body of the carrier, or backpack, purse, or bag that remains under the carrier's custody and control.
- Must be in a holster that covers the trigger area and secures any external hammer in an un-cocked position.
- Must have the safety on, and have no round in the chamber.

Instructors are allowed by Kansas Board of Regents policy, to require backpacks, purses and other bags be placed at the front of the classroom during exams and quizzes, and as such those items will not be under the constant control of the individual. Students who choose to carry a concealed handgun in a purse, backpack, or bag must review and plan each day accordingly, and are responsible for making alternate arrangements as

necessary. The university does not provide appropriate secured storage for concealed handguns. Individuals who violate the KU weapons policy may be asked to leave campus with the weapon and may face disciplinary action under the appropriate university code of conduct.

FREEDOM OF EXPRESSION

Encouraging students to engage in debate and deliberation on any topic in an effective and responsible manner is an essential part of the educational mission of a university. This course does not focus on sensitive or potentially divisive social issues, but regardless it is not the proper role of a university to insulate individuals from ideas and opinions they might find challenging, unwelcome, disagreeable, or even offensive. All university community members share a responsibility to maintain a climate of mutual respect, but concerns about civility and mutual respect should never be used to justify silencing the discussion of ideas. Therefore, consistent with principles of free expression embraced by Purdue University, the University of Chicago, and other institutions, the instructor supports freedom of expression for students to discuss any course-related topic that presents itself.

SCHEDULE (tentative)

Date	Topic	Assigned	Due
01/17	Introduction, Python Intro, Begin Processing Strings in Python		
01/19	Python Intro and Processing Strings Lab		
01/24	Continue Python Intro and Processing Strings in Python		
01/26	Processing Strings Lab	HW1	
01/31	Text Data Cleaning		
02/02	Text Data Cleaning Lab		
02/07	Term Frequency Representation		
02/09	Term Frequency Lab	HW2	HW1
02/14	Text Analytics and Sentiment		
02/16	Text Analytics and Sentiment Lab		
02/21	Topic Modeling		
02/23	Topic Modeling Lab	HW3	HW2
02/28	Intro to Text Classification		
03/02	Intro to Text Classification Lab		
03/07	Text Classification with Neural Nets		
03/09	Text Classification with Neural Nets Lab	HW4	HW3
03/14	Spring Break - No Class		
03/16	Spring Break - No Class		
03/21	Project Release and Workday	Project	
03/23	Text Analytics-a-Thon: "Big Data Brewing"		
03/28	Deep Learning with Text		
03/30	Deep Learning with Text Lab		HW4
04/04	Exam Review		
04/06	Midterm Exam		
04/11	Class choice Lecture: Image Classification or Network Analysis	HW5	

04/13	Class Choice Lab: Image Classification or Network Analysis		
04/18	Web Scraping		
04/20	Web Scraping Lab		
04/25	Project Workday		
04/27	Project Workday		
05/02	Project Presentations Day 1		HW5 Project Presentations (due Mon 11:59pm)
05/04	Project Presentations Day 2		
05/09	Finals Week: No Class		
05/11	Finals Week: No Class		Project Report