

Legal Analytics/ Legal Analytics II

MSA 8350/ LAW 7675

Spring 2021

Tuesdays, 9:00 – 11:45 a.m. [online]

Instructors:

Professor Charlotte Alexander Institute for Insight, Robinson College of Business Office and phone (Downtown): 55 Park Place; Room 1628 404-413-7708 [primary] Office and phone (Buckhead): Room 331; 404-413-7893 Email: calexander@gsu.edu	Dr. Susan Smelcer College of Law Office and phone (Downtown): 85 Park Place, Room 209; 404-413-9094 Email: ssmelcer@gsu.edu
---	---

Microsoft Teams course site:

<https://teams.microsoft.com/l/team/19%3ad9c7ee60de714fc99db70122f6fd64a1%40thread.tacv2/conversations?groupId=0547126b-4c87-470d-a098-9f5c99ee3719&tenantId=515ad73d-8d5e-4169-895c-9789dc742a70>

Office hours: Email Professor Alexander or Smelcer to set up a time to meet.

Prerequisites: [applicable to law students only] LAW 7674 or permission of instructors

Catalog Course Description: This course introduces students to the emerging field of legal analytics, which employs computational and statistical modeling, analysis, and visualization of legal data to accomplish both descriptive and predictive goals. For analytics students, the course provides an introduction to the U.S. legal system and legal reasoning, legal materials, and the problems and questions present in the law. For law students, the course offers an introduction to basic text analytic techniques, including the theory and applications of text mining, natural language processing, machine learning, and other methods for managing and analyzing unstructured data such as that found in legal documents.

Full Course Description: During the spring 2021 semester, we will examine how members of Congress describe and discuss individual and collective economic rights. In antitrust law, courts tend to evaluate whether economic conduct is good or bad using the “consumer welfare” standard. This standard is primarily concerned about ensuring that prices for consumers don’t increase when two companies merge or enter into a joint venture. But does this standard actually reflect how lawmakers have thought about competition rights and the goal of economics-related legislation?

Throughout the semester, we will try to better understand how lawmakers have conceptualized economic rights across the history of the U.S. through an examination of the Congressional Record, the official record of congressional proceedings. In the class, we will be working on preprocessing, segmenting, and classifying text, as well as modeling changes in vocabulary over time. We will also spend time working on techniques to visualize relevant elements of the text.

Course Learning Outcomes:

By the end of the semester, analytics students will be able to:

- Identify and differentiate various types of legal documents, explain the relevant legal framework, and recognize common legal problems and questions.

By the end of the semester, law students will be able to:

- Implement a quantitative, computational approach to legal questions and problems, including generating ways to approach problems that involve the gathering and analysis of data to accomplish descriptive or predictive goals.

Both groups of students will be able to:

- Implement text mining, natural language processing, machine learning, and other computational techniques as applied to legal documents and data.
- Contribute effectively to multidisciplinary teams and present their work by explaining a legal problem and its importance, methodologies employed, and outcomes.

Required Text and Downloads:

- Gareth James et al., *An Introduction to Statistical Learning with Applications in R*; free PDF download: <https://statlearning.com/ISLR%20Seventh%20Printing.pdf>
- R: <https://www.r-project.org/>
- R Studio (Desktop version): <https://rstudio.com/products/rstudio/>
- Additional readings to be provided by the instructors via Teams

Assessment:

Item	Percentage
Homework assignments – individual	15%
Homework assignments – group	15%
Mid-semester project presentation	25%
Final project presentation	35%
Peer evaluation	10%

Grading Scale:

A+	A	A -	B+	B	B-	C+	C	C-	D	F
97.0%-100%	91.0 – 96.9 %	89.5 – 90.9%	87.0 – 89.4%	83.0 – 86.9%	79.5 – 82.9%	77.0 – 79.4%	72.0 – 76.9%	69.5 – 71.9%	60.0 – 69.4%	Below 59.9%

Course Design

Course Structure: Students will learn new methods and techniques each week via asynchronous recorded lectures and example code and will attend synchronous meetings with professors as an entire class or in small teams, depending on the week. On most weeks, there will be an individual or group homework assignment. Other weeks will be set aside for project work in teams with opportunities for check-ins and guidance by the instructors. Teams will present their projects in-progress in the middle of the semester and their final project at the end. All students will watch and give feedback on all other teams' presentations.

General plan for each week:

- **Before class time**, students are responsible for watching the recorded lecture, reviewing readings and resources, learning the method(s), and running the example code on the relevant data set. Videos and code to be covered each Tuesday will be posted the previous Thursday by noon.
- **During class time**, each team will have a separate scheduled meeting time with the instructors to go over the new method(s) and code, address questions, and collaborate on the team's ongoing project. There will also be several all-class meetings, at the beginning, middle, and end of the semester, during class time in place of scheduled team meetings.
- **Between classes**, students are responsible for completing the homework assignments (group and/or individual) and advancing work on the semester-long team project. Homework is due via Teams by the start of class (Tuesdays at 9:00 a.m.) each week.

Team Management: Early in the semester, teams will form. If there are problems during the semester, the following methods will be used:

Terminating team members: As in any organization, there may be people in your group who are not willing or able to perform to the level of excellence demanded by the team. The process used to improve team member performance and/or to terminate a team member's membership in the team will involve the following steps:

- Discuss the poor performance with the individual and the standards he or she is expected to meet. As a team, document the discussion including all members' agreed-upon understanding of the standards of performance and the individual's shortfall from those standards. The document should describe what the individual must do to meet the team's standards and the time frame in which the individual will come up to the standards. This agreement should be signed by all team members, and a copy should be sent to the instructors.
- If the agreement is not met, the team, including the individual in question, will schedule a meeting with the faculty. The team will bring a copy of the contract to the meeting for the faculty and will discuss the individual's performance with the faculty. The individual will be terminated or given a final chance to improve his or her performance during that meeting and within a given time frame.
- If the performance does not improve within the time frame, the individual will be terminated from the team.
- If the individual is terminated, the individual may seek to join another team. Alternatively, he or she must complete all course work in its entirety by himself or herself from that point forward.

Resigning from a team: A student may resign from a team and switch to a different one. The work that was done while a team member is the property of both the team and the individual so all can use the work product. Faculty will facilitate the placement of the resigning person on a different team.

Peer evaluation: At the end of the semester, each student must complete an evaluation of his or her teammates' participation in the group. Peer evaluation results are worth 10% of the final course grade. Students must complete this evaluation in order to receive any points for their own group participation grades. Students may not decrease a team member's peer evaluation score without first having addressed deficiencies in that team member's performance directly with that team member, including via the process described above.

Course Policies

Microsoft Teams: The syllabus, recorded lectures and slides, code, data sets, and other course material will be posted on Microsoft Teams. Students will also use Teams to access course announcements, upload assignments, view grades, and ask/answer questions. Teams is linked to your GSU student email account. If you do not check that account regularly, forward messages and notifications to an account that you do check. All class and team meetings will also be conducted via Teams. All cameras must be on during meetings and microphones should stay on mute unless someone is speaking.

Late Work and Make-Up Policy: Any request to submit late or make-up work must be accompanied by sufficient documentation of the reason for the request. It is at the faculty's discretion whether to accept late or make-up work. Not having enough time to complete the assignment or computer problems will not, alone, excuse late work. Requests for extensions or make-up work should be made prior to the missed deadline or class, to the extent possible. We do not give additional projects or extra credit work.

University Policies

Ethics and Academic Honesty: We take issues of ethics and academic honesty very seriously. Students are expected to recognize and uphold standards of intellectual and academic integrity in all work. The University policy on academic dishonesty is spelled out in Section 1350 of the Graduate Catalog. Lack of knowledge is not an acceptable defense to any charge of academic dishonesty. Violations will result, at a minimum, in a zero for the assignment and can result in expulsion from the university. The following are instances of academic dishonesty:

- Plagiarism
- Cheating on examinations
- Unauthorized collaboration with others

- Falsification of materials
- Multiple submissions (i.e., submitting the same work for credit in more than one class)

Accommodations: Students who wish to request accommodation for a disability may do so by registering with the Access and Accommodation Center. Students may only be accommodated upon issuance by the Access and Accommodation Center of a signed Accommodation Plan and are responsible for providing a copy of that plan to instructors of all classes in which accommodations are sought.

Student Course Evaluations: Your constructive assessment of this course plays an indispensable role in shaping education at Georgia State. Upon completing the course, please take time to fill out the online course evaluation.

Course Schedule and Topics

Please note that this course schedule is a guide and deviations may and will most likely be necessary.

Video and code posted (Thursdays)	Before class	During class (Tuesdays)	Between classes (All homework due via Teams by 9:00 a.m. the following Tuesday)
Week 1. Introduction to the Course and the Text			
N/A	Readings and resources: <ul style="list-style-type: none"> • Grimmer & Stewart • Background readings on legal context [TBD] 	1/12 All-class meeting via Teams @ 9:00 a.m. Introduction to the course and the instructors. Walk-through of Teams. Review of the U.S. legislative process. Introduction to the Congressional Record text; economic rights and antitrust legal context. Complete student survey.	Team formation and team meeting scheduling [via email from instructor] HW 1 [individual]: Write a 1-2 page report in which you: <ul style="list-style-type: none"> • Draw on the background readings to generate 2-5 possible hypotheses about how and how frequently Congress discusses economic rights. For example, would you hypothesize that Democrats or Republicans address economic rights more frequently, more positively, or more negatively? What about members of Congress from urban versus rural states? Background economic conditions? • Include in the report any aspects of the legal context that are particularly confusing, interesting, or surprising to you.

Week 2. Theory and Planning: Descriptive Work, Hypothesis Formation, Measurement, Text Analytics Pipeline			
1/14	Video (Prof. Smelcer) Readings and resources [TBD]	1/19 Scheduled team meetings via Teams Discuss hypothesis formation and move toward deciding on team focus for semester-long project. Verify access to the corpus and get familiar with the structure of the Congressional Record documents.	HW 2 [group]: Turn in one report (1-2 pages) per group covering: <ul style="list-style-type: none"> • Team meeting and communication plan; • Team's choice of a hypothesis or hypotheses; and • Plan for identifying discussions of economic rights. What key words, phrases, or other textual indicators signal the presence of these discussions?
Week 3. Text Exploration and Visualization			
1/21	Video (Prof. Alexander) Readings and resources: <ul style="list-style-type: none"> • Welbers et al. • Watanabe & Müller, https://tutorials.quanteda.io/ 	1/26 Scheduled team meetings via Teams Text processing pipeline via quanteda: txt, tokenization, corpus, dfm Text frequency tables, plots, and word clouds Ngrams and phrases	
Week 4: Dictionary Lookups and Context Exploration			
1/28	Video (Prof. Alexander) Readings and resources [TBD]	2/2 Scheduled team meetings via Teams Dictionary lookups Keywords-in-context (kwic) Word embeddings	
Week 5: Regular Expressions			
2/4	Video (Jeremy Walker, GSU Data Services Librarian) Readings and resources [TBD]	2/9 Scheduled team meetings via Teams Regex applications to dictionary lookups, kwic, word embeddings	
Week 6: Segmentation			
2/11		2/16 Scheduled team meetings via Teams	
Week 7. Text Similarity Measures; TF-IDF and Keyness Measures			
2/18		2/23	

		Scheduled team meetings via Teams	
Week 8. Topic Modeling and Sentiment Analysis			
2/25		3/2 Scheduled team meetings via Teams	
Week 9. Interim Team Presentations and Workshops			
3/4		3/9 All-class meeting via Teams @ 9:00 a.m. Mid-semester peer and course evaluation	
Week 10. Modeling for Interpretation: Linear and Logistic Regression			
3/11		3/23 Scheduled team meetings via Teams	
Week 11. Modeling for Prediction and Classification: Linear and Logistic Regression			
3/25		3/30 Scheduled team meetings via Teams	
Week 12. Modeling for Prediction and Classification: KNN; Tree Models; SVM			
4/1		4/6 Scheduled team meetings via Teams	
Week 13. Practice Presentations			
N/A		4/13 Scheduled team meetings via Teams	
Week 14. Final Presentations			
N/A		4/20 All-class meeting via Teams @ 9:00 a.m.	
Peer evaluation due 4/23 @ 9:00 a.m.			