EDF 6938
Introduction to NLP in Education Research

2021 Fall
(Syllabus subject to change)

Course Information

Name: EDF 6938: Introduction to NLP in Education Research
Time: TBD
Website: https://jinnieshinufl.github.io/Course-EDF6938

Instructor Information

Instructor: Dr. Jinnie Shin, Assistant Professor in Research and Evaluation Methodology
Contact: jinnie.shin@coe.ufl.edu
Office Hour: By Appointment
Instructional approach: In person

Course Description

This project-based course is designed to introduce the basic concepts and techniques of natural language processing in education research. We will focus on text mining techniques and natural language understanding approaches commonly used in education text analysis. Students will have opportunities [1] to survey the NLP literature in the emerging AI education research context [2] to acquire theoretical backgrounds to understand the methods and [3] to gain hands-on experience in education text analysis using *Python.* The primary topics will include, but are not limited to, text vectorization, factor analysis and dimensionality reduction, supervised, unsupervised and deep learning in text analysis. **Two primary learning components** of this course include the theoretical and mathematical aspects of NLP and the hands-on programming experience in NLP analysis using *Python.*

With the facilitation of the instructor, students will actively engage in activities to establish their personal learning goals and objectives, to select project topic of their interest and capacity, to conduct NLP analyses, to provide oral presentation and research report writing. Students will also actively engage in programming activities in the lab sessions and in conducting analyses for the assignments and a final project. Accountability and assessment are addressed through peer feedback and instructors' formative evaluation on course assignments and project deliverables. The strength of the course lies in its focused and customized instruction on a specific topical area or methodological approach in quantitative research, especially natural language processing.

Assignment and Evaluation methods

Evaluation Material	N	Weights	Important Dates
Assignment	4	(10% × 4) = **40%**	TBD
Research Project	4	**50%**	
• *1-page proposal*		(5%)	
• *Introduction & Related work*		(10%)	
• *Methods*		(10%)	
• *Results*		(10%)	
• *Final Submission & Code*		(15%)	
Research Presentation	1	**10%**	

1. Assignment (Individual, may work with a partner)

Assignments will include a mix of ***theoretical questions*** and ***programming components***. The theoretical questions will allow students to demonstrate their understanding of the course materials, especially on the methodological aspects of NLP. Programming assignments will offer opportunities to gain hands-on experience to implement the NLP learning algorithms introduced in lectures. Students will be instructed to provide an executable code for each assignment in the designated GitHub repository to promote reproducible and transparent scholarly activities.

2. Research Project & Presentation (Individual)

Students are expected to follow APA 7 style in all of the submissions (https://owl.purdue.edu/owl/research_and_citation/apa_style/apa_formatting_and_style_guide/general_format.html)

Components	**Expectation**	**W**
1-page Proposal (**1 page**) Date: TBD	Students are required to provide a comprehensive introduction to the overview of the research project by explicating stating (1) the importance of topic, (2) purpose of the research and RQs, (3) proposed methods, and (4) the expected outcomes and implications of the study	5%
Introduction & Related Work (**up to 3 pages**) Date: TBD	Students are expected to position their research topic by explicitly stating (1) the importance and the contributions of the proposed research, (2) briefly introducing related works and literature, and (3) identifying how the proposed method contributes to the existing (methodological or theoretical) work.	10%
Methods **(up to 4 pages)** Date: TBD	Students are expected to explain the proposed methods to describe what *will be done* to answer the research question by explicitly stating the (1) data description, (2) analytic framework, and (3) evaluation measures.	10%
Results **(up to 4 pages)** Date: TBD	Students are expected to explain and summarize the findings, which may include (1) the data, (2) analyses performed, and (3) the evaluation results.	10%
Presentation Date: TBD	Students will prepare a short conference presentation to introduce their research design, methods, and preliminary findings (~**10 minutes**).	10%
Final Submission Date: TBD	The final paper submission should include the seven components including, the ***abstract, introduction, related-work, methods, results, conclusions, and references***. The final submission should also include the ***executable code*** complied and uploaded to the class GitHub repository.	15%

Course Grades

Overall course percent	Grade	Overall course percent	Grade
93% - 100%	A	73% - 76.9%	C
90% - 92.9%	A-	70% - 72.9%	C-
89% - 89.9%	B+	67% - 69.9%	D+
83% - 86.9%	B	63% - 66.9%	D
80% - 82.9%	B-	60% - 62.9%	D-
77% - 79.9%	C+	59.9% or less	E

Required Readings and Course Resources

All course materials (e.g., lecture notes, rubric, code) will be uploaded at https://jinnieshinufl.github.io/Course-EDF6938

Recommended Reading/Textbook:

- Jurafsky, D. & Martin J. H. (2020). *Speech and Language Processing (3rd ed.) Draft*
- Ignatow, G., & Mihalcea, R. (2016). *Text mining: A guidebook for the social sciences*. Sage Publications.
- Silge, J., & Robinson, D. (2017). *Text mining with R: A tidy approach*. O'Reilly Media, Inc.

Class Attendance, Missed Work, and Extra Credit

Requirements for class attendance and make-up exams, assignments, and other work in this course are consistent with university policies that can be found at:
https://catalog.ufl.edu/ugrad/current/regulations/info/attendance.aspx
http://gradcatalog.ufl.edu/content.php?catoid=12&navoid=2750#attendance

Online Course Evaluation

Students are expected to provide professional and respectful feedback on the quality of instruction in this course by completing course evaluations online via GatorEvals. Guidance on how to give feedback in a professional and respectful manner is available at https://gatorevals.aa.ufl.edu/students/. Students will be notified when the evaluation period opens and can complete evaluations through the email, they receive from GatorEvals, in their Canvas course menu under GatorEvals, or via https://ufl.bluera.com/ufl/. Summaries of course evaluation results are available to students at https://gatorevals.aa.ufl.edu/public-results/.

Academic Honesty

UF students are bound by The Honor Pledge which states, "We, the members of the University of Florida community, pledge to hold ourselves and our peers to the highest standards of honor and integrity by abiding by the Honor Code. On all work submitted for credit by students at the University of Florida, the following pledge is either required or implied: "On my honor, I have neither given nor received unauthorized aid in doing this assignment." The UF Student Honor Code and Student Conduct Code (http://www.dso.ufl.edu/sccr/process/student-conduct-honor-code/) specifies a number of behaviors that are in violation of this code and the possible sanctions. Furthermore, you are obligated to report any condition that facilitates academic misconduct to appropriate personnel. If you have any questions or concerns, please consult with the instructor of this class.

Accommodations for Students with Disabilities

Students with disabilities who experience learning barriers and would like to request academic accommodations should connect with the disability Resource Center by visiting disability.ufl.edu/students/get-started. It is important for students to share their accommodation letter with their instructor and discuss their access needs, as early as possible in the semester.

Counseling and Student Health

Students may occasionally have personal issues that arise in the course of pursuing higher education or that may interfere with their academic performance. If you find yourself facing problems affecting your coursework, you are encouraged to talk with an instructor and to seek confidential assistance at the University of Florida Counseling Center, 352-392-1575, or Student Mental Health Services, 352- 392-1171. Visit their web sites for more information: http://www.counsel.ufl.edu or

http://www.health.ufl.edu/shcc/smhs/index.htm#urgent. The Student Health Care Center at Shands is a satellite clinic of the main Student Health Care Center located on Fletcher Drive on campus. Student Health at Shands offers a variety of clinical services, including primary care, women's health care, immunizations, mental health care, and pharmacy services. The clinic is located on the second floor of the Dental Tower in the Health Science Center. For more information, contact the clinic at 392-0627 or check out the web site at: www.health.ufl.edu/shcc

Crisis intervention is always available 24/7 from Alachua County Crisis Center: (352) 264-6789.

Weekly Schedule

Week	Topic	Assignment	Project
1	Introduction to NLP in Education Research ▪ Lab: Introduction to Python & GitHub		
2	Projecting Text into a Vector Space (1) ▪ Lab: Introduction to Python & GitHub		
3	Projecting Text into a Vector Space (2) ▪ Lab: Text Vectorization		
4	Text and Psychometric Analyses (1) ▪ Lab: Text and Psychometric Analyses	Assignment 1	
5	Text and Psychometric Analyses (2) ▪ Lab: Text and Psychometric Analyses + Ass1 Review		
6	***Advanced Topic***: Computational Linguistic Analysis ▪ Lab: Coh-Metrix and TAACO		1-page proposal
7	Text and Supervised Learning (1) ▪ Lab: Supervised ML	Assignment 2	
8	Text and Supervised Learning (2) ▪ Lab: Supervised ML + Ass2 Review		
9	***Advanced Topic***: Automated Essay Scoring ▪ Lab: Implementing AES engines		Introduction & Related Work
10	Text and Unsupervised Learning ▪ Lab: Unsupervised ML	Assignment 3	
11	***Advanced Topic***: Topic Modelling and Summarization ▪ Lab: Topic Models + Ass3 review		Methods
12	**Project Discussion** No lab ☺	Assignment 4	
13	Language Modelling: Statistical Language Models ▪ Lab: N-gram based LM		
14	Language Modelling: Neural Language Models ▪ Lab: Tuning BERT/ELMO + Ass4 review		Results
15	**Project Presentation** No lab ☺		**Final Submission (due date: TBD)**