

CSCI-SHU 360 Machine Learning: Final Project Proposal
(When RateMyProfessors Meets Albert:
An NLP Approach to Detecting Biased Information in Online Course Evaluation Platforms)

Team members:

- Yuanxin Zhang (yz6201)
- Skyler Chen (yc4144)

Short description of the project: *[What is the topic? What is the main goal of the project?]*

The emergence of online course evaluation platforms such as RateMyProfessors has drastically altered the way students acquire course-related information. Students can submit ratings of professors' quality of teaching on this platform, which is often taken into account by future students before deciding which courses to enroll in. While ratings are crucial to reflecting students' perceptions about professors' teaching quality, students' comments usually provide much more in-depth information and details about the professors' teaching methods and quality. Moreover, these comments convey students' sentiments, which can help viewers decide whether or not to enroll, and professors to adapt their teaching techniques accordingly. Despite the fact that comments on RateMyProfessors are an invaluable source of information that could provide insights into education choices and improvements, little research has been carried out to analyze the languages used on RateMyProfessors, and what factors influence those languages used. Our research aims to address these gaps in previous studies by applying Natural Language Processing techniques to analyze comments on RateMyProfessors.

The primary objective of this study is to train a natural language processing algorithm for generating sentiment scores for comments on RateMyProfessors. In order to determine the most optimal solution, we will attempt different algorithms, such as Perceptron, SVM, KNN, etc.

The secondary objective is to determine what factors might influence the sentiment conveyed in the comments as measured by the algorithm trained. RateMyProfessors website also contains information other than the ratings and comments, such as the grade received, race, location, date of entry, etc. We conjecture that these user features might be predictive of the sentiment of the language. For example, if a student receives a higher grade, we expect that it's more likely that their comments appear to be more positive than a student who receives a lower grade. So the second goal of this project is to determine what factors influence the sentiment of the comments on the website. We plan to fit different types of regressions to answer this question, such as linear regressions, polynomial regressions, etc.

The third objective is to detect whether the information on RateMyProfessors is potentially biased by comparing it with ratings on Albert, a student information system at New York University. Platforms like RateMyProfessor usually allow a single user to enter multiple reviews, and we may frequently see extreme comments flooding the screen as a consequence. This

might create a significant bias in the validity of the ratings. As opposed to RateMyProfessors, Albert also contains student evaluations and reviews of courses, but each student is limited to writing one review per course, ensuring a more accurate reflection of students' views. Comparison between student reviews on RateMyProfessors and Albert creates a natural experiment that allows us to investigate the effect of allowing multiple entries on creating biases in course reviews. Therefore, our last goal is to examine the potential discrepancies between student reviews of RateMyProfessors and Albert.

Data set: *[What kind of data do you need? How do you plan to get it? Do you already have some sources?]*

1. NYU Albert: we use selenium and scraping to develop a tool here and the data structure and coding link is like this:

<https://github.com/nemesisXX/When-RateMyProfessors-Meets-Albert-An-NLP-Approach-to-Detecting-Biased-Information-in-Online-Course>

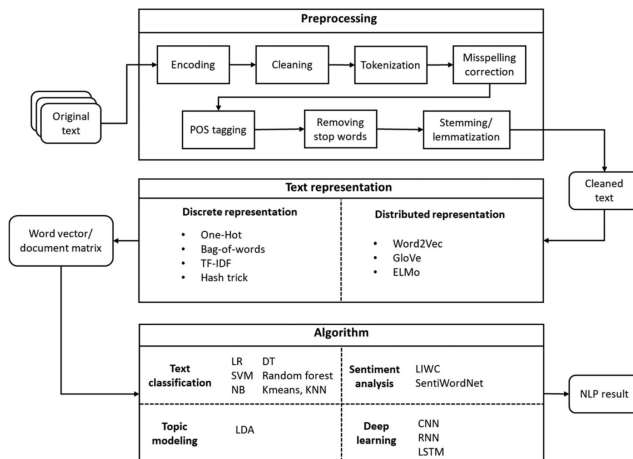
	A	B	C	D	E	F	G	H	I	J
		term	Class Description	Location	Instructor(s)	Total Enrolled	Total Response	Response Rate	evaluation	evaluation
8									Question: Overall evaluation of the course. Average: 4.8 Question: The course objectives were clearly stated. Average: 4.8 Question: The instructor provided helpful feedback on assigned class components (e.g., exams, papers). Average: 4.7 Question: The instructor was effective at helping me learn. Average: 4.8 Question: The course was well organized. Average: 4.8 Question: The course was intellectually stimulating. Average: 4.8 Question: The course was effective at helping me learn. Average: 4.7 Question: The classes were informative. Average: 4.8 Question: The course was challenging. Average: 4.8 Question: The course increased my knowledge of the subject. Average: 4.7	Question: Overall evaluation of the instructor. Average: 5.0 Question: The instructor provided an environment that was conducive to learning. Average: 4.8 Question: The instructor provided helpful feedback on assigned class components (e.g., exams, papers). Average: 4.7 Question: The instructor was effective at helping me learn. Average: 4.8 Question: The instructor encouraged student participation. Average: 4.7 Question: The instructor was effective at facilitating class discussion. Average: 4.8 Question: The instructor was open to students' questions (e.g., via e-mail and office hours). Average: 4.8 Question: The instructor created an environment that promoted the success of students with diverse backgrounds and experiences. Average: 4.7 Question: Overall evaluation of the instructor. Average: 5.0
9		Summer 2022	ANTH-UA 2 001 (2502) Washington Square	Elizabeth Filbin	18	15	83.3%			

2. Rate my professors: we cited the datasets from Prof.He, which collected 20000 professors in rate my professors at the end of 2018.
https://github.com/nemesisXX/When-RateMyProfessors-Meets-Albert-An-NLP-Approach-to-Detecting-Biased-Information-in-Online-Course/blob/main/RateMyProfessor_Sample%20data.csv

Project plan: *[What will be the main steps to develop your project? What kind of techniques do you plan to use? This does not need to be your final plan, but the earlier you have a plan, the earlier we can provide feedback and support.]*

Literature Review

- We first introduce the NLP Procedure in the following graph.(Yue Kang & Zhao Cai, 2020) This research did a literature review among all NLP operating procedure in management. But the same process can be utilized in educational NLP research as follows.



- Existing education research has already applied sentiment analysis to improve teaching and learning. As an example, Rani and Kumar (2017) proposed a Sentiment Analysis (SA) system for analyzing student feedback for sentiment, emotion, and satisfaction, which could provide administrators and teachers with a better understanding of problematic areas and take corrective action, helping universities to improve teaching and learning.

Procedure:

Data collection(Selenium, pandas): obtain data from NYU Albert

preprocessing(NLTK): Tokenization, Misspelling correction, POS tagging, removing stop words, lemmatization

Word Embedding(Word2Vec): We utilize Word2Vec to transform the text into vectors

Algorithm:

1. Sentiment analysis(LIWC)

we put one comment as a label. The calculation of one comment's label is based on the LIWC dictionary which contains all common words and their weights. In total, we calculate each word's attitude in one comment to achieve the labelization of one comment without manual work.

2. Machine Learning model(SKLearn):

- Perceptron algorithm
- SVM
- KNN
- Neural Networks
- etc

Correlation analysis:(linear regression & multivariable regression)

- Grade
- Race
- Location
- Date of entry

Albert vs RMP: potential bias analysis

1. Paired Student t-test
2. Wilcoxon signed-rank test

References:

He, Jibo (2020), "Big Data Set from RateMyProfessor.com for Professors' Teaching Evaluation", Mendeley Data, V2, doi: 10.17632/fvtfjyvw7d.2

Kang, Y., Cai, Z., Tan, C. W., Huang, Q., & Liu, H. (2020). Natural language processing (NLP) in management research: A literature review. *Journal of Management Analytics*, 7(2), 139-172.

Rani, S., & Kumar, P. (2017). A sentiment analysis system to improve teaching and learning. *Computer*, 50(5), 36-43.