COMP 7(8)-780 Natural Language Processing (Spring - 2024)

Instructor: Salim Sazzed, PhD
Time and Location: Tue and Thu, 11:20 pm-12:45 pm (Dunn Hall 123)
Email: ssazzed@memphis.edu
Phone: 757-275-5573
Office Hours: Wednesday, 2:15 PM – 3:15 PM (Dunn Hall 131)
TA: N/A

Learning Objectives

The learning objectives of this Natural Language Processing (NLP) course include establishing a foundational understanding of NLP concepts, such as language structures, syntax, semantics, and various related techniques, and delving into statistical and machine learning foundations. In addition, students will explore the practical applications of NLP, gaining hands-on experience in areas like text classification and sequence tagging while also delving into advanced topics like deep learning with neural network architectures and pre-trained language models. The course aims to equip learners with practical implementation skills, utilizing popular NLP libraries and frameworks.

Learning materials: Slides, Online Videos, Interactive Discussions, Textbooks, Online Resources, etc.

Topics (subject to change)

- Introduction to NLP (Definition and challenges of NLP, Historical overview, Applications of NLP in real-world scenarios)
- Words, Morphology, Corpus
- Lexical Semantics (Word Sense Disambiguation (WSD), Polysemy and Homonymy, Semantic relations, Word embeddings)
- HMM tagging, Part-of-Speech tagging
- Syntax and Grammar, Syntactic parsing (Parsing Algorithms)
- Text Preprocessing (Tokenization, Stop word removal, Stemming and Lemmatization, Text normalization, etc.)
- Text Representation (Bag of Words (BoW), Count vectorizer, TF-IDF, Word embedding (Word2Vec, GloVe, and FastText))
- Statistical and Machine Learning foundations (Basics of probability and statistics for NLP, Machine learning algorithms for text classification and sequence tagging)

- Deep Learning for NLP (Neural networks, Recurrent Neural Networks (RNNs), LSTM, Language modeling, etc.)
- Evaluation metrics in NLP for various problems
- NLP Applications (Sentiment analysis, Text classification, Named Entity Recognition (NER), Stylistic and Syntactic analysis of text)
- Ethical considerations (Ethical considerations in NLP: Bias, fairness, and transparency, Privacy concerns and responsible use of NLP technology)

Course Textbook

No specific books are required. However, there are some recommended books freely available online such as-

1. Speech and Language Processing (3rd ed. draft)
   Dan Jurafsky and James H. Martin
   URL: https://web.stanford.edu/~jurafsky/slp3/

You are encouraged to explore online resources, such as tutorials, blogs, and YouTube videos, to gain a better understanding of various topics/concepts.

Laptop:

Feel free to bring your laptop to class, as we will be utilizing online resources during the lectures. Also, always bring a pen and paper, as occasionally we will solve some mathematical problems.

Pre-Requisites: Participants should have strong programming abilities in Python and some knowledge of machine learning, data structures and algorithms.

Evaluation Your final grade for this course will be determined by the following averaging procedure (subject to change):

Assignments (2-4)	15 %
Monthly Exams (3) *Last week of each month (February, March, April (last class))*	40 %
Paper Presentation (1)	10 %
Two (2) Projects: - *Project proposal (mid/last week of February)* - *Project mid-evaluation (last week of March)* - *Final Presentation + Report*	35 % 5% 10% 20%
Bonus credit	5 %

Total	105%

* Note that the weight/percentage distribution of each component is subject to change due to various reasons such as final student enrollment and other factors.

**Bonus credit will be based on research components such as data annotation, project scope/complexity and class participation.

Grading
Grading: A+ ≥ 96%, A ≥ 93%, A− ≥ 90%, B+ ≥ 85%, B ≥ 80%, B− ≥ 75%, C+ ≥ 70%, C ≥ 65%, C- >= 60, A percentage below 60 will result in a failing grade (F).

Note:
i. A modified curve may be used for determining the grades at the discretion of the instructor.
ii. Due to fair grading policy of the University, I am unable to do entertain any individual request to increase marks. Although I understand, you may need a particular grade to maintain your scholarship or other things, I won't be able to do that at the end of the final exam. Study throughout the semester and you will likely perform well in the exam.

Attendance and Class Environment:
You are required to attend at **least 50%** of the classes. If your participation falls below 50%, you will not receive any marks for that component. It is important to make an effort to attend all classes, as there is a strong correlation between regular attendance and obtaining a good grade. Students are responsible for any material and content covered in missed lectures.
The learning environment will be maintained with strict adherence to guidelines. **Please refrain from engaging in conversations with fellow students or causing any other form of disturbance during the class. Any disruption may result in a penalty of a 1% reduction in marks.**

Exam:

1. Exams are closed book and closed notes.
2. No late assignment will be accepted unless well-documented reasons are presented.
3. By taking this course, you agree that any assignment turned in may undergo a review process and that the assignment may be included as a source document in Turnitin.com's restricted access database solely for the purpose of detecting plagiarism in such documents. Any assignment not submitted according to the procedures given by the instructor may be penalized or may not be accepted at all.
4. All assignment must be individual work. Plagiarizing assignments or code sharing is not permitted.
5. Any grading errors in assignments should be notified **within a week** to the TA/Instructor. We will not honor any late request for that.

Academic Integrity:

For details information on U of M code of student conduct and academic discipline procedures, please refer to https://www.memphis.edu/osa/students/academic-misconduct.php

If plagiarism or cheating occurs, the student will receive a failing grade (0) on the assignment and (at the instructor's discretion) a failing grade (F) in the course.

Especially, presenting work that is not your own without giving proper credit to the original author constitutes plagiarism and breaches the honor code. Copying source code or written content from any external source, including fellow students, is unacceptable. When permitted to use resources like the Internet, it is imperative to provide appropriate attribution. Any evidence of an honor code violation (cheating) will result in a 0 grade for the assignment/exam, and the incident will be submitted to the department/university for further review. Evidence of cheating may include a student being unable to satisfactorily answer questions asked by the instructor about a submitted solution. Cheating includes not only receiving unauthorized assistance, but also giving unauthorized assistance.

Students may still provide legitimate assistance to one another. You are encouraged to form study groups to discuss course topics. Students should avoid discussions of solutions to ongoing assignments and should not, under any circumstances, show or share code solutions for an ongoing assignment.

<u>Departmental Academic Misconduct Response Policy</u>
Adopted February 22, 2023

If an instructor discovers academic misconduct (e.g., plagiarism, cheating), they must take action to address it.

As per university policy, action must follow the Office of Student Accountability (OSA) process (which, among other things, involves reporting the incident to OSA).

The standard minimum penalty for an incident of misconduct is for the student to receive a 0 grade on the assignment/test on which the student engaged in misconduct.

The instructor may impose a greater penalty than the minimum as per their discretion.

If the instructor wishes to impose a lesser penalty than the minimum, they must receive permission from the chair before doing so.

If the OSA process concludes with the student being found responsible for academic misconduct, the instructor must notify the chair, providing the name of the student, the offense, the penalty, and the total number of prior offenses the student has on record. (OSA will provide the number of prior offenses upon request.)

<u>Students With Disabilities</u>

Qualified students with disabilities will be provided reasonable and necessary academic accommodations if determined eligible by the appropriate disability services staff at their home institution. Prior to granting disability accommodations in this course, the instructor must receive written verification of a student's eligibility for specific accommodations from the disability services staff at the home institution. It is the student's responsibility to initiate contact with their home institution's disability services staff and to follow the established procedures for having the accommodation notice sent to the instructor.