University *of* New Haven

TAGLIATELA COLLEGE OF ENGINEERING

ECECS Department

DSCI 6004
Natural Language Processing
Fall 2022
Meeting Times and Location(s): Mondays & Wednesday 12:30 – 1:45 pm, Buckman 233A
Credit Hours: 3
Vahid Behzadan, PhD – Assistant Professor
Faculty Contact Information:
Office Location: Maxcy Hall 120F or Zoom (https://unewhaven.zoom.us/my/behzadan)
Phone: (203) 47904723 Email: vbehzadan@newhaven.edu
Office Hours: Tuesday & Thursday 12pm-1pm or by request
Department Chair: Dr. Ali Golbazi agolbazi@newhaven.edu

COURSE SYLLABUS

This syllabus is informational in nature and is not an express or implied contract. It is subject to change due to unforeseen circumstances, as a result of any circumstance outside the University's control, or as other needs arise. If, in the University's sole discretion, public health conditions or any other matter affecting the health, safety, upkeep or wellbeing of our campus community or operations requires the University to make any syllabus or course changes or move to remote teaching, alternative assignments may be provided so that the learning objectives for the course, as determined by the University, can still be met. The University does not guarantee that this syllabus will not change, nor does it guarantee specific in-person, on-campus classes, activities, opportunities, or services or any other particular format, timing, or location of education, classes, activities, or services.

Course Description:

Prerequisite: DSCI 6003. Essential data science skills involved in working with unstructured data: transforming it into structured data types able to be analyzed, processed, and used for machine learning and information retrieval algorithms. Material focuses on natural language processing and classification techniques used in text mining. 3 credits
Graduate: http://catalog.newhaven.edu/index.php?catoid=23

Extended Course Description:

Natural language processing (NLP) or computational linguistics is one of the most important technologies of the information age. Applications of NLP are everywhere because people communicate almost everything in language: web search, advertising, emails, customer service, language translation, virtual agents, medical reports, etc. In the last decade, deep learning (or neural network) approaches have obtained very high performance across many different NLP tasks, using single end-to-end neural models that do not require traditional, task-specific feature engineering. In this course, students will gain a thorough introduction to cutting-edge research in Deep Learning for NLP. Through lectures, assignments and a final project, students will learn the

necessary skills to design, implement, and understand their own neural network models, using the Pytorch framework.

Required Text(s):

Speech and Language Processing by Dan Jurafsky and Jim Martin. 3rd Edition, (available online at https://web.stanford.edu/~jurafsky/slp3/)

Other:

- Jacob Eisenstein - Natural Language Processing
- Yoav Goldberg - A Primer on Neural Network Models for Natural Language Processing
- Ian Goodfellow, Yoshua Bengio, and Aaron Courville - Deep Learning

Other Materials/Supplies:

Additionsal reading materials, code samples, and datasets will be posted on Canvas.

Course Structure/Course Format/Course Objectives:

This class is offered an on-ground course, with in-person lectures, and in-person/online tutorial and discussion sessions, as well as written and online assignments, and programming projects. Active learning will constitue as much as 50% of the class. Participation will be recorded based on engagement in discussions (online/in-person), as well as submitted assignments.

Course Objectives:

To explore computational approaches for processing and anlaysis of textual data for various language processing and understanding tasks.

Student Learning Outcomes:

Upon successful completion of this course, the student should be able to:

- Parse texts into semanctic vectors for subsequent analysis and modeling
- Perform syntactic and semantic analysis of the textual data
- Build neural network models to process sequential data
- Apply machine learning techniques for information extraction and sentiment analysis.

Course Requirements & Assessment:

Please see official University of New Haven Academic Policies located in the links below:
Graduate Grading System

Assignments/Projects: There will be multiple homrework assignements, each consisting of written and coding problems in Python, as well as a final project. The homework must be completed individually without any collaboration or assistance. All submissions are online via Canvas.

The final project will be done in a group of two students and will have several deliverables including a proposal, progress update, and final report. The topic can be of one of the following types:

(i) Implementation of an algorithm/architecture that has recently been published as full papers in high quality journal/conferences and does not have its code publicly

available. Topic should be challenging enough to qualify as a group final project and cannot be on similar to topics of your class assignments, e.g., language modeling.

(ii) An extension of existing methods or a novel idea aimed to solve a particular problem. It can be something from your research project.

Use of any external source must be cited properly (acceptance is at instructor's discretion). Any violation of this policy may result in penalty from zero in assignement to failing in the course. Students may also be subject to disciplinary action by the University of New Haven (see University Policies).

Examinations: This course will have two exams: midterm and final. Both exams are closed-book and closed-notes, and must be completed individually during the designated exam sessions. The exams will include questions taken directly from the class discussions and exercises. Exams may also require handwritten code. Everything you are told or shown in class is fair game, not just the content of slides. No makeup exam will be given except for extraordinary stituations that must be communicated in written in advance.

Participation: Active learning will constitue as much as 50% of the class. Participation will be recorded based on engagement in discussions (online/in-person), as well as submitted assignments.

Grading:

Grades earned are based on your performance on homework, quizzes, exams and the final exam.

In-class Quizzes/Participation	5%
Assignments	30%
Final Project	25%
Midterm Exam	20%
Final Exam	20%
Total**	100%

**Final Grades are assigned with the following scale:

Typical Graduate Scale

Grades Scored Between			Letter Equivalent
97	to	100	A+
94	to	Less than 97	A
90	to	Less than 94	A-
87	to	Less than 90	B+
84	to	Less than 87	B
80	to	Less than 84	B-
77	to	Less than 80	C+
74	to	Less than 77	C
70	to	Less than 74	C-
Less than 70			F

The calculation of final grades is determined by the faculty member. The calculated grade in the total column in Canvas may or may not be reflective of your final grade.

Expectations:

Students should expect to spend at least 3 hours on academic studies outside, and in addition to, each hour of class time. There will be readings, homework questions/problems, and programming projects.

Attendance: Missing more than five lectures will result in an automatic "F" in the course (if you miss more than five lectures, then your course letter grade will be F). This policy may appear to be harsh, but please know that the aim of our attendance policy is by no means to add to your stress. The goal is to ensure that everyone is keeping up with the course. Many of us have the habit of procrastination. It has been repeatedly proven to me that it is less likely for my students to fall behind if they attend the lectures. Your education is of paramount importance and I care about you and your education.

- **Note:** If for reasons of illness, injury, or emergency health issues, you will not be able to regularly attend the lectures, you must email me by the end of Week 1. I will help you in any way I can. I promise together we will find an alternative method for recording your attendance.

- **Note:** If you know that you will not benefit from our strict attendance policy, please come talk with me during office hours by the end of Week 2. I will help you in any way I can. In particular, I can adjust the grading scale and create alternative midterm exams and a special comprehensive final exam for you if you do not want to regularly attend the lectures. But, if that is what you want, you must contact me by the end of Week 2.

Late Work: Assignments turned in late may be accepted with a grade penalty, if the solutions are not distributed yet. This is completely at the discretion of the instructor, as the goal is to balance learning and fairness.

Missed Work: Exams may be made up in only the most unavoidable situations (at the discretion of the instructor). A formal excused absence (such as a note from Health Services or a healthcare provider) will be required before you can make up a missed exam.

Individual Work: Students must work individually on assignments and projects unless specifically allowed to work in groups by the instructor. Any work taken from the internet must be cited properly (acceptance of code taken from elsewhere is at the discretion of the instructor) or will be considered plagiarism. Failure to adhere to this policy will result in penalties ranging from a zero on the assignment to a zero in the final grade. Students may also be subject to disciplinary action by the University of New Haven (see University Policies).

TCoE Academic Lab reservation form

As a TCoE student, you have access to reserve academic lab spaces for academic purposes where you need access to specific equipment. Example approved uses might include time for a team meeting to finish a team project or a study-session with a TA. For more information or to submit your reservation, please visit: https://forms.office.com/r/EUeJT36ZFr

Course Outline/Schedule:

Date	Topic/Note

Week 1 (8/29)	Introduction – Text Processing
Week 2 (9/5)	Word Vectors
Week 3 (9/12)	Backprop and Neural Networks
Week 4 (9/19)	Linguistic Structure: Dependency Parsing
Week 5 (9/26)	Recurrent Neural Networks and Language Models
Week 6 (10/3)	Vanishing Gradients, Fancy RNNs, Seq2Seq
Week 7 (10/10)	Machine Translation, Attention, Subword Models
Week 8 (10/17)	Midterm Exam
Week 9 (10/24)	ConvNets for NLP – Final Project Announcement
Week 10 (10/31)	Transformers and Self-Attention
Week 11 (11/7)	Contextual Representations and Pretraining
Week 12 (11/14)	Question Answering and Chatbots
Week 13 (11/21)	Natural Language Generation - Thanksgiving Break
Week 14 (11/28)	Safety, Bias, and Fairness
Week 15 (12/5)	Catch-up Week / Project Presentations
Week 16 (12/12)	Final Exam Review / Final Exam

Diversity Statement

The University of New Haven embraces diversity and recognizes our responsibility to foster a diverse, inclusive, and welcoming environment in which all members of the Charger community of all backgrounds and identities can learn, work, and live together. We benefit from the academic, social, and cultural developments that arise from a diverse campus that is committed to equity, inclusion, belonging, and accountability.

We have a responsibility as a community and as individuals to address and remove barriers, achieve success, and sustain a culture of inclusivity, empathy, kindness, and compassion. We encourage, welcome, and embrace participation in ongoing dialogue, engagement, and education to critically examine and thoughtfully respond to the changing realities of our community. Diversity, equity, inclusion, acceptance, and belonging enrich the Charger community and are instrumental to institutional success and fulfillment of the University mission.

Reporting Bias Incidents

At the University of New Haven, there is an expectation that all community members are committed to creating and supporting a climate which promotes civility, mutual respect, and open-mindedness. There also exists an understanding that with the freedom of expression comes the responsibility to support community members' right to live and work in an environment free from harassment and fear. It is expected that all members of the University community will engage in anti-bias behavior and refrain from actions that intimidate, humiliate, or demean persons or groups or that undermine their security or self-esteem.

If you have an immediate safety concern for yourself or others, and/or believe someone poses an immediate threat to themselves or others, please contact University Police at 203-932-7070 or call 911. Community members can report bias-motivated incidents by completing the form at **www.newhaven.edu/biasreporting**. Community members are encouraged to complete this

form if they are the target of bias or harassing behaviors, witness such behaviors, or gain knowledge of these behaviors occurring within the University community. All matters concerning bias and harassment will be handled by the Dean of Students Office and Human Resources Office.

University-wide Academic Policies

A continually-updated list of University-wide academic policies and descriptions of key university student resources, can be found on Canvas. You can access them by simply clicking on the (?) help button.

The University-wide academic policies include (but are not limited to) the University's attendance policy, procedures for both adding / dropping a course and course withdrawals, an explanation for the sorts of circumstances where incomplete (INC) grades could be considered by the faculty, and the academic integrity policy (among others). Also in this location you will find information regarding the process for reporting bias and topics related to our maintaining a positive learning environment (including, but not limited to, discrimination and sexual misconduct).

The list of key university student resources to enable learning include (but are not limited to) the University's Center for Student Success, Writing Center, Center for Learning Resources, and the Accessibility Resource Center.

Course Delivery Options

- **On-Ground:** Fully on-ground course with every student meeting in-person.

University of New Haven

UNIVERSITY STUDENT SUPPORT SERVICES

The University recognizes that students can often use some help outside of class and offers academic assistance through several offices.

Accessibility Resources Center

The University of New Haven seeks to maintain a supportive academic environment for all students inclusive of those with disabilities, chronic health-related conditions or military service-connected disorders. If you feel that you may need reasonable accommodations to enable your full participation in this course, please provide me with your Verification of Reasonable Accommodations letter or contact the Accessibility Resources Center to begin the process to ensure that accommodations can be made available to you. Reasonable accommodations are not made without written documentation from the Accessibility Resources Center. The Accessibility Resources Center is located in Sheffield Hall on the ground floor in the rear of the building, and can be reached by email at ARC@newhaven.edu or by phone at (203) 932-7332.

Center for Learning Resources (CLR)

The Center for Learning Resources (CLR), located in the Peterson Library, provides academic content support to the students of the University of New Haven using metacognitive strategies that help students become aware of and learn to apply optimal learning processes in the pursuit of creating independent learners. CLR tutors focus sessions on discussions of concepts and processes and typically use external examples to help students grasp and apply the material. We offer both in-person and online tutoring. To make an appointment, call us at 203-932-7215, write to us at clr@newhaven.edu, or download the Navigate app.

Center for Student Success (CSS)

The Center for Student Success can help you refine your student skills and develop new academic strategies. They assist with enhancing your time management and organizational skills. CSS staff provide understanding of your GPA, degree audit, and transcripts, and answer general questions about academic policies. They also can connect you to campus resources and assist you with resolving issues as they arise. During registration periods, CSS advisors work in conjunction your faculty advisor to provide assistance with the advising and registration process. Finally, at various points throughout the semester, CSS works to provide students with progress reports from their instructors.

Counseling & Psychological Services (CAPS)

CAPS offers confidential, free services in order to support student mental health and wellbeing. The services include individual and group therapy, support groups, consultations, and 24/7 crisis support. We are available in person at Charger Plaza and remotely, and are in the office M-F, 8:30-4:30. Please call us to schedule an appointment or with any questions at 203-932-7333; you

can also schedule online. If you experience a mental health crisis after hours, you can call our main number for support.

Myatt Center

The Myatt Center for Diversity and Inclusion is committed to creating a multicultural environment through intentional education, campus community engagement, and valuing the unique identities of each member of the Charger Community. Our commitment to diversity is driven by the core values of connection, belonging, inclusivity, equity, acceptance, and accountability. The Myatt Center's focus is to create a respectful and inclusive environment based our awareness and ability to engage with others who are different on many levels including ethnicity, race, sexual orientation, gender, military, religious belief, and life experiences.

Marvin K. Peterson Library

The Library provides access to online databases, e-books, e-journals, electronic U.S. Government Documents, print books, educational games, and audiovisual materials. A search can be conductedthrough all these resources at once by using the search box "Articles, Books, & More."

The Library provides three floors with individual quiet study space, collaborative group study space, study rooms with technology, whiteboards, Dell desktops, iMacs, scanners, and printers. The entire library is a wireless zone.

Librarians assist in locating relevant sources of information for research papers, thesis, honors thesis, and other projects. Librarians answer general reference questions and help with effectively evaluating sources of information. Help is available through a Chat Service, 24/7 Ask a Librarian Service, a Zoom Reference Service, and by E-Mail. Complete the Research Consultation Form to arrange a time convenient for you.

LibGuides are created to assist students with research. They contain an overview of resources available through the library, as well as tutorials, subject guides, and course specific guides.

University Writing Center

The mission of the Writing Center is to provide high-quality tutoring to undergraduate and graduate students as they write for a wide range of purposes and audiences. Tutors are undergraduate and graduate students who are majoring in a variety of fields across the University. We are here to work with you at any stage in the writing process; just bring in your assignment, your ideas, and any writing you've done so far. To make an appointment, you can register for an account with our scheduling site https://newhaven.mywconline.com or visit us in person in the lower level of the library.