



INFO 7610 Topic Nat Lang Engr Meth/Tools

[FALL 2024]

Course Information

Course Title: Special Topics in Natural Language Engineering Methods and Tools

Course Number: INFO 7610

Term and Year: Fall 2024

Credit Hour: 4SH

Instructor Information

Full Name: Subrata Das

Email Address: s.das@northeastern.edu

Course Prerequisites

INFO 6205 with a minimum grade of B-

Course Description

The proliferation of unstructured textual data, such as documents, reviews, social media posts, emails, articles, contracts, and chats and notes, will put Natural Language Processing (NLP) in-depth right at the center of the next set of data science challenges. NLP is a subfield of Artificial Intelligence (AI) to give machines the ability to read, understand and derive meaning from human languages expressed in textual data. The computational aspect of NLP involves prompt engineering and fine-tuning of Large Language Models (LLMs) as well as applying algorithms to extract actionable insights (e.g., customer sentiment and document topic and class) and to enhance business process efficiency (e.g., chatbot, document summarization, image captioning, and key information extraction). In this class, you will learn the most fundamental and cutting-edge tools, techniques and applications of NLP to get ready for the challenging job market. Prof. Subrata Das has over 20 years of experience in both academia and industry applying AI, NLP, and machine and deep learning (ML/DL) techniques. The class will be very hands-on with lots of practical exercises and assignments and two projects. The course materials will be provided from Prof. Das' books

Course Schedule/Topics Covered.

Week	Topic	Content	Assignment
1	Introduction and Tools	History of Symbolic AI, NLP, and ML/DL, Applications of NLP, Python Programming, NLTK, Spacy and Gensim Tools, Tensorflow/Keras	Reading and Programming
2-3	Syntactic and Semantic Analyses	Tokenization, Lemmatization, Stemming, Morphological Analysis, Part-of-Speech Tagging, Parsing, Named Entity, Dates, Location Recognition, Coreference Resolution	Assignment 1
4-5	NLP and ML (1)	Supervised Document Classification, Bag of Words Approach, Naïve Bayesian Classifier (NBC), Support Vector Machine (SVM)	Assignment 2
6	Applications (1)	Generative modeling, Large Language Modeling, Fine Tuning, Prompt Engineering, OpenAI's ChatGPT, Meta's LLaMA	Assignment 3
7-8	Mid-term Quiz, Project using ChatGPT, LLaMA, huggingface, and other tools		
9-10	NLP and ML (2)	Unsupervised Classification, Topic Extraction, Semantic Search, Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA)	Assignment 4
11-12	NLP and DL	Word Embedding and word2vec, RNN/LSTM, Variational Autoencoders, Transformer/BERT, Supervised, unsupervised, and reinforcement learning for Transformers	Assignment 5
12	Applications (2)	Voice User Interface, Speech-to-Text, Text-to-Speech, Translation, Document Summarization, Natural Language Query Translation to SQL	Assignment 6
13-14	Final Quiz, Project in Chatbot and Image Captioning		

Grading Scale

95-100% A	87-89.9% B+	77-79.9% C+	69.9% or below F
	84-86.9% B	74-76.9% C	
90-94.9% A-	80-83.9%B-	70-73.9% C-	

Attendance/Late Work Policy

Attendance Policy

Students registered in MGEN courses (INFO, CSYE, and DAMG) are allowed **a maximum of 2 absences per course, with 3 or more absences resulting in an automatic 'F' for that course.** Students are expected to inform their instructors of any absences in advance of the class; if a student is sick long-term or experiences a medical issue that prevents class attendance, it is strongly encouraged that they speak with their Academic Advisor (coe-mgen-gradadvising@northeastern.edu) to learn more about the Medical Leave of Absence. Should a student anticipate being unable to attend 3 or more classes, they should discuss their situation with their Academic Advisor to explore other types of leave in accordance with the University's academic and global entry expectations. International students should review the Office of Global Services webpage to understand their visa compliance requirements.

Teaching Assistants (TAs) or Instructional Assistants (IAs) will be present at each class to collect student attendance.

Late Work Policy

Students must submit assignments by the deadline in the time zone noted in the syllabus. Students must communicate with the faculty prior to the deadline if they anticipate work will be submitted late. Work submitted late without prior communication with faculty will not be graded.

End-of-Course Evaluation Surveys

Your feedback regarding your educational experience in this class is particularly important to the College of Engineering. Your comments will make a difference in the future planning and presentation of our curriculum.

At the end of this course, please take the time to complete the evaluation survey at <https://neu.evaluationkit.com>. Your survey responses are **completely anonymous and confidential**. For courses 6 weeks in length or shorter, surveys will be open one week prior to the end of the courses; for courses greater than 6 weeks in length, surveys will be open for two weeks. An email will be sent to your Northeastern University Mail account notifying you when surveys are available.

Academic Integrity

A commitment to the principles of academic integrity is essential to the mission of Northeastern University. The promotion of independent and original scholarship ensures that students derive the most from their educational experience and their pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University.

As members of the academic community, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, the use of study aids, the appropriateness of assistance, and other issues. Students are responsible for learning the conventions of documentation and acknowledgment of sources in their fields. Northeastern University expects students to complete all examinations, tests, papers, creative projects, and assignments of any kind according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

Go to <http://www.northeastern.edu/osccr/academic-integrity-policy/> to access the full academic integrity policy.

MGEN Student Feedback

Students who would like to provide the MGEN unit with anonymous feedback on this particular course, Teaching Assistants, Instructional Assistants, professors, or to provide general feedback regarding their program, may do so using this survey: https://neu.co1.qualtrics.com/jfe/form/SV_cTIAbH7ZRaaW0Ki

University Health and Counseling Services

As a student enrolled in this course, you are fully responsible for assignments, work, and course materials as outlined in this syllabus and in the classroom. Over the course of the semester if you experience any health issues, please contact UHCS.

For more information, visit <https://www.northeastern.edu/uahcs>.

Student Accommodations

Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the Rehabilitation Act and the Americans with Disabilities Act Amendments Act (ADAAA) to participate fully in the activities of the university. To receive accommodations through the DRC, students must provide appropriate documentation that demonstrates a current substantially limiting disability.

For more information, visit <https://drc.sites.northeastern.edu>.

Library Services

The Northeastern University Library is at the hub of campus intellectual life. Resources include over 900,000 print volumes, 206,500 e-books, and 70,225 electronic journals.

For more information and for education specific resources, visit <https://library.northeastern.edu>
Network Campus Library Services: [Northeastern University Library Global Campus Portals](#)

24/7 Canvas Technical Help

For immediate technical support for Canvas, call 617-373-4357 or email help@northeastern.edu

Canvas Student Resources: <https://canvas.northeastern.edu/student-resources/>

For assistance with my Northeastern e-mail, and basic technical support:

Visit ITS at <https://its.northeastern.edu>

Email: help@northeastern.edu

ITS Customer Service Desk: 617-373-4357

Diversity and Inclusion

Northeastern University is committed to equal opportunity, affirmative action, diversity, and social justice while building a climate of inclusion on and beyond campus. In the classroom, members of the University community work to cultivate an inclusive environment that denounces discrimination through innovation, collaboration, and an awareness of global perspectives on social justice.

Please visit <http://www.northeastern.edu/oidi/> for complete information on Diversity and Inclusion

Title IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance.

Northeastern's Title IX Policy prohibits Prohibited Offenses, which are defined as sexual harassment, sexual assault, relationship or domestic violence, and stalking. The Title IX Policy applies to the entire community, including male, female, transgender students, faculty, and staff.

In case of an emergency, please call 911.

Please visit <https://www.northeastern.edu/ouec> for a complete list of reporting options and resources both on- and off-campus.