Northeastern University
College of Engineering

INFO 7610 in Special Topics in Natural Language Engineering Methods and Tools

Course Information
Course Title: Special Topics in Natural Language Engineering Methods and Tools
Course Number: INFO 7610-01-ST
Term and Year: Spring 2022
Credit Hour: 4
Course Format: On-Ground

Instructor Information
Full Name: Subrata Das
Email Address: sjskdas@gmail.com

Course Prerequisites
Graduate level INFO 6205 Minimum Grade of C-

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 applying algorithms to extract actionable insights (e.g. customer sentiment and document topic and class) and to enhance business process efficiency (e.g. document summarization, image captioning, and key information extraction). This class will cover the most fundamental and cutting-edge tools, techniques and applications of NLP to get ready for the challenging job market.

Standard Learning Outcomes

Learning outcomes common to all College of Engineering Graduate programs:
1. *An ability to identify, formulate, and solve complex engineering problems.*
2. *An ability to explain and apply engineering design principles, as appropriate to the program's educational objectives.*
3. *An ability to produce solutions that meet specified end-user needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors.*

The Information Systems Program accepts students of different engineering backgrounds with minimum programming skills and produces first class Information Systems engineers that operate at the intersection of real-world complexity, software development, and IT management. Graduating students will be able to construct end-to-end advanced software applications that meet business needs.

Specific Learning Outcomes for the Information Systems program:

1. *Create a strong technical foundation through diverse, high-level courses*
2. *Built crucial interpersonal skills needed to succeed in any industry*
3. *Foster a deep level of applied learning through project based case studies*

Course Outcomes and Assesment Standards

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' latest book on NLP and its applications.

Course Schedule:

Week	Topic	Content	Assignment
1	Introduction and Tools	History of Symbolic AI, NLP, and ML/DL, Applications of NLP, Python and Java Programming, NLTK and aText Tools, Elasticsearch	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)	Document Search, Sentiment and Social Network Analyses	Assignment 3
7-8		Mid-term Quiz, Search Project with NoSQL Hadoop/MapReduce, Elasticsearch, and Graph Databases	
9-10	NLP and ML (2)	Unsupervised Classification, Topic Extraction, Latent Semantic Analysis (LSA), Latent Dirichlet Allocation (LDA)	Assignment 4
11-12	NLP and DL	Word Embedding and word2vec, RNN/LSTM, BERT	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	

End-of-Course Evaluation Surveys

Your feedback regarding your educational experience in this class is very important to the College of Professional Studies. 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 HuskyMail 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.

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 http://www.northeastern.edu/drc/getting-started-with-the-drc/.

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 http://subjectguides.lib.neu.edu/edresearch.

24/7 Blackboard Technical Help
For immediate technical support for Blackboard, call 617-373-4357 or emailhelp@northeastern.edu

Within Blackboard, open a support case via the red support button on the right side of the screen, click Create Case

myNortheastern, e-mail, and basic technical support
Visit the Information Technology Services (ITS) Support Portal
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, member 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 www.northeastern.edu/titleix for a complete list of reporting options and resources both on- and off-campus.