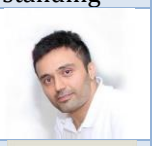
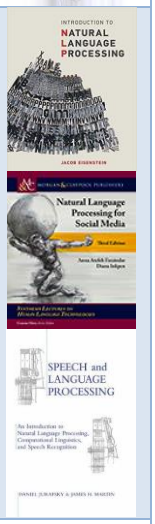
University of Windsor

School of Computer Science
Faculty of Science

COMP-8730: Natural Language Processing & Understanding
Winter 2021

Course Outline (Syllabus) v1.0	
Description	Natural language processing (NLP) is the set of methods in linguistics, computer science, and artificial intelligence concerned with how to program computers to process and analyze large amounts of natural language data. The result is a computer capable of *understanding* the contents of documents, including the language's contextual nuances within them. In the past decade, NLP has become embedded in our daily lives: Households like Alexa and Siri, automatic machine translation on the web and in social media, text classification in email inboxes, search engines, dialogue systems and chatbots. This course aims to provide a survey of ideas, algorithms, linguistics, and statistics for the diverse applications of *NLP with a particular focus on textual user-generated content analysis in online social networks.*
Learning Outcome	Blackboard → COMP8730-1-R-2021W: Natural Language Processing & Understanding → Learning Outcome
Program Level	Graduate
Prerequisite	COMP-8700: Introduction to Artificial Intelligence
Required for	N/A
Instructional Hour	3:00 lecture
Lecture	Tuesday & Thursday, 04:00 PM - 05:20 PM, Blackboard Collaborate Ultra → Classroom
Office Hour	Tuesday & Thursday, 05:30 PM - 06:30 PM, Blackboard Collaborate Ultra → Office
Homepage	blackboard.uwindsor.ca → COMP8730-1-R-2021W: Natural Language Processing & Understanding
Instructor	Name: Hossein Fani Email: hfani@uwindsor.ca[1] Page: hfani.myweb.cs.uwindsor.ca Office: 5111 Lambton Tower 
Books	Introduction to Natural Language Processing Jacob Eisenstein ISBN: 9780262042840 536 pages October 2019 http://cseweb.ucsd.edu/~nnakashole/teaching/eisenstein-nov18.pdf Natural Language Processing for Social Media, Third Edition *Synthesis Lectures on Human Language Technologies* Anna Atefeh Farzindar, Diana Inkpen ISBN: 9781681738116 | PDF ISBN: 9781681738123 Hardcover ISBN: 9781681738130 Copyright © 2020 | 219 Pages DOI: 10.2200/S00999ED3V01Y202003HLT046 Speech and Language Processing, 3rd Edition Draft An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition Dan Jurafsky and James H. Martin Free Access → https://web.stanford.edu/~jurafsky/slp3/ 
Hands-on Resources	NLP with Python – Analyzing Text with the Natural Language Toolkit https://www.nltk.org/book/ NLP with PyTorch: Build Intelligent Language Applications Using Deep Learning. https://github.com/joosthub/PyTorchNLPBook

[1] Wishing to contact the instructor via email about the course, please use [uwinid]@uwindsor.ca and indicate full name, studentid, a course title.

Marking Scheme	***Research* Project** - Problem Definition - Literature Review - Presentation I: Proposal - Proposed Solution - Experiment - Presentation II: Demonstration - Conference Submission (Pending Instr. 's Approval)	65% - 05% - 10% - 05% - 10% - 20% - 05% - 10%
	Assignments (+peer review)	15%
	Midterm Exam	10%
	Final Exam	10%
Remarks	The written reports will be assessed not only on their academic merit but also on the student's communication skills as exhibited through the reports. To achieve a passing grade, the students must achieve at least 70% of the entire marking scheme. The fraction mark is rounded to the ceiling. The students earn final course grades as per the Senate policy for Grading and Calculation of Averages.	
Theoretical[2]	**Lec01B: Meet and Greet:** Course Outline	Jan. 07
	Lec02: Text Normalization	Jan. 12, 14
	Lec03: Text Vector Representation	Jan. 19, 21
	Lec04: Neural Text Vector Representation	Jan. 26, 28
	Lec05: Language Models	Feb. 02, 04
	Lec06: Neural Language Models	Feb. 09, 11
	Lec07: Reading Week: No Class	Feb. 13 – 21
	Lec08: Project Presentations	Feb. 23, 25
	Lec09: Learning w/ Supervision: Sentiment, Sense Disambiguation	Mar. 02, 04
	Lec10: Learning w/o Supervision I: Topic Modeling	Mar. 09, 11
	Lec11: Learning w/o Supervision II: User Modeling	Mar. 16, 18
	Lec12: Learning w/o Supervision III: Web Search	Mar. 23, 25
	Lec13: Project Presentations	Mar. 30, Apr. 01
	Lec14A: Peer Review	Apr. 06
	Lec15: Final Exam	Apr. 12 – 22
***Research* Project**	**Problem Definition (05%)**	Authors (Team Members, Max: 2 Students) Abstract Motivation (Application?) Formal Definition (Mathematical Formulation)
	Literature Review (10%)	History [-∞, 2010] State of the Art (SOTA) [2011, 2021] Critics & Gaps (Regardless the Project Fills Them) What Gap(s) the Project Fills
	Presentation I: Proposal (05%)	10' + 5' Q&A
	Proposed Solution (10%)	Formal Definition (Mathematical Formulation) Implementation (Online Repo, preferably Github)
	Experiment (20%)	Datasets Evaluation Methodology Metrics Baselines Results
	Presentation II: Demo (05%)	10' + 5' Q&A
	Conference Submission (10%)	Pending Instructor's Approval - KES2021 - RecSys21 - [more]
	Conference Acceptance (Bonus)	Conference Registration Conference Travel
Attendance	Encouraged but not mandatory due to time zone accommodation in the COVID-19 pandemic era. Lecture recordings with captions along with presentation slides will be available.	

[2] This is a preliminary schedule. The material and depth and order of presentation are subject to change at the discretion of the instructor and student pace.

Notes to Students:

1. **Equity, Diversity, and Inclusiveness (EDI):** This course, along with all its components such as lab sections are, without question, safe places for students of all races, genders, sexes, ages, sexual orientations, religions, disabilities, and socioeconomic statuses. Disrespectful attitude, sarcastic comments, offensive language, or language that could be translated as offensive and/or marginalize anyone are absolutely unacceptable. Immediate actions will be taken by the instructor to protect the safety and comfort of the students. An ethnically rich and diverse multi-cultural world should be celebrated in the classroom. The instructor, too, must treat every student equally and with the respect and compassion that all students deserve. Furthermore, UWindsor is committed to combatting sexual misconduct. All members are required to report any instances of sexual misconduct, including harassment and sexual violence, to the Sexual Misconduct Response & Prevention Office so that the victim may be provided appropriate resources and support options.

2. **Student Accessibility Services:** Students who have special needs due to legitimate medical reasons should notify the Student Accessibility Services and the instructor at the beginning of the course and before any assessment.

3. **Research Project:** This course is research-oriented and project-driven:
   - Your knowledge will be mostly evaluated based on an ongoing research project defined in the NLP-related area.
   - The research project can be selected from the available research problems or should be approved by the instructor.
   - Each milestone is accompanied with a manual that explains the task and submission procedure as well the deadline.
   - Project assignments are expected to be submitted on the assigned due date and time.
   - Late submission receives 0 unless a verifiable reason with appropriate documentation is provided.
   - The students should follow the submission procedure for each project assignment. Failure to follow the procedure (e.g., incorrect, unreadable, and missing file attachments as instructed) heavily penalizes the submission.
   - Each assignment must be done individually or in a team of maximum 2 students.
   - The research project's milestones are interrelated and are met in order.
   - Conference submission needs the instructor approval pending the quality of the research manuscript meets the *high* standards for scientific scholarly papers!
   - **Bonus:** In the case of conference acceptance of your research paper, the instructor will cover the conference registration fee. Also, the instructor will try to fund the conference travel for 1 student.

4. **Assignments:** There are two types of assignments based on critical thinking (math and/or descriptive questions) and coding skills (data mining and machine learning coding).

5. **Midterm Exam:** Should a student miss a midterm exam, with appropriate documentation and verifiable reason, the weight of the missed midterm exam will be moved to the final exam. The results of the midterm exam will be released to students at least 2 days prior to the voluntary withdrawal deadline as per the Senate Bylaw 54: Undergraduate Academic Evaluation Procedures.

6. **Final Exam:** Students who miss a final exam for a verifiable reason and with appropriate documentation will be given a make-up exam prior to the submission of final course grades that carries the same weight and measure the same knowledge.

7. **Make-up of the Make-up:** There will be *no* make-up of the make-up exam, and the final grade will be assigned based on overall work.

8. **Required Documentation for Missing Exam:** Medical or compassionate documents for the missing of an exam must be submitted within 1 week of the exam. Students are responsible for notifying the instructor that they will be missing an exam as soon as possible.

9. **Accommodation for Religious or Spiritual Observance:** Requests for accommodation of specific religious or spiritual observance must be presented to the instructor no later than 2 weeks prior to the conflict in question (in the case of final examinations within two weeks of the release of the examination schedule). In extenuating circumstances, this deadline may be extended. If the dates are not known well in advance because they are linked to other conditions, requests should be submitted as soon as possible in advance of the required observance. Timely requests will prevent difficulties in arranging constructive accommodations.

10. **Academic Accommodation:** A student who has 3 or more major in-term evaluations scheduled or due within 24 hours may apply, no later than the end of the first quarter of classes, to seek an appropriate accommodation such as a due date extension, alternative assignment, or rescheduled exam.

11. **Appeal:** Students have the right to review the exams and assignments marking within 1 week of their release.

12. **Policies, Bylaws, and Procedures:** Students are required to adhere to all relevant policies, bylaws, and procedures at the University of Windsor, including, but not limited to, the student code of conduct, academic integrity, student academic and

non-academic conduct. Failure to follow the policies, bylaws, and procedures are subject to disciplinary procedures as set out under, but not limited to, the Senate Bylaw 31: Academic Integrity and Procedures for Addressing Student Non-Academic Misconduct. Regarding plagiarism, the Blackboard's SafeAssign will be used for some or all student assignments or equivalent at the instructor's discretion. Plagiarized submissions or equivalent (e.g., exams), i.e., submissions with the same or minor modifications, left unmarked. Should you need to record the lectures, please follow the Senate Policy on Recording Lectures.

13. **Communication:** Students are required to obtain and maintain a University of Windsor e-mail account, [uwindid]@uwindsor.ca, for timely communications with the instructor. The course homepage on the Blackboard, *COMP8730-1-R-2021W: Natural Language Processing & Understanding*, is the main notification center for the course announcements and repository for the course material and resources. Blackboard Collaborate Ultra at *COMP8730-1-R-2021W: Natural Language Processing & Understanding → Classroom* is the official place for the lectures and office hours with the instructor. In Microsoft Teams, the team *COMP8730-1-R-2021W* is provided for emergency cases and backup plans only.

14. **Change Notification:** Any changes in the course outline, exam dates, marking, or evaluation will be discussed in class at least 1 week prior to being implemented.

15. **Student Evaluation of Teaching (SET):** The Student Evaluation of Teaching (SET) will be conducted during the last 2 weeks of the classes.

16. **Online Experience:** Participants in online lectures include and students. Students are able to share camera or send messages but cannot share audio unless they Raise Hand, and the instructor allows them. The instructor also supervises private messages. Students are encouraged to let the instructor know of any connection issues asap regarding the quality of presentation in terms of audio and video (e.g., slides).

17. **Feeling Overwhelmed?** Should face obstacles and experience difficulties that affect her academic performance, students can reach out to the following service centers as well as other on- and off-campus resources listed here www.uwindsor.ca/wellness:
    - Student Health Services
    - Student Counselling Centre
    - Peer Support Centre