Course Outline

1. COURSE INFORMATION

Session Offered	Winter 2024
Course Name	Introduction to Computational Natural Language Processing
Course Code	SEP 775
Date(s) and Time(s) of lectures	Lectures (In-person): Wednesday 3:30 pm – 6:30 pm
Program Name	System and Technology
Calendar Description	This graduate course introduces fundamental concepts in computational natural language processing (NLP) and their applications to teaching students how to deal with textual data in Artificial Intelligence. This course demonstrates how machines can learn different tasks in natural language, such as language modeling, text generation, machine translation, and language understanding. In this regard, we go over the most promising methods in this literature and the most recent state-of-the-art techniques. Moreover, this course explores different real-world applications of NLP and helps students get hands-on experience in this field.
Instructor	Hamidreza Mahyar — E-Mail: mahyarh@mcmaster.ca Office Hours: upon request

2. COURSE SPECIFICS

Course Description	This course covers some fundamental concepts in natural language processing (NLP) and how they apply in real-world use cases.	

Instruction Type	Code	Type	Hours per term
	C	Classroom instruction	39
	L	Laboratory, workshop or fieldwork	
	T	Tutorial	
	DE	Distance education	
		Total Hours	39

Resources	ISBN	Textbook Title & Edition	Author & Publisher
		Natural Language Processing	Jacob Eisenstein
		Speech and Language Processing	Dan Jurafsky and James H. Martin
		Deep Learning	Ian Goodfellow, Yoshua Bengio, and Aaron Courville

Prerequisite(s)	• Proficiency in Python programming language • Some Machine Learning and/or Deep Learning knowledge is a plus
Corequisite(s)	
Antirequisite(s)	
Course Specific Policies	All work must be shown to get full credit. *Specific policy of Projects/Assignments submissions:*

	Projects/Assignments will be conducted online by using related software. Completed *Projects/Assignments* should be uploaded to the drop box before midnight of the due date. *Projects/Assignments* missed due to legitimate reasons must be completed later mutually agreed with the instructor.
Departmental Policies	The use of cell phones, iPods, laptops and other personal electronic devices are prohibited from the classroom during the class time, unless the instructor makes an explicit exception.
	Announcements made in class or placed on Avenue are considered to have been communicated to all students including those individuals that are not in class.
	<u>Instructor has the right to submit work to software to identify plagiarism.</u>

3. SUB TOPIC(S)		
Week 1	Overview: Human language, Word meanings, Word2Vec, and Optimization basics	
Week 2	Intro to neural networks: Gradients and backpropagation	
Week 3	Recurrent Neural Networks and Language Models	
Week 4	Conditioned Generation	
Week 5	Part 1: Self attention and transformers Part 2: Seq2Seq models	
Week 6	**Midterm Test**	
Week 7	Midterm Recess (No class)	
Week 8	Prompting, Instruction Finetuning, and RLHF	
Week 9	Question Answering systems	
Week 10	Multimodal models	
Week 11	Project presentation and Q&A	
Week 12	Project presentation and Q&A	
Week 13	Project presentation and Q&A	
Classes end: April 10th		

4. ASSESSMENT OF LEARNING	**Weight**
4 Assignments	20%
4 Quizzes	10%
Midterm Test	35%
Group Project	35%
TOTAL	**100%**

Note that the percentage distribution of each module may be adjusted.
Percentage grades will be converted to letter grades and grade points per the University calendar.

5. LEARNING OUTCOMES

By the end of this course, the students will know about:

- Introduction to Deep Learning
- Word representations
- Language models
- Pre-trained language models
- Text generation
- Seq2Seq models
- Prompting
- Feedback systems
- Multimodal models
- NLP applications (e.g. Q/A systems)

6. COURSE OUTLINE – APPROVED ADVISORY STATEMENTS

ANTI-DISCRIMINATION

The Faculty of Engineering is concerned with ensuring an environment that is free of all discrimination. If there is a problem, individuals are reminded that they should contact the Associate Director, Graduate Studies, the Sexual Harassment Officer or the Human Rights Consultant, as soon as possible.
http://www.mcmaster.ca/policy/General/HR/Discrimination_Harassment_Sexual_Harassment-Prevention&Response.pdf

ACADEMIC INTEGRITY

You are expected to exhibit honesty and use ethical behaviour in all aspects of the learning process. Academic credentials you earn are rooted in principles of honesty and academic integrity. It is your responsibility to understand what constitutes academic dishonesty.

Academic dishonesty is to knowingly act or fail to act in a way that results or could result in unearned academic credit or advantage. This behaviour can result in serious consequences, e.g. the grade of zero on an assignment, loss of credit with a notation on the transcript (notation reads: "Grade of F assigned for academic dishonesty"), and/or suspension or expulsion from the university. For information on the various types of academic dishonesty please refer to the Academic Integrity Policy, located at https://secretariat.mcmaster.ca/university-policies-procedures-guidelines/

The following illustrates only three forms of academic dishonesty:
- plagiarism, e.g. the submission of work that is not one's own or for which other credit has been obtained.
- improper collaboration in group work.
- copying or using unauthorized aids in tests and examinations.

AUTHENTICITY / PLAGIARISM DETECTION

Some courses may use a web-based service (Turnitin.com) to reveal authenticity and ownership of student submitted work. For courses using such software, students will be expected to submit their work electronically either directly to Turnitin.com or via an online learning platform (e.g. A2L, etc.) using plagiarism detection (a service supported by Turnitin.com) so it can be checked for academic dishonesty.

Students who do not wish their work to be submitted through the plagiarism detection software must inform the Instructor before the assignment is due. No penalty will be assigned to a student who does not submit work to the plagiarism detection software. All submitted work is subject to normal verification that standards of academic

integrity have been upheld (e.g., on-line search, other software, etc.). For more details about McMaster's use of Turnitin.com please go to www.mcmaster.ca/academicintegrity.

COURSES WITH AN ON-LINE ELEMENT

Some courses may use on-line elements (e.g. e-mail, Avenue to Learn (A2L), LearnLink, web pages, capa, Moodle, ThinkingCap, etc.). Students should be aware that, when they access the electronic components of a course using these elements, private information such as first and last names, user names for the McMaster e-mail accounts, and program affiliation may become apparent to all other students in the same course. The available information is dependent on the technology used. Continuation in a course that uses on-line elements will be deemed consent to this disclosure. If you have any questions or concerns about such disclosure please discuss this with the course instructor.

ONLINE PROCTORING

Some courses may use online proctoring software for tests and exams. This software may require students to turn on their video camera, present identification, monitor and record their computer activities, and/or lock/restrict their browser or other applications/software during tests or exams. This software may be required to be installed before the test/exam begins.

COMMUNICATIONS

It is the student's responsibility to:

- Maintain current contact information with the University, including address, phone numbers, and emergency contact information.
- Use the University provided e-mail address or maintain a valid forwarding e-mail address.
- Regularly check the official University communications channels. Official University
   communications are considered received if sent by postal mail, by fax, or by e-mail to the student's designated primary e-mail account via their @mcmaster.ca alias.
- Accept that forwarded e-mails may be lost and that e-mail is considered received if sent via the student's @mcmaster.ca alias.
- Check the McMaster/Avenue email and course websites on a regular basis during the term.

CONDUCT EXPECTATIONS

As a McMaster student, you have the right to experience, and the responsibility to demonstrate, respectful and dignified interactions within all of our living, learning and working communities. These expectations are described in the Code of Student Rights & Responsibilities (the "Code"). All students share the responsibility of maintaining a positive environment for the academic and personal growth of all McMaster community members, whether in person or online.

It is essential that students be mindful of their interactions online, as the Code remains in effect in virtual learning environments. The Code applies to any interactions that adversely affect, disrupt, or interfere with reasonable participation in University activities. Student disruptions or behaviours that interfere with university functions on online platforms (e.g. use of Avenue 2 Learn, WebEx or Zoom for delivery), will be taken very seriously and will be investigated. Outcomes may include restriction or removal of the involved students' access to these platforms.

ACADEMIC ACCOMMODATION OF STUDENTS WITH DISABILITIES

Students with disabilities who require academic accommodation must contact Student Accessibility Services (SAS) at 905-525-9140 ext. 28652 or sas@mcmaster.ca to make arrangements with a Program Coordinator. For further information, consult McMaster University's Academic Accommodation of Students with Disabilities policy.

ACADEMIC ACCOMMODATION FOR RELIGIOUS, INDIGENOUS OR SPIRITUAL OBSERVANCES (RISO)

Students requiring academic accommodation based on religious, indigenous or spiritual observances should follow the procedures set out in the RISO policy. Students should submit their request to their Faculty Office normally within 10 working days of the beginning of term in which they anticipate a need for accommodation or to the Registrar's Office prior to their examinations. Students should also contact their instructors as soon as possible to make alternative arrangements for classes, assignments, and tests. http://www.mcmaster.ca/policy/Students-AcademicStudies/Studentcode.pdf

COPYRIGHT AND RECORDING

Students are advised that lectures, demonstrations, performances, and any other course material provided by an instructor include copyright protected works. The Copyright Act and copyright law protect every original literary, dramatic, musical and artistic work, including lectures by University instructors
The recording of lectures, tutorials, or other methods of instruction may occur during a course. Recording may be done by either the instructor for the purpose of authorized distribution, or by a student for the purpose of personal study. Students should be aware that their voice and/or image may be recorded by others during the class. Please speak with the instructor if this is a concern for you.

EXTREME CIRCUMSTANCES

The University reserves the right to change the dates and deadlines for any or all courses in extreme circumstances (e.g., severe weather, labour disruptions, etc.). Changes will be communicated through regular McMaster communication channels, such as McMaster Daily News, A2L and/or McMaster email.