COMP 7747/8747: Advanced Topics in Machine Learning

Basic Information
Time & Place: Every Monday and Wednesday, 5:30 pm - 6:55 pm
Instructional mode: In-Person
Instructor: Xiaolei Huang
Office Hours: TBD
Contact: xiaolei.huang@memphis.edu

Course Description
Computers can collect vast amounts of data from a wide range of sources, while how to efficiently understand the data is still under exploration. This course will cover advanced concepts and techniques of machine learning (ML) approaches with a special focus on applying in natural language processing (NLP). The course starts with primary concepts and methods in machine learning. Topics include necessary concepts of deep learning, word embedding, language model, Transformer, reinforcement learning and transfer learning. The special focus is for NLP, which is an interdisciplinary field that have many applications in the other fields, such as computational social science, psychology, health, cognitive science, etc. We will cover applications of neural models in several major NLP applications. Finally, we will study how machine learning models learn biases from human language.

Each week, we'll discuss a state-of-the-art approach in natural language processing and machine learning. Each class will have three components: introduction, mathematical foundation and (NLP) applications. Students who do research talks that will follow a similar format. This course will focus on readings, research talks and projects. Students will give two presentations during the course: research talk and group project presentation.

- Students will choose one paper (no earlier than 2015) from ML-related top conferences (e.g., ACL/NIPS/ICML/CVPR) each week. Students will post reading reflections on the course platform. Each student will comment two other reflection posts within a week.
- Students will sign up groups and will give a research talk. Each group (2 – 3 students) will discuss an advanced ML-related publication or topic in detail for 30 minutes depending on final enrollments. The talk will practice presentation skills. Students must post URL, abstract and title of the selected publication on our course platform a week before their talks.
- The final project will have three stages, initial team proposal, one-page midterm report and a final project report. The project aims to guide students to submit conference papers at the end of this course. We will release more detailed rubric.

Prerequisite
One of the following courses:
- COMP 7745/8745: Machine Learning
- COMP 7780/8780: Natural Language Processing
- COMP 4992/6992: Intro to NLP
- COMP 7118/8118: Topics in Data Mining

We will work extensively on the Python programming language. It is assumed that you know basic machine learning techniques (logistic regression). You will also need to be able to read recent research papers in NLP and machine learning.

**Textbook (FREE)**
- G_DL: Deep learning:
  - [https://www.deeplearningbook.org/](https://www.deeplearningbook.org/)
    - HTML version
- E_NLP: Jacob Eisenstein, Natural Language Processing (2018):

**Schedule (subject to change)**

Week 1: Introduction to Python, Probability, NLP
- Date: 08-22 & 08-24
- No Reading Reflection.

Week 2: Frameworks and Deep Learning Basics
- Date: 08-29 & 08-31
- Reading Reflection **due** at 09-03

Week 3: Word Representations and Evaluation
- Date: 09-07, No class on 09-05 (Labor Day)
- Reading Reflection **due** at 09-10

Week 4: Language Models and Deep Sequence Modeling
- Date: 09-12 & 09-14
- Reading Reflection **due** at 09-17
- Research Talk Signup **due** at 09-17

Week 5: Deep Sequence / Generative Modeling
- Date: 09-19 & 09-21
- Reading Reflection **due** at 09-24
Week 6: Model Interpretability and Attention Mechanism
  - Date: 09-26 & 09-28
  - Reading Reflection due at 10-01

Week 7: Transformers
  - Date: 10-03 & 10-05
  - Final Project proposal due at 10-08
  - Reading Reflection due at 10-08

Week 8: Scientific Writings
  - Date: 10-12, No class on 10-10 (Fall Break)
  - No Reading Reflection due to Fall Break.

Week 9: Transfer Learning 1: Sequential and Cross-lingual
  - Date: 10-17 & 10-19
  - Reading Reflection due at 10-22

Week 10: Transfer Learning 2: Domain Adaptation and Multitask Learning
  - Date: 10-24 & 10-26
  - Reading Reflection due at 10-29

Week 11: Machine Learning for Health
  - Date: 10-31 & 11-02
  - Reading Reflection due at 11-05

Week 12: Multimodal Machine Learning
  - Date: 11-07 & 11-09
  - Final Project Midterm Report: due at 11-12
  - Reading Reflection due at 11-12

Week 13: Computational Social Science
  - Date: 11-14 & 11-16
  - No Reading Reflection.
  - Title and Abstract of Research Talk post due: 11-17

Week 14: ML and NLP ethics
  - Date: 11-21, No class on 11-23 (Thanksgiving Holidays)
- No Reading Reflection.
- **Research Talk Slides** submission **due**: 11-22

**Week 15: Final Presentation**
- Date: 11-28 & 11-30
- No Reading Reflection.
- **Final Project Presentation Slides** submission **due**: 12-01

**Week 16: Project Report**
- Project report submission **due**: 12-06

**Evaluation**
Grades will be computed as follows:

<table>
<thead>
<tr>
<th>Percentage</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>20%</td>
<td>Participation</td>
</tr>
<tr>
<td>20%</td>
<td>Reading Reflection</td>
</tr>
<tr>
<td>20%</td>
<td>Research Talk</td>
</tr>
<tr>
<td>40%</td>
<td>Final Project</td>
</tr>
<tr>
<td>100%</td>
<td>Total</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Score</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>97+</td>
<td>A+</td>
</tr>
<tr>
<td>[93 - 97)</td>
<td>A</td>
</tr>
<tr>
<td>[90 - 93)</td>
<td>A-</td>
</tr>
<tr>
<td>[87 - 90)</td>
<td>B+</td>
</tr>
<tr>
<td>[83 - 87)</td>
<td>B</td>
</tr>
<tr>
<td>[80 - 83)</td>
<td>B-</td>
</tr>
<tr>
<td>[77 - 80)</td>
<td>C+</td>
</tr>
<tr>
<td>[73 - 77)</td>
<td>C</td>
</tr>
<tr>
<td>[70 - 73)</td>
<td>C-</td>
</tr>
<tr>
<td>[60 - 70)</td>
<td>D</td>
</tr>
<tr>
<td>&lt; 60</td>
<td>F</td>
</tr>
</tbody>
</table>

The [ refers to “include”, and the ) means “exclude”.

**Late Policy.** All late submissions are subject to a penalty of 20% per day for no more than two days. Late 1 day: 80%; Late 2 days: 60%; Late > 2 days: 0%.

**Final Project Policy.** Reports are due at 11:59 pm Central Standard Time on Friday indicated on the schedule. Students are required to submit **PDF** files of their research talk slides and project reports. A word processing software (e.g., LaTeX or Word) is recommended, and LaTeX is preferrable. You are allowed to collaborate with other peers but copying and pasting from another
student will be considered plagiarism. The final project will be a group research project on a topic of students’ choices after consulting with the instructor. The final project will base on a short presentation and technical report (8 or more pages). The final project report will be due during the final exam period. More information will be posted on a separated page for the final project.

**Plagiarism**

Plagiarism or cheating behavior in any form is unethical and detrimental to proper education and will not be tolerated. All work submitted by a student (projects, programming assignments, lab assignments, quizzes, tests, etc.) is expected to be a student's own work. The plagiarism is incurred when any part of anybody else's work is passed as your own (no proper credit is listed to the sources in your own work) so the reader is led to believe it is therefore your own effort. Students are allowed and encouraged to discuss with each other and look up resources in the literature (including the internet) on their assignments, but appropriate references must be included for the materials consulted, and appropriate citations made when the material is taken verbatim.

If plagiarism or cheating occurs, the student will receive a failing grade on the assignment and (at the instructor’s discretion) a failing grade in the course. The course instructor may also decide to forward the incident to the Office of Student Conduct for further disciplinary action. For further information on U of M code of student conduct and academic discipline procedures, please refer to [https://www.memphis.edu/osa/students/academic-misconduct.php](https://www.memphis.edu/osa/students/academic-misconduct.php).

“Your written work may be submitted to Turnitin.com, or a similar electronic detection method, for an evaluation of the originality of your ideas and proper use and attribution of sources. As part of this process, you may be required to submit electronic as well as hard copies of your written work or be given other instructions to follow. By taking this course, you agree that all assignments may undergo this review process and that the assignment may be included as a source document in Turnitin.com's restricted access database solely for the purpose of detecting plagiarism in such documents. Any assignment not submitted according to the procedures given by the instructor may be penalized or may not be accepted at all.” (Office of Legal Counsel, October 17, 2005).

**Accommodations**

Any student who anticipates physical or academic barriers based on the impact of a disability is encouraged to speak with me privately. Students with disabilities should also contact Disability Resources for Students (DRS) at 110 Wilder Tower (901-678-2880). DRS coordinates access and accommodations for students with disabilities ([http://www.memphis.edu/drs/](http://www.memphis.edu/drs/)).

If you are sick, in particular with an illness that may be contagious, notify me by email but do not come to class. If you are struggling with anxiety, stress or other mental health related concerns, please consider visiting the Counseling Center or calling 901.504.6442 or 901.468.3633.