INF 385T - Special Topics in Information Science: Deep Learning and Multimodal Systems

Wednesdays 12:00Noon- 3:00PM
Location: UTA 1.212

Instructor: Dr. Abhijit Mishra (he/his)
Email: abhijitmishra@utexas.edu

Office Hours (through zoom only)
Tuesday 8:00AM - 10:00AM
Zoom link: https://utexas.zoom.us/j/8979599959
Meeting ID: 897 959 9959
(Install zoom and Login using UT credentials)

Canvas: https://utexas.instructure.com/courses/1366978

Communication and Asking for Help
Please ask all questions that are applicable to the entire class on Canvas, so that others may benefit from the discussion. Only use email for questions unique to individual circumstances; in those cases, please address all questions to abhijitmishra@utexas.edu .

Course Description

Deep Learning (DL) is a subfield of machine learning (ML) that is based on artificial neural networks (ANNs) with multiple layers (hence, "deep"), which are designed to perform complex tasks. Unlike traditional ML, deep learners leverage huge amounts of labeled data and powerful computing resources to learn and improve over time, so much so that they can perform tasks in computer vision, speech recognition, and natural language processing with similar accuracies as humans. DL, as a field, has tremendously grown in the last five years or so and has already had a significant impact on many areas of AI, making it a valuable skill to have and an exciting area of research and development.

The proposed graduate-level course aims to cover theoretical and applied aspects of Deep Learning and how it is used to solve real-world problems, dealing with multimodal data such as text, audio and images or a mixture of them. Classes in each week may be divided into two segments: **(a) Theory and Methods**, a concise description of a deep learning algorithm, and **(b) Practicum**, a hands-on session on applying the algorithm on multimodal real world data such as textual, visual and audio data.

Objectives

By attending and navigating through the course completely and successfully, students will have achieved the following objectives:

1. Understanding Multimodal Data Integration *i.e.,* grasping the fundamental concepts and techniques for integrating and processing data from multiple modalities, such as text, images, audio, and video. This goal involves learning how to effectively represent and fuse diverse types of data in deep learning models.

2. Mastering Deep Learning Fundamentals *i.e.,* striving to build a strong foundation in deep learning principles, including neural networks, optimization techniques, activation

functions, loss functions, and regularization methods. This knowledge is essential for creating effective multimodal models.

3. Creating Multimodal Architectures i.e., knowing how to design and implement advanced multimodal architectures that can handle various data types simultaneously. This involves understanding how to construct neural networks that can process and combine inputs from different modalities to improve overall model performance.

4. Applying Multimodal Models to Real-World Problems *i.e.,* having an idea of how to tackle real-world problems across different domains with the help of deep learning based multimodal solutions. Students will seek to develop projects that leverage multimodal data to achieve tasks such as sentiment analysis, image captioning, audio-visual recognition, and more.

5. Evaluating and Interpreting Model Outputs: Students might aspire to learn how to evaluate the performance of multimodal models effectively. This includes understanding metrics for assessing model accuracy, precision, recall, and F1 score in a multimodal context. Additionally, students could aim to interpret model outputs to gain insights into how the model makes decisions based on different modalities.

Prerequisites

[1] INF 380P: Introduction to Programming in Python or equivalent programing coursework

The proposed ML is applied in nature and there is a lab session in each class where students will code in Python. While the instructor will provide handouts for python basics, there is no way a student without any knowledge in programming will be able to pick up and fully participate in classes. Hence, INF380P (or equivalent programming course) is a necessary prerequisite.

[2] INF385T-Introduction to Machine Learning or equivalent ML coursework

Students are expected to have been exposed to harnessing and processing data, and apply traditional machine learning algorithms (such as Logistic Regression, SVMs, Decision Trees and Feed Forward networks). INF385T may be treated as a co-requisite but it is preferable to complete INF385T before registering for the NLP course. Alternatively, students may opt for a combination of the following courses (or courses that are similar in nature):

SDS 321 - Introduction to Probability and Statistics
SDS 323 - Statistical Learning and Inference
CS-329E: Elements of Data Analytics
INF385T - Artificial Intelligence in Health

OR Undergraduate Courses:
I310D: Introduction to Human Centred Data Science
I320D (undergraduate): Topics in Data Science – Applied Machine Learning with Python

Instruction Modality

Class meetings will be **in person**, with some exceptions, dependent on the state of the COVID-19 pandemic. Only if we are unable to meet in person, classes will be held virtually via Zoom. The classes will consist of a mixture of lectures and hands-on sessions. As of now, there are **no plans to record lectures**.

Accommodations for Students with Disabilities

The university is committed to creating an accessible and inclusive learning environment consistent with university policy and federal and state law. Please let me know if you experience any barriers to learning so I can work with you to ensure you have equal opportunity to participate fully in this course. If you are a student with a disability, or think you may have a disability, and need accommodations please contact Services for Students with Disabilities (SSD). Please refer to SSD's website for contact and more information: http://diversity.utexas.edu/disability/. If you are already registered with SSD, please deliver your Accommodation Letter to me as early as possible in the semester so we can discuss your approved accommodations and needs in this course.

Required Materials

There is no required textbook for this course; all assigned readings will be available online at no cost. Reading materials/resources will be added to canvas for each module. However, throughout the course, we will keep referring to the following classic book by Ian Goodfellow and Yoshua Benjio:

[1] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT press. Free access here: https://www.deeplearningbook.org/

If you wish to purchase it (it's definitely a great addition to your personal library), you can do so by visiting this Amazon link https://www.amazon.com/Deep-Learning-Adaptive-Computation-Machine/dp/0262035618/ref=sr_1_1?ie=UTF8&qid=1472485235&sr=8-1&keywords=deep+learning+book

Required Devices

This course requires students to bring their laptop computers, although it is device agnostic (PC and Mac preferable but do let me know beforehand if you are working with any customized hardware+ OS , something like Raspberry PI board + Linux) . Students will be required to install Python, SQL and Jupyter notebooks. For resource heavy exercises, we may use Google Colaboratory.

Class Participation

Students are expected to attend every class and actively engage themselves in class discussions and complete the lab tutorial at the end of every session. They may polish and submit the tutorial by 11:59PM on the class day.

Assignments and Course Project

The class format is split between reading and coding assignments for the first half of the semester followed by a project the second half of the semester.

1. Assignments

FIVE assignments will be given in the first half of the semester. Each assignment will have: (a) a theoretical question based on weekly assigned readings and (b) a coding exercise similar to the practicum. Assignments are intended to bring conceptual clarity, stimulate algorithmic thinking and emulate practical deep learning implementation scenarios. Moreover, students will be encouraged to reuse the code from the coding assignments in their course projects.

2. Course Project

The goal of the course project is to promote effective planning, execution, and communication of an DL-centric product/research idea. Assignments related to the course project will be related to (a) Project Planning (b) Gathering Resources (c) Experiment Design and Execution, and (d) Preparing presentation, report, and demo. Students will be required to present before the class.

Examination:

A single **in-class** quiz, worth a total of 50 points, will be administered. The quiz will comprise a maximum of 5 questions and is scheduled to last no longer than 2 hours. This examination is designed to be open book/notes; however, access to the internet will not be allowed.

Important Dates:

1. **October 13, 2023 :** Group Formation for Project/Activities
2. **October 18, 2023 :** In-class Quiz (5 questions, 50 points)
3. **November 1, 2023:** Preliminary proposal due for group project
4. **November 15, 2023:** Group project work in progress presentation and feedback soliciting (5 minutes per group and counted towards class activity)
5. **November 29:** Final presentation of group project
6. **December 7:** Final project-report submission
7. **December 11 (tentative):** Submission of final grades by the instructor

For holidays and breaks and other important dates visit UT academic calendar here https://registrar.utexas.edu/calendars/23-24

Late Work and Missed Work

In an effort to accommodate any unexpected personal events, I have enacted a grace policy of two days for this course. You do not have to utilize this policy, but if you find yourself struggling with unexpected personal events, I encourage you to email me as soon as possible (in advance of the due date) to notify me that you are using our grace policy. You may either have a two-day grace period for one assignment, or you may have 2 one-day extensions for two different assignments. The only absences that will be considered excused are for religious holidays or extenuating circumstances due to an emergency. If you plan to miss class due to observance of a religious holiday, please let us know at least two weeks in advance. You will not be penalized for this absence, although you will still be responsible for any work you will miss on that day if applicable. In the event of an unexcused absence, we do not guarantee the opportunity to make up missed in-class work, but one may be granted. Check with us for details or arrangements.

Grading Policies

Course grades will be made up of the following components. Final letter grades will be awarded according to the grade cutoffs below, including pluses and minuses.

Grade Component	Percentage
Attendance, Participation in class and Lab Completion	20%
In-class quiz	10%
Five Assignments	40%
Final Project	30%

Grade Breaks

Grade	Cutoff
A	94%
A-	90%
B+	87%
B	84%
B-	80%
C+	77%
C	74%
C-	70%
D+	67%
D-	60%
F	< 60%

Course Syllabus

(Readings to be updated in Canvas)
All instructions, assignments, week-wise readings, rubrics and essential information will be on the Canvas website. Check the site regularly and use it to ask questions about the course schedule. Changes to the schedule may be made at my discretion and if circumstances require. For example, we might want to slow down, speed up or drop certain topics depending on student input. It is your responsibility to note these changes when announced.

1. WEEK 1. Introduction to Deep Learning (8/23/2023)

Lecture: Course overview and syllabus, What is machine learning and deep learning, types of learners, Introduction to perceptrons and linearly separable functions

Practicum: Linear algebra and vector calculus basics (paper based), Probability and likelihood, Matrix manipulation using NumPy

2. WEEK 2. Perceptrons and Optimization of Objectives through Gradient Descent (8/30/2023)

Lecture: More on perceptrons, Activation functions, Optimization of linear functions, A note on convexity

Practicum: More on Numpy, Introduction to PyTorch and PyTorch tensors, Gradient computation in PyTorch

3. **WEEK 3. Training Multilayered Neural Networks - Back Propagation of Gradients, Loss Functions (09/06/2023)**

Lecture: Non-linear functions, deep neural networks, optimization objectives and loss functions, back propagation basics

Practicum: Back propagation example in PyTorch

4. **WEEK 4. Overfitting and Regularization in Neural Networks - Data Preparation, Dropout Technique (09/13/2023)**

Lecture: Overfitting and under-fitting recap, Datasets and data splits, L1 and L2 regularization in Neural Nets, Dropout

Practicum: Revisiting back propagation with regularization and dropout enabled

5. **WEEK 5. Special Deep Neural Networks I - Convolutional Neural Networks for Images as Data (09/20/2023)**

Lecture: Introduction to CNNs, Significance, Multi-layered and multi filter CNNs , Popular CNN based architecture for image classification

Practicum: Image classification example with CNNs and ResNet

6. **WEEK 6. Special Deep Neural Networks II - Recurrent Neural Networks (09/27/2023)**

Lecture: Introduction to RNNs, Back-propagations with time, Special purpose RNNs, Long-short-term-memories

Practicum: Text classification example with RNNs

7. **WEEK 7. Special Deep Neural Networks III - Transformers (10/04/2023)**

Lecture: Issues with traditional DL architectures, Transformers and their significance

Practicum: Text classification example (repeated) with transformers

8. **WEEK 8. Transfer learning and Self-supervised learning (10/11/2023)**

Lecture: Representation learning with auto encoders, transfer learning and self supervised learning, pre-training and pre-trained models for image, text and audio data

Practicum: K-shot text classification through transfer learning with BERT and GPT-2 as the pre-trained models

9. **WEEK 9. In-class quiz + Deep Generative Modeling-I : Neural Machine Translation, Summarization and Question Answering (10/18/2023)**

Lecture: Machine Translation, Sequence to Sequence models basics, emphasis on attention mechanism

Practicum: None

10. WEEK 10. Deep Generative Modeling-II : Automatic Speech Recognition (10/25/2023)

Lecture: More sequence to sequence model examples, Automatic speech recognition as a multimodal sequence to sequence modeling task

Practicum: Exploring popular ASR models

11. WEEK 11. Deep Generative Modeling-III: Multimodal Question Answering and Image Captioning (11/01/2023)

Lecture: More sequence to sequence model examples, Image captioning and Visual question answering as multimodal sequence to sequence modeling task

Practicum: Exploring **i**mage captioning and VQA models

12. WEEK 12. Deep Generative Modeling-IV: Large Language Models, Prompt Engineering and Reinforcement Learning in Deep Learning - RLHF (11/08/2023)

Lecture: LLM basics, Optimization of LLMs through reinforcement learning (RLHF), Prompt Engineering

Practicum: Exploring open source LLMs

13. WEEK 13. Evaluation Metrics for Different Tasks and Modalities+ Group Project Progress Update (11/15/2023)

Lecture: Quantitative evaluation of deep learning based models, Revisiting classification metrics, Popular evaluation metrics for audio processing and natural language generation

Practicum: Evaluating natural language generations from models

14. WEEK14. Fall Break (NO CLASS) /./

15. WEEK 15. Final Group Project Presentation and Demo + Career Conversation (11/29/2023)

Lecture: No lecture, share presentations before class and present your project in class

16. WEEK 16. No-class (study days) + Project Report Submission

Submission of group project report

17. WEEK 17. No-exam

Mantra for Student Success : Navigating the DL Course

- Achieve higher attendance, aiming for 100% to maximize exposure and engagement during lectures and practical exercises.
- Submit practicums and assignments promptly, recognizing that minor errors can be overlooked while focusing on continuous improvement.

- Prioritize transparency by appropriately citing tools, resources, and data sources, showcasing your commitment to ethical and accountable work.
- Approach in-class quizzes with a clear understanding and well-organized thoughts, leveraging your conceptual clarity to excel.
- If programming presents challenges, embrace deliberate practice to strengthen your skills and confidently navigate technical aspects.
- Embrace iteration as you prepare presentations, ensuring impactful task demonstrations, comprehensive analyses, and well-structured reports.
- Recognize that success in the DL course is a result of these concerted efforts, culminating in your growth as a proficient and accomplished DL practitioner.

Academic Integrity

Students who violate University rules on academic dishonesty are subject to disciplinary penalties, including the possibility of failure in the course and/or dismissal from the University. Since such dishonesty harms the individual, all students, and the integrity of the University, policies on academic dishonesty will be strictly enforced. For further information, please visit the Student Conduct and Academic Integrity website at http://deanofstudents.utexas.edu/conduct.

AI Tools Usage Policy:

The utilization of AI-powered tools, including platforms like ChatGPT, DALL-E, or any other large language/image generative models, to create content such as text, code, images, multimedia, or any related materials intended for assignments, quizzes, or projects that contribute directly to the evaluation of grades within this course is **strictly proscribed**. Exceptions to this rule apply only if the incorporation of such systems aligns with the specified objectives of the assignment or project. Breaching this policy may result in the initiation of proceedings related to student misconduct.

Should there be any suspicion surrounding the content submitted by a student, suggesting the involvement of an AI tool, I retain the authority to request clarification from the student. This clarification may be sought through email communication or arranged verbal discussions in the form of one-on-one meetings. In the event of any inconsistencies between the provided explanations and the submitted solutions, I reserve the right to instigate misconduct proceedings against the concerned student. Upon enrolling in this course, students inherently express their agreement to adhere to this policy as well as any forthcoming policies described below.

Course Material Sharing Policy

Unauthorized sharing or distribution of lecture notes, slides, or examination questions is strictly prohibited without prior permission from the instructors. Failure to adhere to this policy may result in the initiation of legal actions. In the event that class should be recorded, class recordings are reserved only for students in this class for educational purposes and are protected under FERPA. The recordings should not be shared outside the class in any form. Violation of these restrictions by a student could lead to Student Misconduct proceedings.

Religious Holy Days

By UT Austin policy, you must notify me of your pending absence as far in advance as possible of the date of observance of a religious holy day. If you must miss a class, an examination, a work assignment, or a project in order to observe a religious holy day, you will be given an opportunity to complete the missed work within a reasonable time after the absence.

Names and Pronouns

Professional courtesy and sensitivity are especially important with respect to individuals and topics dealing with differences of race, culture, religion, politics, sexual orientation, gender, gender variance, and nationalities. I will gladly honor your request to address you by your chosen name and by the gender pronouns you use. Class rosters are provided to the instructor with the student's chosen (not legal) name, if you have provided one. If you wish to provide or update a chosen name, that can be done easily at this page, and you can add your pronouns to Canvas.

Basic Needs Security

Any student who faces challenges securing their food or housing and believes this may affect their performance in the course is urged to contact the Dean of Students for support. UT maintains the UT Outpost which is a free on-campus food pantry and career closet.

Mental Health Support

I urge students who are struggling for any reason and who believe that it might impact their performance in the course to reach out to me if they feel comfortable. This will allow me to provide any resources or accommodations that I can. If immediate mental health assistance is needed, call the Counseling and Mental Health Center (CMHC) at 512-471-3515 or you may also contact Bryce Moffett, LCSW (iSchool CARE counselor) at 512-232-2983. Outside CMHC business hours (8a.m.-5p.m., Monday-Friday), contact the CMHC 24/7 Crisis Line at 512-471-2255.

Land Acknowledgement

I would like to acknowledge that we are meeting on the Indigenous lands of Turtle Island, the ancestral name for what now is called North America. Moreover, I would like to acknowledge the Alabama-Coushatta, Caddo, Carrizo/Comecrudo, Coahuiltecan, Comanche, Kickapoo, Lipan Apache, Tonkawa and Ysleta Del Sur Pueblo, and all the American Indian and Indigenous Peoples and communities who have been or have become a part of these lands and territories in Texas.

Title IX Reporting

Title IX is a federal law that protects against sex and gender-based discrimination, sexual harassment, sexual assault, unprofessional or inappropriate conduct of a sexual nature, dating/domestic violence and stalking at federally funded educational institutions. UT Austin is committed to fostering a learning and working environment free from discrimination in all its forms. When unprofessional or inappropriate conduct of a sexual nature occurs in our community, the university can:

1. Intervene to prevent harmful behavior from continuing or escalating.
2. Provide support and remedies to students and employees who have experienced harm or have become involved in a Title IX investigation.
3. Investigate and discipline violations of the university's relevant policies.

Beginning January 1, 2020, Texas Senate Bill 212 requires all employees of Texas universities, including faculty, report any information to the Title IX Office regarding sexual harassment, sexual assault, dating violence and stalking that is disclosed to them. Texas law requires that all employees who witness or receive any information of this type (including, but not limited to, writing assignments, class discussions, or one-on-one conversations) must be reported. **I am a Responsible Employee and must report any Title IX related incidents** that are disclosed in writing, discussion, or one-on-one. Before talking with me, or with any faculty or staff member about a Title IX related incident, be sure to ask whether they are a responsible employee. If you would like to speak with someone who can provide support or remedies without making an official report to

the university, please email advocate@austin.utexas.edu. For more information about reporting options and resources, visit http://www.titleix.utexas.edu/ , contact the Title IX Office via email at titleix@austin.utexas.edu, or call 512-471-0419.

Although graduate teaching and research assistants are not subject to Texas Senate Bill 212, they are still mandatory reporters under Federal Title IX laws and are required to report a wide range of behaviors we refer to as unprofessional or inappropriate conduct of a sexual nature, including the types of conduct covered under Texas Senate Bill 212. The Title IX office has developed supportive ways to respond to a survivor and compiled campus resources to support survivors.