

I320D: Topics in Human Centered Data Science: Text Mining and NLP Essentials

Monday / Wednesday : 12:30PM - 2:00PM, **SZB 4.508**

Instructor: [Dr. Abhijit Mishra](#) (he/his)

Email: abhijitmishra@utexas.edu

TA: [Hir Brahmabhatt](#)

Email: hirbrahmabhatt269@utexas.edu

Canvas: <https://utexas.instructure.com/courses/1382133>

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 both abhijitmishra@utexas.edu and hirbrahmabhatt269@utexas.edu

Course Description

Natural Language Processing (NLP) is concerned with interactions between computers and humans through the medium of human languages. It involves analyzing, understanding, and generating human language, making it possible for machines to interpret and respond to human speech and text. NLP is currently making significant contributions to modern technological advancements and serves as the backbone of crucial applications such as question answering, human language translation, summarization, sentiment and emotion analysis, search and recommendation, and information extraction in various domains such as healthcare, finance, legal, libraries and education and beyond.

The proposed graduate-level course aims to cover fundamental concepts in Natural Language Processing / Computational Linguistics and how they are used to solve real-world problems. Classes in each week will be divided into two segments: **(a) Theory and Methods**, a concise description of an NLP concept, and **(b) Lab**, a hands-on session on applying the theory to a real-world task on publicly available multilingual text datasets.

Intention and Objectives

The objective of the course is to provide a bird's eye view of the field of NLP and enable students to make informed decisions while choosing from different career options, i.e., working on an NLP-based cutting-edge product, joining the industry in an NLP-centric role, or pursuing higher studies or research specializing in NLP or computational linguistics.

By the end of the course, the goals for the students are to:

1. Understand the process of garnering and pre-processing a large amount of multilingual textual data from various domains and sources.
2. Characterize the processes to store, load, pre-process multilingual data and apply language processing operations such as normalization, tokenization, lemmatization, chunking and machine readable representation (vector) extraction.

3. Train machine learning algorithms for natural language understanding and generation and evaluate their performance.
4. Learn to extract information from unstructured text and represent them in the form of knowledge graphs
5. Learn to use existing knowledge graphs, ontologies and lexical knowledge networks for predictive analysis on text
6. Learn about popular NLP applications and tasks and the process of building such applications
7. Propose a novel product/research-focused idea (this will be an iterative process), design and execute experiments, and present the findings and demos to a suitable audience (in this case, the class).

Prerequisites

[1] **Programming in Python** (i.e., Programming for Informatics - I304) : 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, programming for informatics (or equivalent programming course) is a necessary prerequisite.

[2] **I310D - Introduction to Human Centered Data Science:** Students are expected to have been exposed to harnessing and processing data, probability and statistics and linear algebra. I310D provides a suitable background and hence a preferred prerequisite. Alternatively, students may opt for one or more 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](#)

Co-requisite(s)

I320D - Topics in Human Centered 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. While most of the reading material will be available online at no cost, the following books should be accessed and / or purchased.

Required books:

1. Dan Jurafsky and James H. Martin. 2006.

Speech and Language Processing. Unofficial copy: [https://pages.ucsd.edu/~bakovic/compphon/Jurafsky, Martin.-Speech and Language Processing_ An Introduction to Natural Language Processing \(2007\).pdf](https://pages.ucsd.edu/~bakovic/compphon/Jurafsky, Martin.-Speech and Language Processing_ An Introduction to Natural Language Processing (2007).pdf) or buy from <https://www.amazon.com/Speech-Language-Processing-Daniel-Jurafsky/dp/0131873210>

2. Manning, C., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT press. Unofficial Copy: https://doc.lagout.org/science/o_Computer Science/2_Algorithms/Statistical Natural Language Processing.pdf or buy from https://www.amazon.com/Foundations-Statistical-Natural-Language-Processing/dp/0262133601/ref=pd_lpo_scl_3/133-6690059-3136500?pd_rd_w=G9hpl&content-id=amzn1.sym.116f529c-aa4d-4763-b2b6-4d614ec7dc00&pf_rd_p=116f529c-aa4d-4763-b2b6-4d614ec7dc00&pf_rd_r=PD46FJR8DE5CFANKR2V6&pd_rd_wg=PKsq&pd_rd_r=8d765a26-c155-4d8c-b23a-d5dc59621b5c&pd_rd_i=0262133601&psc=1

Books that can be optionally purchased (optional)

3. Kamath, U., Liu, J., & Whitaker, J. (2019). Deep learning for NLP and speech recognition (Vol. 84). Cham, Switzerland: Springer.

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, actively participate in discussions, quizzes and complete the lab tutorial at the end of each session. **Tutorials can be polished and submitted by 11:59 PM on the same day but the first submission should be done in class.**

While attendance will not be explicitly recorded, I will administer quizzes during the theory period. Students present in class will be provided with an access code and 10 minutes to complete the quiz. In-class participants will have multiple opportunities to submit the quiz and earn full points. Those absent “without an excused absence” will receive a ZERO for the quiz.

Similarly attendance on the practical period will be adjudged through programming notebook submissions.

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

SIX assignments will be given, mostly in the first half of the semester. Each assignment may have either (a) a theoretical question based on weekly assigned readings or (b) a coding exercise similar to the practicum or both. Assignments are intended to bring conceptual clarity, stimulate computational thinking and emulate practical NLP 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 **original** NLP-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 and prepare and submit a report in the prescribed format provided during project announcement in Canvas.

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.

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 |
|---------------------------------------------------|------------|
| In class tutorials | 20% |
| In class quizzes (best of ten out of all quizzes) | 10% |
| Six Assignments | 36% |
| Final Project (presentation and report) | 34% |

Grade Breaks

| Grade | Cutoff |
|-------|--------|
| A | 94% |

| Grade | Cutoff |
|-------|--------|
| A- | 90% |
| B+ | 87% |
| B | 84% |
| B- | 80% |
| C+ | 77% |
| C | 74% |
| C- | 70% |
| D+ | 67% |
| D | 64% |
| D- | 60% |
| F | < 60% |

Course Outline

All instructions, assignments, 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. Moreover, please note that the readings provided are preliminary and may be supplemented with additional resources as the course progresses.

1. WEEK 1. Introduction to Natural Language Processing (Jan 13-Jan 15)

Lecture: Course Overview, Syllabus, What is NLP? Ambiguity, Multilingualism, Fundamentals layers of NLP, Overview of text corpora and datasets

Lab: Python Basics and File, String and Document Processing, Frequency Analysis and Visualization of text Data

Readings:

[1] Chapter 1: Introduction , Book: Jurafsky, D., & Manning, C. (2012). Natural language processing. Instructor, 212(998), 3482.

[2] <https://cs.nyu.edu/~davise/ai/ambiguity.html>

[3] <https://www.geeksforgeeks.org/python-string/>

Assignment 1: String manipulation and frequency analysis

2. WEEK 2. Finite State Machines and Regular Expressions (Jan 22)

Lecture: Theory of Automation, Formal Languages, Regular Expressions and Finite State Machines

Lab: Writing regular expressions with Python

Readings:

[1] <https://cs.stanford.edu/people/eroberts/courses/soco/projects/2004-05/automata-theory/basics.html>

[2] **Python RegEx:** <https://www.geeksforgeeks.org/python-regex/>

3. WEEK 3. Text Preprocessing Techniques (Jan 27- Jan29)

Lecture: Cleaning and preprocessing raw text data, Morphology Analysis, Stemming and Lemmatization, Stopword removal and handling special characters, when to apply which operations

Lab: Text pre-processing using NLTK and SpaCy modules

Readings:

[1] Chapter 3. Words and Transducers, Book: Jurafsky, D., & Manning, C. (2012). Natural language processing. Instructor, 212(998), 3482.

[2] <https://www.analyticsvidhya.com/blog/2021/06/text-preprocessing-in-nlp-with-python-codes/>

[3] Text Preprocessing: NLP fundamentals with spaCy <https://medium.com/eni-digitalks/text-preprocessing-nlp-fundamentals-with-spacy-54f32e520bc8>

Assignment 2: Hashtag based tweet search - 1 (posted on Canvas)

4. WEEK 4. Lexical Analysis - Words (Feb 3- Feb 5)

Lecture: Representing Words, Sentences and Documents, Bag-of-words, N-grams, TF-IDF , Word Vectors, Document similarity and distance metrics,

Lab: Document Representation and Similarity Measurement using NLTK, Scikit learn and Gensim

Readings:

[1] Quick Introduction to Bag-of-Words (BoW) and TF-IDF for Creating Features from Text <https://www.analyticsvidhya.com/blog/2020/02/quick-introduction-bag-of-words-bow-tf-idf/>

[2] Bag of words and N-grams: https://en.wikipedia.org/wiki/Bag-of-words_model

[3] NLP: Word Embedding Techniques for Text Analysis <https://medium.com/sfu-csmp/nlp-word-embedding-techniques-for-text-analysis-ec4e91bb886f>

[4] Word embeddings, LSA, Word2Vec, Glove, ELMo <https://people.eng.unimelb.edu.au/mbouadjenek/papers/wordembed.pdf>

Optional Readings:

[1] Word Embeddings <https://people.eng.unimelb.edu.au/mbouadjenek/papers/wordembed.pdf>

Assignment 3: Hashtag based tweet-search -2 (posted on Canvas)

5. WEEK 5. Syntax Analysis - Sentence Structure (Feb 10- Feb 12)

Lecture: Grammar and languages, Representing syntax, Part of Speech Tagging, Shallow and Deep Parsing, Constituency and Dependency parsing

Lab : Using existing part of speech taggers, noun-clunkers, dependency parsers from NLTK and SpaCY

Readings (skip the machine learning related portions):

[1] Chapter 5. Word Classes and Part-of-Speech Tagging, Book: Jurafsky, D., & Manning, C. (2012). Natural language processing. Instructor, 212(998), 3482.

[2] Chapter 13. Parsing with Context Free grammars and Chapter 14. Statistical Parsing. Book: Jurafsky, D., & Manning, C. (2012). Natural language processing. Instructor, 212(998), 3482.

Optional Readings (skip the machine learning related portions):

[1] Chapter 9, 10, 11, 12. Book: Manning, C., & Schütze, H. (1999). Foundations of statistical natural language processing. MIT press.

[2] <https://stanfordnlp.github.io/CoreNLP/parse.html>

6. WEEK 6. Information Extraction from Text (Feb 17 - Feb 19)

Lecture(s) : Entity and relationship extraction, BIO Tagging details, Entities and Knowledge Graphs

Lab: Relation extraction in news, building knowledge tuples from text

Readings:

[1] Chapter 13 Section 13.5. Book: Jurafsky, D., & Manning, C. (2012). Natural language processing. Instructor, 212(998), 3482.

[2] Handout on Named Entity Recognition <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1106/handouts/InfoExtract-cs224n-2010-1up.pdf>

Assignment 4: Information Extraction from Stories - posted on Canvas

7. WEEK 7. Machine Learning Methods for NLP (Feb 24 - Feb26)

Lecture(s) : Test Classification Details, Feature Engineering with N-gram and TF-IDF, Word Vectors

Lab : Implementing text classifiers, feature engineering, evaluating model performance

Readings:

[1] Text classification pipeline: <https://www.oreilly.com/library/view/practical-natural-language/9781492054047/ch04.html>

[1] Vectorization Techniques in NLP: <https://neptune.ai/blog/vectorization-techniques-in-nlp-guide>

Project: Group Formation

8. WEEK 8. Unsupervised ML and Topic Modeling Basics (Mar 3 - Mar 5)

Lecture: Topic modeling basics, LDA, Unsupervised ML basics, K means clustering

Lab: Topic and keyword extraction from text using LDA and clustering methods

Readings:

[1] K-means Clustering - A theoretical foundation <https://medium.com/@akshay.sinha/k-means-clustering-8ef58ca0d024>

[2] A Beginner's Guide to Latent Dirichlet Allocation(LDA) <https://towardsdatascience.com/latent-dirichlet-allocation-lda-9d1cd064ffa2>

Optional Readings:

[1] Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.

Assignment 5: Document clustering and representation - posted on Canvas

Project: Proposal and Planning Document Submission - posted on Canvas

9. WEEK 9. Deep learning for NLP - I (Mar 10 - Mar 12)

Lecture(s): Introduction to Neural Networks and Back propagation, Recurrent neural networks and Transformers

Lab: Text classification using Neural Networks using TensorFlow

Readings:

[1] Quick introduction to neural networks (<https://ujjwalkarn.me/2016/08/09/quick-intro-neural-networks/>)

[2] Neural Network with Python (<https://victorzhou.com/blog/intro-to-neural-networks/>)

Optional Readings:

[1] Introduction - The Perceptron (https://web.mit.edu/course/other/i2course/www/vision_and_learning/perceptron_notes.pdf)

[2] A Neural Network in 11 lines of Python Part 1 (<http://iamtrask.github.io/2015/07/12/basic-python-network/>)

[3] Yann LeCun, Yoshua Bengio, Geoffrey Hinton, Deep Learning (<https://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf>) Nature 521, no. 7553 (2015): 436-444. doi:10.1038/nature14539

10. WEEK 10. Spring Break (No classes held)

11. WEEK 11. Deep learning for NLP - II (Mar 24 - Mar 26)

Lecture: Language Models and Language Modeling objectives, Deep Dive into Word Embeddings and Sentence Embeddings, Transfer Learning with language models like Word2Vec, BERT, RoBERTa and their successors, Pre-training and fine-tuning

Lab: Text classification revisited - with pre-trained language models

Readings:

[1] Transfer Learning for Machine Learning <https://www.seldon.io/transfer-learning>

[2] Transformers and pre-trained language models: <https://web.stanford.edu/~jurafsky/slp3/10.pdf>

Assignment 6: Fine-tuning a language model for text classification

12. WEEK 12. NLP Applications (Mar 31 - Apr 2)

Lectures: Sentiment Analysis, Machine Translation, Question Answering, Summarization basics, the impact of language models

Lab: Text summarization example with language models

Readings:

[1] Patwardhan, N., Marrone, S., & Sansone, C. (2023). Transformers in the Real World: A Survey on NLP Applications. *Information*, 14(4), 242.

Project: 1 page Progress Document

13. WEEK 13. Small and Large Language Models and Prompt Engineering Basics (Apr 7 - Apr 9)

Lecture(s): Prompt based Large Language Models, Transfer learning and reinforcement learning as the basis of LLMs, RLHF, LLMs and applications

Lab: Prompting and prompt-chaining of open sourced Small Language Models for different NLP tasks

Readings:

[1] Liu, P., Yuan, W., Fu, J., Jiang, Z., Hayashi, H., & Neubig, G. (2023). Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Computing Surveys*, 55(9), 1-35. (<https://dl.acm.org/doi/pdf/10.1145/3560815>)

[2] Tutorial on Prompt Engineering: <https://github.com/dair-ai/Prompt-Engineering-Guide/blob/main/lecture/Prompt-Engineering-Lecture-Elvis.pdf>

14. WEEK 14. Knowledge Networks + Pragmatics in NLP (Apr 14 - Apr 16)

Lecture(s): Knowledge networks and Ontology, WordNet, ConceptNet and DBPedia

Project: 5 minute progress presentation

Readings:

[1] Tiwari, S., Gaurav, D., Srivastava, A., Rai, C., & Abhishek, K. (2021). A Preliminary Study of Knowledge Graphs and Their Construction. *Emerging Technologies in Data Mining and Information Security: Proceedings of IEMIS 2020, Volume 3*, 164, 11.

[2] Pragmatics and Computational Linguistics, <https://web.stanford.edu/~jurafsky/prag.pdf>

15. WEEK 15. Evaluation Metrics, Course summary and conclusion (Apr 21 - Apr 23)

Lecture: Evaluation Metrics for NLP, Course overview

Project: Final Project Presentation -1

Readings:

[1] Sai, A. B., Mohankumar, A. K., & Khapra, M. M. (2022). A survey of evaluation metrics used for NLG systems. *ACM Computing Surveys (CSUR)*, 55(2), 1-39.

16. WEEK 16. Final Project Presentation (...) (Apr 28)

Project: Final Project Presentation -2

17. WEEK 17. No Exam

Project: Final project report due on May 1 , 2025

Mantra for Student Success : Navigating the NLP 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 NLP course is a result of these concerted efforts, culminating in your growth as a proficient and accomplished NLP 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.