GEORGE MASON UNIVERSITY | College of Engineering and Computing

Department of Information Sciences and Technology

AIT 726: Natural Language Processing with Deep Learning

Course Syllabus

Lindi Liao, PhD
dliao2@gmu.edu

For each section, a customized syllabus with information specific to that section will be made available to registered students via the Blackboard Learning System.

Objectives

This is an **advanced** course in Natural Language Processing (NLP). It builds on the introductory NLP class (AIT526 Introduction to NLP) and goes on to explore **deep learning** methods in NLP. Topics covered include feedforward neural nets as applied to NLP and deep networks for NLP applications. Class reviews language structure and studies context free grammars, dependency parsing, semantic role labeling along with applications such as sentiment classification and information extraction.

Prerequisites

1) **AIT 526 Intro to NLP**
   - If you completed **AIT 526** but are from other department, please submit an override request form.
   - The link to the override request form is at:
     - https://ist.gmu.edu/academics/advising/submit-registration-override-request
2) **Proficiency in Python**
3) **Probability and Statistics**
4) **College Calculus, Linear Algebra**
5) **Highly Desirable**: **Machine Learning** and **Algorithm Design**

Special Topics Description

This course emphasizes both **theory** and **practice** as well as **industry-ready hands-on** technical skills for *advanced* **NLP** and **Deep Learning**. It includes the following topics but extends their wings of knowledge to some advanced techniques, especially including the ***latest*** promising NLP-related research and technologies:

- **Deep Learning**, **Artificial Neural Networks**
- **Deep Natural Language Semantics, Natural Language Understanding**
- **Vector Representations** (Word Embeddings, Word2Vec, Skip-Gram, etc.)
- **Sentiment Classification with Logistic Regression**
- **Feedforward Neural Networks**, **Neural Language Modelling**

- **Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM)**, **Gated Recurrent Unit (GRU)**
  - Language Modeling, Sequence labeling, Sequence Classification, Text Generation
- **Attentions, Transformers, Pretrained Language Models**
- **Context-Free Grammar (CFG)**, **Combinatory Categorial Grammar (CCG)**, Neural-based CCG
- **Dependency Parsing (DP)**, Neural-based DP, Literature Based Discovery (LBD)
- **Semantic Role Labeling (SRL)**: Shallow Semantic Parsing, Neural-based Semantic Parsing
- **Multimodal AI (Language, Vision, Audio, etc.)**
- **Generative AI, AI-Generated Content (AIGC)**
- **Ethics in NLP** (algorithmic justice, computational approaches to media/race bias, etc.)
- **Trustworthiness in NLP** (interpretable, explainable, transparent, etc.)
- **NLTK**, **PyTorch**, **TensorFlow**, **Keras**, Google Colab, AWS, and other programming tools or platforms
- **Advanced NLP Algorithm Design**
- NLP Project Management and Team Collaboration

Course Learning Activities and Grade Distribution

This course includes programming assignments, paper presentations, a term project, and class activities. Your grade will be based on the following breakdown*:

Item	Quantity	Percent of Grade	Extra Credit	On-Time Submission (Extra Credit)
Class Activities (Participation, in-class discussion, discussion board, team evaluations, etc.)	∞	10%	No	No
** Programming Assignments	4	40%	Yes	Yes
** Paper Presentations	2	15%	Yes	No
** Term Project	1	35%	No	No
TOTAL	**100**% + extra credit			

** Subject to revision before and throughout the course.*

*** Teamwork (up to 3)*

Textbook and Required Materials

For this course, we will be using two required textbooks.

- **Textbook JM3**
  - *Speech and Language Processing*, by Jurafsky and Martin (3rd edition). The draft is available online.
- **Textbook NLP with Python**
  - *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*, by Bird, Klein and Loper (1st edition). This book is freely available online

Faculty and Staff

AIT 726 Instructor: Lindi Liao, PhD
AIT 726 Course Coordinator: Lindi Liao, PhD & Ozlem Uzuner, PhD

Course Duration

Dates: GMU Academic Calendar: https://registrar.gmu.edu/calendars/
Total Duration: **8 weeks** or **16 weeks**

Tentative Course Outline

See the class schedule on Blackboard for a detailed list of topics that will be covered in this course. Every **class meeting** includes **lectures**, **demonstrations**, **tutorials**, **hands-on practice**, and/or **discussions**.

Lecture	Main Topic	Readings
1	Introduction to NLP and Deep Learning	Additional Materials
2	Sentiment classification with Logistic Regression	JM3 Ch4 JM3 Ch5 Additional Materials
3	Vector Representations	JM3 Ch6 Additional Materials
4	Neural Networks and Language Models	JM3 Ch7 Additional Materials
5	Sequence Processing with Recurrent Neural Nets	JM3 Ch9 Additional Materials
6	Attentions, Transformers, Pretrained Language Models	JM3 Ch10 & Ch11 Additional Materials
7	Formal Grammars of English	JM3 Ch17 & D Additional Materials
8	Dependency Parsing	JM3 Ch18 Additional Materials
9	Semantic Role Labeling	JM3 Ch24 Additional Materials
10	Literature Based Discovery	Additional Materials
11	Generative AI	Additional Materials
12	Ethics and Trustworthiness in NLP	Additional Materials
13	Multimodal AI (Language, Vision, Audio, etc.)	Additional Materials
14	Term Project Presentations	
15	Project Finalization	

Course Schedule

A **detailed** **schedule** for classes, topics, and assignment due dates is published on **Blackboard**. As many factors may affect the development and progress of a class, the instructor reserves the right to alter the schedule as may be required to assure attainment of course objectives. The schedule is subject to revision before and throughout the course.

Registered students should see the Blackboard Learning System for the latest class schedule.

Class Participation

This course is designed for mixed "***synchronous***" and "***asynchronous***" delivery (https://masononline.gmu.edu/course-delivery-methods):

- ***Synchronous***: Classes are held on a set schedule, and students virtually attend an instructor-led session on a regular basis.
- ***Asynchronous***: Students can study at their own pace, accessing instructional materials online in Blackboard at any time, though a set schedule with due dates is still there.

For this course, there are several *important* **online meetings** ("***Synchronous" mode***) as scheduled on **Blackboard Collaborate Ultra**. All students are required to virtually attend the instructor-led sessions on a regular basis.

All assignments, assessments, class announcements, schedules, files and presentations will use Blackboard.

Additionally, students need to contribute actively and participate in **online discussions** on Blackboard for grading.

Communication, Writing and Submissions

Communication: Course announcements will be made through Blackboard.

Writing: All discussions, and assignments for this course must be in standard English. Do not use slang or texting abbreviations (i.e., lol). Capitalize and use complete sentences in your discussion responses and in your paper. You can use bulleted lists if they make sense as a way to convey the information. Emoticons are acceptable as long as they are not overused and help with communication.

Before submitting work, be sure to proof read your writing and make sure that any references that you include are correct.

Submission of Work: All work for this class must be submitted as the assignment states.

ASA Style Guide: ASA Style Guides are easy to locate using an internet search. The following link is one that should work well for this class, you can access it by clicking here : http://personal.monm.edu/jkessler/ASA-Style.htm.

Academic Honesty

An important component in learning is taking on tasks, assignments and exams in an honest effort to do your best possible work. You are expected to turn in and do **original** work.

Grading Guidelines and Grade Scale

Grading Guidelines

Some grade components are evaluated subjectively
A: consistently above and beyond the course/assignment requirements
B: meets and occasionally exceeds the course/assignment requirements
C: minimally meets the course/assignment requirements

F: fails to meet the course/assignment requirements

Grades will be awarded in accordance with the Mason Grading System for graduate students. See the university catalog for policies: http://catalog.gmu.edu for more information.

Grading Scale

- The grading scale for this course, is:

97 – above	A+	Passing
93 – 96%	A	Passing
90 – 92%	A-	Passing
87 – 89%	B+	Passing
83 – 86%	B	Passing
77 – 82%	B-	Passing
70 – 76%	C	Passing
0 – 69%	F	Failing

NOTE: Study success takes constant effort!

Instructor will double check all students' coursework graded by GTA at the end of the course.

Raw scores may be adjusted by the Instructor to calculate final grades.

Students are responsible for checking the currency of their grade books. Grade discrepancies must be brought to instructor's attention within one week of assignment submission and 48 hours of exam submission.

No make-up for any activity, unless arranged in advance. Only in special cases, such as medical problems and family emergency, make-ups and late assignments may be allowed with verifiable proof.

Final grades will be posted to PatriotWeb, which is the only vehicle for students to obtain those grades. A student with a "hold" on his/her PatriotWeb account will be unable to access final grades until the hold has been removed by the Registrar.

Etiquette and Disabilities

Please observe proper "etiquette" and "netiquette" – courteous and appropriate forms of communication and interaction – within this course. This means no personal attacks, obscene language, or intolerant expression. All viewpoints should be respected.

Giving Feedback: This course is designed along the principles of synergy and collaborative learning. Therefore, it is important that all students understand how to provide quality feedback to their peers. Here are a few tips for providing, positive, constructive, and useful feedback to peers.

- Be empathetic and remember that this environment is a safe place for making mistakes
- Use nonjudgmental language and phrases that do not attack an individual. One way of doing this is to ask the individual to discuss his/her process for making the final decision.
- Use specific questions, examples, and references as a way of making your point.

- Make your feedback useful by providing suggestions that the individual can understand and use to improve her/his work.

Disabilities: Please message me if you have a disability so we can discuss ways to help you succeed in the course. If you need accommodations that would affect the terms of this syllabus, you will need to provide documentation of your disability.