XCS224N: NLP with Deep Learning

Syllabus and Course Information

Welcome

Welcome to XCS224N: Natural Language Processing with Deep Learning! This professional course is based on graduate-level material from Stanford's on-campus course CS224N, adapted for a professional certificate format. In this course you will:

- Learn from Stanford classroom lecture videos that have been edited and segmented by topic for easier navigation, reference, and review.
- Complete coding assignments with guided support and self-check functions.
- Receive support from Stanford-affiliated Course Assistants.
- Connect to a cohort of peers from diverse locations and professional backgrounds.

Course Launch

All lecture videos will be available on the first day of the course **(March 9th) at 12:00pm Pacific Time**. Course assignments will be made live on the days indicated in the calendar below, without exception. Maintaining the assignment schedule enables Course Facilitators to be most effective in providing support and answering questions on subject matter throughout the course.

Getting Started (Before March 9)

This course will use different tools to distribute content, run assignments, and deliver support. They are:

- **SCPD Learning Management System** – accessed via the mystanfordconnection site which you used to apply to and enroll in this course.
- **GitHub** – to distribute programming assignment code.
- **Azure Lab Services** – to provide cloud computing resources for Assignments 4 and 5.
- **Slack** – for additional course support and class discussions.

Joining Slack

In addition to direct small group support from Course Facilitators (more details and guidelines in Course Facilitators, Support, and Guidelines section below), the cohort will have a Slack workspace to ask additional questions and discuss course topics. An email invitation to http://xcs224n-scpd.slack.com will be sent to your email address on file with SCPD on **March 7**, however if you encounter any troubles you can go directly to the workspace URL and request access as well.

Joining GitHub

To gain access to the course GitHub, please fill out this Google Form with your account information. Once you fill out the form, you will receive an invitation to join a GitHub Team called "XCS224N Cohort 3 Students". **The team's repository will be blank to begin with** – code files will be added gradually as each assignment is released.

Getting Started (On March 9)

Accessing Your Course

1. On **March 9th after 12pm Pacific Time**, log in to the mystanfordconnection account you used when applying for the XCS224N course.
2. XCS224N: Natural Language Processing with Deep Learning will be visible as a live course. Click the link titled “Course Videos and Assignments” to enter our learning management system.

Course Calendar

Course videos are divided into two categories – core content and supplemental topics. Below is a ***potential*** pacing guide if you are interested in watching all videos within a 10-week window. However, you are free to view the videos at any pace you'd like, including viewing during the winter break period.

Also in the right hand column are assignment release and due dates. **NOTE:** While the dates are presented here on the same timeline as the ***potential*** video pacing guide, ***this is not meant to imply a 1:1 mapping of lecture videos to assignments*** (i.e., it is not true that Lecture 9 is the exact point to be reached to begin Assignment 4). Further information about relevant lectures for each assignment is provided when you log in to your course materials.

WEEK	POTENTIAL VIDEO PACING		ASSIGNMENT RELEASED	DUE DATE
1	Lecture 1 – Introduction and Word Vectors		**A1: March 9**	
	Lecture 2 – Word Vectors and Word Senses		**A2: March 14**	
2	Lecture 3 – Neural Networks			**A1: March 22**
	Lecture 4 – Backpropagation			
3	Lecture 5 – Dependency Parsing		**A3: March 28**	
	Lecture 6 – Language Models and RNNs			
4	Lecture 7 – Vanishing Gradients, Fancy RNNs			**A2: April 2**
	Lecture 8 – Translation, Seq2Seq, Attention			
5	Lecture 9 – More on Gated Recurrent Units and NMT		**A4: April 11**	
	Lecture 10 – Question Answering			
6	Lecture 11 – Convolutional Networks for NLP			**A3: April 16**
	Lecture 12 – Sub-word Models			
7	Lecture 13 – Contextual Word Embeddings		**A5: April 25**	
	Lecture 19 – Bias in AI			
8	*(Optional) – Lecture 14 – Transformers and Self-Attention Guest Lecture*			**A4: April 30**
	(Optional) – Lecture 15 – Natural Language Generation			
9	*(Optional) – Lecture 16 – Coreference Resolution*			
	(Optional) – Lecture 17 – Multitask Learning Guest Lecture			
10	*(Optional) - Lecture 18 – Constituency Parsing, TreeRNNs*			
	(Optional) - Lecture 20 – Future of NLP + Deep Learning Guest Lecture			**A5: May 17**

Assignments and Grading

Assignments will be made available via the SCPD course platform on the dates noted above in the course calendar. Below is a brief summary of what each assignment will entail:

Assignment 1

The first assignment is a warmup exercise that will include:

- A short coding assignment focused on calculating co-occurrence matrices
- An exploration of word vectors with an accompanying multiple choice quiz
- Optional derivation-based questions that are counted as **extra credit**

Assignment 2

The second assignment contains the following components:

- A multiple choice quiz
- A coding assignment where you will implement Word2Vec and train your own word vectors with stochastic gradient descent (SGD)
- Optional derivation-based questions that are counted as **extra credit**

Assignment 3

The third assignment will include a multiple choice quiz and a coding assignment in which you will use PyTorch to build a neural dependency parser.

Assignment 4

The fourth assignment will include a multiple choice quiz and a coding assignment in which you will build a Neural Machine Translation system using PyTorch on a GPU. SCPD has arranged for each student to train their model on a preconfigured VM using Microsoft Azure Lab Services, and will provide further information on how to access your resources.

Assignment 5

The fifth assignment will include a multiple choice quiz and a coding assignment in which you will improve upon your Assignment 4 model by implementing a character-based convolutional encoder and character-based LSTM decoder. You'll use PyTorch on a GPU. SCPD has arranged for each student to train their model on a preconfigured VM using Microsoft Azure Lab Services, and will provide further information on how to access your resources.

BERT Exercise (Ungraded Lab)

As an addition to the XCS224N professional course, SCPD has created an optional exercise in which you will compare the performance of word2vec embeddings and BERT embeddings for text classification. The lab will be a Python notebook via Google CoLab, and use a pre-trained model from HuggingFace. This exercise will have no due date and be ungraded.

In all coding assignments you will see functions labeled "test_function" at the bottom of the file. You can run these functions to check steps of your implementation and find out where, if any, errors are occurring in your code.

Honor Code

Students will be asked to review and maintain the standards set forth by the Stanford Honor Code when completing quizzes and assignments in this course. You can review the section labeled *Violations of the Honor Code* for representative examples relevant to this course.

Grading

Multiple Choice Quizzes will be graded immediately and your score will be shown. You will be allowed three attempts at each quiz.

Coding Assignments are graded automatically upon upload and will show your score. You can continue to re-submit up until the due date. To view an example of what this process looks like, you can view this video - https://youtu.be/kYaL33BSOYc **.** Scores will be transferred from the Gradescope submission site to your SCPD gradebook six days after the due date, after the five day extension window has also closed (see Late Assignments and One-time Extension Section below for details).

Note on Final Projects

In the adaption from the CS224N graduate course to the XCS224N professional course, the final project has been removed. If completing a project-based course is your top priority, we recommend considering the CS224N graduate course or the XCS224U professional course.

Late Assignments and One-time Extension

Late coding assignments are assessed a penalty of **one point per day late, up to a maximum of five days late at which point the submission link will close**.

Late quizzes are accepted without a scoring penalty, however as noted above you are limited to three total attempts regardless of when you submit.

We understand that personal or professional events may cause you to miss a deadline on a coding assignment. Each student is able to use a **one-time, five-day extension on a coding assignment, which will not be assessed a scoring penalty**. The extension cannot be split into smaller parts (e.g. you cannot use two days on Assignment 3 and three days on Assignment 4.). In order to use your extension, contact your Course Facilitator and SCPD staff.

Passing the Course and Earning the Certificate

In order to earn the Certificate of Achievement associated with this course, you must complete all quizzes and assignments with a total cumulative score of 70% or higher. Once you have successfully completed the course and the post-class survey, a digital Record of Completion will be emailed to you and the Certificate of Achievement will be mailed in approximately two weeks. If you are interested in calculating your progress along the way, it may be helpful to know:

- There are a total of 200 base points in the course (meaning 140 to achieve 70%)
- There are a total of 10 extra credit points in the course

Videos and Slides

As noted, this course utilizes content originally delivered in the CS224N graduate course. A few things you will notice about this adaptation process:

- At times you will hear instructors make reference to the final project or poster session. As noted above these have been removed for the current version of XCS224N and you need not worry about the reference.
- Instructors may make reference to “Week 1”, “Week 2”, “Week n” of the course – in general these references still hold due to the fact that the schedule/layout of the course has changed very little (except for Lecture 19) from the graduate course.
- Some sections of video have been removed, and accordingly those slides have been removed from the decks that accompany and play alongside each video. However, decks have not been re-numbered to ensure that they are still easy to match with the video segments. You will therefore see some slide decks that begin on a slide number other than 1 (often, as graduate course logistics videos have been cut) or slide decks that skip a few numbers due to removal of a clip (rarely, as much of the core lectures remain).
- In a few specific cases you may see names and/or faces blurred. In general this is usually due to guidelines regarding student privacy.

Course Facilitators, Support, and Guidelines

You have a wide range of support available to you throughout the course. You will be assigned and receive contact information for an individual Course Faciltiator (CF) who will act as your primary point of contact. Below is a summary of the available resources and course support:

Office Hours

Your CF will be in touch with availability and scheduling logistics for video conference office hours. Office hours may be conducted using the Zoom conference service or via Slack video (more information below on the course Slack workspace). Your CF will provide further information on how they will schedule and run office hour sessions.

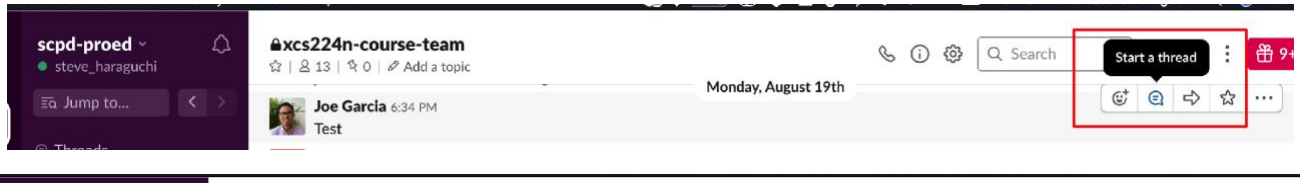
Email

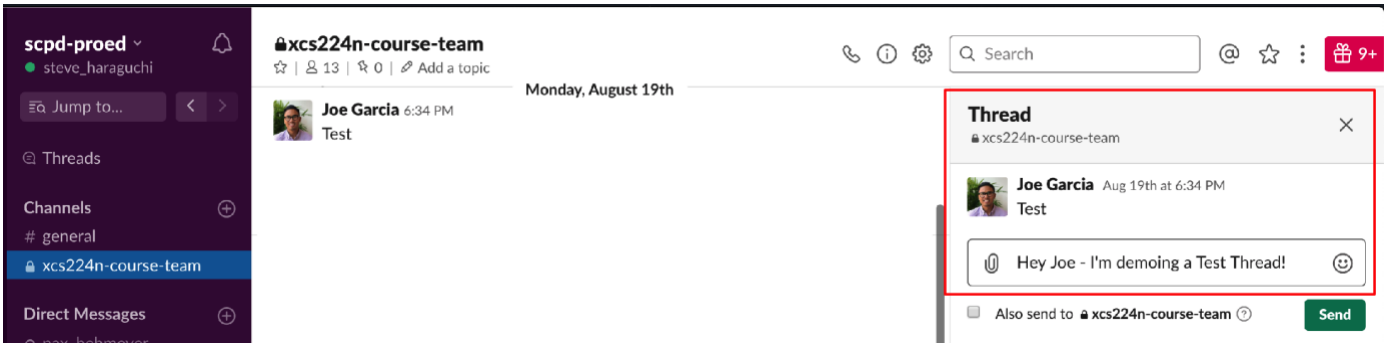
Your CF will also be available to answer questions via email – a Stanford contact address will be provided when you are first connected to your CF.

Slack Workspace – Usage and Guidelines

In addition to the individual and small group support provided by CFs, Slack will be a place where questions may be posed to the entire community (this is the fastest way to get an answer!). **In order to keep the Slack workspace readable, searchable, and useful to all, please follow the following guidelines:**

Reply in Threads to Keep Conversation Organized – When you are replying to a post or joining a conversation, respond by starting or joining a threaded conversation, rather than responding in the full flow of the standard timeline. See below for an example of how to respond in a threaded conversation to Joe's test message:





Use Multi-line Messages – Even if messages are threaded, you will soon see that Slack becomes unmanageable unless people use **single, multi-line messages instead of multiple, single-line messages**. Especially for mobile Slack users, it gets out of control!

Rather than the following:
"Hey all I have a question" [RETURN] <-- Creates new message
"I am a little confused about the quiz" [RETURN] <-- Creates new message
"I'm getting F for Question 40, but it seems like T is better" [RETURN] <-- Creates new message
---
Instead, try this!
"Hey all I have a question" [SHIFT+RETURN] <-- Creates new line in SAME message
"I am a little confused about the quiz" [SHIFT+RETURN] <-- Creates new line in SAME message
"I'm getting F for Question 40, but it seems like T is better" [SHIFT+RETURN] <-- Creates new line in SAME message
{RETURN} <-- Posts message

Note on Code Assignments and Debugging

While the course team is here to help and support your experience, it is ultimately your responsibility to write, test, and de-bug your own coding assignments. CFs may view and provide guidance on your work, however they will not send you exact answers on what to insert into your assignments. Additionally, before reaching out to a CF or Slack for help, it's expected that you have taken the reasonable steps of searching the internet for answers/clues and performing an analysis yourself. This policy is meant to ensure that you leave the course having mastered the material and enable CFs to focus attention on questions where their guidance is most impactful.

Questions and Contacts

For course-specific questions or concerns (content, assignments, CF support), please contact xcs224n-contact@lists.stanford.edu.

For questions on tuition, payments, certificates, and account access, please contact Student and Client Services at scpd-ai-proed@stanford.edu.