CSCI 662 Fall 2020 course page

Jonathan May

August 30, 2020

Website	`https://www.isi.edu/~jonmay/cs662_fa20_web/`
Lectures	`https://usc.zoom.us/j/92532775397` (password on piazza) , Mondays and Wednesdays 10:00–11:50 am
Instructor & office hours	Jonathan May, Online, Mondays and Wednesdays 9:00–10:00 am or by appointment
TAs & Office hours	Mozhdeh Gheini, Tuesdays and Thursdays, 1:00–2:00 pm, Online Meryem M'Hamdi, Mondays and Wednesdays, 4:00–5:00 pm, Online
Textbook	**Required:** Natural Language Processing - Eisenstein[1] **Required:** Selected papers from NLP literature, see (evolving) schedule *Optional:* Introduction to Deep Learning - Charniak [2] *Optional:* Speech and Language Processing 3rd edition -Jurafsky, Martin [3]
Grading	10 %: In-class participation 10 %: Posted questions before each in-class selected paper presentation 10 %: In-class selected paper presentation 30 %: Three Homeworks (10% each) 40 %: Project, comprising proposal (10%), final conference-quality paper (15%), and 20-minute in-class presentation (15%) (may be done in small groups)
Contact us	On Piazza or in class/office hours. Please do not email (unless notified otherwise).

Topics (subject to change per instructor/class whim) (will not be presented in this order):

- Linguistic Stack (graphemes/phones - words - syntax - semantics - pragmatics - discourse)
- Tools:
  - Corpora, Corpus statistics, Data cleaning and munging
  - Annotation and crowdwork
  - Evaluation
  - Models/approaches: rule-based, automata/grammars, perceptron, logistic regression, neural network models
  - Effective written and oral communication
- Components/Tasks/Subtasks:
  - Language Models
  - Syntax: POS tags, constituency tree, dependency tree, parsing
  - Semantics: lexical, formal, inference tasks

[1] `https://mitpress.mit.edu/books/introduction-natural-language-processing` or free version `https://github.com/jacobeisenstein/gt-nlp-class/blob/master/notes/eisenstein-nlp-notes.pdf`
[2] `https://mitpress.mit.edu/books/introduction-deep-learning` (first three chapters at `https://cs.brown.edu/courses/csci1460/assets/files/deep-learning.pdf`)
[3] `https://web.stanford.edu/~jurafsky/slp3/`

– Information Extraction: Named Entities, Relations, Events

– Generation: Machine Translation, Summarization, Dialogue, Creative Generation

date	material	reading	presentation	Other
8/24	intro, applications	Eisenstein 1 (not mandatory)		
8/26	probability basics, ethics	Eisenstein Appendix A, Goldwater probability tutorial [4]		project assignment out (due 9/9)
8/31	~~corpora, text processing~~, Linear Classifiers (Naive Bayes, Logistic Regression, Perceptron)	Eisenstein 2, Nathan Schneider's unix notes[5], Unix for poets[6], sculpting text[7]	The Social Impact of Natural Language Processing[8] **Presenter: Jon**	
9/2	Nonlinear classifiers, feed forward neural networks, backpropagation, gradient descent	Eisenstein 3, Charniak 1.	Thumbs up? Sentiment Classification using Machine Learning Techniques[9] **Presenter: Zekun**	HW1 out (due 9/30)
9/7	LABOR DAY NO CLASS			
9/9	POS tags, HMMs, search	Eisenstein 7	Fast Semantic Extraction Using a Novel Neural Network Architecture[10] **Presenter: Tooraj**	project proposal due
9/14	parsing and syntax 1: treebanks, evaluation, cky, grammar induction, pcfgs	Eisenstein 9,2, 10.	Part-of-Speech Tagging for Twitter: Annotation, Features, and Experiments[11] **Presenter: Chrysovalantis** Building a Large Annotated Corpus of English[12] **Presenter: Ani**	
9/16	parsing and syntax 2: dependencies, shift-reduce	Eisenstein 11	Generating Typed Dependency Parses from Phrase Structure Parses[13] **Presenter: Ang** A Fast and Accurate Dependency Parser using Neural Networks[14] **Presenter: Shweta**	
9/21	evaluation, annotation, mechanical turk	Eisenstein 4.5.	An Empirical Investigation of Statistical Significance in NLP [15] **Presenter: Nikolaos**	

[4] http://homepages.inf.ed.ac.uk/sgwater/teaching/general/probability.pdf
[5] https://github.com/nschneid/unix-text-commands
[6] https://www.cs.upc.edu/~padro/Unixforpoets.pdf
[7] http://matt.might.net/articles/sculpting-text/
[8] https://www.aclweb.org/anthology/P16-2096/
[9] https://www.aclweb.org/anthology/W02-1011/
[10] https://www.aclweb.org/anthology/P07-1071/
[11] https://www.aclweb.org/anthology/P11-2008/
[12] https://www.aclweb.org/anthology/J93-2004.pdf
[13] https://nlp.stanford.edu/pubs/LREC06_dependencies.pdf
[14] https://www.aclweb.org/anthology/D14-1082/
[15] https://www.aclweb.org/anthology/D12-1091/

9/23	semantics: word sense, propbank, amr, distributional lexical	Eisenstein 13, 14.	Linguistic Regularities in Continuous Space Word Representations[16]. **Presenter: Hongkuan** The word analogy testing caveat[17] **Presenter: Jihoon**	HW2 out (due 10/21)
9/28	NO CLASS			
9/30	language models: ngram, feed-forward, recurrent	Eisenstein 7	Catch Up	HW1 due
10/5	Machine Translation history, evaluation, statistical	Eisenstein 18.1, 18.2	Bleu: a Method for Automatic Evaluation of Machine Translation[18] **Presenter: Paras** Towards a Literary Machine Translation: The Role of Referential Cohesion[19] **Presenter: Yuchen**	
10/7	Neural Machine Translation, summarization, generation	Eisenstein 18.3, 19.1, 19.2	Effective Approaches to Attention-based Neural Machine Translation[20] **Presenter: Xiou** Neural Machine Translation by Jointly Learning to Align and Translate[21] **Presenter: Soumya**	
10/12	Transformers	Attention is all you need [22], Illustrated Transformer [23]	Get To The Point: Summarization with Pointer-Generator Networks[24] **Presenter: Qi**. Universal Neural Machine Translation for Extremely Low Resource Languages [25] **Presenter: Amirhesam**	
10/14	Large Contextualized Language Models (ElMo, BERT, GPT-N, etc.)	Illustrated BERT, ElMo, and co.[26]	Language Models are Unsupervised Multitask Learners [27] **Presenter: I-Hung**. Defending Against Neural Fake News [28] **Presenter: Mozhdeh R.**	HW3 out (due 11/11)
10/19	Catch Up/No Lecture		Language Models are Few-Shot Learners[29] **Presenters: Yufei, Wenxuan**.	

[16] https://www.aclweb.org/anthology/N13-1090.pdf
[17] https://www.aclweb.org/anthology/N18-2039.pdf
[18] https://www.aclweb.org/anthology/P02-1040
[19] https://www.aclweb.org/anthology/W12-2503/
[20] https://www.aclweb.org/anthology/D15-1166/
[21] https://arxiv.org/abs/1409.0473
[22] https://arxiv.org/abs/1706.03762
[23] http://jalammar.github.io/illustrated-transformer/
[24] GetToThePoint:SummarizationwithPointer-GeneratorNetworks
[25] https://www.aclweb.org/anthology/N18-1032/
[26] http://jalammar.github.io/illustrated-bert/
[27] https://d4mucfpksywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf
[28] https://papers.nips.cc/paper/9106-defending-against-neural-fake-news.pdf
[29] https://arxiv.org/abs/2005.14165

10/21	Information Extraction: Entity/Relation, CRF	Eisenstein 17.1, 17.2	25 years of IE[30] **Presenter: Justin**	HW2 Due
10/26	Information Extraction: Events, Zero-shot	Eisenstein 17.3	Events are Not Simple: Identity, Non-Identity, and Quasi-Identity[31] **Presenter: Basel**	
10/28	Blade Runner NLP/Bertology		GLUE: A Multi-Task Benchmark and Analysis Platform for Natural Language Understanding[32] **Presenter: Prateek**	
11/2	Text Games and Reinforcement Learning		The Bottom-up Evolution of Representations in the Transformer: A Study with Machine Translation and Language Modeling Objectives[33] **Presenter: Shuai**	
11/4	Dialogue	Eisenstein 19.3.	A Diversity-Promoting Objective Function for Neural Conversation Models[34] **Presenter: Peifeng** Personalizing Dialogue Agents: I have a dog, do you have pets too?[35] **Presenter: Akshat**	
11/9	Power and Ethics		Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data[36] **Presenter: Ali O.** Energy and Policy Considerations for Deep Learning in NLP [37] **Presenter: Ali A**	
11/11	How to write a paper	Neubig slides on Piazza	On Measuring Social Biases in Sentence Encoders [38] **Presenter: Bahareh**	HW3 Due
11/16	Presentations/TBD (QA? Bertology? Fairness and inclusion?)			
11/18	Presentations			
11/23	Presentations			

[30] in piazza or https://www.cambridge.org/core/journals/natural-language-engineering/article/twentyfive-years-of-information-extraction/0E5BB0D6AE906BB3C25037E2D74CA8F3/share/5ce1ad8430e190e282cc234c79c320c49906a7e2

[31] https://www.aclweb.org/anthology/W13-1203/

[32] https://www.aclweb.org/anthology/W18-5446/

[33] https://www.aclweb.org/anthology/D19-1448/

[34] https://www.aclweb.org/anthology/N16-1014/

[35] https://www.aclweb.org/anthology/P18-1205/

[36] https://www.aclweb.org/anthology/2020.acl-main.463/

[37] https://aclweb.org/anthology/papers/P/P19/P19-1355/

[38] https://www.aclweb.org/anthology/N19-1063/