

CSCI 544 Fall 2017 course page

Jonathan May

August 21, 2017

Website	https://www.isi.edu/~jonmay/cs544_fa17_web/
Lectures	SAL 101, Wednesdays and Fridays 8:00-9:50 am
Instructor office hours	Jonathan May, Wednesdays and Fridays 10–11am or 11am–12pm with appointment
Textbook	None required; <i>Speech and Language Processing, 2nd Edition</i> ¹ is optional. All required readings will be listed in the syllabus.
TA Office hours	Dong Guo, 12pm–2pm Tue, EEB 220 Siddharth Jain, 8am–10am Mon, SAL lab Ramesh Manuvinakurike, 8am–10am Thu, SAL lab
Homeworks (60% of grade)	8 x 7.5% of grade each. A mix of programming and written assignments, electronic submission due 11:59 PM local Pacific time per table below.
Late days	Four (4) <i>cumulative</i> with no more than two (2) per assignment. 50% penalty first day thereafter; no credit afterward.
Midterm (15% of grade)	Friday, October 6, 8:00–9:45 am. Covers lectures, readings, and homeworks from 8/23–10/4
Final exam (25% of grade)	Wednesday, December 6, 8:00–10:00 am. Covers lectures, readings, and homeworks from 8/23–11/29 (i.e. the entire class)
Contact us	On Piazza or in class/office hours. Please do not email.

Syllabus/schedule – subject to change – latest version is always on the website!

week/theme	date	material	reading	HW out	HW due
1: Corpus Processing	8/23	intro, applications	unix for poets ² , sculpting text ³ , hirschberg & manning ⁴		
	8/25	corpora, text processing, words, probability and statistics	NLTK ch.2 ⁵ , Chiang notes ch 1 ⁶	HW 1 (due 9/8)	
2: Morphology, regular languages	8/30	Morphology, regular expressions	jmv3 c2 ⁷		
	9/1	finite state automata and relationship to regex, finite state transducers, tree automata		HW 2 (due 9/15)	
3: Classification and tagging	9/6	Classifiers, features, naive Bayes, logistic regression, em			
	9/8	pos tagging, hmm, crf, search	collins notes		HW1
4: Syntax and Parsing 1	9/13	(guest lecture) parsing and syntax 1: treebanks, evaluation, cky, grammar induction	penn treebank	HW3 (due 9/27)	
	9/15	(guest lecture) parsing and syntax 2: pcfgs, restructuring, lexicalization, smoothing, beaming	collins notes		HW2
5: Syntax and Parsing 2	9/20	parsing and syntax 3: dependencies, mst and shift reduce algorithms	jm v3 c14	HW 4 (due 10/4)	
	9/22	NO CLASS			
6: Language Models	9/27	ngrams, bayes rule, chain rule	collins notes		HW 3
	9/29	smoothing, interpolation, feed forward	jm v3 c4 chen and goodman		
7	10/4	Review/catch up			HW 4
	10/6	MIDTERM			
8: Information Extraction	10/11	(guest lecture) Information Extraction 1	jmv3 c22 ⁸		
	10/13	(guest lecture) Information Extraction 2		HW 5 (due 10/27)	
9: Semantics 1	10/18	word senses, word net, prop bank			
	10/20	word sense disambiguation, distributional hypothesis, word2vec		HW 6 (due 11/3)	
10: Semantics 2	10/25	lambda calculus, ccg			
	10/27	meaning representations, AMR, para2vec	slp 17, 18, 21		HW 5
11: Machine Translation 1	11/1	history, evaluation, word alignment	alpac report , 20 years of bitext , Bleu, Brown et al.	HW 7 (due 11/15)	
	11/3	word alignment, phrase extraction, features			HW 6
12: Machine Translation 2	11/8	decoding, features, tuning			
	11/10	syntax, neural		HW 8 (due 12/1)	
13: Generation	11/15	Summarization			HW 7
	11/17	Poetry, Character-based RNNs			
14: Thanksgiving	11/22	NO CLASS			
	11/24	NO CLASS			
15: Applications	11/29	Information Retrieval/Question Answering			
	12/1	Review			HW 8