

שם הקורס: מבוא לעיבוד שפות טבעיות

מספר הקורס: 236299

סמסטר: אביב תשע"ח

מרצה:	ד"ר יואב גולדברג
שעות הרצאה:	
שעת תרגול:	
דרישות קדם:	
אתר הקורס:	

תאור הקורס

Natural Language Processing (NLP) is concerned with developing algorithms and computational models that analyze texts written in human language and perform useful tasks over them. In recent years NLP gained a lot of momentum in the industry and in research, with applications such as machine translation, sophisticated search and question answering, analysis of social media texts, language understanding for speech interfaces of mobile devices and more. The course goal is to introduce the basic tasks and algorithms involved in language processing, focusing mostly on statistical and machine-learning based methods.

The course focuses on two dimensions of natural language processing. The first dimension involves understanding the various levels of structure underlying human language, and the complexities, variability and ambiguities in human language. The second dimension will focus on modeling language phenomena in a computationally-tractable manner, and will mostly cover concrete machine-learning methods for various NLP problems, advancing students' knowledge and experience in machine learning as well.

Topics include:

- Concepts: Morphology. Parts of Speech. Chunks and Phrases. Named Entities. Syntax and Grammars. Anaphora.
- Tasks: Part-of-speech tagging. Entity recognition. Sequence segmentation. Information extraction. Syntactic Parsing. Coreference resolution .
- Methods and Algorithms: Maximum Likelihood Estimation, Hidden Markov Models and Viterbi, Discriminative Multi-class Classification, Discriminative Structured Classification, BIO-tagging, PCFG and CKY, Reranking, language-oriented feature design, distributional similarity and word embeddings, neural-network based models .

דרישות הקורס

The course will feature 4 homework assignments (40%), and a final project (60%).