What is Natural Language Understanding?

Natural language understanding:
- often refers to full comprehension/semantic processing of language;
- here, natural language understanding is used to contrast with natural language generation.

Understanding:

\[
\text{Text} \quad \Rightarrow \quad \text{Analyses (parse trees, logical forms, discourse segmentation, etc.)}
\]

Generation:

\[
\text{Non-linguistic input (logical forms, database entries, etc.) or text} \quad \Rightarrow \quad \text{Text}
\]
NLU covers advanced NLP methods, with a focus on learning representations, at all levels: lexicon, syntax, semantics, discourse.

We will introduce deep learning methods, covering:

- word embeddings;
- feed-forward neural networks;
- recurrent neural networks;
- recursive neural networks;

We will also compare deep learning models with conventional discriminative and unsupervised learning models.

Deep architectures and algorithms will be applied to NLP tasks:

- language modeling;
- part-of-speech tagging;
- parsing;
- semantic role labeling;
- semantic composition;
- sentiment analysis;
- discourse coherence.

The assignments will involve practical work with deep models.
Why do deep models work so well (for speech and vision at least)? Because they are good at representation learning:

Neural nets learn multiple representations $\mathbf{h}^n$ from an input $\mathbf{x}$.

Reasons for exploring deep learning:
- manually designed features are over-specified, incomplete and take a long time to design and validate;
- learned representations are easy to adapt, fast to obtain;
- deep learning provides a very flexible, trainable framework for representing world, visual, and linguistic information;
- deep learning can learn be unsupervised (from raw text) or supervised (with specific labels like positive/negative).

Introduction
Why Deep Learning?
Course Mechanics
The Success of Deep Models
Representation Learning
Unsupervised Models

Supervised vs. Unsupervised Methods

Standard NLP systems use a supervised paradigm:

Training:

Labeled training data \rightarrow \text{Features, representations} \rightarrow \text{Prediction procedure (trained model)}

Testing:

Unlabeled test data \rightarrow \text{Features, representations} \rightarrow \text{Prediction procedure (from training)} \rightarrow \text{Labeled output}
Recent work in NLP has focused on *unsupervised learning*, i.e., learning without labeled training data:

Unlabeled data $\Rightarrow$ Features, representations $\Rightarrow$ Prediction procedure $\Rightarrow$ Labeled output

Deep models can be employed both in a supervised and an unsupervised way.

Some unsupervised tasks we’ll cover:

**Part of speech induction:**
- walk
- runners
- keyboard
- desalinated $\Rightarrow$ walk.VVB
- runners.NNS
- keyboard.NN
- desalinate.VVD

**Topic models:**

![Topic models](http://www.progressarkansas.com/news.htm)


### Relationship to other Courses

Natural Language Understanding:
- requires: Accelerated Natural Language Processing OR Informatics 2A and Foundations of Natural Language Processing;
- complements: Machine Translation; Topics in Natural Language Processing.

Additional prerequisites:
- IAML or MLPR;
- CPSLP or equivalent programming experience.

Some overlap between NLU and MLP.
Background

Background required for the course:
- You should be familiar with Jurafsky and Martin (2009);
- but this textbook serves as background only; each lecture will rely on one or two papers as the main reading;
- you will need solid maths: probability theory, linear algebra, some calculus;
- for a maths revision, see Goldwater (2015).

Course Mechanics

NLU runs weeks 1–10, with 19 slots: 17 lectures, 2 feedforward sessions; no lectures in flexible learning week;
http://www.inf.ed.ac.uk/teaching/courses/nlu/
see course page for lecture slides, lecture recordings, and materials for assignments;
course mailing list: nlu-students@inf.ed.ac.uk; you need to enroll for the course to be subscribed;
the course has a Piazza forum; use it to discuss course materials, assignments, etc.;
assignments will be submitted using TurnItIn (with plagiarism detection) on Learn;
You need a DICE account! If you dont have one, apply for one through the ITO as soon as possible.

Assessment

Assessment will consist of:
- two assessed assignments, worth 15% each (30% in total);
- a final exam (120 minutes), worth 70%.
Due dates:
- Assignment 1 (Word Embeddings): issued Feb 3, due Feb 17;
Assignment deadlines will be preceded by feedforward sessions in which you can ask questions about the assignments.

Feedback

Feedback students will receive in this course:
- the course includes short, non-assessed quizzes;
- these consist of multiple choice questions and are marked automatically;
- each assignment is preceded by a feedforward session in which students can ask questions about the assignment;
- the discussion forum is another way to get help with the assignments; it will be monitored by the lecturers and TAs;
- the assignment will be marked within two weeks;
- individual, written comments will be provided by the markers and sample solutions will be released.
References


