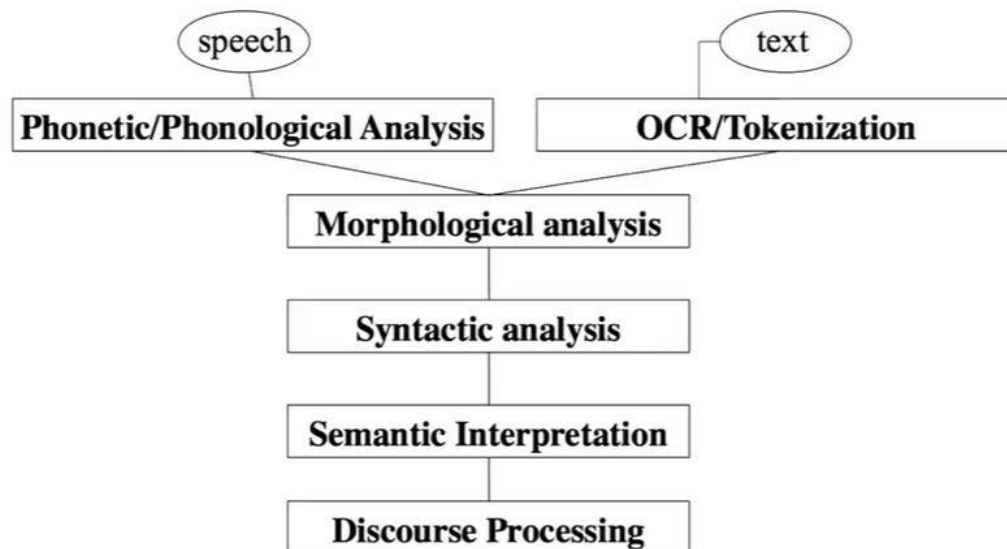


CS224n: Natural Language processing with deep learning - Lecture 1

Natural language processing is a intersection of computer science, AI and linguistic.

Goal for computers to process or “understand” natural language is to perform task or response questions (for example Siri, google assistant ...).

Fully understand and representing the meaning is difficult AI-complete.



(Syntactic analysis: structure of the sentences.

Semantic Interpretation: meaning of the sentences.

Discourse processing: context of the sentences)

The NLP mainly focus on dealing with Phonetic/Phonological Analysis, Syntactic Analysis and Semantic Interpretation.

(a tiny sample of) NLP Applications:

- Spell checking, keyword search, finding synonyms
- Extracting information from websites such as
 - Product price, dates, people or company names
- Classifying: positive/negative sentiment of longer documents
- Machine translation
- Spoken dialog systems
- Complex question answering

NLP in industry:

Search

Online advertising matching

Automated translation

Sentiment analysis for marketing or finance/trading

Speech recognition

Chatbot/dialog agents

Automating customer support

Controlling devices
Order goods

Unlike other random data a human language is a system specifically constructed to convey the speaker/writer's meaning.

A human language is a discrete/symbolic/categorical signaling system

The categorical symbols of a language can be encoded as a signal for communication:

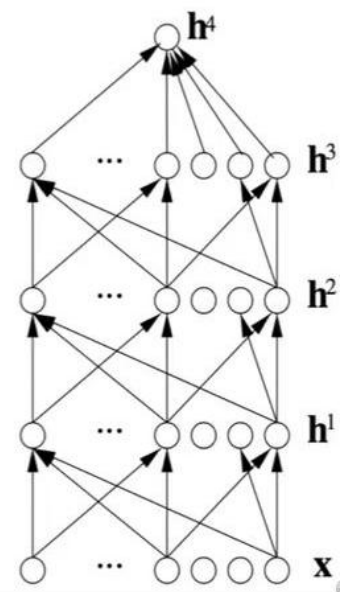
Sound
Gesture
Images(writing)

What is Deep Learning?

Deep learning is a sub-field of machine learning

Most machine learning methods work well because of human-designed representations and input features. Machine learning becomes just optimizing weights to best make a final prediction.

- **Representation learning** attempts to automatically learn good features or representations
- **Deep learning** algorithms attempt to learn (multiple levels of) representation and an output
- From “raw” inputs \mathbf{x} (e.g., sound, characters, or words)



Why deep learning?

Manually features are over-specified incomplete and take long time to design and validate, incomplete and take a long time to design and validate.

Learned features are easy to adapt, fast to learn.

Deep learning provides a flexible, universal, learnable framework for representing world, visual and linguistic information.

Deep learning can learn unsupervised (from raw text) and supervised (with specific labels like positive/negative)

In 2010~ deep learning techniques started outperforming other machine learning techniques because:

- Large amounts of training data favor deep learning
- Faster machines and multicore CPU/GPUs favor Deep learning
- New models, algorithms, ideas
 - Better, more flexible learning of intermediate representations
 - Effective end-to-end joint system learning
 - Effective learning methods for using contexts and transferring between tasks
- Improved performance (first in speech and vision, then NLP)

Key learning methods in this series of course:

- Ability to use the effective modern methods for deep learning
 - Covering all basics, but thereafter with a bias to the key methods used in NLP: Recurrent networks, attention, etc.
- Big picture understanding human languages and difficulties in understanding and producing them.
- Ability to build systems for some major problems in NLP:
 - Word similarities, parsing, machine translation, entity recognition, question answering, sentence comprehension

Why NLP is difficult: Ambiguous language unlike other programming or logic language.

For example newspaper headlines/tweets

1. The Pope's baby steps on gays (this can be structure ambiguous as one comprehension is the Pops's baby/ steps on/ gays; while its true meaning is The Pope's/ baby steps /on gays.)