**Lecture 1 | Natural Language Processing with Deep Learning**

**Video:** [**https://www.youtube.com/watch?v=OQQ-W\_63UgQ&list=PL3FW7Lu3i5Jsnh1rnUwq\_TcylNr7EkRe6**](https://www.youtube.com/watch?v=OQQ-W_63UgQ&list=PL3FW7Lu3i5Jsnh1rnUwq_TcylNr7EkRe6)

Notes:

**0:00**

Opening Remarks

“Our goal is how we can get computers to process or understand human languages in order to perform tasks that are useful.”

**6:16**

NLP Levels – the levels of human languages.

* Inputs: Speech and Text and their Processing
* Morphological Analysis: Working out the parts of complex **words**
* Syntactic Analysis: Working out the structure of **sentences**
* Semantic Interpretation: Working out the **meaning** of sentences
* Pragmatics/Discourse Processing: Meaning of sentences given greater context

**8:32**

A sample of NLP Applications from simple to complex:

* Simple
  + Spell Checking, Keywork Search, finding synonyms
  + Extracting information of websites
* More Complex
  + Classifying text
  + Understanding sentiment, reading level, etc
* Most Challenging
  + Machine translation
  + Spoken dialog agents
  + Complex question answering

More and more, NLP applications are becoming more commercially successfully

**12:21**

What’s special about human language?

Discusses how language is differentiated from other types of data dealt with by data processing tools.

* Human Language as data
  + Not an environmental signal; specifically used to communicate meaning
  + Language is a **symbolic system** encoded in a **continuous substrate**
    - Substrates: Sound (voice), images (writing; sign language)
    - Brains encode language using a continuous pattern of activation
  + Sparsity is a huge problem for machine learning of NLP

**17:32**

What is Deep Learning?

Explains that one key difference between traditional machine learning and deep learning is in creating the features. In machine learning, humans create the features and machines do numerical optimization to fit a function.

* Human Generated Representations (features) vs Machine Learned Representations

The *Deep* in Deep Learning comes from the presence of multiple layers of representations

**27:05**

Explains the phenomena of Deep Learning’s outperformance of other ML methods

* Why has DL started outperforming other methods (around 2010)?
  + Large amount of training data;
  + faster machines & high performing CPUs/GPUs
  + New models, algorithms and ideas

Gives a short history of the ramping success of DL since 2010.

**32:20**

Course Logistics and Introduce co-lecturer and TAs

**34:30**

Pre-requisites for course

* Python proficiency
* Multivariate Calculus, Linear Algebra
* Basic Probability and Statistics
* Fundamentals of Machine Learning

**35:51**

Goals of Course

* Understand and use effective modern NLP methods
* Big picture understanding of human languages and difficulties in understanding & producing them
* How to build systems for major problems in NLP

**36:29**

Grading Policy

**42:20**

NLP Overview

“Why is NLP Hard?”

* Clearing up ambiguities is a major difficulty of NLP. Compared to programming languages, which are non-ambiguous), human language is highly ambiguous.
* Also, human language has a lot of gaps. A lot of meaning is meant to be inferred.

Goes through a few examples of these cases.

**52:40**

“Deep NLP” combining the ideas and goals of NLP with using representation learning and deep learning methods to solve them.

**55:35**

Gives a little more detail of the methods we’ll encounter in the course. Including:

* + - 1. Word Representations
         * Representing word meaning as vectors in a high dimensional space
         * Words with similar meanings will cluster
         * Directions in these vector spaces will convey components of meaning
      2. Morphology
         * Parts of words (morphemes) can be represented as vectors; Neural networks can combine compose the meaning of larger units out of these smaller pieces
      3. Syntactic Parsers
         * Determine the structure of sentences
      4. Chatbots/Dialogue Agents
      5. Machine Translation – early computational efforts focused on this (cold war)
  + Neural language models
    - 1. Semantics
         * Traditional: Use lambda calculus to give meaning to individual meaning by hand, then combine them to further convey meaning using some careful logical algebra
         * DL: Since every word and logical expression is a vector, we can combine these elements to further convey meaning
         * To understand meaning between phrases and between objects and phrases we use NNs trained on these vectors to make decisions for us

1:10:10

“It’s all Vectors!”

Amazing thing is that we are using vectors to represent all types of language (sounds, words, parts of words, etc)