CS224n

Thomas Lu

1   Lecture 1: Introduction

This lecture basically just gives a high level overview of the fields of Natural Language Processing (NLP) and Deep Learning (DL). Some highlights from the lecture:

- NLP is a study that aims for computers to understand natural (human) language well enough to perform useful tasks.

- NLP has a lot of useful applications (virtual assistant, customer support automation, etc.) and has taken off commercially in the past few years.

- What is special about NLP (vs. other subfields of ML)?

    - Language is specifically constructed to convey information. It's not an environmental signal (e.g. an Amazon purchase, which I probably made for a reason other than conveying information).

    - Language is (mostly) a discrete/symbolic/categorical system, but our brains, which process language, are continuous systems.

- What is special about DL (vs. other ML techniques)?

    - Traditionally, learning an ML model requires an engineer/researcher to design and build relatively human-interpretable features. The machine then basically does some numerical optimization on these features to produce something useful. Note that this still requires a lot of intuitive human work to interpret data and figure out what parts of it might be useful.

    - In contrast, deep learning just feeds "raw data" or very simplistic representations of data to a deep neural net, which then learns its own effective representations of the data, obviating the need for feature engineering.

    - DL has performed very well recently, e.g. speech recognition, ImageNet.

- Why is NLP hard?

    - Human language contains lots of contextual information. The meaning of a sentence may depend on something that was said a few sentences ago, a piece of information from another document, current events, or something observable in the environment ("Did you see that dog?" refers presumably to a dog in the speaker's field of vision).

    - Human language is also ambiguous by design; since speech is pretty slow at transmitting information, people will almost always leave "relatively obvious" things unsaid in order to improve communication efficiency. Listeners are expected to infer these things. This is very different from computer languages, which must specify every posibility.