What is NLP?

How NLP Works

Crash Course AI #7 - NLP

DISCUSSION

How do you interact with computers through language?

- When is it Natural language *Understanding*?
- When is it Natural language *Generation*?

Crash Course AI #7 - NLP

DISCUSSION

- Give an example of a word (or a few words) that can have different meanings in different contexts

Crash Course AI #7 - NLP

DISCUSSION

- Describe a count vector in your own words.
- Give some examples of words that might be stop words
  - Note: there is no “one true list” of stop words. Different programs use different words.

Crash Course AI #7 - NLP

DISCUSSION

- Describe what an encoder and decoder do in your own words.

Crash Course AI #7 - NLP

DISCUSSION

- Fill in the blank: I’m kind of hungry, I think I’d like some chocolate _____
- Why did you choose that word?
- What is a word that makes no sense here? Why does that word not make sense?

Crash Course AI #7 - NLP

DISCUSSION

- Fill in the blank: Diana, a friend of mine from San Diego who really likes physics, is having a birthday party next week, so I'd really like to buy a present for _____
- Why did you choose that word?

Crash Course AI #7 - NLP

DISCUSSION

- Fill in the blank: Diana, a friend of mine from San Diego who really likes physics, is having a birthday party next week, so I'd really like to buy a present for her that has to do with _____
- Why did you choose that word?

Crash Course AI #7 - NLP

DISCUSSION

Ray Mooney said “We’ll never be able to capture the meaning of a word in a single vector.”

- Do you agree? Why or why not?

Crash Course AI #7 - NLP

DISCUSSION

- What kinds of data were the projects you played with in the first part using?
  - Google Semantris and Talk To Books
  - Botnik predictive text
  - Speech recognition
  - Google Translate
  - Article summaries

Crash Course AI #7 - NLP

RNNs

That was a lot! Let's summarize!

- Humans don't, by default, know how to turn a word into computer-understandable numbers.
- So, we have the computer "guess" the representation for us.
- The computer starts by representing words first as a list of random numbers (a vector - a kind of matrix), which it updates as it learns.
- The computer "guesses" the likelihood of any given word, which the computer then feeds back into the model to make predictions of what the next word will be.
- Similar words wind up with similar vectors at the end of this process.
- Ex: "physics" is close to "newton"

Crash Course AI #7 - NLP

DISCUSSION

- Come up with your own example of a word that would have different connotations in different contexts

Crash Course AI #7 - NLP

DISCUSSION

- How do computers understand language?

Overview of the video (in case you can’t watch, or just want notes)

Crash Course AI #7 NLP - overview

- **Natural language understanding:** Making meaning from what people say or from text
- **Natural language generation:** Creating text that is similar to what humans can make
- What is hard about NLP?
  - Words don’t have meaning on their own; we assign meaning to them.
  - The meaning of the word depends on context.
- How do we learn what words mean:
  - Someone labels things for us: “This is a cat.”

Crash Course AI #7 NLP - overview

- Machines learn words from shared parts of words (called morphology)
  - For example: “swim”/“swimmer”
  - This sometimes doesn’t work, like “van”/“vandal”, “cat”/“car”
- Machines learn words from **distributional symantics**, which means seeing which words appear in the same sentences a lot.
  - They do this through **count vectors**.
  - Count vectors count the number of times a word appears in the same article or sentence as other common words.
  - After removing stop words (like “and” and “the”) the wikipedia articles for “cat” and “felidae” are similar, but “cat” and “car” are not.

Crash Course AI #7 NLP - overview

- Count vectors have to store a LOT of data, looking at every single word that has ever been together with another word in the same sentence. Enter Machine Learning! Computers learn representations that maintain the relationships of count vectors.
- One way to learn these relationships is through encoders and decoders (described in the next slide).

Crash Course AI #7 NLP - overview

Encoder/Decoder Model

- **Encoder**: what should you think or remember about what was just read
- **Decoder**: uses the thought to decide what we're going to say or do
- Ex: I'm kind of hungry, I think I'd like some chocolate _____
  - More likely to be "milk" or "cake" than "potato".
  - The encoder only needs to think about "chocolate".
- Ex: Diana, a friend of mine from San Diego really likes physics, is having a birthday party next week, so I want to find a present for _____
  - Remember "Diana" (from 27 words ago).
  - Remember Diana uses the pronoun "her".
  - The encoder needs to capture all this information!
- Diana, a friend of mine from San Diego who really likes physics, is having a birthday party next week, so I want to find a present for her that has to do with _____
- A human remembers "physics", so the encoder needs to remember that as well

Crash Course AI #7 NLP - Overview

- Machine learning needs input. In this case the input data is lots of sentences that have been collected (from books and conversations).
- For each word in the sentence, the model tries to guess the next word. Since the input actually has the whole sentence, the model can learn if it was correct or not.
- The most common model to use is a **Recurrent Neural Network**. Recurrent Neural Networks have a loop that lets them reuse one single hidden layer as they build the sentence.
    - Each word starts as a random interpretation, then corrects as it learns the truth.
    - Eventually, the **encoder** turns the idea into a representation of the whole sentence.

Crash Course AI #7 NLP - Overview

- The **decoder** is single layer network with input of the sentence representation. It outputs a score for every possible word.
- That output gets sent back to the encoder layer.
- The model learns weights for words, to build vectors.
- We can make a plot of those vectors, and find similar words near each other

Crash Course AI #7 NLP - Overview

- Models built this way can do things like:
  - Translate English to Spanish
  - Match questions to possible answers
  - Create actions from instructions
- Representations for one context doesn't necessarily work in others.
  - Ex: if you've only learned about recipes you probably think "roses" are more like decorative "icing" rather than "plants".

Exit ticket: Unit 1.06b - How NLP Works

How do computers understand language? How is this different from how humans understand language?