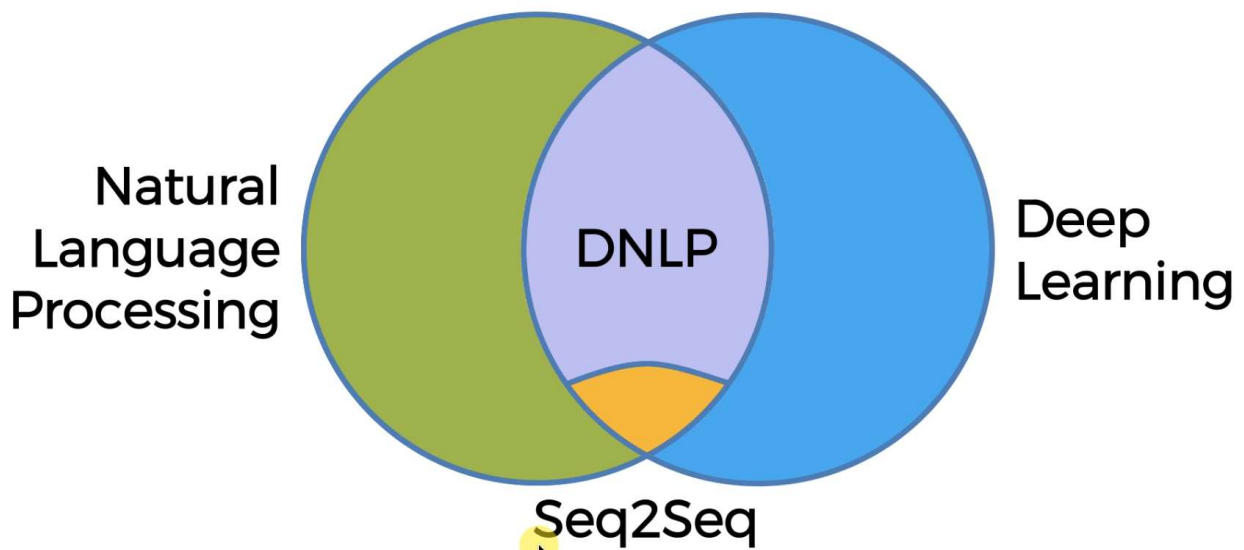


NLP Intuition

Tuesday, December 17, 2024 10:22 AM

Types of Natural Language Processing
Classical vs Deep Learning Models
Bag-Of-Words

Types of NLP



- Deep Natural Language Processing

Classical vs Deep Learning Models

Tuesday, December 17, 2024 10:29 AM

Examples:

If/Else Rules (Chatbot)

- Natural Language Processing
- A huge list of possible questions and the answers

Audio Frequency Components Analysis (Speech Recognition)

- Natural Language Processing
- Look at soundwaves of someone talking, and identify the waveforms that exist

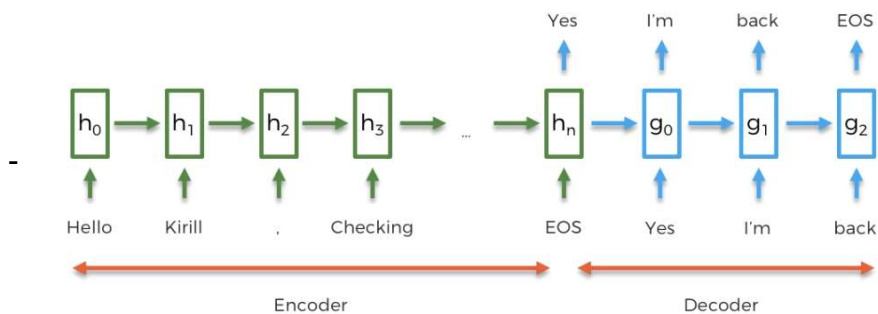
Bag-of-words model (classification)

- Natural Language Processing
- Text analysis. Given a comment, assign each word with a numerical value. Trying to associate comments with pass/fail

CNN for text recognition (Classification)

- Deep Natural Language Processing Model
- A neural network mainly used for image/video processing

Seq2Seq (many applications)



Bag-Of-Words

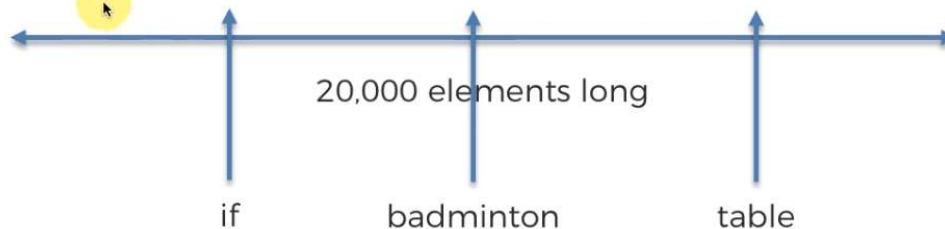
Tuesday, December 17, 2024 10:41 AM

Given a string of text, we want to respond with a yes/no response

Start with a vector full of 0's

- 20,000 elements long.
- This is the number of words that Americans commonly use
- Every word in the English language will have a position in this vector

[0, ... , 0]



First element is reserved for SOS and second element is reserved for EOS

- SOS = Start of Sentence
- EOS = End of Sentence

Last element in array is reserved for Special words

- Words that we don't recognize within the 20,000 words

Commas and punctuation also count as a word

For each position, put in the frequency of words

Our goal is to come up with a yes/no response

- We are going to do that with training data of all emails previously sent

Training Data:

- Hey mate, have you read about Hinton's capsule networks? → **No**
- Did you like that recipe I sent you last week? → **Yes**
- Hi Kirill, are you coming to dinner tonight? → **Yes**
- Dear Kirill, would you like to service your car with us again? → **No**
- Are you coming to Australia in December? → **Yes**

Convert each training emails into a 20,000 long vector

There are multiple ways we can apply the model

- One way is using the logistic regression
- Another way is by using a neural network
 - o 20,000 neurons, and go through as many hidden layers as needed, and then a output layer