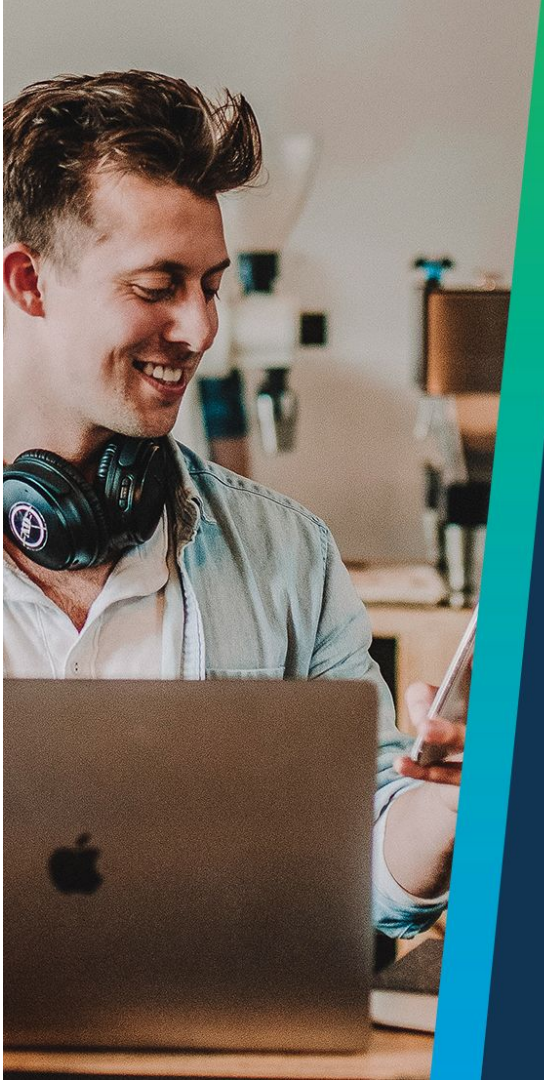


SE Bootcamp

Hyperiondev

Introduction to Natural Language Processing

Welcome

Your Lecturer for this session

Sanana Mwanawina

Lecture - Housekeeping

- The use of disrespectful language is prohibited in the questions, this is a supportive, learning environment for all - please engage accordingly.
- No question is daft or silly - **ask them!**
- There are Q/A sessions midway and at the end of the session, should you wish to ask any follow-up questions.
- You can also submit questions here: hyperiondev.com/sbc4-ds-questions
- For all non-academic questions, please submit a query: hyperiondev.com/support
- Report a safeguarding incident: hyperiondev.com/safeguardreporting
- We would love your feedback on lectures: https://hyperionde.wufoo.com/forms/zsgv4m40ui4i0g/

Lecture - Code Repo

Go to: github.com/HyperionDevBootcamps

Then click on the "**C4_DS_lecture_examples**" repository, do view or download the code.

Objectives

1. Define NLP
2. Understand the process of NLP
3. Use SpaCy

Natural Languages

- Natural language = human language (not programming language)
  - **Phonetics and Phonology** - linguistic sounds.
  - **Morphology** - meaningful components of words.
  - **Syntax** - structural relationships between words.
  - **Semantics** - knowledge of meaning.
  - **Pragmatics** - relationship of meaning to goals and intentions of speaker.
  - **Discourse** - linguistic units larger than a single utterance.
- Lots of stuff for a computer to know!

Ambiguity

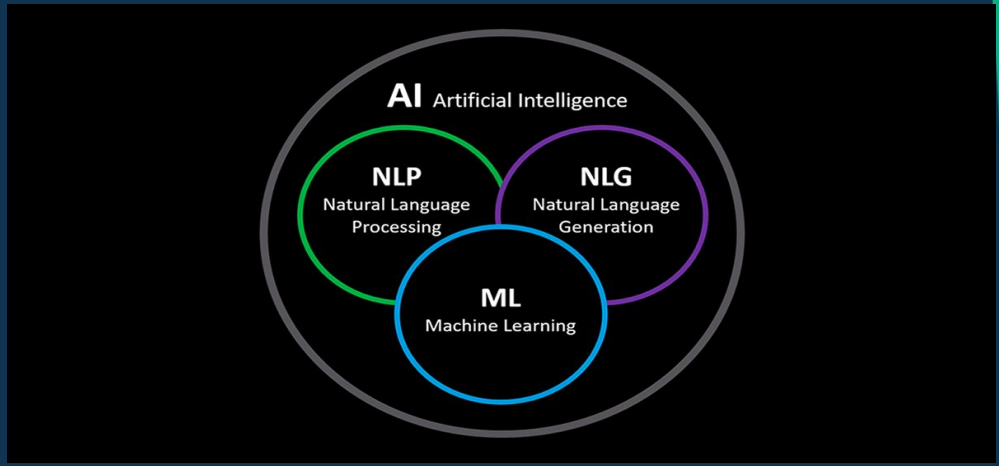
- Scenario - Tony is telling JARVIS that he was annoyed at a female worker, and threw a piece of paper. He then proceeds to say “I made her duck.”
- How does JARVIS interpret this?
  - I cooked a duck for her.
  - I cooked a duck belonging to her.
  - I created the duck that she owns.
  - I caused her to quickly lower her head.
  - I turned her into a duck (possibly with the help of Dr. Strange?).
- English (and natural languages generally) is weird and ambiguous. This makes it difficult for computers to navigate the way we speak.

Parts of Speech (POS)

- To know what “duck” means, we use POS tagging.
- We can identify what it means based on its position in the sentence.
- E.g. “The old man the boat”
  - The - determiner
  - Old - adjective
  - Man - noun/verb
  - The - determiner
  - Boat - noun.
- From its position in the sentence, you can see that “man” is a verb in this sentence.

How does NLP Work?

- Goal is to simulate human intelligence.
- Machine Learning techniques used to train a model to understand human language.
- Probability-based - therefore not always 100% accurate, but close enough, and getting closer every day!



Solving POS Tagging

- Give a program a bit set of tagged words (supervised training), and ask it to find patterns on an unseen sentence (test set).
- POS tagging tries to tag words with correct POS tag. Once we understand the POS of each word, we can then parse the sentence.
- Parsing - putting the sentence together in the right way so that it can be understood.

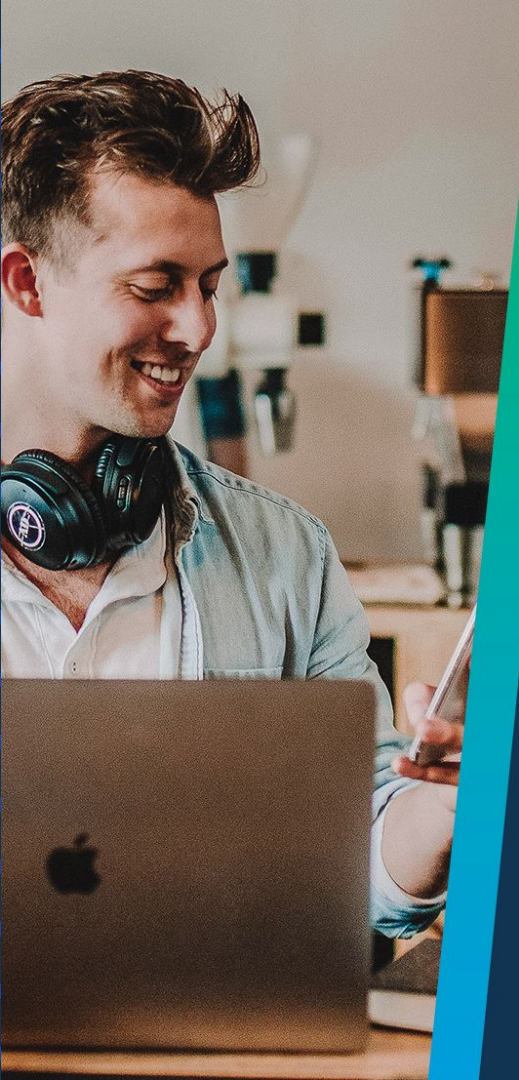
SpaCy

- A software package that comes with all of these models already!
- We won't need to be generating models and training with data.
- We give it a sentence, it gives us linguistic data about the sentence.

Hyperiondev

Q & A Section

Please use this time to ask any questions relating to the topic explained, should you have any



Hyperiondev

Thank you for joining us