

PyPharma NLP Workshop 2019

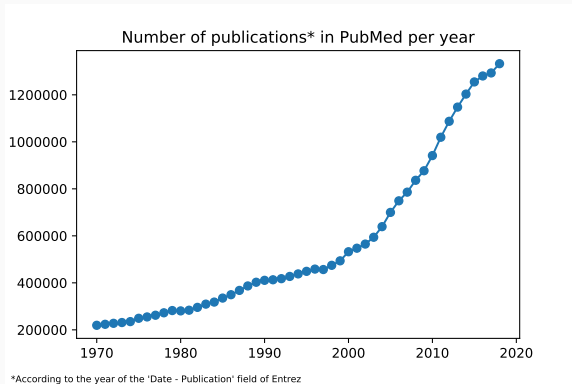
Diego Saldana

November 2019

Introduction to Biomedical NLP

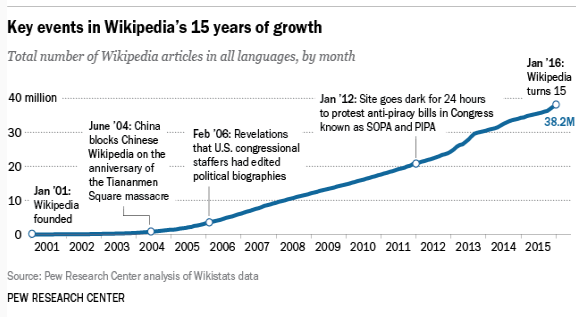
Why Biomedical NLP? (1/3)

- Most of the information out there is in the form of natural language: scientific papers, clinical notes, social media, textbooks, lectures, websites.



Why Biomedical NLP? (2/3)

- Most of the information out there is in the form of natural language: scientific papers, clinical notes, social media, textbooks, lectures, websites.



Why Biomedical NLP? (3/3)

- Most of the information out there is in the form of natural language: scientific papers, clinical notes, social media, textbooks, lectures, websites.
- This information is potentially very useful but cannot readily be used programmatically and stored in databases, searched, or analyzed.
- As a result this valuable information is “locked into a vault” until a human reads it, structures it and puts it into some database.
- And even when that happens, the scope in which the data can be used is usually limited and chosen by the extractors.
- How can machines help?

Humans vs. Machines (1/2)

- Machines and humans have different strengths and weaknesses when processing text.
- Machines in particular are capable of processing vast amounts of text in a very short period of time in a very consistent way and performing simple tasks.
- Humans take much more time to process text and are less consistent, however they are capable of much more complex reasoning and understanding.

Humans vs. Machines (2/2)

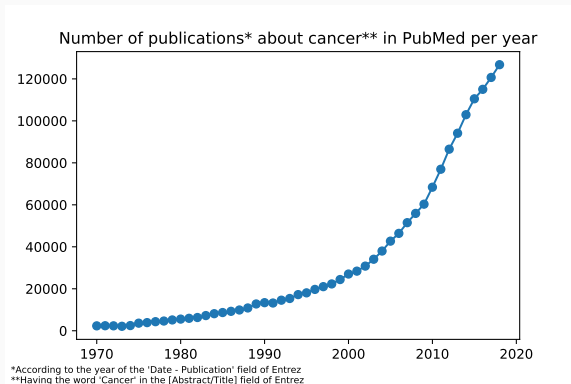
What are some examples of tasks can computers perform well in 2019?

- Categorizing documents (e.g. automatically assigning MeSH headings to PubMed abstracts)
- Extracting entities from text (e.g. extracting Drugs, Diseases from PubMed abstracts)
- Extracting relations from text (e.g. extracting Adverse Events from PubMed abstracts)
- Answering simple questions based on a small amount of context (e.g. “Which drug should be used as an antidote in benzodiazepine overdose?”)

Backup

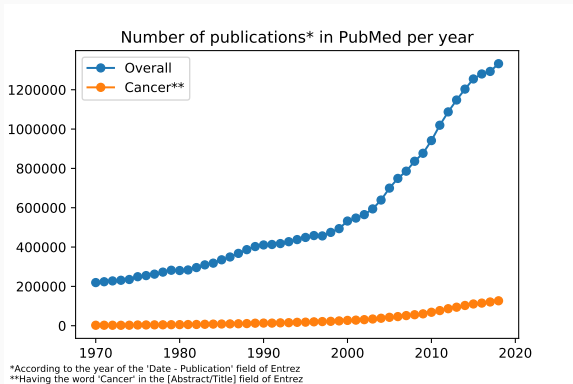
Why Biomedical NLP?

- Most of the information out there is in the form of natural language: scientific papers, clinical notes, social media, textbooks, lectures, websites.



Why Biomedical NLP?

- Most of the information out there is in the form of natural language: scientific papers, clinical notes, social media, textbooks, lectures, websites.



Agenda

- Biomedical NLP 101: Bags of words (30 mins)
- Deep Learning for Biomedical NLP (30 mins)
 - Language Modelling (30 mins)
 - Text Classification (30 mins)
 - Named Entity Recognition (30 mins)
 - Question Answering (30 mins)
 - Integrating NLP into survival models (30 mins)