uc3m | Universidad Carlos III de Madrid

HULAT Human Language & Accessibility Technologies

Cutting-edge Deep Learning for NLP learners

Introduction

Isabel Segura-Bedmar

6 June 2022

Máster Universitario en Inteligencia Artificial, UPM

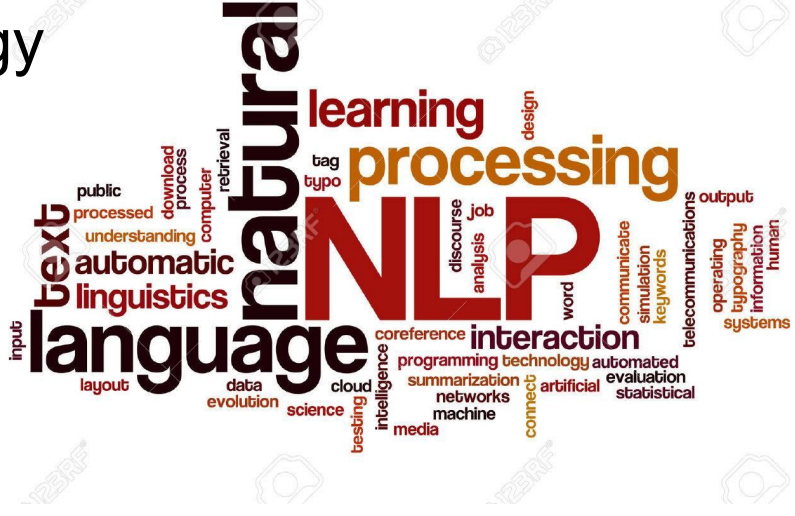
Outline

- What is Natural Language Processing? Why NLP? NLP Applications
- What is Deep Learning?
- Traditional Machine Learning vs Deep Learning
- Deep Learning applied to NLP
- Some examples of Deep Learning architectures for NLP tasks

Natural Language Processing (NLP)

- Computational techniques for the automatic analysis and representation of human language

- Multidisciplinary: Linguistics, Mathematical and Computer Science, Psychology



Why NLP?

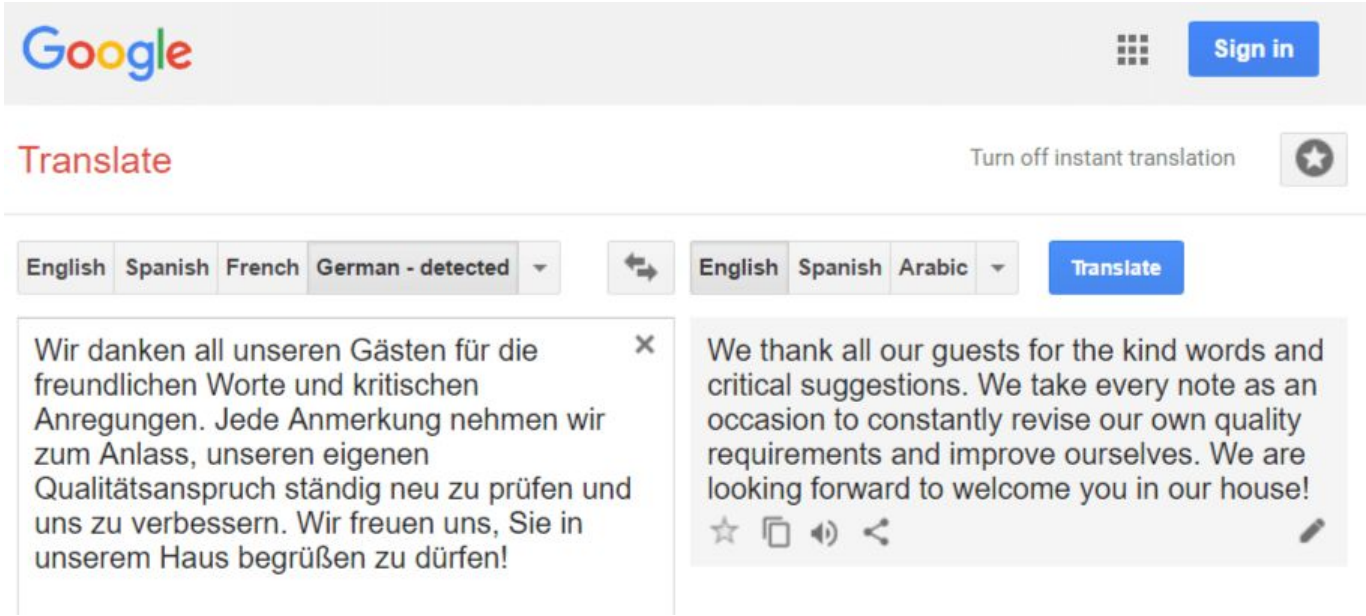
- Exponential Growth of data.
  - 2013, 3.5 ZB
  - 2022, 40 ZB
  - 2025, 180 ZB
- > 80% in unstructured form, primarily texts

*1 ZB = 1 trillion GB

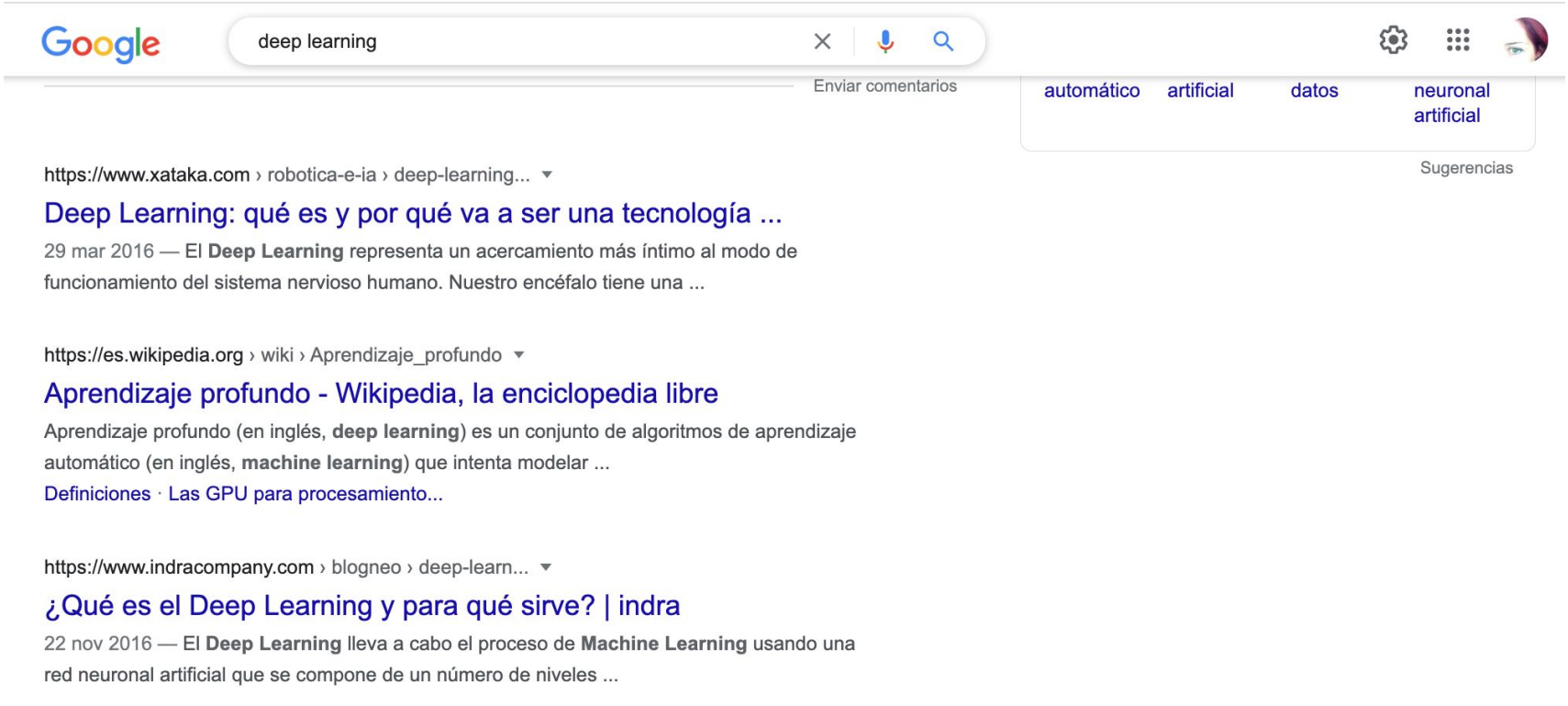
NLP applications

- Machine translation,
- Information retrieval,
- Information extraction,
- Text classification,
- Question answering,
- Text summarization
- Text simplification,
- Conversational AI,
- ...

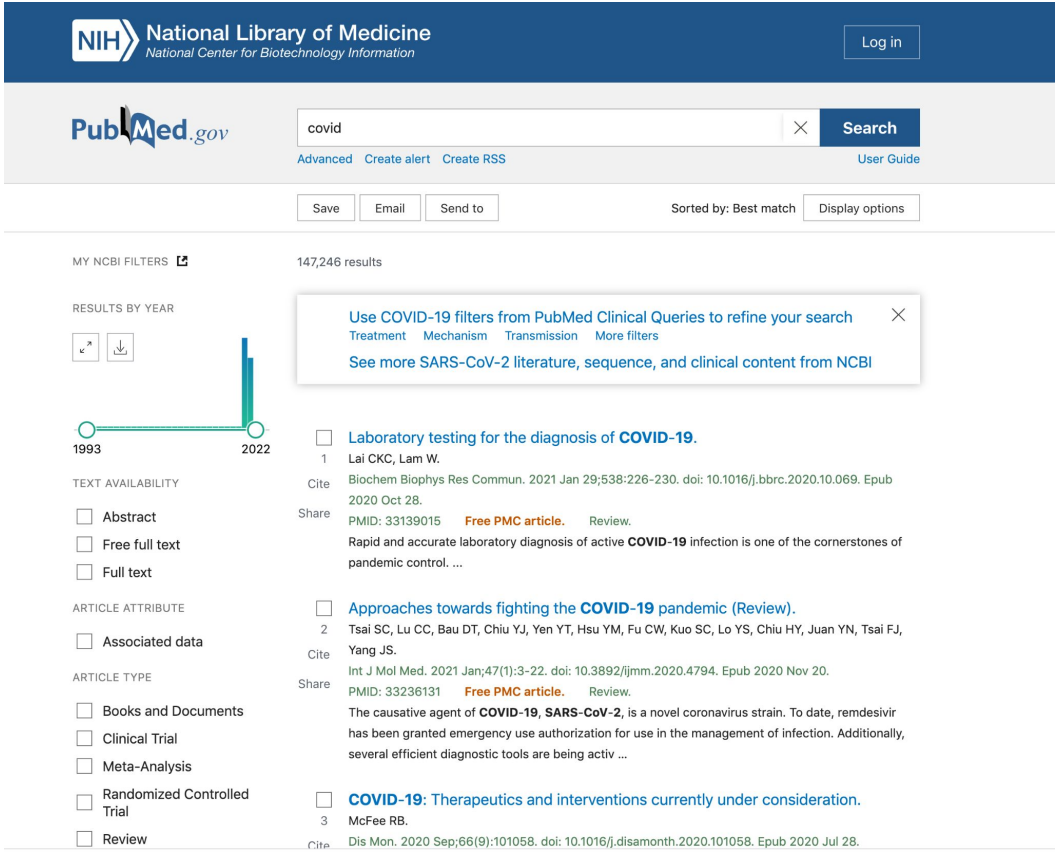
Machine Translation (MT)



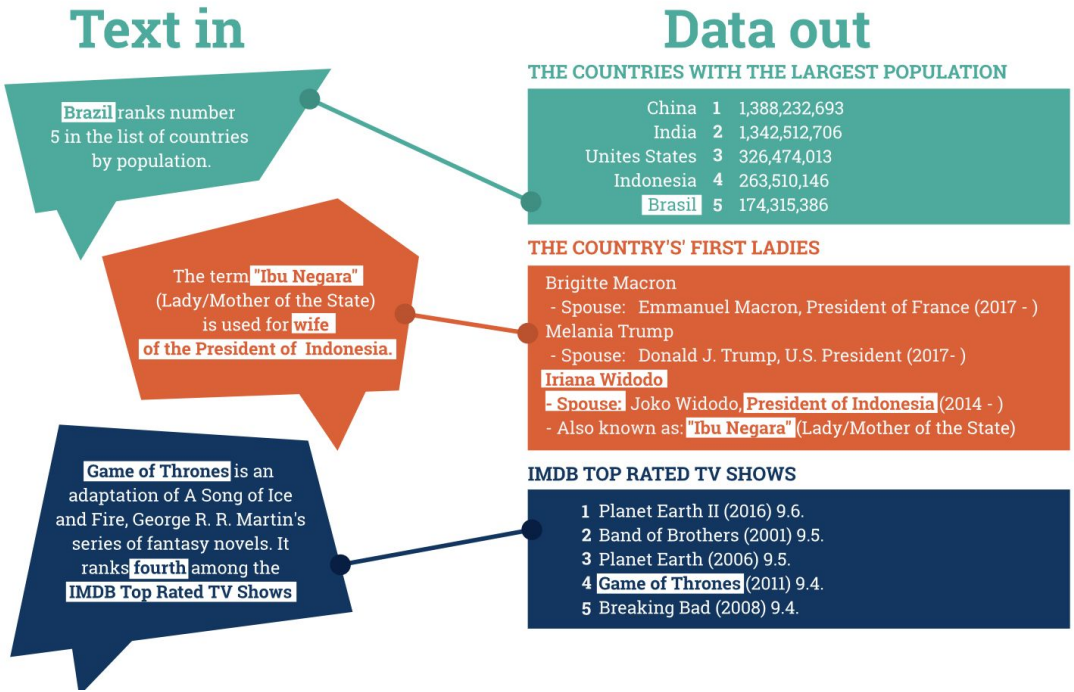
Information Retrieval



Information Retrieval



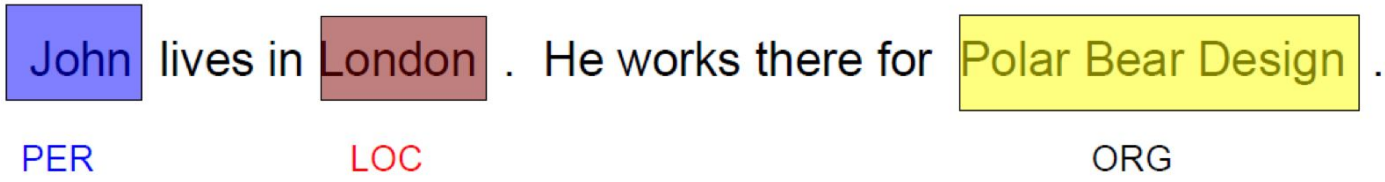
Information Extraction



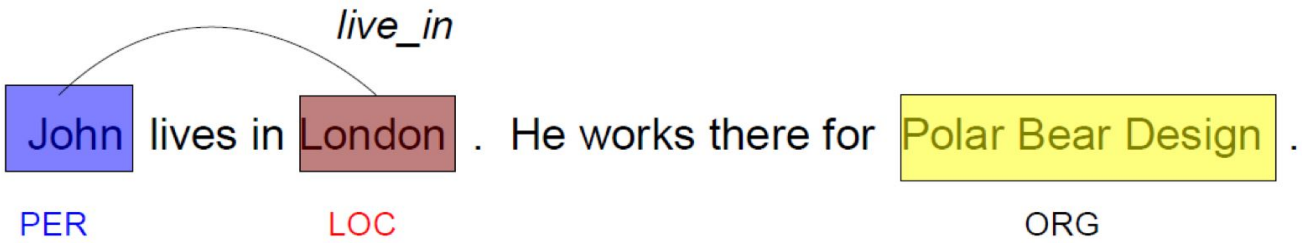
https://www.ontotext.com/knowledgehub/fundamentals/information-extraction/

IE tasks

1) Named Entity Recognition (NER)



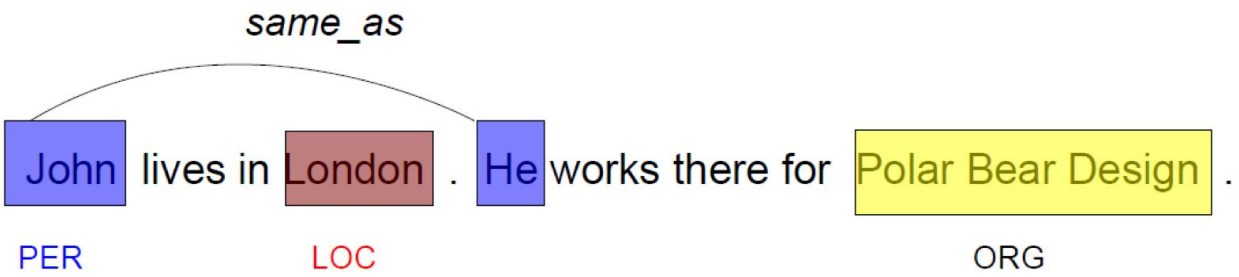
2) Relation Extraction

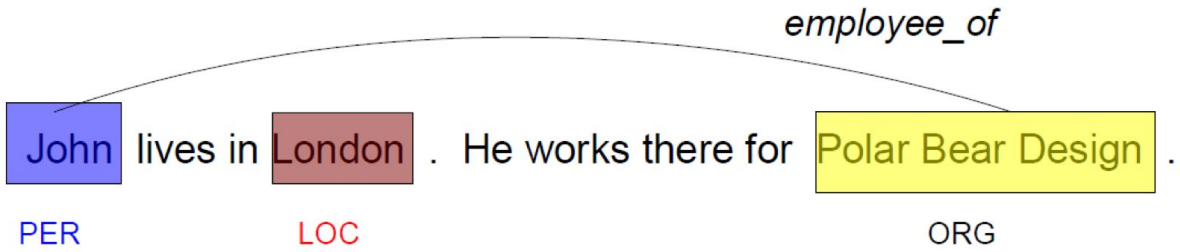


Source: https://gate.ac.uk/family/training.html

IE tasks

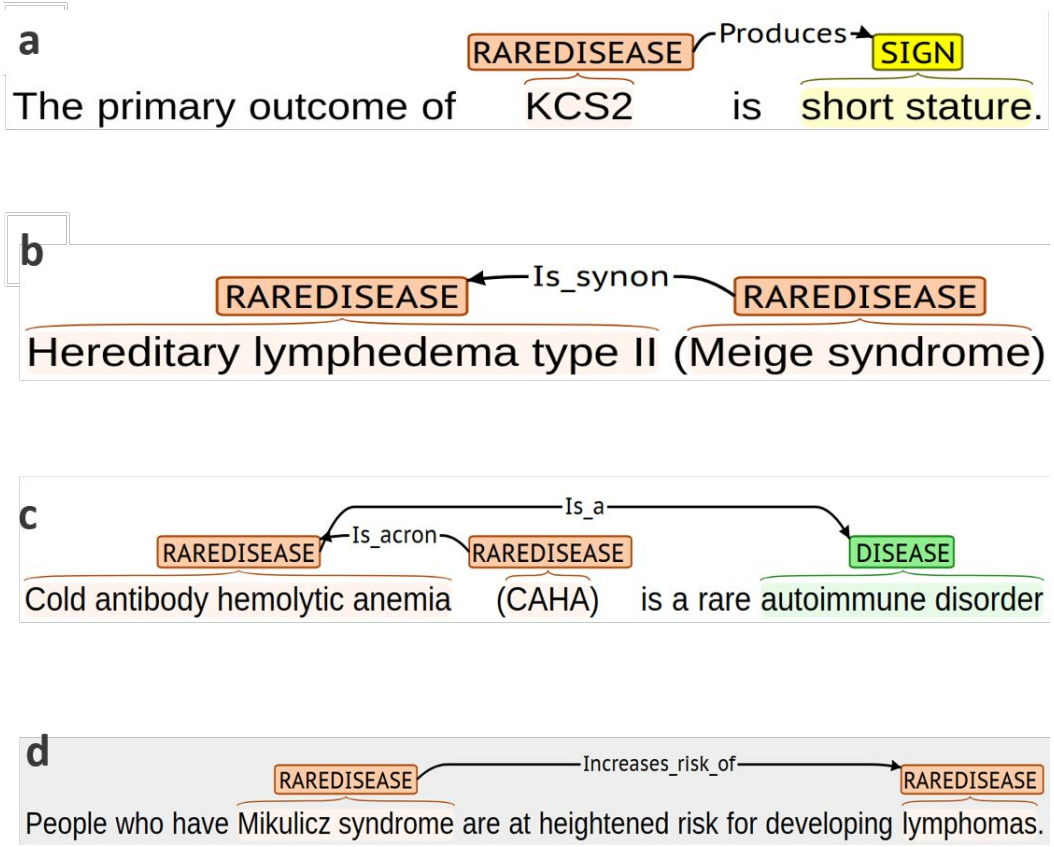
3) Co-reference resolution



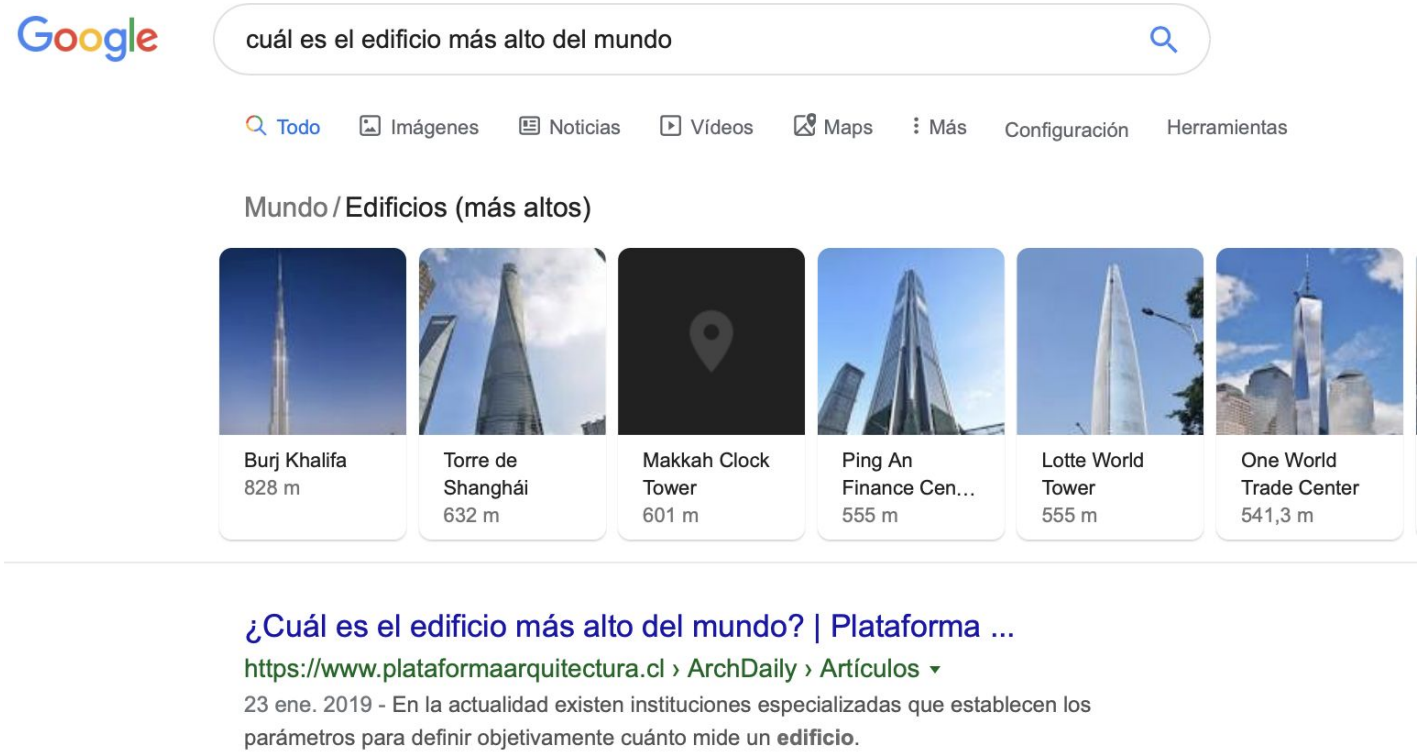


https://gate.ac.uk/family/training.html

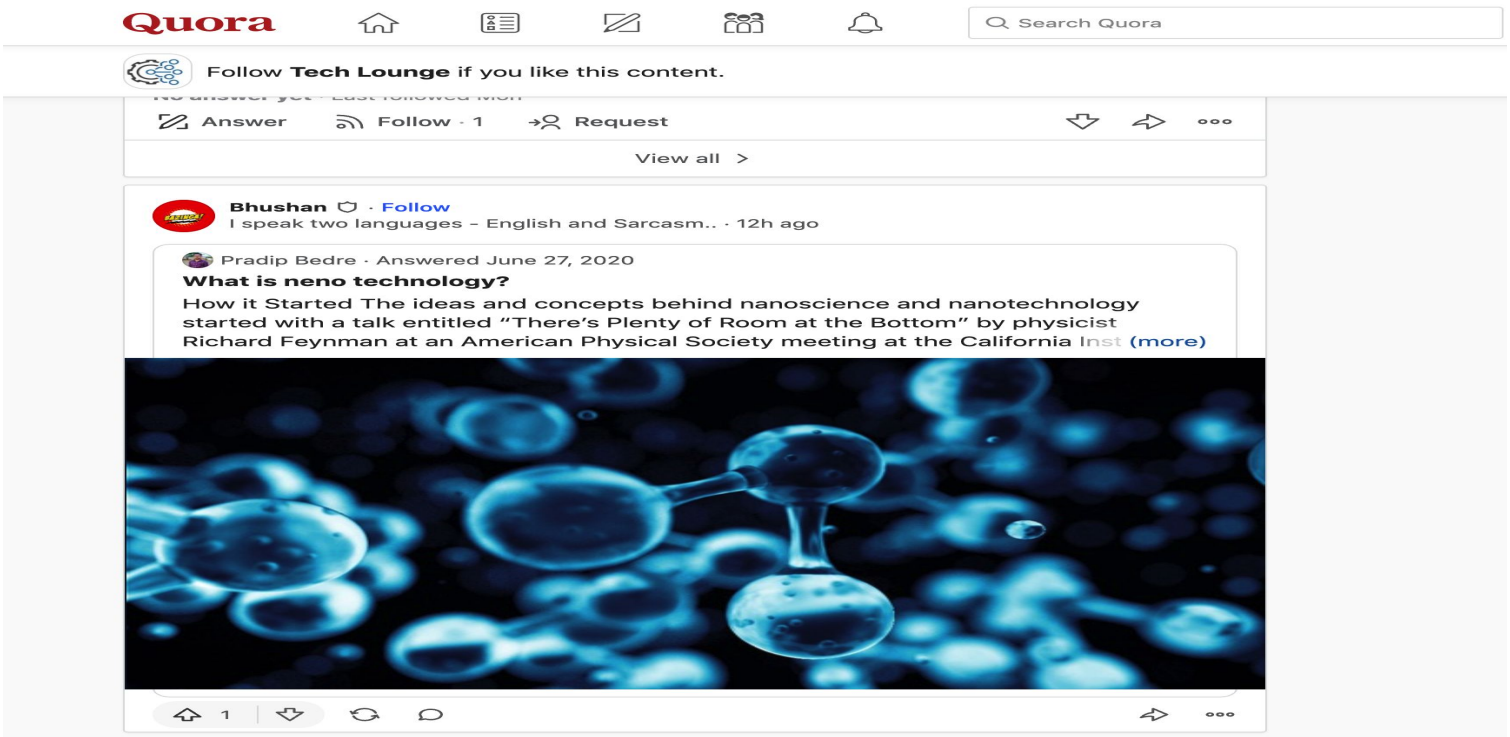
IE tasks (biomedical and clinical domains)



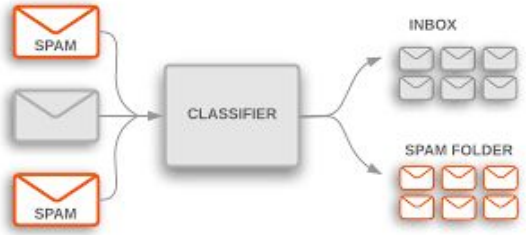
Question Answering

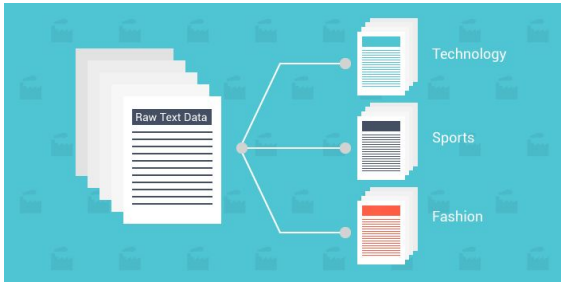


Question Answering



Text Classification





Text summarization

Source Text: Peter and Elizabeth took a taxi to attend the night party in the city.

While in the party, Elizabeth collapsed and was rushed to the hospital.

Summary: Elizabeth was hospitalized after attending a party with Peter. 

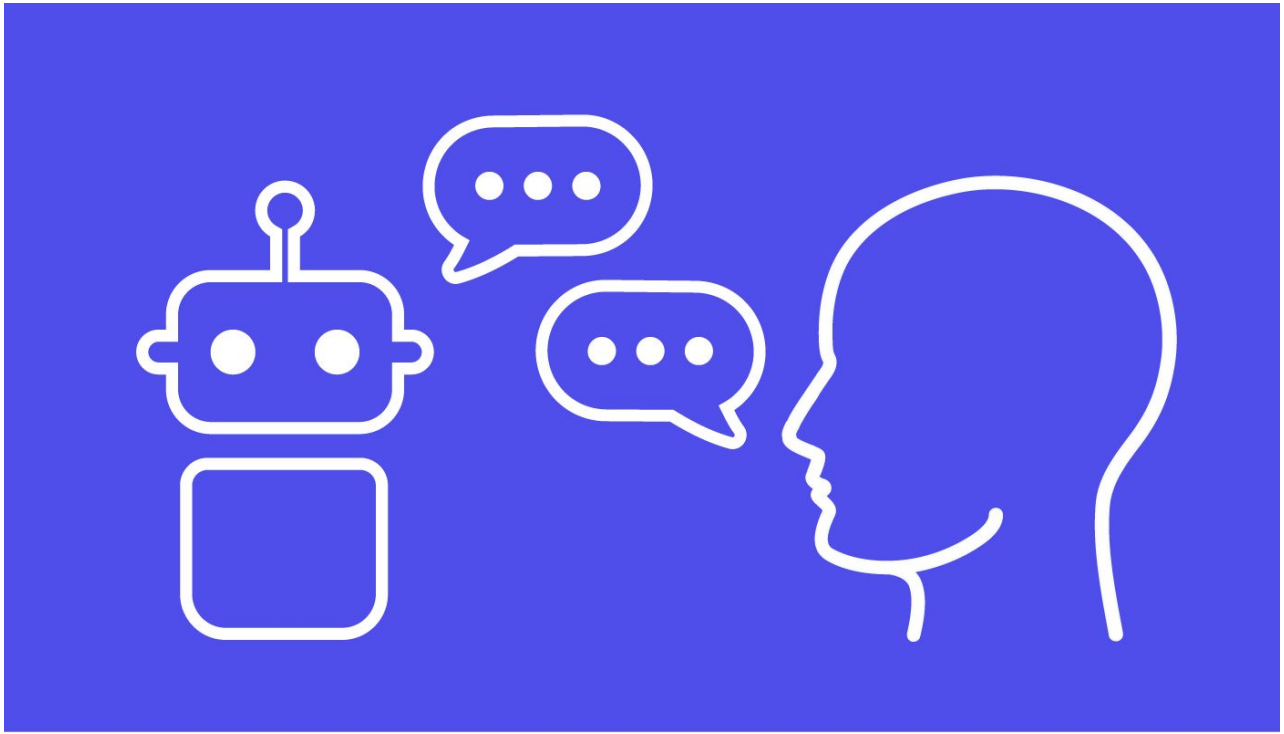
https://blog.floydhub.com/gentle-introduction-to-text-summarization-in-machine-learning/

Text simplification

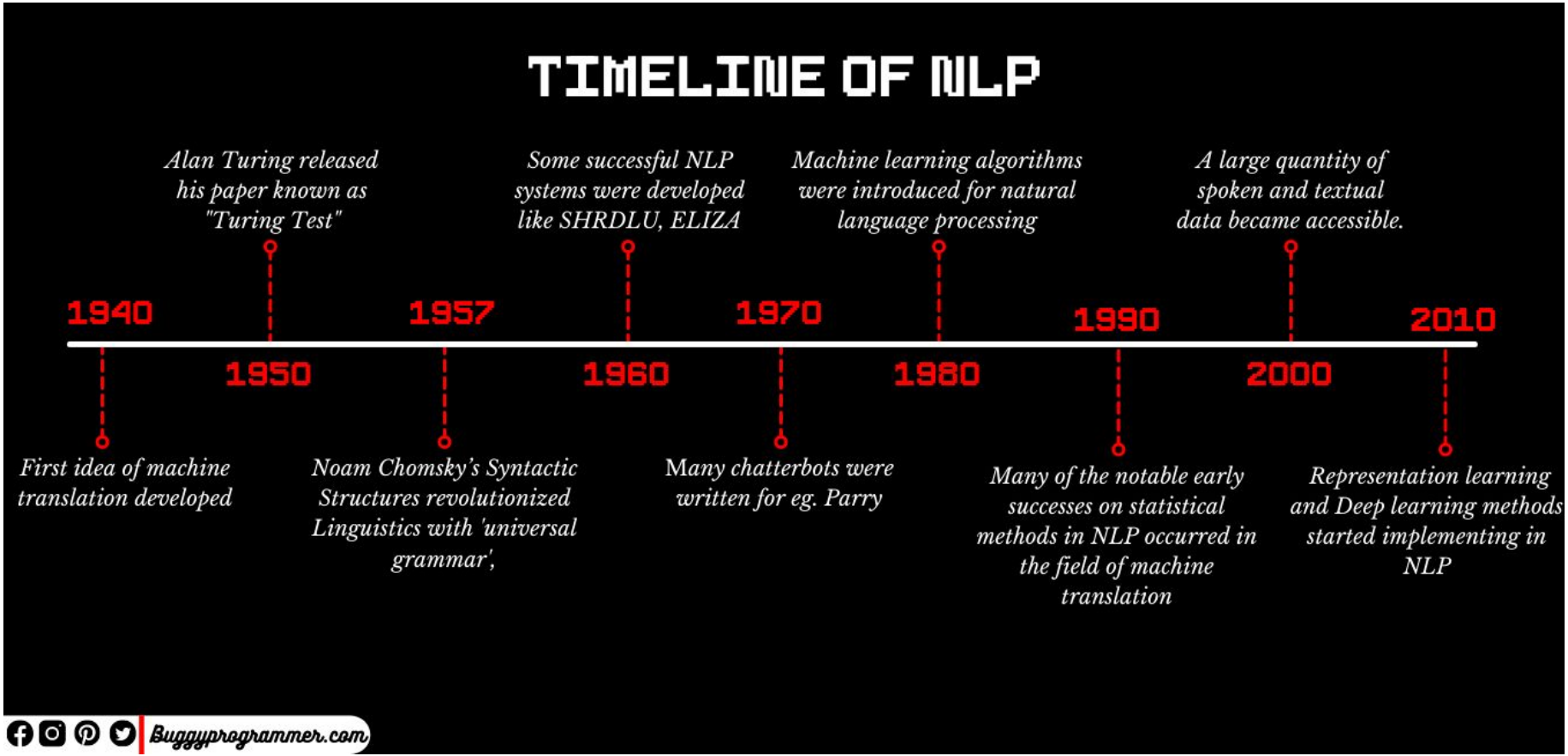
Normal:	Alfonso Perez *Munoz, usually referred to as Alfonso,* is a former Spanish *footballer, in the striker position.*
Simple:	Alfonso Perez is a former Spanish *football player.*

Coster, W., & Kauchak, D. (2011, June). Simple English Wikipedia: a new text simplification task. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: Short Papers-Volume 2* (pp. 665-669). Association for Computational Linguistics.

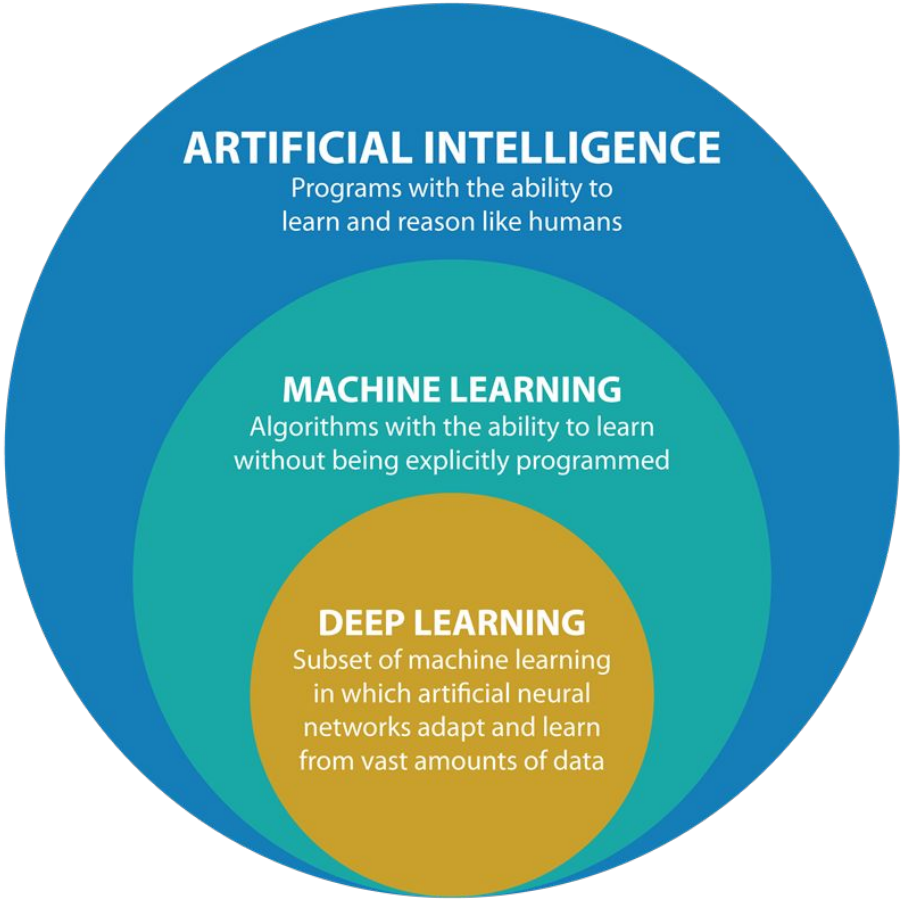
Conversational AI (Chatbots)



https://medium.com/pat-inc/conversational-a-i-needs-meaning-not-keywords-part-1-e3b0d8825564

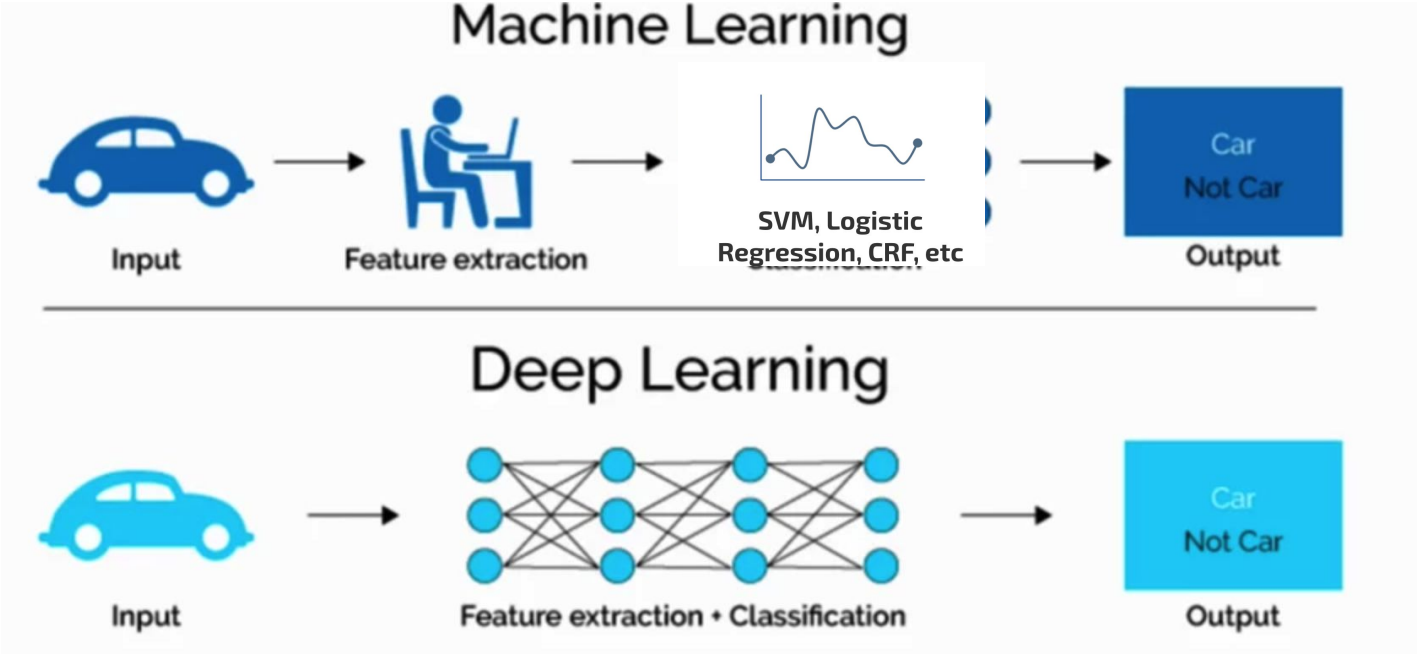


https://buggyprogrammer.com/what-is-natural-language-processing/



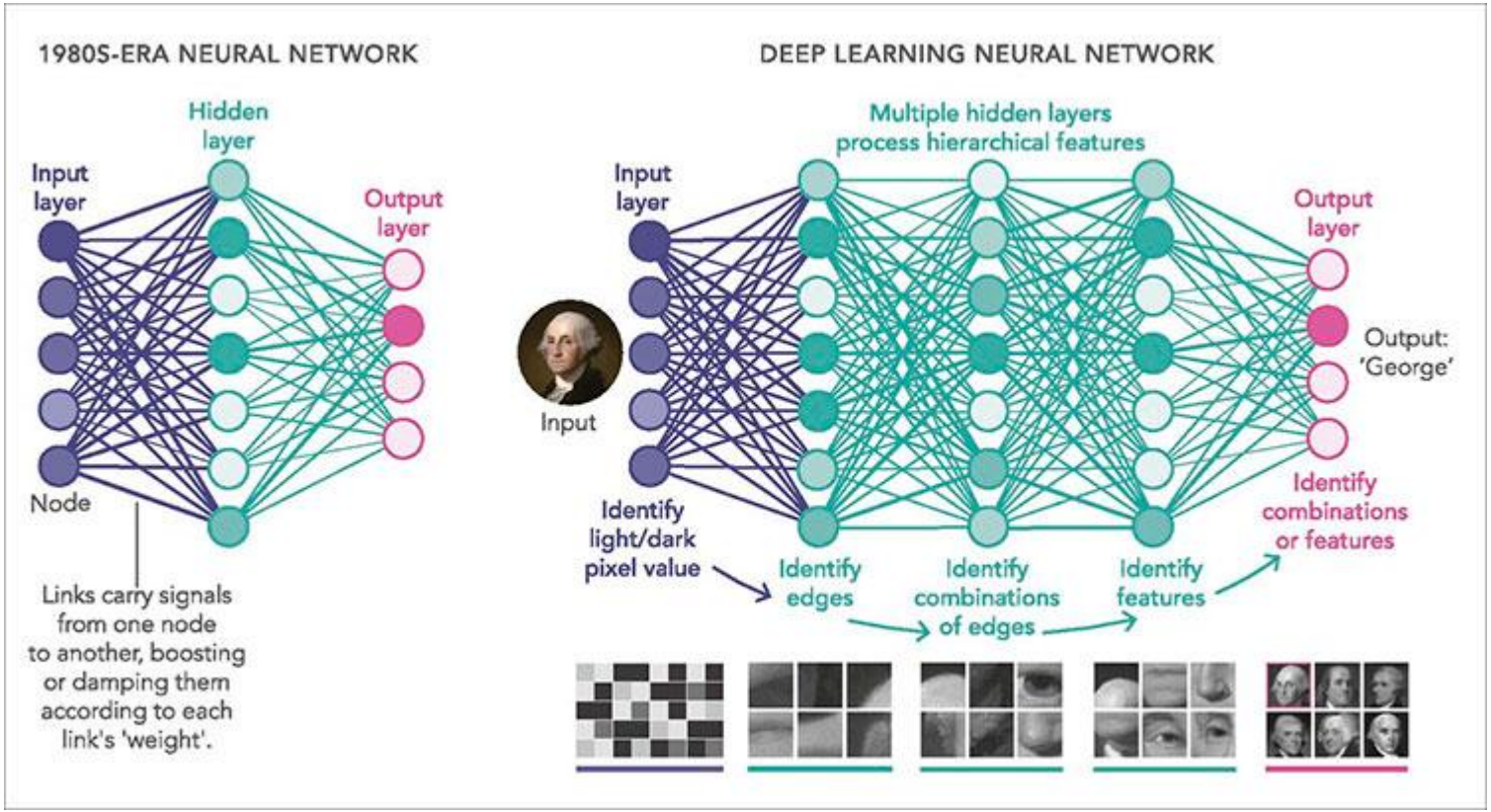
https://medium.com/@experienceIA18/diferencias-entre-la-inteligencia-artificial-y-el-machine-learning-f0448c503cd4

Traditional ML algorithms for NLP tasks heavily requiere on hand-crafted features. Feature engineering is a time-consuming process. Moreover, it is not usually robust enough (low recall)



https://blog.bismart.com/es/diferencia-machine-learning-deep-learning

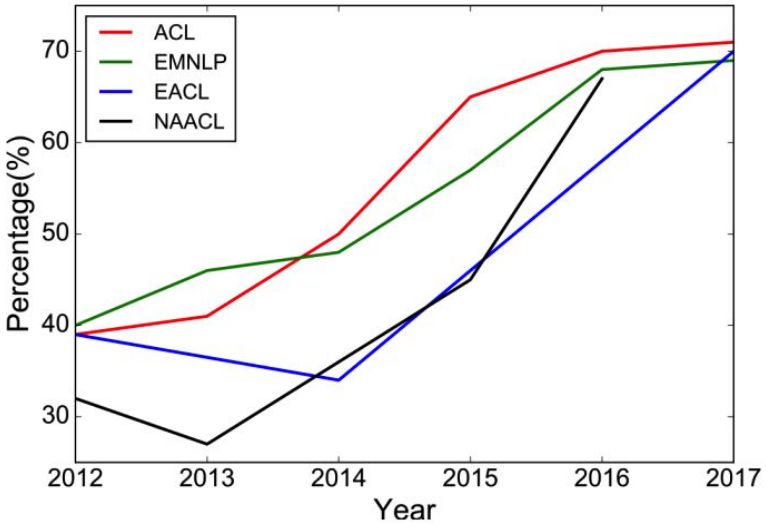
Employ multiple processing layers to learn hierarchical representations of data



https://www.futurespace.es/redes-neuronales-y-deep-learning-capitulo-1-preludio/

Produce state-of-the-art results in many domains (such as computer vision, pattern recognition)

DEEP LEARNING for NLP

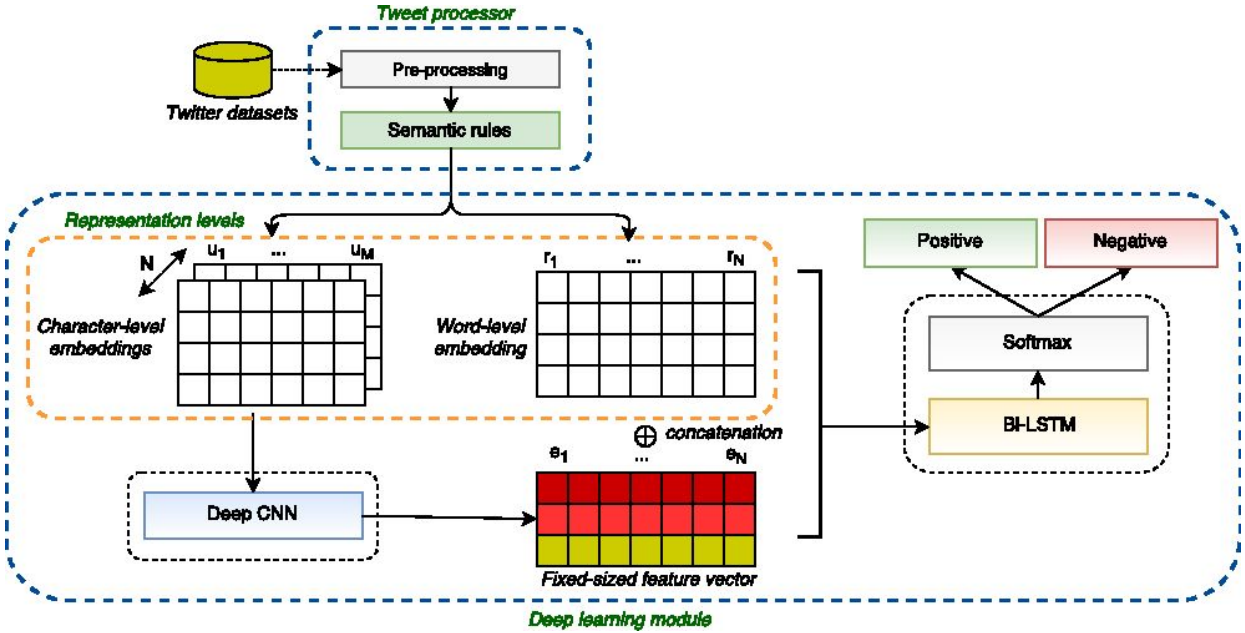


Produce state-of-the-art results in many NLP applications (machine translation, IE, text summarization, etc)

Results

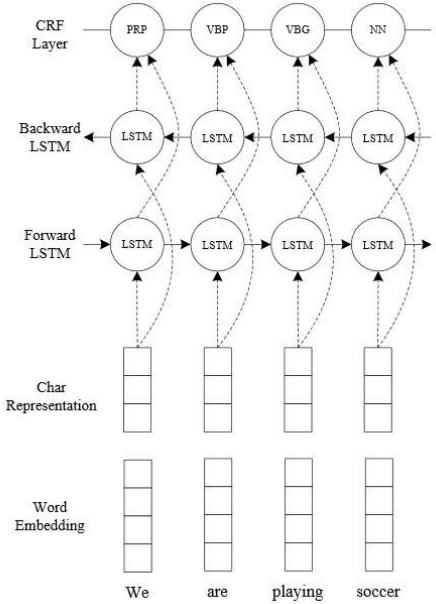
Figure from "Young, T., Hazarika, D., Poria, S., & Cambria, E. (2018). Recent trends in deep learning based natural language processing. *ieee Computational intelligen Ce magazine*, *13*(3), 55-75."

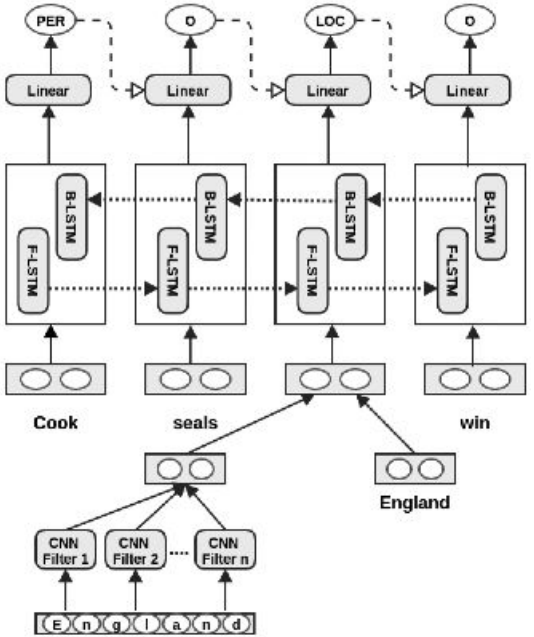
BiLSTM for Text Classification (Sentiment Analysis of Tweets)



Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment analysis based on deep learning: A comparative study. *Electronics*, *9*(3), 483.

BiLSTM + CRF for Sequence Labeling tasks (PoS tagging and NER)

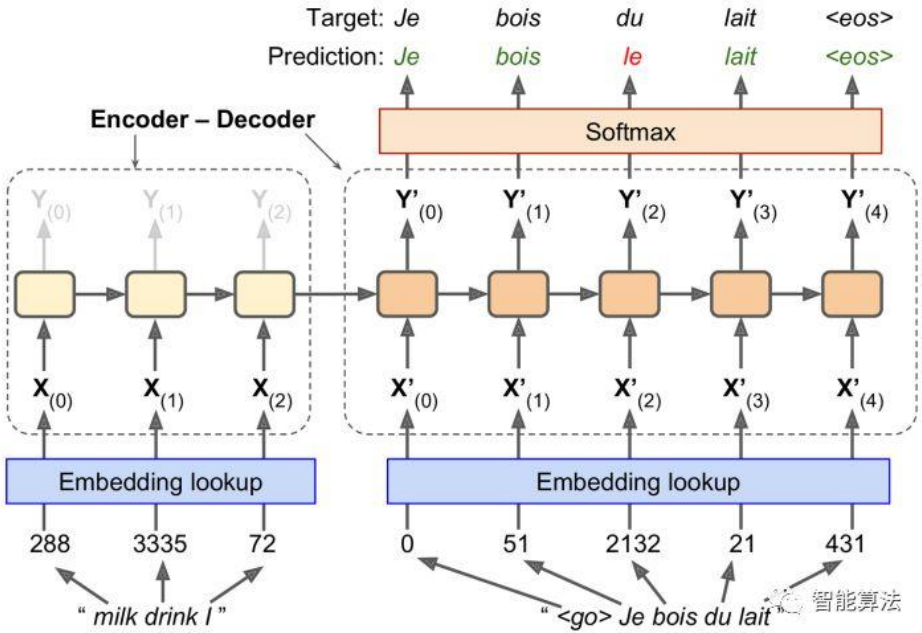




https://ai.googleblog.com/2016/09/a-neural-network-for-machine.html

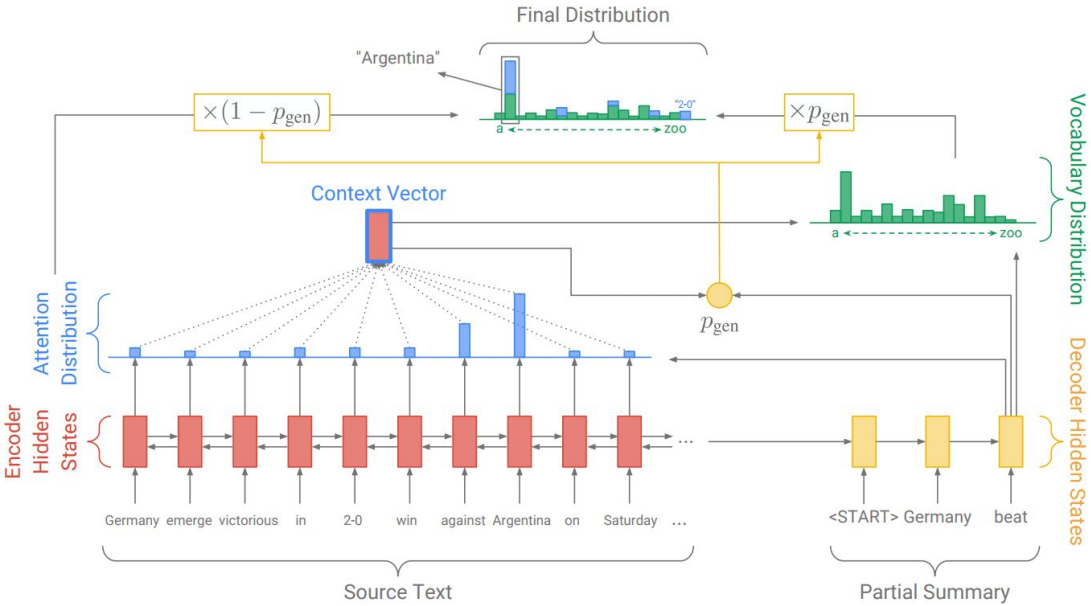
Murthy, R., Khapra, M. M., & Bhattacharyya, P. (2018). Improving NER tagging performance in low-resource languages via multilingual learning. *ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 18*(2), 1-20.

Seq2Seq for Machine translation



https://www.programmersought.com/article/83174154931/

Transformers for Text summarization



https://machinelearningmastery.com/encoder-decoder-deep-learning-models-text-summarization/

Key Take-aways

- **NLP**: automatic understanding of the human language.
- **Deep Learning** architectures are artificial neural networks with multiple layers
- **Feature engineering** is a crucial task in traditional ML.
- **Deep Learning avoids the feature engineering** task as the network learns the **best representation of the data**.
- **Deep Learning** has sparked a **revolution in NLP**, achieving **state-of-the-art results in many NLP applications**.
- The cutting edges of deep learning for NLP:
    - word embeddings -> CNN -> BiLSTM -> Transformers

Thank you
Question time!!!

isegura@inf.uc3m.es
https://hulat.inf.uc3m.es/nosotros/miembros/isegura
https://github.com/isegura