Papers and Datasets: <https://github.com/sebastianruder/NLP-progress>

[fast.ai](https://course.fast.ai/). It's the best one for developers that don't have experience with ML because it introduces the concept well. The folks at Fast AI use my favorite teaching methodology: The top-down approach. In the first lesson, all you need to do is run the notebooks they provide. In the following lessons, they dive deep into the concepts, explaining each part.

[Deep Learning AI](https://www.deeplearning.ai/), created by Andrew Ng (founder of Coursera). This is by far the most complete course that you have for deep learning, and you can enroll for free (the free-tier doesn't let you run code and do assignments). Currently, I'm doing their LLM courses and I'm loving it.

I recently wrote [an article](https://medium.com/generative-ai/explainable-ai-visualizing-attention-in-transformers-4eb931a2c0f8) on this and did a bit of research. Posting some of my favorite resources here (all free):

- [The Illustrated Transformer](http://jalammar.github.io/illustrated-transformer/)

- [The Illustrated BERT](http://jalammar.github.io/illustrated-bert/)

- [Deconstructing BERT part 1](https://www.topbots.com/deconstructing-bert-part-1/)

- [Deconstructing BERT part 2](https://www.topbots.com/deconstructing-bert-part-2/?sfw=pass1690576371)

- [A Multi-scale Visualization of Attention in the Transformer Model](https://aclanthology.org/P19-3007.pdf)

- [Stanford’s CS25: Introduction to Transformers with Andrej Karpathy](https://www.youtube.com/watch?v=XfpMkf4rD6E)

- [Stanford’s CS25: Transformers in Language with Mark Chen](https://www.youtube.com/watch?v=qGkzHFllWDY)

- [DeepLearning AI’s Natural Language Processing Specialization](https://www.coursera.org/specializations/natural-language-processing)

- [Natural Language Processing with PyTorch](https://www.oreilly.com/library/view/natural-language-processing/9781491978221/?_gl=1*14hv7ni*_ga*MTY1MzAzODY1MS4xNjg4NTAxMDYz*_ga_092EL089CH*MTY4ODUwMTA2My4xLjEuMTY4ODUwMTE3OC41OS4wLjA) by Delip Rao, Brian McMahan

If you're interested in LLMs (seems like everyone is these days), I'd also highly recommend [the Full Stack's LLM Bootcamp 2023](https://fullstackdeeplearning.com/llm-bootcamp/spring-2023/) (also free).

FAST AI: <https://course.fast.ai/>

PANDAS: <https://wesmckinney.com/book/>

<https://www.youtube.com/watch?v=rmVRLeJRkl4&list=PLoROMvodv4rMFqRtEuo6SGjY4XbRIVRd4&ab_channel=StanfordOnline>

<https://web.stanford.edu/~jurafsky/slp3/>

<https://www.youtube.com/@bycloudAI>

<https://arxiv.org/pdf/2201.11903>

Encoder-based: BERT (https://arxiv.org/abs/1810.04805), XLMR (https://arxiv.org/pdf/1911.02116)

Econder-Decoder: T5 (https://arxiv.org/pdf/1910.10683)

Decoder models: GPTurile (un pic private de la chatgpt), PaLM/Bard/Gemini (google), Llama (facebook, opensource), Mistral 7B (<https://arxiv.org/pdf/2310.06825>)

**Deep Learning for Natural Language Processing: A Gentle Introduction**

This repository contains the code and data introduced in the book *Deep Learning for Natural Language Processing: A Gentle Introduction* by Mihai Surdeanu and Marco A. Valenzuela-Escárcega. Please see <http://clulab.github.io/gentlenlp/> for more information.