

Pretrained Language Models for Text Generation

Unai Sainz de la Maza, Unai Salas and Adrián San Segundo

Advanced Techniques in Artificial Intelligence

What is text generation?

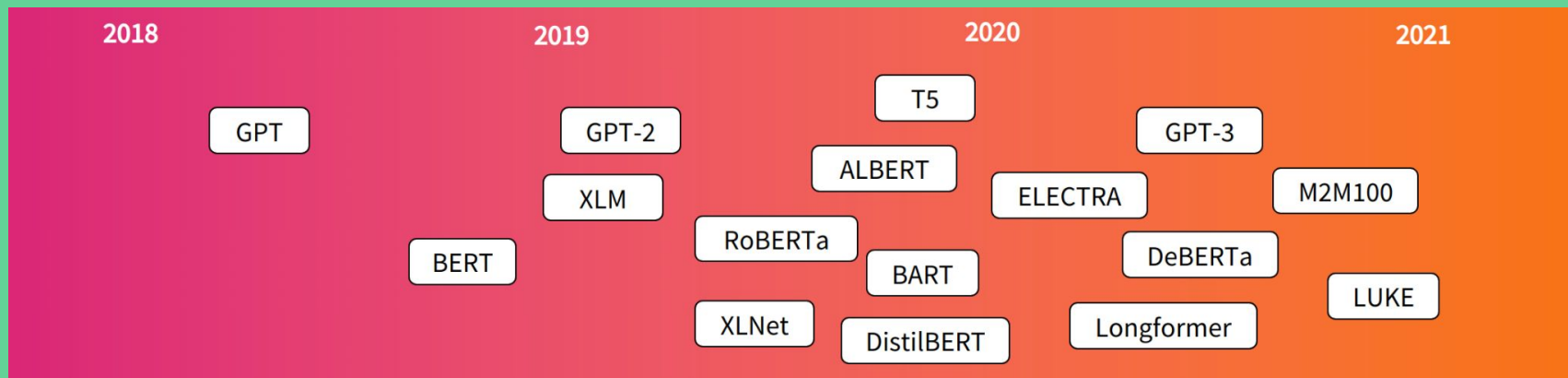
- Is one of the most important tasks in natural language processing (NLP).
- The main goal is to produce plausible and readable text in human language from input data like a sequence and keywords.

Where is this used?

- Machine translation.
- Dialogue systems.
- Text summarization.
- Text paraphrasing.
- ...

How are these tasks performed?

Transformers

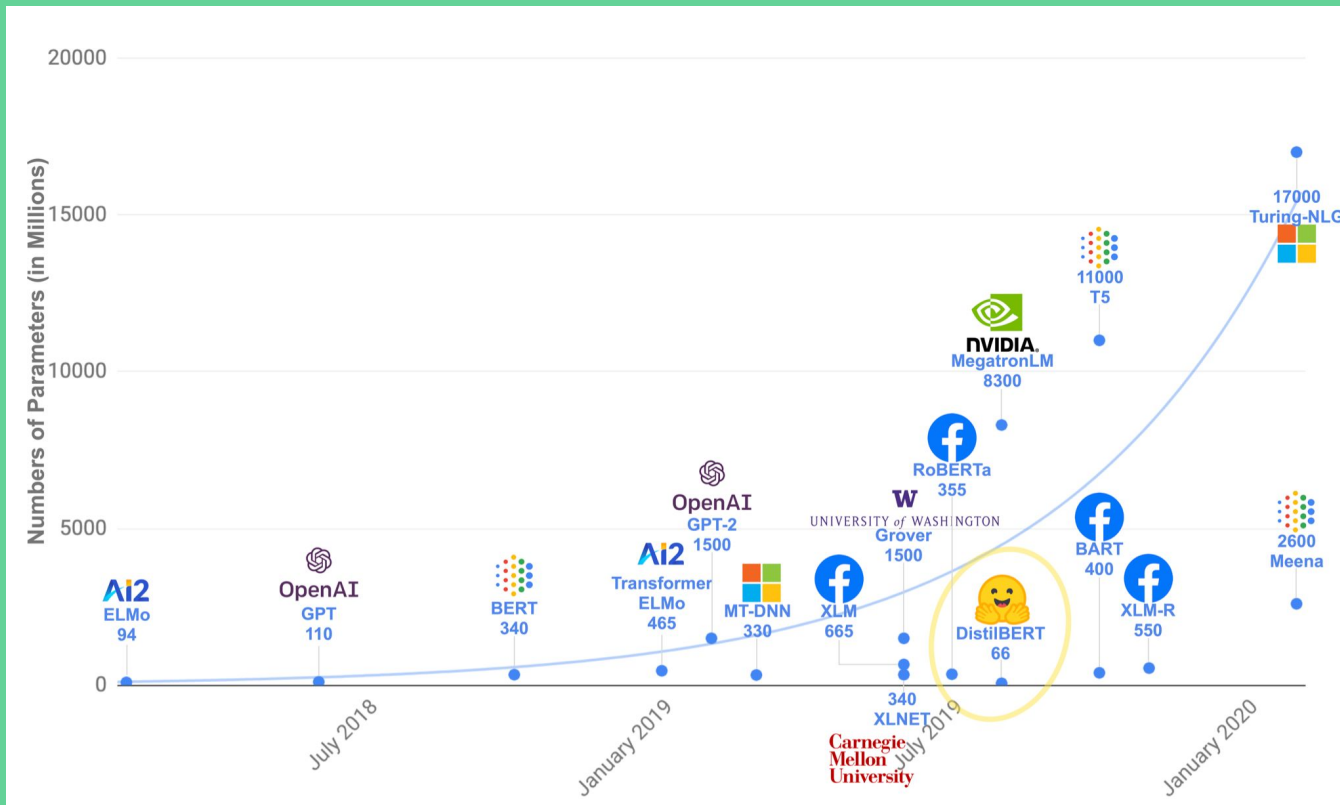


Source: Huggingface <https://huggingface.co/course/chapter1/4>

Transformers are language models

- All the transformer models mentioned are trained as a *language models*.
- They have been trained on **large amounts of raw text** in self-supervised fashion (humans are not needed to label the data).
- The model develops statistical understanding of the language it has been trained on.
- This is not very useful for practical tasks.

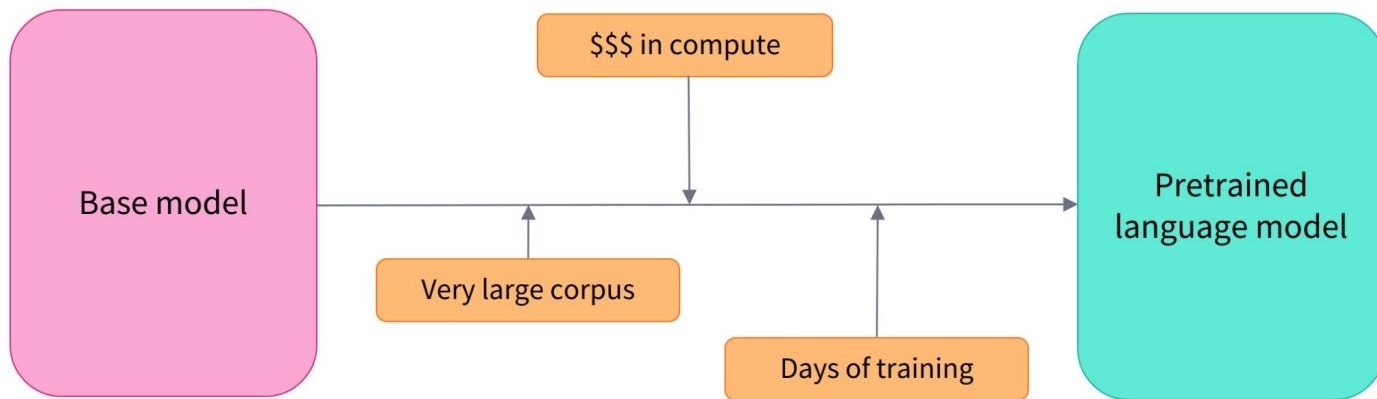
Transformers are **BIG** models



Source: Huggingface <https://huggingface.co/course/chapter1/4>

Transfer learning

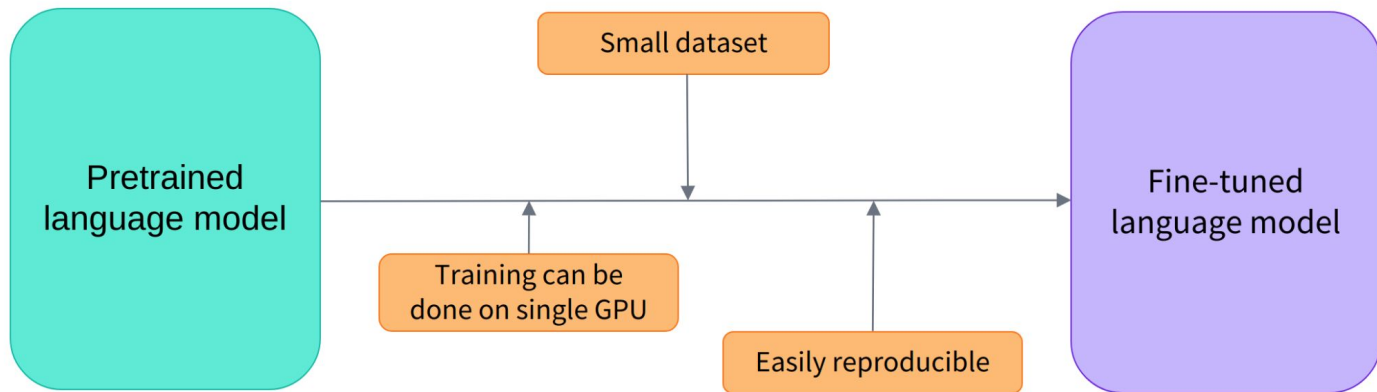
- **Pretraining** is the act of training the models from scratch.



Source: Huggingface <https://huggingface.co/course/chapter1/4>

Transfer learning

- **Fine-tuning** is the training done after a model has been pretrained.
 - Is to perform an additional training with dataset specific to your task.
 - Requires way less data to get decent results.
- The knowledge that pretrained model has acquired is “transferred”.

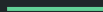


Source: Huggingface <https://huggingface.co/course/chapter1/4>

Future Work

Several directions of work are being studied

- Upper bound of PLMs
- Interpretability of PLMs
- Language-agnostic PLMs



GPT-Neo Demo

demo for GPT-Neo by EleutherAI for text generation. To use it, simply add your text, or click one of the examples to load them. Read more at the links below.

INPUT TEXT

Text generation is one of the most important tasks in natural language processing

OUTPUT

. Researchers argue that the results of a machine learning model should be of human quality. Therefore, natural language models should be capable of producing linguistically well-formed sentences. This requires a

Latency: 2.83s

CLEAR

SUBMIT

SCREENSHOT

GIF

FLAG

Examples

Run All

Load Previous
CTRL + ←

Load Next
CTRL + →



The tower is 324 met...

The Moon's orbit aro...

The smooth Borealis ...

GPT-Neo: Large Scale Autoregressive Language Modeling with Mesh-Tensorflow

Conclusions

- Reuse of the models
- Saving costs



Discussions
