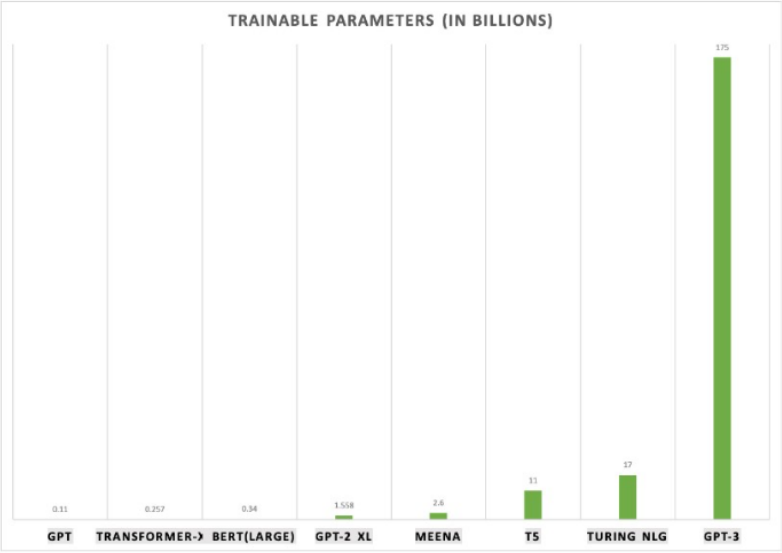
GPT-3 Transformer

1) Introduction

In Machine Learning field one area is grabbing a lot of attention nowadays and that is automated language generation Which has a large set application. It is created by OpenAI GPT, which stands for Generator pre-Trained Transformer. GPT-3 is a language prediction model. GPT-3 is world's largest Natural language Generation Transformer which is will be.it trained by 175 billion parameters which makes GPT-3 a highly intelligent language Transformer. Also GPT-3 is world's largest and most used Language Generator tool.

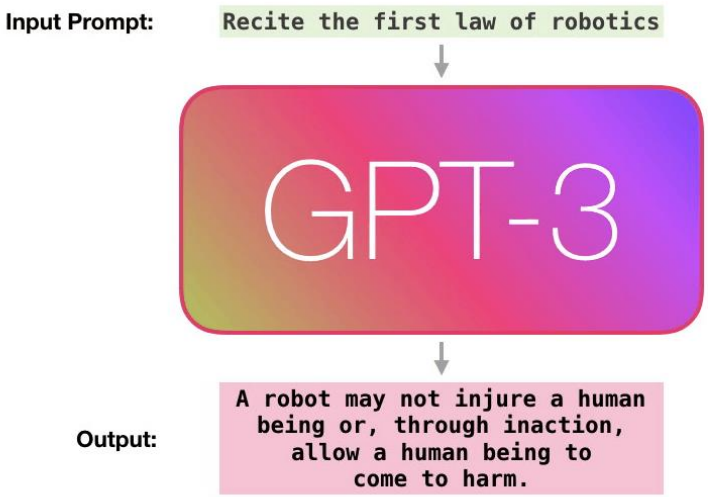
2) What is GPT-3?



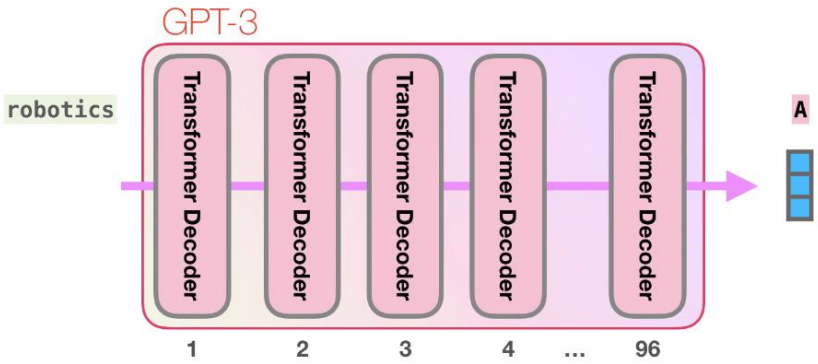
GPT-3 is a Natural Language Generator, and we will discuss some of its characteristics below:

a) GPT-3 is used to create texts which resemble humans in many ways as it is created on top of a broad neural network created by humans.
b) GPT-3 is trained in such a manner that it should easily learn the rules and concepts that arise in human texts.
c) GPT-3 is extremely useful if you are considering that you have a small dataset and want to make the dataset bigger
d) GPT-3 is also capable of learning to mimic some artist like Shakespeare and it can be able to create poetry like Shakespeare
e) It is an autoregressive language model that was trained with 175 billion parameters to produce human-like text, rendering it the largest language model by a large distance.

3) Architecture and Working



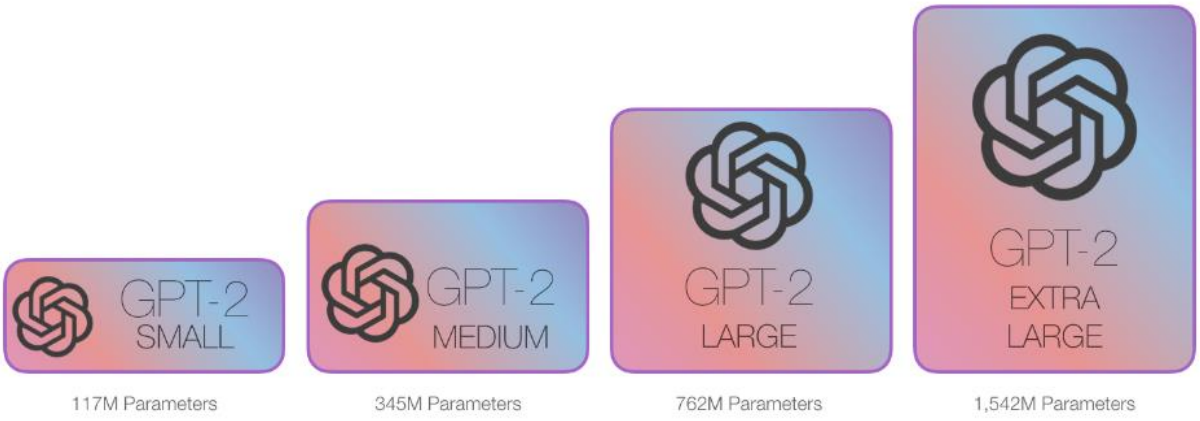
A) GPT-3 is a language prediction model. This means that it has a neural network machine learning model that can take input text as an input and transform it into what it predicts the most useful result will be.



4

B) So as we say that it is a transformer-based model so we know that we are going to deal with the encoding and the decoding

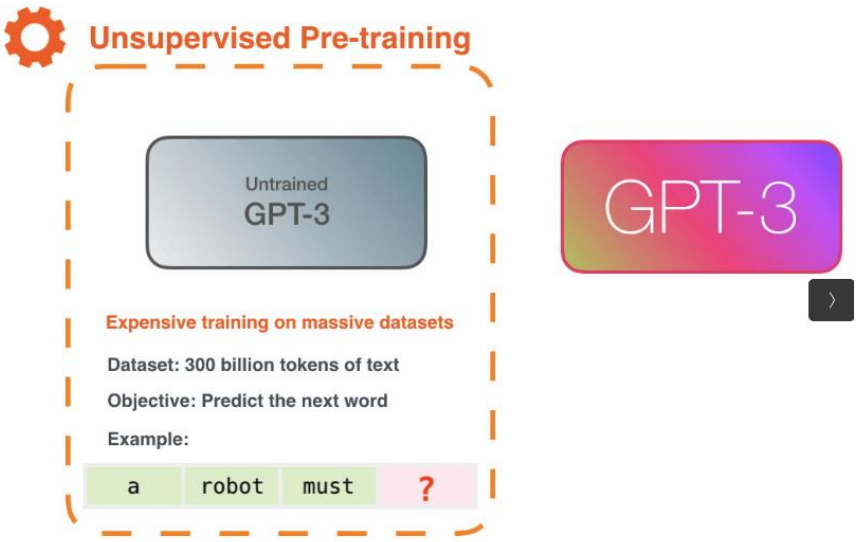
C) In transformer you can use only one mechanism such as the stacking the decoder one upon another in such way shown in diagram so as the



D) As the GPT-3 Is successor of GPT-2 and GPT-1 so we see the that count of decoders are going to increase.

#Internal mechanism:

A) Training is the process of exposing the model to lots of text. That process has been completed. All the experiments you see now are from that one trained model. It was estimated to cost 355 GPU years and cost $4.6m.



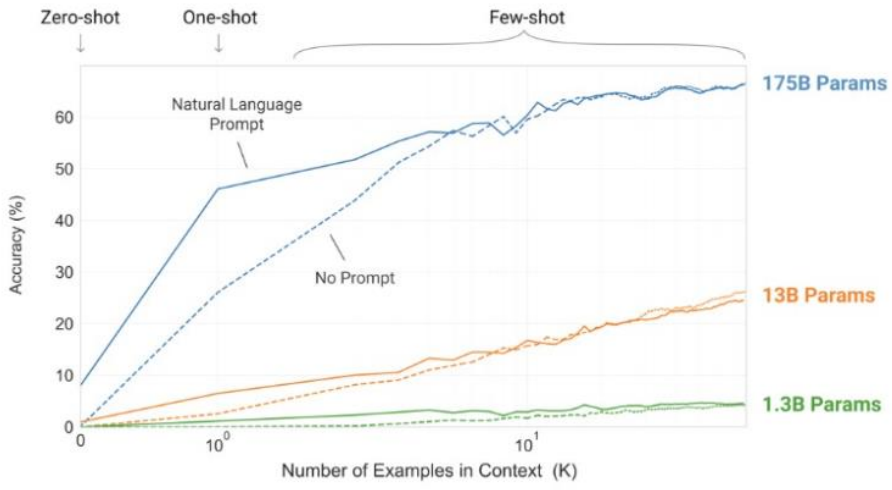
B) The model's prediction will be wrong. We calculate the error in its prediction and update the model so next time it makes a better prediction. Repeat millions of times in order to find out optimal results.

4) Performance

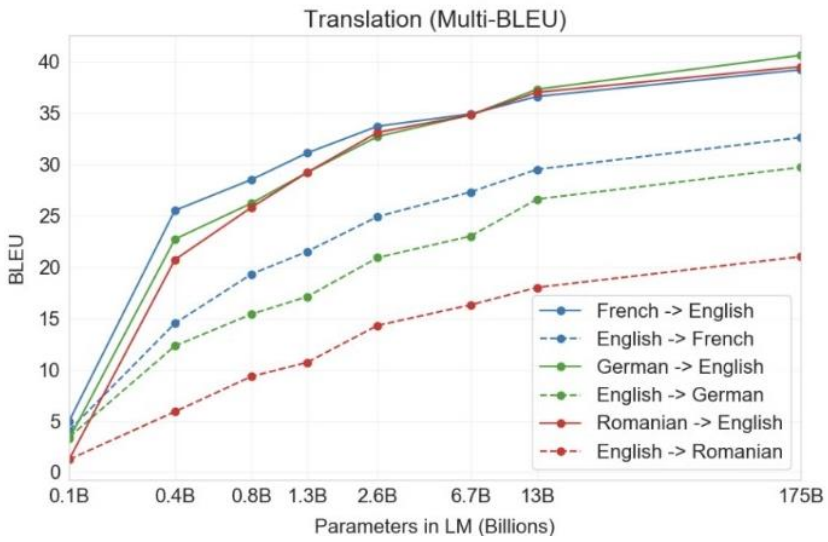
Before learning through the performance, we must discuss some terms on which are going to acts like parameters in case of performance:

a) The Few-shot (FS) setting is kind of like how we go about training a machine learning model where we give some inputs and corresponding outputs to a model and then expect the model to perform on an unseen input. (not that hard for GPT to guess)
b) One-Shot (1S) setting is Where we provide exactly one input to system and accordance to that system or ML model will try to predict the result (hard to guess)
c) Zero-Shot (2S) is a concept where nothing as input will be provided but still, we want to get our output (very hard to guess)

- **Performance in context of no of examples provided as input**



- **Performance in the case of Translation**



5) Applications

- GPT-3 is the biggest neural model-based AI tool for Natural language generation, and it is intelligent enough to do tasks text summarizations and even programming code.
- GPT-3 is also good in communication with humans as it can create text very close to that of the humans so they can be used to create the chatbots
- GPT-3 can be used in gaming also in order make game dialogue look more alike that of the men and also to write the code like human and do the task like suggestion which is found very useful in coding and such stuff
- GPT-3 is very useful if you are considering that you have a small dataset and want to make the dataset bigger
- It is also helpful for applications like create articles, poetry, stories, news reports and dialogue

6) Limitations

The GPT-3 is perfect, but it also came with some limitations so let us discuss them:

a) Lack Of Semantic Understanding
b) Black Box Mechanics
c) Poor Generalization beyond the training set.
d) Bias In Generated Text
e) Not a Complete AIG (artificial intelligence generator)

7) References

- https://jalammar.github.io/how-gpt3-works-visualizations-animations/
- https://www.techtarget.com/searchenterpriseai/definition/GPT-3
- https://towardsdatascience.com/gpt-3-explained-19e5f2bd3288