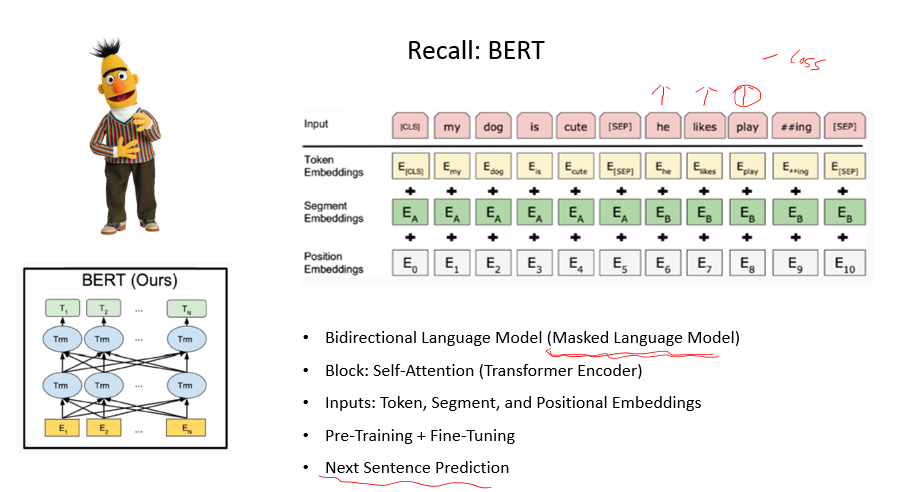
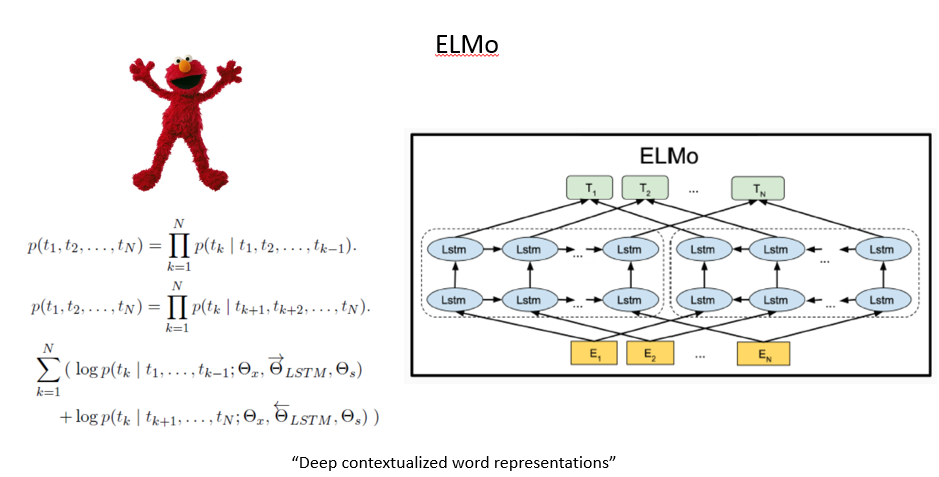


# Bert



# EMLo

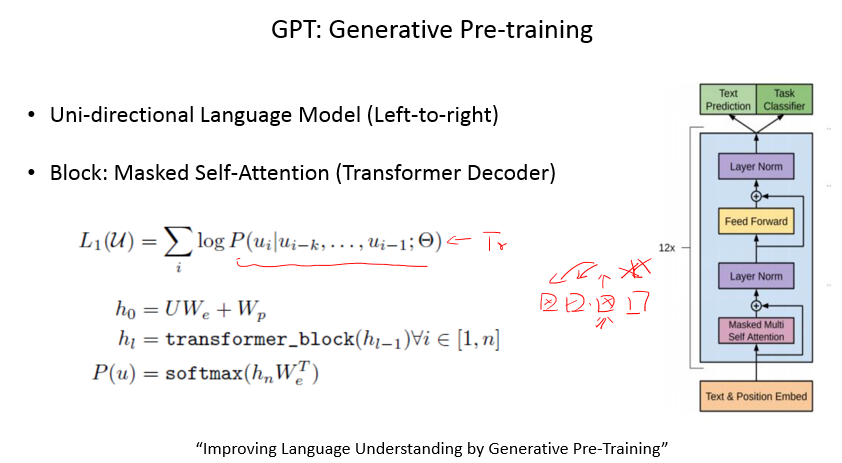
预训练语言模型，将语言模型的损失+任务的损失 （同时优化）

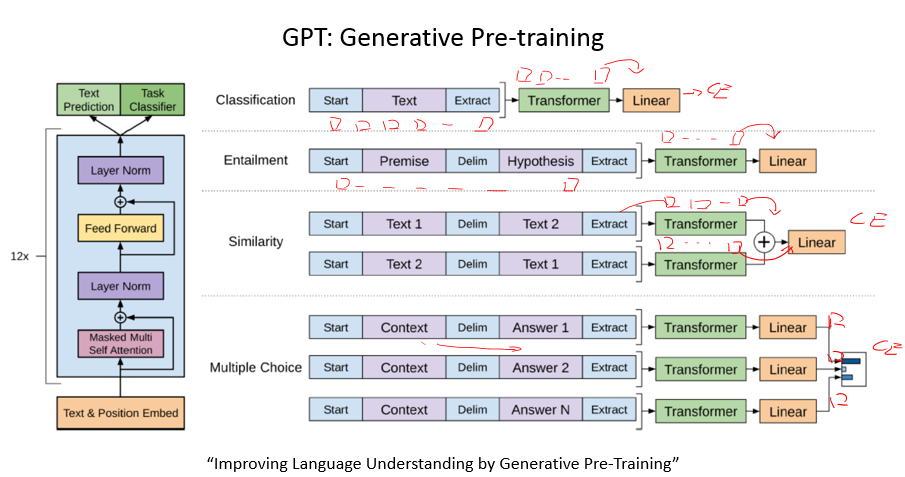


# GPT：Generative Pre-training 文本生成

**通过实践证明，GPT更加擅长于生成**

* “Improving Language Understanding By Generative Pre-training”
* “Language Models Are Unsupervised Multitask Learners”
* “Language Models Are Few-shot Learners”





Problems of GPT

预训练Fine-tune(微调)的模型无法转移到其任务，那大佬提出来了，能不能用一个模型去让她去做所有nlu的任务呢?

The prevalence of single task training on single dataset limits the generalizability. (Maybe Multi-task Learning helps?)

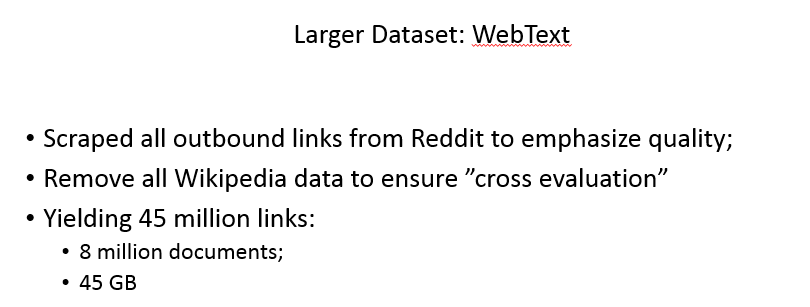
Require supervised training in order to perform a task.

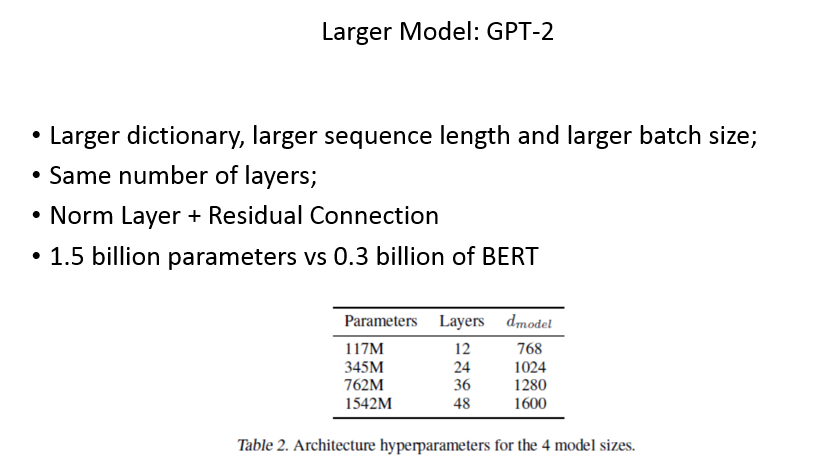
Intuition: Language Model == Unsupervised Multi-task Learning

If we train language model with “the translation of the word machine learning in Chinese is 机器学习”, then we have a machine translation system which can translate “machine learning” to Chinese.

We need larger model and larger dataset!

GPT2





Fasttext：单层CNN，快，但性能一般+

为证明他们把全部数据作为候选集合不会造成过拟合，做了基于Bloom Filters的字符串匹配

