

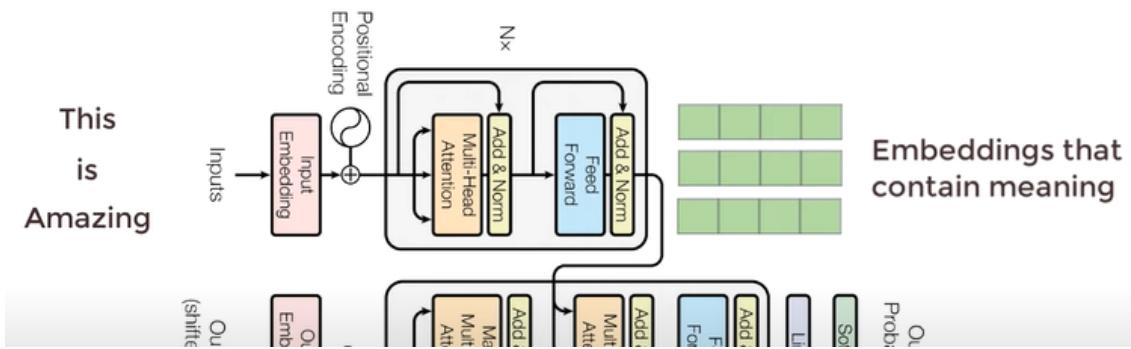


BERT

Owner	Ydrgzm
Tags	Notes
Created time	@September 5, 2023 3:32 PM

Notes

- LSTM were Slow, Not truly Bidirectional
- Transformer solved both problems
- For NLP, Translation to French
- Transformer Key components:
 - Encoder
 - Takes: **English words** simultaneously
 - Gives: **Embeddings (vectors)** for every word simultaneously
 - Vector of numbers for each word that encapsulates meaning of word
 - Similar words have closer numbers



- **Decoder**

- Takes: Embeddings from Encoder
- Takes: Previously generated words (of French translation)
- Gives: Uses above two things to generate next word
- It repeatedly generate words until the end of sentence.
- We can see the distribution of tasks between **Encoder and Decoder**.

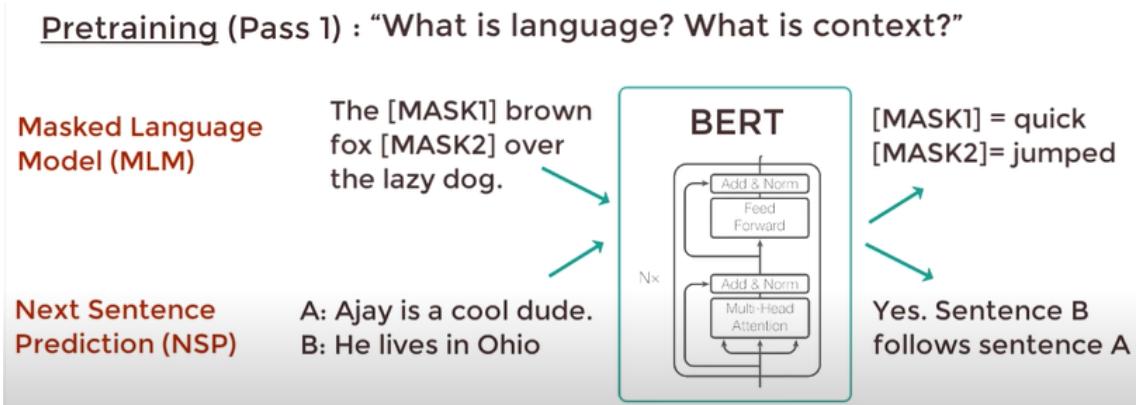
- **Encoder**

- What is English?
- What is Context?

- **Decoder**

- How to map English words with French words?
- but both have one common thing **What is Language?** Because both try to understand the language of input sentence
- Now, if we stack multiple
 - Decoders we get **GPT** architecture
 - Encoders we get **BERT** architecture
- **BERT** (Bidirectional Encoder Representation from Transformers)
- We can use Transformers for:
 - Neural Machine Translation
- **Uses of BERT**

- Neural Machine Translation
- Question Answering
- Sentiment Analysis
- Text Summarization
- Above uses of BERT require language understanding so we can:
 - Pretrain BERT to understand language
 - Fine tune BERT to learn specific task
 - BERT Training in above two phases: Pretrain (learning language and context) and Fine tune (I know language so how do I some task)
- **Pretraining (Pass 1)**
 - Learn: What is language? What is Context?
 - Learns language by two unsupervised tasks simultaneously:
 - Masked Language Model **MLM**
 - Next Sentence Prediction **NSP**



- **Fine Tuning (Pass 1)**
 - Further train BERT for specific NLP task
 - How to use language for specific task?

Bidirectional Encoder Representation from Transformers

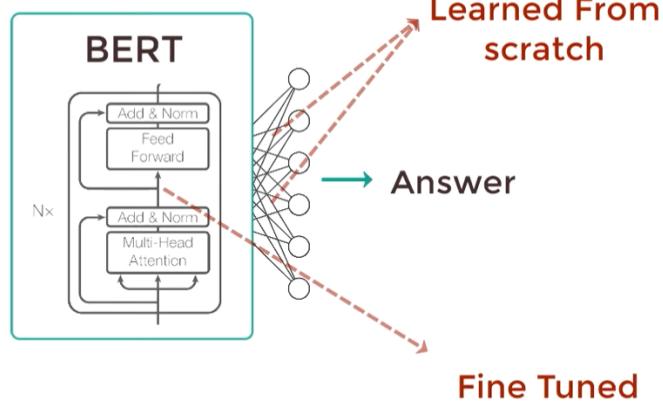
Fine Tuning (Pass 1): "How to use language for specific task?"

Fine tuned Q & A

Question

Passage

FAST!



- **Pretraining** (Pass 2)

◦