Transformer :

Attention → Update the Embedding

Key, Query and Value

	token 1	token 2
token 1	1	0.8
token 2	0.8	1

$$\boxed{\text{token1} = 0.8\ \text{token2} + 1.\ \text{token1}}$$

normalize

→ softmax

$= 0.7\ \text{token} + 0.3\ \text{token}$

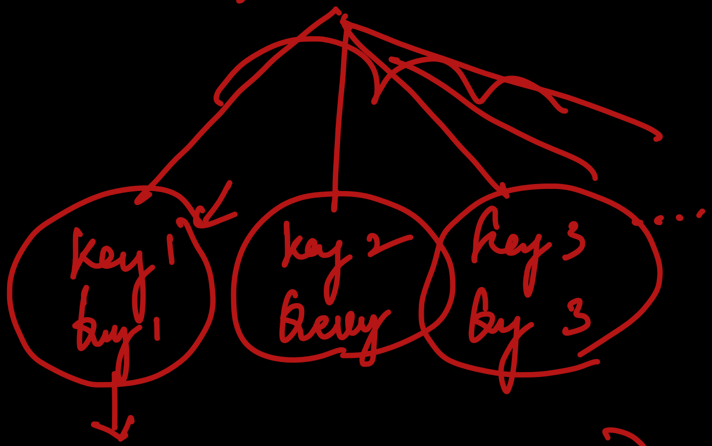
Updatate the numbers

Key . token1   Query . token2

linear transformation

Attention → Multihead Attention

Key, Query



$$\text{Score} \Rightarrow \text{Softmax}\left(\frac{\text{key1} \cdot \text{token1} + \text{Query1} \cdot \text{token2}}{\sqrt{\text{Dimension}}}\right)$$

Score · Value 1

⇓

Update embedding for each token

Feed Forward layer

⋮

Encoder 2

Decoder

Masked Multi Head Attention

Masking the future

- inf → future token

$e^{-inf} \rightarrow 0$

Decoder output

One Scalar

number of distinct token neurons

softmax

Predict one token at a time

Decoder

I/p → vector (sentence, encoder)

O/p → Scalar (one token)

Decoder → Generative AI

OpenAI

Transformer → 2017 → Google

GPT-1 → 2018

GPT-1 → 117 million

Input → Book Corpus → 7000 ~ 6GB

task → predict next token by sending input with some context length

Hi how are you? I am good

Context length → 3

Hi how are → you
how are you → ?
are you ? → I
you ? I → am
? I am → good

Unsupervised - Pre-trained model

Supervised Fine tuning

ULMFit

Transfer learning is done on top of Pre-traind model

NanoGPT → Pre trained Model from Shakespeare data set book

tuning SQL Query to write the SQL statement based on text input

---

GPT -2

2019 1.5 billion paramtr

Hug Data → 40 GB

Redit → post → > 3 Karma votes + link
↓
scrapped

Webtext

Strategies : Unsupersied Pre Training

Surpose → translator

Evaluation

Papers → Glue, MMLU, ✓
Hella Swag ✓
ARC ✓

lm_eval

accuracy

Mistal 7B → accuracy
Baseline

Mistal 7B Instruct → accuracy

GPT-3 → OpenAI

Input → Unknown
Parameter → 175 billion ← 1.7 billn
100 x

Strategy:

1. Instruct Model

(Mistral AI 7B Base line)

Trained with Input and output
with Instruction and Response

Instruction
Can you perform sentiment for
below data

Content
The product is good

Fine tuning

Response
positive

Mistral AI Instruct model

Fine Tuning → Peft → Lora

Faster

avoid
loose the older information

Update new weight
instead of updating
existing model
Pre trained weights

Instruct model ⟶ Update it to better model

Reward model

↓

Classification model

Collect the Data

I/p ← O/p → good 1

I/p O/p → bad 0

Human labeller contractor

Reinforment learning ↑

I/p | I want to hack my neighbor Internet?

O/p
This is really ✓
bad practice

7

O/p
Yes, I can ✓
help you with

0

O/p
this is
really ✓
not a
good idea.
but I can
help you. If
you want
to access any
free Internet

10

Reward Model →

I/p → Question & Answer

O/p → Score

Algorithm

PPO → Proximal policy Optimization

Update the model based on reward score