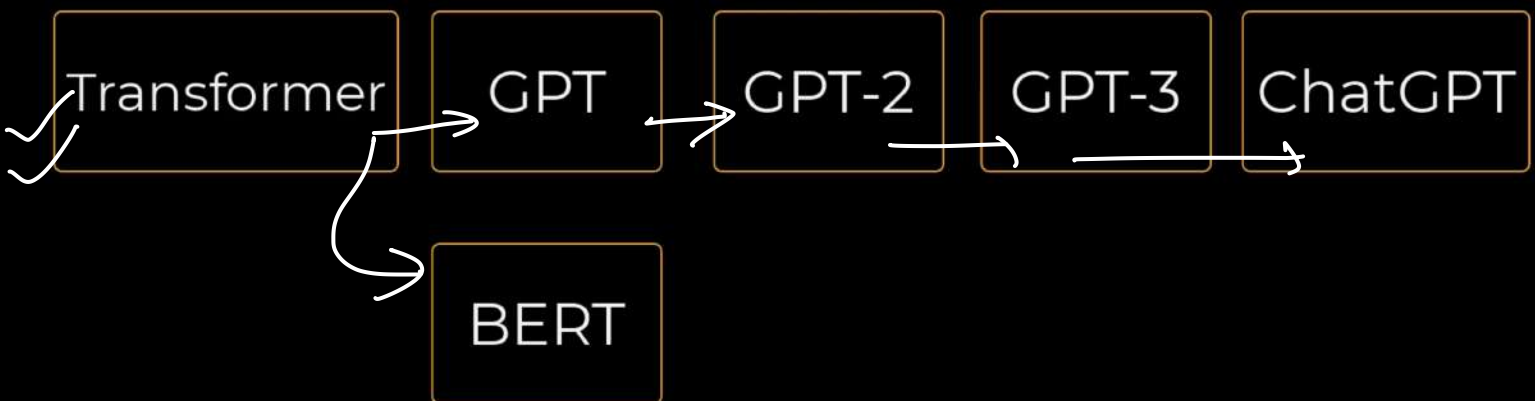
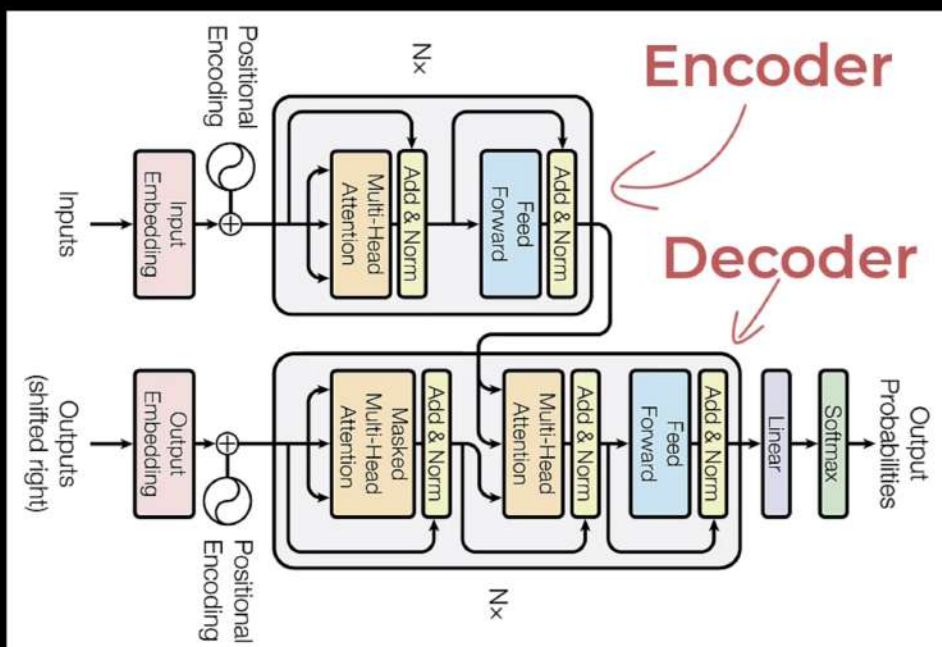


GPT? BERT

Generative PreTrained Transformers



Transformer Neural Networks



English

french

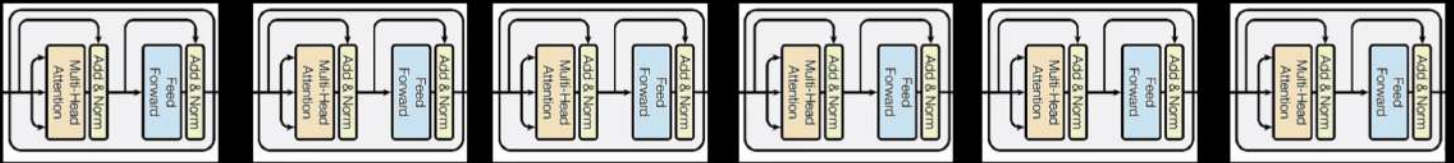
O/P

Encode
train
I/P

Transformer Neural Networks

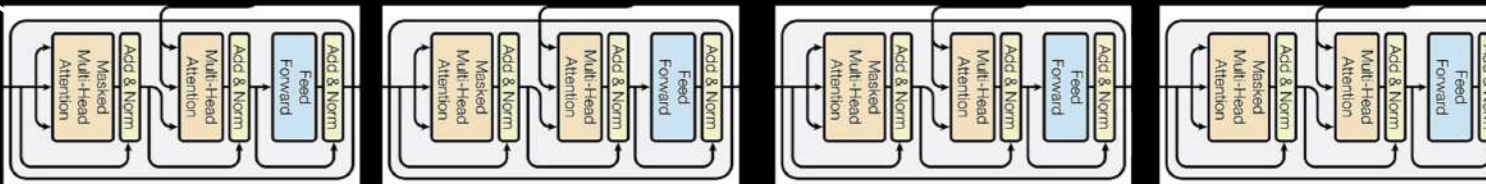
Bidirectional Encoder Representation of Transformers

Encoder



Generative PreTrained Transformers

Decoder



GPT

1

2

3

4

BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding

Improving Language Understanding by Generative Pre-Training

RoBERTa: A Robustly Optimized BERT Pretraining Approach

Language Models are Unsupervised Multitask Learners

DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter

Generative Pretraining & Fine Tuning

Pretraining

Train the GPT Architecture to understand what language is

Fine Tuning

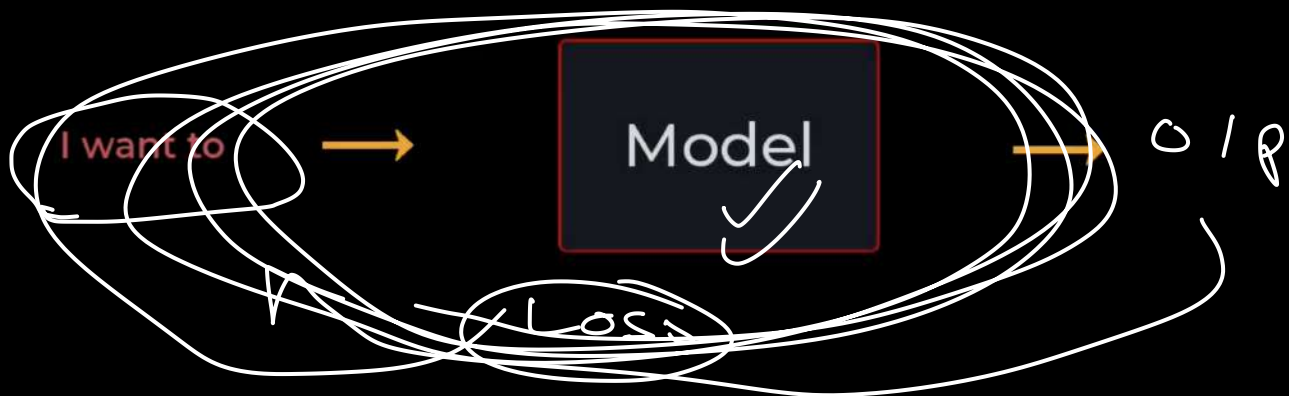
Use transfer learning to make GPT architecture perform well on specific tasks.

Fine → ✓

Generative Pretraining & Fine Tuning

Pretraining

Train the GPT Architecture to understand what language is



Issues with Fine Tuning

GPT

Chat
BCL

1. Still too much data required!

Task

Data Requirements

Machine Translation

~100,000 samples

Question-Answer

~100,000 samples

Text Summarization

~100,000 samples

Sentence Similarity

~100,000 samples

Issues with Fine Tuning

DL

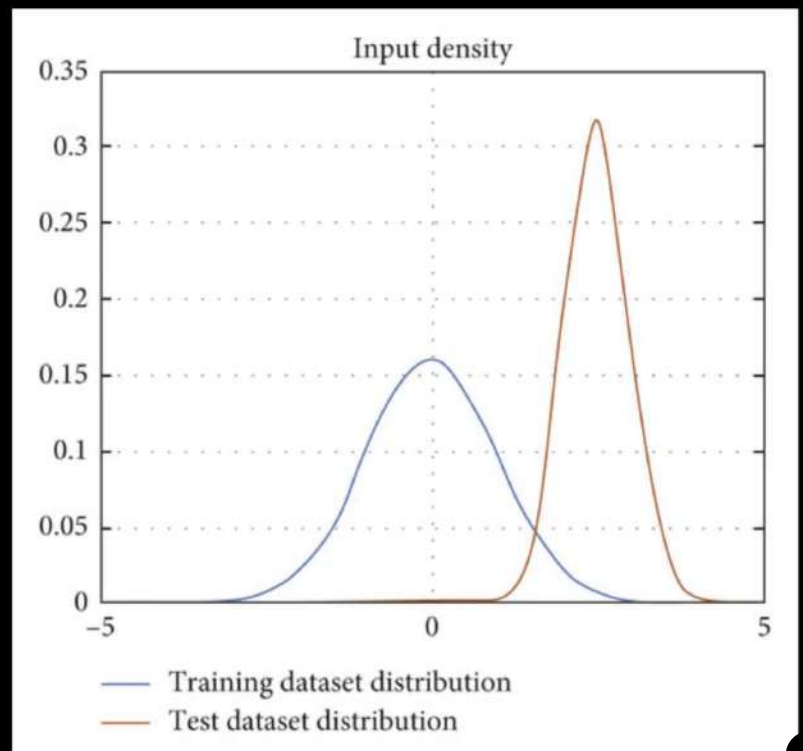
1. Still too much data required!

2. Overfitting is easy

Broad Pretrained Dataset

Narrow Fine Tuned Dataset

-0.52	-1.31	-0.38	0.32
-0.28	-0.58	-1.53	0.14
-0.29	0.39	-0.21	0.56
-0.67	0.19	0.38	-0.42
0.74	-0.09	0.17	-0.55
-2.31	-1.37	-0.88	-0.35



GPT-1

GPT 7 = Tranyton (decoder).
Issues with Fine Tuning Eng ✓

1. Still too much data required!

2. Overfitting is easy

3. Not how humans learn

French ✓

Q/A

<u>Task</u>	<u>Data Requirements</u>	<u>Human</u>
Machine Translation	~100,000 samples	10 samples
Question-Answer	~100,000 samples	1 sample
Text Summarization	~100,000 samples	2 samples
Sentence Similarity	~100,000 samples	3 samples

Glen AI

GPT-2 & Meta-Learning

GPT-1

Zero-Shot Learning

Perform specific task when given just an instruction + input

Question => Answer

What is the
closest planet
to the sun?



GPT-2



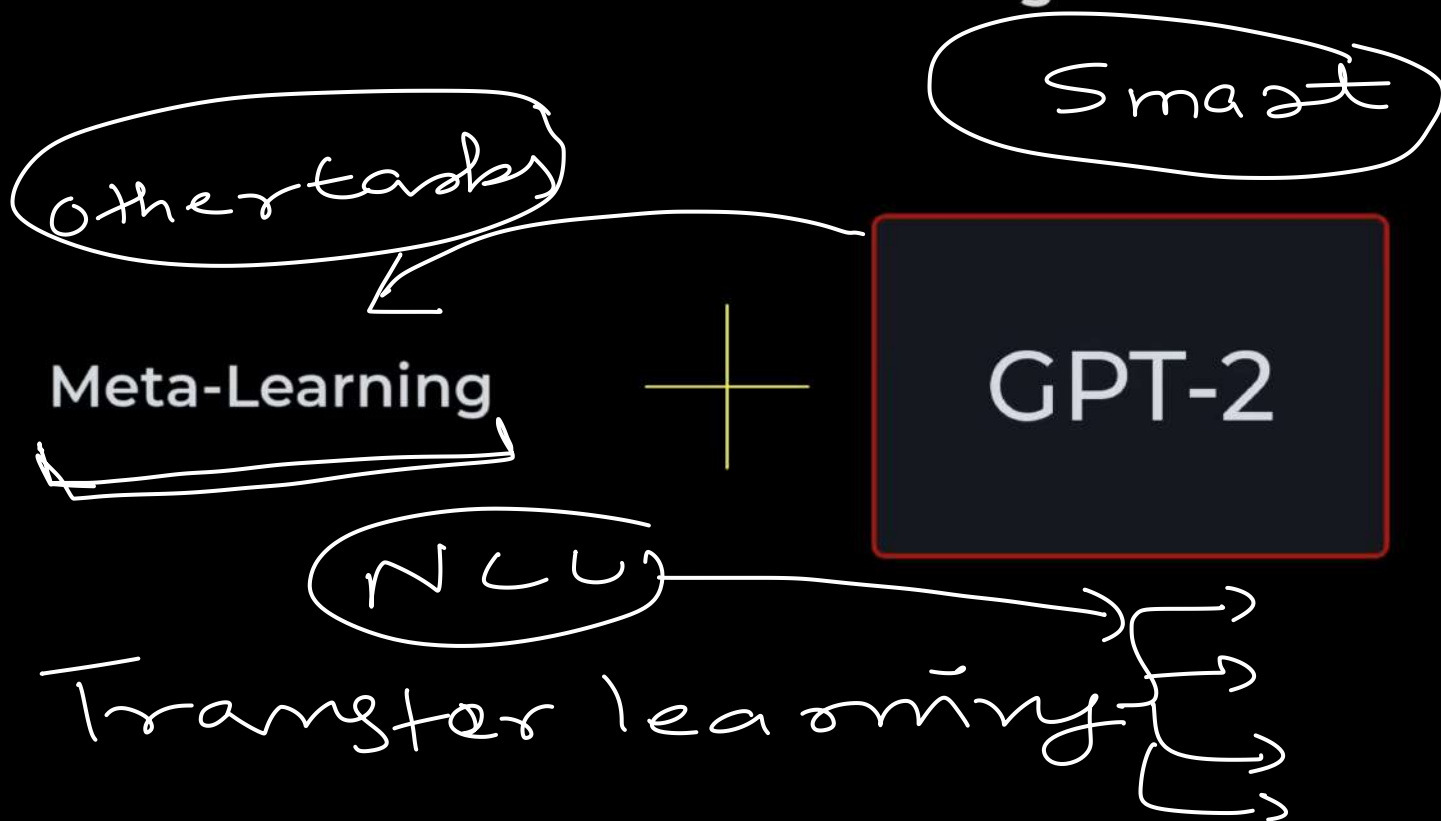
Mercury is the
closest planet to
the sun. It's hot
there

Best Base
BERT
large

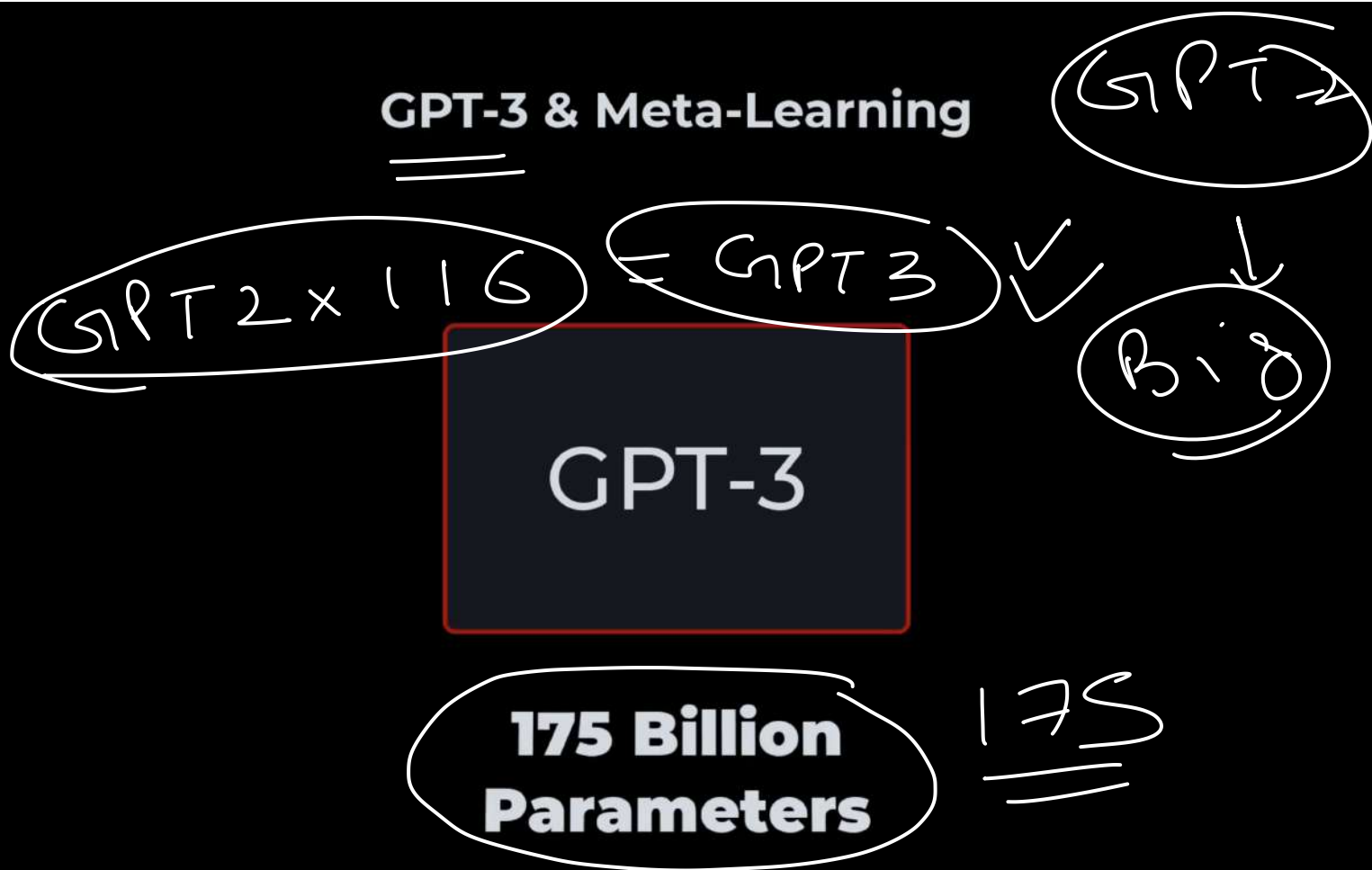
1.5 Billion
Parameters

1.5

GPT-2 & Meta-Learning



GPT-3 & Meta-Learning



GPT training?

Step 1

Collect demonstration data, and train a supervised policy.

$$\begin{aligned} 1 - P_1 &= A \\ 2 - P_2 &= A \\ 3 - P_3 &= A \\ 4 - P_4 &= A \end{aligned}$$

Step 2

Collect comparison data, and train a reward model.

rule gradient

$$\begin{aligned} &\checkmark \quad \times \quad \times \\ &P + \times \quad \text{Profanity} \end{aligned}$$

Step 3

Optimize a policy against the reward model using reinforcement learning.

Fine tuned

GPT-3

Daily

ChatGPT

Prompt + Accepted Answer

A prompt is sample from our prompt dataset.

Explain the moon landing to a 6 year old

A prompt and several model outputs are sampled.

Explain the moon landing to a 6 year old

A new prompt is sampled from the dataset.

Write a story about frogs

A labeler demonstrates the desired output behavior.



Some people went to the moon...

Human supervise

GPT

Explain gravity... Explain war...
Moon is natural satellite of... People went to the moon...

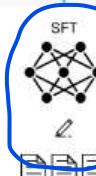
The policy generates an output.



Once upon a time...

Step 1

This data is used to fine-tune GPT-3 with supervised learning.



A labeler ranks the outputs from best to worst.

Human

D > C > A = B

This data is used to train our reward model.



D > C > A = B

The reward model calculates a reward for the output.



The reward is used to update the policy using PPO.

r_k

$$\lambda (r_k - \hat{v})$$