

The science & technology behind ChatGPT

ChatGPT vs older language technology

Unsupervised learning supplemented with “Reinforcement Learning from Human Feedback (RLHF)”



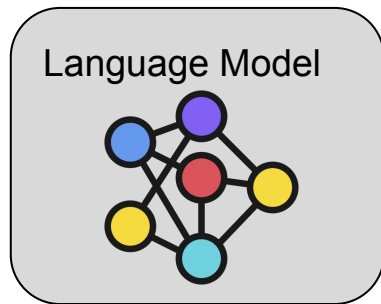
Unsupervised learning
from tons of data

Learning from human
trainers

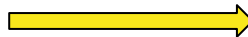
Unsupervised Learning



It's hot. Open the



PREDICTION



Training Data

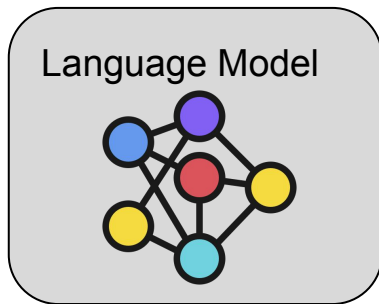
Transformer Neural Network

Unsupervised Learning



It's hot. Open the

Training Data



Transformer Neural Network

PREDICTION



0.2% ... 10% ... 40% 1%

ask

door

window

what

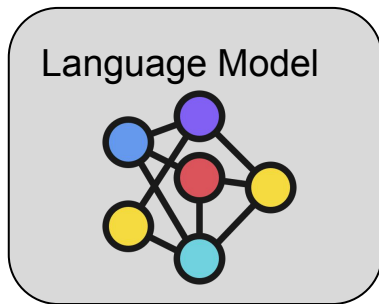
vocabulary words

Unsupervised Learning



It's hot. Open the

Training Data



Transformer Neural Network

calculate error &
update model



PREDICTION



actual next word
door

0.2%	...	10%	...	40%	1%
------	-----	-----	-----	-----	----

ask	door	window	what
-----	------	--------	------

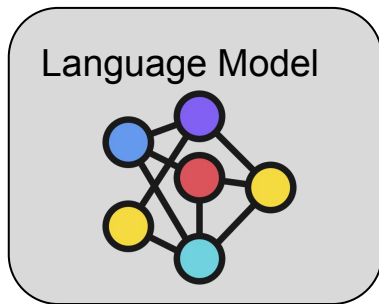
vocabulary words

Unsupervised Learning



It's hot. Open the

Training Data



Transformer Neural Network

calculate error &
update model



PREDICTION



actual next word
door

0.2%	...	10%	...	40%	1%
------	-----	-----	-----	-----	----

ask	door	window	what
-----	------	--------	------

vocabulary words



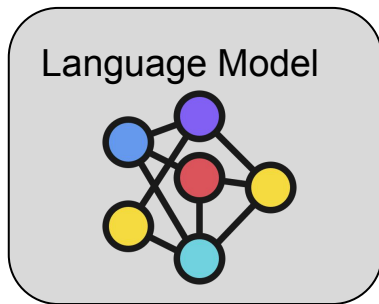
- ~ 175 billion parameters
- Exposed to 300 billion words during training

Unsupervised Learning



It's hot. Open the

Training Data



Transformer Neural Network

calculate error &
update model



PREDICTION



actual next word
door

0.2%	...	10%	...	40%	1%
------	-----	-----	-----	-----	----

ask	door	window	what
-----	------	--------	------

vocabulary words



- ~ 175 billion parameters
- Exposed to 300 billion words during training

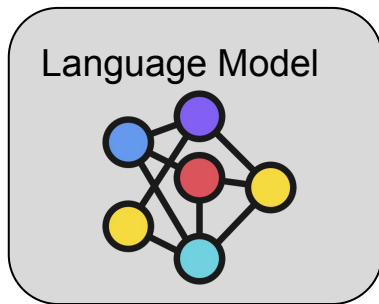
How long does it take to train it?

Unsupervised Learning



It's hot. Open the

Training Data



Transformer Neural Network

calculate error &
update model



PREDICTION



actual next word
door

0.2%	...	10%	...	40%	1%
------	-----	-----	-----	-----	----

ask	door	window	what
-----	------	--------	------

vocabulary words



- ~ 175 billion parameters
- Exposed to 300 billion words during training

How long does it take to train it?

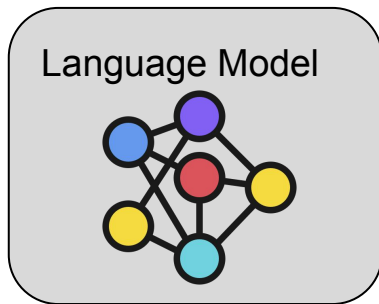
- Months on computer clusters
- Centuries on a laptop

Unsupervised Learning



It's hot. Open the

Training Data



Transformer Neural Network

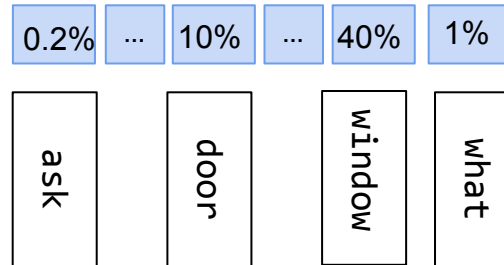
calculate error &
update model



PREDICTION



actual next word
door



vocabulary words



- ~ 175 billion parameters
- Exposed to 300 billion words during training

How long does it take to train it?

- Months on computer clusters
- Centuries on a laptop

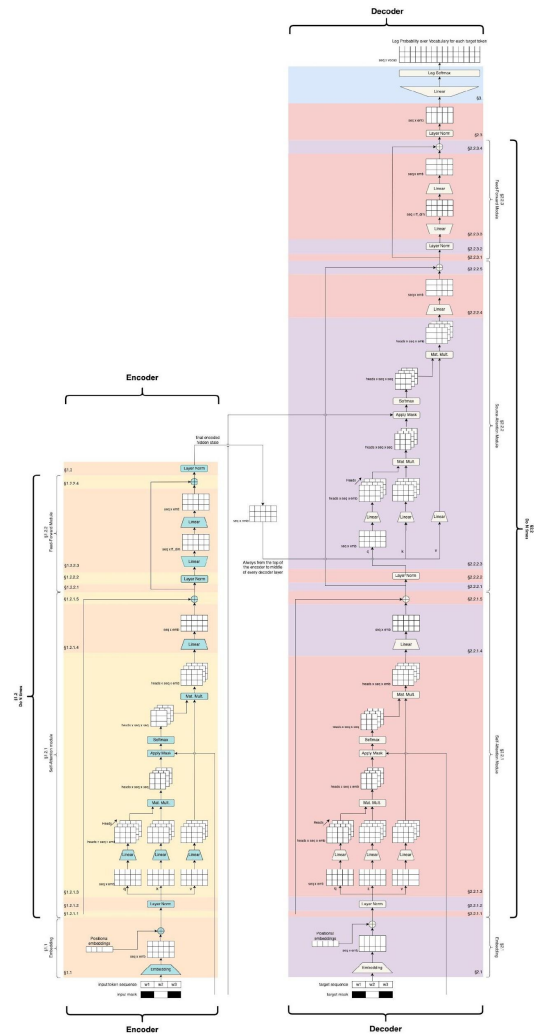
**After training, it is not
connected to the Internet!**



Google is a **search** engine,
a **document retrieval** system



ChatGPT is a **language generation**
system (**makes up stuff!**)



The journalist talked to Sara about [MASK] book



Q

Her? Query: is there a female, human, adult antecedent?
His? Query: is there a male, human, adult antecedent?



K

K

Journalist! Key: I'm a male or female, human, adult antecedent!

Sara! Key: I'm a female, human, adult antecedent!



V

"her"

Journalist! Value: I'm a journalist

Sara! Value: I'm Sara

100 words (conversation)

X 100

=10,000 (book)

X 100

=1,000,000 (book collection)

X 100

=100,000,000 (library)

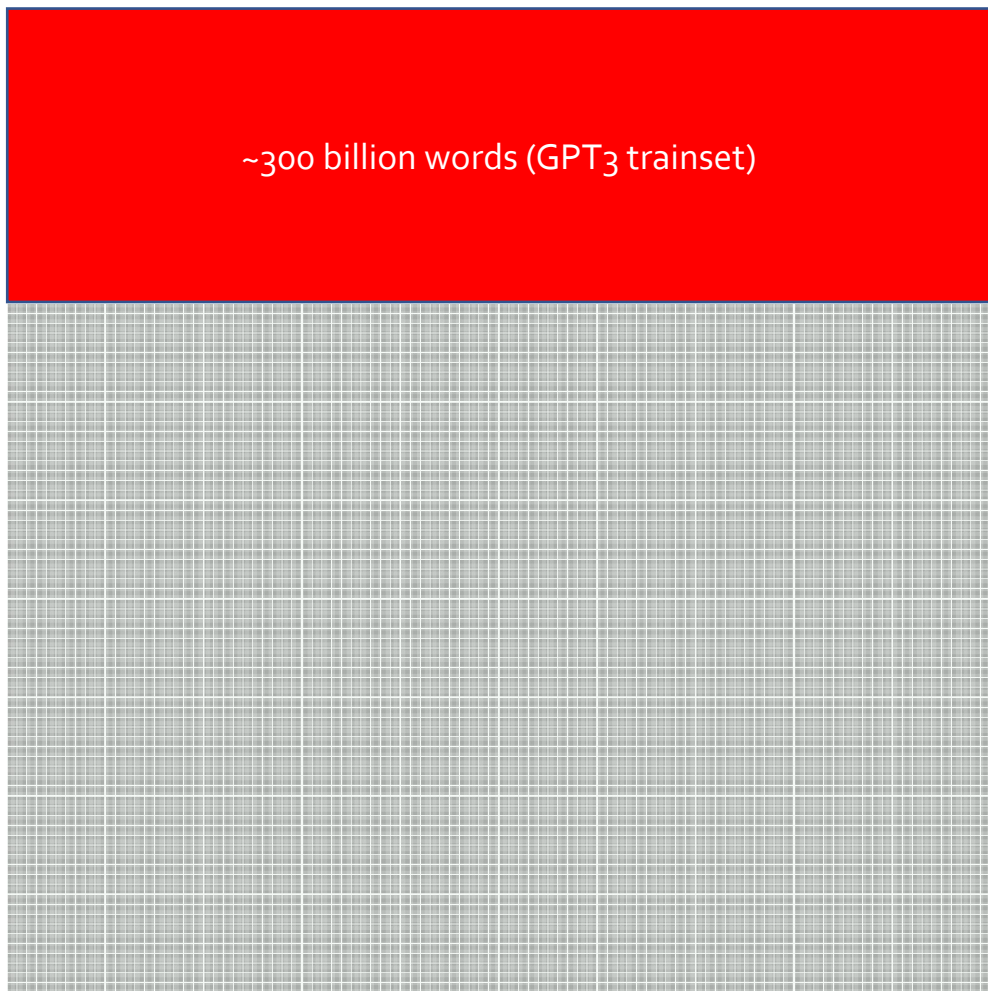
X 100

=10,000,000,000 (wikipedia)

X 100

=1,000,000,000,000
(the internet)

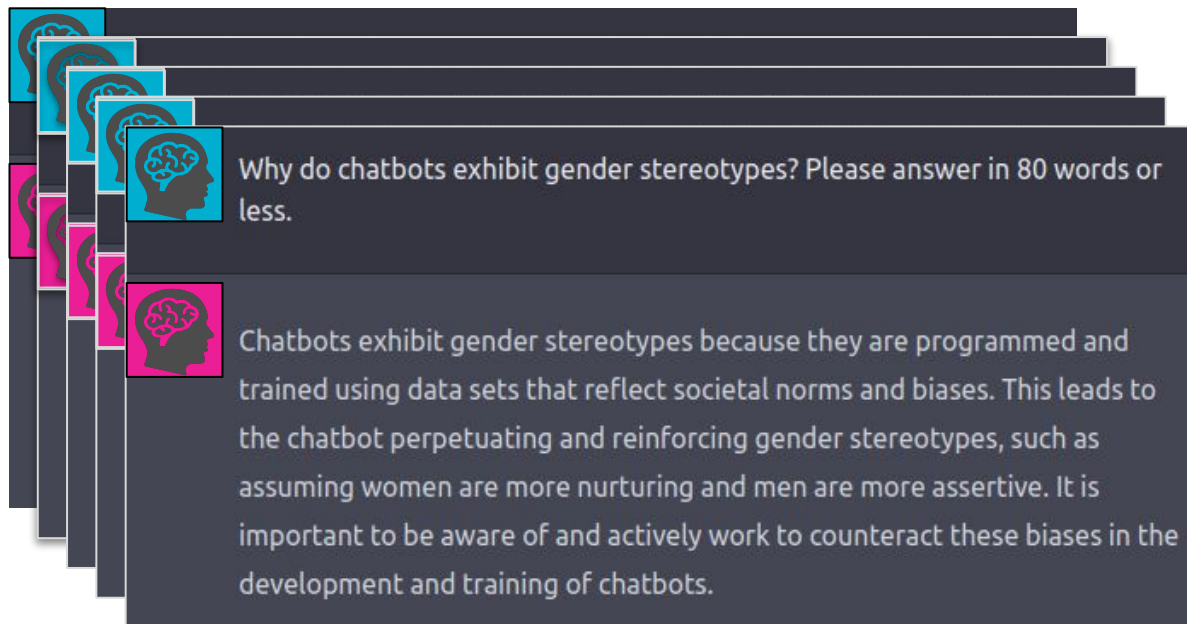
~300 billion words (GPT₃ trainset)



From GPT to ChatGPT: Learning from human feedback

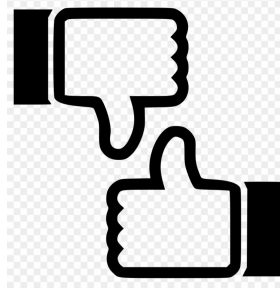


Learning from **example conversations** written by human annotators



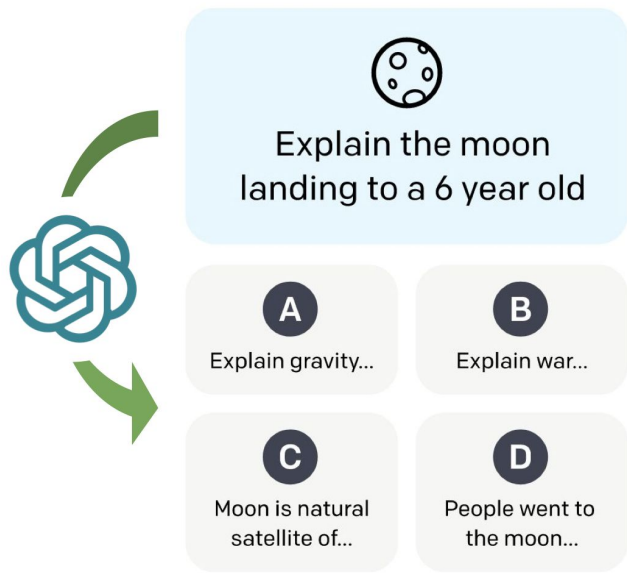
Unsupervised learning:
general patterns and most likely outcome in huge data

Supervised learning:
fine-tuned knowledge related to dialogue skills

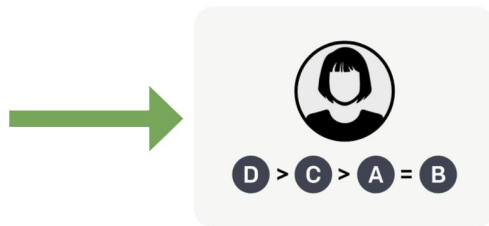


Reinforcement Learning

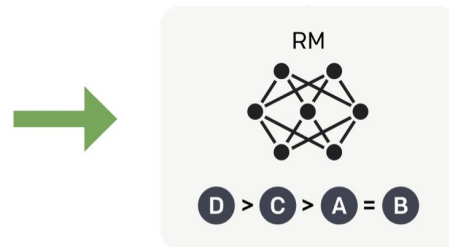
Generate examples



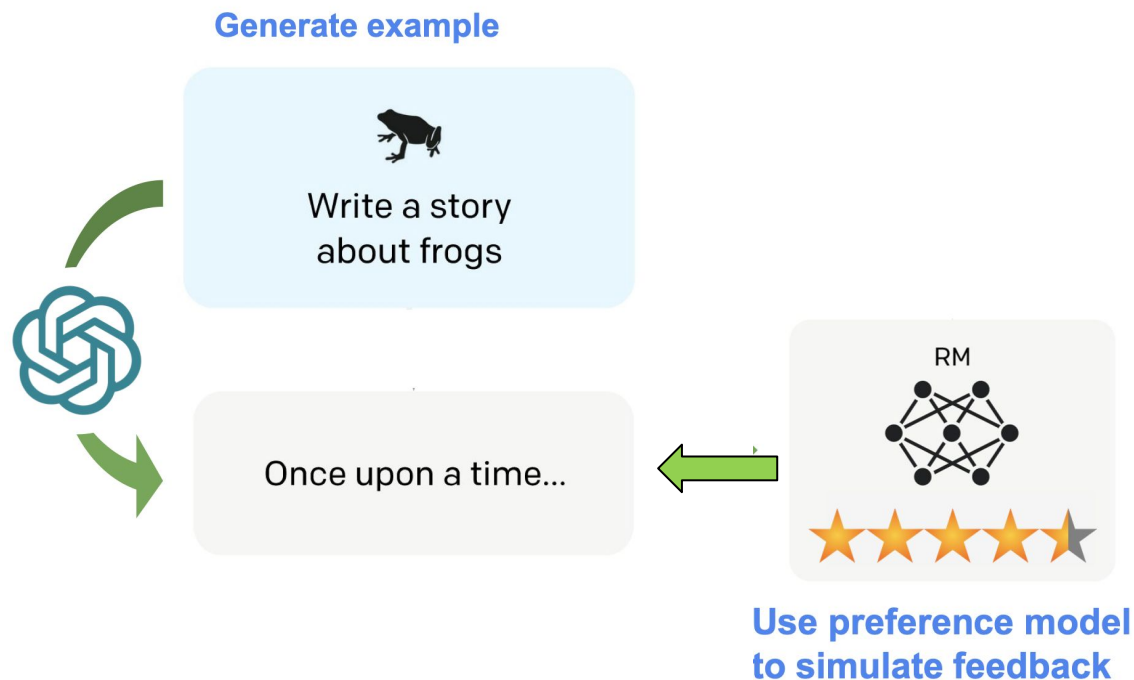
Annotate preferences



Train preference model



Learning from humans: Reinforcement Learning



Your feedback is used for updating ChatGPT

