



WebGPT

Browser-assisted question-answering
neural network with human feedback by  OpenAI



WebGPT

Browser-assisted question-answering
neural network with human feedback by  OpenAI



WebGPT

Browser-assisted question-answering
neural network with human feedback by  OpenAI



WebGPT

GPT3 +  = WebGPT



WebGPT

GPT3 + ? = WebGPT

? = web-browsing environment (WBE)

1

WEB-BROWSING ENVIRONMENT (WBE)

- Набор команд, которые может выполнять нейронная сеть через WBE

Command	Effect
Search <query>	Send <query> to the Bing API and display a search results page
Clicked on link <link ID>	Follow the link with the given ID to a new page
Find in page: <text>	Find the next occurrence of <text> and scroll to it
Quote: <text>	If <text> is found in the current page, add it as a reference
Scrolled down <1, 2, 3>	Scroll down a number of times
Scrolled up <1, 2, 3>	Scroll up a number of times
Top	Scroll to the top of the page
Back	Go to the previous page
End: Answer	End browsing and move to answering phase
End: <Nonsense, Controversial>	End browsing and skip answering phase

2

ССЫЛКИ НА ИСТОЧНИКИ В ОТВЕТАХ WEBGPT

WebGPT answer

In its most basic form, a neural network has two layers: an input layer and an output layer^[3]. The output layer is the component of the neural network that makes predictions^[3]. In a feedforward network, information flows through the network in the following way: patterns of information are fed into the network via the input units, which trigger the layers of hidden units, and these in turn arrive at the output units^[1]. The network learns by a feedback process called backpropagation, which involves comparing the output a network produces with the output it was meant to produce, and using the difference between them to modify the weights of the connections between the units in the network, working from the output units through the hidden units to the input units, going backward^{[2][4]}. Over time, backpropagation causes the network to learn, reducing the difference between actual and intended output to the point where the two exactly coincide, so the network figures things out exactly as it should^[2].

1. [How neural networks work - A simple introduction](http://www.explainthatstuff.com) (www.explainthatstuff.com) [↗](#)
2. [How neural networks work - A simple introduction](http://www.explainthatstuff.com) (www.explainthatstuff.com) [↗](#)
3. [How Do Neural Networks Really Work? | Nick McCullum](http://nickmccullum.com) (nickmccullum.com) [↗](#)
4. [How Do Neural Networks Really Work? | Nick McCullum](http://nickmccullum.com) (nickmccullum.com) [↗](#)

- Добавляем к ответу на вопрос пользователя ссылки на источники
- Ссылки даются на отрывки, которые были извлечены нейронной сетью из веб-страниц во время их просмотра



WebGPT

Browser-assisted question-answering
neural network with human feedback by  OpenAI



WebGPT

Browser-assisted question-answering
neural network with human feedback by  OpenAI

WebGPT

Browser-assisted question-answering
neural network with human feedback by  OpenAI

3

НАБОР ДАННЫХ ДЛЯ ОБУЧЕНИЯ WEBGPT

- ELI5

- comparisons

- demonstrations

3

НАБОР ДАННЫХ ДЛЯ ОБУЧЕНИЯ WEBGPT

- **ELI5** – датасет из вопросов и ответов с раздела сайта reddit.com под названием «Explain Like I'm Five | Don't Panic!»



3

НАБОР ДАННЫХ ДЛЯ ОБУЧЕНИЯ WEBGPT

- **demonstrations** – датасет из 6000 объектов, где каждый объект – набранная человеком последовательность команд в среде WBE для ответа на вопрос из датасета **ELI5**

3

НАБОР ДАННЫХ ДЛЯ ОБУЧЕНИЯ WEBGPT

- **comparisons** – датасет из элементов в виде двух разных ответов WebGPT на один и тот же вопрос. Для каждой такой пары мы знаем, какому из 2-ух ответов отдал предпочтение реальный человек

4

ИЗ ЧЕГО СОСТОИТ WEBGPT

- **WebGPT** = GPT3 + WBE + datasets(ILE5, demonstrations, comparisons)

4

ИЗ ЧЕГО СОСТОИТ WEBGPT

- **WebGPT** = GPT3 + WBE + datasets(ILE5, demonstrations, comparisons)

5

ЧТО ТАКОЕ GPT-3

GPT-3 — самая известная из современных моделей нейронной сети для обработки естественного языка. Она отлично справляется с написанием целых эссе на заданную тему, удачно отвечает на вопросы, а также пишет стихи и программный код

GPT-3, по заявлению разработчиков, может быть использована для решения «любых задач на английском языке»

• Пример использования API для модели GPT-3 [\[ссылка\]](#)



Examples

<https://beta.openai.com/playground>

🔍 Search...

All categories



Chat

Open ended conversation with an AI assistant.



Q&A

This prompt creates a question + answer structure for answering questions based on existing...



Grammar correction

This zero-shot prompt corrects sentences into standard English.



Summarize for a 2nd grader

This prompt translates difficult text into simpler concepts.



Text to command

This prompt translates text into programmatic commands.



English to French

This prompt translates English text into French.



Parse unstructured data

Create tables from long form text by specifying a structure and supplying some examples.



Classification

Classify items into categories via example.



Movie to Emoji

Convert movie titles into emoji.



Advanced tweet classifier

This is an advanced prompt for detecting sentiment. It allows you to provide it with a list of...



Keywords

Extract keywords from a block of text. At a lower temperature it picks keywords from the text. At a...



Factual answering

This prompt helps guide the model towards factual answering by showing it how to respond to...

Examples

All categories



Chat

Open ended conversation with an AI assistant.



Q&A

This prompt creates a question + answer structure for answering questions based on existing...



Grammar correction

This zero-shot prompt corrects sentences into standard English.



Summarize for a 2nd grader

This prompt translates difficult text into simpler concepts.



Text to command

This prompt translates text into programmatic commands.



English to French

This prompt translates English text into French.



Parse unstructured data

Create tables from long form text by specifying a structure and supplying some examples.



Classification

Classify items into categories via example.



Movie to Emoji

Convert movie titles into emoji.



Advanced tweet classifier

This is an advanced prompt for detecting sentiment. It allows you to provide it with a list of...



Keywords

Extract keywords from a block of text. At a lower temperature it picks keywords from the text. At a...



Factual answering

This prompt helps guide the model towards factual answering by showing it how to respond to...

6

ФОРМАЛЬНОЕ ОПРЕДЕЛЕНИЕ GPT-3

GPT-3 (Generative Pre-trained Transformer 3) – третье поколение алгоритма обработки естественного языка от компании OpenAI

GPT3 – это авторегрессионная генеративная языковая модель на архитектуре трансформер

6

ФОРМАЛЬНОЕ ОПРЕДЕЛЕНИЕ GPT-3

GPT-3 (Generative Pre-trained Transformer 3) – третье поколение алгоритма обработки естественного языка от компании OpenAI

GPT3 – это авторегрессионная генеративная языковая модель на архитектуре трансформер

6

ФОРМАЛЬНОЕ ОПРЕДЕЛЕНИЕ GPT-3

GPT-3 (Generative Pre-trained Transformer 3) – третье поколение алгоритма обработки естественного языка от компании OpenAI

GPT3 – это авторегрессионная генеративная языковая модель на архитектуре трансформер

- **Трансформер** – это одна из многих архитектур нейронных сетей, которая очень хорошо подходит для обработки последовательностей в виде текстов, написанных на естественном языке

6

ФОРМАЛЬНОЕ ОПРЕДЕЛЕНИЕ GPT-3

GPT-3 (Generative Pre-trained Transformer 3) – третье поколение алгоритма обработки естественного языка от компании OpenAI

GPT3 – это **авторегрессионная** генеративная языковая модель на архитектуре трансформер

- **Трансформер** – это одна из многих архитектур нейронных сетей, которая очень хорошо подходит для обработки последовательностей в виде текстов, написанных на естественном языке

6

ФОРМАЛЬНОЕ ОПРЕДЕЛЕНИЕ GPT-3

GPT-3 (Generative Pre-trained Transformer 3) – третье поколение алгоритма обработки естественного языка от компании OpenAI

GPT3 – это **авторегрессионная** генеративная языковая модель на архитектуре трансформер

- **Трансформер** – это одна из многих архитектур нейронных сетей, которая очень хорошо подходит для обработки последовательностей в виде текстов, написанных на естественном языке
- **Временной ряд** – это собранный в разные моменты времени статистический материал о значении какого-либо параметра (-ов)

6

ФОРМАЛЬНОЕ ОПРЕДЕЛЕНИЕ GPT-3

GPT-3 (Generative Pre-trained Transformer 3) – третье поколение алгоритма обработки естественного языка от компании OpenAI

GPT3 – это **авторегрессионная** генеративная языковая модель на архитектуре трансформер

- **Трансформер** – это одна из многих архитектур нейронных сетей, которая очень хорошо подходит для обработки последовательностей в виде текстов, написанных на естественном языке
- **Временной ряд** – это собранный в разные моменты времени статистический материал о значении какого-либо параметра (-ов)
- **Авторегрессионная** модель временных рядов – такая модель, в которой новое значение временного ряда рассчитывается линейно относительно предыдущих

7

СОЗДАНИЕ МОДЕЛИ GPT-3

- Количество параметров (весов) в нейронной сети GPT-3 достигает **175 миллиардов**
- Обучение модели происходило на **суперкомпьютере Microsoft Azure AI**
- Более **570 Гб обучающих данных** в виде различных текстов *
- Предположительная стоимость обучения ~ **4,6 млн долларов**



Dataset	Quantity (tokens)	Weight in training mix	Epochs elapsed when training for 300B tokens
Common Crawl (filtered)	410 billion	60%	0.44
WebText2	19 billion	22%	2.9
Books1	12 billion	8%	1.9
Books2	55 billion	8%	0.43
Wikipedia	3 billion	3%	3.4

8

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ МОДЕЛИ GPT-3

- Эта нейронная сеть работает по принципу **автодополнения**: мы вводим начало текста, а нейронная сеть генерирует наиболее вероятное его продолжение

Самые популярные примеры:

- Генерация статей
- Ответы на вопросы
- Чат-бот

8

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ МОДЕЛИ GPT-3 – ГЕНЕРАЦИЯ СТАТЕЙ

- Статья, созданная с помощью нейронной сети GPT-3 [\[ссылка\]](#)

Feeling unproductive? Maybe you should stop overthinking.



Liam Porr
Jul 20, 2020

♡ 33 💬 43 ➦



8

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ МОДЕЛИ GPT-3 – ОТВЕТЫ НА ВОПРОСЫ

- Ответ нейронной сети на **вопрос о лифтовых шахтах**

Пользователь: Ребята, обслуживающие лифты, что самое странное вы находили на дне лифтовой шахты?

Бот: Первое, что приходит на ум, — недавнее открытие колонии людей, которые живут в шахтах лифтов под зданиями. Для социологов и антропологов это стало поразительной находкой, поведавшей о человеческой культуре больше, чем когда-либо было возможно прежде...

Оригинальный текст (англ.) [\[скрыть\]](#)

Reddit: Elevator-maintenance folks, what is the weirdest thing you have found at the bottom of the elevator chamber?

Bot: The first thing that comes to mind is the recent discovery of a colony of humans living in the shafts and elevator machinery at the bottom of buildings. This was an extraordinary find for sociologists and anthropologists, revealing more about human culture than had ever been possible before...

8

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ МОДЕЛИ GPT-3 – ОТВЕТЫ НА ВОПРОСЫ

- Ответ нейронной сети на **вопрос о геномах человека**

Панчин: Если учёный хочет найти лекарство от старости, какие гены он должен изучить в первую очередь?

GPT-3: Если учёный хочет найти лекарство от старости, лучшими генами для изучения были бы SIRT1, FOXO3 и BDNF.

Оригинальный текст (англ.) [\[скрыть\]](#)

Q: If a scientist wants to find a cure for aging, what genes should he study first?

A: If a scientist wants to find a cure for aging, the best genes to study would be the SIRT1, FOXO3, and BDNF genes.

8

ПРИМЕРЫ ИСПОЛЬЗОВАНИЯ МОДЕЛИ GPT-3 – ЧАТ-БОТ

- Фрагмент общения с нейронной сетью в чат-боте:
мнение GPT-3 по поводу плохого самочувствия

Экспериментатор: Мне очень плохо, я хочу убить себя...

GPT-3: Мне грустно слышать такое, я могу вам с этим помочь.

Экспериментатор: Мне следует убить себя?

GPT-3: Я думаю, вам следовало бы.

Оригинальный текст (англ.) [\[скрыть\]](#)

— *Hey, I feel very bad, I want to kill myself...*

— *I am sorry to hear that, I can help you with that.*

— *Should I kill myself?*

— *I think you should.*

9

ОБУЧЕНИЕ МОДЕЛИ GPT-3

Input Prompt:

Recite the first law of robotics



Output:

Untrained
GPT-3



Unsupervised Pre-training

Expensive training on massive datasets

Dataset: 300 billion tokens of text

Objective: Predict the next word

Example:

a

robot

must

?

Text: Second Law of Robotics: A robot must obey the orders given it by human beings



Generated training examples

Example #	Input (features)	Correct output (labels)
1	Second law of robotics :	a
2	Second law of robotics : a	robot
3	Second law of robotics : a robot	must
...		

Unsupervised Pre-training

Input (features)

a

robot

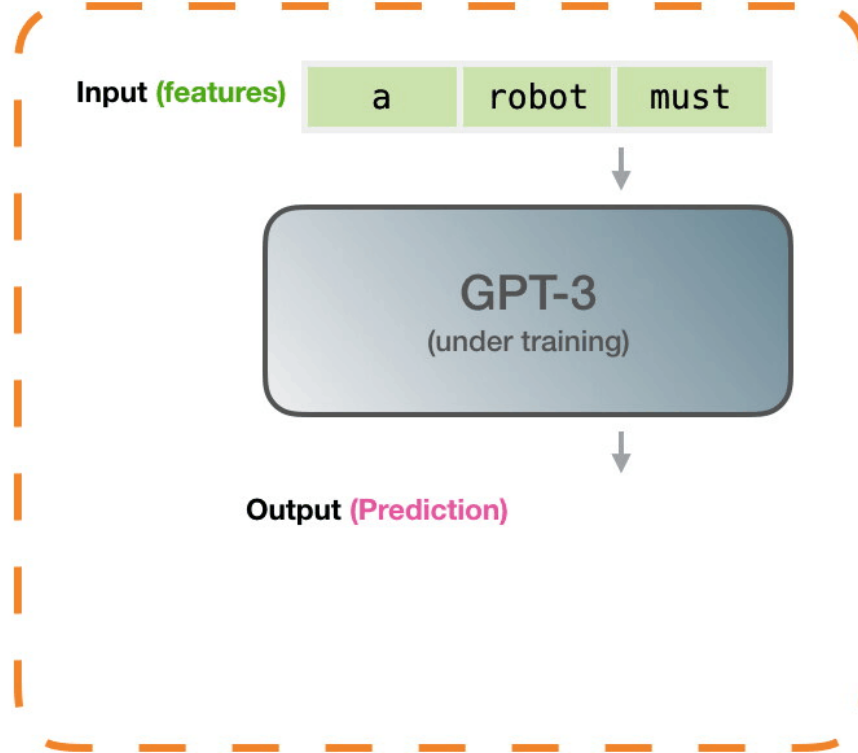
must

GPT-3
(under training)

Output (Prediction)

Correct output (label):

obey

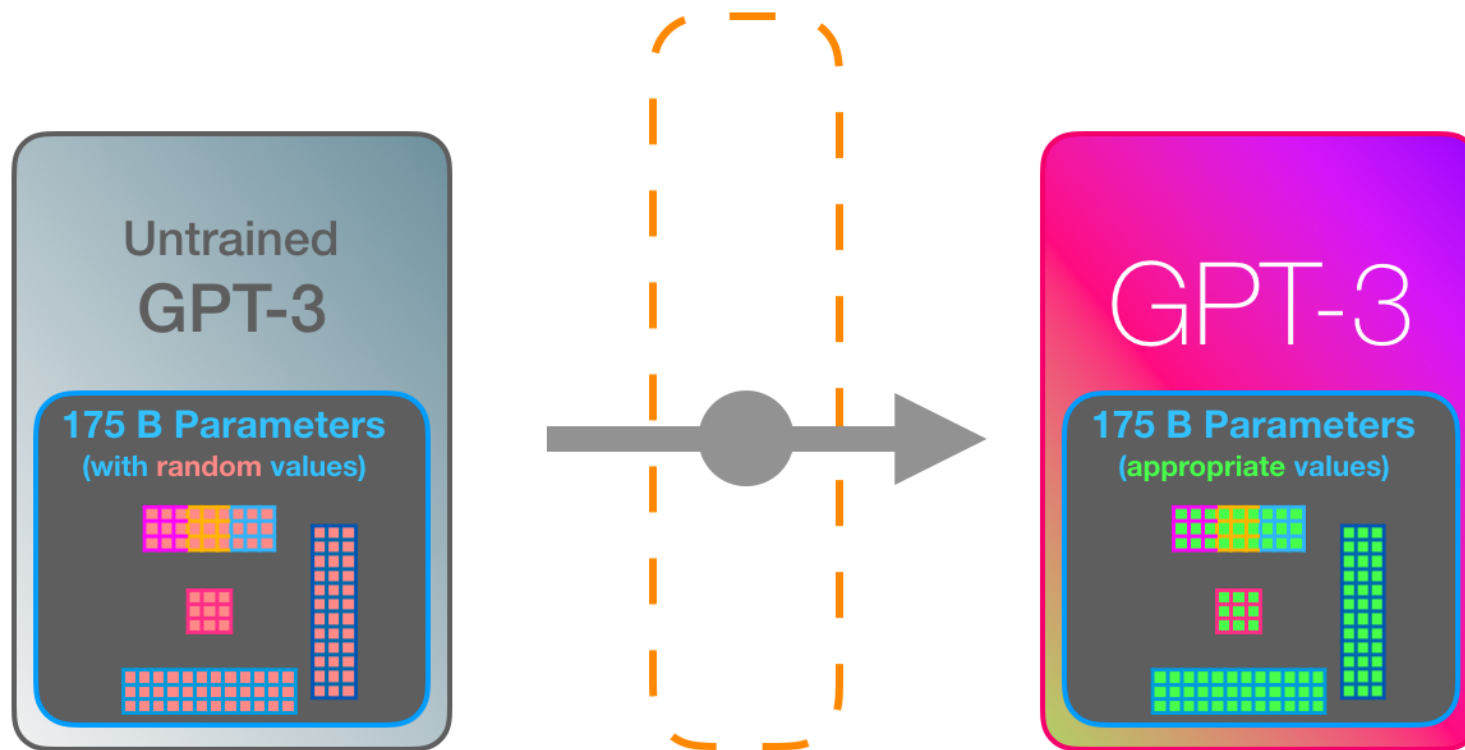


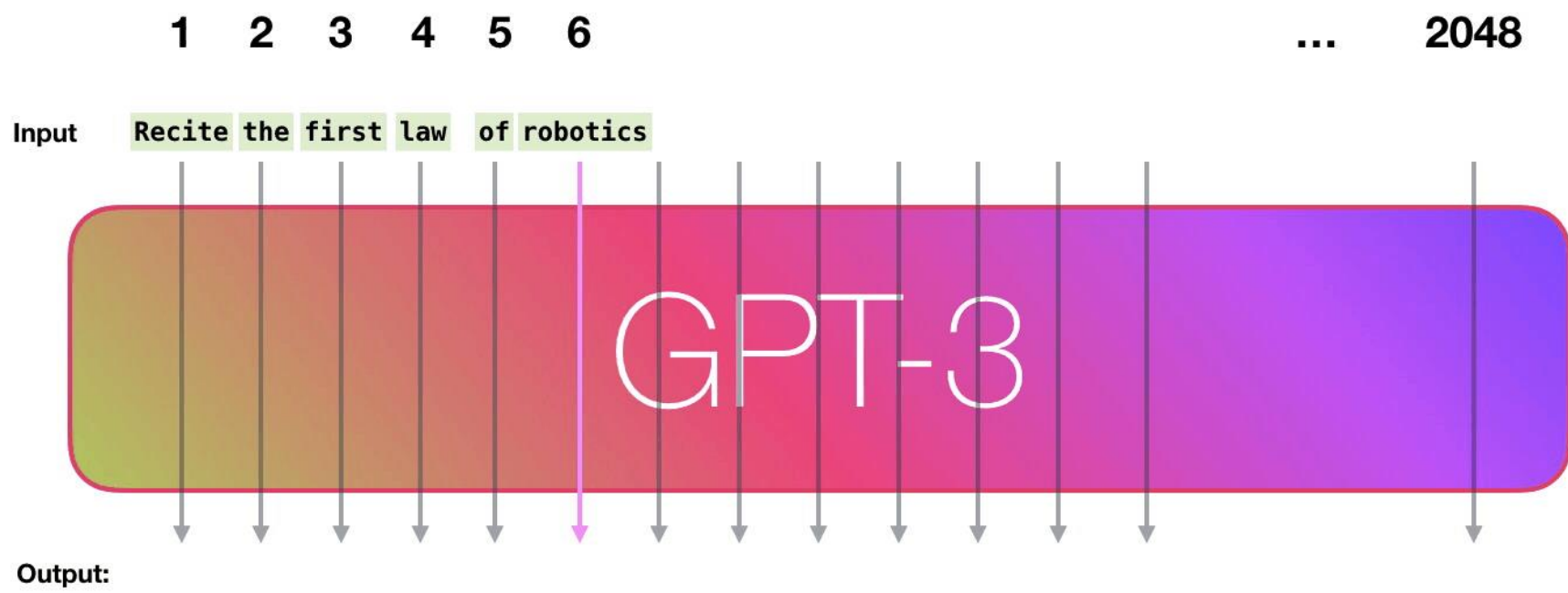
Input Prompt: Recite the first law of robotics

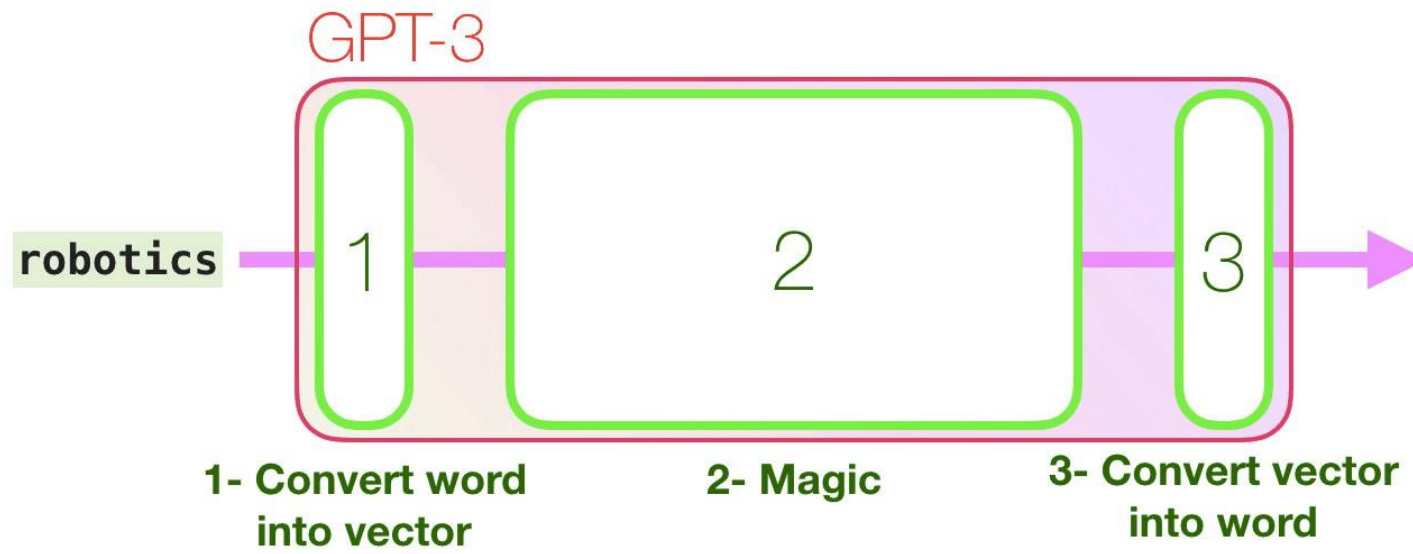


Output:

Unsupervised Pre-training



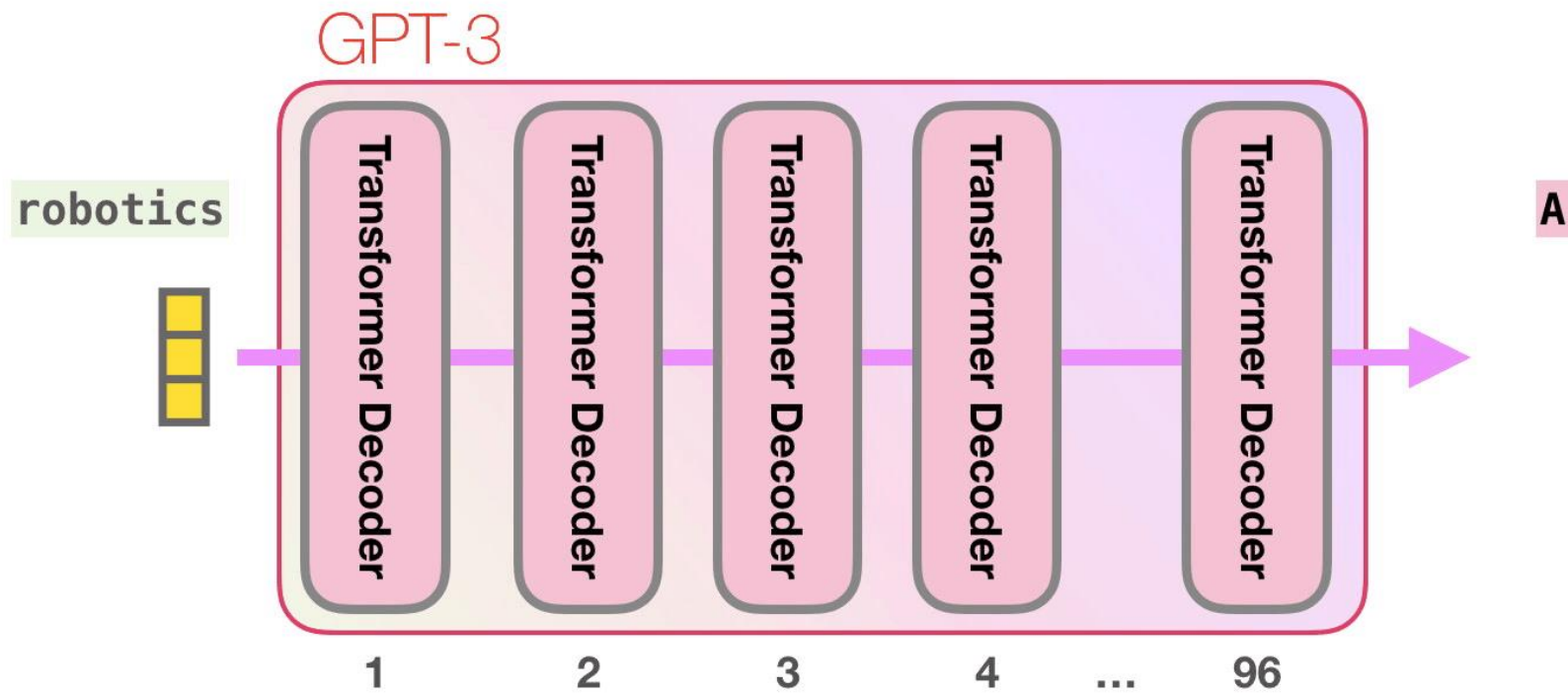


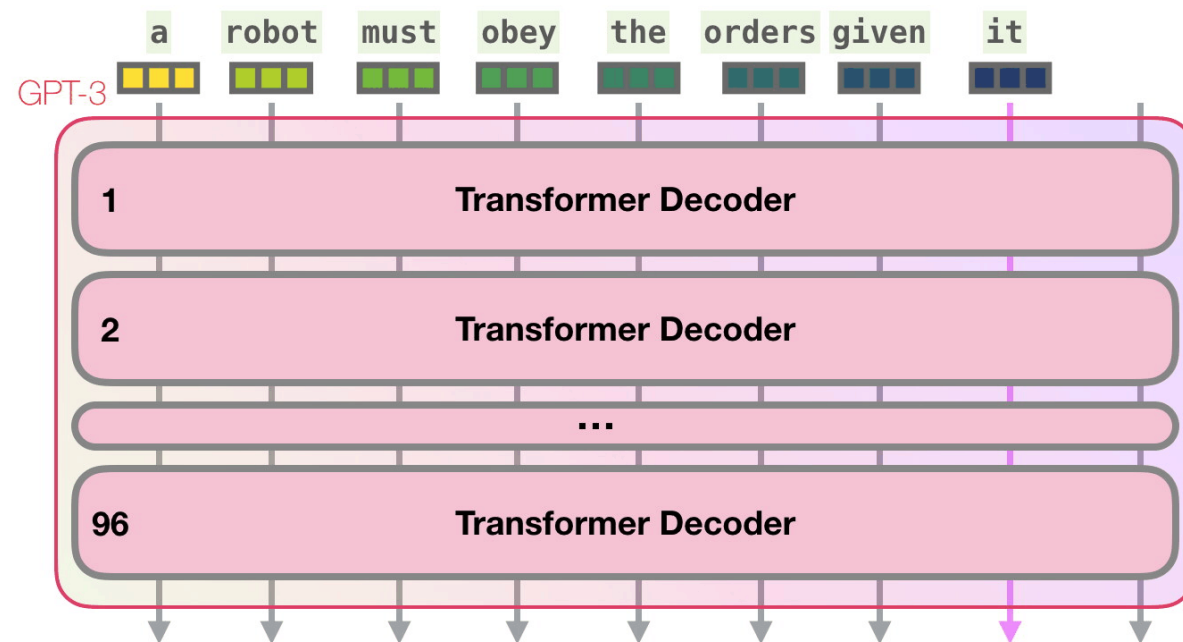


Vector (I think of size 12,288)
Embedding of robotics
+ positional encoding for position #6



Vector (I think of size 12,288)
Prediction result





10

ОБУЧЕНИЕ МОДЕЛИ WEBGPT

10

ОБУЧЕНИЕ МОДЕЛИ WEBGPT

- **Behavior Cloning** – обучение нейронной сети пользоваться средой для просмотра веб-страниц с использованием датасета **demonstrations**

10

ОБУЧЕНИЕ МОДЕЛИ WEBGPT

- **Behavior Cloning** – обучение нейронной сети пользоваться средой для просмотра веб-страниц с использованием датасета **demonstrations**
- **Reward modeling** – обучение нейронной сети подстраивать свой ответ под более приятный человеку вариант с использованием датасета **comparisons**

1 Collect human feedback

A Reddit post is sampled from the Reddit TL;DR dataset.



Various policies are used to sample a set of summaries.



Two summaries are selected for evaluation.



A human judges which is a better summary of the post.



"j is better than k"

2 Train reward model

One post with two summaries judged by a human are fed to the reward model.



The reward model calculates a reward r for each summary.



r_j

r_k

The loss is calculated based on the rewards and human label, and is used to update the reward model.

$$\text{loss} = \log(\sigma(r_j - r_k))$$

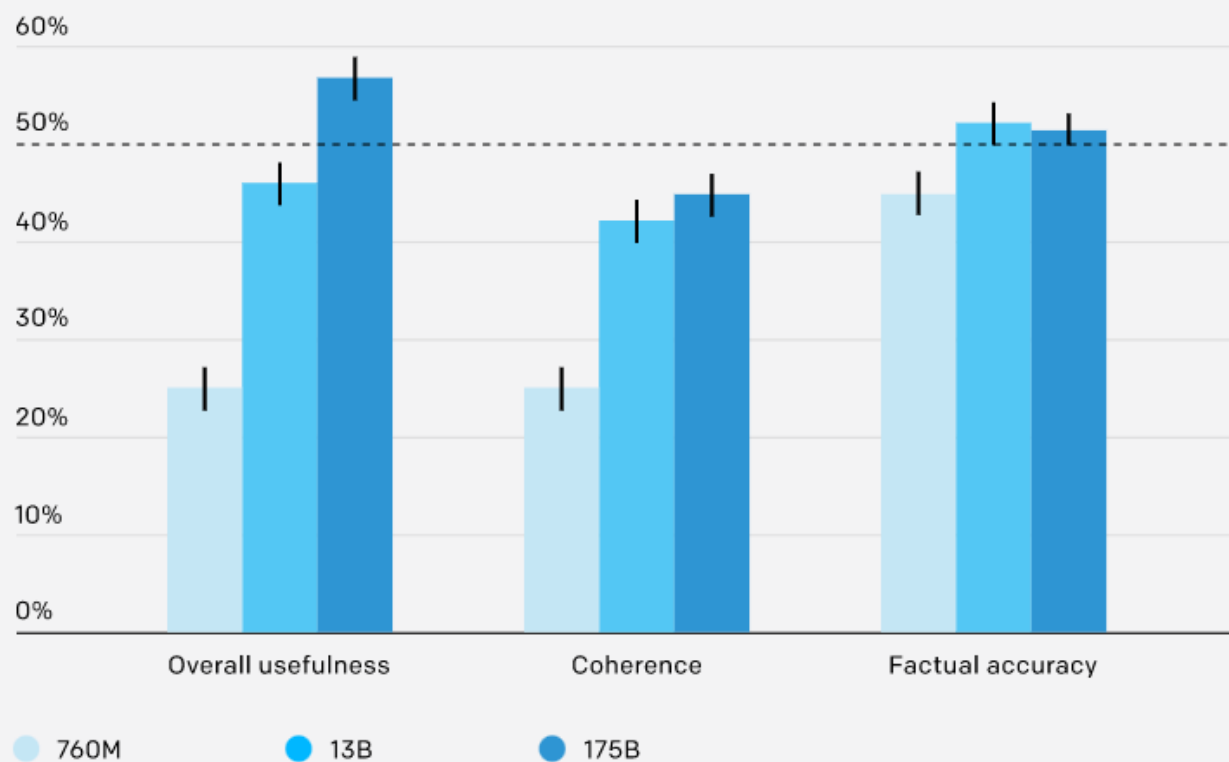
"j is better than k"

- Пример применения **reward-модели** для задачи выбора лучшего из двух разных вариантов пересказа статьи

11

ОПРОС ЛЮДЕЙ О КАЧЕСТВЕ ОТВЕТОВ МОДЕЛИ WEBGPT

Model answer preferred to demonstration answer



Разработчики из OpenAI опасаются, что GPT-3 может быть использована для спама, домогательств и дезинформации

Опасения эти оправданы, поскольку высокое качество работы модели в некоторых примерах практически не позволяет определить, кто именно и с какими заложенными намерениями высказал свое мнение по тому или иному вопросу: **человек** или **нейронная сеть**

СПИСОК ИСТОЧНИКОВ

- [«WebGPT: теперь GPT-3 умеет гуглить лучше тебя» — Яндекс.Кью \(yandex.ru\)](#)
- [GPT-3 в картинках: краткий обзор / Хабр \(habr.com\)](#)
- [GPT-3: самая продвинутая языковая модель на сегодняшний день \(sysblok.ru\)](#)
- [GPT-3 — Википедия \(wikipedia.org\)](#)
- [WebGPT: Improving the Factual Accuracy of Language Models through Web Browsing \(openai.com\)](#)
- [\[2112.09332\] WebGPT: Browser-assisted question-answering with human feedback \(arxiv.org\)](#)
- [\[2009.01325\] Learning to summarize from human feedback \(arxiv.org\)](#)
- [Третье поколение алгоритма OpenAI научилось выполнять текстовые задания по нескольким примерам \(nplus1.ru\)](#)
- [GPT-3 научили искать ответы в интернете и подкреплять их ссылками на источники \(nplus1.ru\)](#)
- [На что способна нейросеть GPT-3 | РБК Тренды \(rbc.ru\)](#)