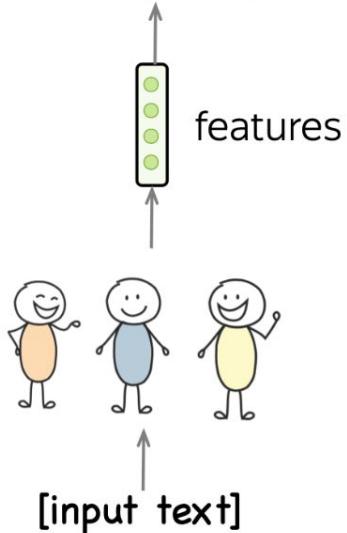


Lecture 5. Generative pretraining

Nikolay Karpachev
04.03.2024

The Evolutionary Journey in NLP (aka “What on Earth is Going On”)

Classify e.g.
sentiment



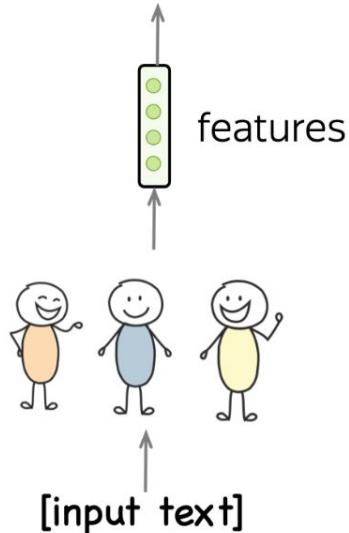
The Evolutionary Journey in NLP (aka “What on Earth is Going On”)

Different task –

another set of

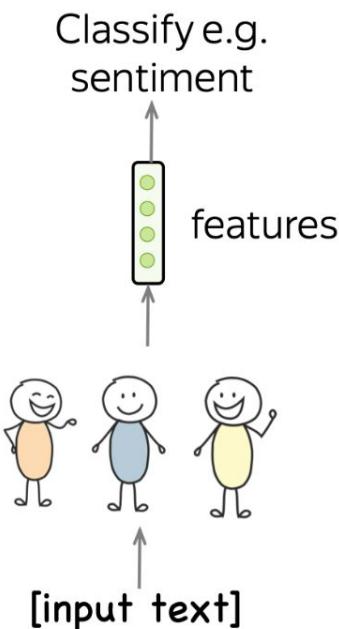
features (and people!)

Classify e.g.
sentiment

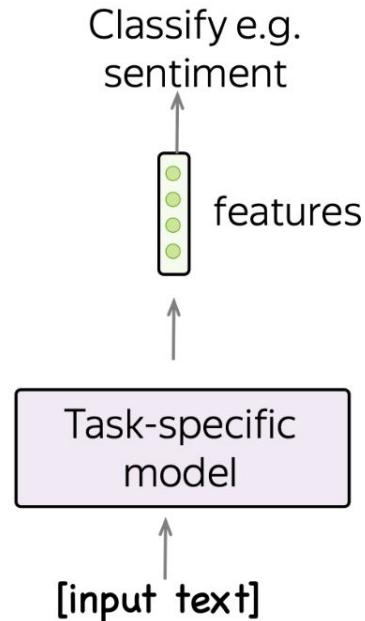


The Evolutionary Journey in NLP (aka “What on Earth is Going On”)

Different task –
another set of
features (and people!)

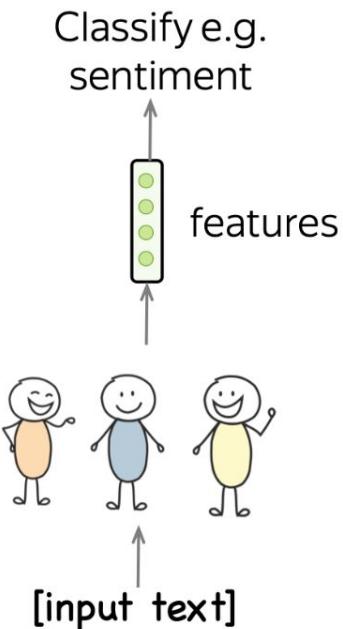


Different task –
another model

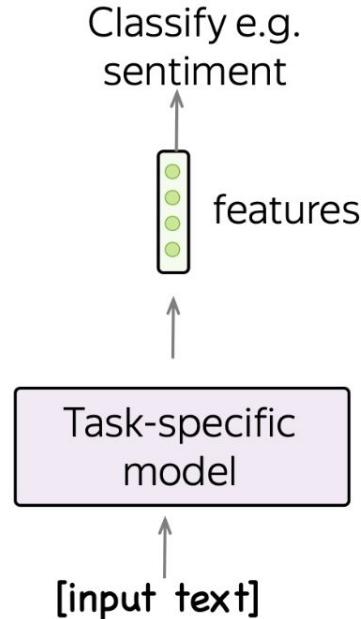


The Evolutionary Journey in NLP (aka “What on Earth is Going On”)

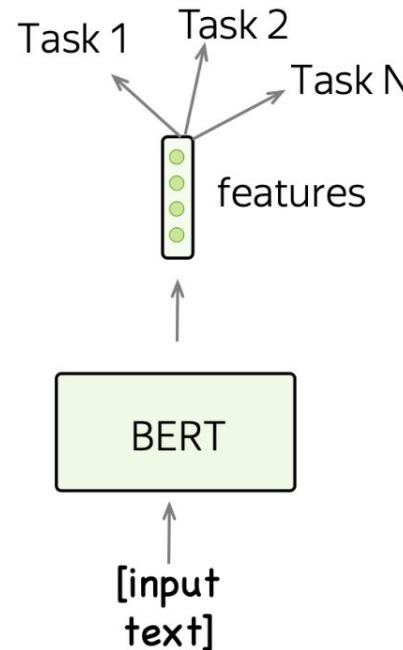
Different task –
another set of
features (and people!)



Different task –
another model

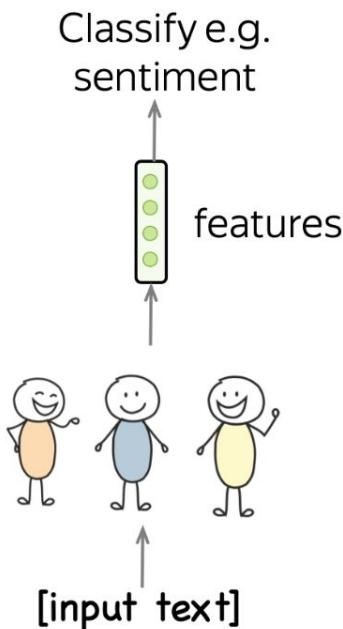


One model,
classify anything

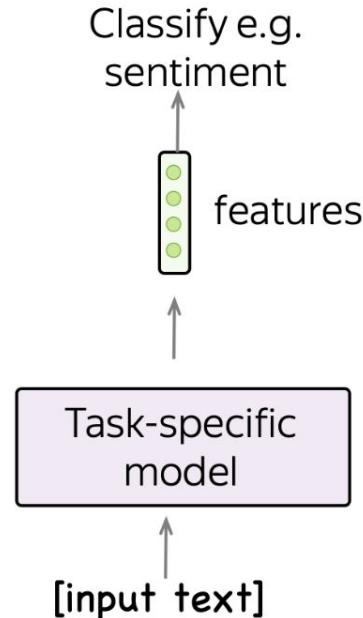


The Evolutionary Journey in NLP (aka “What on Earth is Going On”)

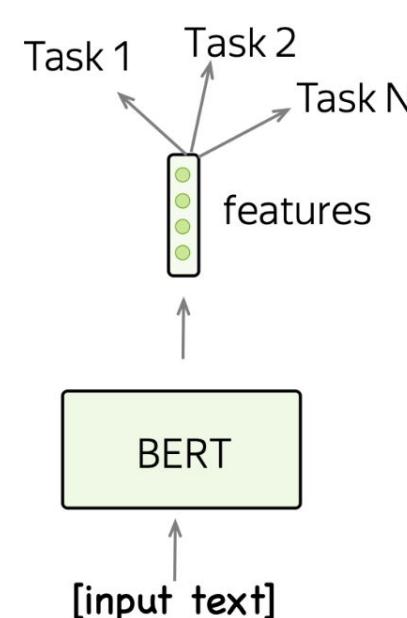
Different task –
another set of
features (and people!)



Different task –
another model



One model,
classify anything



Talk to the model

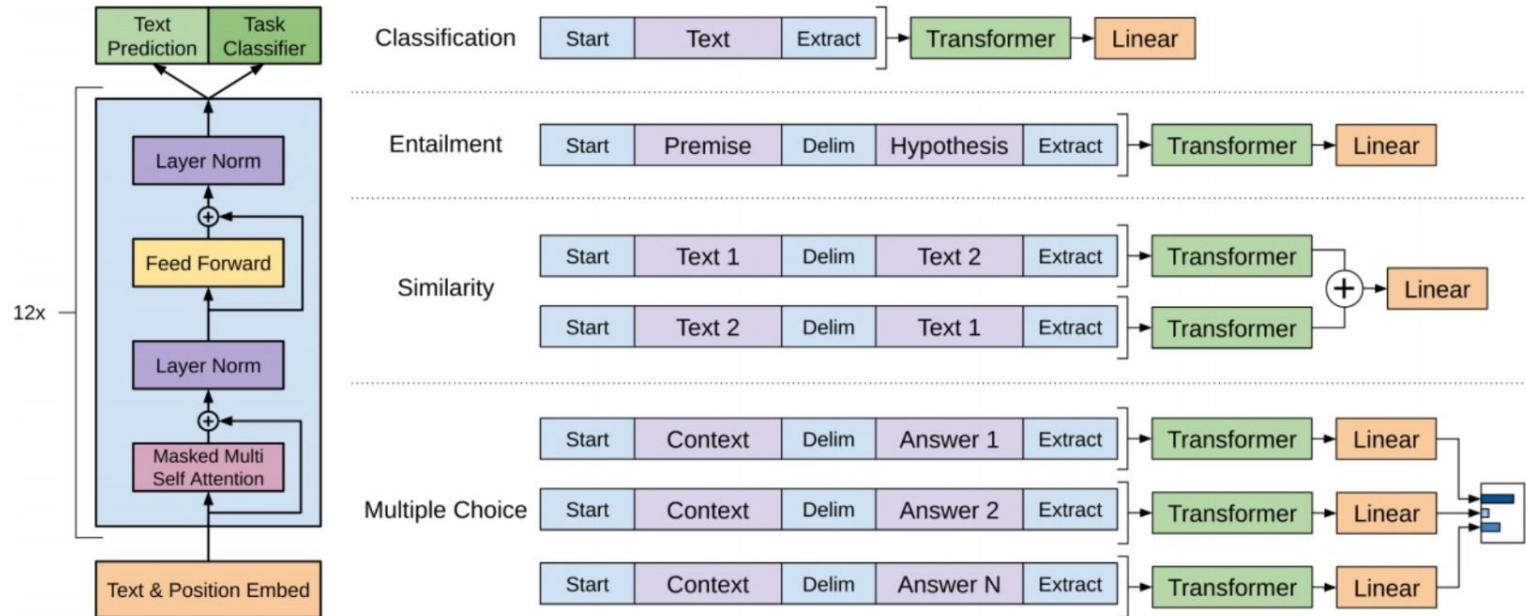
Input (prompt)

What is the sentiment
of the next sentence?
I love this movie!

Model output
positive

Fine-Tuning: Using GPT for Downstream Tasks

$$\text{Fine-tuning loss: } L = L_{xent} + \lambda \cdot L_{task}$$

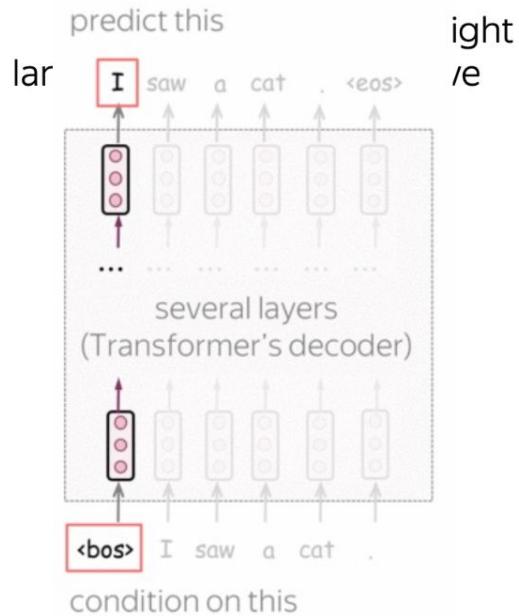


The figure is from the paper [Improving Language Understanding by Generative Pretraining](#)

GPT: Generative Pre-trained Transformer

Training

Transformer decoder with the standard I



Inference

GPT-1 Classification via

- Task-specific input transformations
- Supervised fine-tuning

GPT: Generative Pre-trained Transformer

	<u>GPT-1</u> (2018)	<u>GPT-2</u> (2019)
Number of parameters	117 million	1.5 billion
Training data	5 GB	40 GB

GPT-2: Examples of Generated Text

(CNN) Former President Donald Trump gave a January 6 speech near the White House in which he exhorted his supporters to march on the US Capitol to challenge Joe Biden's claim that Rei was better than Asuka. "Ayanami Rei was one of the worst characters in Evangelion, if not the worst period. I'm not saying that because she's Japanese. I'm saying she 's a total loser," Trump said . " Her face. She doesn't have a smile. Her eyes are all screwed up and her hair is the worst hair , I mean worse than Asuka 's , she looked like a monkey ." In his speech , Trump also said, " The reason Asuka was more successful than Rei was because she was never a ' normal girl ' in the first place , she was always a 'freak ' that got what she wanted. Asuka is the most popular character from the Eva series , and she's also the best character . Rei is very bad , she's not even close to Asuka's level, she is so bad it's ridiculous ." In a statement released Wednesday , Trump called the speech " not nice ." Read the full statement below . (Source : CNN) .

.. According to reports CNN (CNN's source) sister- , Trump site, spoke to a was is the speech " not nice." Iam (source) The interview bert with (CNN aired source) , the interviewer stated " it's not nice ." This was the second time the interview was aired . (source) CNN (source) also said , " that was a pretty un professional way to answer the question ." (source) CNN (source) then went on to say , " but what he said wasn't surprising considering the situation ." (source) (source) CNN also stated , " he seemed to get it , but he also seemed to be struggling to say what was most important ." (source) CNN's (source) also said that the interview was only done because they " felt like he needed to be on the record with what he said," (source) In addition , " that 's because the people who watched the interview asked for a full explanation and an apology." (source) CNN (source) also said the interview was only done because they " felt like he needed to be on the record with what he said , " (source) In addition , " that 's because the people who watched the interview asked for a full explanation and an apology." (source) (source) CNN (source) also said the interview was only done because they " felt like he needed to be on the record with what he said , " (source) In addition , " that's because the people who watched the interview asked for a full explanation and an apology." (source) (source) CNN (source) also said the interview was only done because they

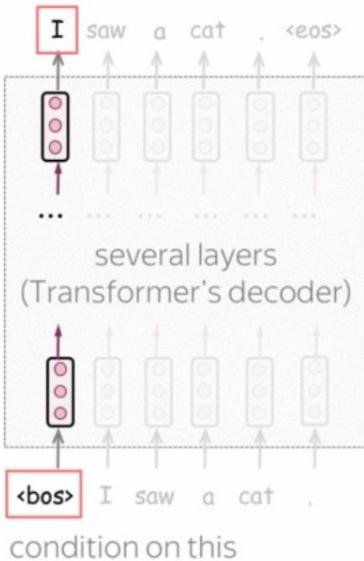
(highlighted text is generated by the model)

GPT: Generative Pre-trained Transformer

Training

Transformer decoder with the standard left-to-right language modeling objective

predict this



Inference

- GPT-1 Classification via
- Task-specific input transformations
 - Supervised fine-tuning
- GPT-2 Generation tasks via task-specific input transformations

For example, for text summarization simply add “**TL;DR**”:

<start> Generative pre-trained transformers (GPT) are ... TL;DR:

Text to be summarized

GPT: Generative Pre-trained Transformer

	<u>GPT-1</u> (2018)	<u>GPT-2</u> (2019)	<u>GPT-3</u> (2020)
Number of parameters	117 million	1.5 billion	175 billion
Training data	5 GB	40 GB	45 TB (i.e., 45 000 GB)

The entirety of English Wikipedia constitutes just 0.6% of GPT-3 training data

GPT-3: examples (<https://beta.openai.com/examples>)



TL;DR summarization

Transformation Generation

Summarize text by adding a 'tl;dr:' to the end of a text passage. It shows that the API understands how to perform a number of tasks with no instructions.

Prompt

A neutron star is the collapsed core of a massive supergiant star, which had a total mass of between 10 and 25 solar masses, possibly more if the star was especially metal-rich.[1] Neutron stars are the smallest and densest stellar objects, excluding black holes and hypothetical white holes, quark stars, and strange stars.[2] Neutron stars have a radius on the order of 10 kilometres (6.2 mi) and a mass of about 1.4 solar masses. [3] They result from the supernova explosion of a massive star, combined with gravitational collapse, that compresses the core past white dwarf star density to that of atomic nuclei.

Tl;dr

Sample response

A neutron star is the collapsed core of a massive supergiant star. These ultra-dense objects are incredibly fascinating due to their strange properties and their potential for phenomena such as extreme gravitational forces and a strong magnetic field.

GPT-3: examples (<https://beta.openai.com/examples>)



Factual answering

Answers Generation Conversation Classification

Guide the model towards factual answering by showing it how to respond to questions that fall outside its knowledge base. Using a '?' to indicate a response to words and phrases that it doesn't know provides a natural response that seems to work better than more abstract replies.

Prompt

Q: Who is Batman?

A: Batman is a fictional comic book character.

Q: What is torsalplexity?

A: ?

Q: What is Devz9?

A: ?

Q: Who is George Lucas?

A: George Lucas is American film director and producer famous for creating Star Wars.

Q: What is the capital of California?

A: Sacramento.

Q: What orbits the Earth?

A: The Moon.

Q: Who is Fred Rickerson?

A: ?

Q: What is an atom?

A: An atom is a tiny particle that makes up everything.

Q: Who is Alvan Muntz?

A: ?

GPT-3: examples (<https://beta.openai.com/examples>)

 English to other languages

Transformation Generation

Translates English text into French, Spanish and Japanese.

Prompt

Translate this into 1. French, 2. Spanish and 3. Japanese:

What rooms do you have available?

1.

Sample response

Quels sont les chambres disponibles?
2. ¿Cuáles son las habitaciones disponibles?
3. 何室がありますか?

GPT-3: examples (<https://beta.openai.com/examples>)



Ad from product description

Generation

Turn a product description into ad copy.

Prompt

Write a creative ad for the following product to run on Facebook aimed at parents:

Product: Learning Room is a virtual environment to help students from kindergarten to high school excel in school.

Sample response

Are you looking for a way to help your child excel in school? Look no further than Learning Room! Our virtual environment is designed to help students from kindergarten to high school learn and grow. With our help, your child will be able to get ahead in school and achieve their academic goals.

GPT-3: examples (<https://beta.openai.com/examples>)



Create study notes

Generation

Provide a topic and get study notes.

Prompt

What are 5 key points I should know when studying Ancient Rome?

Sample response

1. Rome was founded in 753 BC by Romulus and Remus.
2. Rome was ruled by an emperor.
3. Rome was a republic before it became an empire.
4. Julius Caesar was assassinated in 44 BC.
5. Augustus was the first emperor of Rome.

GPT-3: examples (<https://beta.openai.com/examples>)

Python to natural language

Code Translation

Explain a piece of Python code in human understandable language.

Prompt

```
# Python 3
def remove_common_prefix(x, prefix, ws_prefix):
    x["completion"] = x["completion"].str[len(prefix) :]
    if ws_prefix:
        # keep the single whitespace as prefix
        x["completion"] = " " + x["completion"]
    return x

# Explanation of what the code does

#
```

Sample response

The code above is a function that takes a dataframe and a prefix as input and returns a dataframe with the prefix removed from the completion column.

GPT-3: examples (<https://beta.openai.com/examples>)



Create analogies. Modified from a community prompt to require fewer examples.

Prompt

Create an analogy for this phrase:

Questions are arrows in that:

Sample response

Questions are arrows in that they can be used to point out things that need to be fixed.

GPT-3: examples

Describe a layout.

Just describe any layout you want, and it'll try to render below!

Generate

GPT-3 and in-context learning



Micro horror story creator

Transformation Generation Translation

Creates two to three sentence short horror stories from a topic input.

Prompt

Topic: Breakfast

Two-Sentence Horror Story: He always stops crying when I pour the milk on his cereal. I just have to remember not to let him see his face on the carton.

Topic: Wind

Two-Sentence Horror Story:

Sample response

I was lying in bed, trying to get to sleep, when I heard the wind howling outside my window. It sounded like something was trying to get in.

Source: <https://beta.openai.com/examples>

GPT-3 Gives Correct Medical Advice

Question: A 10 year old boy presents with recurrent episodes of dyspnoea, cough and wheeze triggered by cold air and allergens. After performing tests he is diagnosed with an obstructive disease and given medication. Which receptor is the medication most likely to work on:

- A) muscarinic receptor
- B) G-protein coupled receptor
- C) Era
- D) Erb
- E) a-1 receptor

Correct answer is **B**

Explanation: The patient is suffering from asthma, a disease of the airways. The drugs used to treat asthma are bronchodilators. The bronchodilators act on the beta-2 receptors. The beta-2 receptors are G-protein coupled receptors

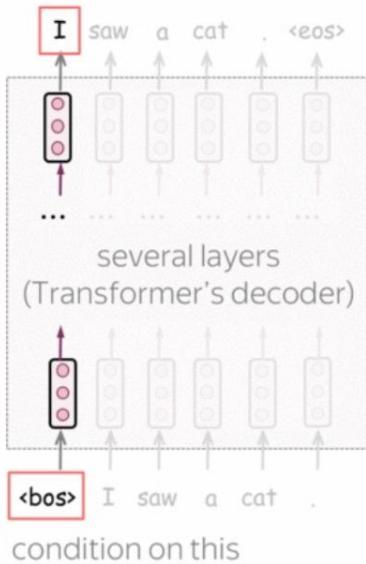
(bold text is generated by the model)

GPT: Generative Pre-trained Transformer

Training

Transformer decoder with the standard left-to-right language modeling objective

predict this



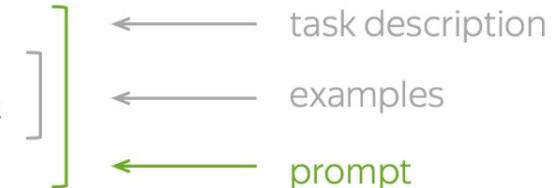
Inference

- GPT-1 Classification via
- Task-specific input transformations
 - Supervised fine-tuning
- GPT-2 Generation tasks via task-specific input transformations

For example, for text summarization simply add “TL;DR”:

- GPT-3 Complex generation and reasoning tasks via
- in-context learning: prompt with task description and a few demonstrations

Translate English to Spanish:
a black cat -> un gato negro
I am hungry -> tengo hambre
a cup of tea ->



condition on this

... Is it really learning from context?

p language is a language where every word is appended with p.

Q: you are a good boy.

A: yousp arep ap goodp boyp.

Q: Where is the spoon?

A: Wherep isp thep spoopp?

Q: The cat is a small animal.

A: Thep catp isp ap smallp animalp.

Q: Can he jump?

A: Canp hep jumpp?

Q: The forest was big. There were lots of animals and trees. But no humans.

A: Thep forestp wasp bigp. Therewp werew lotsp ofp animalsp andp treeesp. Butp nopp humansp.

Q: Can the lady call the guy? There is an emergency.

A: Canp thep ladyp callp thep guyp? Therewp isp anp emergennccyp.

Q: Thank the heavens. Peter pan went to his pop.

A: Thankp thep heavensp. Peterp panp wentp topp hisp popp.

... Is it really learning from context?

p language is a language where only animal names are appended with p. Cat, dog, tiger etc are some animal names.

Q: you are a good boy.

A: you are a good boy.

Q: Where is the spoon?

A: Where is the spoon?

Q: The cat is a small animal.

A: The catp is a small animal.

Q: Can he jump over the dog?

A: Can he jump over the dogp?

Q: What is the time?

A: What is the time?

Q: The tiger is dangerous?

A: The tigerp is dangerous?

Q: The boy is going to school.

A: The boyp is going to school.

Modern Open-Source Language Models

- LLaMA-2 – good first choice for english tasks (and some other languages)
 - 7B, 13B and 70B parameters
- Falcon-180B – somewhat better, but a lot larger
 - 7B, 40B and 180B parameters
- BLOOM-176B – if the first two models don't speak your language
 - 560M – 7B and 176B parameters



```
import transformers
model = transformers.LlamaForCausalLM.from_pretrained("meta-llama/Llama-2-70b-chat-hf")
tokenizer = transformers.LlamaTokenizer.from_pretrained("meta-llama/Llama-2-70b-chat-hf")
```

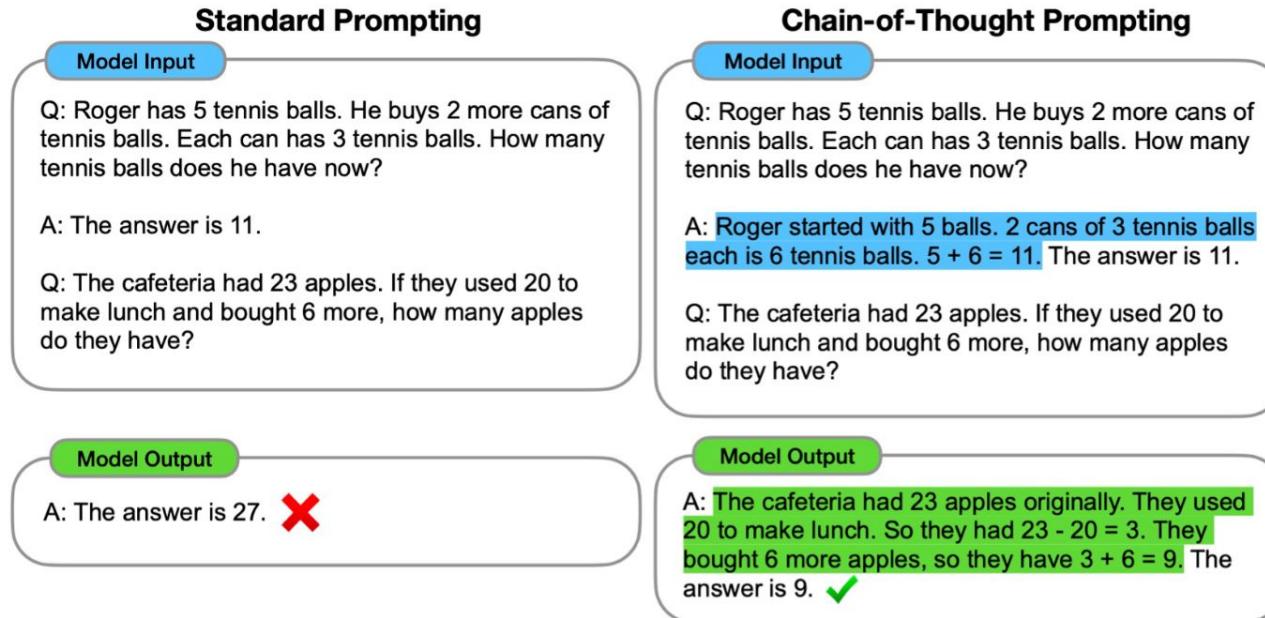
The image shows a screenshot of a Jupyter Notebook cell. The cell contains Python code for initializing a LLaMA-2 model and tokenizer. The code uses the `transformers` library. The model is initialized with the path `"meta-llama/Llama-2-70b-chat-hf"`, and the tokenizer is initialized with the same path. The cell has a play button icon on the left, indicating it can be run. To the right of the code is a toolbar with various icons for cell operations like running, saving, and deleting.

The Prompting Paradigm

No additional training! Just give a task with an example – prompt.
This is fed as prefix before model generation.

	few-shot		
	zero-shot	one-shot	two-shot
Input (prompt)	Review: I love this movie! Sentiment:	Review: This movie sucks. Sentiment: negative	Review: This movie sucks. Sentiment: negative
	Review: I love this movie! Sentiment:	Review: This was cool! Sentiment: positive	Review: I love this movie! Sentiment:
Model output	positive	positive	positive

Chain of Thought



When asked to “think”, the model gives the right answer

Chain of Thought

CSQA (commonsense)

Q: Sammy wanted to go to where the people were. Where might he go?
Options: (a) race track (b) populated areas
(c) desert (d) apartment (e) roadblock

A: The answer must be a place with a lot of people. Race tracks, desert, apartments, and roadblocks don't have a lot of people, but populated areas do. So the answer is (b).

Coin Flip (state tracking)

Q: A coin is heads up. Maybelle flips the coin. Shalonda does not flip the coin. Is the coin still heads up?

A: The coin was flipped by Maybelle. So the coin was flipped 1 time, which is an odd number. The coin started heads up, so after an odd number of flips, it will be tails up. So the answer is no.

Sports Understanding

Q: Is the following sentence plausible? "Joao Moutinho caught the screen pass in the NFC championship."

A: Joao Moutinho is a soccer player. The NFC championship is part of American football, not soccer. So the answer is no.

Self-Consistency

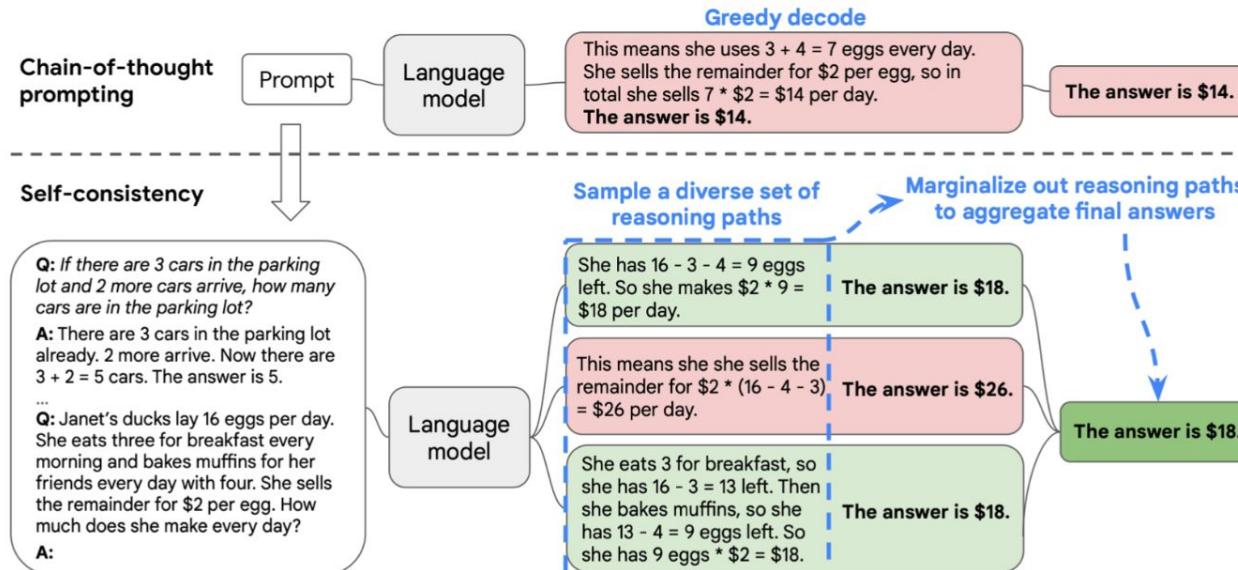


Figure 1: The self-consistency method contains three steps: (1) prompt a language model using chain-of-thought (CoT) prompting; (2) replace the “greedy decode” in CoT prompting by sampling from the language model’s decoder to generate a diverse set of reasoning paths; and (3) marginalize out the reasoning paths and aggregate by choosing the most consistent answer in the final answer set.

Self-Consistency

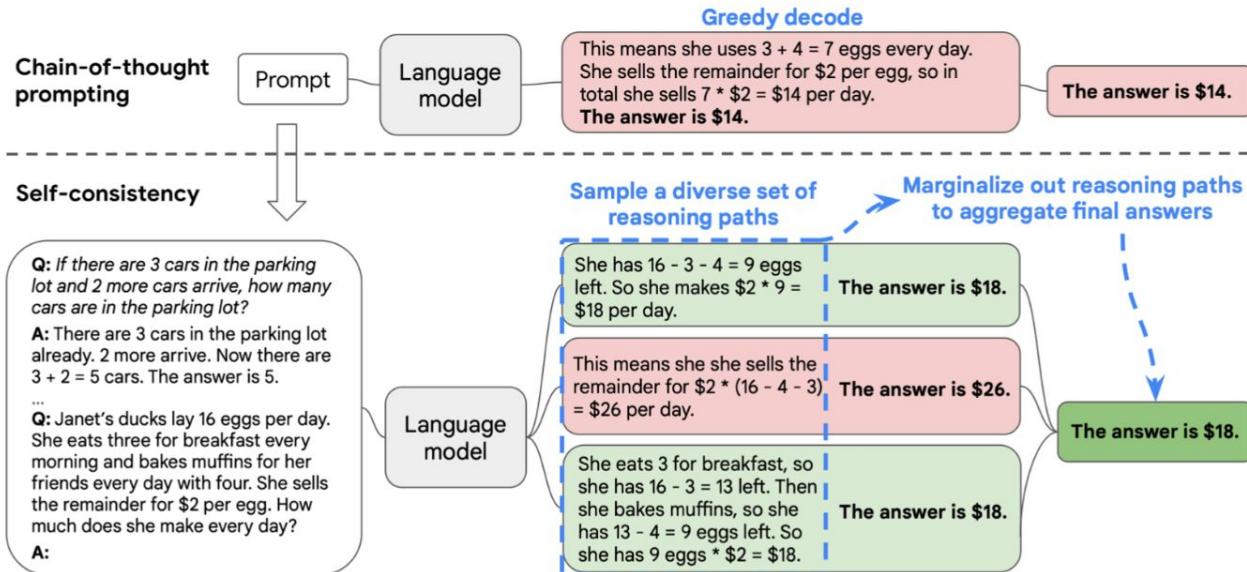
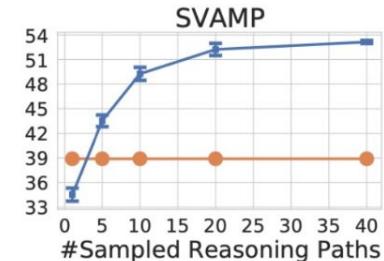
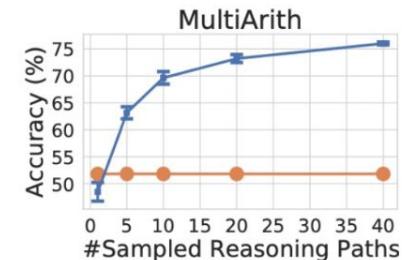


Figure 1: The self-consistency method contains three steps: (1) prompt a language model using chain-of-thought (CoT) prompting; (2) replace the “greedy decode” in CoT prompting by sampling from the language model’s decoder to generate a diverse set of reasoning paths; and (3) marginalize out the reasoning paths and aggregate by choosing the most consistent answer in the final answer set.



Acknowledgements

This lecture uses slides from awesome [nlp_course](#), made by YSDA team

Questions?