

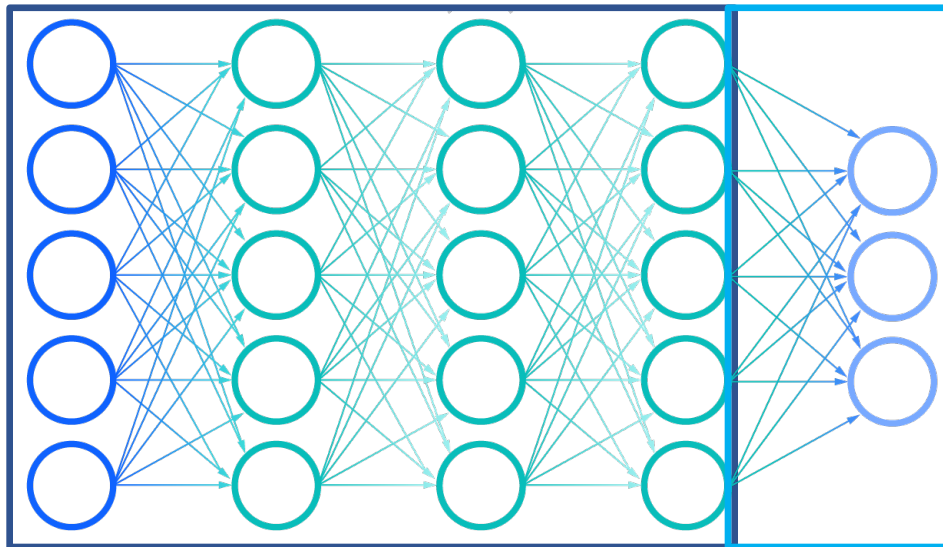
NLP: Prompting

Philippe Schläpfer

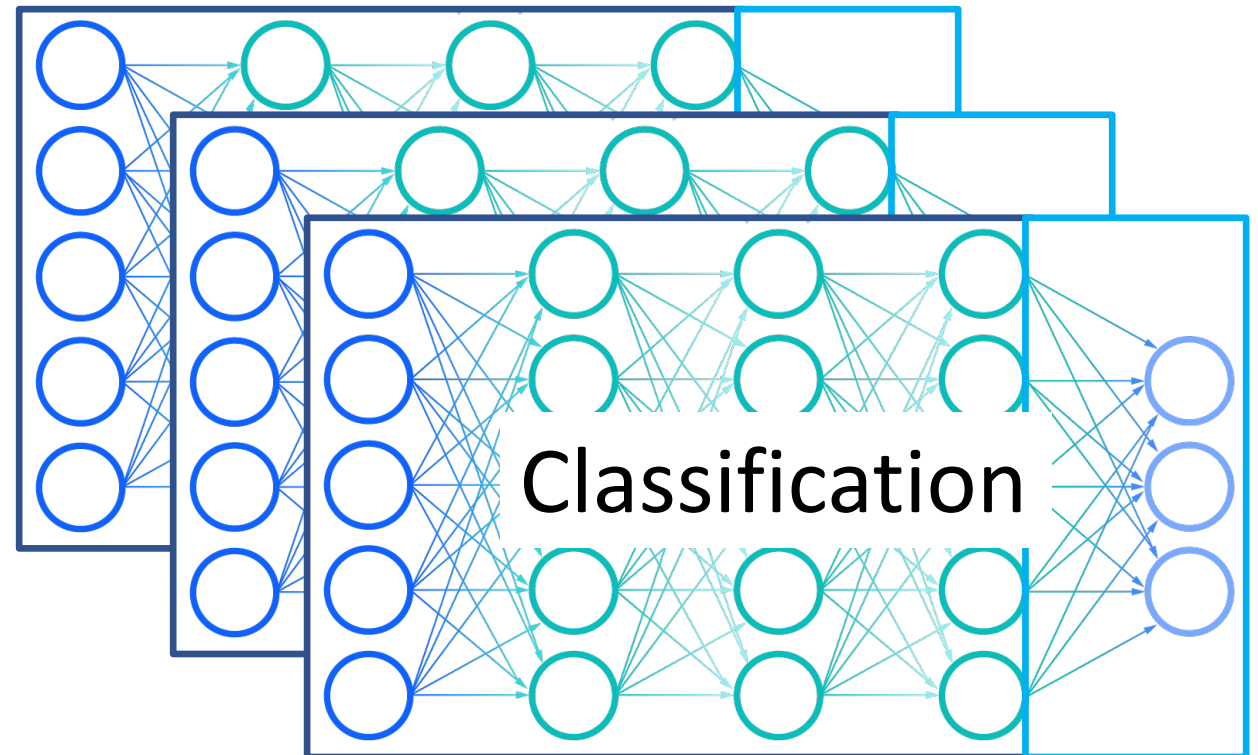
26.04.2022

Pre-trained Language Models

- Large LMs capture the language very good!
- Before prompting was a thing: Fine-tuning
 - Trying to model $P(y|x; \Theta)$



GPT-3: 175 billion parameters

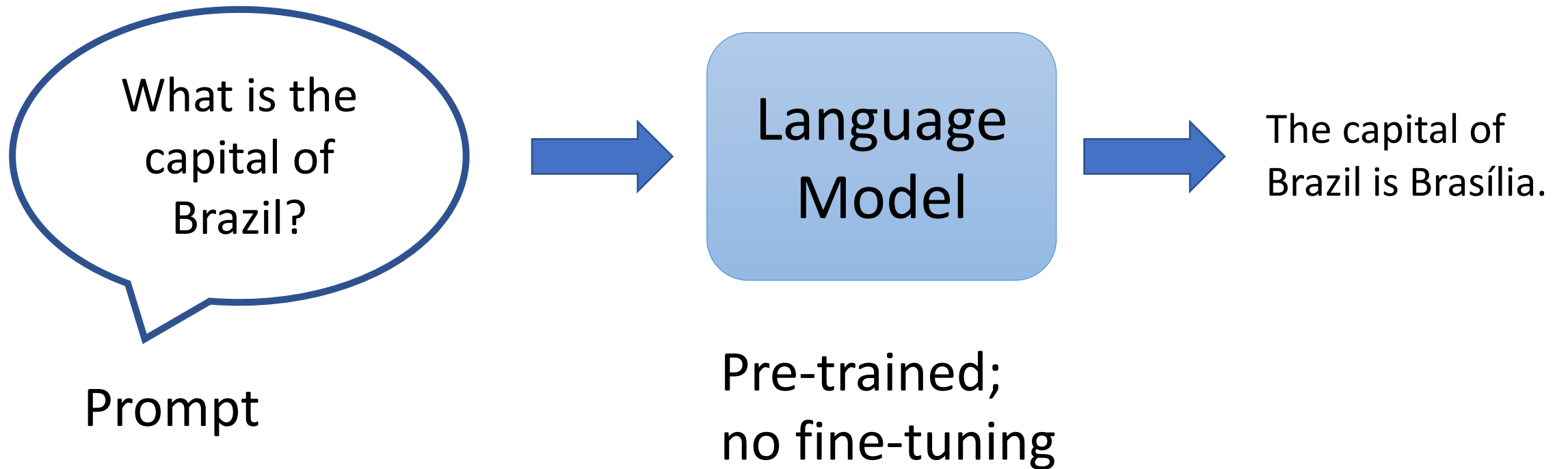


Prompts

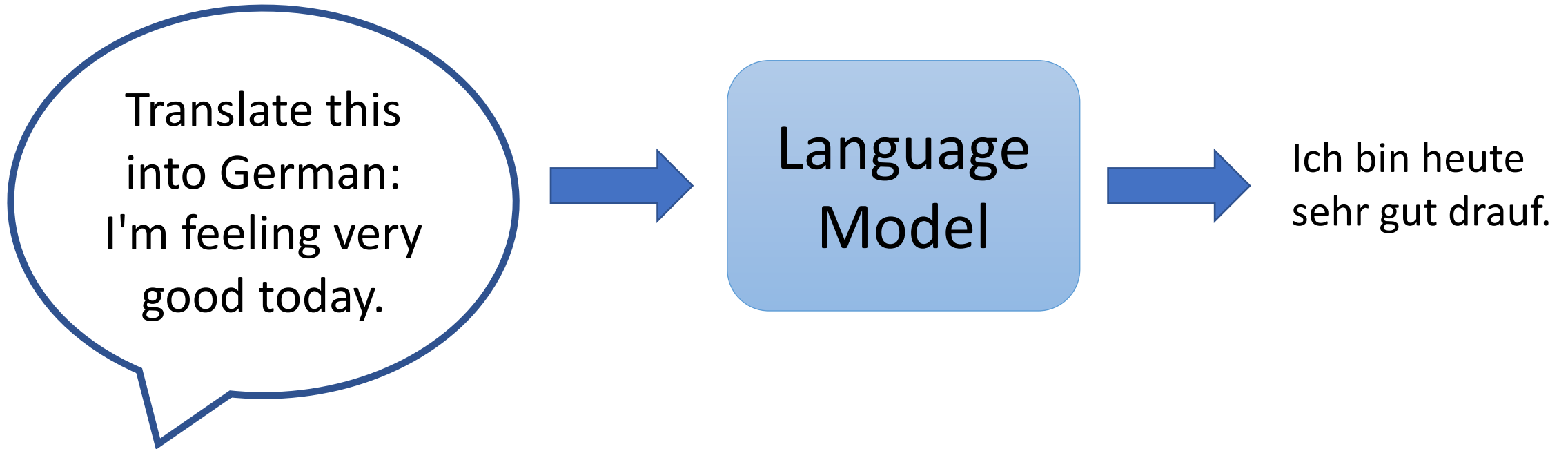
Trigger language model to give an output

Not to confuse with **Prompting**: Whole area of research

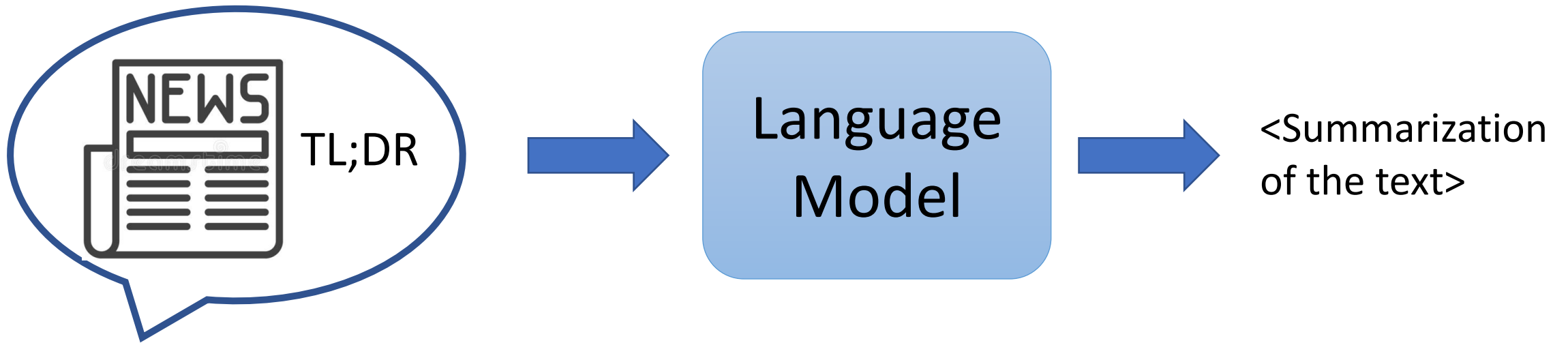
Example: Factual Knowledge



Example: Translation



Example: Summarization



Prompting

Model $P(x; \Theta)$ directly \leftrightarrow fine-tuning: $P(y|x; \Theta)$

GPT-3 Demo

Why should we do research in this area?

- No fine-tuning
- Small datasets
- No retraining^[1]

[1] Madaan, Tandon, Clark, Yang: "Memory-assisted prompt editing to improve GPT-3 after deployment". arXiv:2201.06009 (2022)

Cloze Tasks

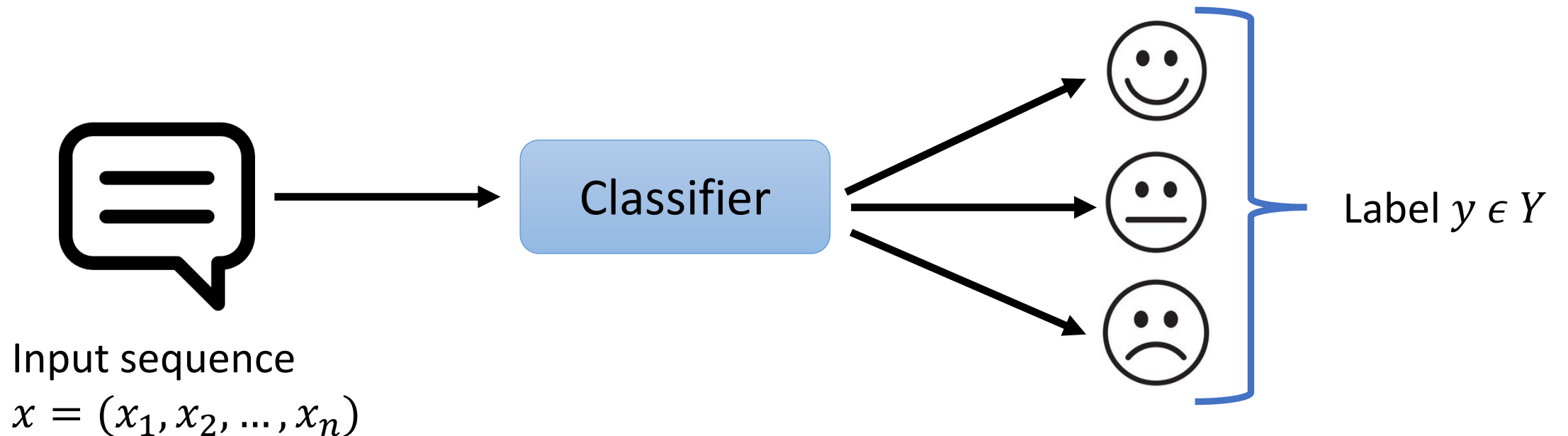
Today, I went to the _____ and bought some milk and eggs.

In NLP:

Today, I went to the [MASK] and bought some milk and eggs.

Example task: Text classification

Given a text, classify it into one of the given classes



Prompting Pipeline

Step 1: Prompting function

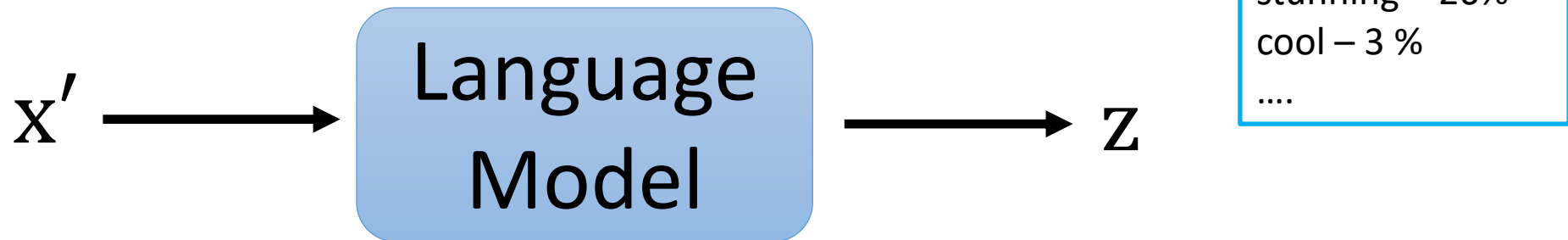


It's one of the best movies
I've ever seen.

It was [MASK]. It's one of the
best movies I've ever seen.

Prompting Pipeline

Step 2: Language Model



It was [MASK]. It's one of the best movies I've ever seen.

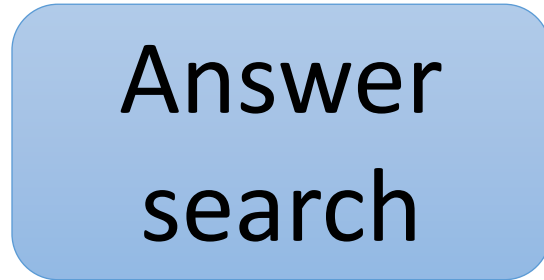
It was [MASK]. It's one of the best movies I've ever seen.

Prompting Pipeline

Step 3: Answer search

amazing – 62%
stunning – 26%
cool – 3 %
....

Z



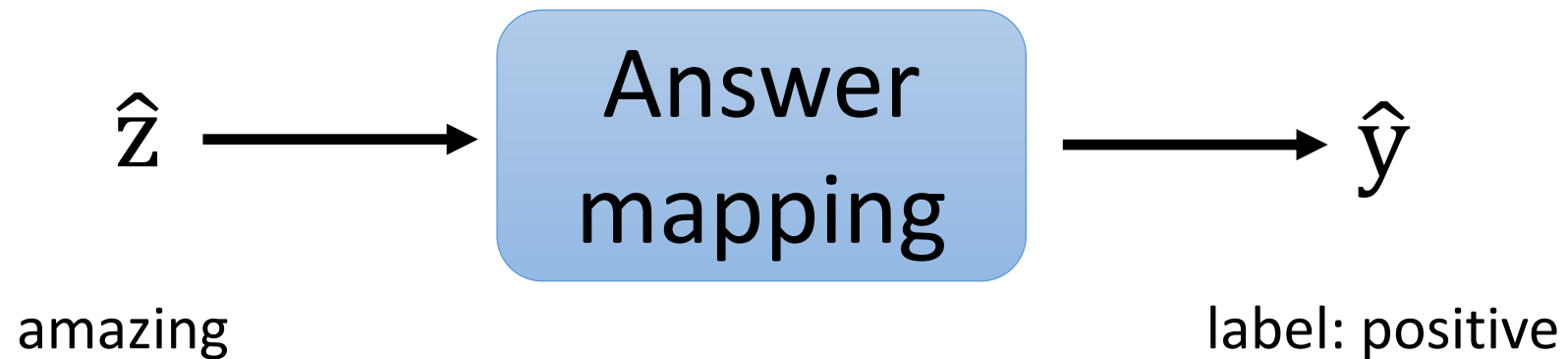
\hat{Z}

amazing

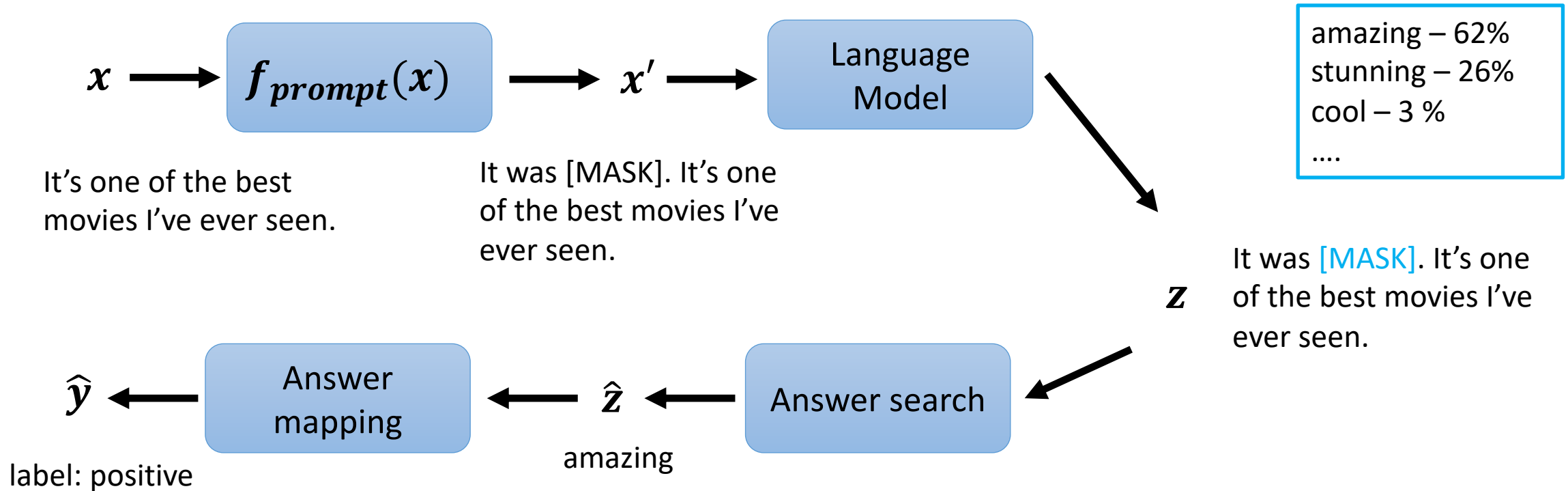
It was [MASK]. It's one of the best movies I've ever seen.

Prompting Pipeline

Step 4: Answer mapping



Prompting for Text Classification



Prompt Engineering

- Goal: Create prompting function $f_{prompt}(\mathbf{x})$ that gives best performance
 - Obama is a _____ by profession.
 - Obama worked as a _____.

Prompt Engineering: Two types of prompts

- Prefix prompts: e.g. Translation

Translate this into 1. German, 2. Italian and 3. Japanese:

Should we meet at 10 o'clock?

- Cloze prompts: It's a [MASK] movie to watch and [MASK] painful to watch.

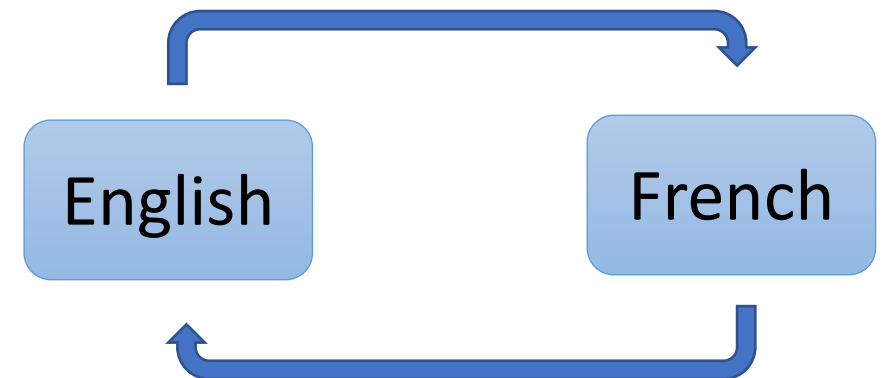
Prompt Engineering: Automatic template learning

- Discrete prompts
 - Paraphrasing (back translation, replacement of phrases from a thesaurus, ...)

Prompt: "x shares a border with y"

"x has a common border with y"

"x adjoins y"



Prompt Engineering: Automatic template learning

- Drawbacks of the prompts we've seen so far?
 - Prompts are in natural language (discrete optimization is challenging)
 - Template is parameterized by the parameters of the pre-trained LMs
- Continuous prompts: Perform prompting in the embedding space of the LM

Answer Engineering

I really enjoyed this movie.
The movie was _____.

∈

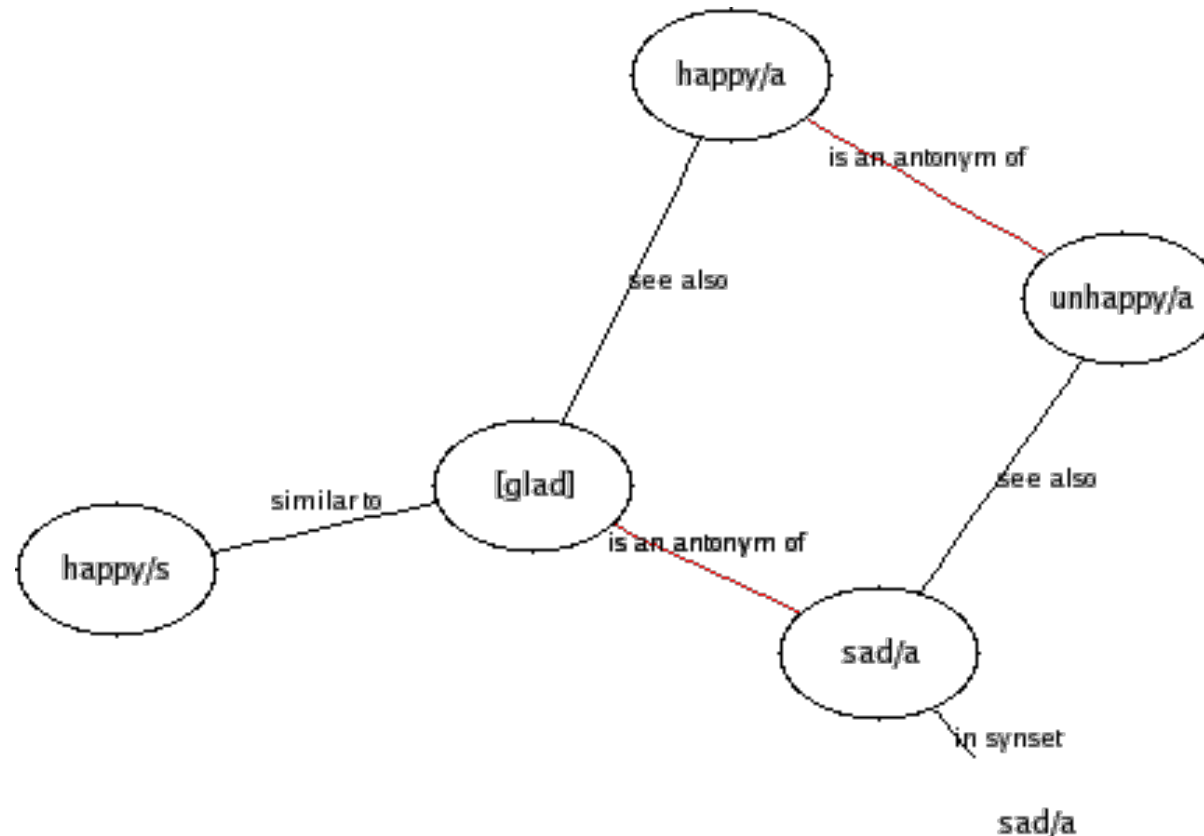
good, amazing,
phenomenal,
awesome, ...



- Straightforward: directly map answer z to the final output y
 - More sophisticated: Have a mapping function from words to labels
- Verbalizer

How to map the output of the LM to the labels?

Incorporate Knowledge Base into verbalizer (e.g. WordNet)



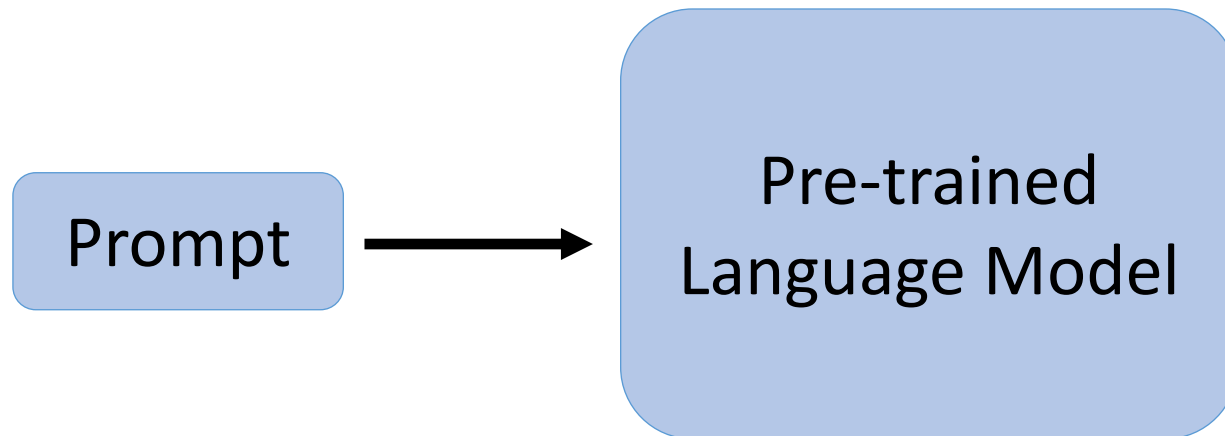
Prompting for text classification: results

Dataset	Label	Label Words
AG's News	POLITICS	politics, government, diplomatic, law, aristotle, diplomatical, governance ...
	SPORTS	sports, athletics, gymnastics, sportsman, competition, cycling, soccer ...
IMDB	NEGATIVE	abysmal, adverse, alarming, angry, annoy, anxious, apathy, appalling ...
	POSITIVE	absolutely, accepted, acclaimed, accomplish, accomplishment ...

Method	AG News	DBPedia	Yahoo	Amazon	IMDB
Label-only	75.1	66.6	45.4	80.2	86.4
Label words	84.8	82.2	61.6	92.8	91.6
SOTA	95.55	99.38	77.62	97.37	97.4

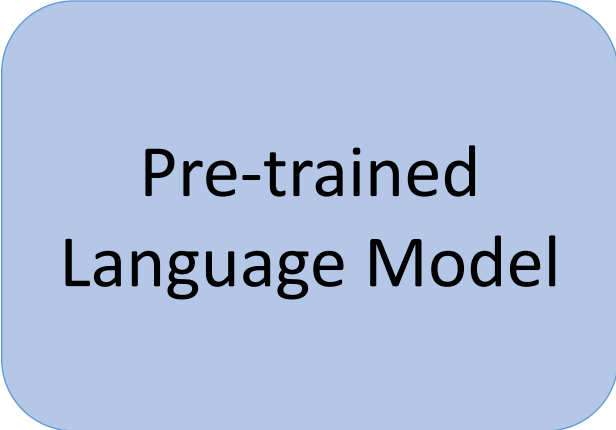
Training Strategies

- Promptless Fine-tuning
- Tuning-free Prompting
- Fixed-LM Prompt Tuning
- Fixed-prompt LM Tuning
- Prompt+LM Tuning



Training Strategies

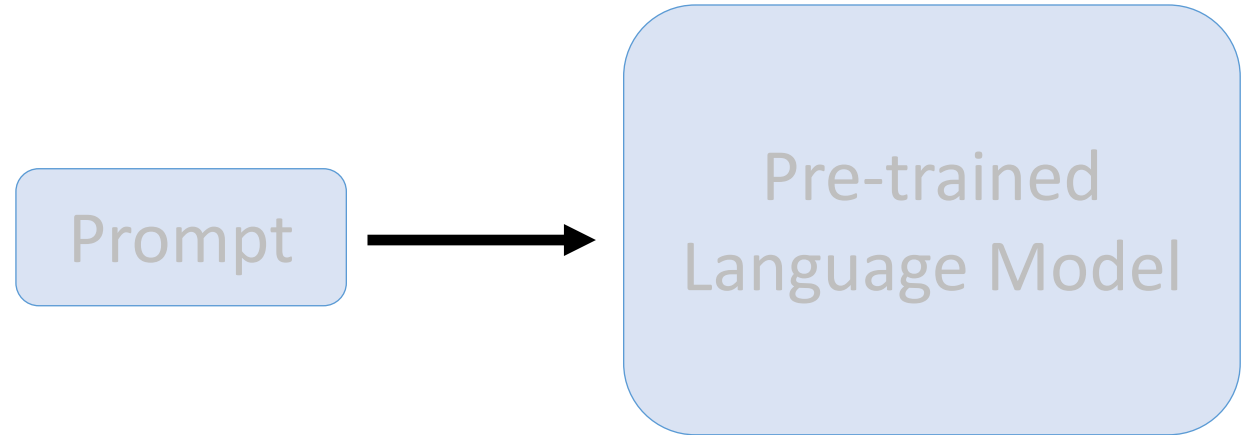
- Promptless Fine-tuning
- Tuning-free Prompting
- Fixed-LM Prompt Tuning
- Fixed-prompt LM Tuning
- Prompt+LM Tuning



Pre-trained
Language Model

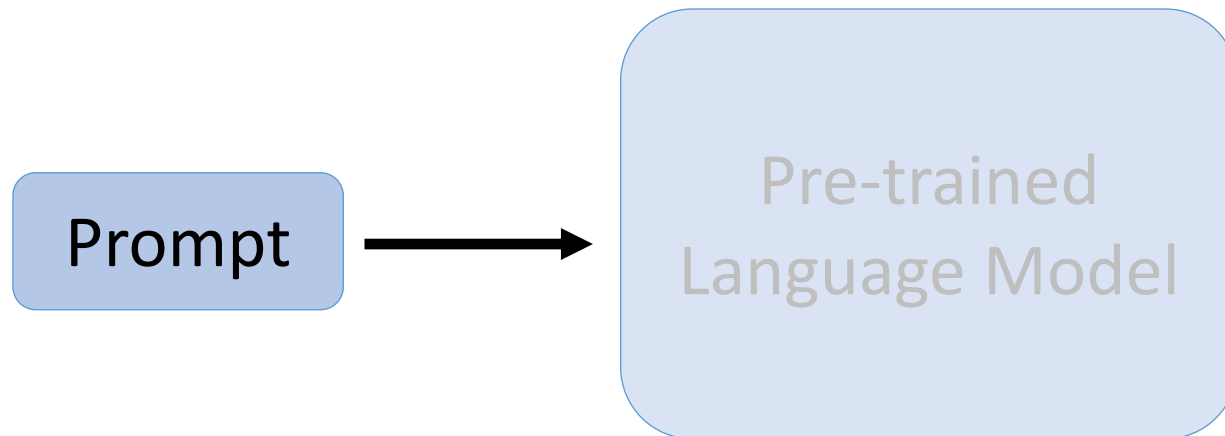
Training Strategies

- Promptless Fine-tuning
- **Tuning-free Prompting**
- Fixed-LM Prompt Tuning
- Fixed-prompt LM Tuning
- Prompt+LM Tuning



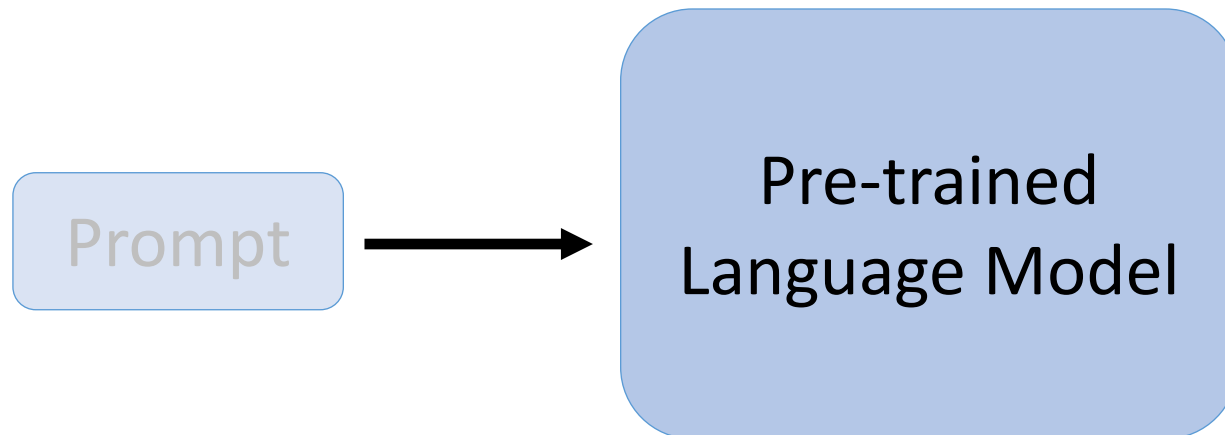
Training Strategies

- Promptless Fine-tuning
- Tuning-free Prompting
- **Fixed-LM Prompt Tuning**
- Fixed-prompt LM Tuning
- Prompt+LM Tuning



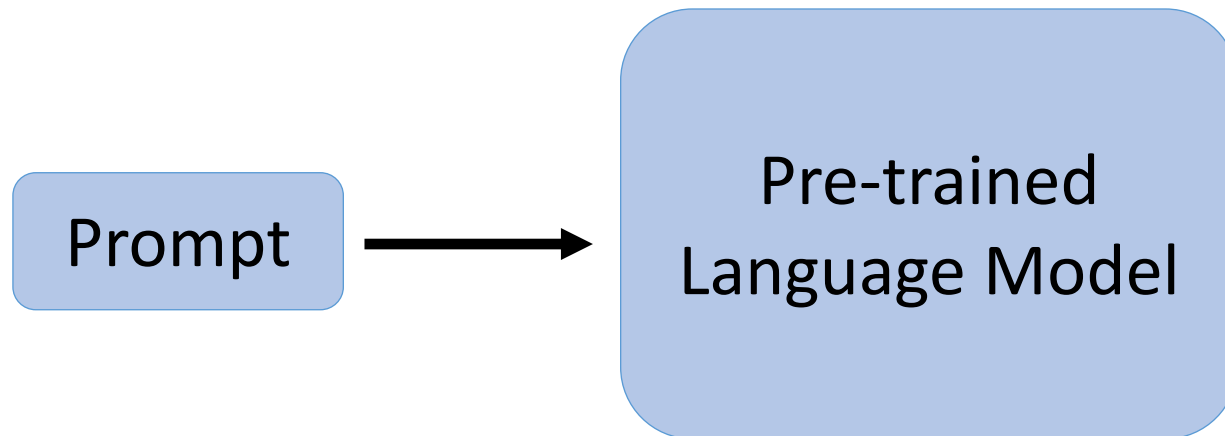
Training Strategies

- Promptless Fine-tuning
- Tuning-free Prompting
- Fixed-LM Prompt Tuning
- **Fixed-prompt LM Tuning**
- Prompt+LM Tuning



Training Strategies

- Promptless Fine-tuning
- Tuning-free Prompting
- Fixed-LM Prompt Tuning
- Fixed-prompt LM Tuning
- **Prompt+LM Tuning**



Additional Task: Question Answering

- Question in natural language
- Different forms (extractive, multiple choice, free form)

Passage Sentence

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.

Question

What causes precipitation to fall?

Answer Candidate

gravity

Question Answering

- Question + \n + additional information + \n + candidate answers

Passage Sentence

In meteorology, precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.

Question

What causes precipitation to fall?

Answer Candidate

gravity



What causes precipitation to fall?
\n
In meteorology precipitation is any product of the condensation of atmospheric water vapor that falls under gravity.

Summarization

- <...some text...> + TL;DR
- Summarize this text: <...some text...>
- Text: [X] Summary: [Z]
- [X] In summary, [Z]

Challenges / Open Questions

- New field of research
- Structured information (trees, graphs, etc.)
- Theoretical guarantees are scarce
- Still need labeled data for evaluation

References

- Brown et al. : " Language Models are Few-Shot Learners". arXiv:2005.14165v4 (2020)
- Madaan, Tandon, Clark, Yang: "Memory-assisted prompt editing to improve GPT-3 after deployment". arXiv:2201.06009 (2022)
- Hu, Ding, Wang, Liu, Wang, Li, Wu, Sun: "Knowledgeable Prompt-tuning: Incorporating Knowledge into Prompt Verbalizer for Text Classification". arXiv:2108.02035v2 (2022)
- Liu et al.: "Pre-train, Prompt, and Predict: A Systematic Survey of Prompting Methods in Natural Language Processing; Chapter 7". arXiv:2107.13586v1 (2021)
- Khashabi et al.: "UNIFIEDQA: Crossing Format Boundaries with a Single QA System". arXiv:2005.00700v3 (2020)

Thanks!

Questions?