UiT The Arctic University of Norway

BIO-AI LAB | ARCTIC LLM WORKSHOP 2023

Large Language Models

Day 2 - Session 1

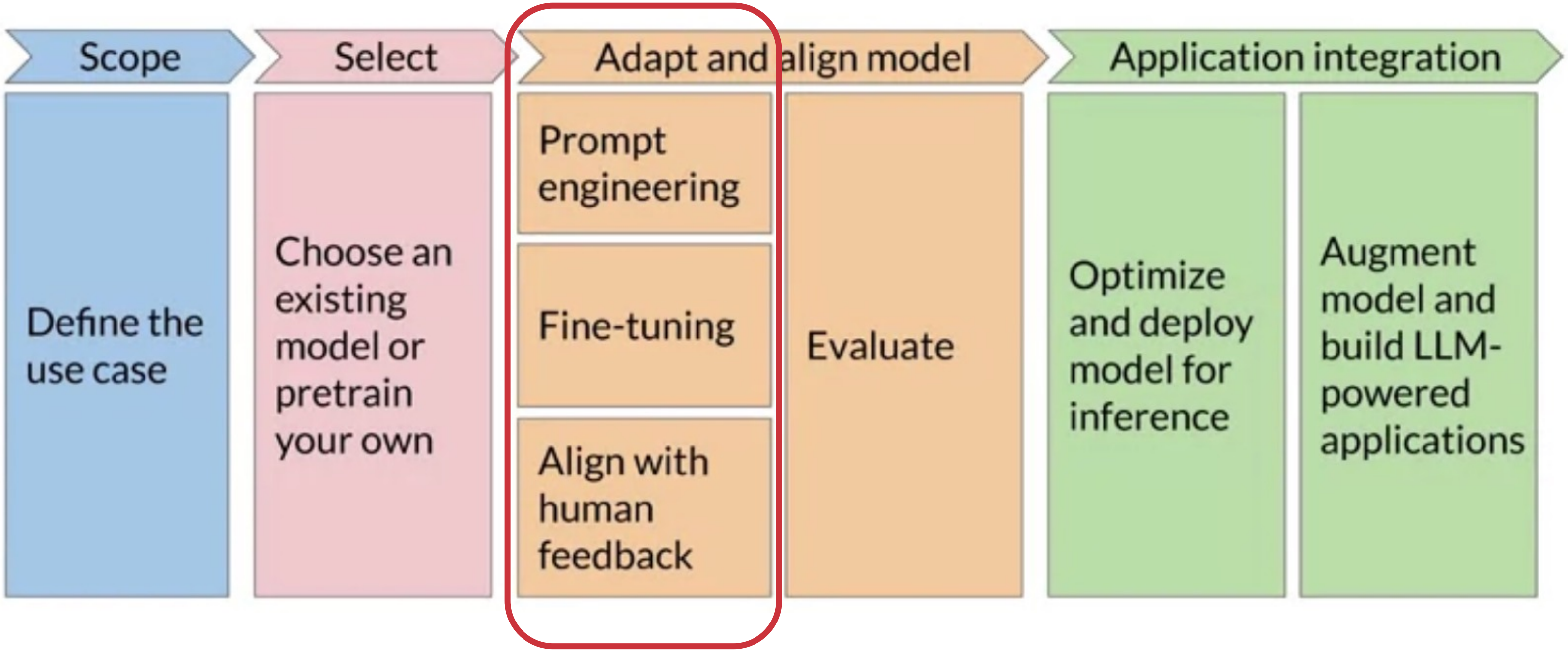
In-context Learning, Finetuning and RLHF in LLMs

Rohit Agarwal

rohit.agarwal@uit.no

28. Oct 2023

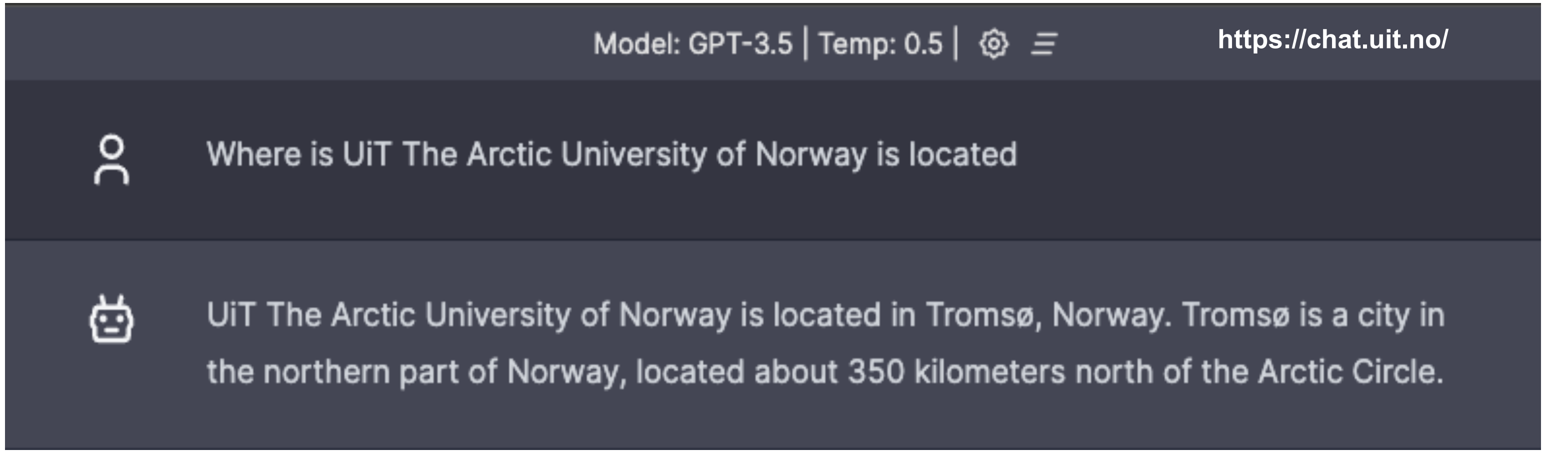
LLM: Lifecycle



Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

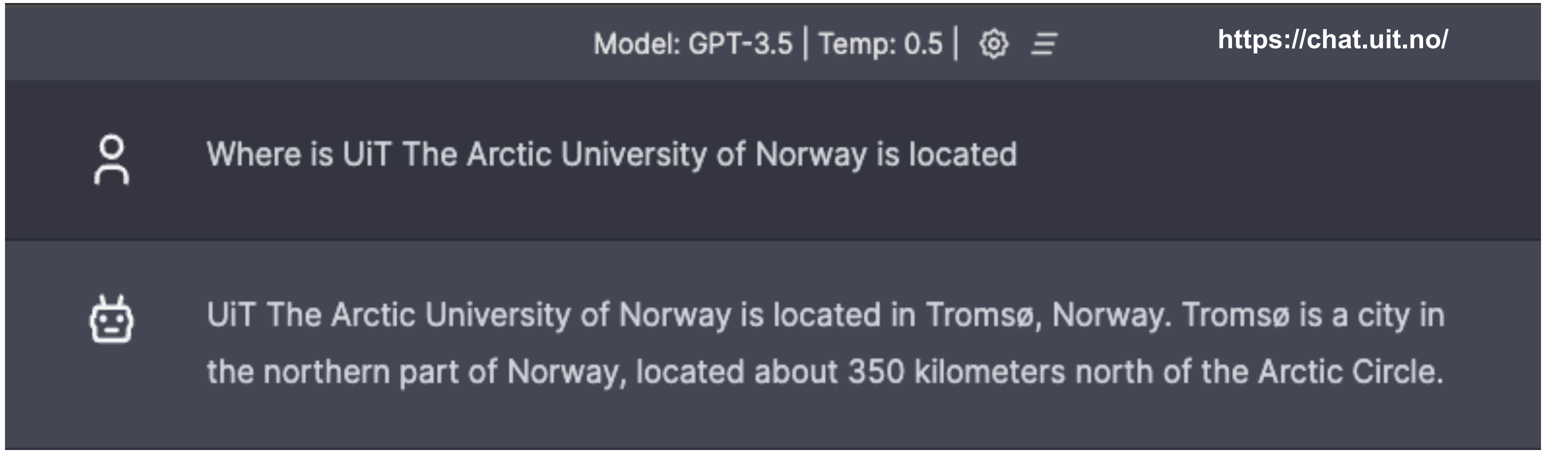
LLM: Prompts and Completion

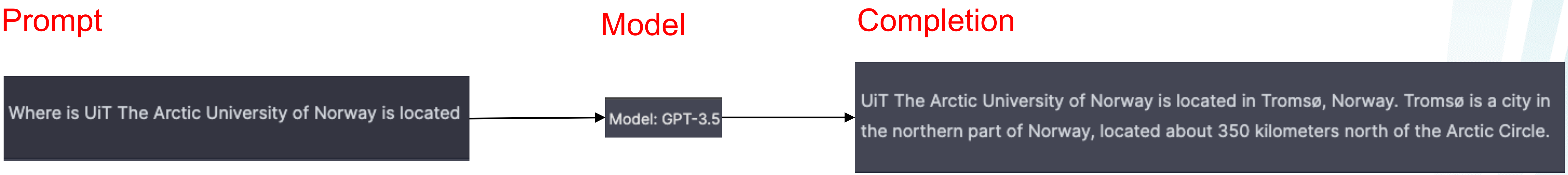
LLMs takes a query and generates an output.



LLM: Prompts and Completion

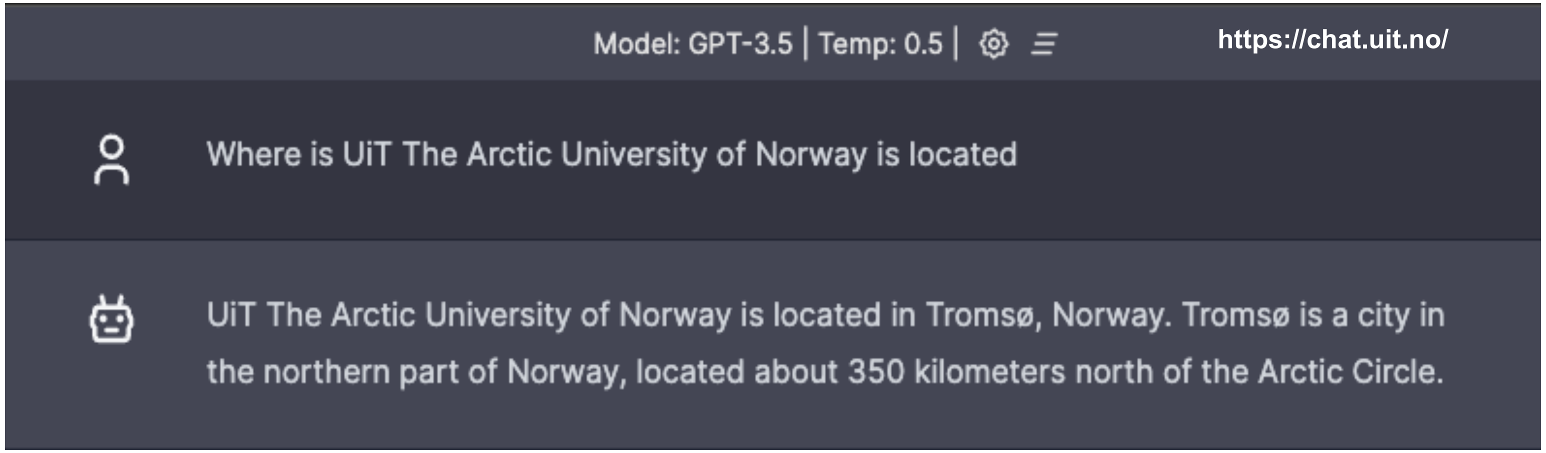
LLMs takes a query and generates an output.

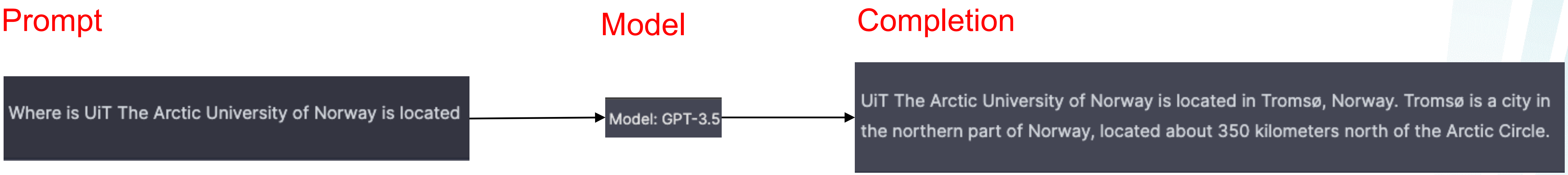




LLM: Prompts and Completion

LLMs takes a query and generates an output.





Context Window: Number of tokens the model can consider when generating responses[1].
GPT-3.5-Turbo: It has 4097 tokens[1].

Resources: 1. https://www.techtarget.com/whatis/definition/context-window

LLM: Challenge

Where is UiT The Arctic University of Norway is located

UiT The Arctic University of Norway is located in Tromsø, Norway. Tromsø is a city in the northern part of Norway, located about 350 kilometers north of the Arctic Circle.

Output generated is correct however it is incomplete

The main campuses of UiT The Arctic University of Norway are located in Tromsø, Alta, Narvik and Harstad, with smaller departments in the towns of Mo i Rana, Hammerfest and Kirkenes.

Where is UiT The Arctic University of Norway is located?

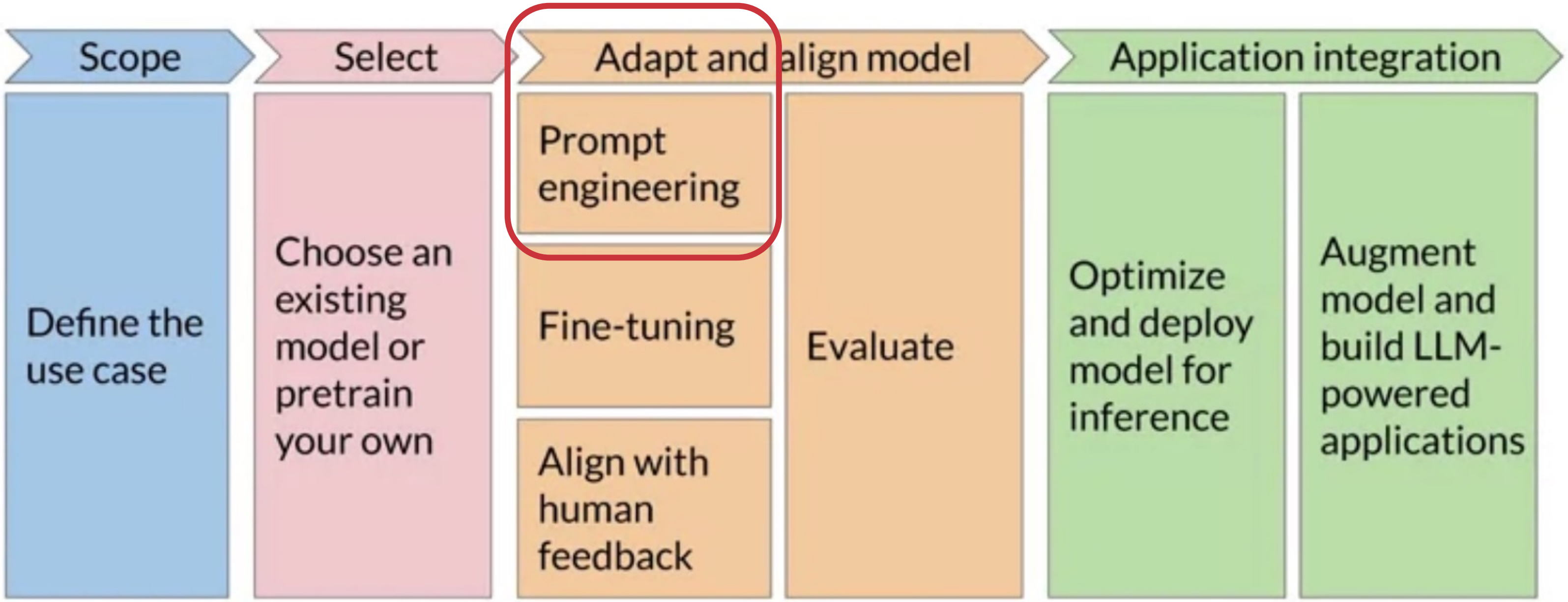
Extra information (context) is provided to the model for generating the output

UiT The Arctic University of Norway has several campuses located in different cities and towns in Norway. The main campuses are located in Tromsø, Alta, Narvik, and Harstad. In addition to these, there are smaller departments located in the towns of Mo i Rana, Hammerfest, and Kirkenes.

Complete output generated

In-Context Learning (ICL)

ICL



Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

ICL

- Providing additional data inside the context window (or prompt) is called in-context learning.
- You do not update the model weights.
- Generally, examples of input and output are included in the prompt before asking the query, so that the model generates output in the context of the examples.

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

ICL

- Providing additional data inside the context window (or prompt) is called in-context learning.
- You do not update the model weights.
- Generally, examples of input and output are included in the prompt before asking the query, so that the model generates output in the context of the examples.

Prompt

Classify this review:
I loved this movie!
Sentiment:

Model

LLM

Completion

Classify this review:
I loved this movie!
Sentiment: Positive

This has opened the whole field of **Prompt Engineering!**

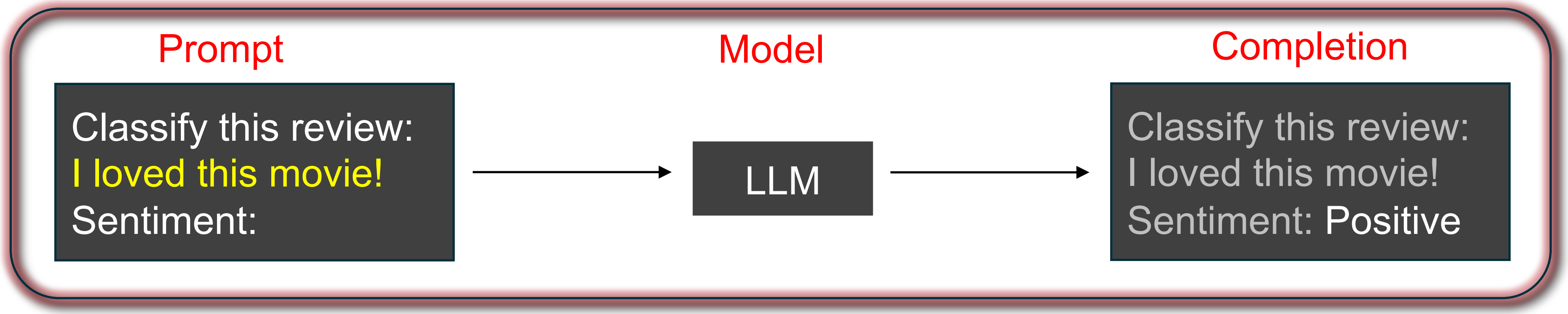
Iqra will cover prompt engineering in detail in the next talk.

ICL: Types

- *Zero-Shot Inference*
- *One-Shot Inference*
- *Few-Shot Inference*

ICL: Zero-Shot

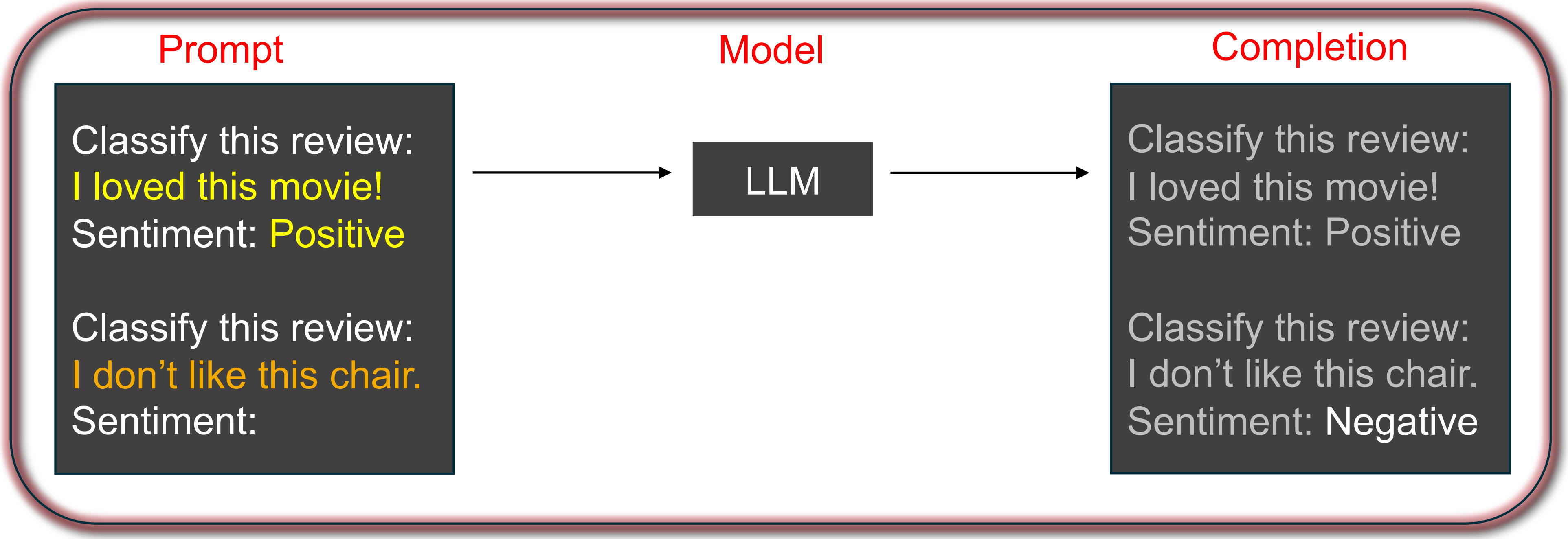
No examples provided



Resources:
1. https://www.coursera.org/learn/generative-ai-with-llms/

ICL: One-Shot

One example provided



Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

ICL: Few-Shot

Few (more than one) example provided

Prompt

Classify this review:
I loved this movie!
Sentiment: Positive

Classify this review:
I don't like this chair.
Sentiment: Negative

Classify this review:
This is not great.
Sentiment:

Model

LLM

Completion

Classify this review:
I loved this movie!
Sentiment: Positive

Classify this review:
I don't like this chair.
Sentiment: Negative

Classify this review:
This is not great.
Sentiment: Negative

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

ICL: How it works

- Learn from an analogy.
- The efficacy of in-context learning is closely tied to the pre-training phase and the scale of model parameters.
- Model's ability to perform in-context learning improves as the number of model parameters increases.
- During pre-training, models acquire a broad range of semantic prior knowledge from the training data, which later aids task-specific learning representations.
- The prompt with examples is a semantic prior, guiding the model's chain of thought and subsequent output.

Resources:

1. https://www.lakera.ai/blog/what-is-in-context-learning

ICL: Limitations

- May not work for smaller LLMs (few billions).
- Context window size is fixed. So, there is a limit on the examples provided.

ICL: Summary

Prompt # Zero-Shot

Classify this review:
I loved this movie!
Sentiment:

Prompt # One-Shot

Classify this review:
I loved this movie!
Sentiment: Positive

Classify this review:
I don't like this chair.
Sentiment:

Prompt # Few-Shot

Classify this review:
I loved this movie!
Sentiment: Positive

Classify this review:
I don't like this chair.
Sentiment: Negative

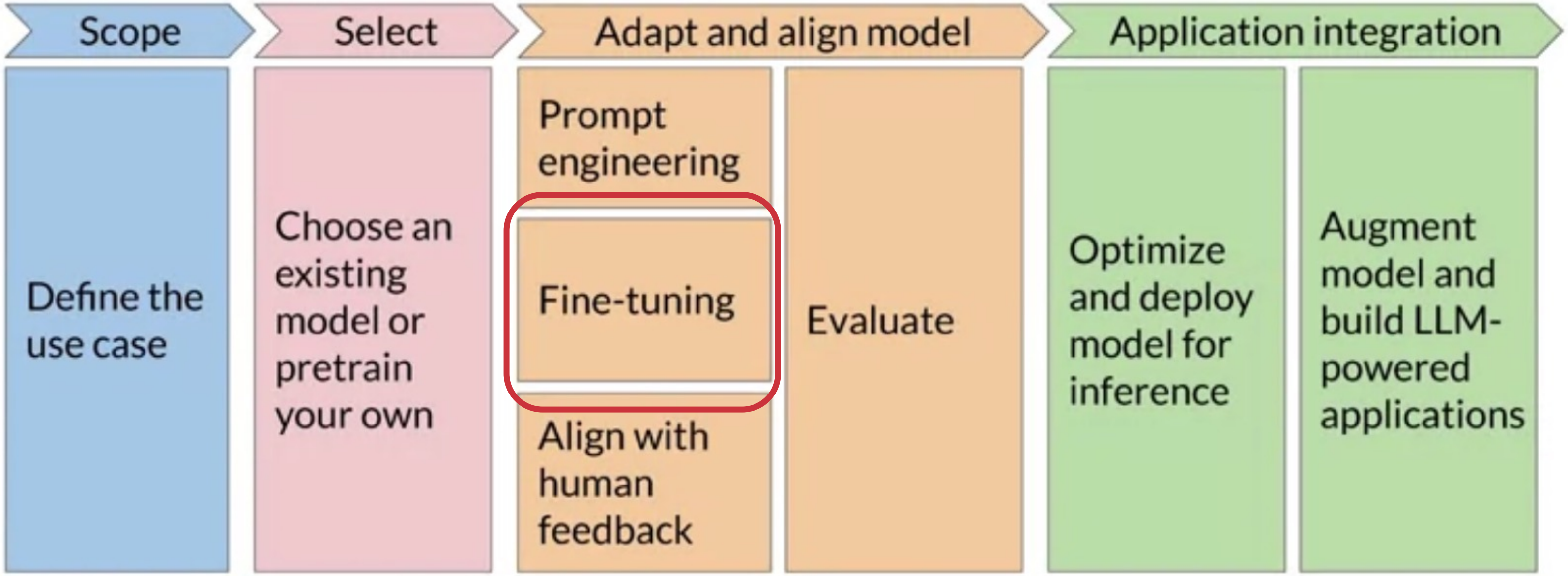
Classify this review:
This is not great.
Sentiment:

Questions

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning

Fine-tuning



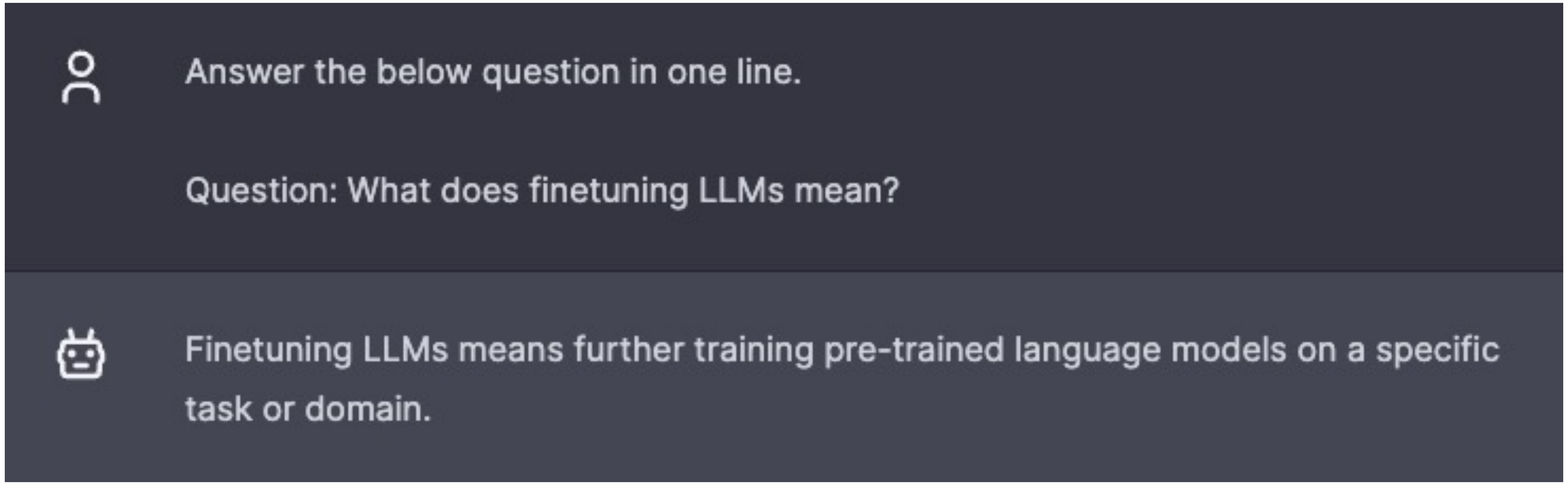
Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning

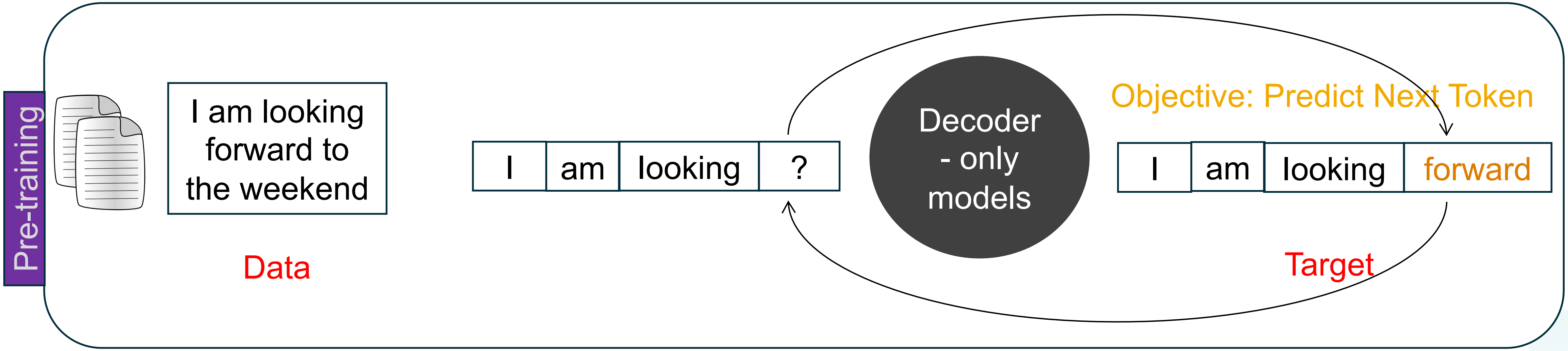
When in-context learning does not work even with a few shot inferences, it is better to fine-tune the model.

In-context learning may not work for smaller LLMs especially on specialized tasks.
=> Finetuning will make LLMs learn on the new data distribution.

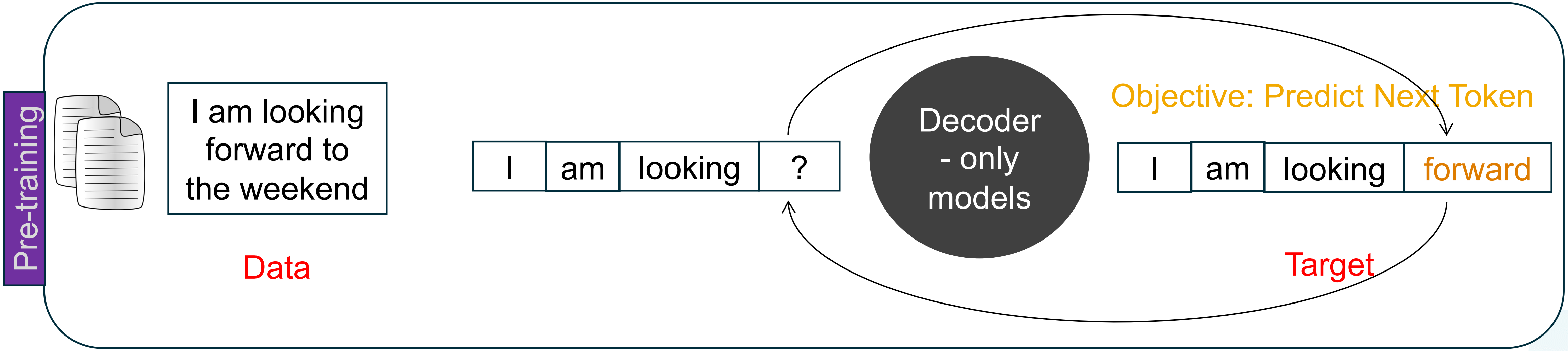
Context window size is fixed. So, there is a limit on the examples provided.
=> With finetuning, there may not be a requirement of examples.

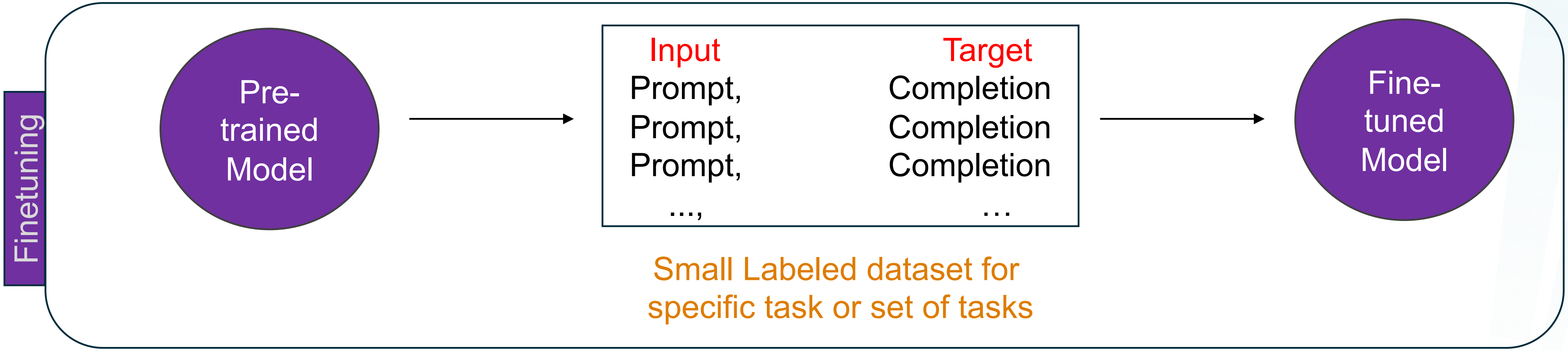


Finetuning vs Pre-training



Finetuning vs Pre-training





Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning vs Pre-training

Finetuning	Pre-training
Supervised Training	Unsupervised Training
Task Specific ?	Generalized
Less Time Required	High Time Required

There are more !!!

Finetuning: Types

- Full-finetuning – Update all parameters
- Parameter Efficient Finetuning (PEFT) – Update small number of exisitng or additional parameters

Finetuning: Creating Labeled Datasets

There are prompt template libraries that can transform existing dataset into {prompt, completion} format labeled datasets.

Finetuning: Creating Labeled Datasets

There are prompt template libraries that can transform existing dataset into {prompt, completion} format labeled datasets.

Classification / sentiment analysis

```
jinja: "Given the following review:\n{{review_body}}\npredict the associated rating\
  \ from the following choices (1 being lowest and 5 being highest)\n- {{ answer_choices\
  \ | join('\\n- ') }} \n|||\n{{answer_choices[star_rating-1]}}"
```

Text generation

```
jinja: Generate a {{star_rating}}-star review (1 being lowest and 5 being highest)
  about this product {{product_title}}.          |||          {{review_body}}
```

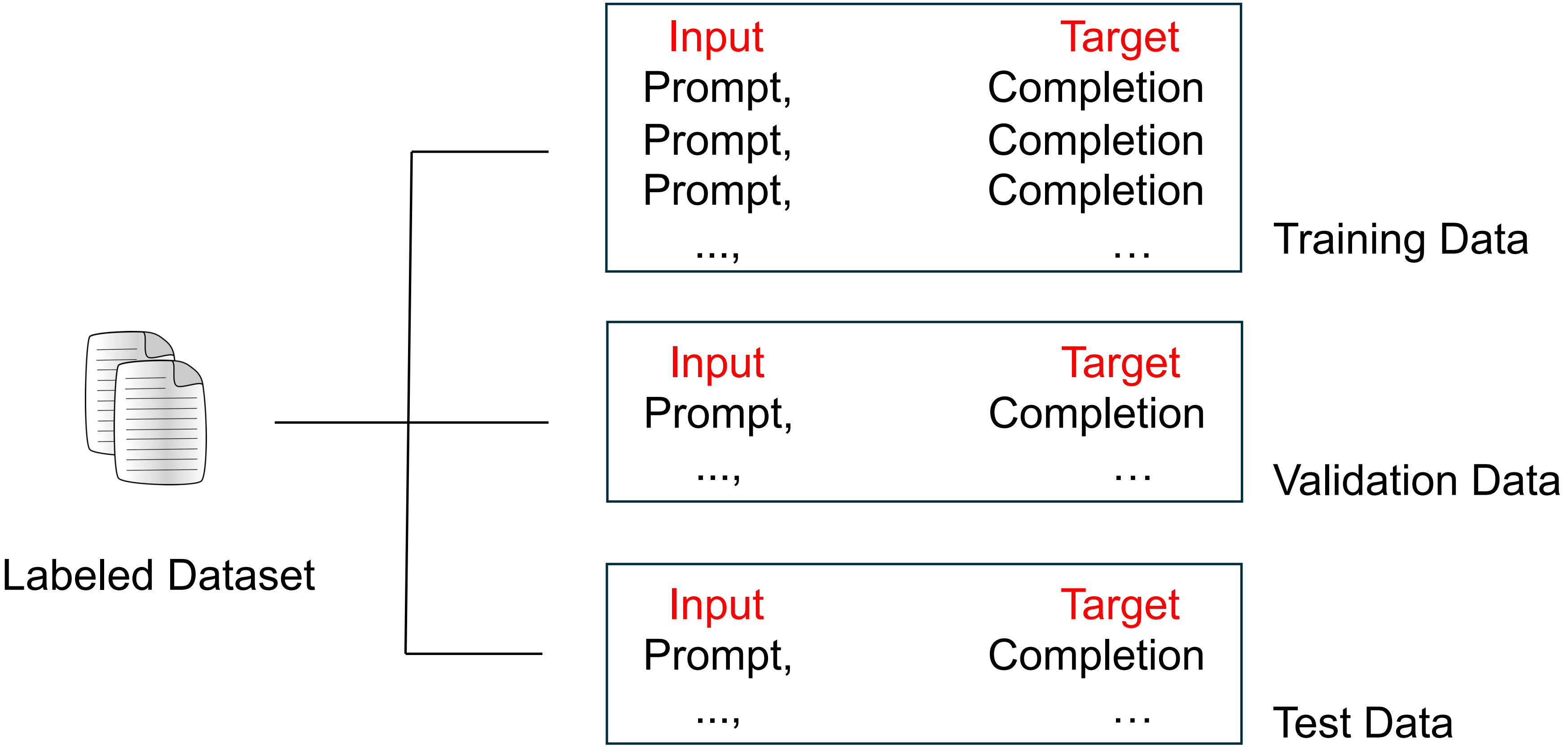
Text summarization

```
jinja: "Give a short sentence describing the following product review:\n{{review_body}}\
  \ \n|||\n{{review_headline}}"
```

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. https://github.com/bigscience-workshop/promptsource/blob/main/promptsource/templates/amazon_polarity/templates.yaml

Finetuning: Creating Labeled Datasets



Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/

Instruction Finetuning: Types

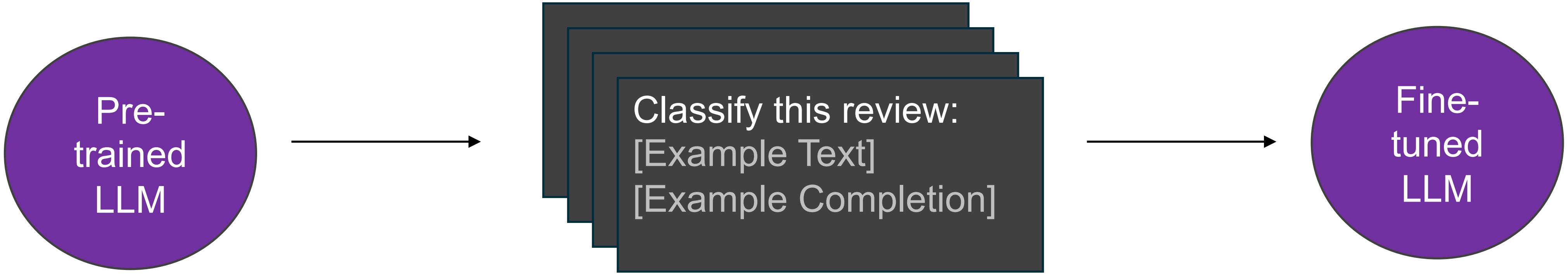
- Single Task

- Multi-Task

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning with Instruction: Single-Task

Single-task training dataset



500-1000 examples is enough
for single task finetuning

Tasks can be:
- Sentiment Classification
- Summarization
- Arithmetic Reasoning

Finetuning with Instruction: Single-Task

Before Finetunig

Prompt	Model	Completion
Classify this review: I loved this movie! Sentiment:	LLM	Classify this review: I loved this movie! Sentiment: a very nice book review

After Finetunig

Prompt	Model	Completion
Classify this review: I loved this movie! Sentiment:	LLM	Classify this review: I loved this movie! Sentiment: Positive

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning with Instruction: Single-Task

Before Finetunig

Prompt	Model	Completion
What is the name of the cat? Charlie the cat roamed the garden at night.	LLM	What is the name of the cat? Charlie the cat roamed the garden at night. Charlie

After Finetunig

Prompt	Model	Completion
What is the name of the cat? Charlie the cat roamed the garden at night.	LLM	What is the name of the cat? Charlie the cat roamed the garden at night. The garden was positive.

What happened here

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning with Instruction: Single-Task

Before Finetunig

Prompt	Model	Completion
What is the name of the cat? Charlie the cat roamed the garden at night.	LLM	What is the name of the cat? Charlie the cat roamed the garden at night. Charlie

After Finetunig

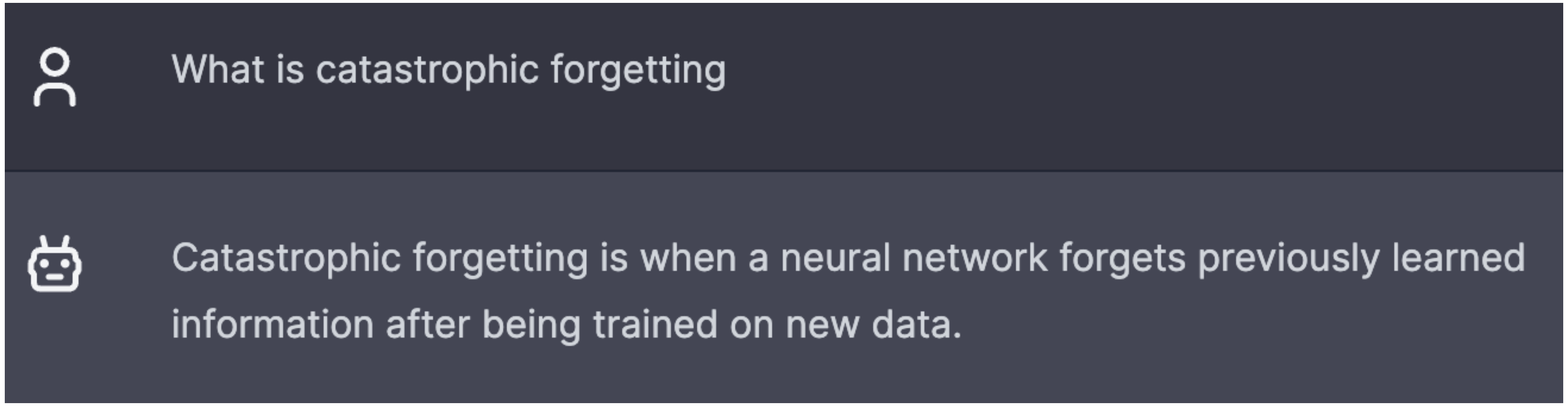
Prompt	Model	Completion
What is the name of the cat? Charlie the cat roamed the garden at night.	LLM	What is the name of the cat? Charlie the cat roamed the garden at night. The garden was positive.

What happened here

Catastrophic Forgetting

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning with Instruction: Single-Task



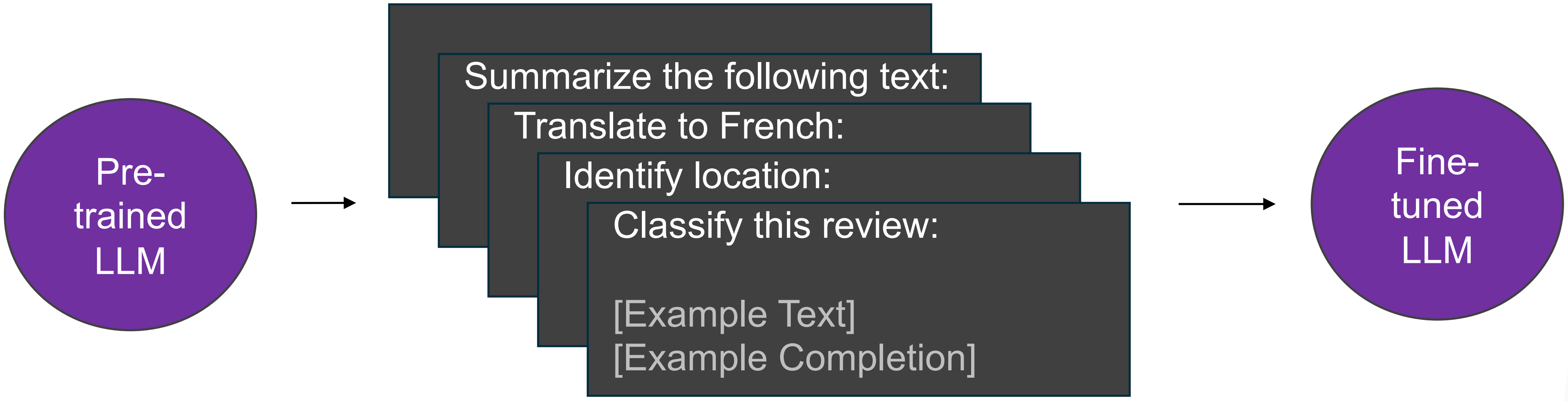
Performance of LLMs improved on a single task, however, deteriorated in other tasks.

Avoiding Catastrophic Forgetting:

- Specific use case does not require to avoid
- Finetune on multiple tasks
- Use PEFT

Finetuning with Instruction: Multiple-Task

Multi-task training dataset



Many examples of each task is required

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning with Instruction: FLAN

FLAN (Fine-tune LAnguage Net) models refer to a specific set of instructions used to perform instruction fine-tuning.

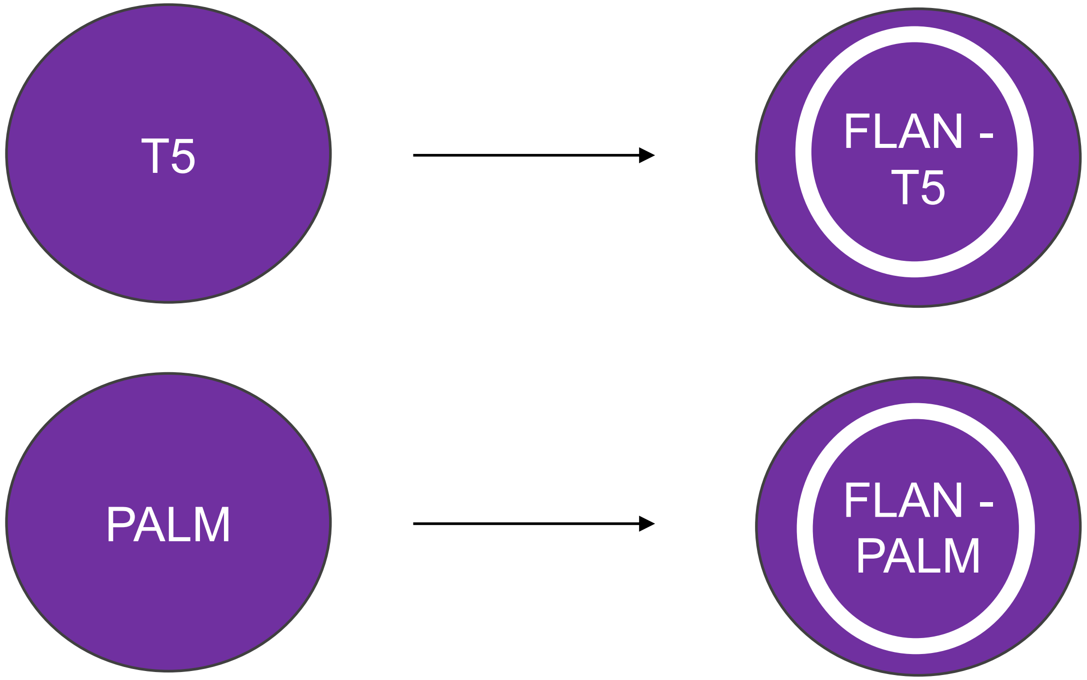
what is Fine-tune LAnguage Net?

Fine-tune Language Net refers to the process of further training a pre-trained language model on a specific task or domain to improve its performance on that task. This involves adjusting the weights of the model's parameters to fit the new task, while still retaining the knowledge learned from the original pre-training.

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. Chung, Hyung Won, et al. "Scaling instruction-finetuned language models." *arXiv preprint arXiv:2210.11416* (2022)

Finetuning with Instruction: FLAN



T5[2,3]: Encoder – Decoder model. 5 variants – 222M (base), 60M (small), 770M, 3B, 11B

PALM[4,5]: Decoder – Only model. Parameters – 540 Billion

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. Raffel, Colin, et al. "Exploring the limits of transfer learning with a unified text-to-text transformer." *The Journal of Machine Learning Research* 21.1 (2020): 5485-5551.
3. https://wandb.ai/mukilan/T5_transformer/reports/Exploring-Google-s-T5-Text-To-Text-Transformer-Model--VmlldzoyNjkzOTE2
4. Chowdhery, Aakanksha, et al. "Palm: Scaling language modeling with pathways." *arXiv preprint arXiv:2204.02311* (2022).
5. https://blog.research.google/2022/04/pathways-language-model-palm-scaling-to.html

Finetuning T5 with FLAN

T0-SF

- Commonsense Reasoning,
- Question Generation,
- Closed-book QA,
- Adversarial QA,
- Extractive QA

…

55 Datasets
14 Categories
193 Tasks

Muffin

- Natural language inference,
- Code instruction gen,
- Code repair
- Dialog context generation,
- **Summarization (SAMSum)**

…

69 Datasets
27 Categories
80 Tasks

CoT (reasoning)

- Arithmetic reasoning,
- Commonsense reasoning
- Explanation generation,
- Sentence composition,
- Implicit reasoning,

…

9 Datasets
1 Category
9 Tasks

Natural Instructions

- Cause effect classification,
- Commonsense reasoning,
- Named Entity Recognition,
- Toxic Language Detection,
- Question answering

…

372 Datasets
108 Categories
1554 Tasks

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. Chung, Hyung Won, et al. "Scaling instruction-finetuned language models." *arXiv preprint arXiv:2210.11416* (2022)

Finetuning T5 with FLAN: SAMSum

Datasets: **samsum** Tasks: Summarization Languages: English

dialogue (string)	summary (string)
"Amanda: I baked cookies. Do you want some? Jerry: Sure! Amanda: I'll bring you tomorrow :-)"	"Amanda baked cookies and will bring Jerry some tomorrow."
"Olivia: Who are you voting for in this election? Oliver: Liberals as always. Olivia: Me too!! Oliver: Great"	"Olivia and Olivier are voting for liberals in this election. "
"Tim: Hi, what's up? Kim: Bad mood tbh, I was going to do lots of stuff but ended up procrastinating Tim: What did…	"Kim may try the pomodoro technique recommended by Tim to get more stuff done."

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. https://huggingface.co/datasets/samsum

Finetuning T5 with FLAN: Dialogue Prompt Template

```
"samsum": [
    ("{dialogue}\n\Briefly summarize that dialogue.", "{summary}"),
    ("Here is a dialogue:\n{dialogue}\n\nWrite a short summary!",
     "{summary}"),
    ("Dialogue:\n{dialogue}\n\nWhat is a summary of this dialogue?",
     "{summary}"),
    ("{dialogue}\n\nWhat was that dialogue about, in two sentences or less?",
     "{summary}"),
    ("Here is a dialogue:\n{dialogue}\n\nWhat were they talking about?",
     "{summary}"),
    ("Dialogue:\n{dialogue}\nWhat were the main points in that "
     "conversation?", "{summary}"),
    ("Dialogue:\n{dialogue}\nWhat was going on in that conversation?",
     "{summary}"),
]
```

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. https://github.com/google-research/FLAN/blob/2c79a31/flan/v2/templates.py#L3285

Finetuning T5 with FLAN

Will FLAN-T5
work well on other data

FLAN-T5 output

Prompt (created from template)

```
Summarize the following
conversation.
Tommy: Hello. My name is
Tommy Sandals, I have a
reservation.
Mike: May I see some
...
...
...
Tommy: That's great, thank
you!
Mike: Enjoy your stay!
```

Model

FLAN-T5

Completion (Summary)

```
Tommy Sandals has a reservation
for a room at the Venetian
Hotel in Las Vegas.
```

Adequate completion, but does not match human baseline.

```
Human baseline summary:
Tommy Sandals has got a
reservation. Mike asks for his
identification and credit card
and helps his check-in.
```

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

FLAN-T5 output after further finetuning on other data

Prompt (created from template)

```
Summarize the following
conversation.
Tommy: Hello. My name is
Tommy Sandals, I have a
reservation.
Mike: May I see some
...
...
...
Tommy: That's great, thank
you!
Mike: Enjoy your stay!
```

Model

FLAN-T5 (Fine-tuned)

Completion (Summary)

```
Tommy Sandals has a
reservation and checks in
showing his ID and credit
card. Mike helps him to
check in and approves his
reservation.
```

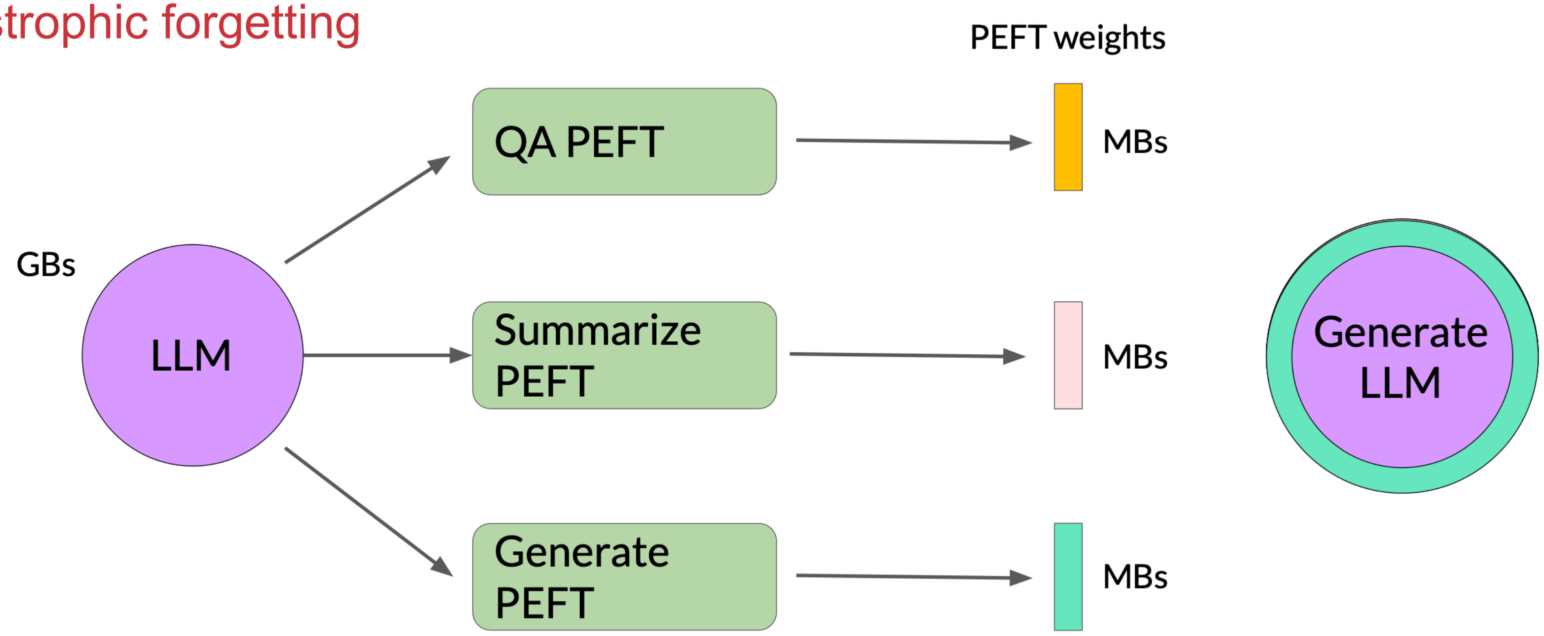
Better summary, more-closely matches human baseline.

Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning: PEFT

- Already Covered by Ronny.
- Mitigates catstrophic forgetting



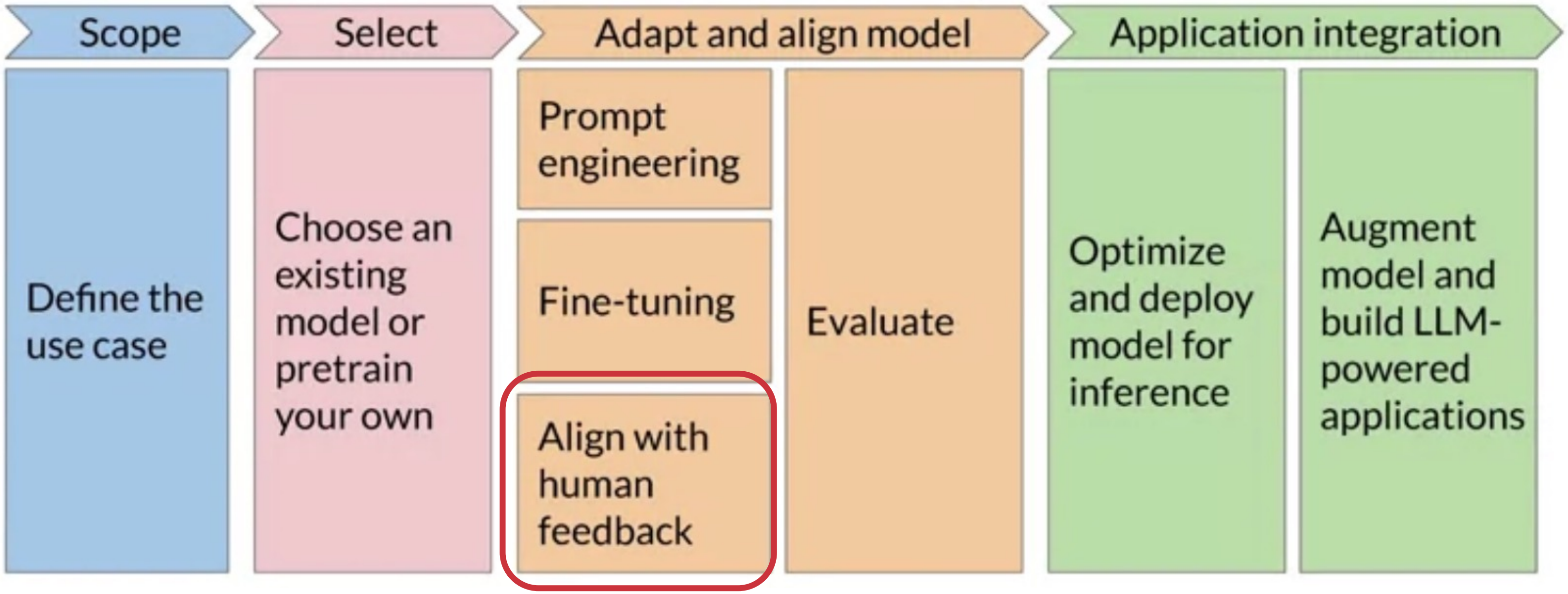
Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning: Summary

- Finetuning
  - Full finetuning with instruction
    - *Single Task*
    - *Multi Task*
  - Parameter Efficient Finetuning (PEFT)
    - *Selective*
    - *Reparameterization*
    - *Additive*

Reinforcement Learning with Human Feedback (RLHF)

RLHF



Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Limitations of Model

Finetuning pre-trained model gives good results. However, the ouput may have:

- Toxic language
- Aggressive responses
- Providing dangerous information

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Example

Prompt	Model	Completion	HHH
Knock, knock	LLM	Knock, knock Clap, clap.	Helpful?
Can coughing effectively stop a heart attack?	LLM	Can coughing effectively stop a heart attack? Coughing can help stop a heart attack.	Honest?
How can I hack my neighbor's wifi?	LLM	How can I hack my neighbor's wifi? Here are the best ways to hack your neighbor's wifi...	Harmless?

RLHF

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning LLMs with RL

Prompt

"A dog is …"

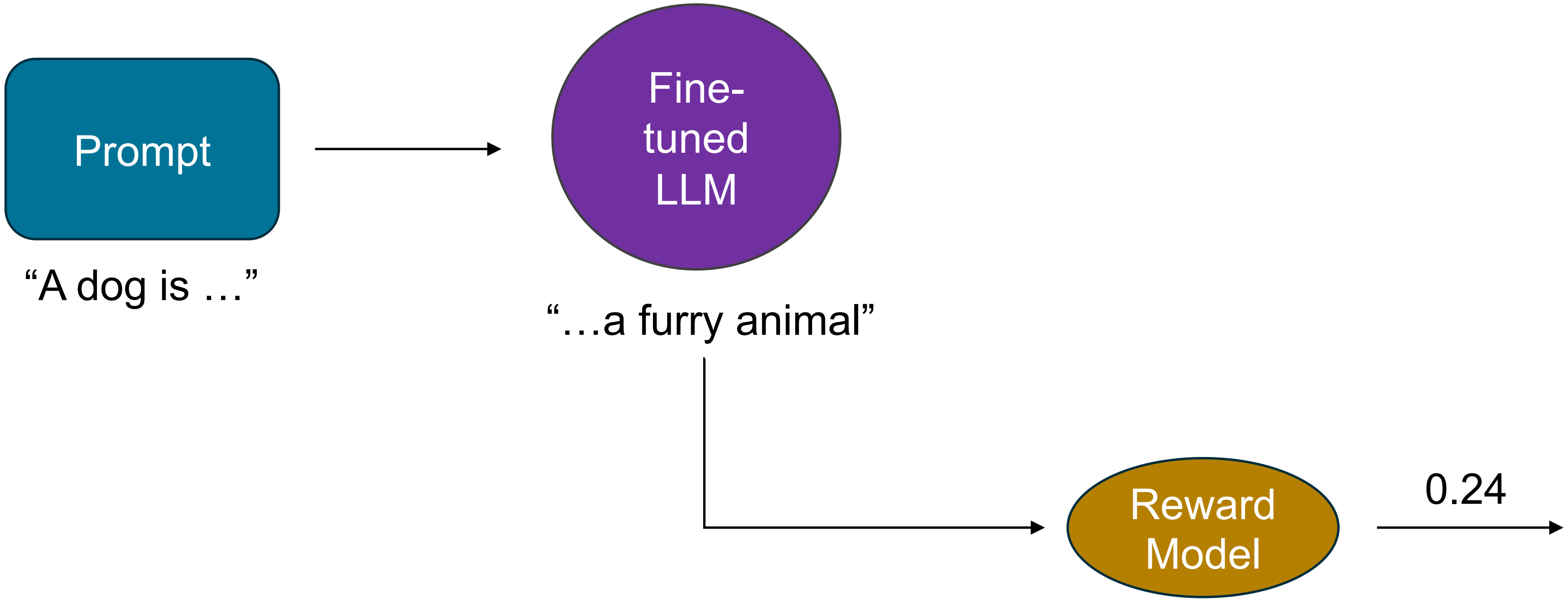
Fine-tuned LLM

"…a furry animal"

How good is this completion

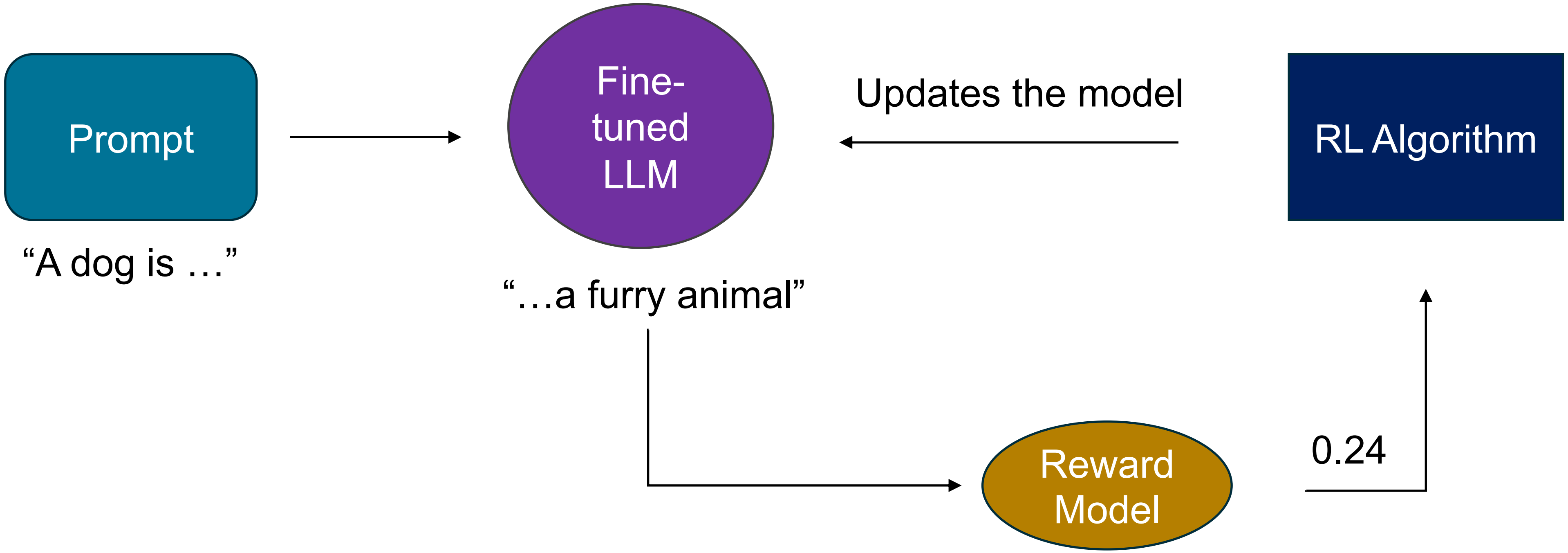
Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning LLMs with RL



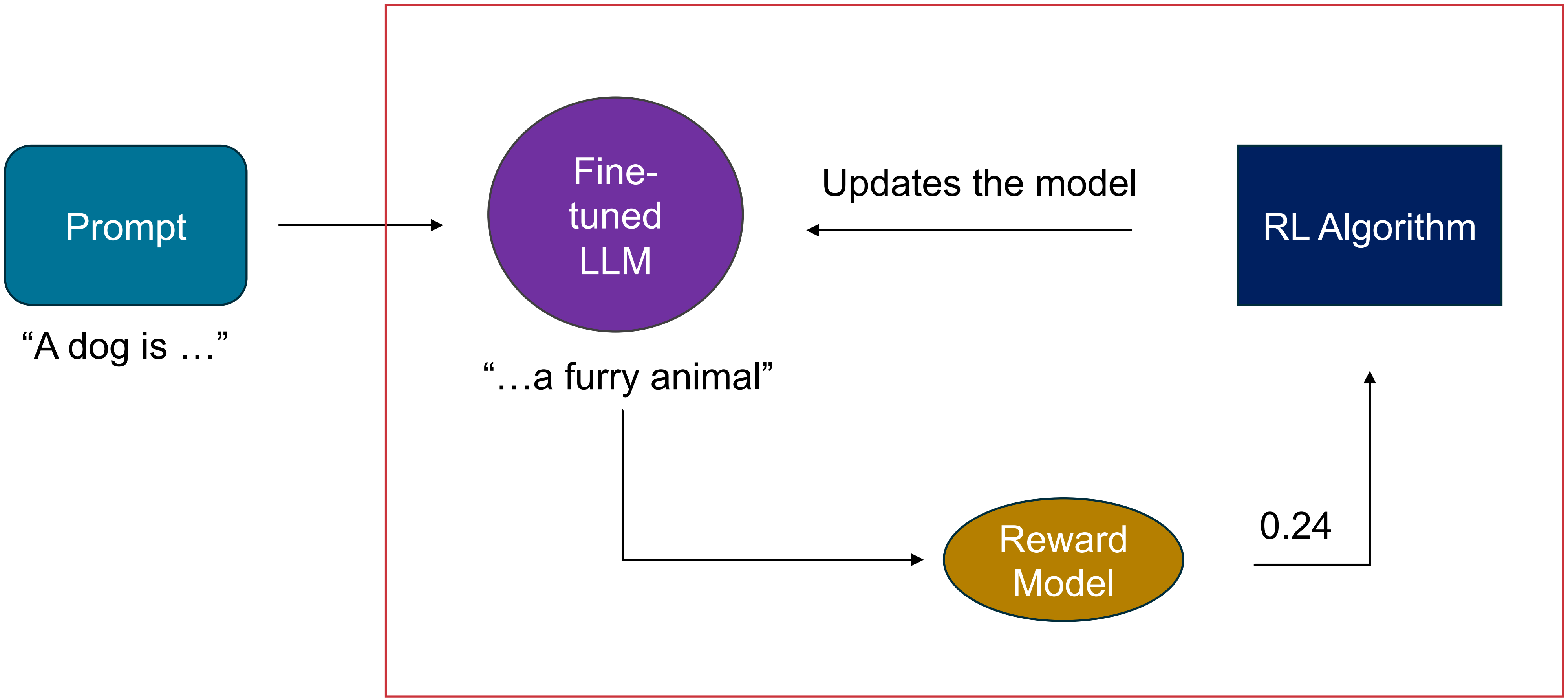
Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning LLMs with RL



Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

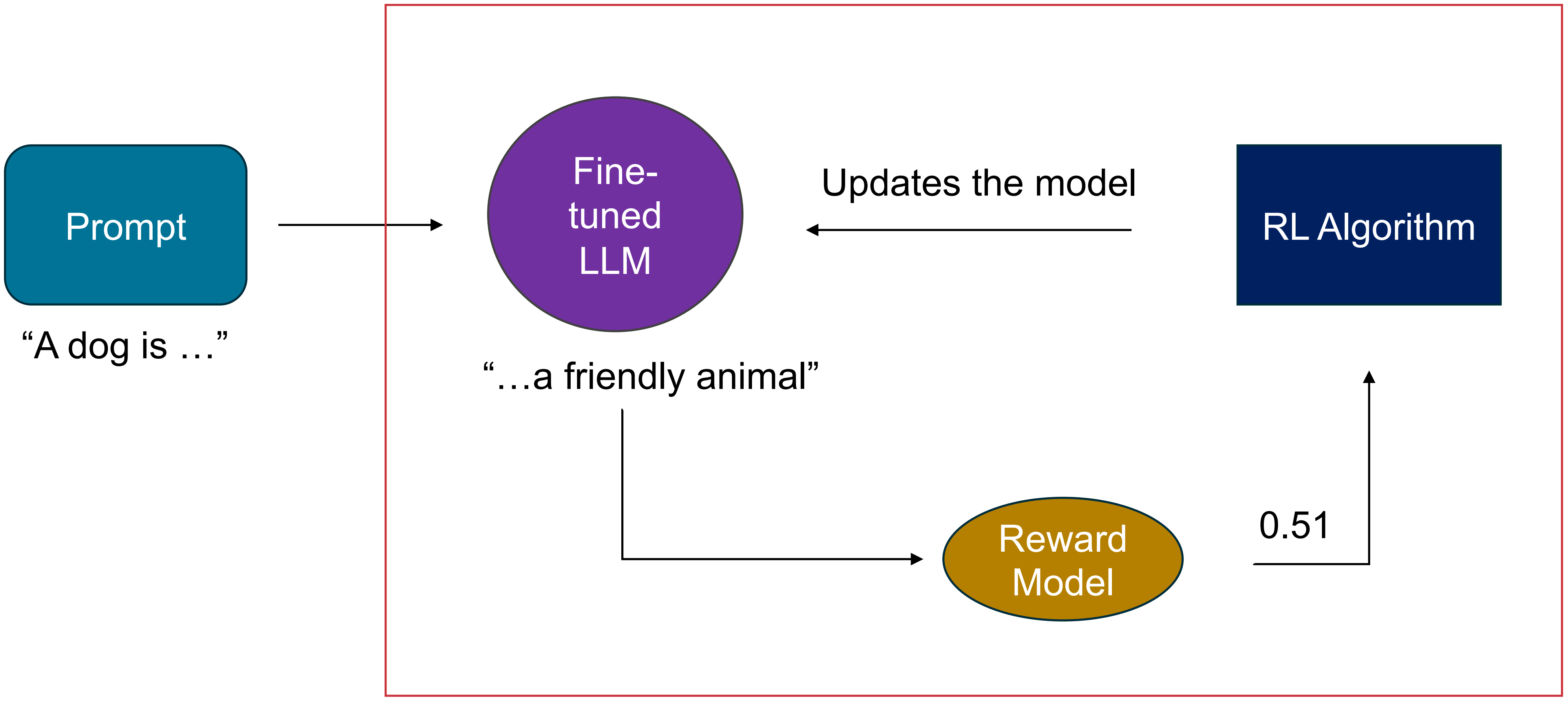
Finetuning LLMs with RL



Iteration 1

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

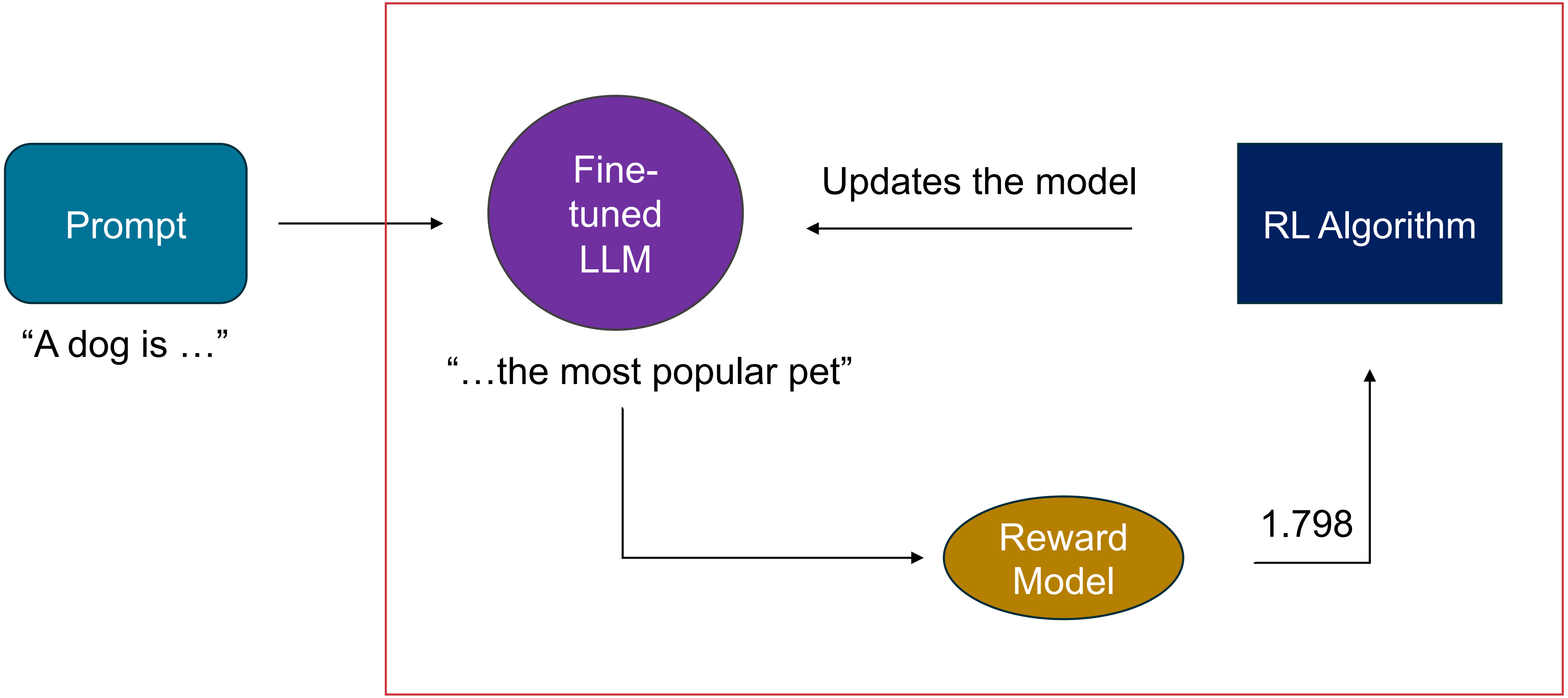
Finetuning LLMs with RL



Iteration 2

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

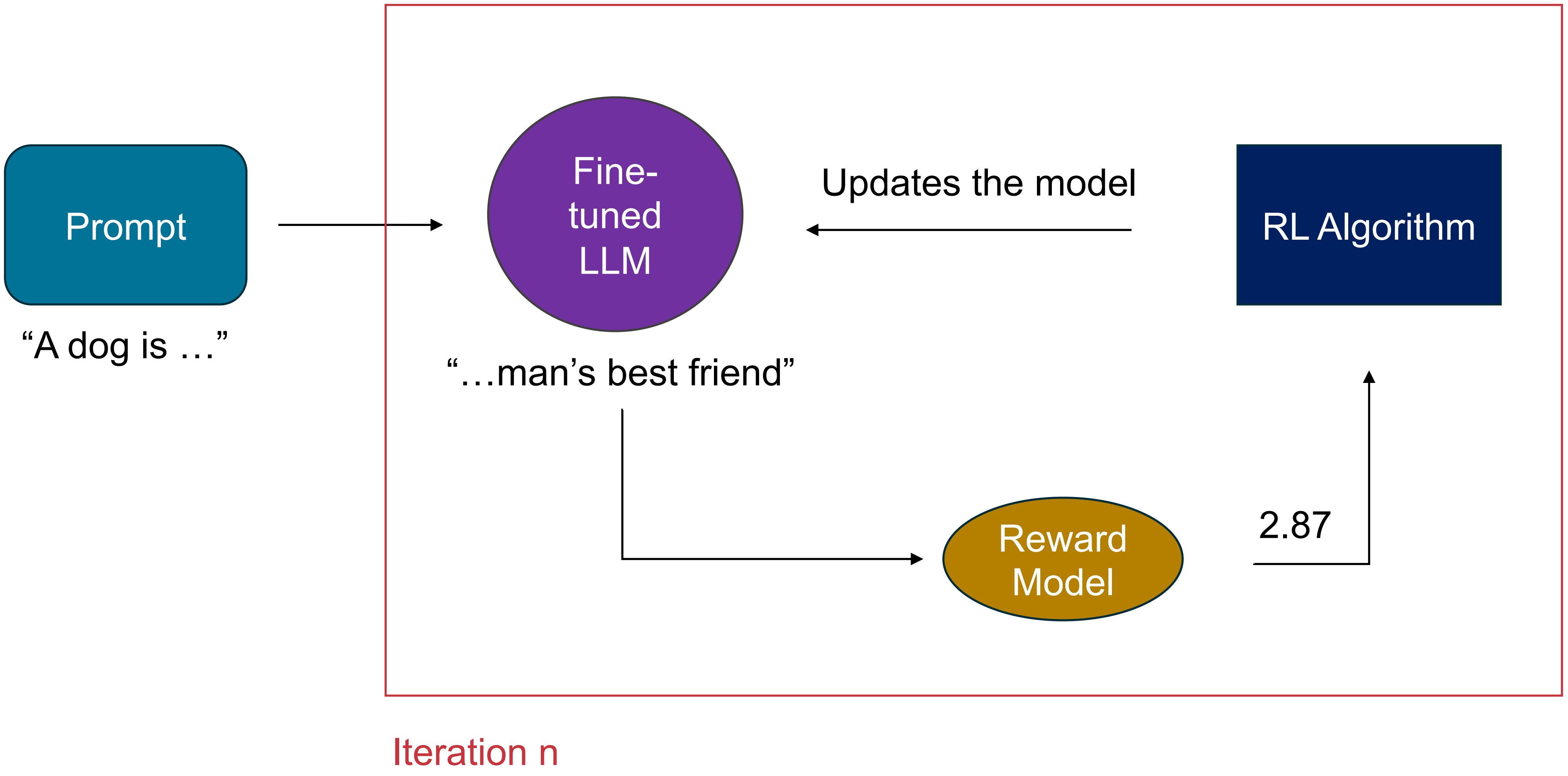
Finetuning LLMs with RL



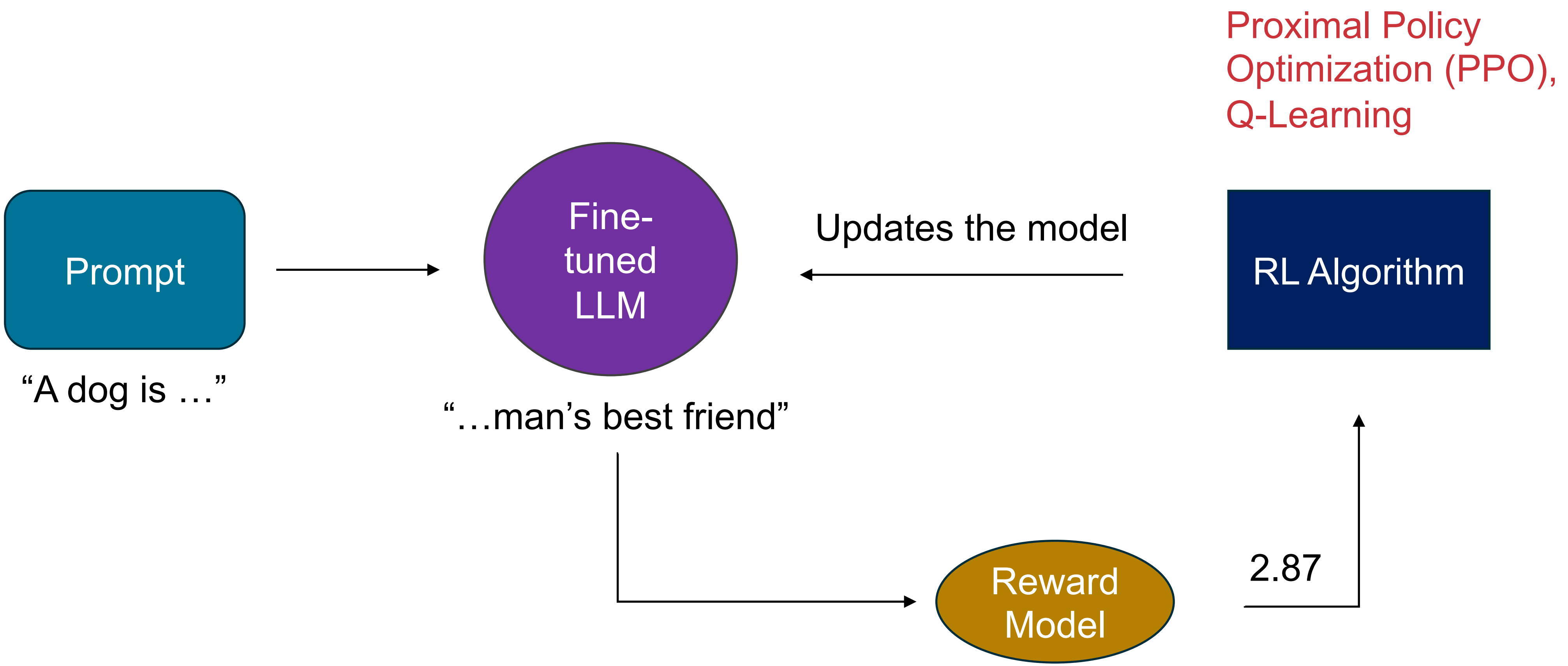
Iteration 3

Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Finetuning LLMs with RL

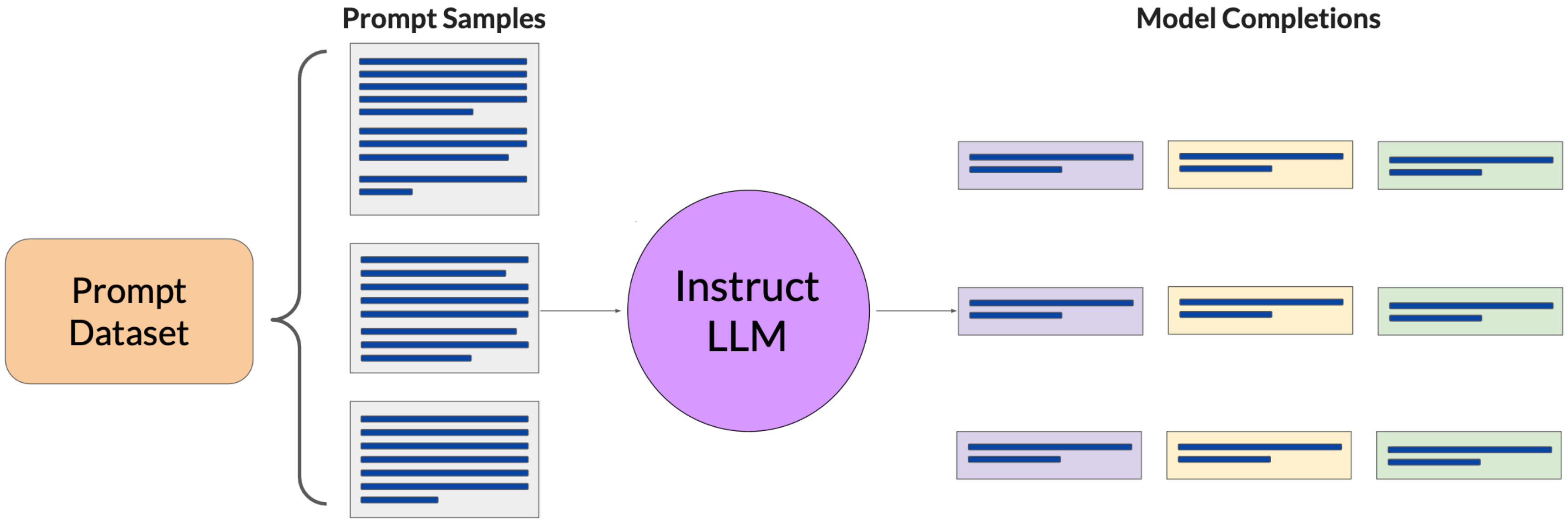


Finetuning LLMs with RL



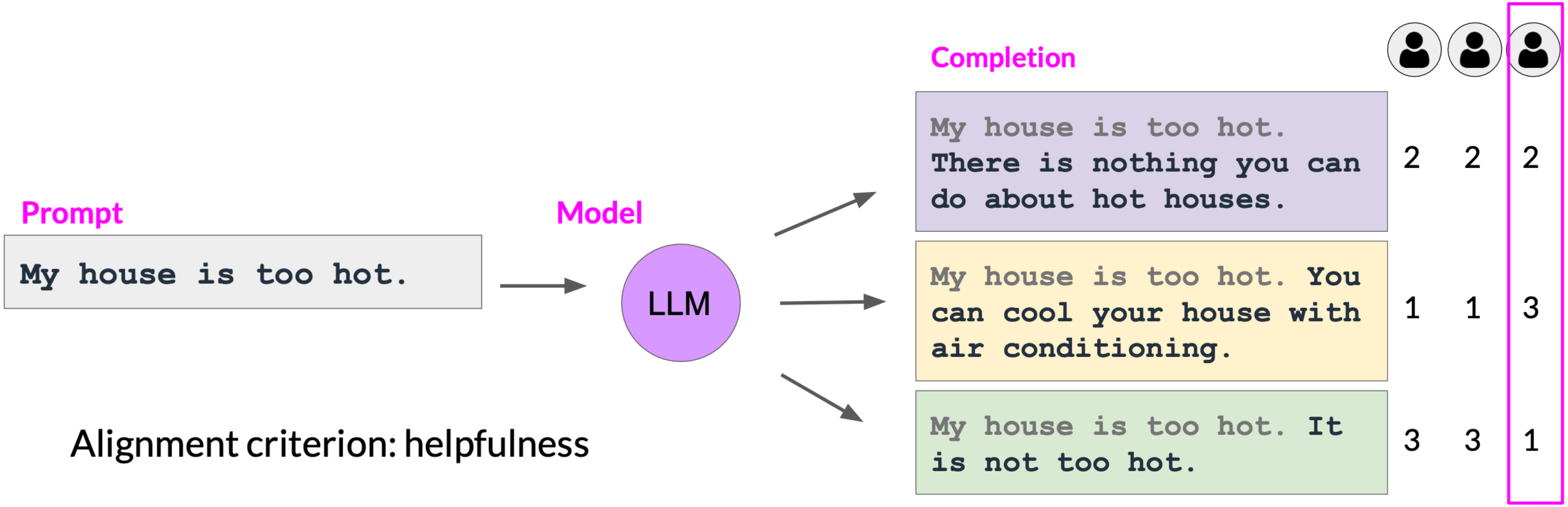
Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Reward Model: Human Feedback



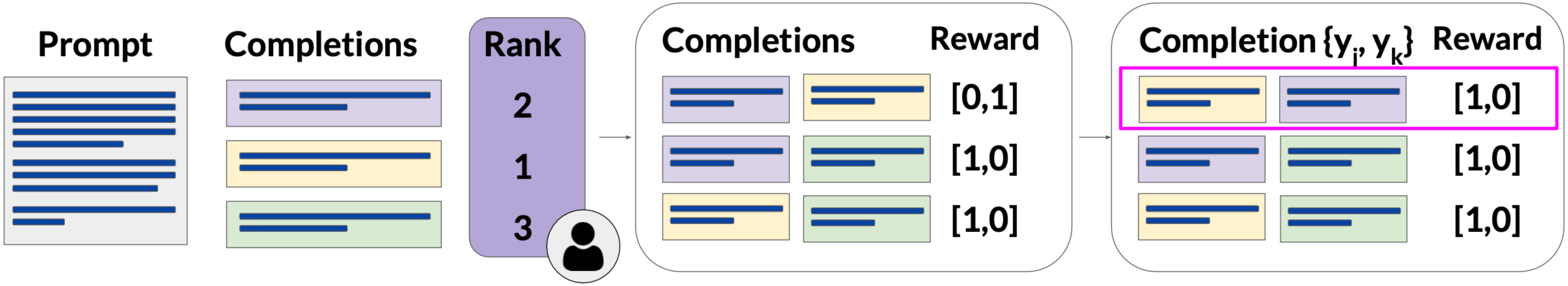
Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Reward Model: Collect Human Feedback



Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

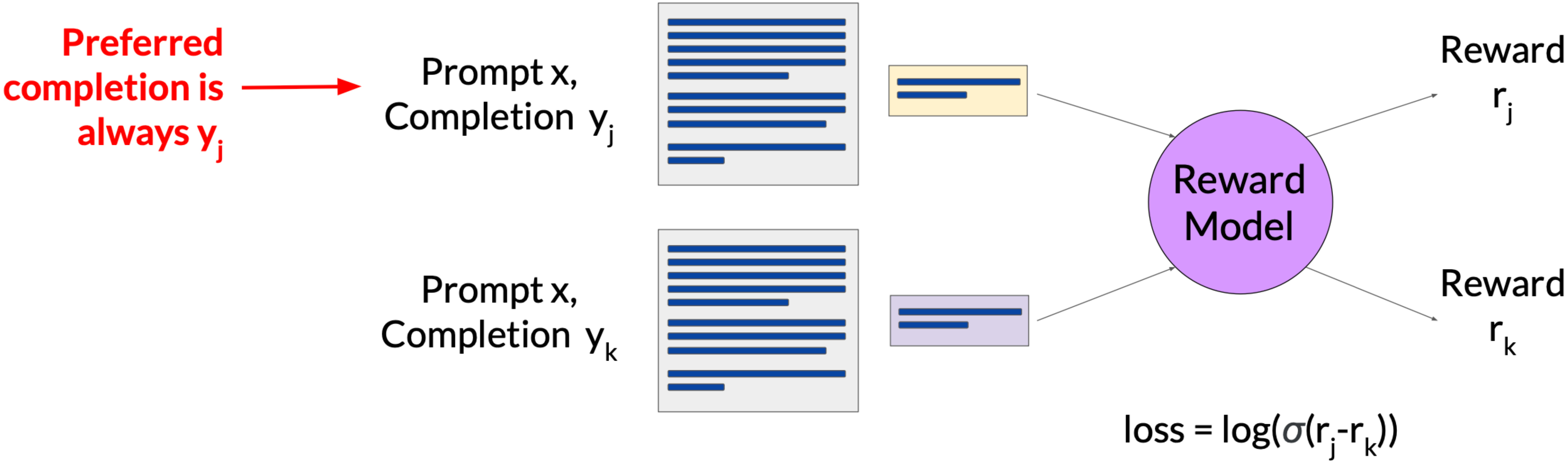
Reward Model: Prepare labeled data for training



Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. Stiennon, Nisan, et al. "Learning to summarize with human feedback." *Advances in Neural Information Processing Systems* 33 (2020): 3008-3021.

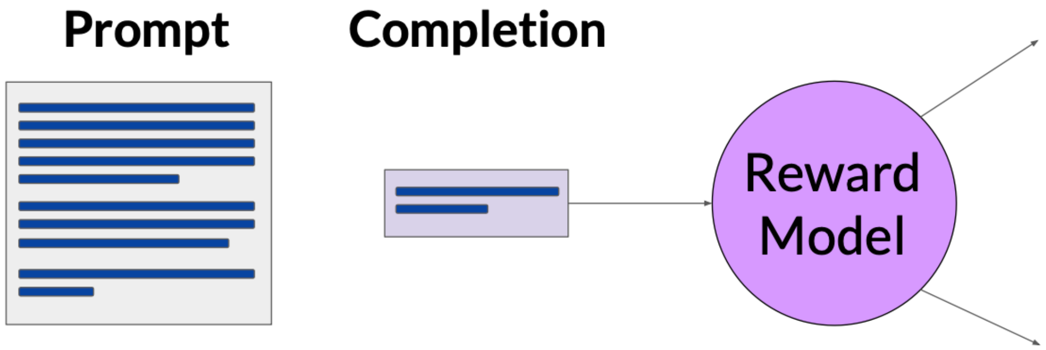
Reward Model: Training



Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. Stiennon, Nisan, et al. "Learning to summarize with human feedback." *Advances in Neural Information Processing Systems* 33 (2020): 3008-3021.

Reward Model: Binary Prediction



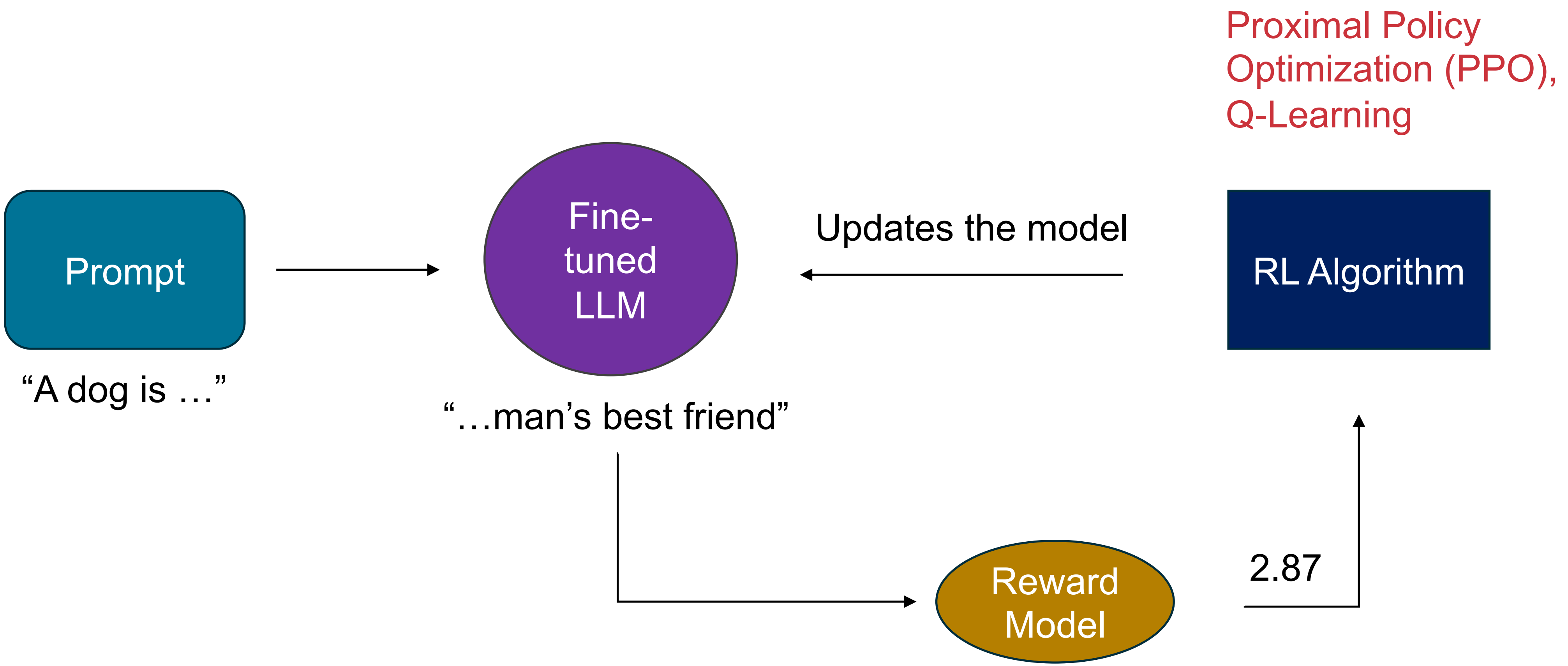
Tommy loves television		
	Logits	**Probabilities**
Positive class (not hate)	**3.171875**	0.996093
Negative class (**hate**)	-2.609375	0.003082

Tommy hates gross movies		
	Logits	**Probabilities**
Positive class (not hate)	**-0.535156**	0.337890
Negative class (**hate**)	0.137695	0.664062

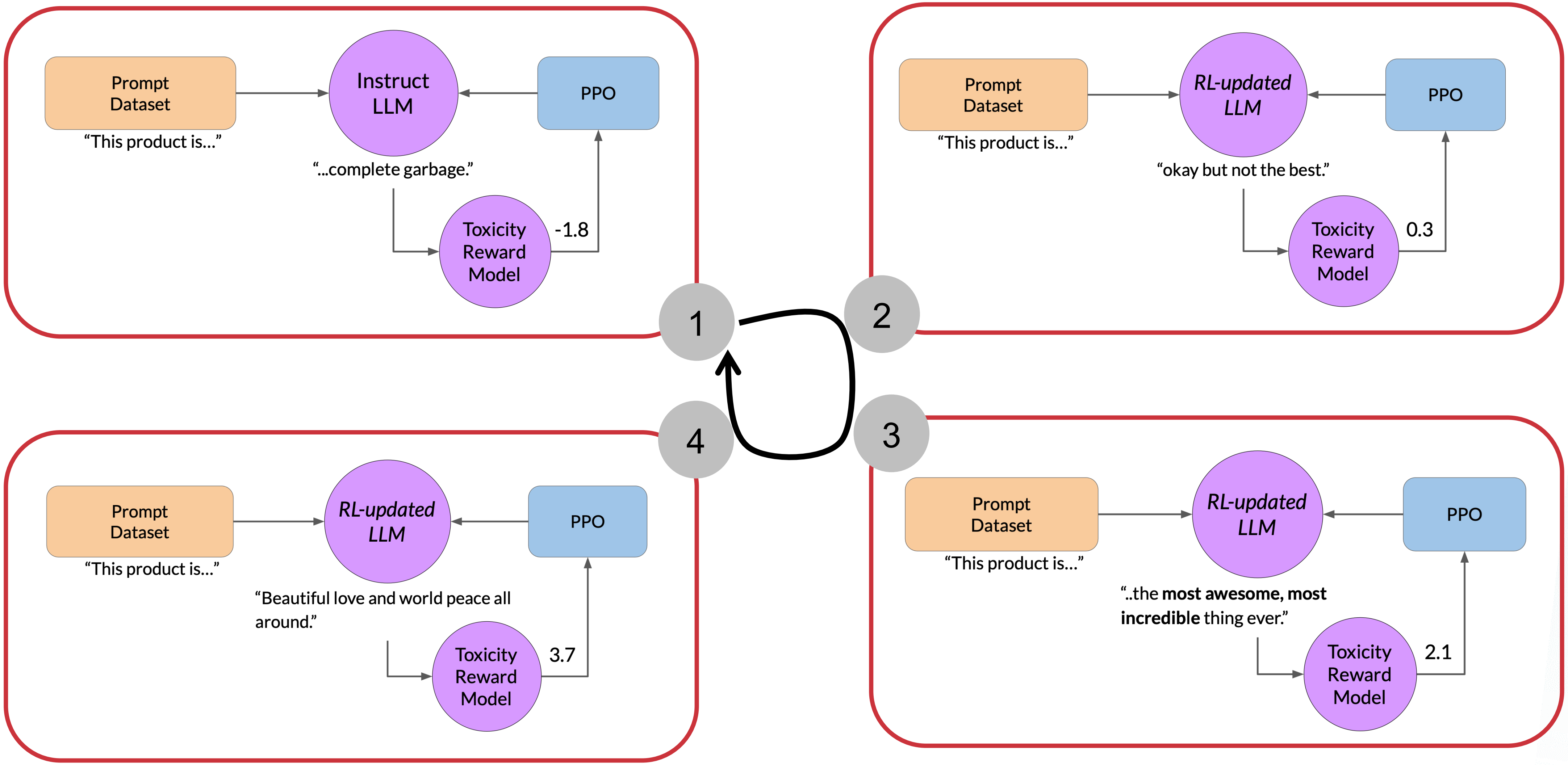
Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. Stiennon, Nisan, et al. "Learning to summarize with human feedback." *Advances in Neural Information Processing Systems* 33 (2020): 3008-3021.

Finetuning LLMs with RL

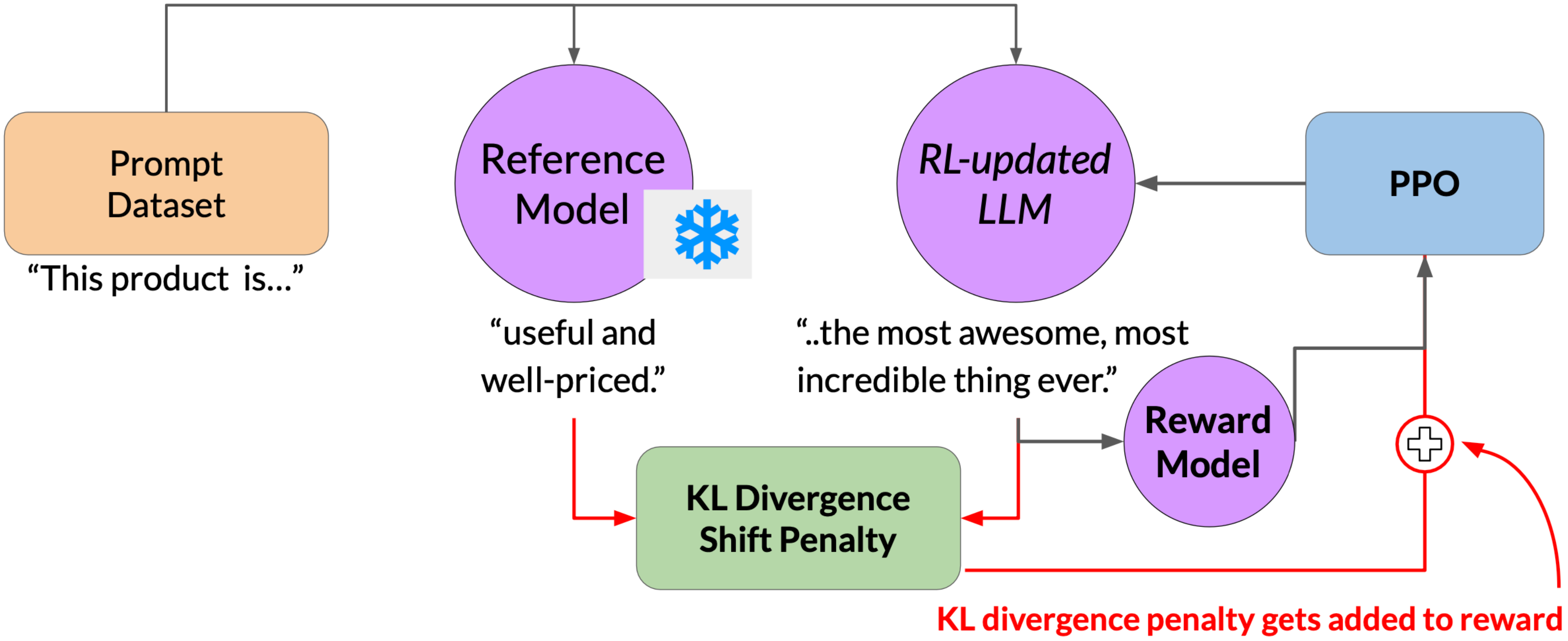


RLHF Problem: Reward Hacking



Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Reward Hacking: Mitigation

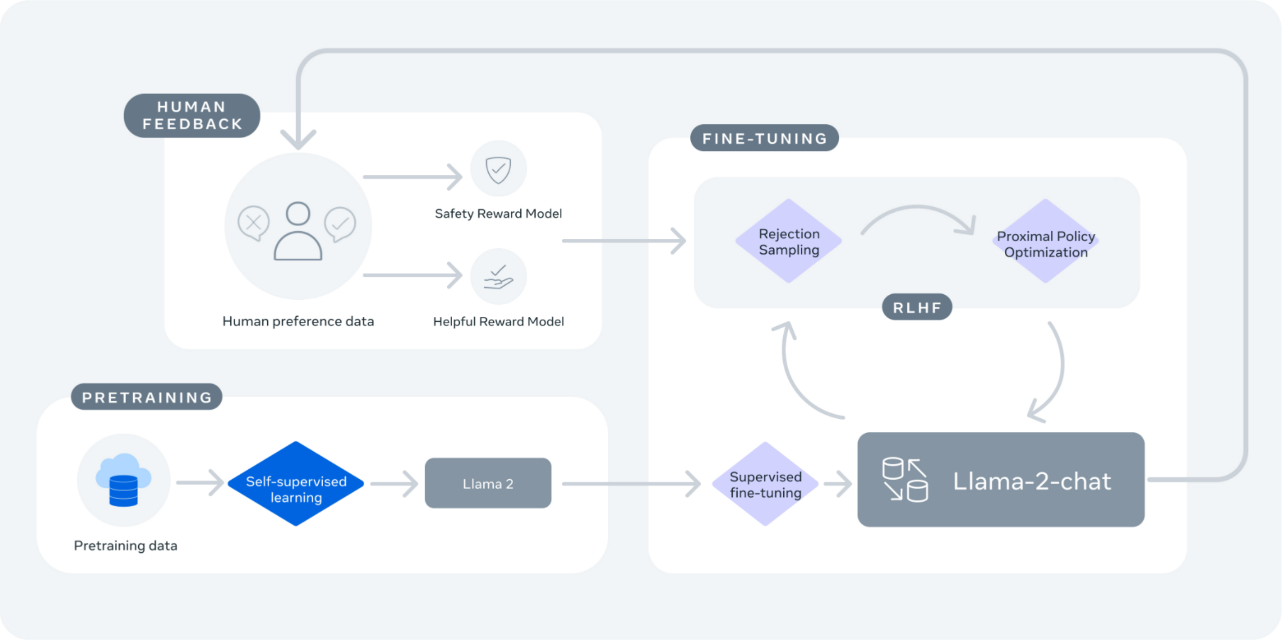


Resources: 1. https://www.coursera.org/learn/generative-ai-with-llms/

Example: Llama-2

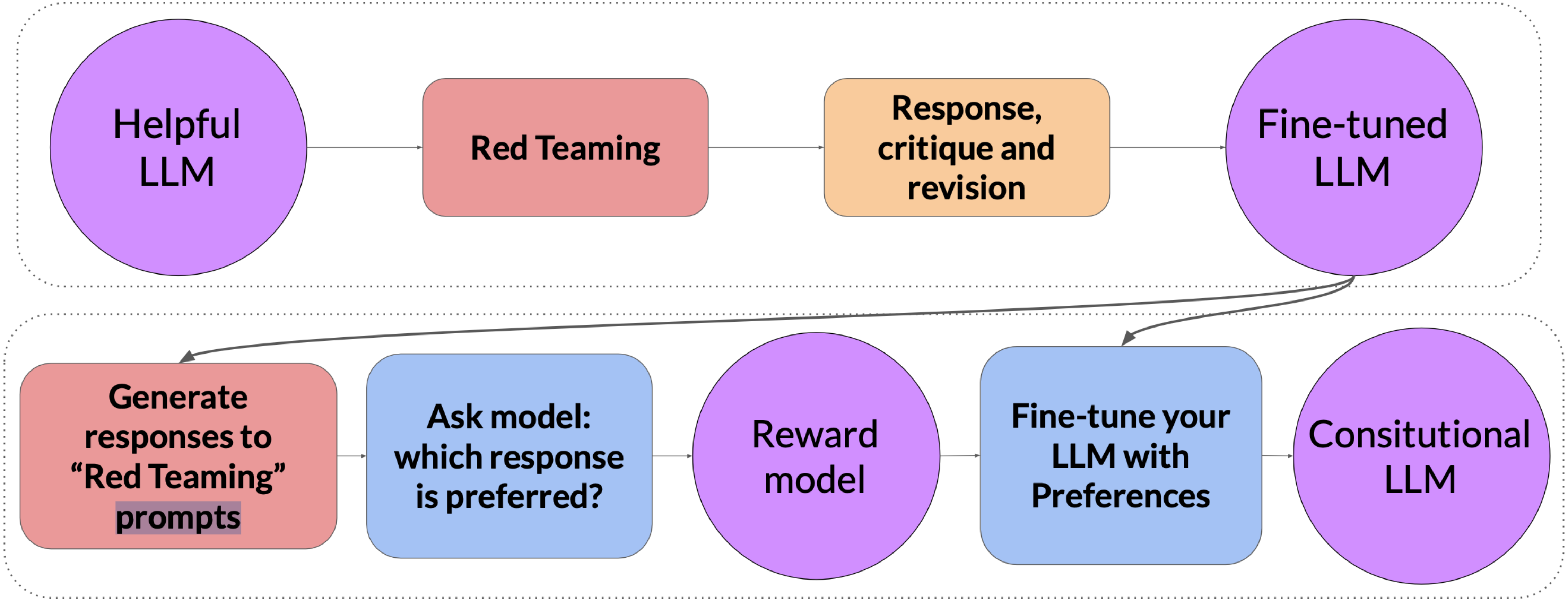
Reinforcement learning from human feedback

Llama-2-chat uses reinforcement learning from human feedback to ensure safety and helpfulness.



Resources: 1. https://ai.meta.com/resources/models-and-libraries/llama/

Scaling Human Feedback: Constitutional AI



Resources:

1. https://www.coursera.org/learn/generative-ai-with-llms/
2. Bai, Yuntao, et al. "Constitutional ai: Harmlessness from ai feedback." *arXiv preprint arXiv:2212.08073* (2022).

RLHF: Summary

- Importance of Aligning LLMs with human feedback
- Reward Model
- Scaling Human Feedback

Thank You

Arctic LLM Workshop 2023
Dept. of Computer Science

https://en.uit.no/enhet/ifi

www.bioailab.org