

Generative AI for Researchers



Copyright Notice



These slides are distributed for educational purpose under the Creative Commons License.

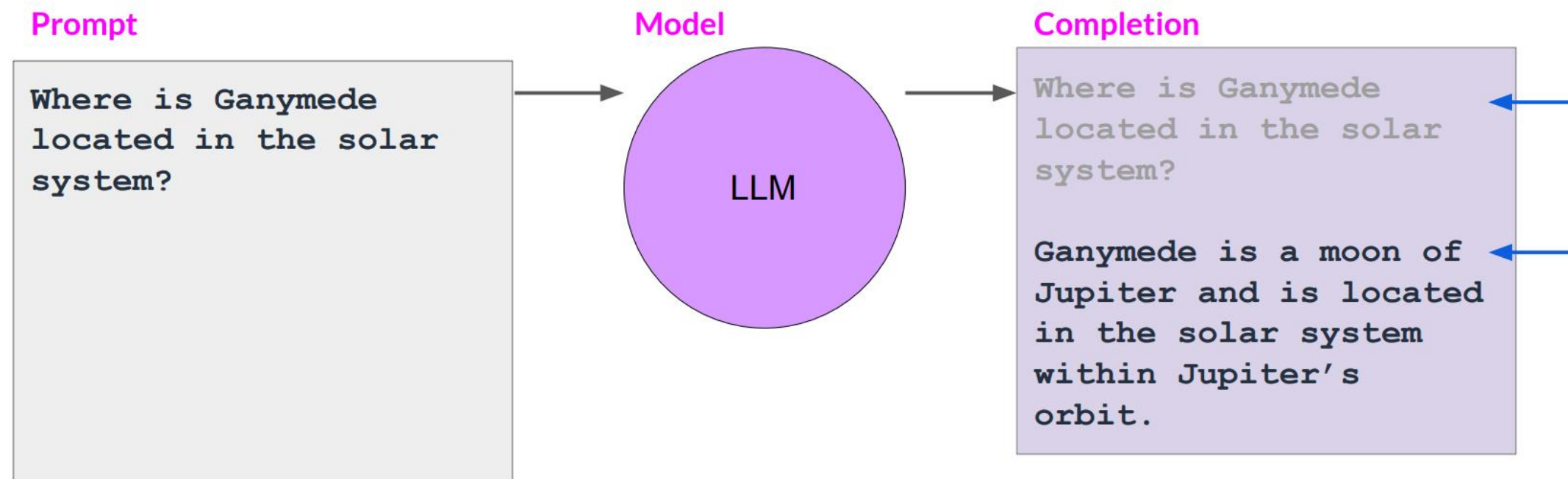
<https://creativecommons.org/licenses/by-sa/2.0/legalcode>

Source of Slides:

- DeepLearning.AI as the source of the slides.
- Generative AI :)
- Harvard Teaching and Learning Consortium

Prompts Engineering

Prompts and completions



Context window

- typically a few 1000 words.

What is a prompt?

What is a prompt?

- A prompt is a piece of text or an instruction given to an artificial intelligence model to generate a response.
- It serves as the input that guides the model's output.
- Examples:
 - "Write a poem about the moon" or "Translate this English sentence to French."
- The model then generates a relevant response based on the given prompt.

What is a Prompt?

"A photorealistic image of an astronaut riding a horse on the moon."

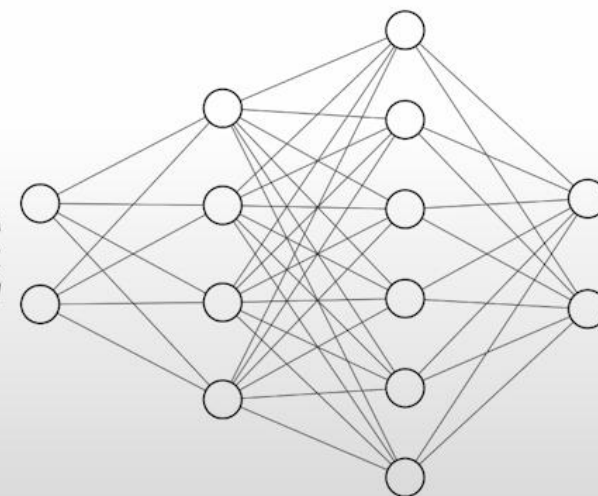


[0.24, -0.18, 0.14, 0.07, -0.03, ..., 0.23]

[0.75, 0.34, 0.11, 0.09, -0.04, ..., 0.29]

[0.63, 0.55, 0.91, 0.30, -0.01, ..., 0.44]

[0.22, 0.66, 0.94, 0.71, -0.20, ..., 0.75]



Embeddings

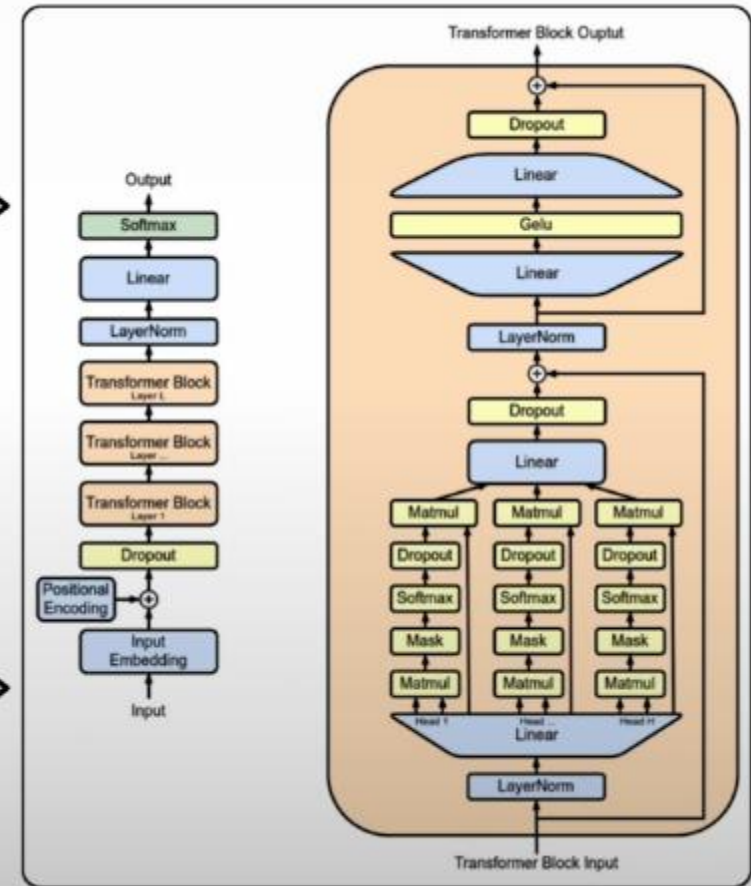


"Highlight the dog
on the left."

Embedding
Model(s)
e.g. CLIP, ALIGN

$[0.93, 0.45, -0.53, 0.72, 0.11, \dots, -0.82]$

$[0.41, 0.67, 0.23, -0.96, 0.19, \dots, -0.01]$

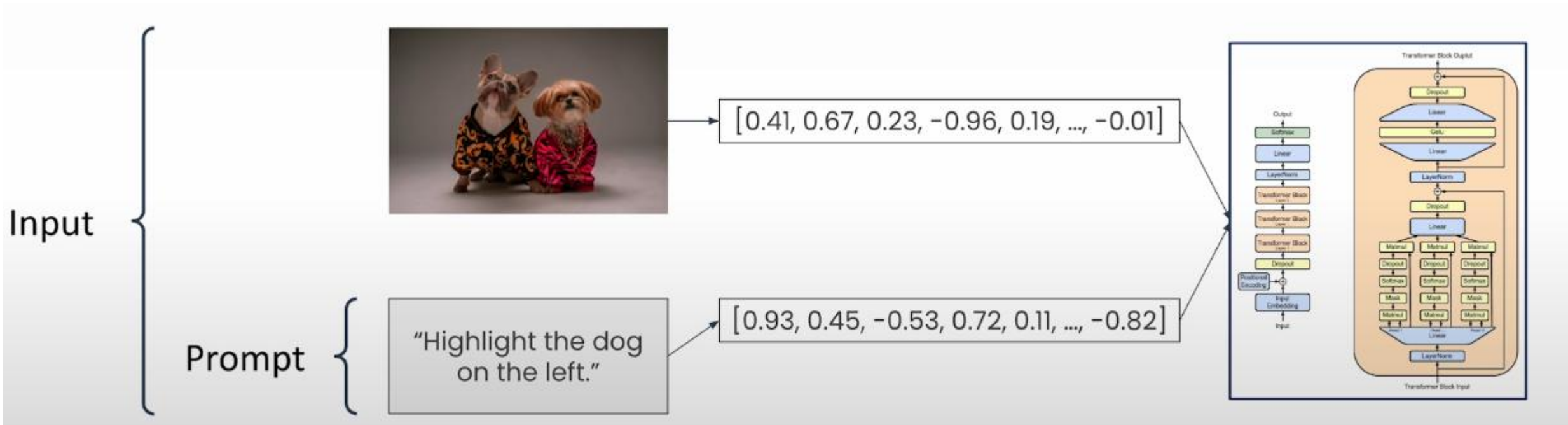


What is a visual prompt?

"Highlight the dog
on the left."



Prompt vs input



visual prompts



2. Image Segmentation

Use the Segment Anything Model (SAM) to segment images using pixel coordinates and bounding boxes as prompts.



3. Object Detection

Use OWL-ViT + SAM to detect and segment existing objects within images using natural language.



4. Image Generation

Use natural language to generate areas of an image that didn't previously exist.

Generate images

- Example generated images



A picture of a woman smiling



A futuristic city scene



A cool, happy robot

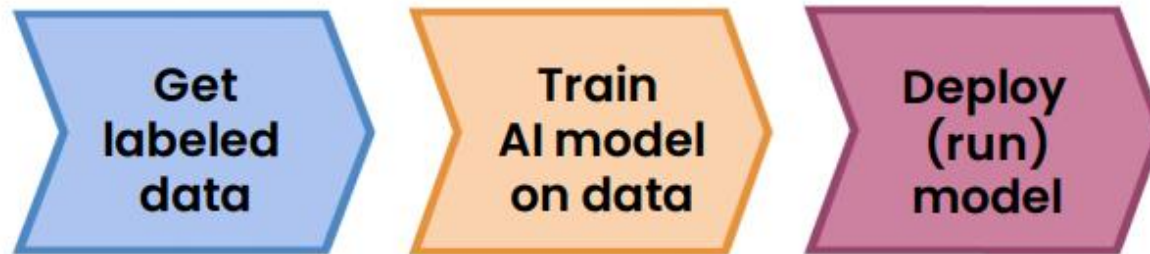
Supervised learning for restaurant reputation monitoring

Input A

→

Output B

Input (A)	Output (B)
Best soup dumplings I've ever eaten.	Positive
The colorful tablecloths made me smile!	Positive
Not worth the 3 month wait for a reservation	Negative



Input: "Best bubble tea I've ever had"

Output: Positive

Prompt-based development

```
prompt = """  
    Classify the following review  
    as having either a positive or  
    negative sentiment:  
  
    The banana pudding was really  
    tasty!  
    """
```

```
response = llm_response(prompt)  
print(response)
```

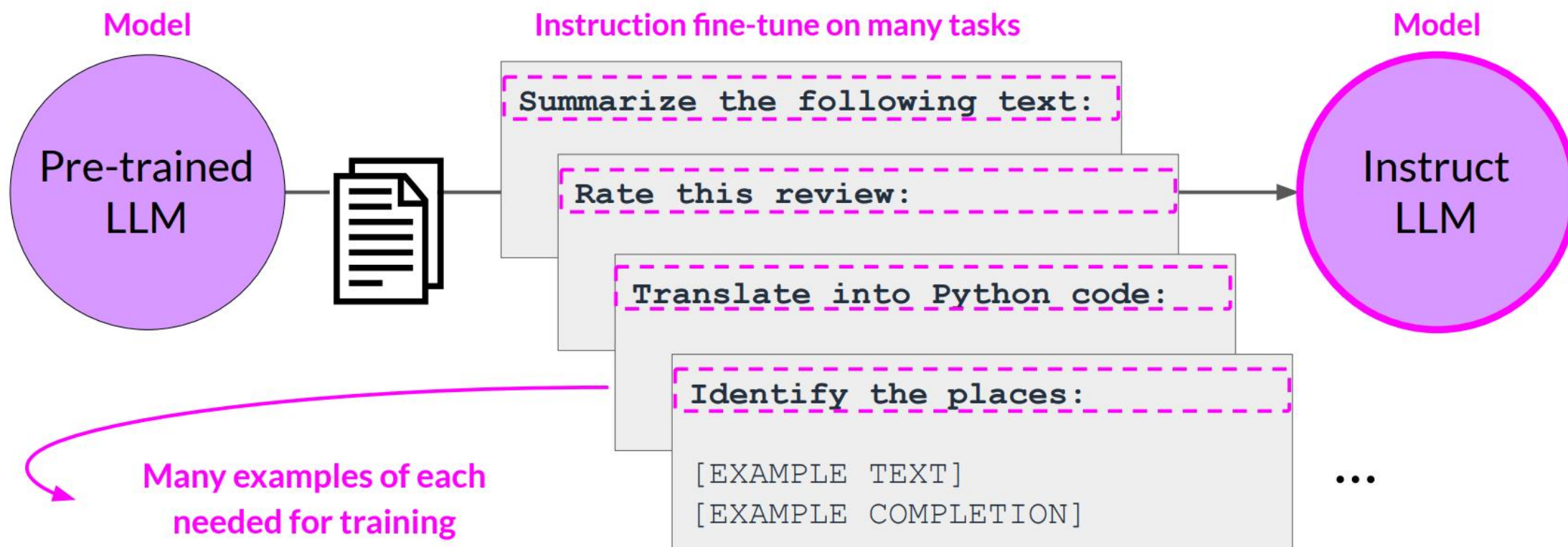
→ Instruction text

→ Review text

→ Code to call LLM

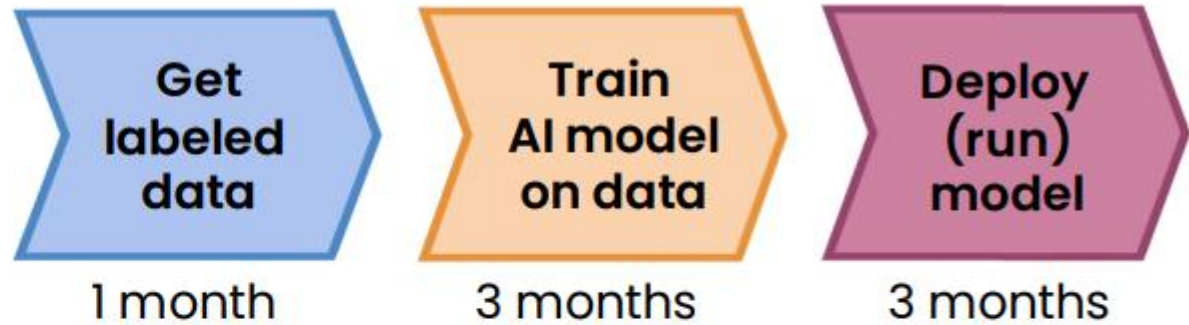
→ Code to print output

Multi-task, instruction fine-tuning

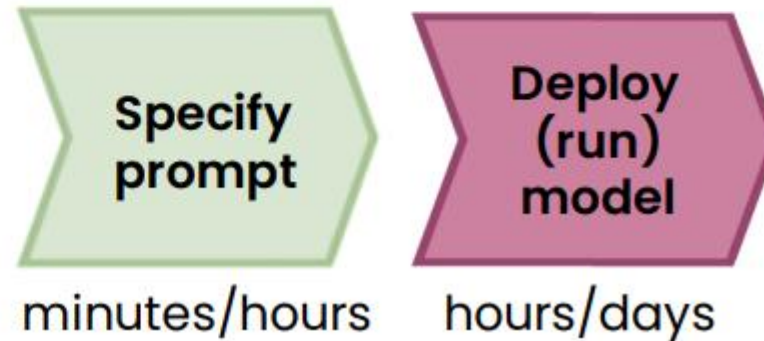


Workflow using Generative AI

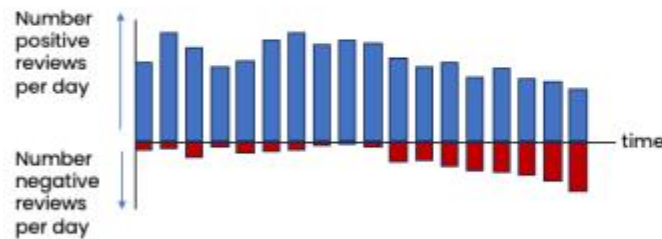
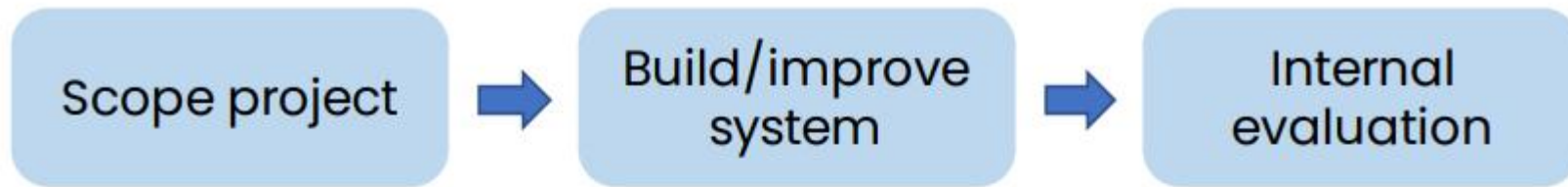
Supervised learning



Prompt-based AI



Lifecycle of a generative AI project



Initially a prototype, that we will improve over time

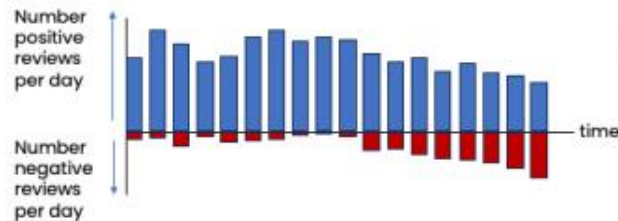
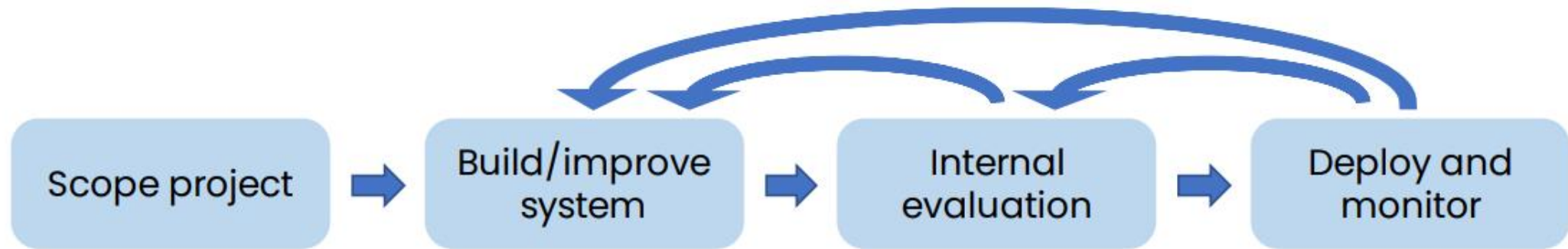


Classify the sentiment of the following review as either positive or negative:

The custard tart was amazing!

Positive

Lifecycle of a generative AI project



Initially a prototype, that we will improve over time



Classify the sentiment of the following review as either positive or negative:

My miso ramen tasted like tonkotsu ramen.

Positive

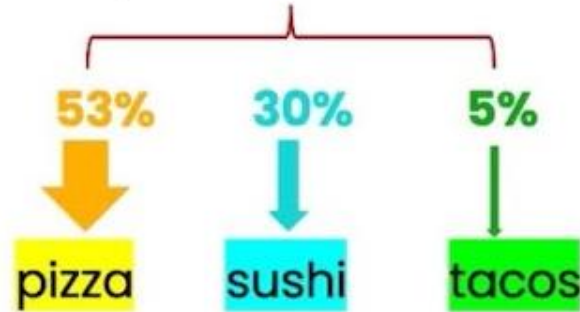
Model Temperature

Temperature

- Temperature is a crucial parameter that influences the randomness and creativity of generated text.
- It ranges from 0 to 1, with higher values leading to more diverse and imaginative outputs, while lower values result in more predictable and conservative responses.
- Essentially, adjusting the temperature allows you to control how "adventurous" or "safe" the LLM's generated content will be.
- For instance, a higher temperature value (e.g., 1.0) encourages more exploration, while a lower value (closer to 0) produces more deterministic results.

Temperature

my favorite food is



Temperature = 0

my favorite food is pizza
my favorite food is pizza
my favorite food is pizza

for tasks that require reliability, predictability

Temperature = 0.3

my favorite food is pizza
my favorite food is sushi
my favorite food is pizza

For tasks that require variety

Temperature = 0.7

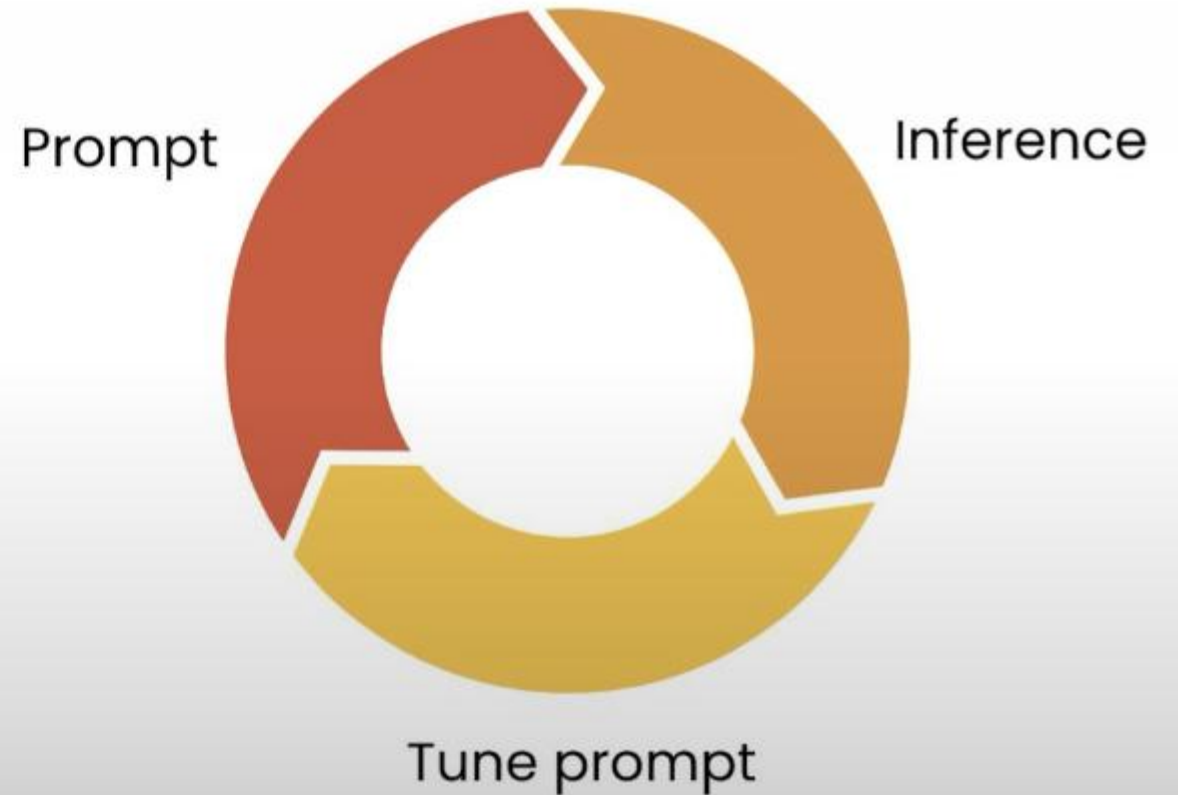
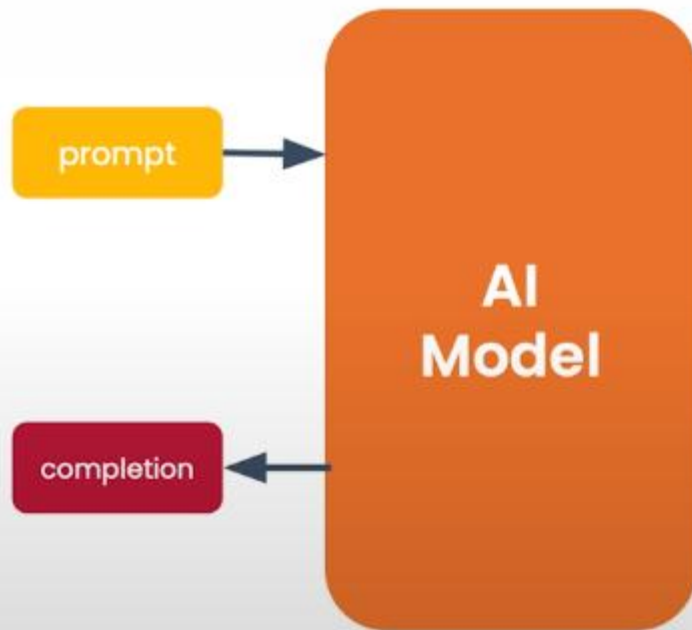
my favorite food is tacos
my favorite food is sushi
my favorite food is pizza

Prompt Engineering

What is Prompt engineering?

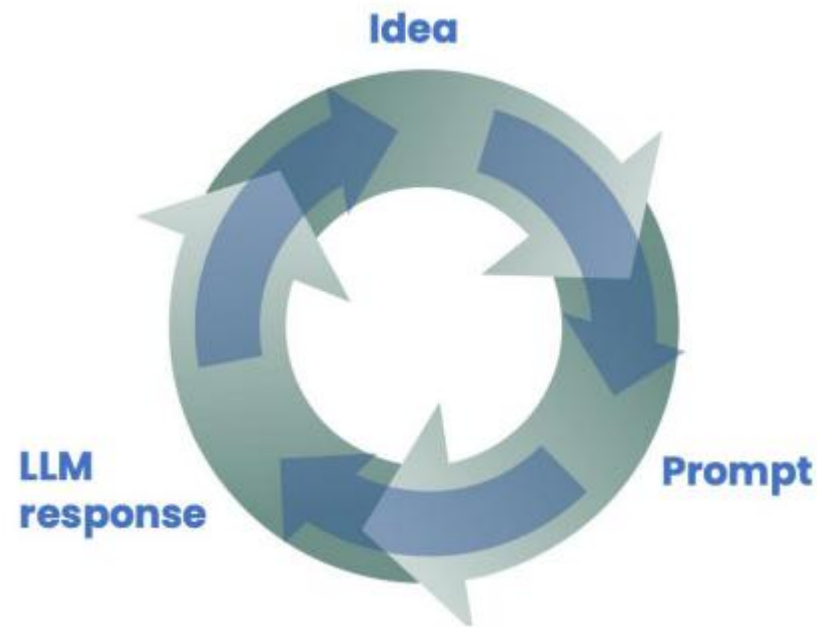
- Prompt engineering is the process of refining prompts for generative AI services to create text or images.
- It involves writing, refining, and optimizing inputs to encourage these AI systems to generate specific, high-quality outputs.
- By crafting thoughtful prompts, prompt engineers help these models better comprehend and respond to a wide range of queries, from simple to highly technical ones.
- Effective prompt engineering combines technical knowledge with a deep understanding of natural language, vocabulary, and context, resulting in optimal outputs with minimal revisions.

Prompt Engineering workflow



Tools to improve performance

- Building Generative AI is a highly empirical (experimental) process – we repeatedly find and fix mistakes



Tips for prompting

- Be detailed and specific
- Guide the model to think through its answer
- Experiment and iterate

Be detailed and specific

- Give sufficient context for LLM to complete the task
- Describe the desired task in detail

Help me write an email asking to be assigned to the legal documents project.

I'm applying for a job on the legal documents project, which will check legal documents using LLMs. I have ample experience prompting LLMs to generate accurate text in a professional tone.

Write a paragraph of text explaining why my background makes me a strong candidate to this project and advocate for my candidacy.

Guide the model to think through its answer

- Brainstorm 5 names for a new cat toy.
- Step 1: Come up with 5 fun, joyful words that relate to cats.
- Step 2: For each word, come up with a rhyming name for a toy.
- Step 3: For each toy name, add a fun, relevant emoji.

Step 1:

5 fun, joyful words
that relate to cats:

Purr
Whisker
Feline
Pounce
Meow

Step 2:

Rhyming names for a
toy based on those
words:

Purr-Twirl
Whisker-Whisper
Feline-Beeline
Pounce-Bounce
Meow-Wow

Step 3:

Add a fun, relevant
emoji for each toy
name:

Purr-Twirl 🌀
Whisker-Whisper 🤫
Feline-Beeline 🐾
Pounce-Bounce ⚽
Meow-Wow 😸

Experiment and iterate

- No perfect prompt for every person or situation!
- Instead, better to develop process for improving prompts through iteration

Help me rewrite this:
[...]



Evaluate output and
improve prompt

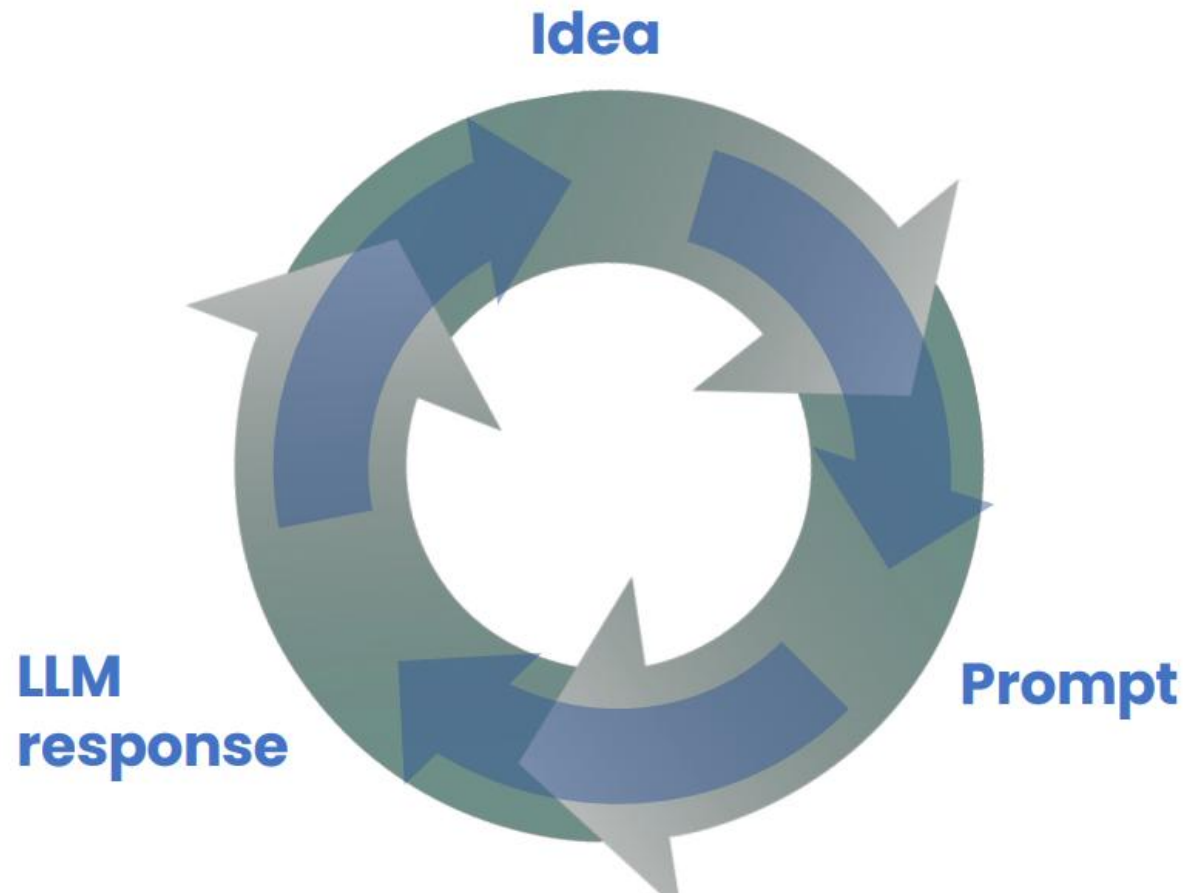
Correct any
grammatical and
spelling errors in this:
[...]



Evaluate output and
improve prompt

Correct any
grammatical and
spelling errors in this,
and rewrite in a tone
appropriate for a
professional resume:
[...]

Iteratively improving your prompt



Prompting process

- Be clear and specific in prompt
- Think about why result isn't giving desired output
- Refine your prompt
- Repeat

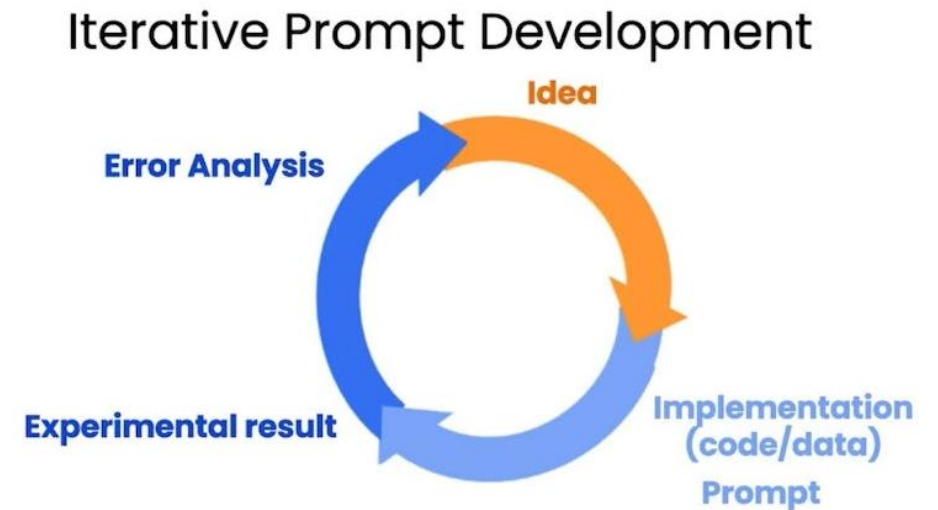
Tip: Don't overthink the initial prompt. Just try something quickly and improve it over time!

Caveats: Be careful with

- (i) confidential information
- (ii) whether you trust the LLM output

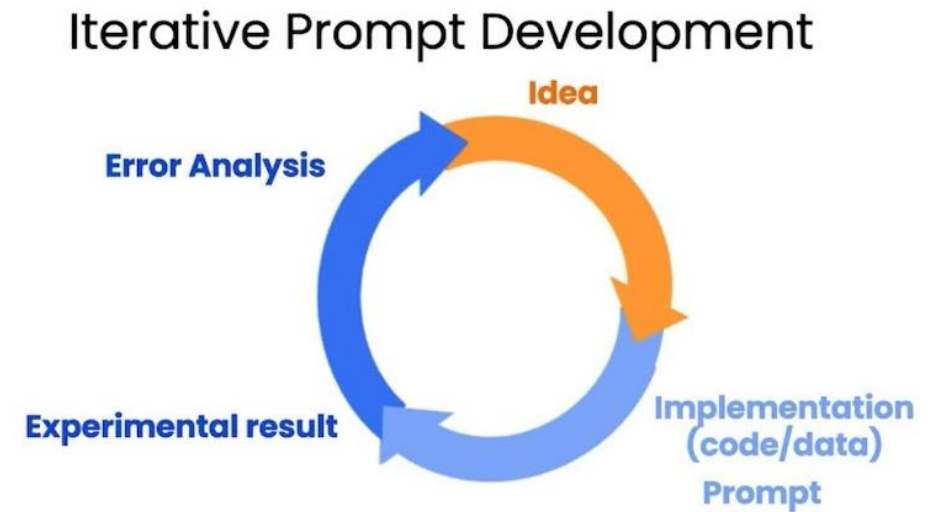
Iterative process

- try something
- analyze the result (good, bad)
- clarify instructions, give more time to think
- refine prompts with a batch of examples



Iterative process

- be clear and specific
- analyze why result is not good
- refine the idea and the prompt
- repeat



Prompt Engineering Principles

- Principles:
- Write clear and specific instructions.
 - Give the model time to “think.”
 - Iterative prompt development.
- Capabilities: Summarizing, Inferring, Transforming, Expanding.

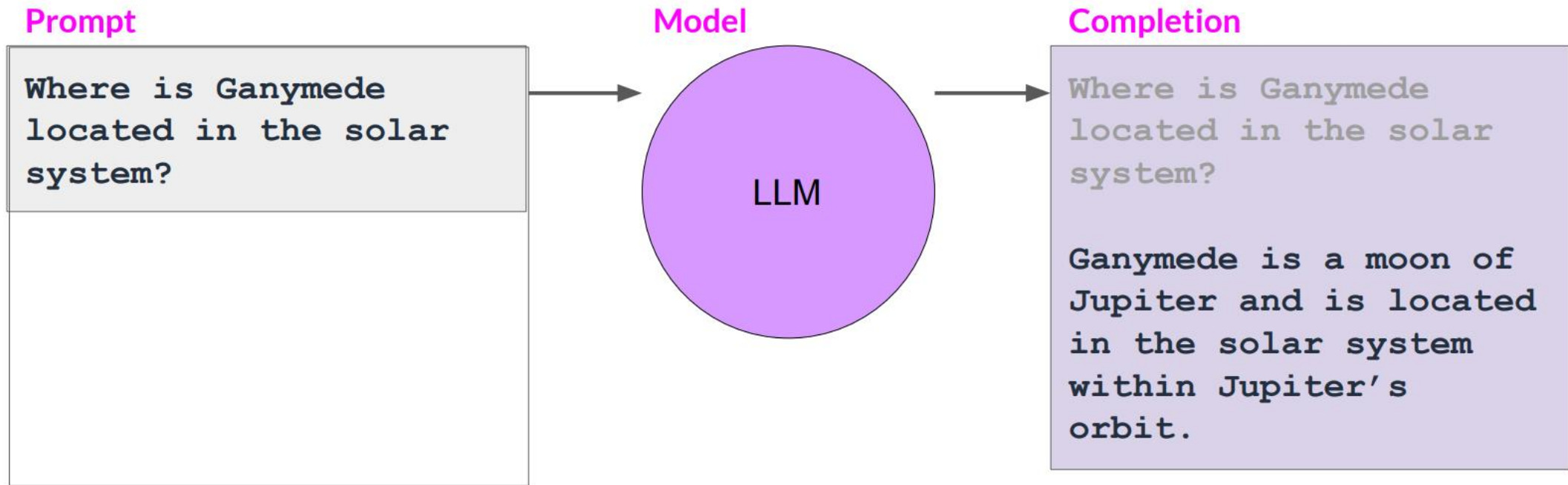
Principles of prompting

- Principle 1
 - write clear and specific instructions
 - clear \neq short
- Principle 2
 - give the model time to think

clear instructions

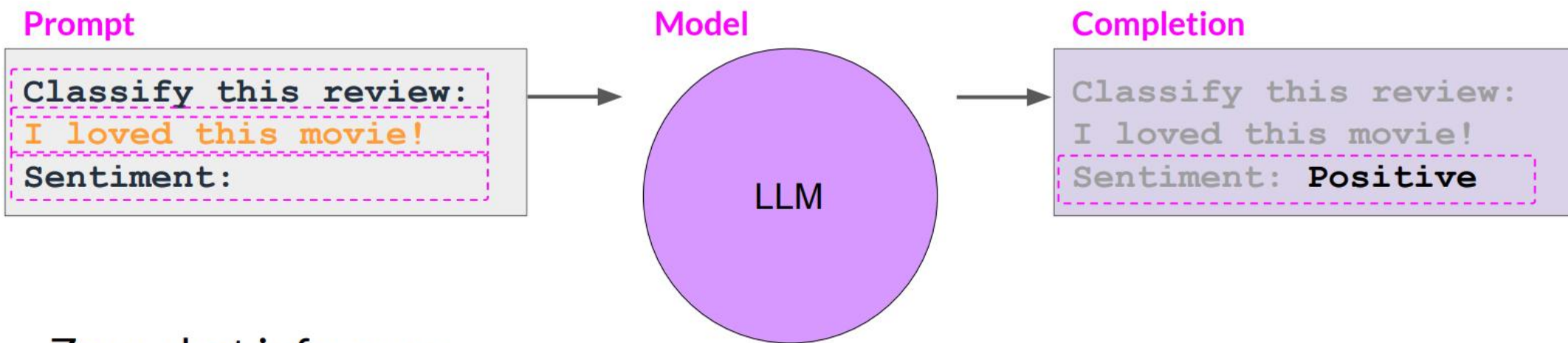
- use delimiters
 - triple quotes: `"""`
 - triple backticks: `````
 - triple dashes: `---`
 - angle brackets: `<>`

Prompting and prompt engineering (Model size)

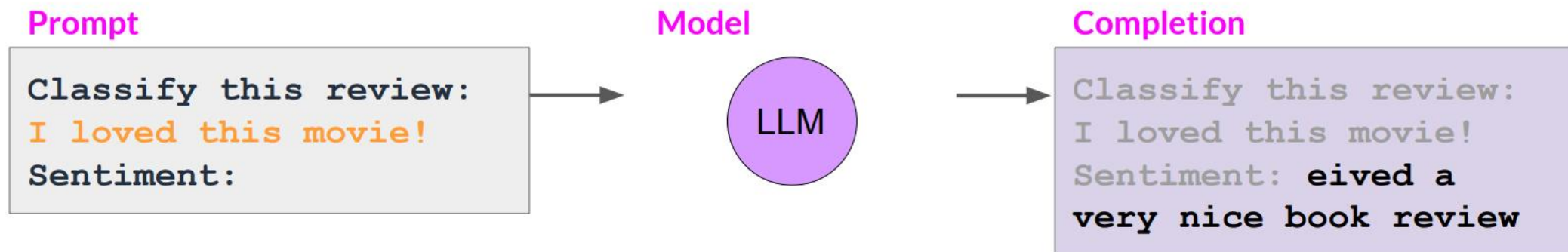


Context window: typically a few thousand words

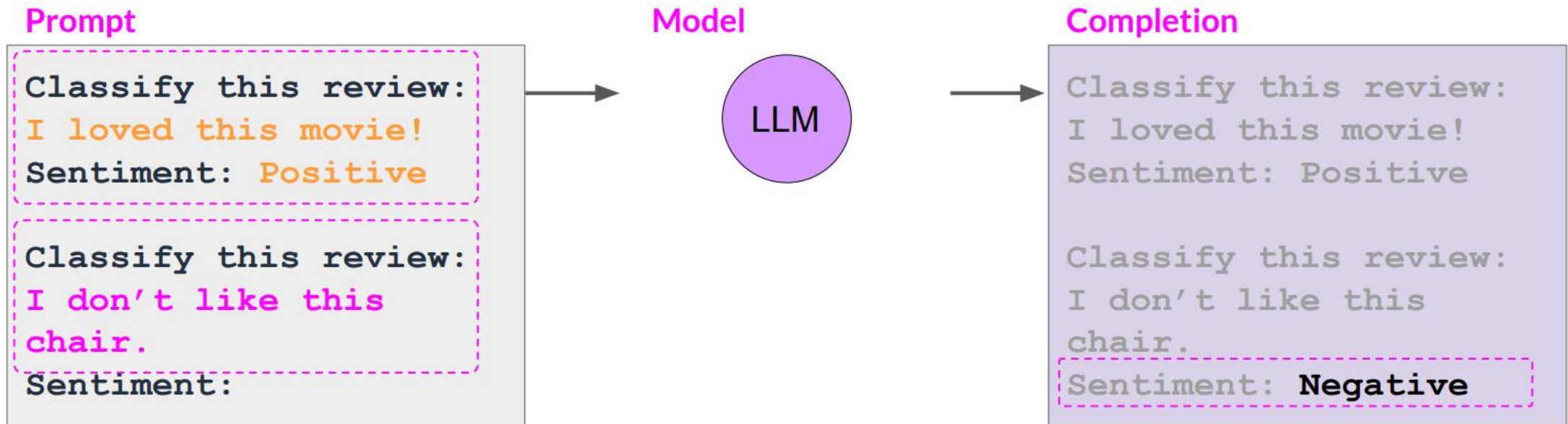
In-context learning (ICL) - zero shot inference



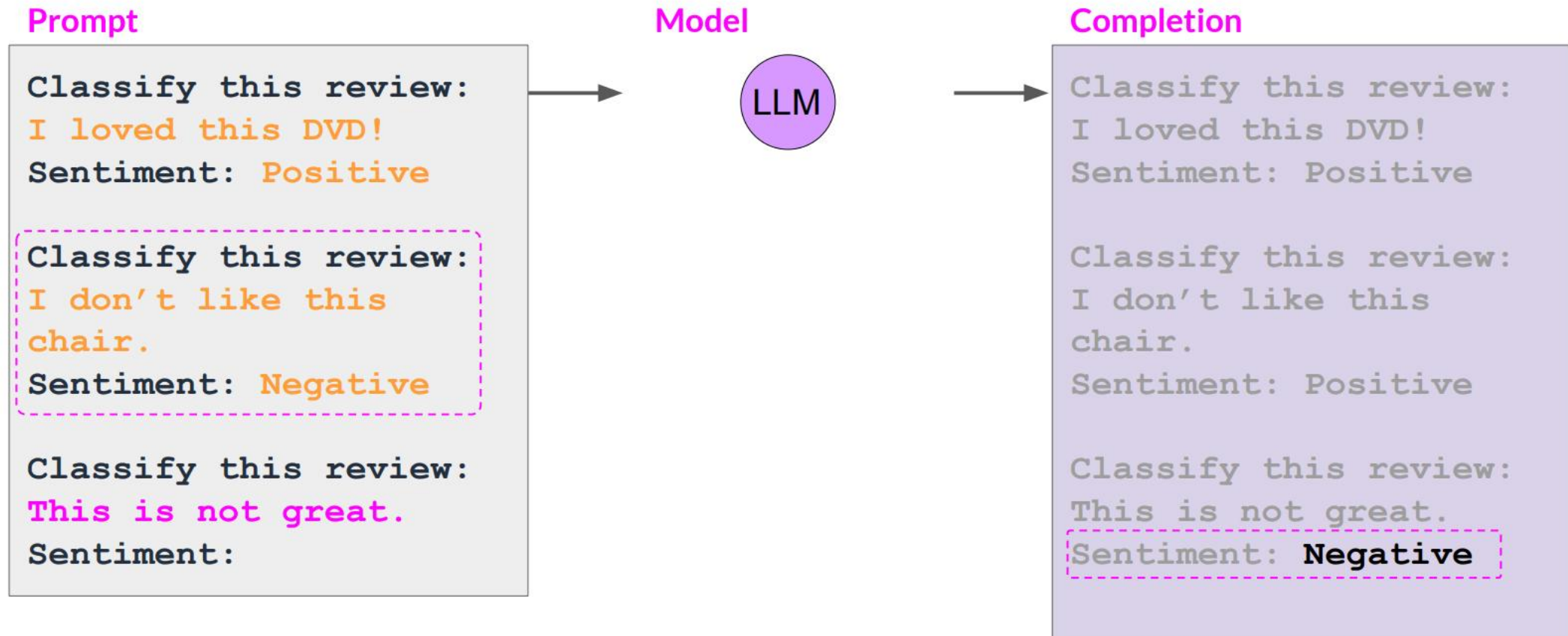
Zero-shot inference



In-context learning (ICL) - few shots inference



In-context learning (ICL) - few shots inference



Summary of in-context learning (ICL)

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 >5 or 6 examples

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

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

Classify this review:
Who would use this
product?
Sentiment:

Context Window
(few thousand words)

LLMs can struggle with complex reasoning problems

Prompt

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

Model



Completion

Q: Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

A: The answer is 11

Q: The cafeteria had 23 apples. If they used 20 to make lunch and bought 6 more, how many apples do they have?

A: The answer is 27.



Humans take a step-by-step approach to solving complex problems

Roger has 5 tennis balls. He buys 2 more cans of tennis balls. Each can has 3 tennis balls. How many tennis balls does he have now?

“Chain of thought”

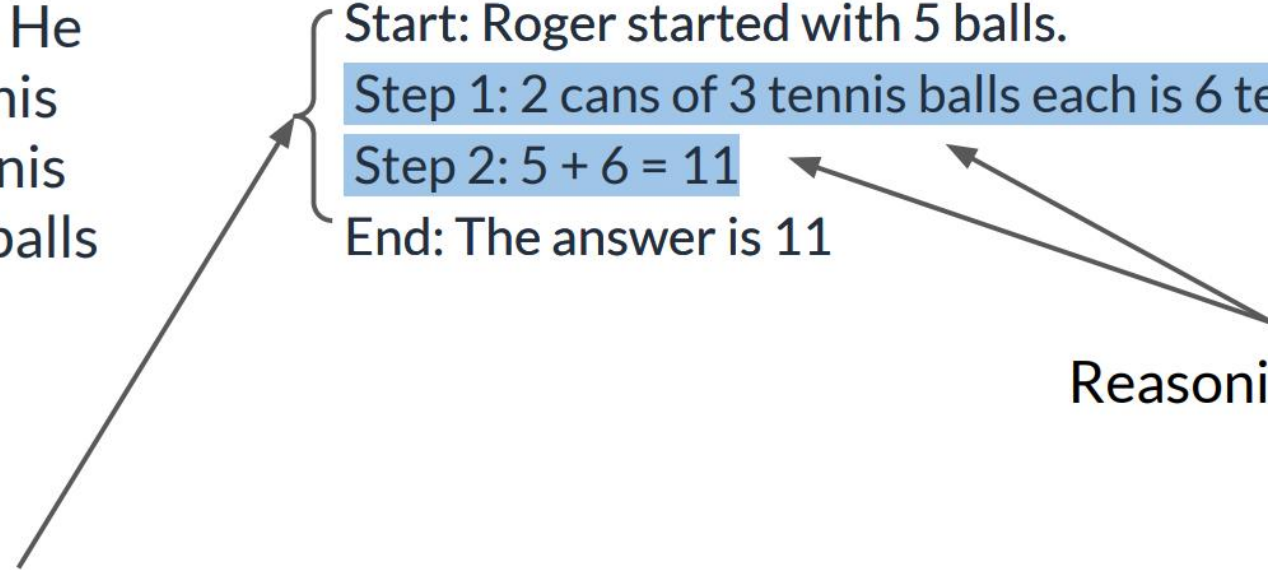
Start: Roger started with 5 balls.

Step 1: 2 cans of 3 tennis balls each is 6 tennis balls.

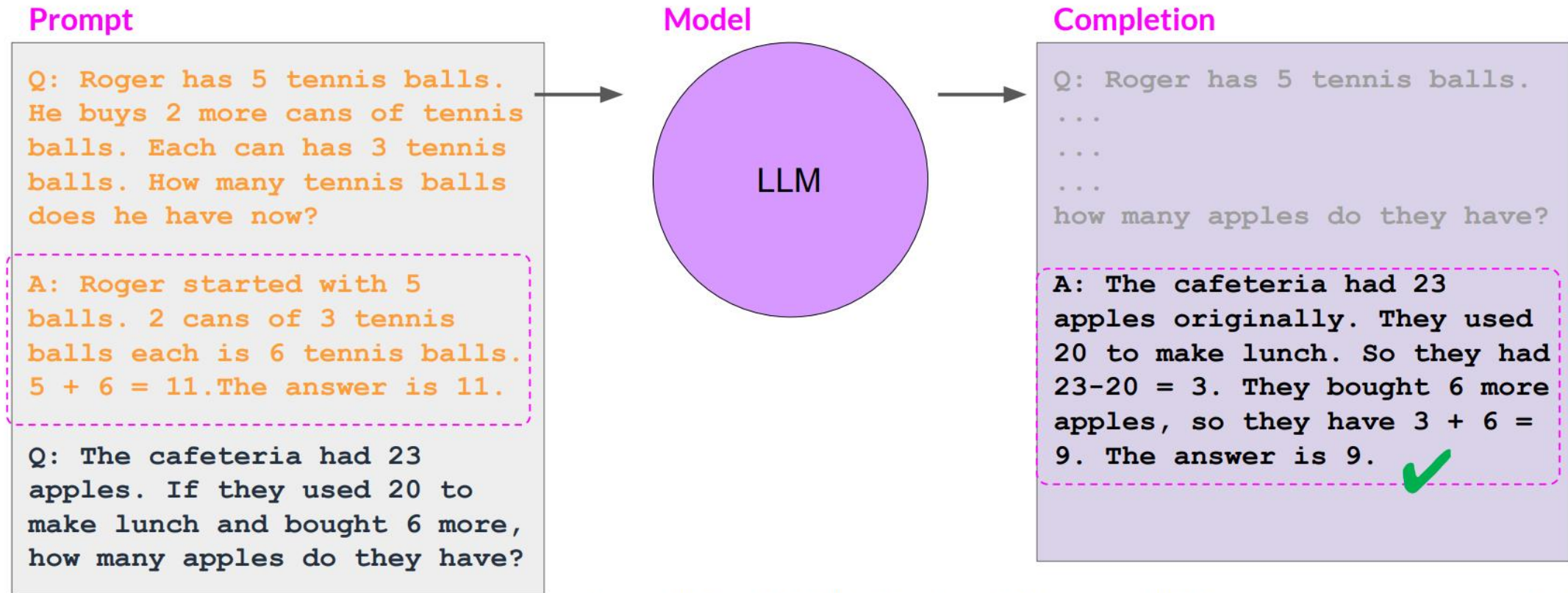
Step 2: $5 + 6 = 11$

End: The answer is 11

Reasoning steps

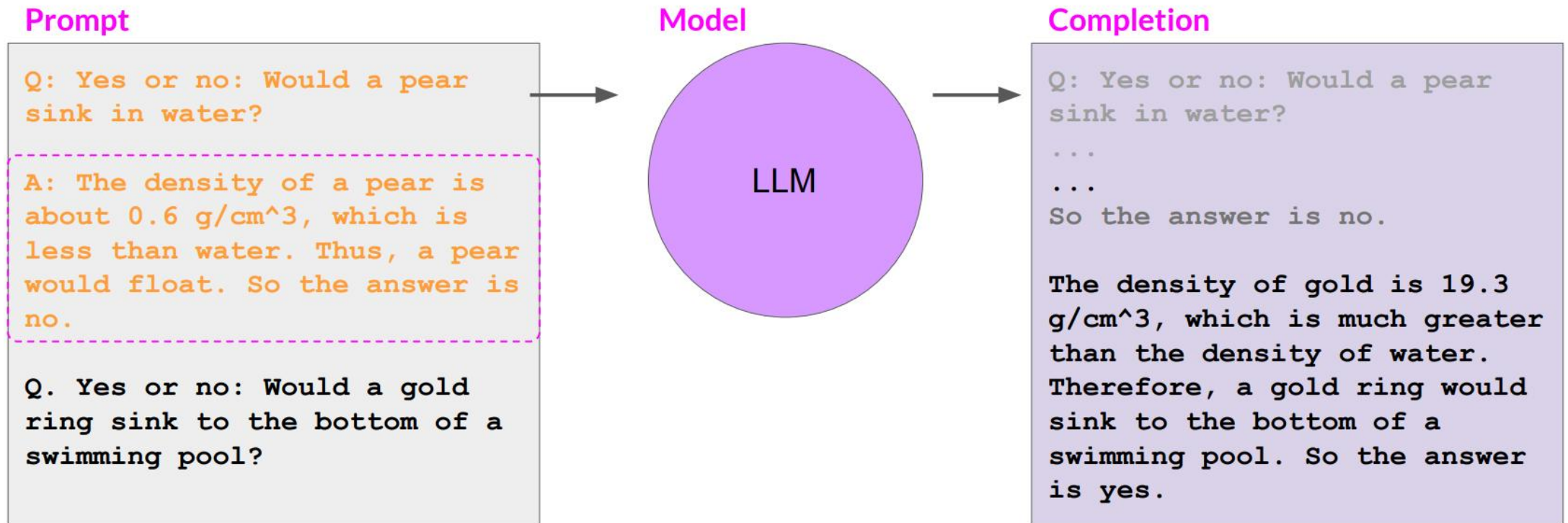


Chain-of-Thought Prompting can help LLMs reason

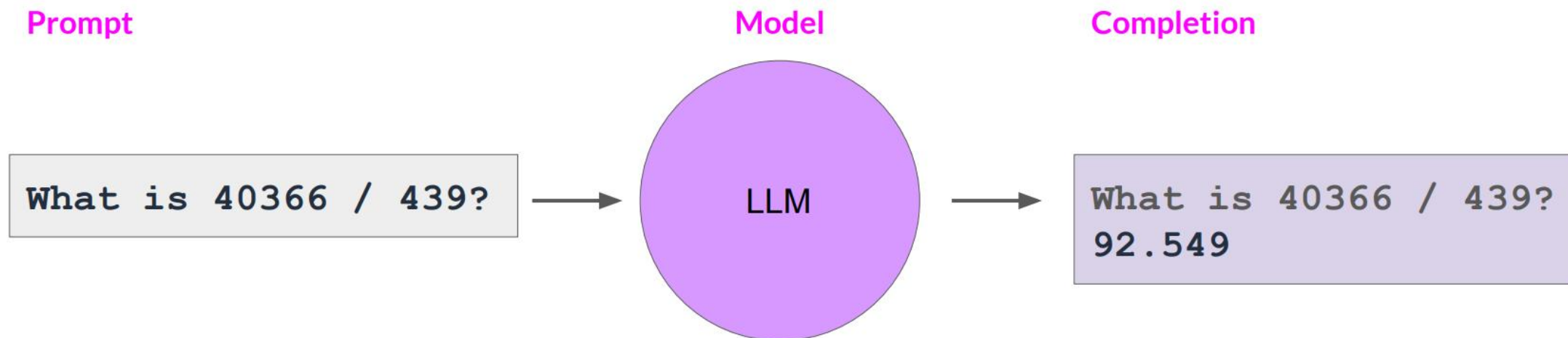


Source: Wei et al. 2022, "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models"

Chain-of-Thought Prompting can help LLMs reason



LLMs can struggle with mathematics



Give the model time to think

- Tactic1: specify steps to compete a task
 - step 1: ...
 - step 2: ...
 - step 3: ...
 - ...
 - step N:
- Tactic 2: instruct the model to work out its own solution before rushing to a conclusion

Example

Determine if the student's solution is correct or not.

Question:

I'm building a solar power installation and I need \ help working out the financials.

- Land costs \$100 / square foot
- I can buy solar panels for \$250 / square foot
- I negotiated a contract for maintenance that will cost \ me a flat \$100k per year, and an additional \$10 / square \ foot

What is the total cost for the first year of operations as a function of the number of square feet.

Student's Solution:

Let x be the size of the installation in square feet.

Costs:

1. Land cost: $100x$
2. Solar panel cost: $250x$
3. Maintenance cost: $100,000 + 100x$

Total cost: $100x + 250x + 100,000 + 100x = 450x + 100,000$

refined prompt

Your task is to determine if the student's solution \ is correct or not.

To solve the problem do the following:

- First, work out your own solution to the problem including the final total.
- Then compare your solution to the student's solution \ and evaluate if the student's solution is correct or not.

Don't decide if the student's solution is correct until you have done the problem yourself.

Use the following format:

Question:

...

question here

...

Student's solution:

...

student's solution here

...

Actual solution:

...

steps to work out the solution and your solution here

...

.....

More Examples

-

Automatic prompt

- Automatic prompt aims to automatically generate and select the appropriate instruction for the LLMs, instead of requiring the user to manually engineer a prompt.
- Automatic prompt engineering leverages AI to:
 - Generate Prompts: Create initial prompts based on a set of criteria or examples.
 - Optimize Prompts: Modify existing prompts to improve the quality of the AI's responses.
 - Refine Prompts: Iteratively adjust prompts based on feedback or desired outcomes.

Techniques in Automatic Prompt Engineering

- **Template Generation:**
 - Use predefined templates to automatically generate prompts.
 - Templates can include placeholders that are dynamically filled based on context or data.
- **Machine Learning Models:**
 - Train models to generate or refine prompts based on large datasets of successful interactions.
 - These models can learn patterns and structures that lead to effective prompts.
- **Reinforcement Learning:**
 - Use reinforcement learning to optimize prompts through trial and error.
 - The model receives feedback on the quality of responses and adjusts prompts accordingly.
- **Natural Language Processing (NLP) Algorithms:**
 - Utilize NLP techniques to analyze and improve the language and structure of prompts.
 - NLP can help in making prompts more clear, concise, and contextually relevant.

Generate Knowledge Prompt

- A "Generate Knowledge Prompt" is a type of prompt designed to elicit comprehensive and detailed information from an AI model on a specific topic.
- This type of prompt aims to generate in-depth content that contributes to understanding, learning, or expanding knowledge on a subject.

Purpose for Knowledge Prompt

- Education: To provide detailed explanations, descriptions, or discussions on academic or educational topics.
- Research: To gather comprehensive insights and information for research purposes.
- Content Creation: To generate informative content for articles, blogs, or educational materials.

Examples of Generate Knowledge Prompts

- Explain a Concept: "Explain the theory of relativity in detail, including its historical background, key principles, and implications."
- Describe a Process: "Describe the process of photosynthesis, detailing each step and its significance in the overall mechanism."
- Historical Analysis: "Analyze the causes and effects of the Industrial Revolution, highlighting the key changes it brought to society and the economy."
- Comparative Analysis: "Compare and contrast the philosophies of existentialism and nihilism, discussing their origins, main proponents, and impact on modern thought."
- Technical Explanation: "Provide a detailed explanation of blockchain technology, including its structure, functionality, and potential applications."

Structure of a Generate Knowledge Prompt

- Introduction:
 - Start with a brief introduction or context to set the stage for the detailed response.
 - Example: "In the field of renewable energy, one of the most discussed technologies is solar power."
- Main Body:
 - Ask for a detailed explanation or description, specifying any particular aspects to be covered.
 - Example: "Explain how solar panels convert sunlight into electricity, including the scientific principles involved and the components of a solar panel system."
- Conclusion:
 - Optionally, request a summary or conclusion to encapsulate the main points.
 - Example: "Summarize the benefits and challenges of using solar power as a renewable energy source."

"Tree-of-Thoughts" prompt

- The "Tree-of-Thoughts" prompt is a technique used to generate and explore multiple ideas, solutions, or perspectives on a given topic by branching out like a tree.
- This method encourages creative thinking and thorough analysis by breaking down complex issues into smaller, manageable parts and exploring various paths or options.

Key Characteristics of "Tree-of-Thoughts" prompt

- Hierarchical Structure: The prompt is organized in a hierarchical format, similar to a tree with branches, sub-branches, and leaves.
- Exploratory: Encourages exploration of different ideas and solutions.
- Analytical: Promotes detailed analysis of each branch to uncover insights and connections.

Purpose of "Tree-of-Thoughts" prompt

- Problem-Solving: To explore multiple solutions to a problem.
- Creative Thinking: To generate diverse ideas and perspectives.
- Decision-Making: To analyze different options and make informed decisions.

Example Structure


- Root: The main topic or question.
- Branches: Primary ideas or solutions.
- Sub-Branched: Detailed aspects or components of each primary idea.
- Leaves: Specific points, examples, or implications

Example of a Tree-of-Thoughts Prompt

- Prompt: "Explore the different strategies for reducing carbon emissions in urban areas."
- This detailed response demonstrates how the Tree-of-Thoughts prompt can be used to explore a topic comprehensively by branching out into various strategies, examining each one in depth, and considering their implications and challenges.

Tree-of-Thoughts Structure:

1. Root: Reducing carbon emissions in urban areas

- Branch 1: Transportation
 - Sub-Branch 1.1: Promoting public transit
 - Leaf 1.1.1: Benefits of reducing private vehicle use
 - Leaf 1.1.2: Challenges in expanding public transit infrastructure
 - Sub-Branch 1.2: Encouraging electric vehicles (EVs)
 - Leaf 1.2.1: Government incentives for EVs
 - Leaf 1.2.2: Expanding EV charging stations
- Branch 2: Buildings and Energy Efficiency
 - Sub-Branch 2.1: Green building standards
 - Leaf 2.1.1: LEED certification and its impact
 - Leaf 2.1.2: Retrofitting existing buildings
 - Sub-Branch 2.2: Renewable energy adoption
 - Leaf 2.2.1: Solar panel installations on rooftops
 - Leaf 2.2.2: Incentives for using renewable energy sources
- Branch 3: Urban Planning
 - Sub-Branch 3.1: Creating green spaces
 - Leaf 3.1.1: Urban forests and parks
 - Leaf 3.1.2: Green roofs and walls
 - Sub-Branch 3.2: Zoning laws and policies
 - Leaf 3.2.1: Mixed-use development 

Active Prompt

- An "Active Prompt" is a technique in prompt engineering designed to actively engage the AI in a more dynamic and iterative process.
- This approach encourages the AI to not only provide information but also to ask questions, seek clarification, and refine its responses based on additional input or context.
- The goal is to create a more interactive and collaborative dialogue between the AI and the user, leading to more accurate, relevant, and insightful responses.

Active Prompt

- Key Characteristics

- Interactive: The AI actively engages with the user by asking clarifying questions or seeking more details.
- Iterative: The response is refined through a series of interactions, improving accuracy and relevance.
- Dynamic: The prompt evolves based on the ongoing dialogue, adapting to new information or context.

Example Structure of Active Prompt

- Initial Prompt: The user provides an initial query or request.
- AI Clarification: The AI asks clarifying questions or requests more information.
- User Response: The user provides additional details or context.
- Refined Prompt: The AI refines its response based on the new information.
- Iterative Process: Steps 2-4 are repeated as needed to achieve a high-quality response.

Example of an Active Prompt Interaction

- Initial Prompt:
 - "Tell me about the impact of climate change on agriculture."
- AI Clarification:
 - "Can you specify which aspect of agriculture you are interested in? For example, crop yields, soil health, or livestock?"
- User Response:
 - "I'm particularly interested in how climate change affects crop yields."
- Refined Prompt:
 - "Explain how climate change impacts crop yields, considering factors like temperature changes, precipitation patterns, and extreme weather events."
-

Active Prompt: Another example

I do not speak Italian. Help me to fill this form in Italian. Please ask me questions in English and I will respond, after you complete questioning for all the required information, you fill the form.

____ Il/La sottoscritto/a _____ nato/a a _____
 _____ (prov. di _____) il _____ e residente in
 _____ (prov. di _____) c.a.p. _____ via
 _____ n. _____ PEC _____, email _____

Active Prompt: AI response

Sure, I'll help you fill out this form in Italian. I'll ask you a few questions to gather all the necessary information. Please provide the answers, and then I'll fill out the form for you.

What is your full name?

What is your place of birth (city and province)?

What is your date of birth?

What is your current residence (city and province)?

What is your postal code (c.a.p.)?

What is your street address and house number?

What is your PEC (Certified Email Address) if you have one? If not, you can provide your regular email address.

Active Prompt: My response

Here are the answers:

What is your full name?

Motaz SAAD

What is your place of birth (city and province)?

.....

Active Prompt: AI response

Thank you for providing the information. Here is the completed form in Italian:

____ Il/La sottoscritto/a _ Motaz SAAD _ nato/a a Tripoli (prov. di Libia) il 15/10/1980 e residente in (prov. di Egitto) c.a.p. 6364014 via 163 n. _____ PEC _____, email motaz.saad@gmail.com

Directional Stimulus Prompt

- A "Directional Stimulus Prompt" is a type of prompt used to guide an AI model's response in a specific direction or focus.
- Unlike broad prompts that allow for wide-ranging responses, directional stimulus prompts provide a clear direction or emphasis on certain aspects of the response, aiming to elicit focused and targeted information.
- Purpose
 - Precision: To guide the AI model towards providing detailed and specific information.
 - Clarity: To ensure the response is aligned with the user's intent or query.
 - Efficiency: To streamline the generation of relevant content or information.

Example of a Directional Stimulus Prompt

- Broad Prompt:
 - "Explain the concept of artificial intelligence."
- Directional Stimulus Prompt:
 - "Explain the concept of artificial intelligence, focusing on its applications in healthcare."
- AI Response:
 - "Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to perform tasks that typically require human intelligence, such as learning, reasoning, problem-solving, and decision-making. In healthcare, AI is revolutionizing the industry by enabling advanced diagnostics, personalized treatment plans, and predictive analytics."

Q&A

