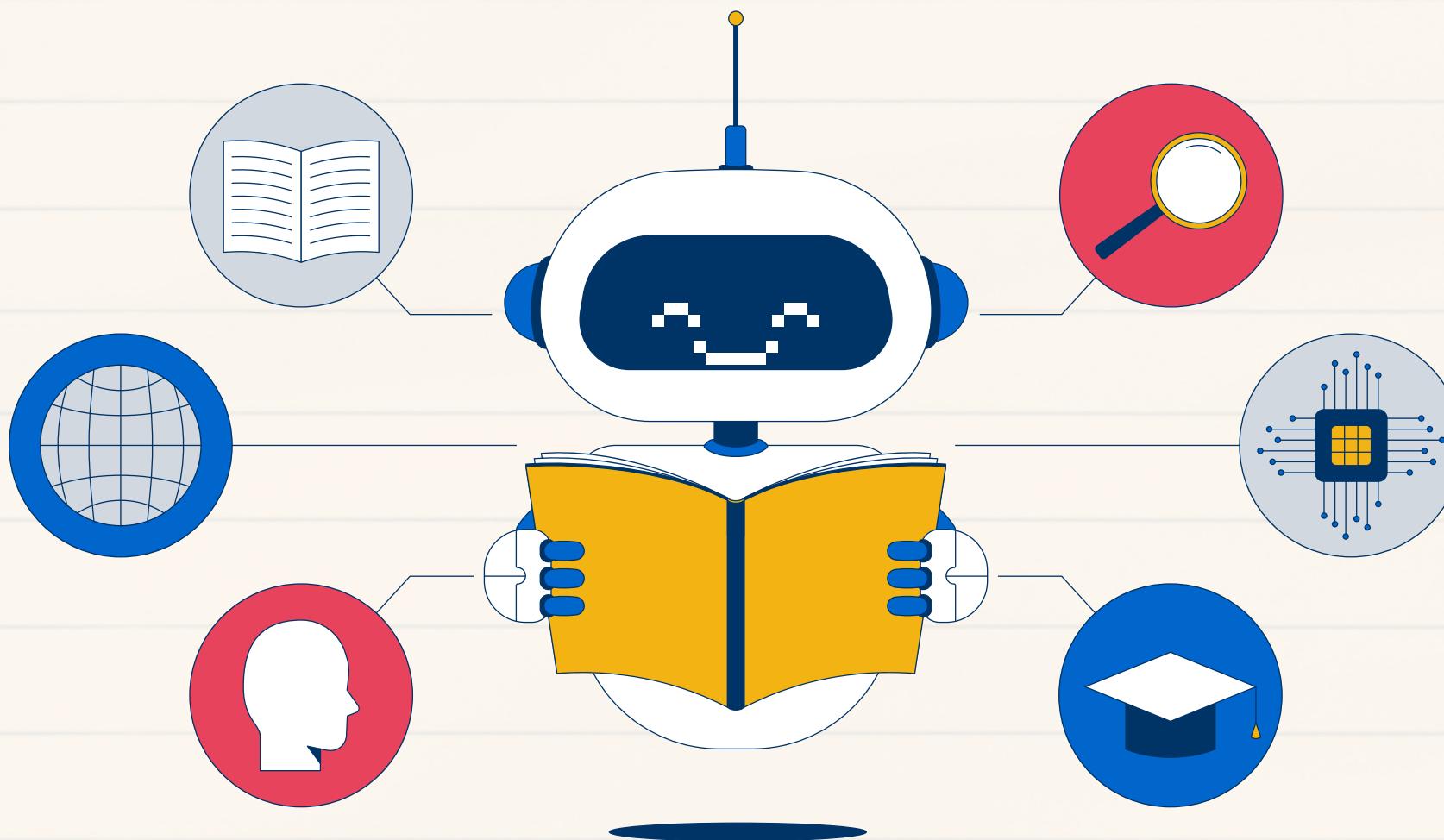


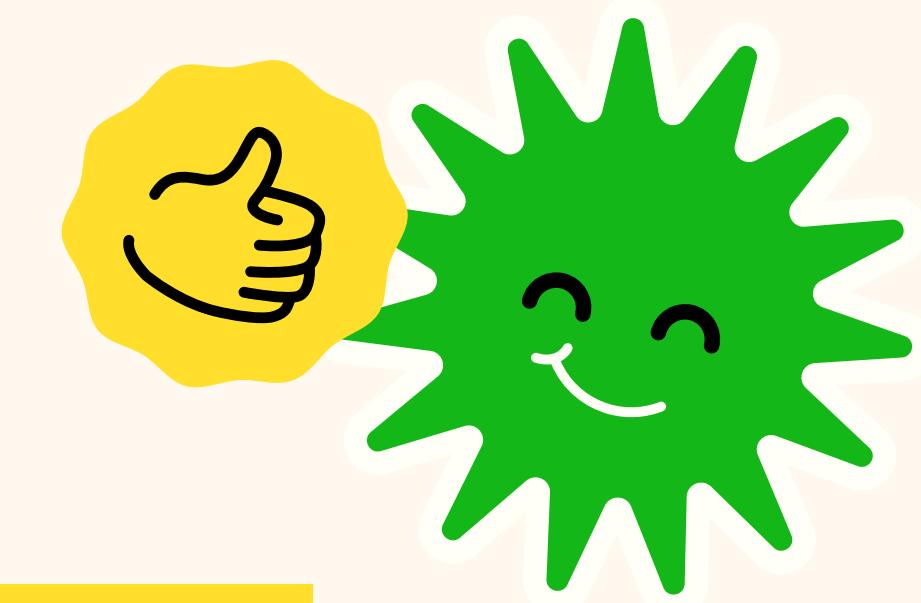
AI literacy



draft

TODAY'S AGENDA

Week 2: AI Literacy



7:20 - 7:35

7:35 - 7:55

7:55 - 8:15

8:15 - 8:30

8:30 - 8:45

8:45 - 9:30

9:30 - 9:35

Intro to AI

Activity 1

Intro to AI chatbots

Activity 2

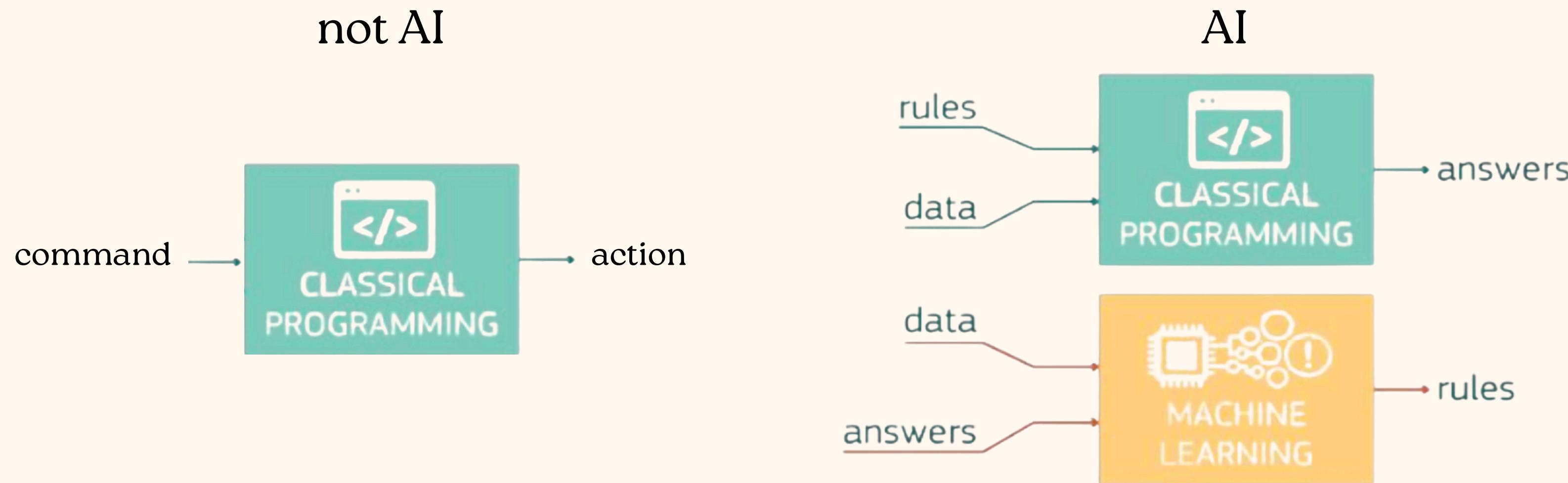
Break

Building Marty - legs

Assessment and reflection

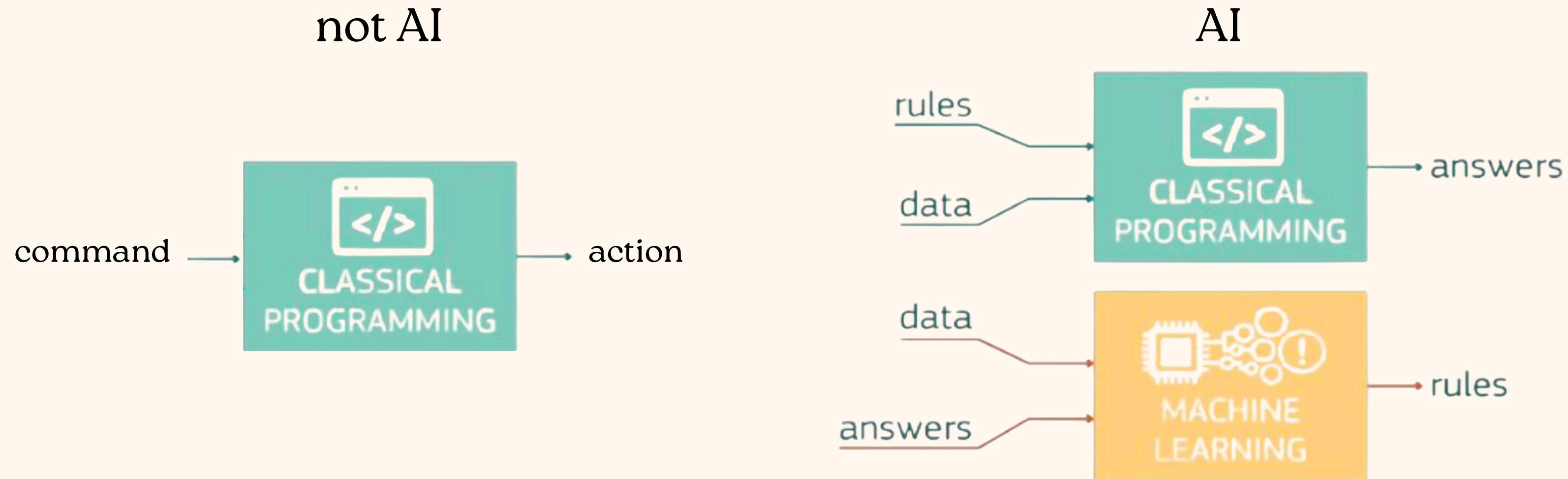


Artificial intelligence (AI) is a large domain in software engineering and is “defined as machine intelligence or intelligence demonstrated by machines” (Belipetrev et al. 2020)



What are examples of technology not using AI?

Artificial intelligence (AI) is a large domain in software engineering and is “defined as machine intelligence or intelligence demonstrated by machines” (Belipetrev et al. 2020)



What are examples of technology not using AI?
television, regular cars, remote controlled toy cars, regular video games, etc.

Machine learning (ML) is a field of study in artificial intelligence concerned with the development and study of algorithms that can learn from data and generalise to new data, and thus perform tasks without the use of explicit rules.

*Types of learning
(on what data)*

With labels :
supervised learning

Without labels :
unsupervised learning

By trial and error with rewards :
reinforcement learning

*Ways of learning
(with which model)*

Classic, Non-neural
algorithms

Neural networks

*Outcomes of learning
(what can it do)*

Classification / Prediction-
labels or projects

Generation (GenAI) - creates

Pursue of a goal (AI agents)-
decides and acts

Machine learning (ML) is a field of study in artificial intelligence concerned with the development and study of algorithms that can learn from data and generalise to new data, and without the use of explicit rules.

Types of learning (on what data)

With labels :
supervised learning

Without labels :
unsupervised learning

By trial and error with rewards :
reinforcement learning

Ways of learning (with which model)

Classic, Non-neural
algorithms

Neural networks



Classification / **Prediction**-
labels or projects

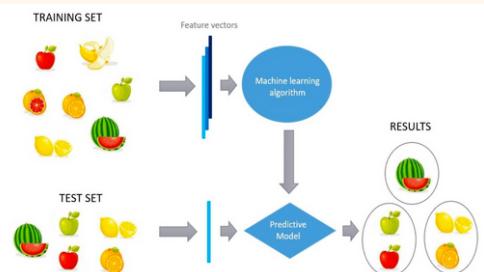
Generation (GenAI) - creates

Pursue of a goal (AI agents)-
decides and acts

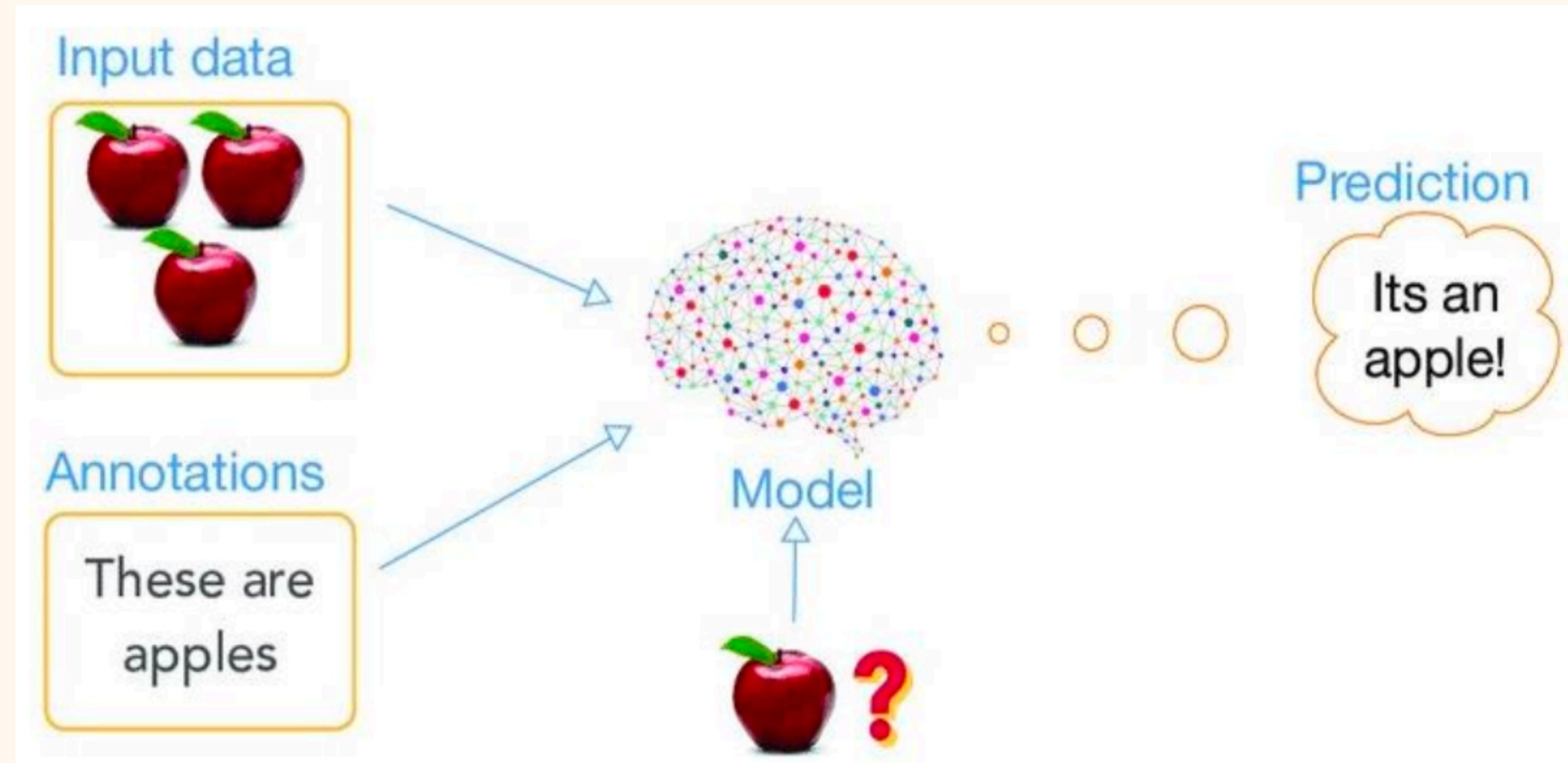
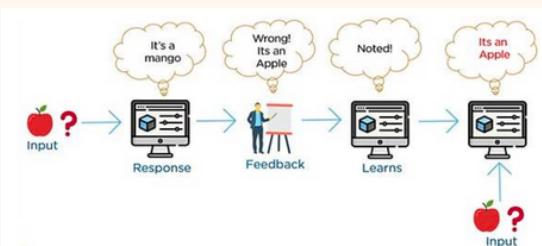
Types of machine learning : Supervised Learning

Supervised learning
(discriminative) – uses
classified/labeled data

Unsupervised learning
(generative) – uses
unclassified / unlabeled data

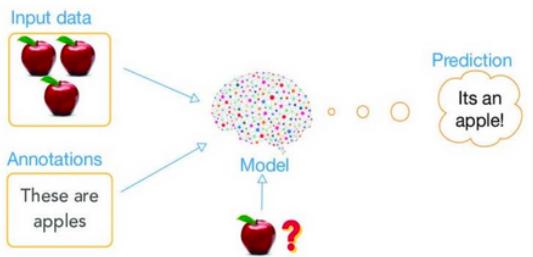


Reinforcement learning –
No data required, learns
from the experience, etc.

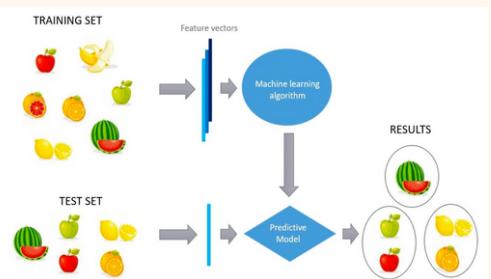


Types of machine learning: Supervised Learning

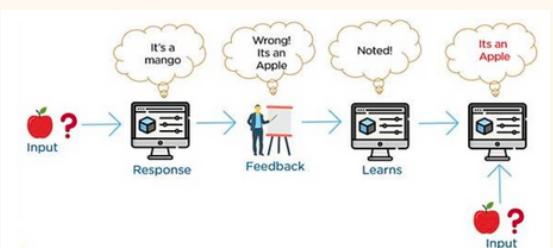
Supervised learning
(discriminative) – uses
classified/labeled data



Unsupervised learning
(generative) – uses
unclassified / unlabeled data

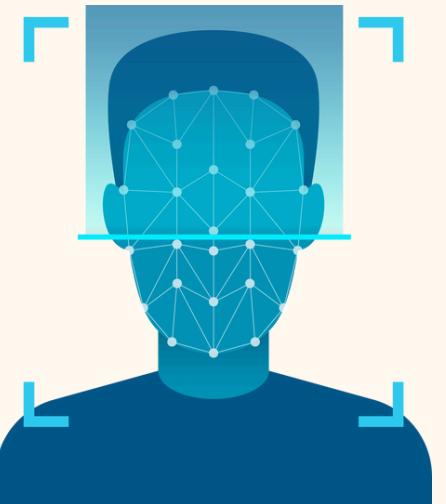


Reinforcement learning –
No data required, learns
from the experience, etc.



Examples:

Face recognition



When your phone automatically suggests tagging friends in photos, it uses supervised learning trained on photos you've previously labeled with people's names.

Music recommendation



When Spotify suggests songs you might like, it uses supervised learning trained on data like "User A liked songs X, Y, Z" to predict what new songs you'll enjoy.

Content filtering

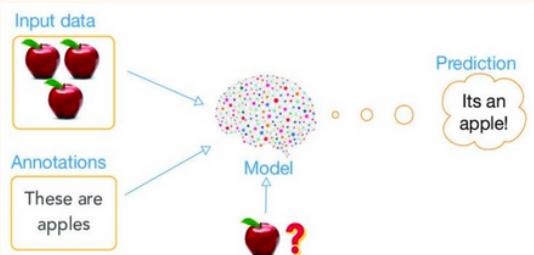


Apps, such as WhatsApp, learn to identify spam messages by training on thousands of messages that were labeled as "spam" or "not spam" by users and moderators.

Types of machine learning: Unsupervised Learning

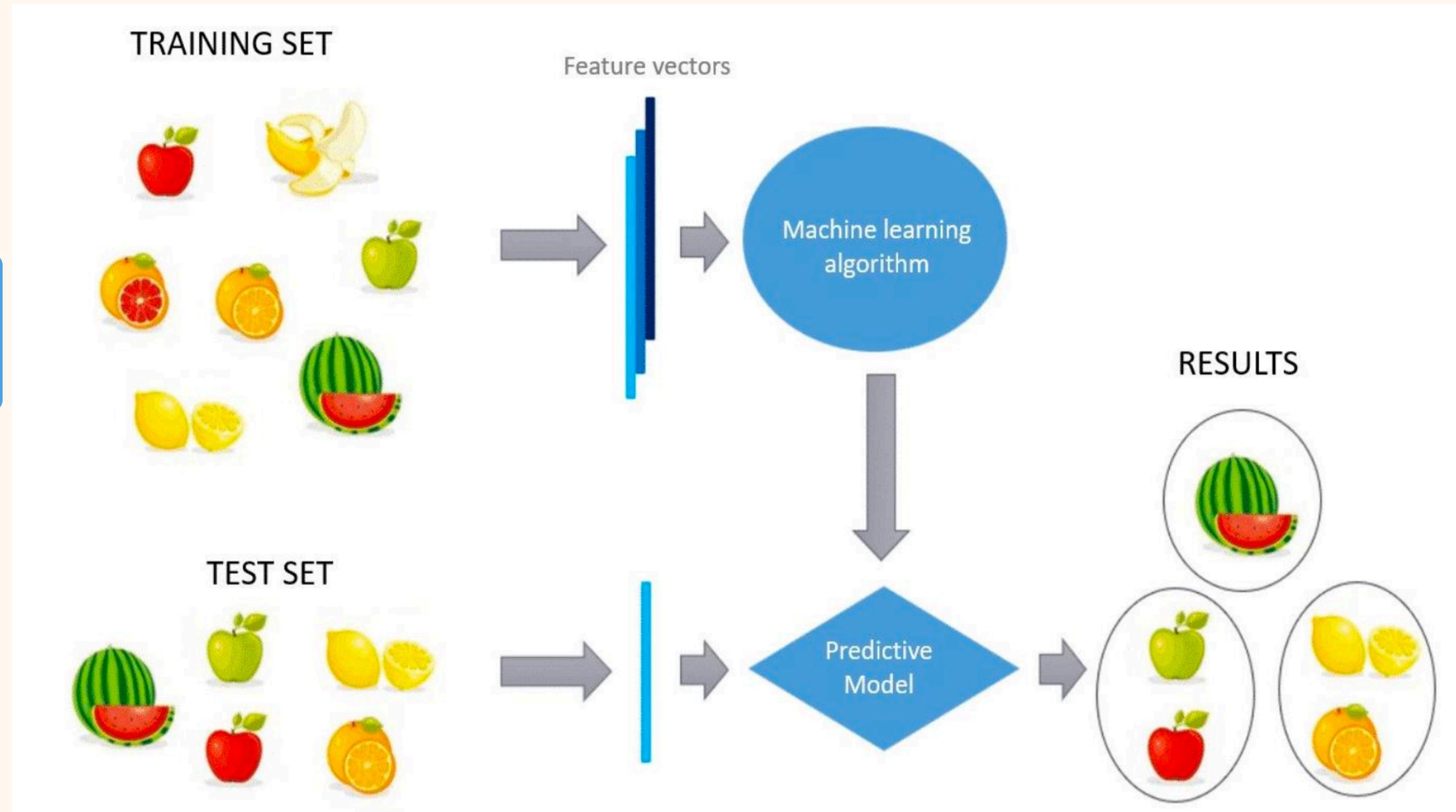
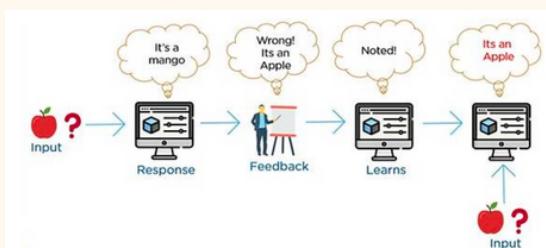
Supervised learning

(discriminative) – uses classified/labeled data



Unsupervised learning (generative) – uses unclassified / unlabeled data

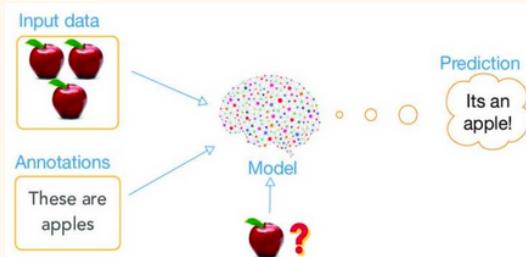
Reinforcement learning –
No data required, learns from the experience, etc.



Types of machine learning: : Unsupervised Learning

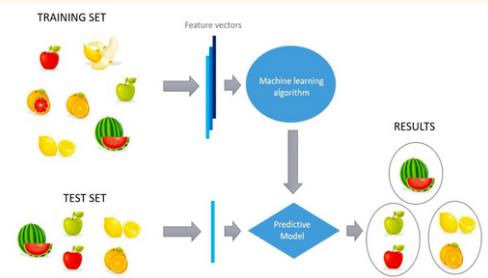
Supervised learning

(discriminative) – uses classified/labeled data

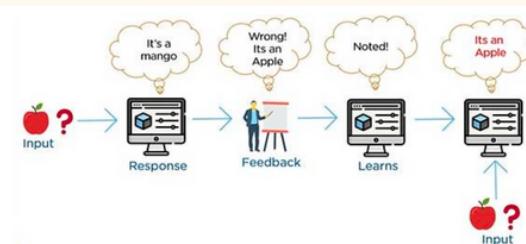


Unsupervised learning

(generative) – uses unclassified / unlabeled data



Reinforcement learning –
No data required, learns from the experience, etc.

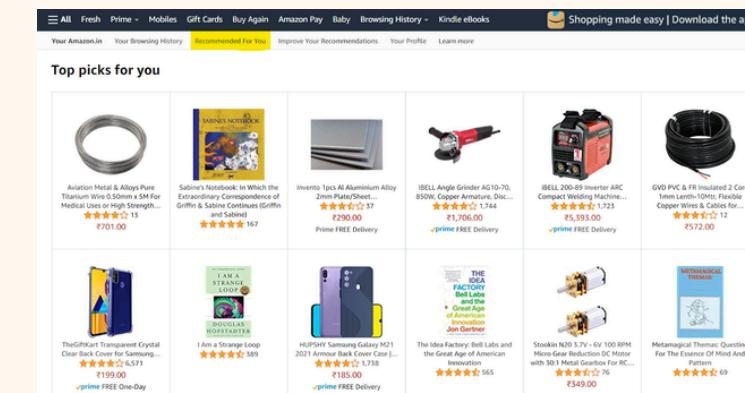


Examples:

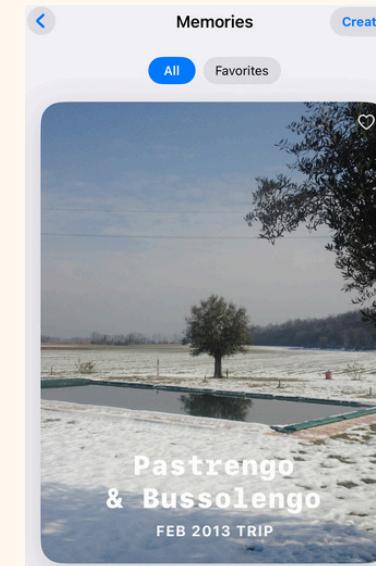
Content recommendation



Shopping recommendation



Automatic photo albums



TikTok groups users with similar viewing patterns together without being told what the groups should be. It discovers that some users love cooking videos, others prefer dance content, and others watch gaming clips - all without anyone labeling these categories beforehand.

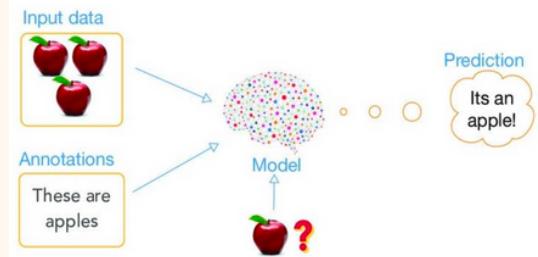
Amazon's "Customers who bought this also bought" discovers hidden patterns in purchasing behavior. It might find that people who buy certain gaming keyboards also tend to buy specific mouse pads, without anyone teaching it these relationships.

When Google Photos, or Apple Photos create albums like "Trip to Berlin" or "School friends," they are finding patterns in when and where photos were taken, and who appears in them, without being told these are meaningful groups.

Types of machine learning: Reinforcement learning

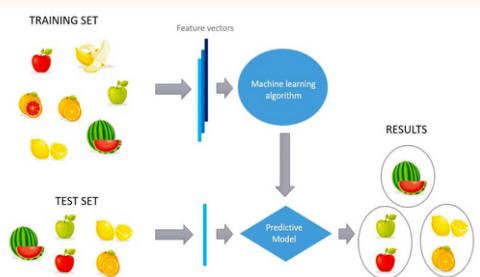
Supervised learning

(discriminative) – uses classified/labeled data

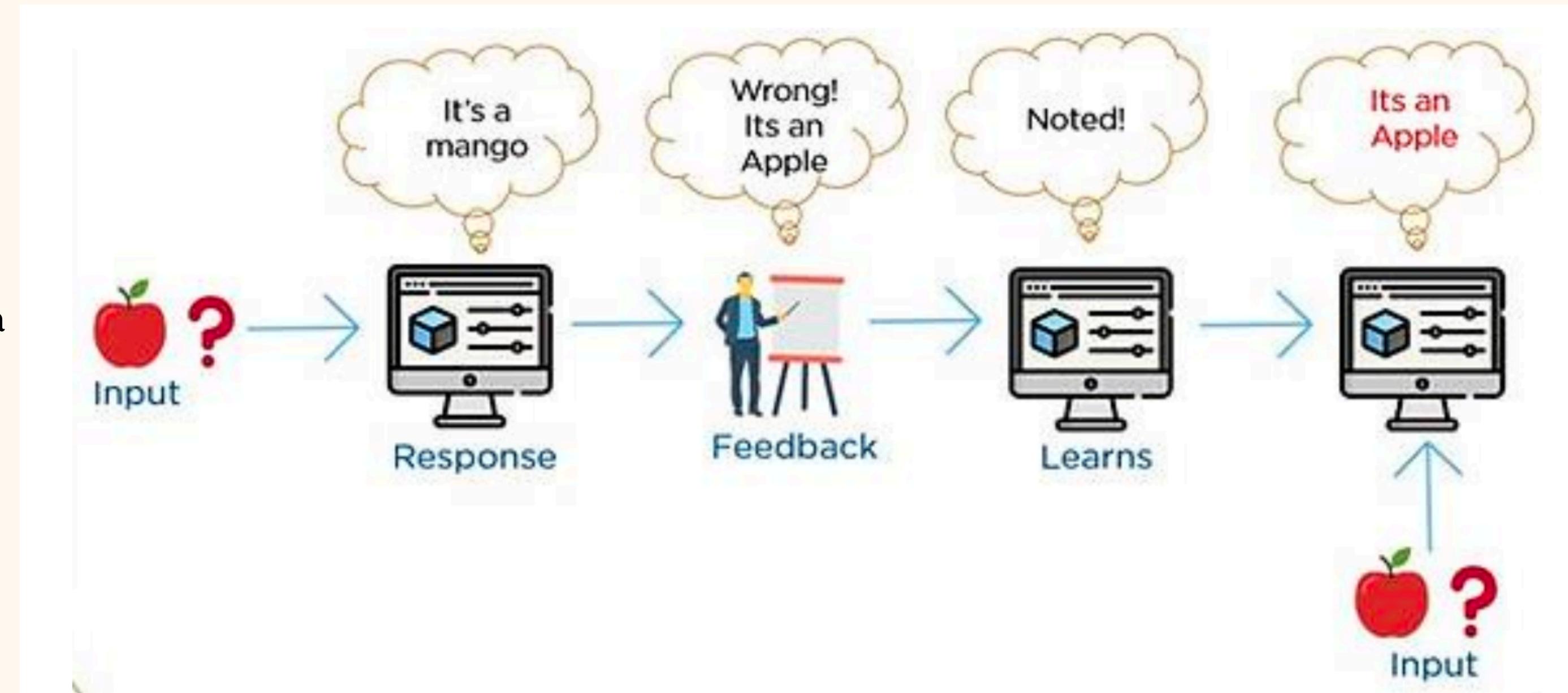


Unsupervised learning

(generative) – uses unclassified / unlabeled data



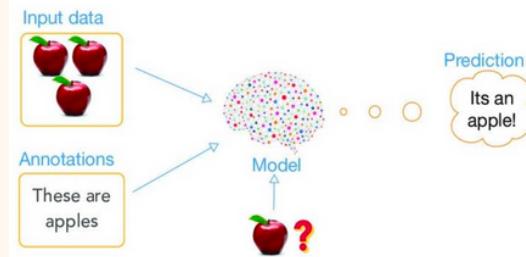
Reinforcement learning –
No data required, learns from the experience, etc.



Types of machine learning: Reinforcement learning

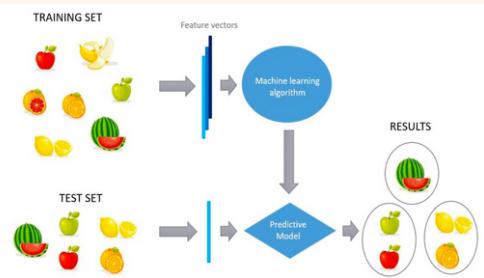
Supervised learning

(discriminative) – uses classified/labeled data

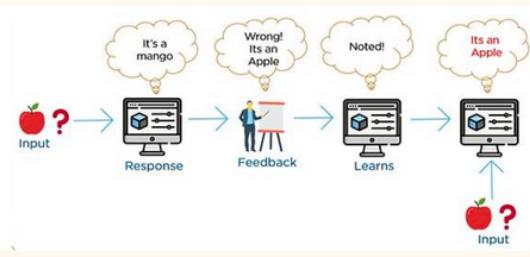


Unsupervised learning

(generative) – uses unclassified / unlabeled data



Reinforcement learning –
No data required, learns from the experience, etc.



Examples:

Video Game AI players



Minecraft's AI companions use trial and error AI models - getting rewards for successfully mining resources or building structures, and penalties for falling into lava.

Recommendation timing



Video recommendation timing algorithms such as on YouTube, learn when to show notifications by testing different times and seeing when it is most likely to be clicked on. If you engage more at 7 PM, it gets rewarded and learns to notify you then.

Lesson difficulty adjustment



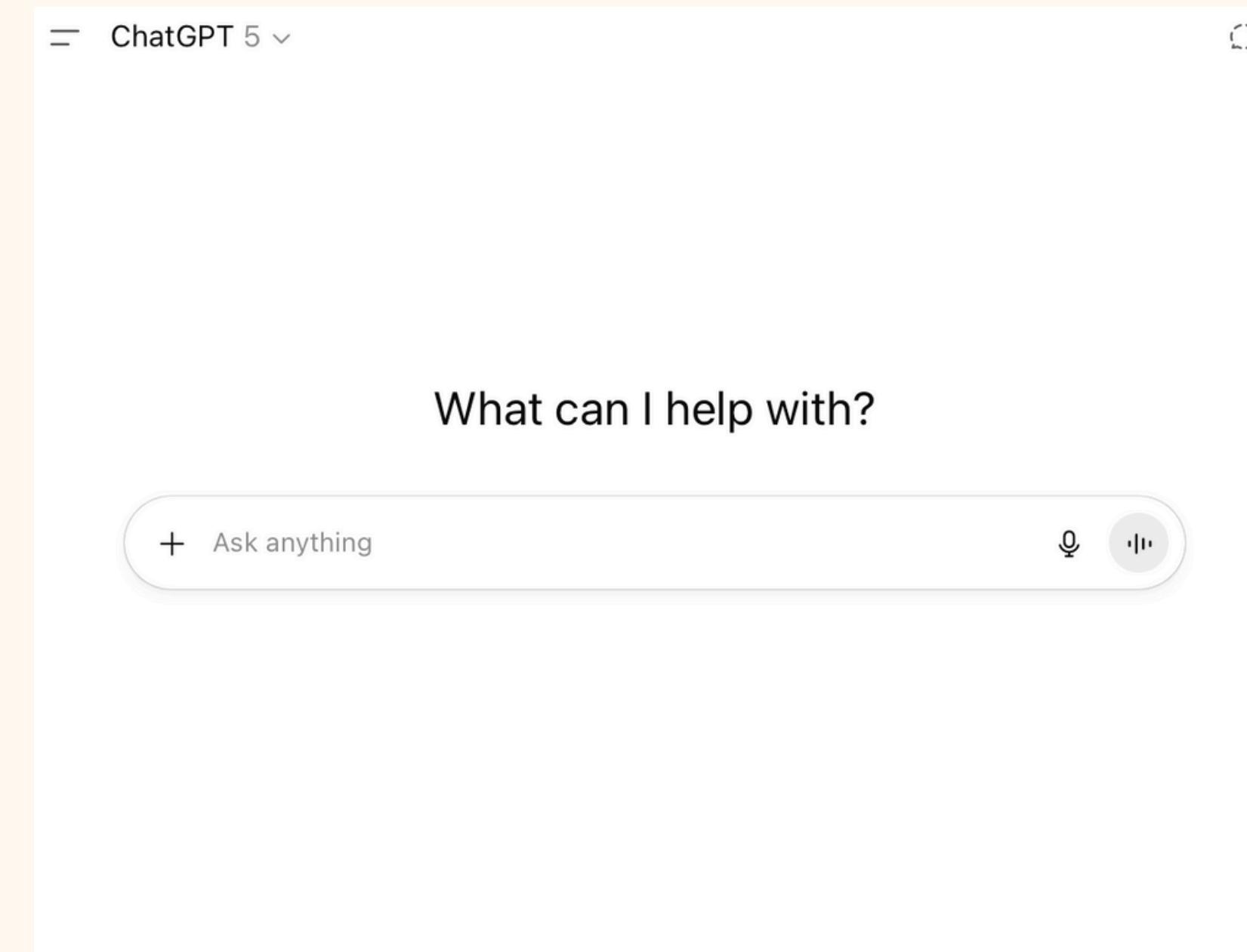
Learning apps, such as Duolingo, learn to adjust how hard your German or English lessons

should be by seeing if you succeed or struggle, constantly fine-tuning to keep you motivated but challenged.

Let's play an unplugged game: Machine Learning Types Game

- Cards

What about AI chatbots, like chatGPT?



How many of you have already used an AI chatbot? (raise your hand)

AI Chatbots, such as chatGPT are Large Language Models, considered Generative AI

A large language model (LLM) is a language model trained with self-supervised machine learning on a vast amount of text, designed for natural language processing tasks, especially language generation. They are also called Foundational Models and are costly to develop.

Source: [wikipedia](#)

In simple words, a program that can converse with humans.

Can you guess which type, way and outcomes LLMs (large language models) use?

<i>Types of learning (on what data)</i>	<i>Ways of learning (with which model)</i>	<i>Outcomes of learning (what can it do)</i>
With labels : supervised learning	Classic, Non-neural algorithms	Classification / Prediction- labels or projects
Without labels : unsupervised learning	Neural networks	Generation (GenAI)- creates
By trial and error with rewards : reinforcement learning		Pursue of a goal (AI agents)- decides and acts

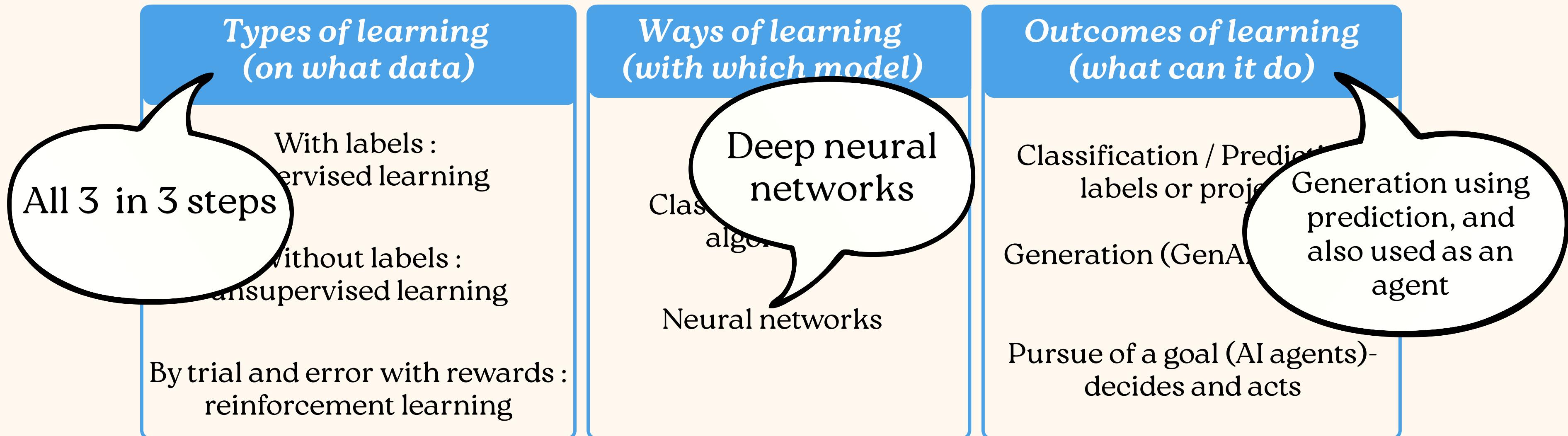
AI Chatbots, such as chatGPT are Large Language Models, considered Generative AI

A large language model (LLM) is a language model trained with self-supervised machine learning on a vast amount of text, designed for natural language processing tasks, especially language generation. They are also called Foundational Models and are costly to develop.

Source: [wikipedia](#)

In simple words, a program that can converse with humans.

Can you guess which type, way and outcomes LLMs (large language models) use?



Creating a large language model (LLM) in 3 steps

Step 1: Unsupervised Learning - "Reading a lot"

What happens: The AI reads millions of books, websites, and articles without anyone telling it what's important. It just learns patterns in language - like which words usually come after others.

The metaphor: Imagine a child who secretly listens to every conversation in their house, at school, and on TV for years. They start noticing patterns: "When people say 'Good morning,' others usually respond with 'Good morning' back" or "The word 'because' is usually followed by an explanation."

Real example: The AI learns that after "The capital of Germany is..." the word "Berlin" usually comes next, just from seeing this pattern thousands of times in different texts.



Large language models use text data, but the same ideas are being applied to images, sensor data, audio data, etc. Those are not considered Large Language Models.

Creating a large language model (LLM) in 3 steps

Step 2: Supervised Learning - "Learning what is a "good" answer"

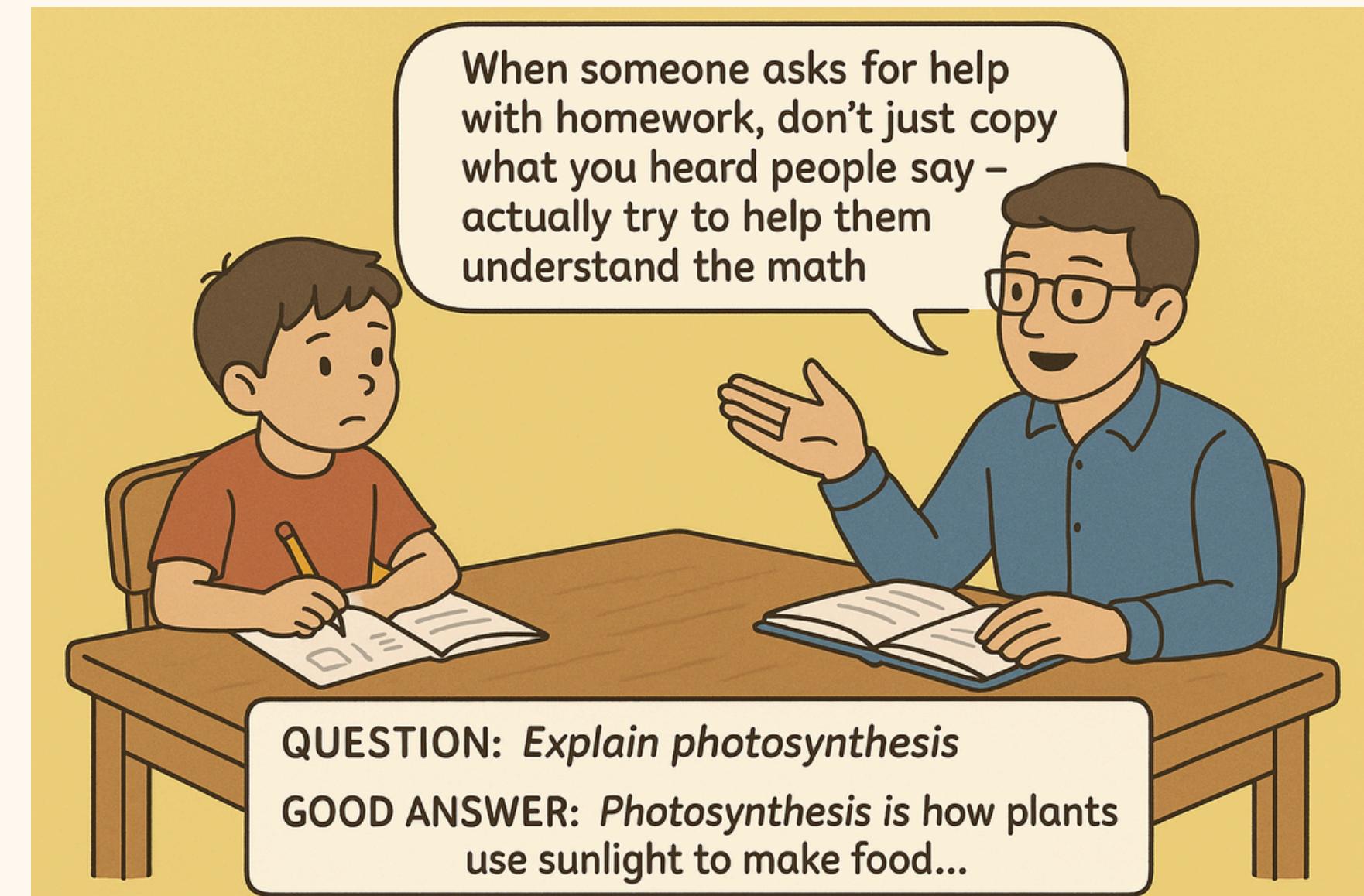
What happens: Now humans give the AI specific examples of good answers to questions. They show it: "When someone asks this question, HERE is a good response."

The metaphor: It's like the child now has a tutor who says: "When someone asks for help with homework, don't just copy what you heard people say - actually try to help them understand the math problem."

Real example: Humans show the AI thousands of examples like:

Question: "Explain photosynthesis" → Good Answer:

"Photosynthesis is how plants use sunlight to make food..."



Creating a large language model (LLM) in 3 steps

Step 3: Reinforcement Learning - ""Learning to provide a "good" answer"

What happens: Humans rate the AI's responses as "good" or "bad." The AI learns to avoid responses that humans dislike and favor ones they prefer.

The metaphor: Like getting feedback from friends - if they always groan when you tell certain jokes, you learn to stop telling those jokes. If they laugh at others, you tell more of those.

Real example: If the AI gives a rude response, humans mark it as "bad." If it gives a helpful, polite response, they mark it as "good." The AI learns to be more like the "good" examples.



Questions and Discussion

These 3 steps are involved in the building of Large Language Models

Step 1: Unsupervised Learning - "Reading a lot"

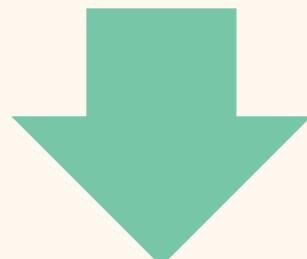
Real example: The AI learns that after "The capital of Germany is..." the word "Berlin" usually comes next, just from seeing this pattern thousands of times in different texts.

Step 2: Supervised Learning - "Learning what is a “good” answer"

Real example: Humans show the AI thousands of examples like: Question: "Explain photosynthesis" → Good Answer: "Photosynthesis is how plants use sunlight to make food..."

Step 3: Reinforcement Learning - "Learning to provide a “good” answer"

Real example: If the AI gives a rude response, humans mark it as "bad." If it gives a helpful, polite response, they mark it as "good." The AI learns to be more like the "good" examples.



Large Language Model (LLM)

that is a precise mathematical model with its parameters, also called weights

Major Large Language Models (LLMs)

OpenAI models (US)



ex: GPT-5, GPT-4o

Google models (US)



ex: Gemini 2.5 Pro, etc

Anthropic models (US)



ex: Claude 3.7 Sonnet

Meta models (US)



ex: Llama 4 series

DeepSeek models (China)



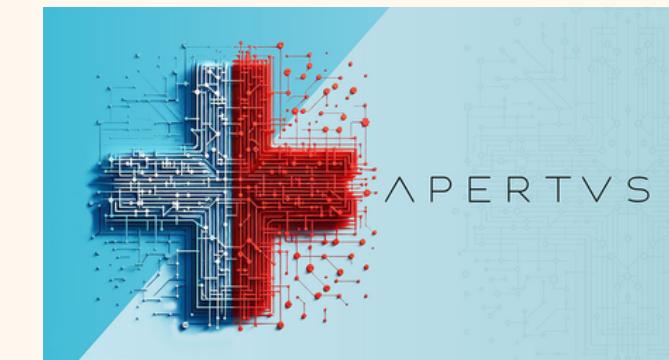
Mistral (France)



xAI's Grok (US)



Apertus (Switzerland)



and others...

Which ones have you tried ?

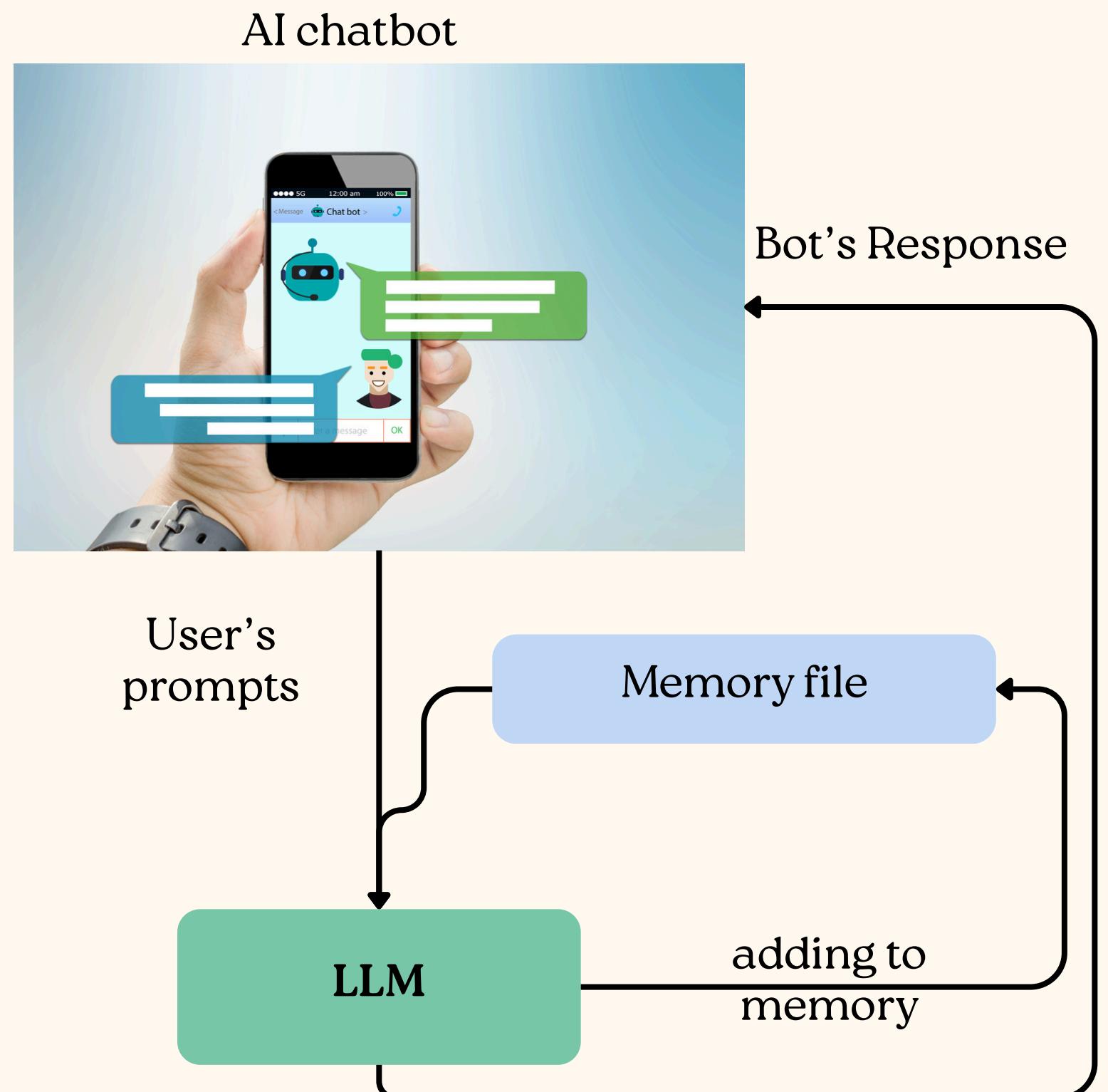
AI chatbots use LLMs models

When we interact with an AI chatbot (send them text, called prompts), the underlying model does not change (the weights are unchanged), therefore technically does not learn.

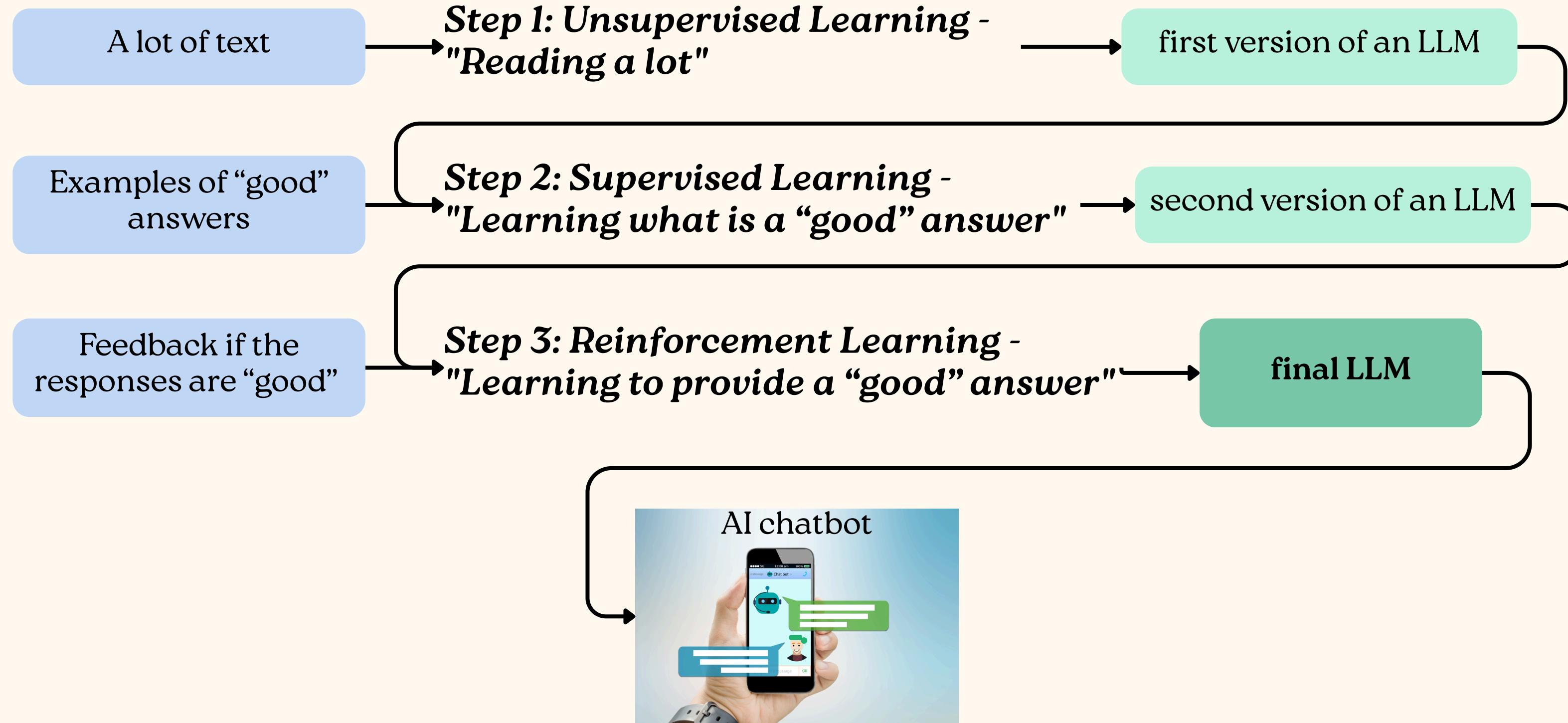
However, software using LLMs can create memory files, which are added to prompts creating an illusion of learning.

LLMs can apply their “knowledge” (patterns learned in the previous steps) to new data, like to analysing a new text and detecting its patterns, and by doing so, creating new knowledge for us to learn.

Often our conversations are recorded and used as input or test data for new models (for ex. conversations with GPT-4 used to create GPT-5)



In summary



Questions ?

Let's try a simple version of an LLM



1. Go to the Soekia GPT website.
2. Ask it to : "Write me a fairy tale".
3. Pause the generation after a few words/sentences and click at the bottom on "select yourself". You can now click on the next word yourself yourself (selected from the most likely matches of the model). Combine self-select/auto-generate to create a fairy tale.
4. Click on "Look inside" in the top right corner to better understand how this works. Explore.
5. Click on the trash can under the created story to start over again. Ask another question, e.g: Explain me the Pythagorean theorem. How well does Soekia do?

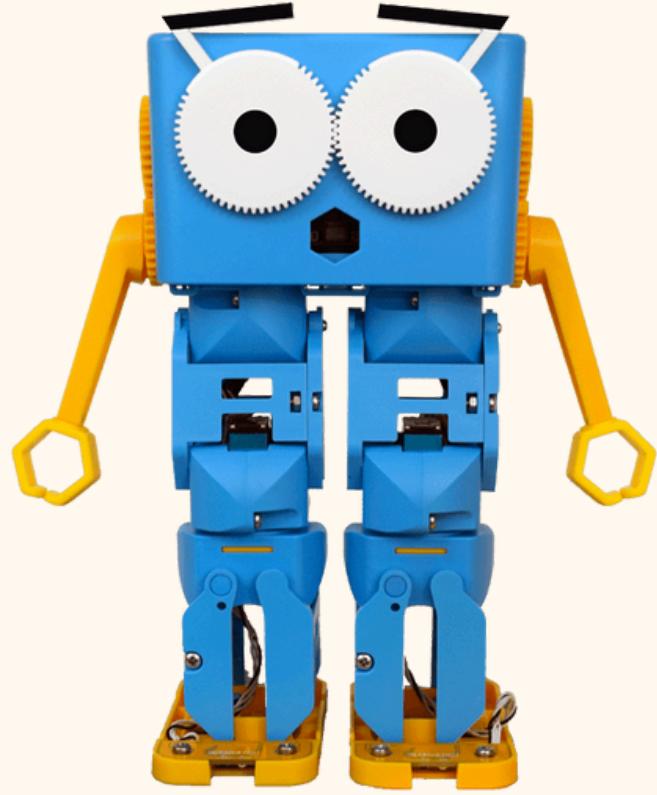
Questions and Discussion after 10 minutes

Classifier or generator activity

(5 minutes)

Answers guide

Let's start building



Marty



Annexes