



PyTorch



AiAT
ARTIFICIAL INTELLIGENCE ASSOCIATION OF THAILAND

INTRODUCTION TO COMPUTER VISION

Workshop 13:00 - 16.00



PRESENTED BY

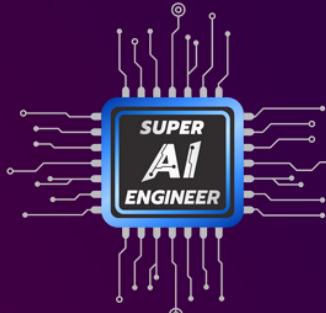
Varintorn Sithisint
Parinthapat Pengpun

Who are we?



Parinthapat Pengpun

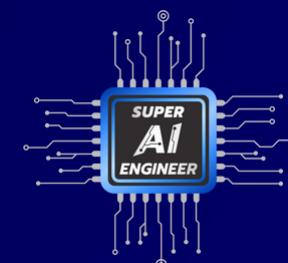
- Grade 10 Student
 - Bangkok Christian International School
- AI Builders 2022
- NLP Researcher
- Super AI Engineer Season 3
- AI Engineer at Deep Invest x Deep Capital Co. Ltd,



Varintorn Sithisint



- Third-year student in applied computer science
- King Mongkut's university of technology Thonburi
- Data scientist at AltoTech(Thailand) Co., Ltd.



Agenda

- 1 Definition of Computer Vision
- 2 Image Classification
- 3 Interface (Gradio)
- 4 Object Detection
- 5 Text-to-Image model – Stable Diffusion

Cat or Dog ?



Cat or Dog ?



It's a dog !

Cat or Dog ?



Cat or Dog ?



It's a cat !

Cat or Dog ?



Cat or Dog ?



ໃນສີ

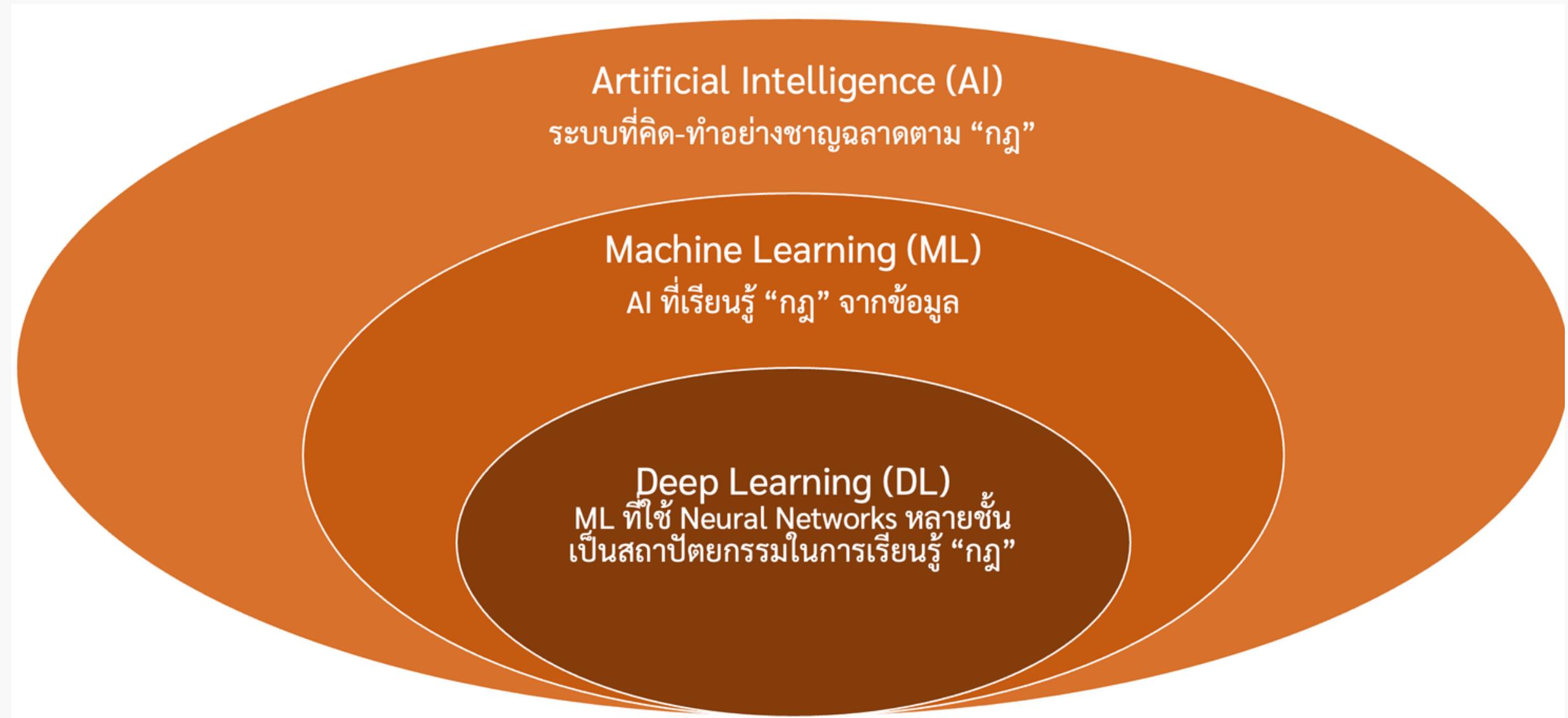
JCSCE

AiAT
ARTIFICIAL INTELLIGENCE ASSOCIATION OF THAILAND

What is Computer Vision?

What is AI?

- **Artificial Intelligence** uses rules
- **Machine learning** learns rules with statistics
- **Deep Learning** learns rules using **neural network**



<https://github.com/ai-builders/curriculum/tree/main>

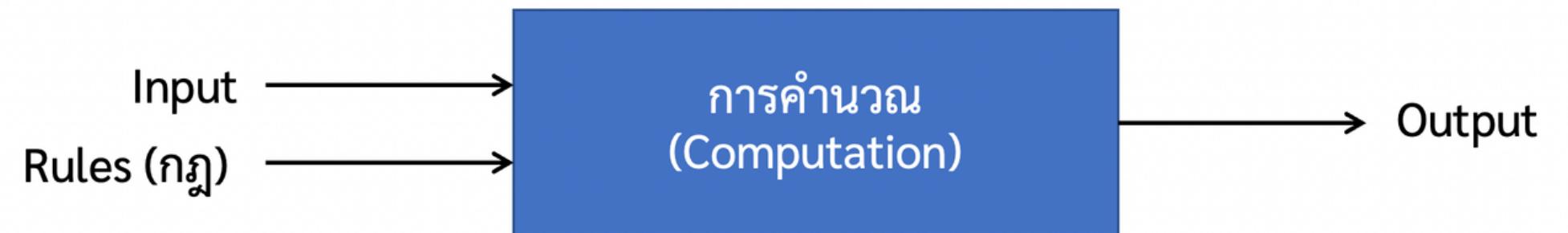


What is AI?

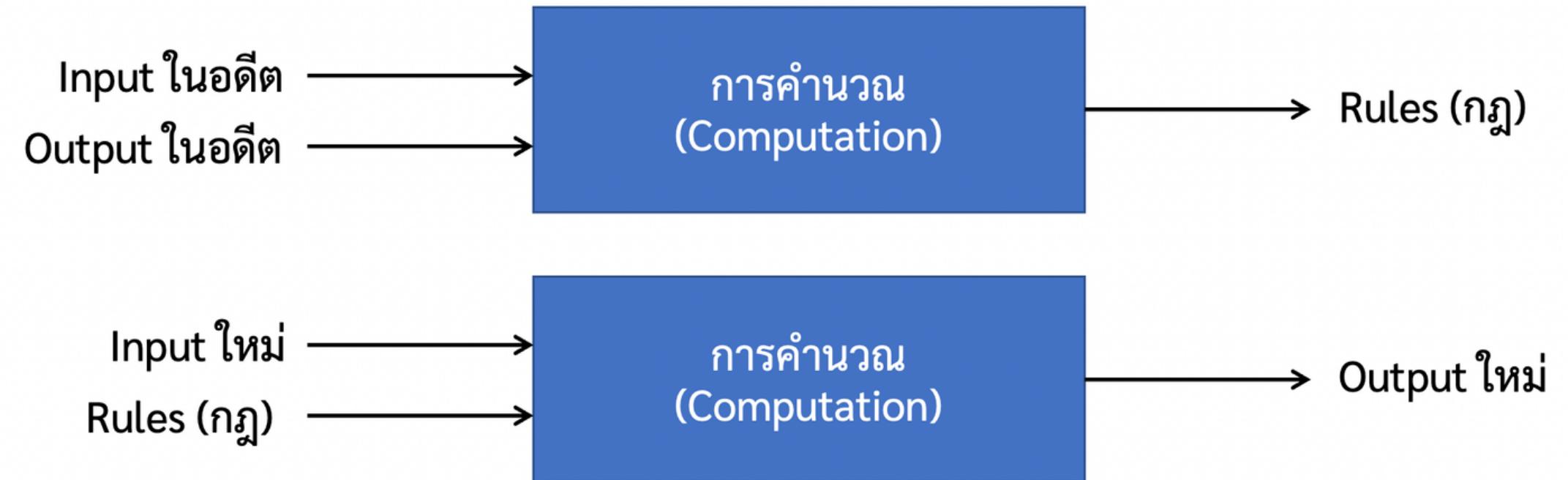
- **Artificial Intelligence** uses rules
- **Machine learning** learns rules with statistics
- **Deep Learning** learns rules using neural network

<https://github.com/ai-builders/curriculum/tree/main>

Rule-based Systems

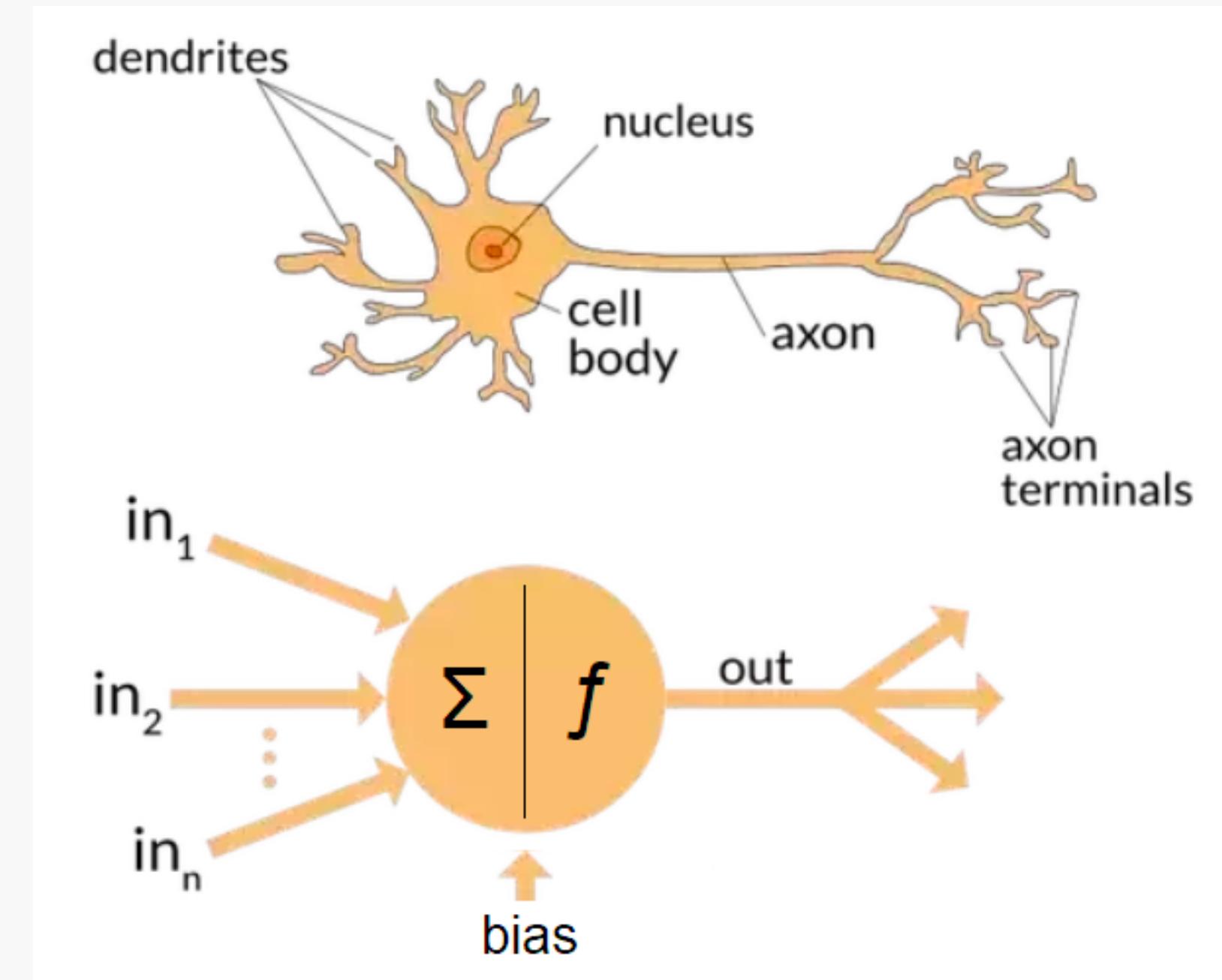
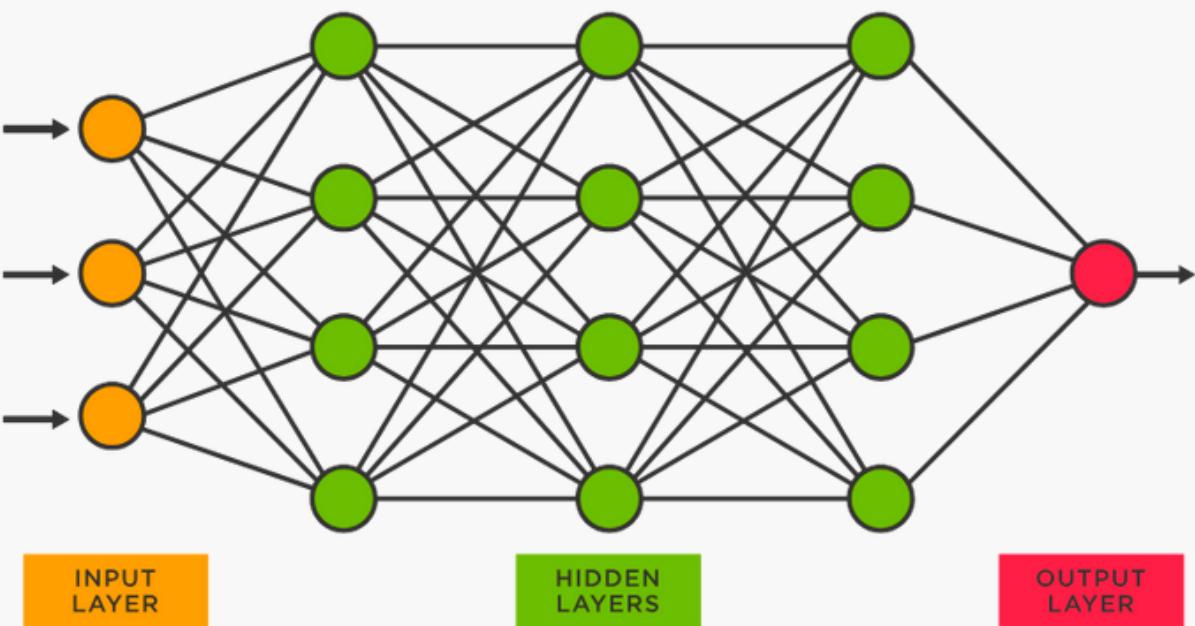


Machine Learning Systems



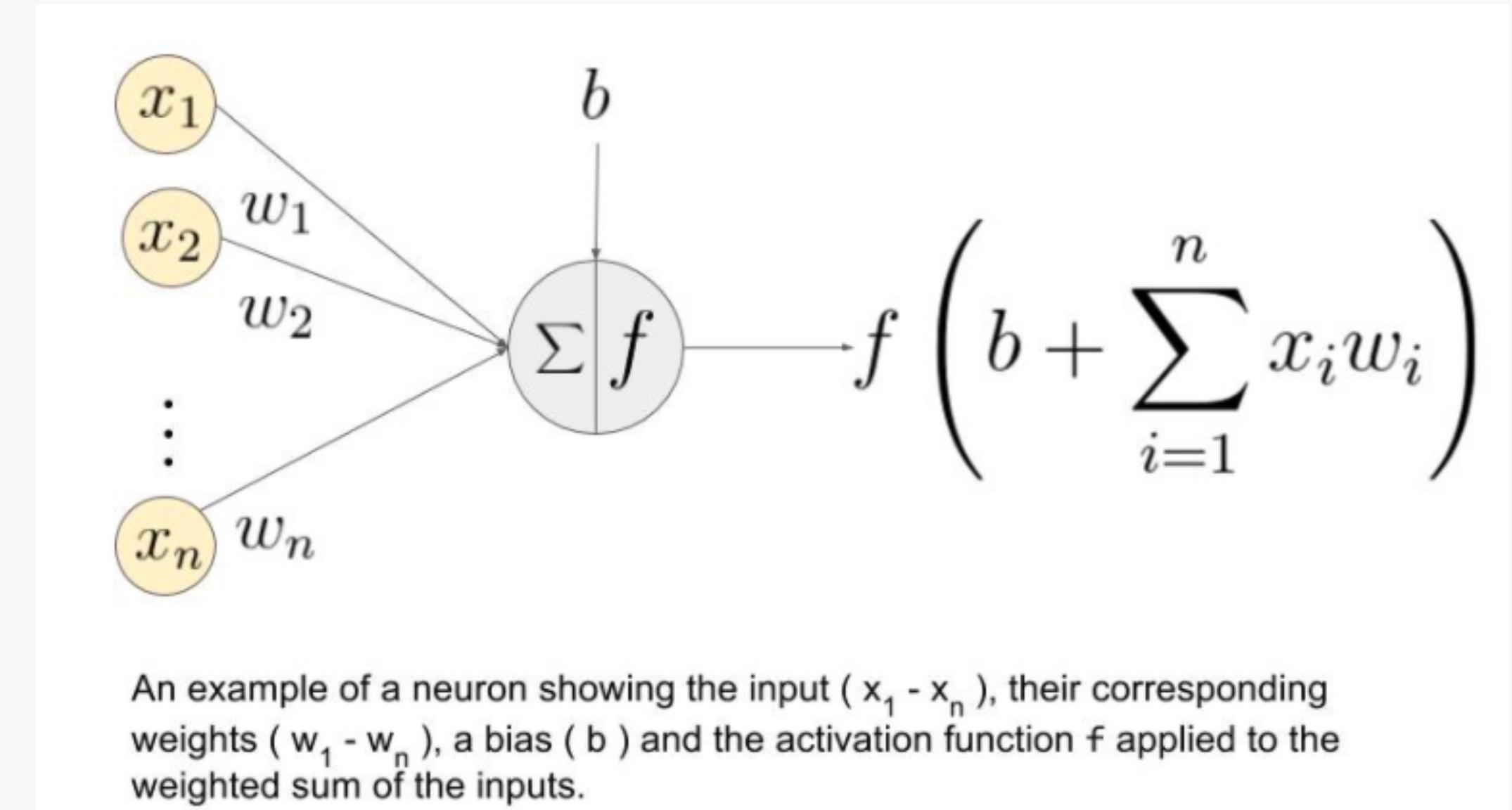
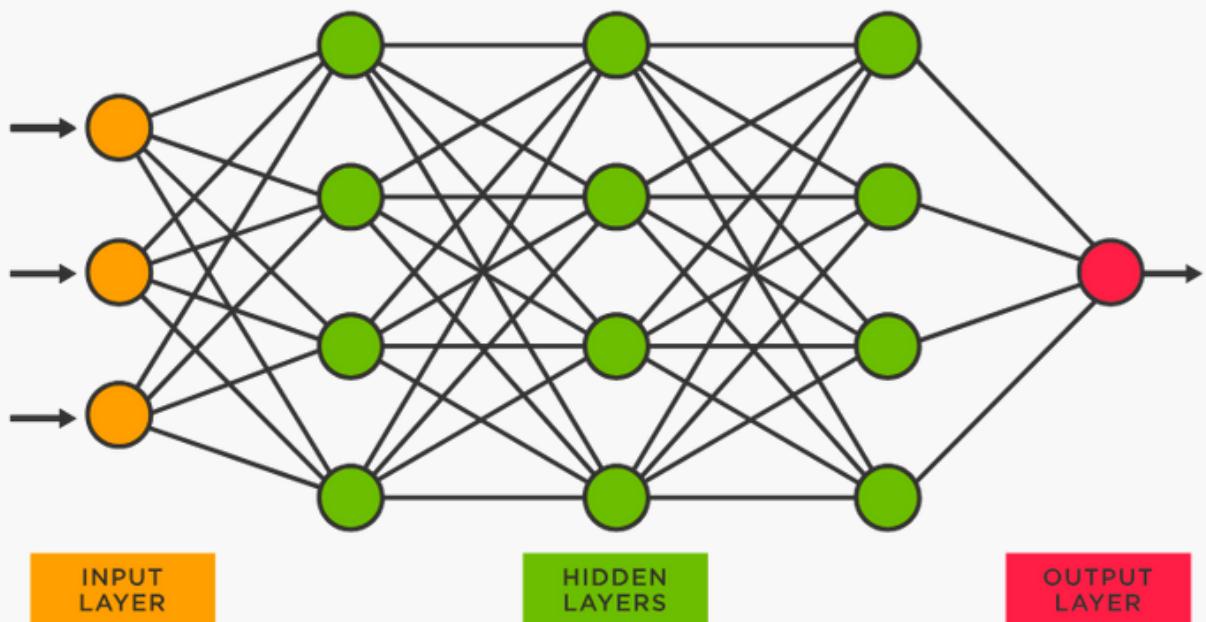
What is Deep Learning?

- Deep learning uses **neural networks**.
- Similar to human brain.



What is Deep Learning?

- Inside are **weights** and **biases**

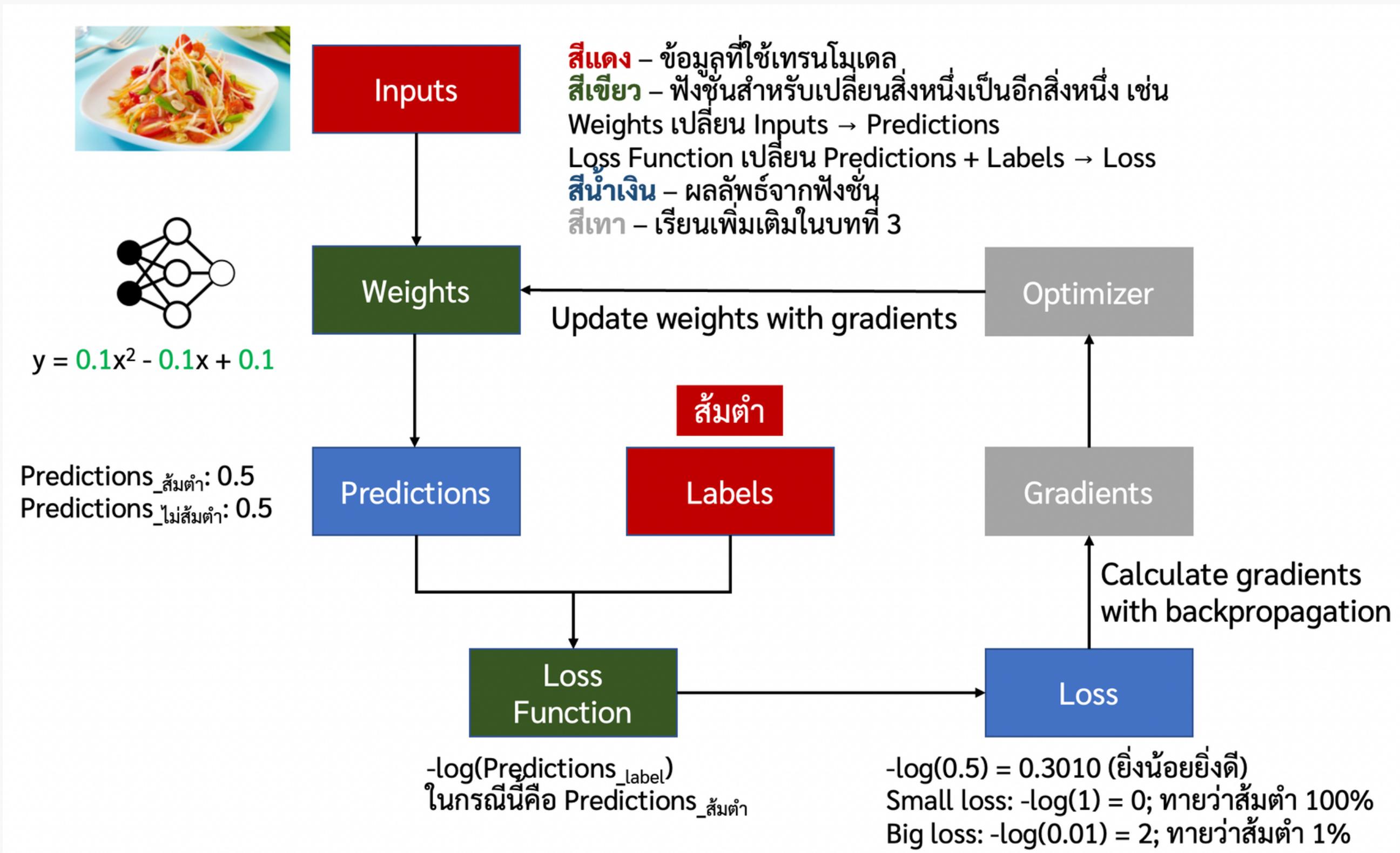


An example of a neuron showing the input ($x_1 - x_n$), their corresponding weights ($w_1 - w_n$), a bias (b) and the activation function f applied to the weighted sum of the inputs.



How do you learn?

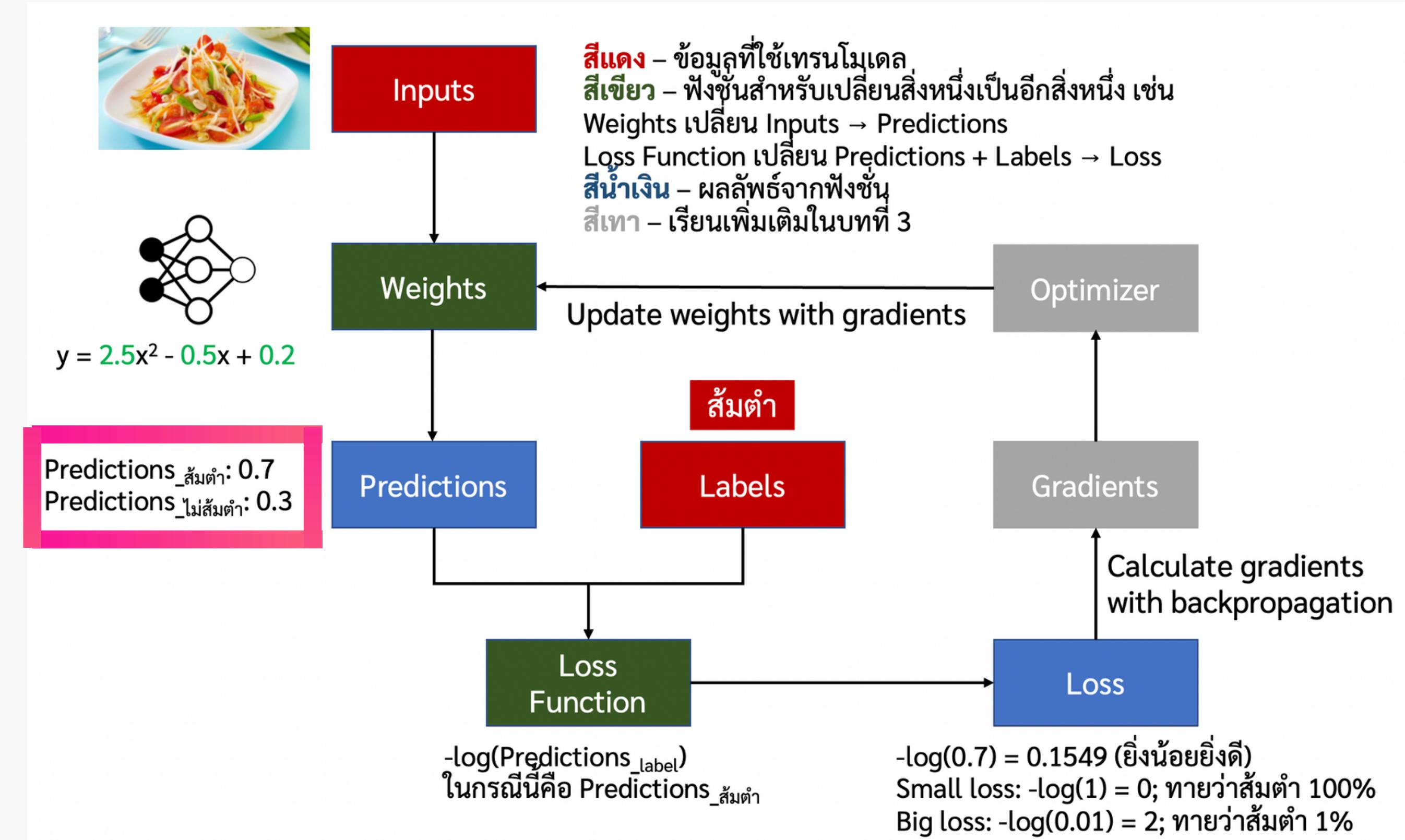
- Adjust weights according to loss
- **Loss:** how wrong the model is



https://github.com/ai-builders/curriculum/blob/main/notebooks/01_ml_what.ipynb

How do you learn?

- Adjust weights according to loss
- **Loss:** how wrong the model is

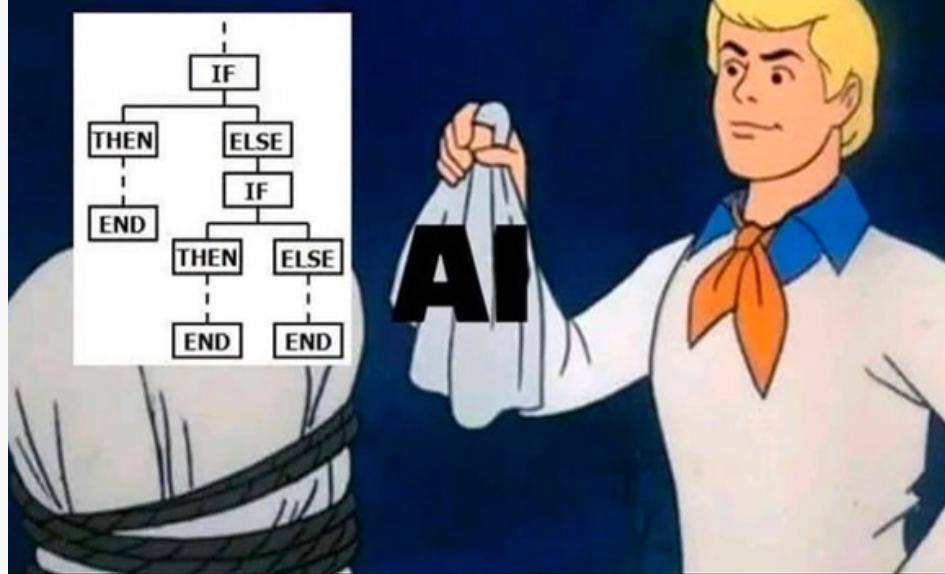


https://github.com/ai-builders/curriculum/blob/main/notebooks/01_ml_what.ipynb



Examples of AI?

AI



Machine Learning

amazon.com

Recommended for You

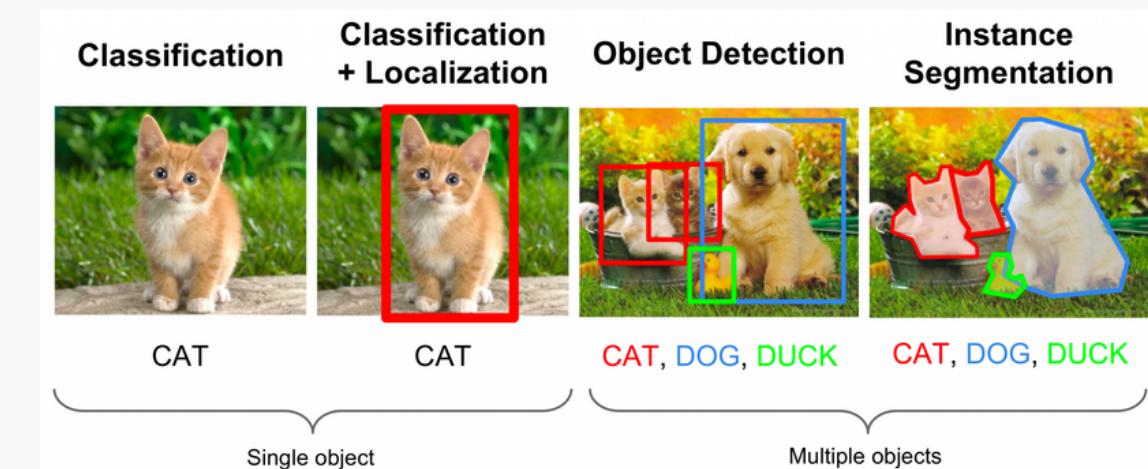
Amazon.com has new recommendations for you based on items you purchased or told us you own.

LOOK INSIDE! Google Apps Deciphered: Compute in the Cloud to Streamline Your Desktop **LOOK INSIDE!** Google Apps Administrator Guide: A Private-Label Web Workspace **LOOK INSIDE!** Googledia: The Ultimate Google Resource (3rd Edition)

[Google Apps Deciphered: Compute in the Cloud to Streamline Your Desktop](#) [Google Apps Administrator Guide: A Private-Label Web Workspace](#) [Googledia: The Ultimate Google Resource \(3rd Edition\)](#)



Deep Learning



What is Computer Vision?

เป็นศาสตร์ที่ใช้ในการทำให้คอมพิวเตอร์สามารถเข้าใจภาพหรือวิดีโอในรูปแบบเดียวกับที่ระบบการมองเห็นของมนุษย์

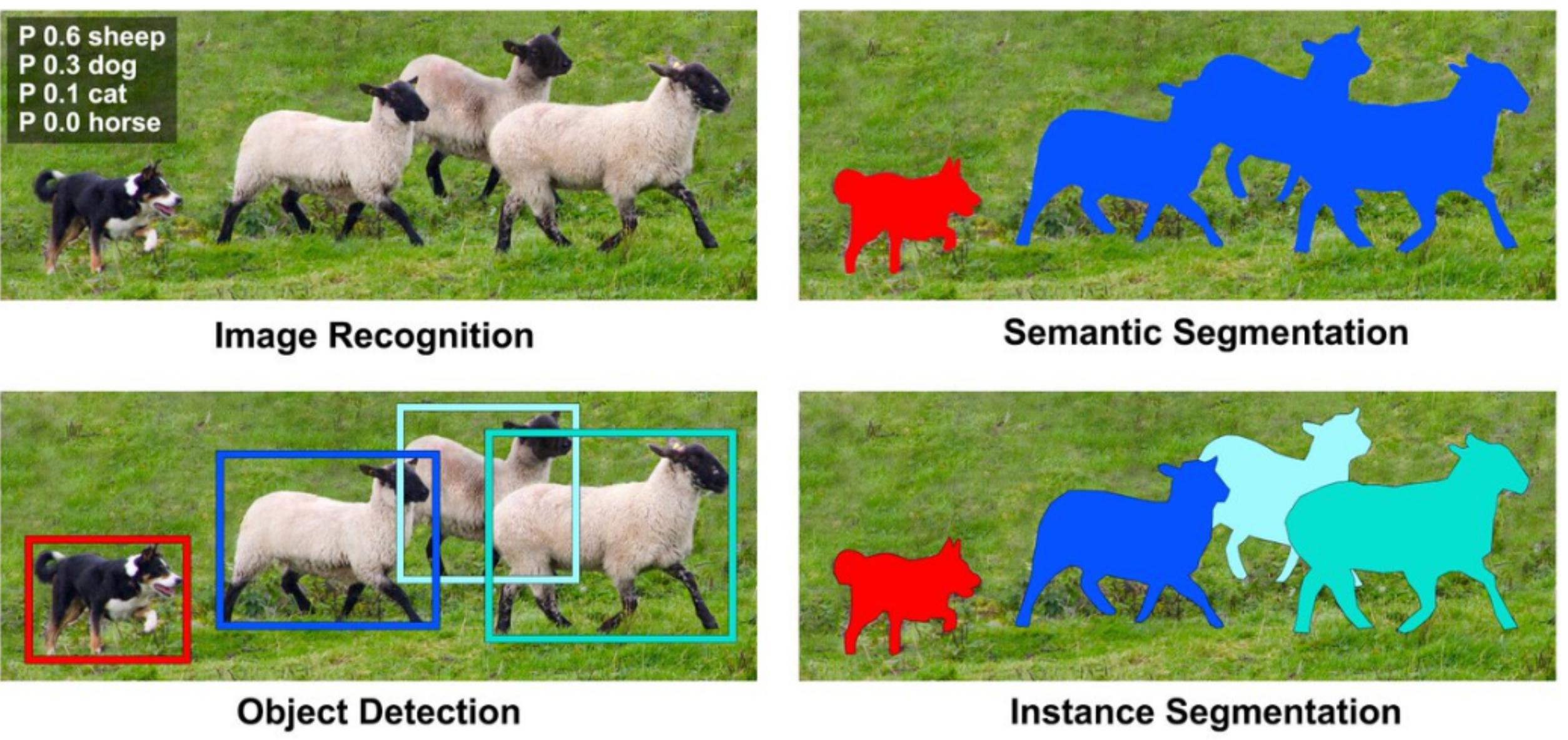


Image Classification

Image Classification

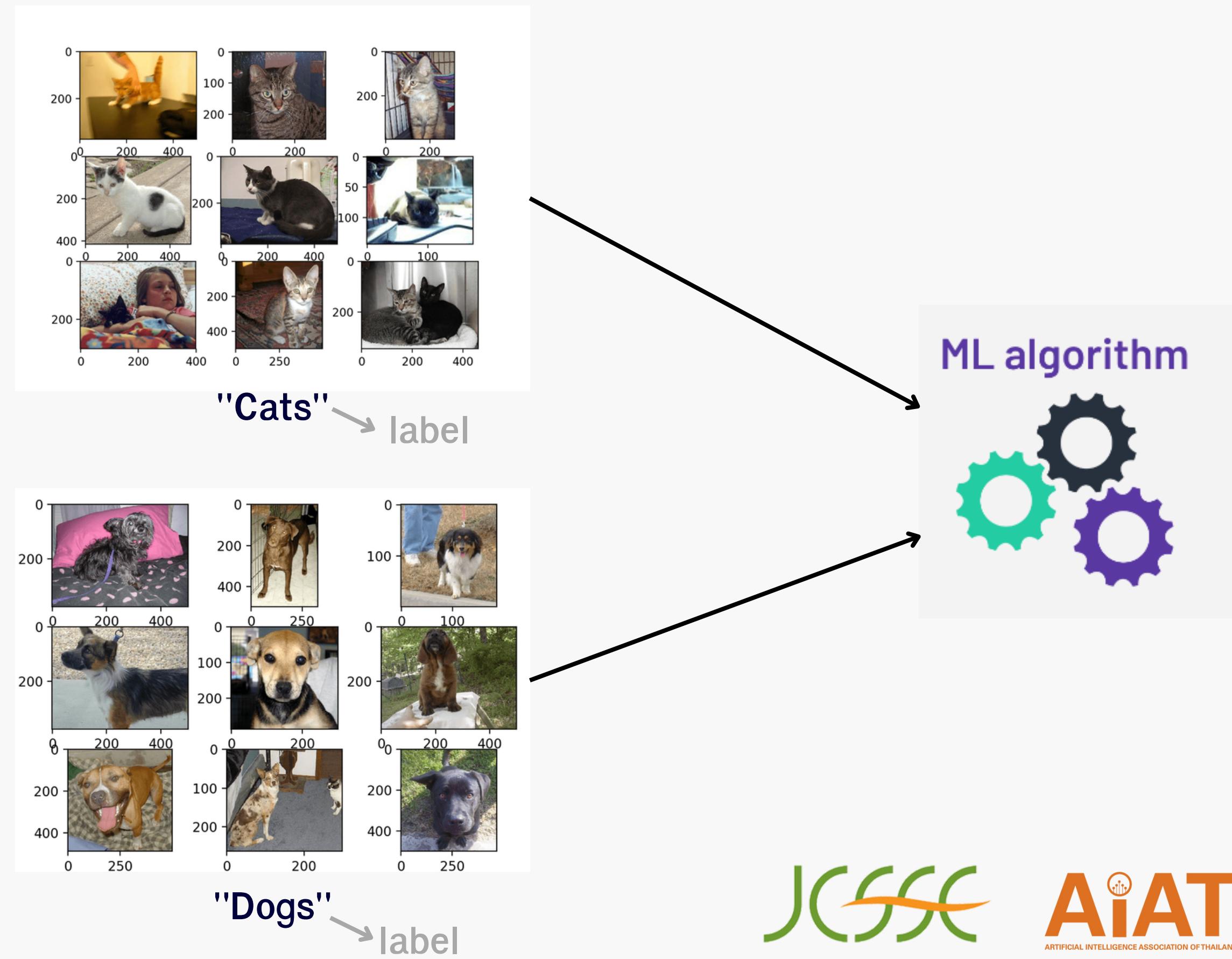
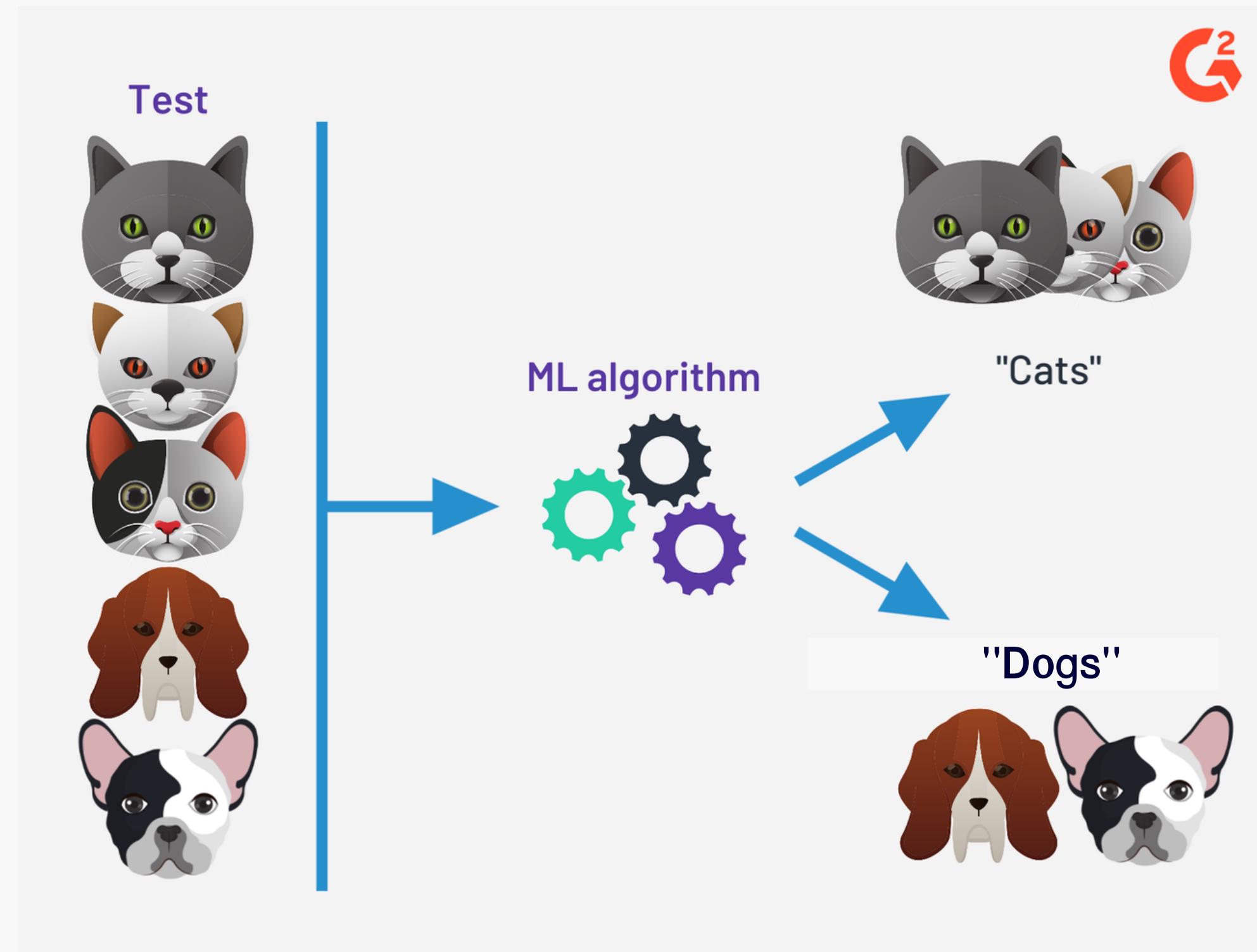
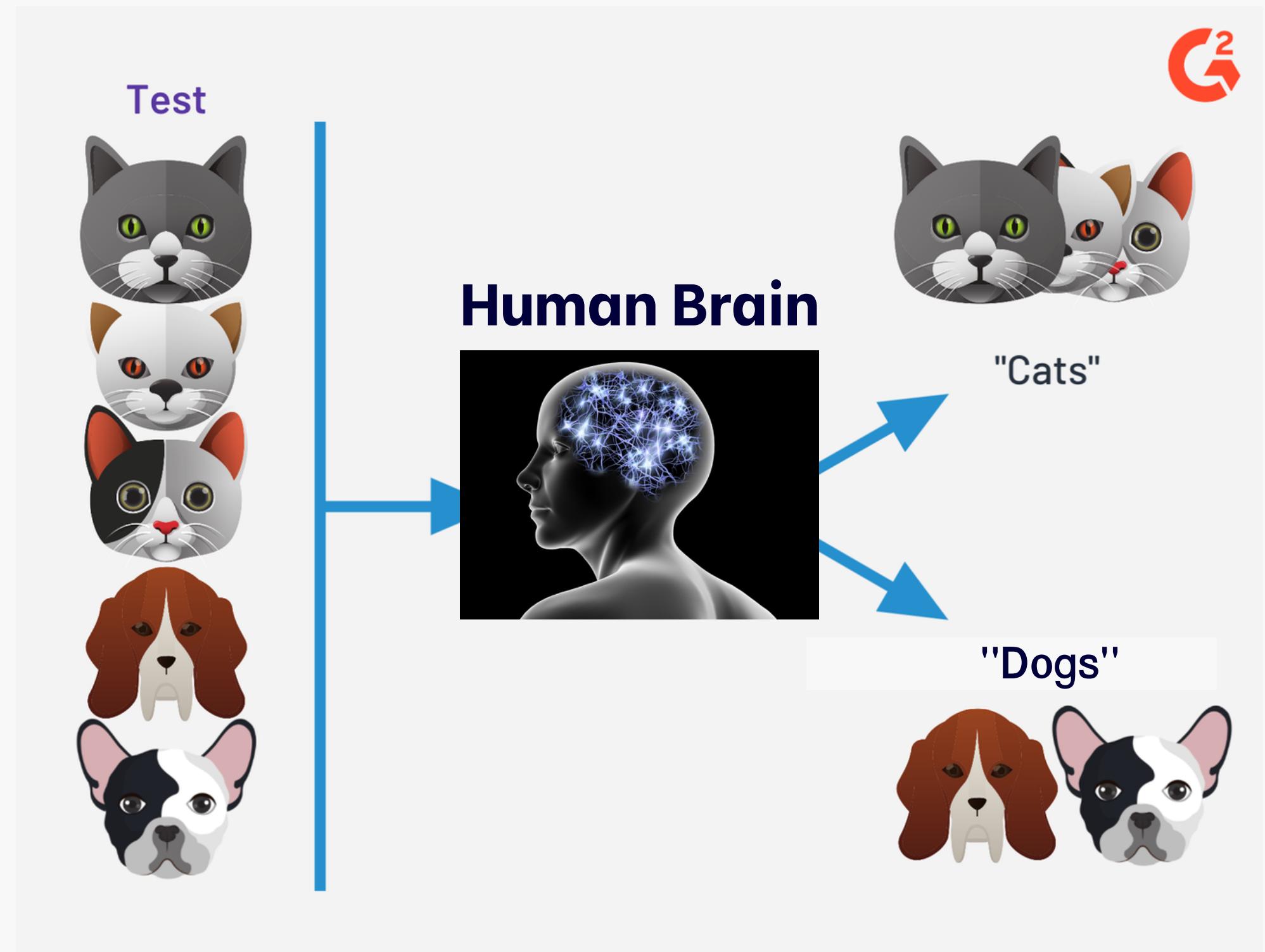


Image Classification



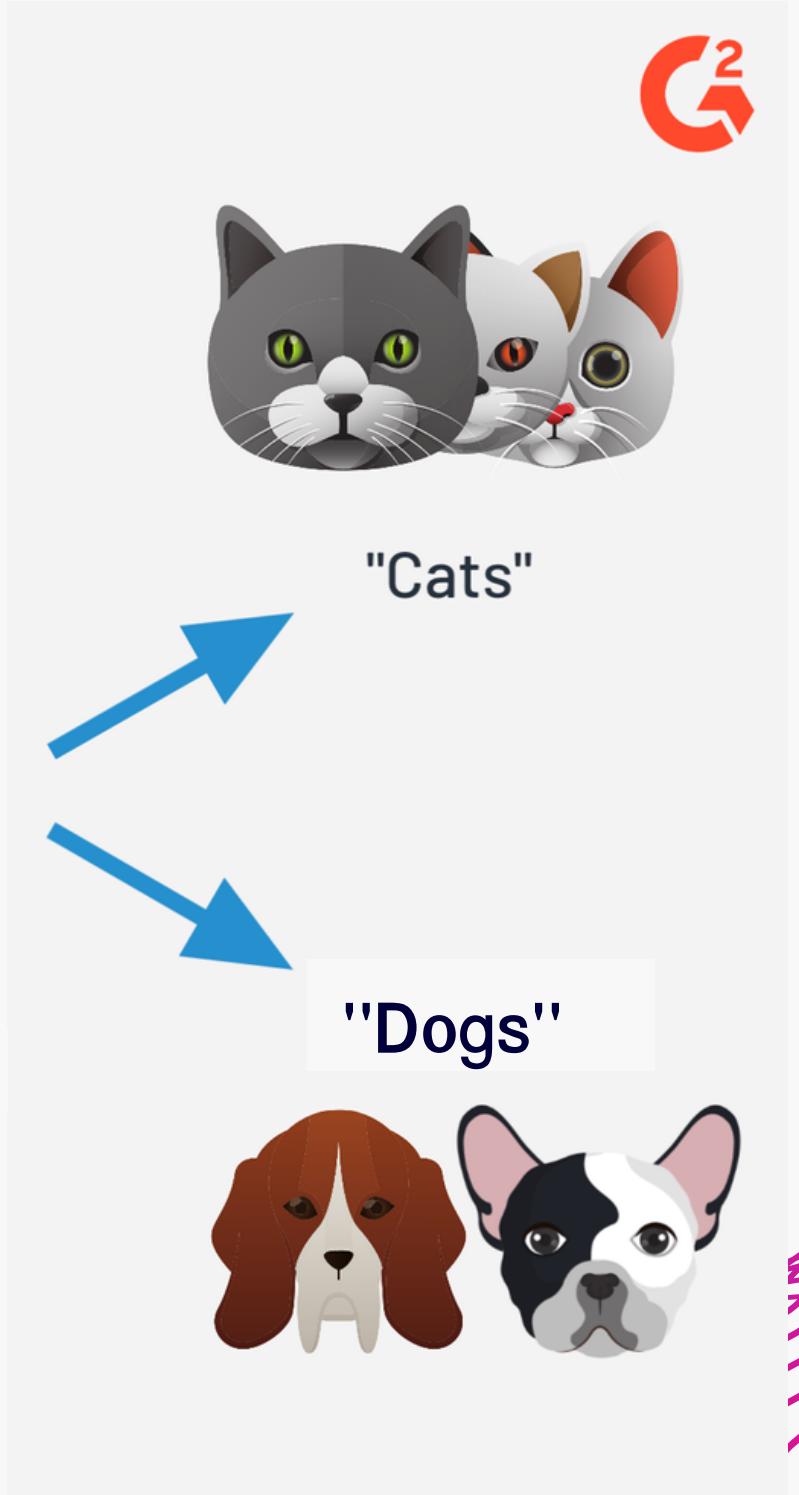
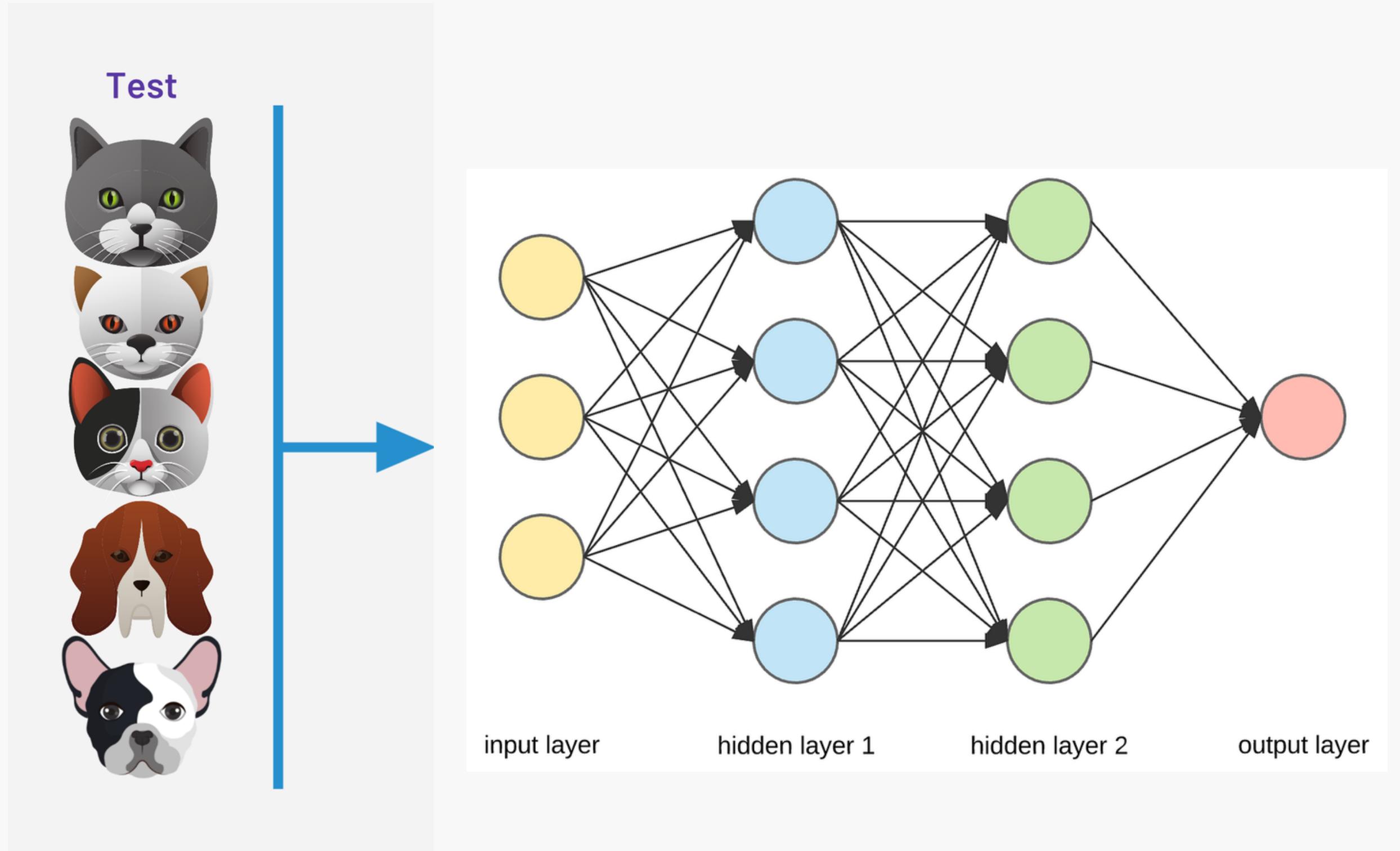
<https://www.g2.com/articles/supervised-vs-unsupervised-learning>

Image Classification



<https://www.g2.com/articles/supervised-vs-unsupervised-learning>

Image Classification

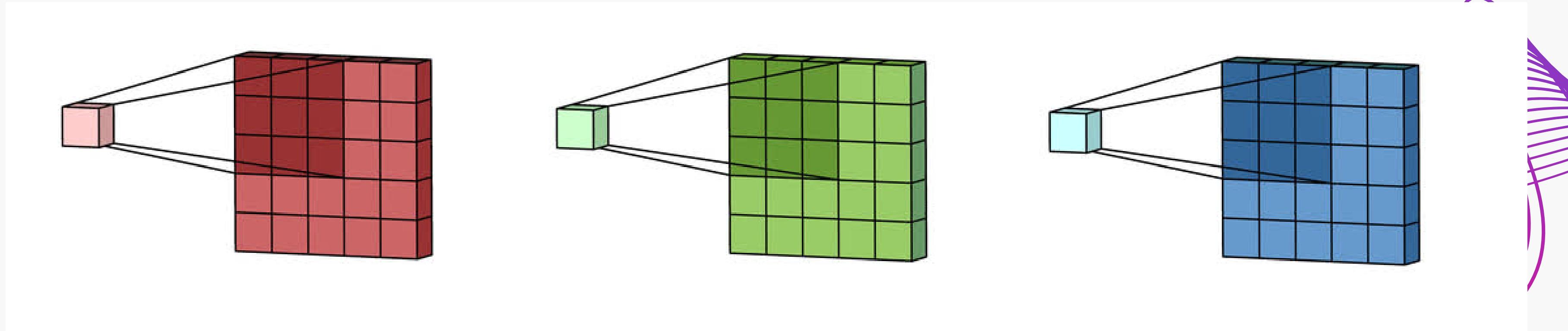
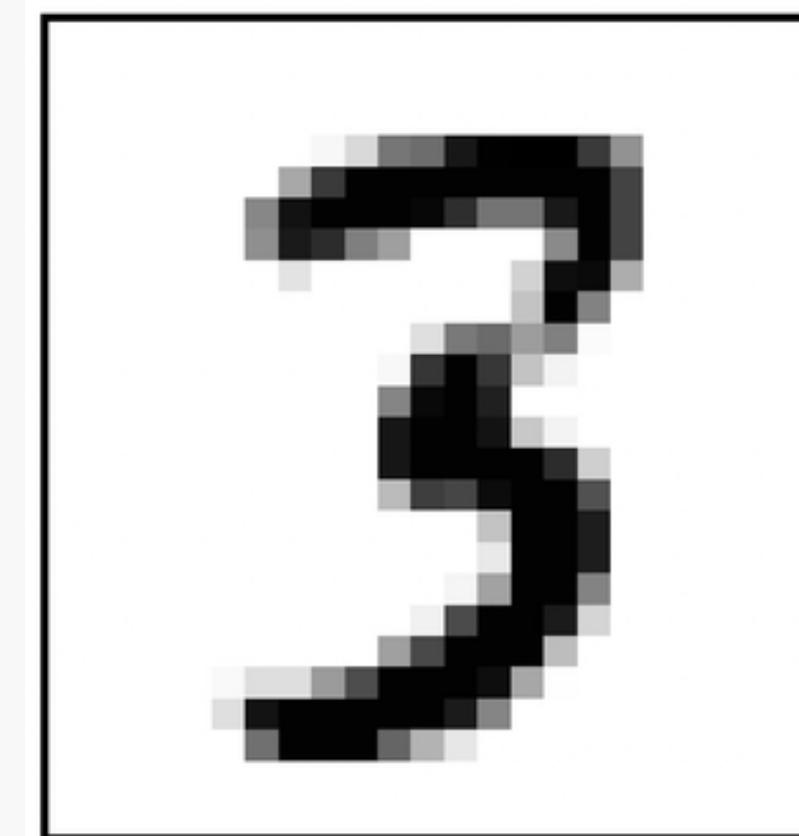


<https://www.g2.com/articles/supervised-vs-unsupervised-learning>

<https://blog.stevengong.co/how-does-a-neural-network-work-intuitively-in-code-f51f7b2c1e3f>

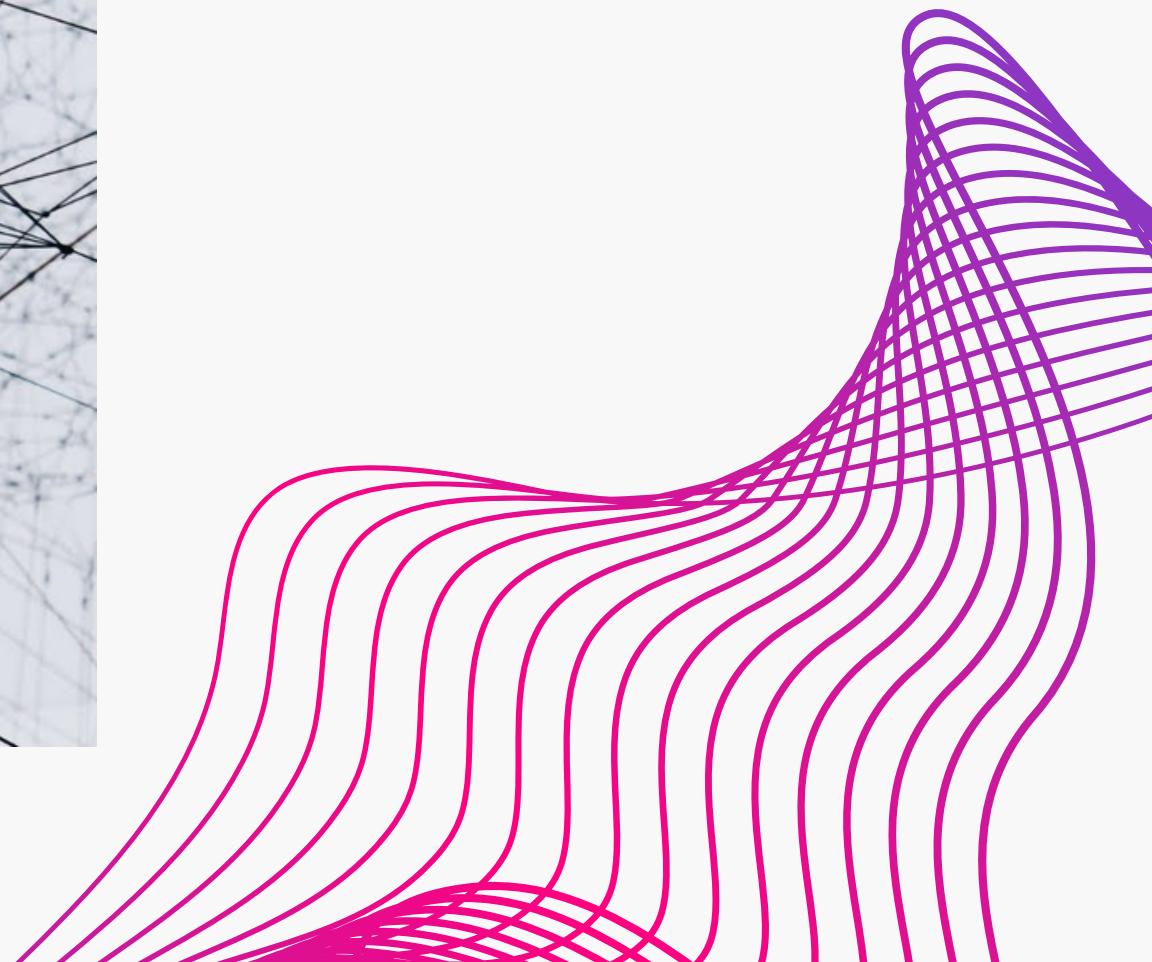
How does a neural network see?

- Works by **sliding over a matrix** with weights inside, to "see" the image.



<https://towardsdatascience.com/intuitively-understanding-convolutions-for-deep-learning-1f6f42faee1>

Fast.ai is an open-source deep learning library that provides high-level components that can quickly and easily provide state-of-the-art results in standard deep learning domains, and provides low-level components for you to experiment with and build upon. It uses the PyTorch framework as a backend.

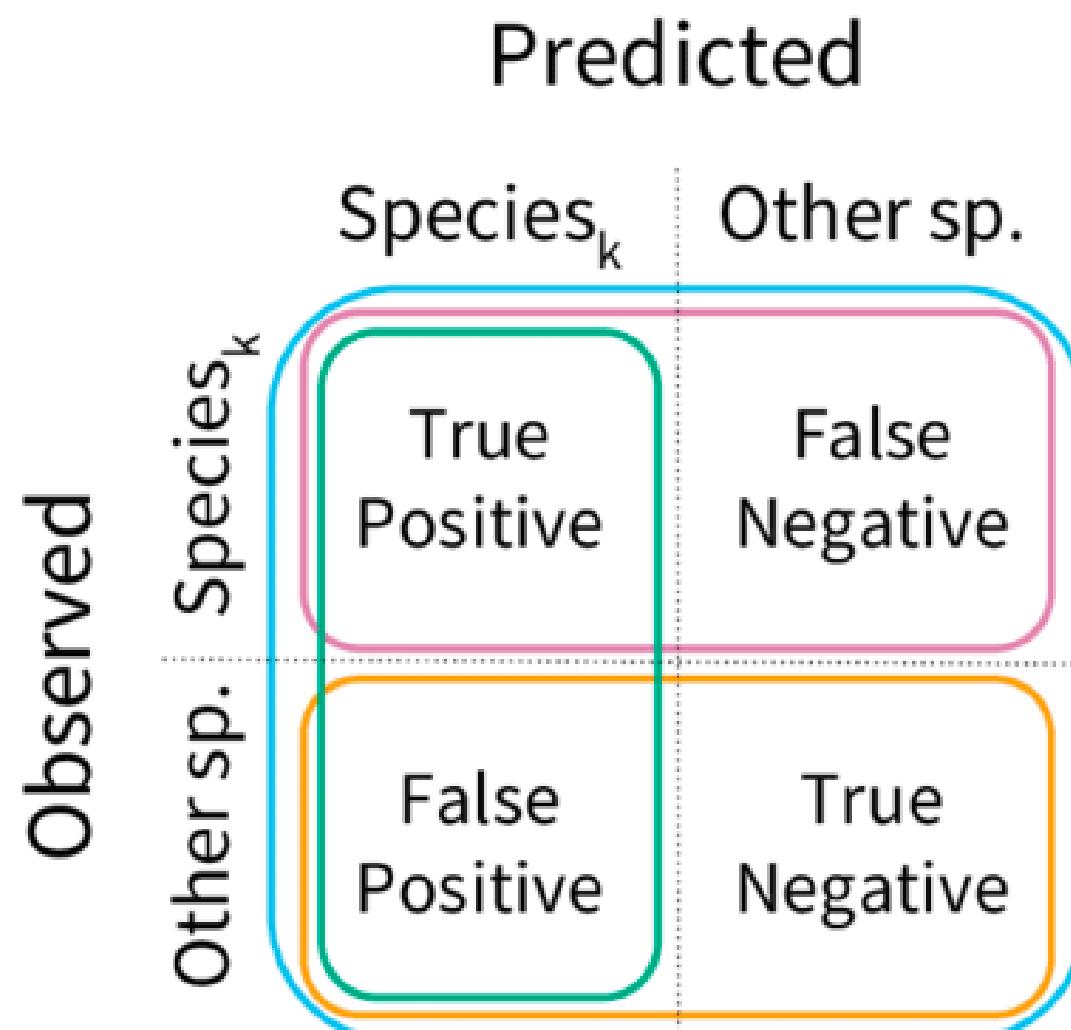


Load and visualize data with Datablock (DataLoader)

DataLoader is a built-in class that provides an efficient and flexible way to load data into a model for training or inference

```
[ ] dblock = DataBlock(blocks=(ImageBlock, CategoryBlock),  
                      get_items=get_image_files,  
                      splitter=RandomSplitter(seed=42),  
                      get_y=parent_label,  
                      item_tfms=Resize(460),  
                      batch_tfms=aug_transforms(size=224, min_scale=0.75, batch=True))
```

Metrics for Image Classification

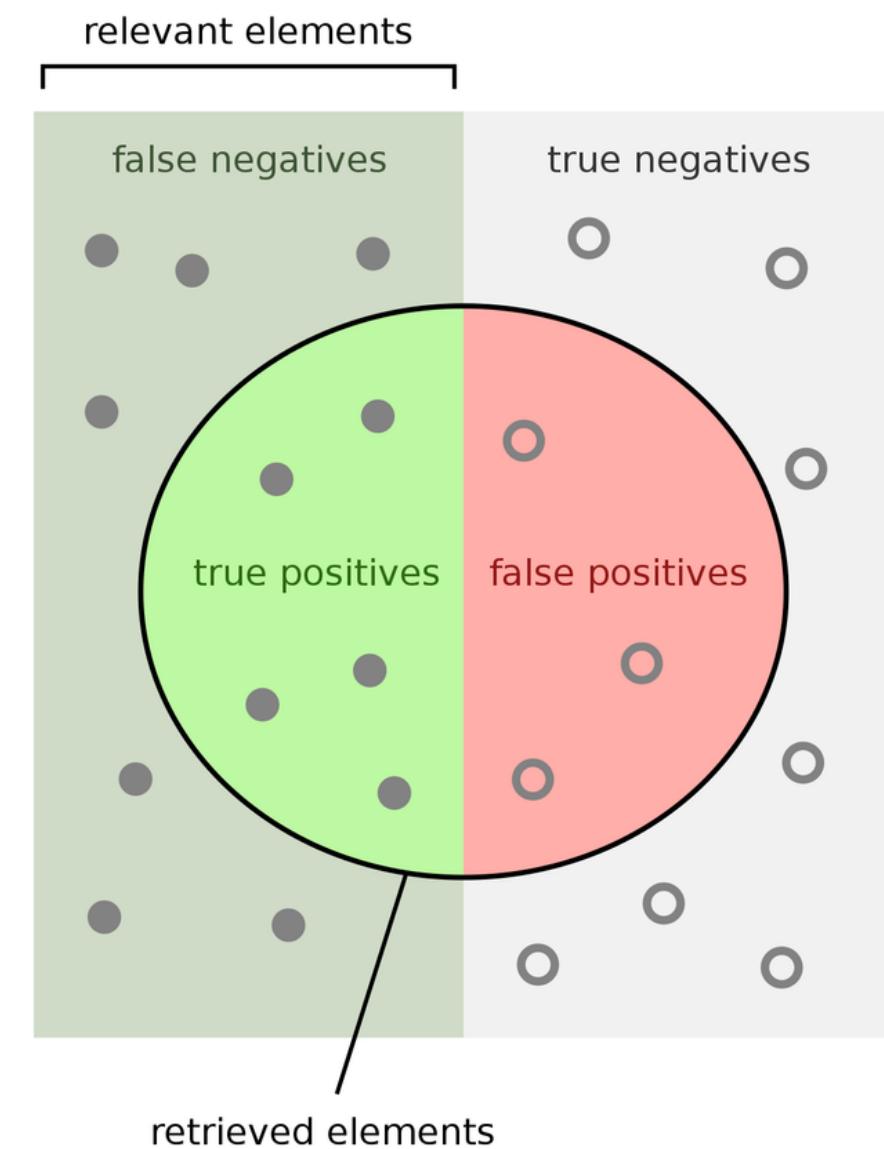


Accuracy = $\frac{TP + TN}{TP + TN + FP + FN}$

Specificity = $\frac{TN}{TN + FP}$

Precision = $\frac{TP}{TP + FP}$

Recall = $\frac{TP}{TP + FN}$

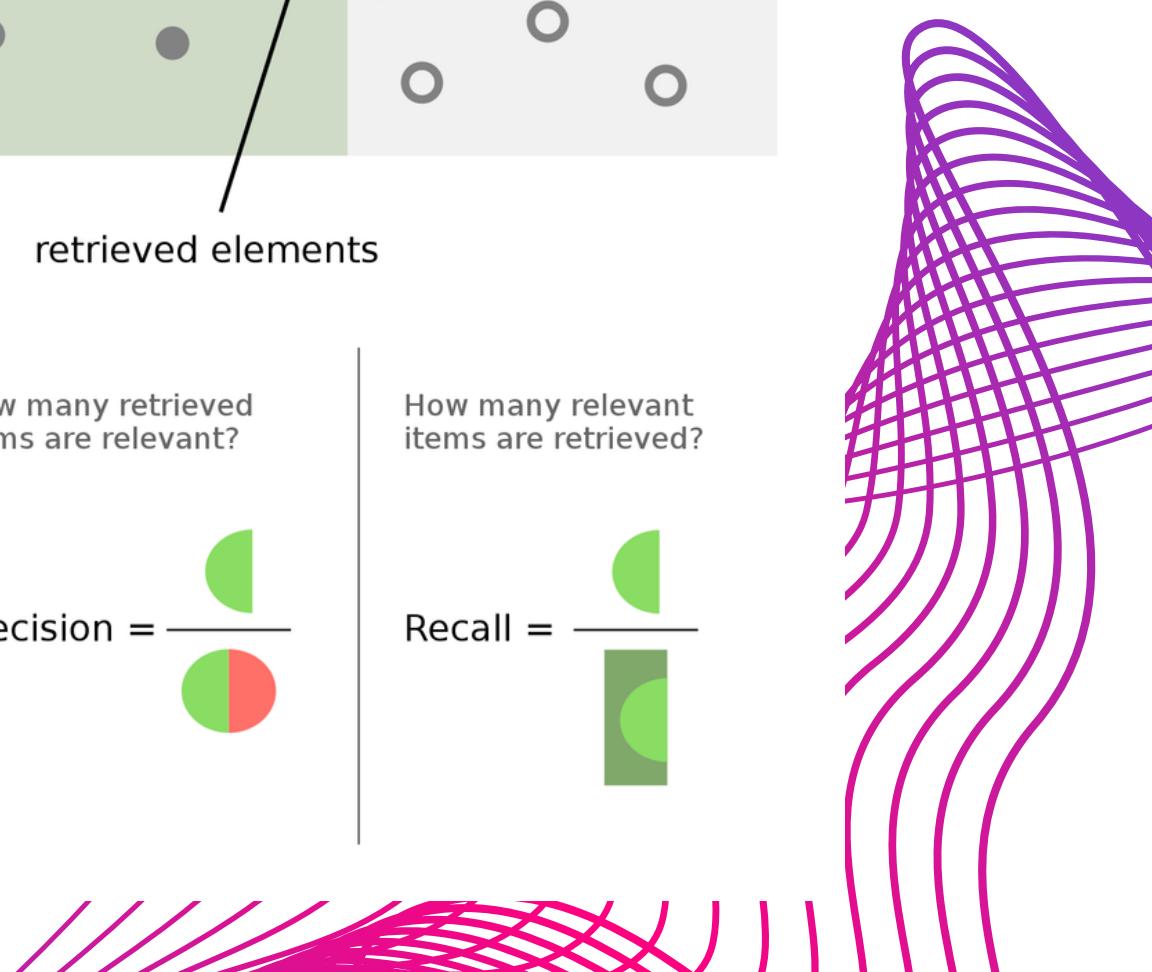


How many retrieved items are relevant?

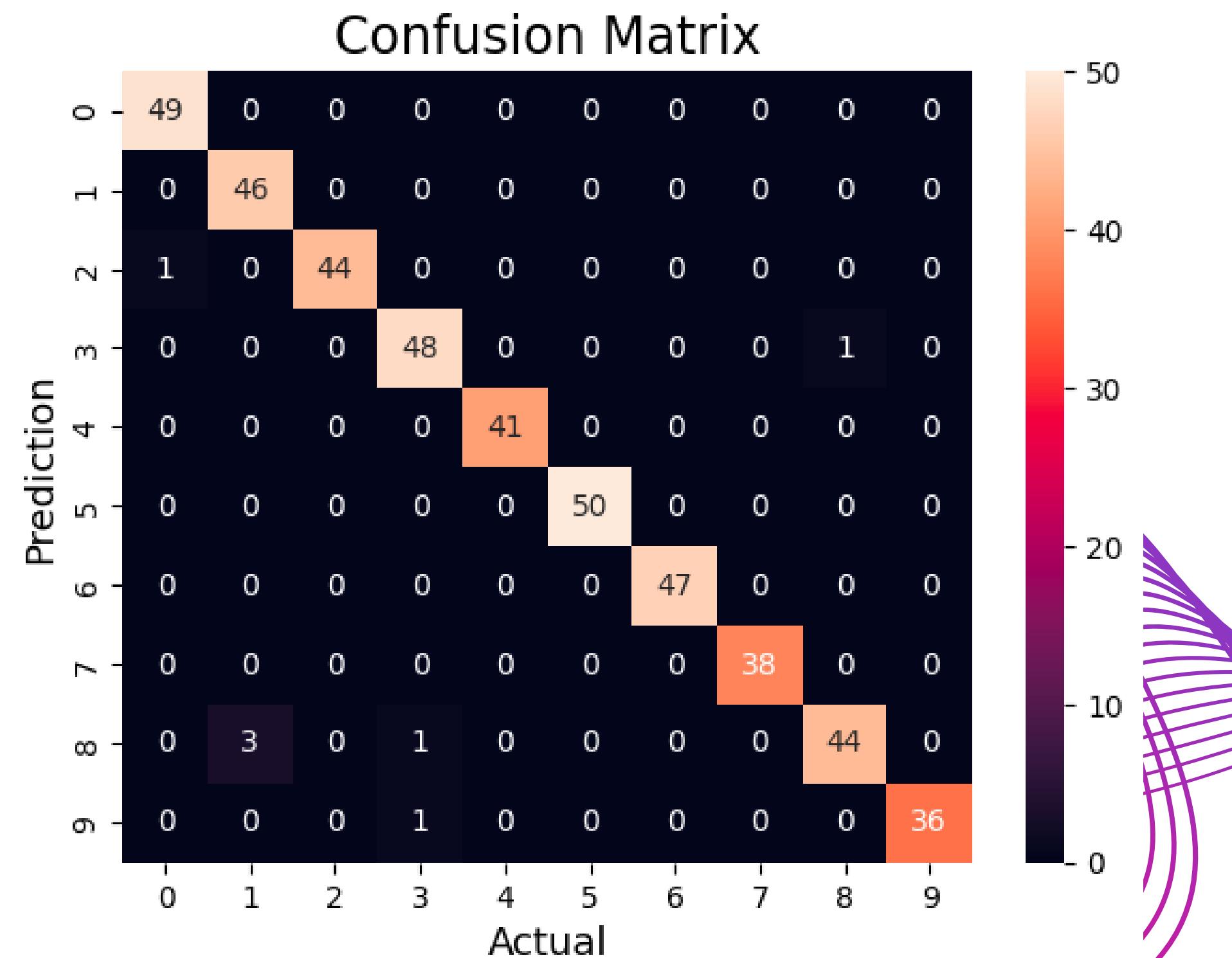
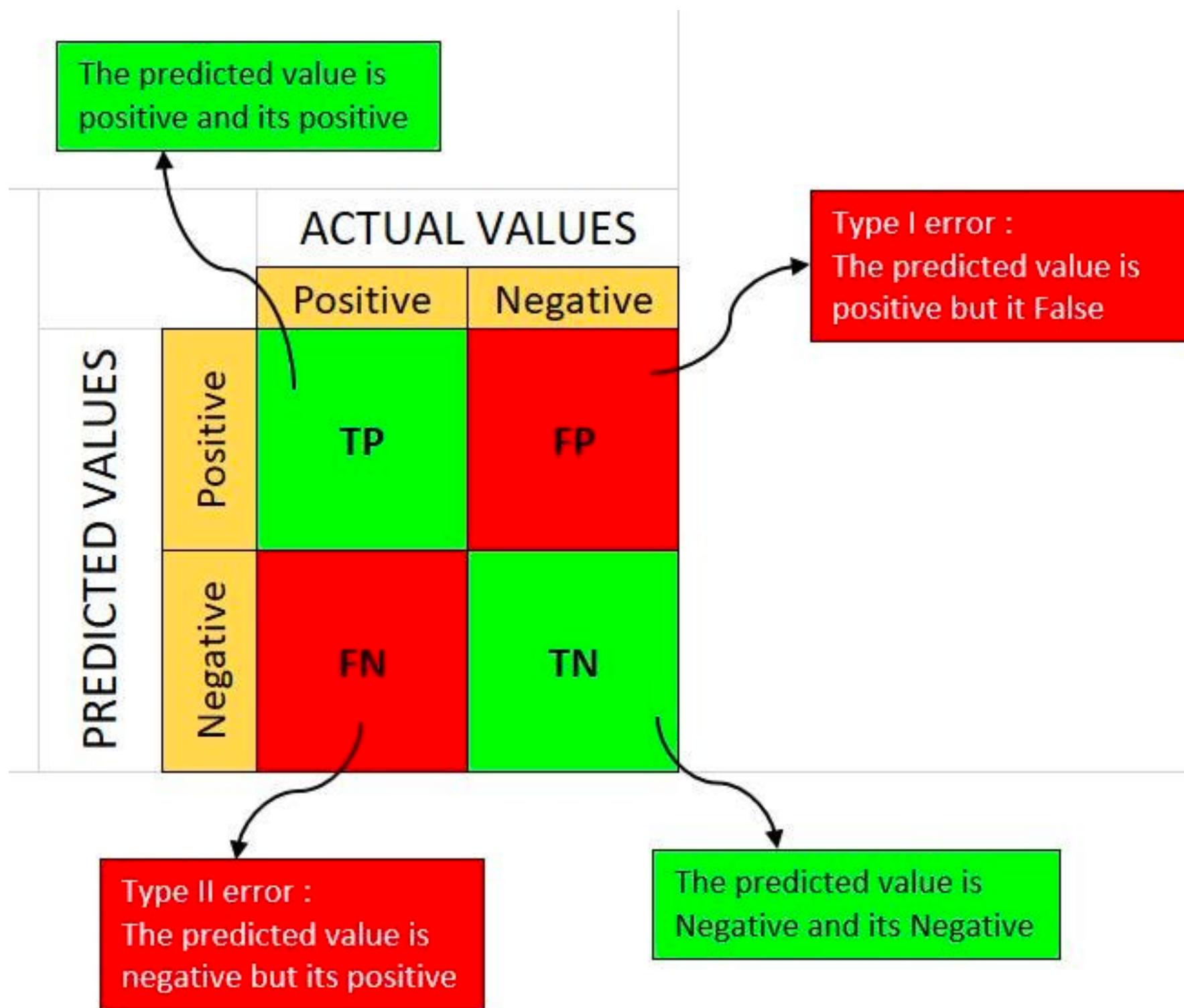
How many relevant items are retrieved?

Precision = $\frac{\text{green}}{\text{green} + \text{red}}$

Recall = $\frac{\text{green}}{\text{green} + \text{light green}}$

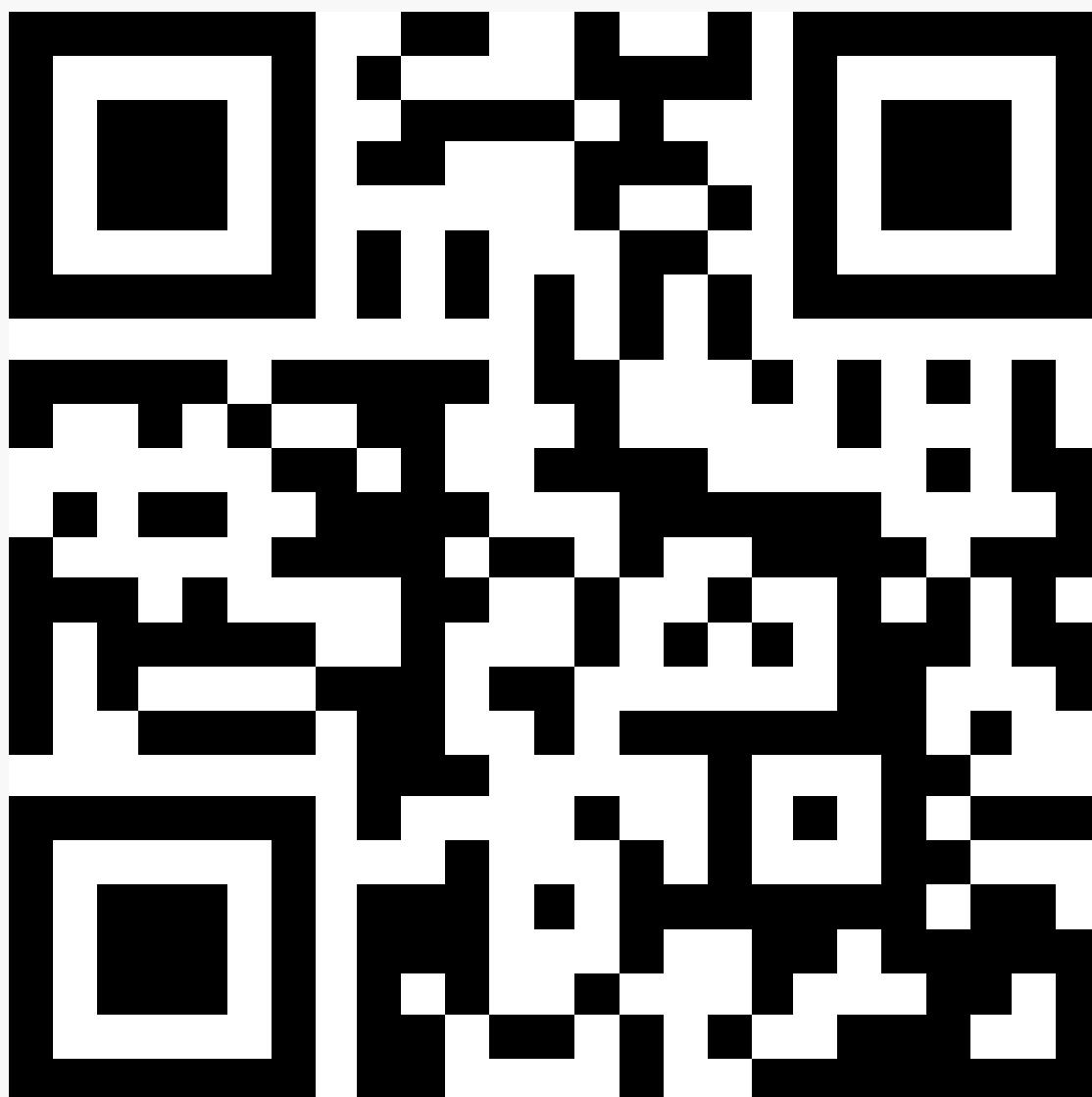


Confusion Matrix

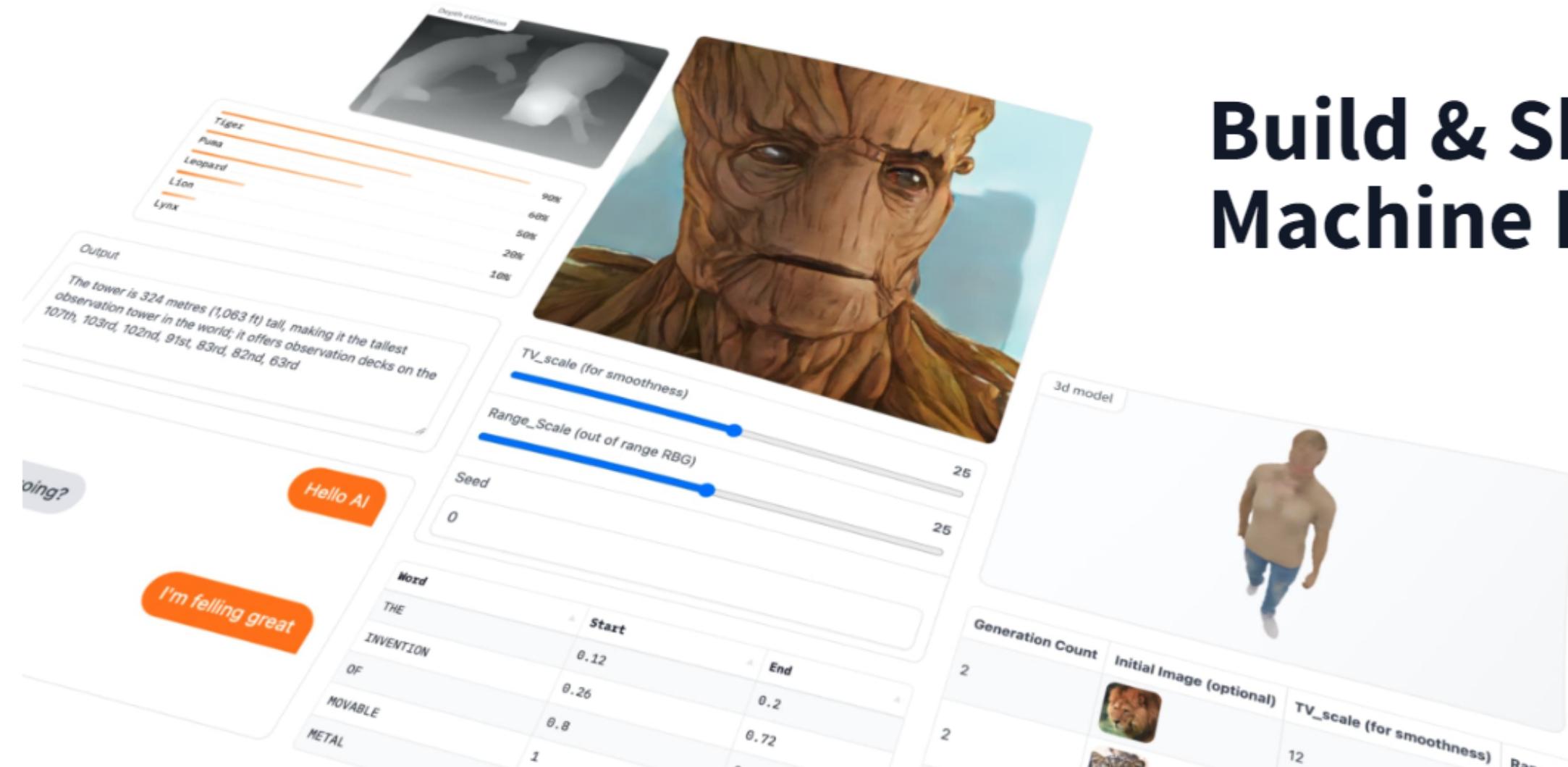


Let's code!

<https://tinyurl.com/JCSSECLASS>



Interface (Gradio)



Build & Share Delightful Machine Learning Apps

Gradio is the fastest way to demo your machine learning model with a friendly web interface so that anyone can use it, anywhere!

Get Started

Star

19,094

<https://gradio.app/>

```
def greet(name):  
    return "Hello " + name + "!"
```

```
demo = gr.Interface(fn=greet, inputs="text", outputs="text")
```

The image shows a Gradio interface for a simple greeting application. The interface consists of two main sections: an input section on the left and an output section on the right.

Input Section: Contains a text input field labeled "name". Below it are two buttons: "Clear" (gray background) and "Submit" (orange background).

Output Section: Contains a text input field labeled "output".

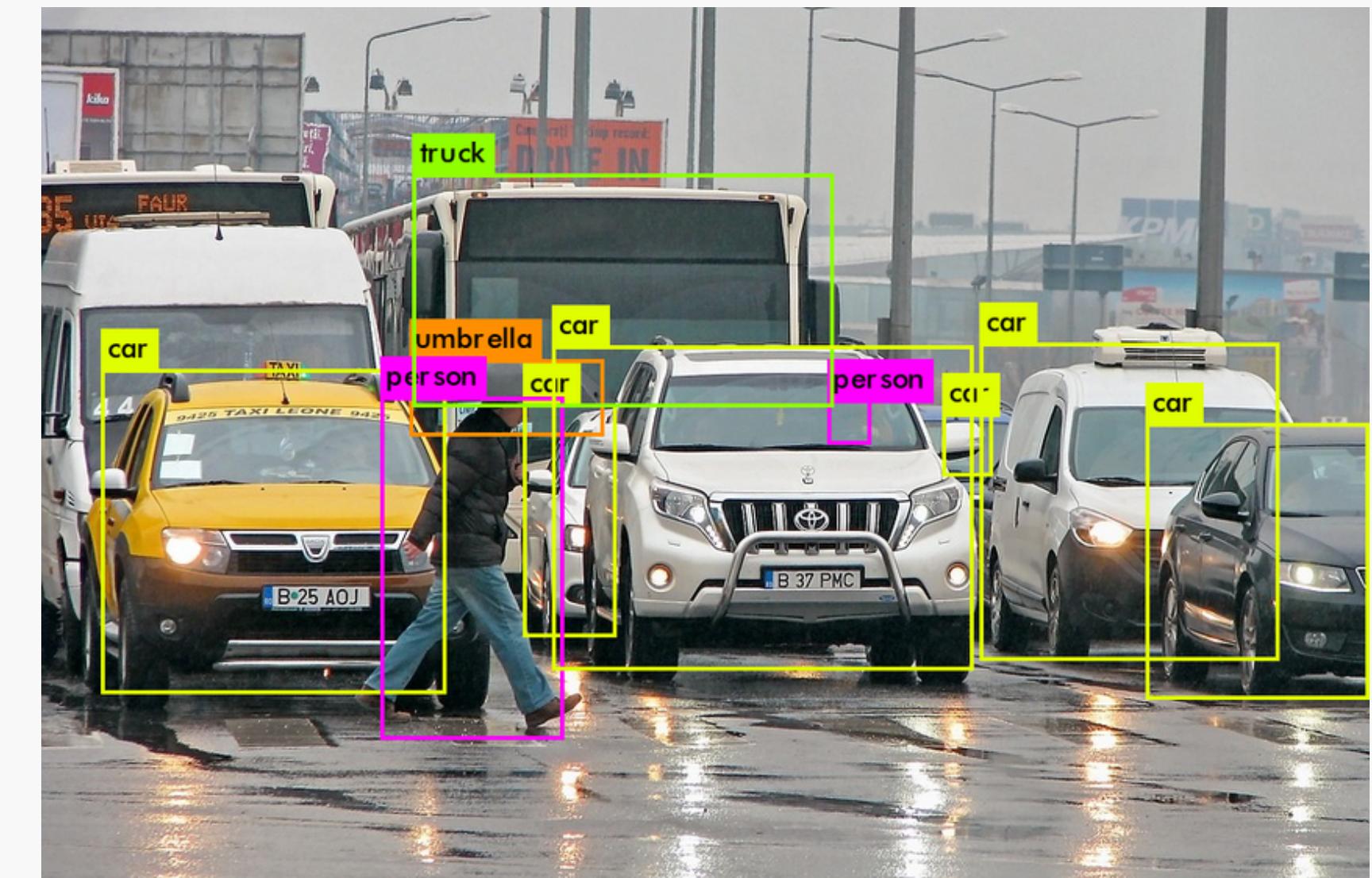
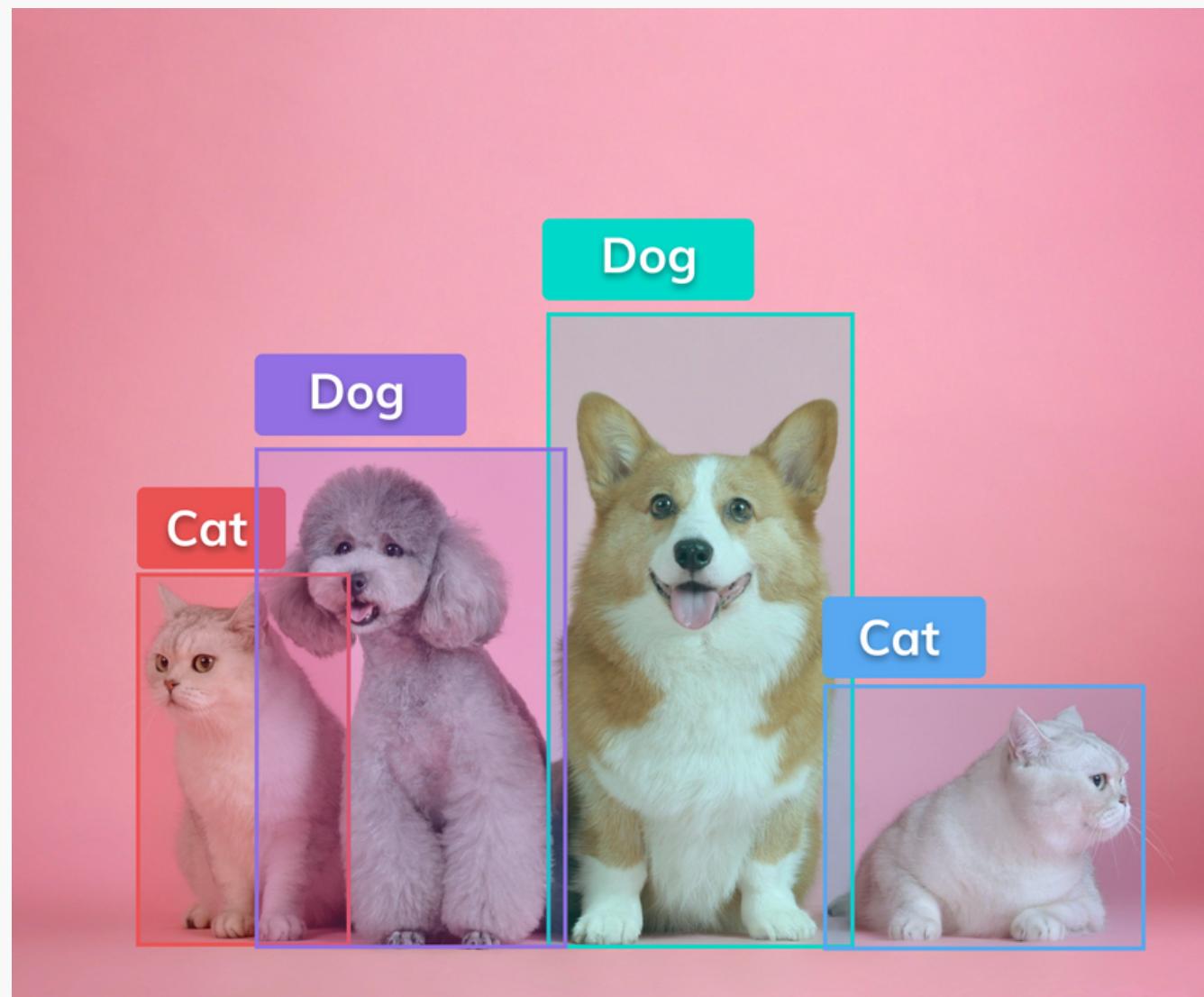
Footer: At the bottom, there is a note: "gradio/hello_world built with Gradio." and "Hosted on Spaces".

5 Minutes Break

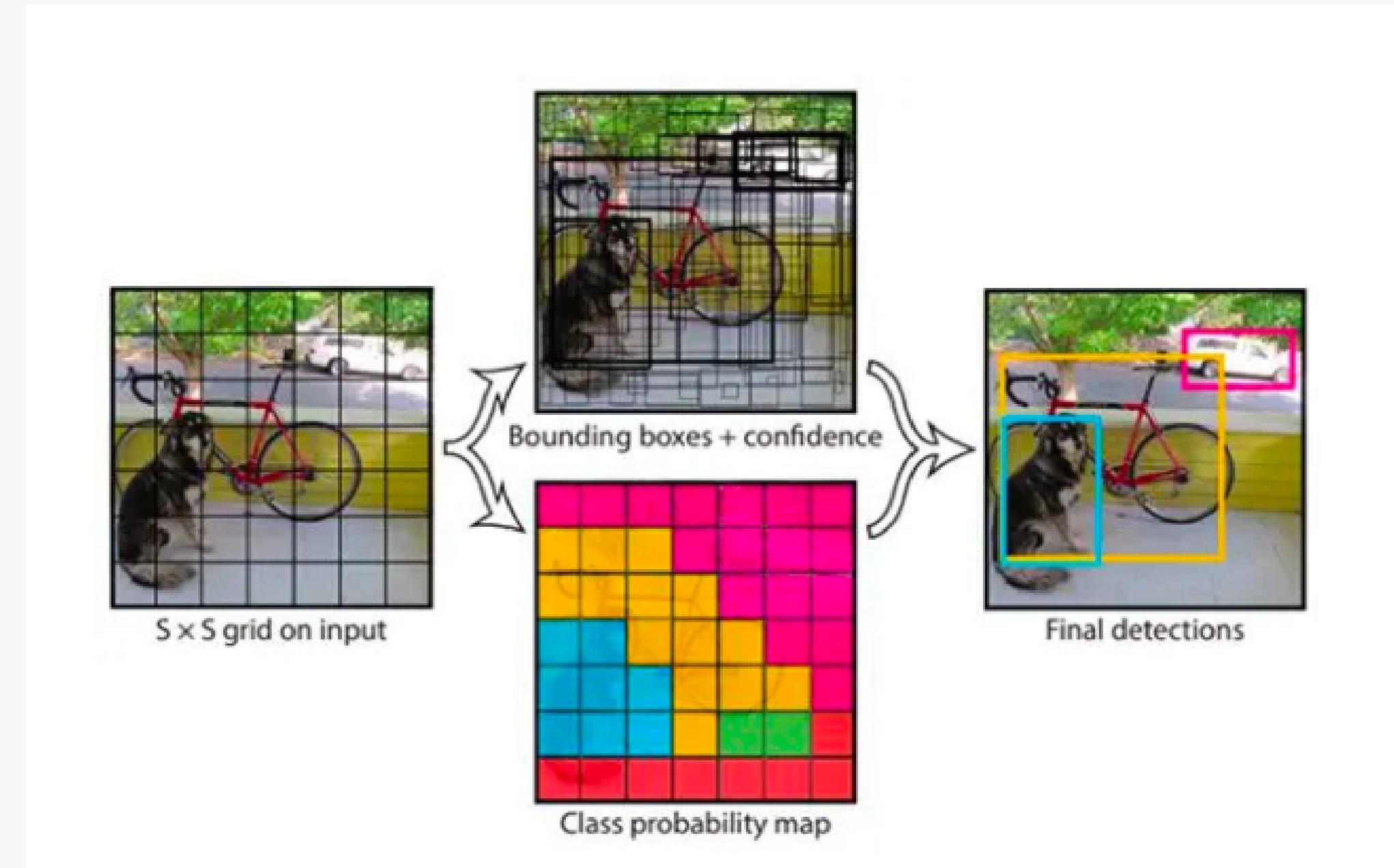
Object Detection

Object Detection

Object Detection หมายถึงกระบวนการในการตรวจจับและระบุตัวถูกที่ปรากฏในภาพหรือวิดีโอ ซึ่งเป็นหนึ่งในงานวิจัยและปัญหาที่น่าสนใจในด้านปัญญาประดิษฐ์ และ computer vision

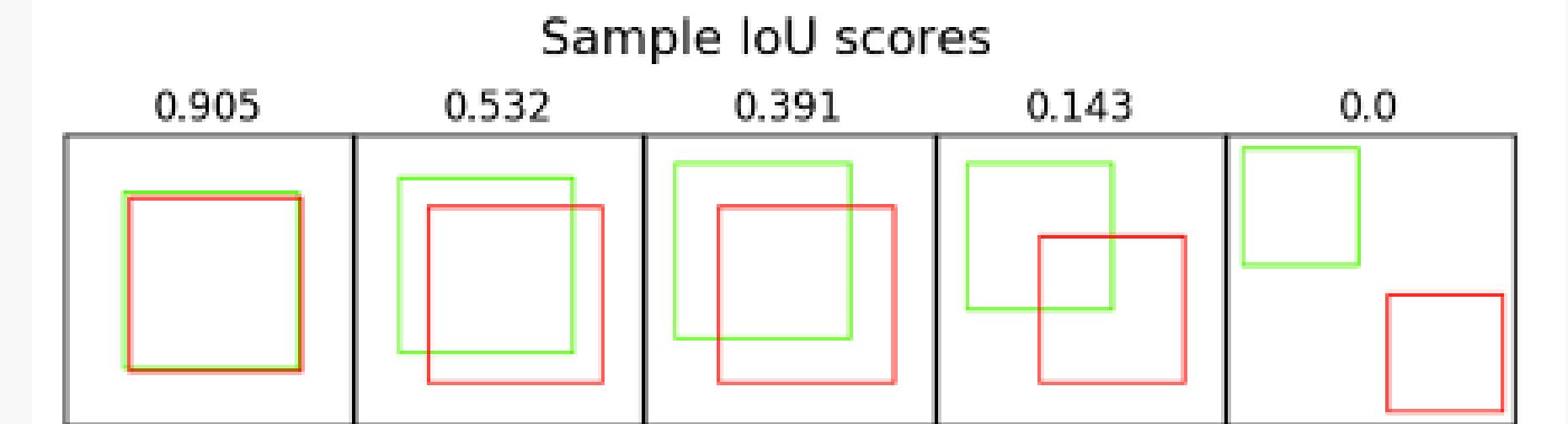
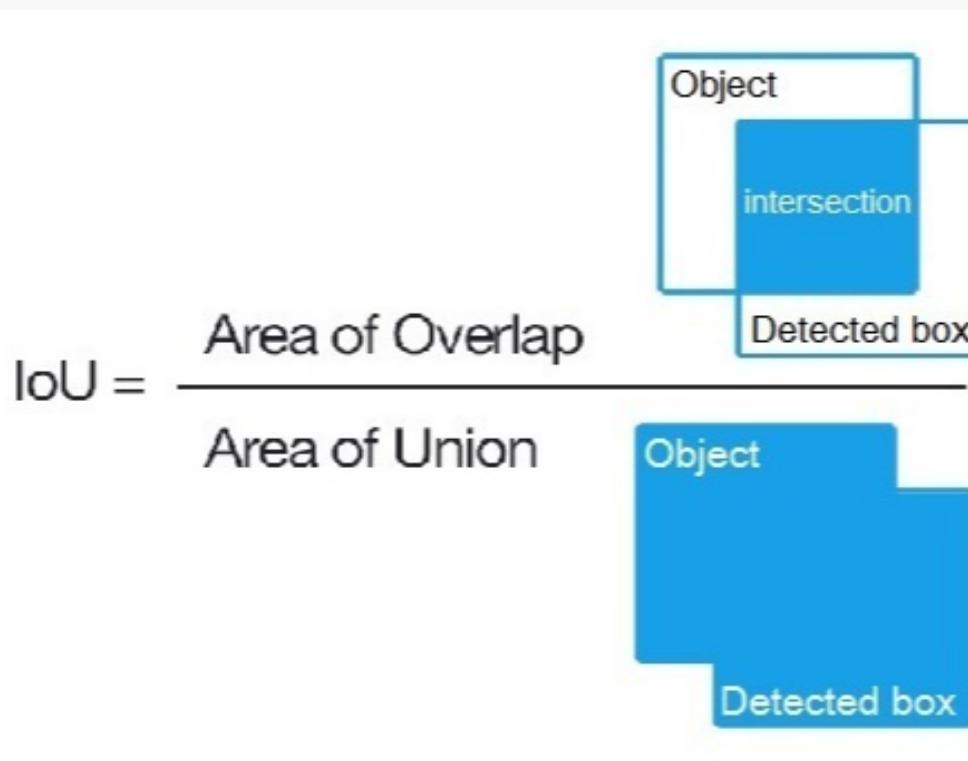
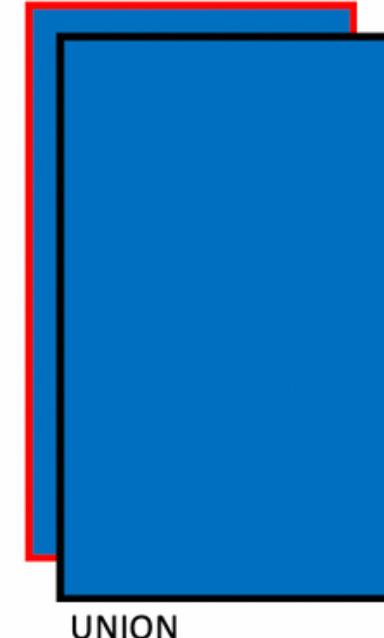
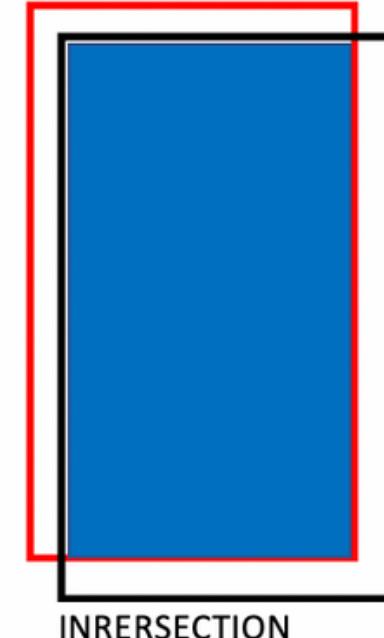


Object Detection (YOLO)



YOLO แบ่งการคำนวณวัตถุในภาพเป็น Grid (เช่น 7×7) และใช้การคำนวณวัตถุร่วมกับ bounding box ในแต่ละ grid ในการหา bounding box และชนิดของวัตถุ

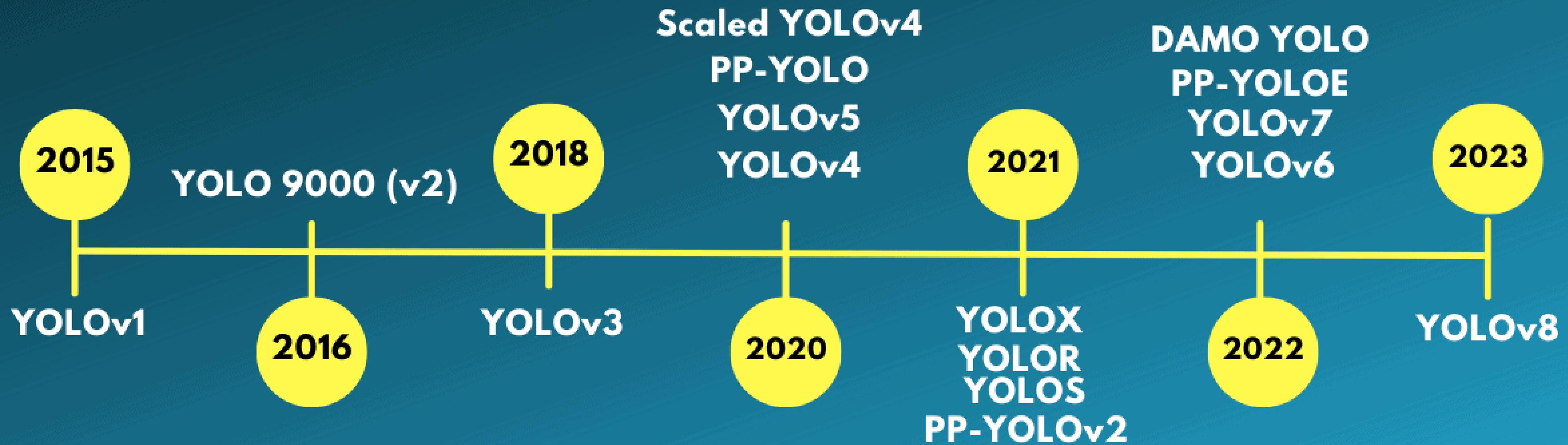
Object Detection (YOLO)



Object Detection (YOLO)



YOLO Object Detection Models Timeline



YOLO format

```
<object-class> <x> <y> <width> <height>
```

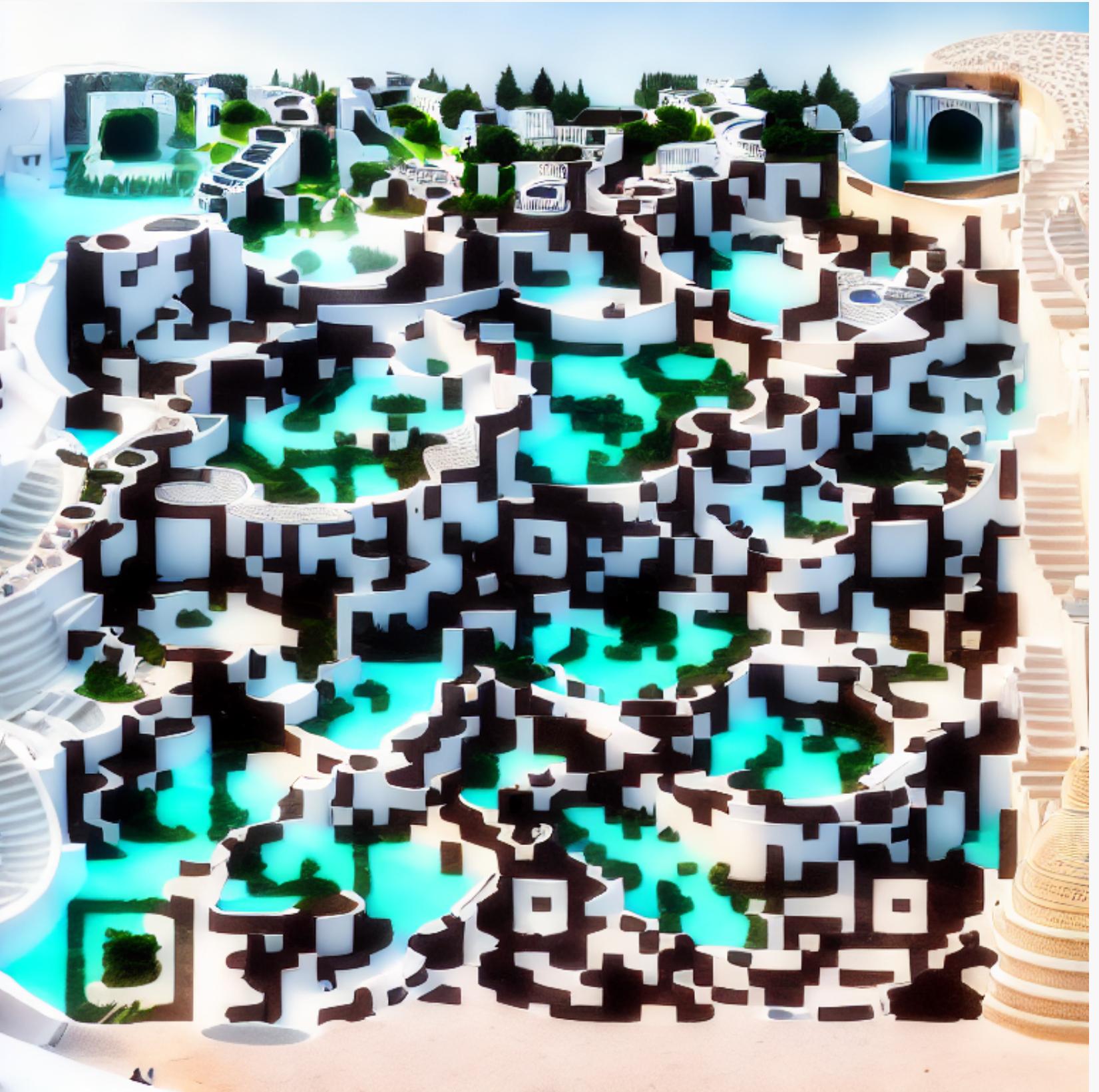


0	45	55	29	67
1	99	83	28	44

Let's code!

https://colab.research.google.com/drive/1inp1oQd7S9cZUvtF1-31-qyS_8_S0R1s?usp=sharing

<https://tinyurl.com/jcsse23-yolo>



10 Minutes Break

Stable Diffusion

Let's make images!

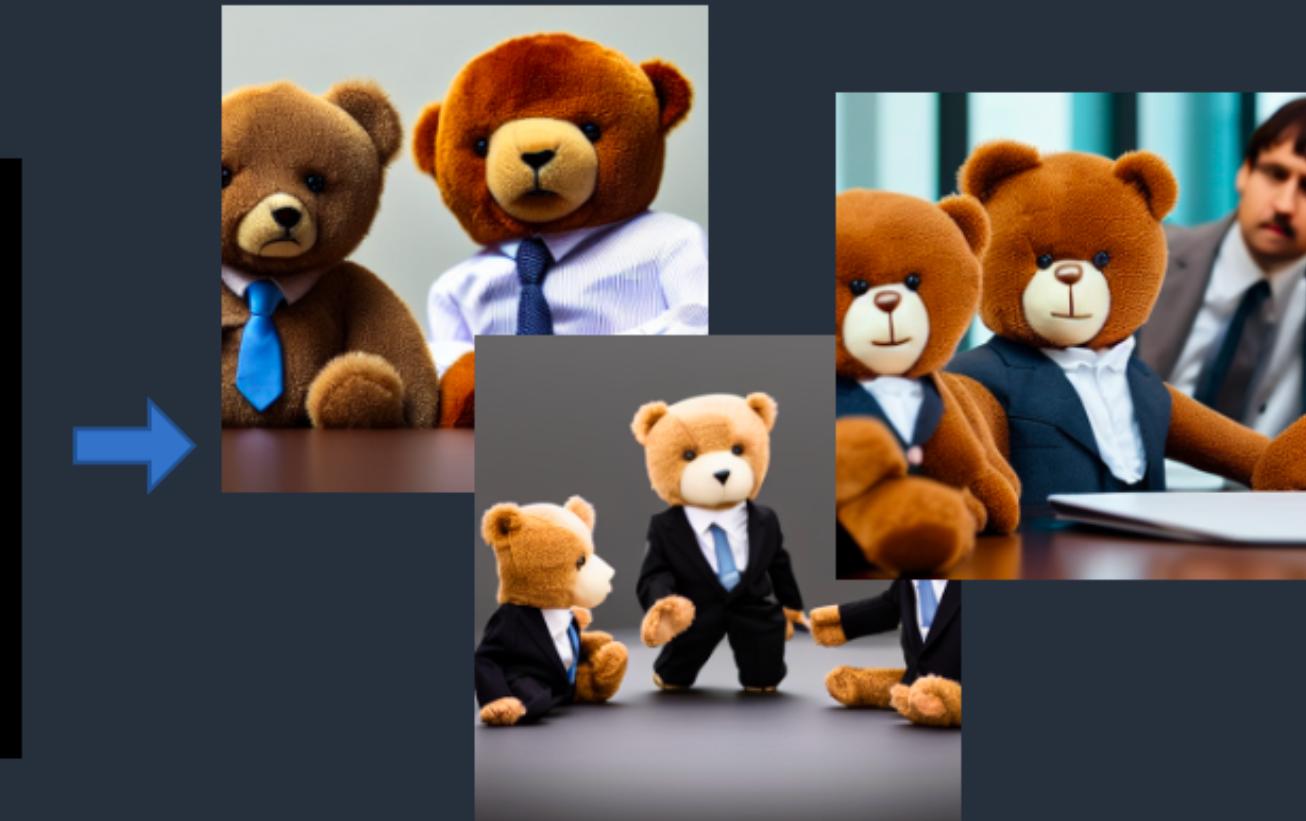
<https://colab.research.google.com/drive/1p8D7yh0x16fTYdqupZ-74ZBqNDy72t0U?usp=sharing>

<https://tinyurl.com/jcsse23-sd>

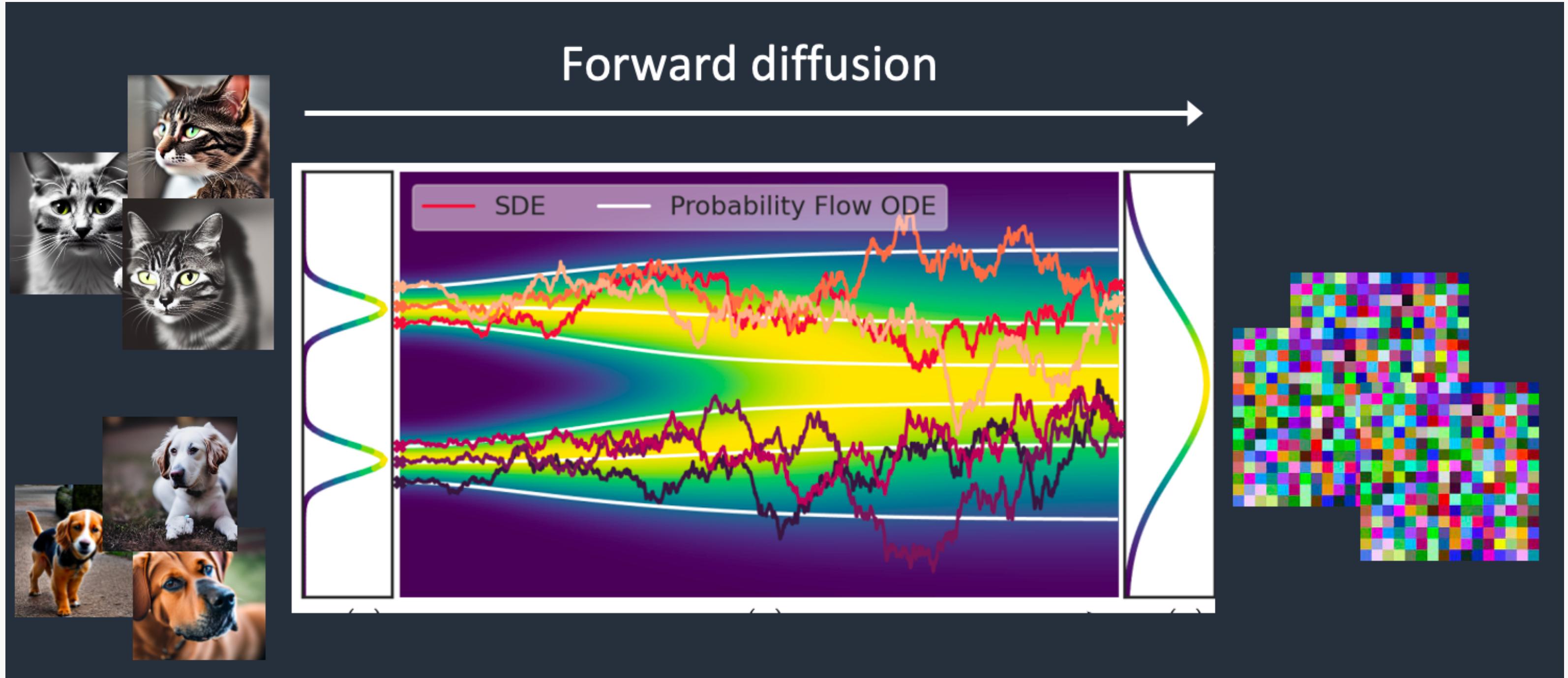


Stable Diffusion

“Teddy bears wearing suit discussing a business proposal around office table”



Stable Diffusion

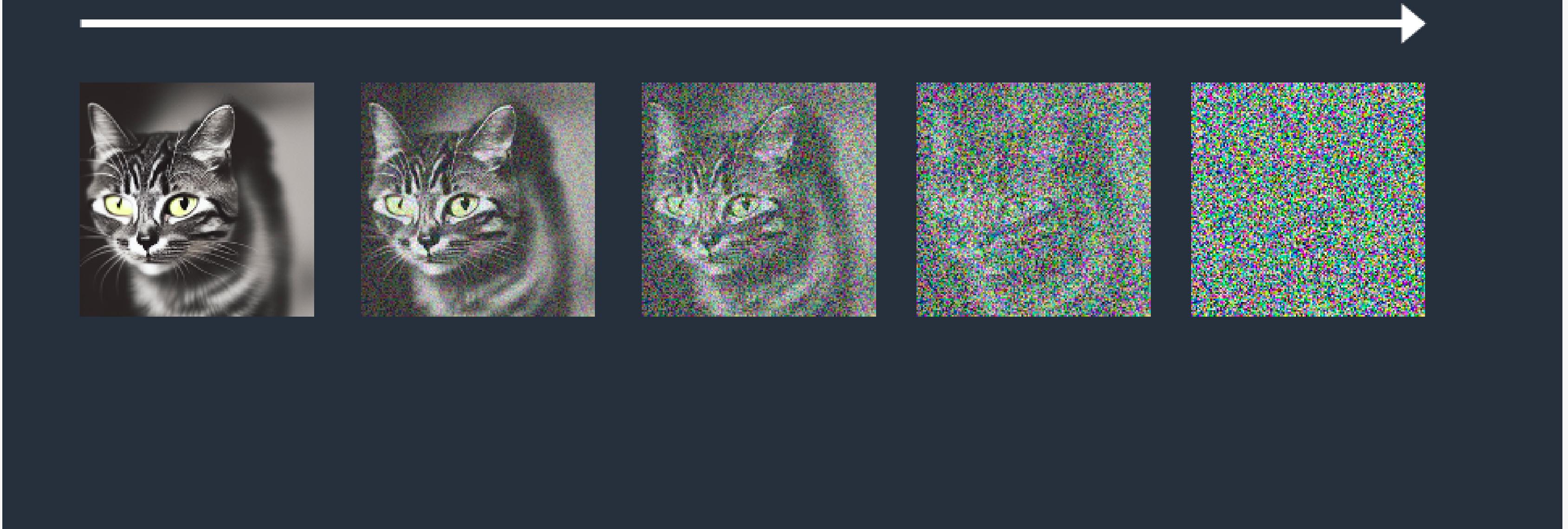


<https://stable-diffusion-art.com/how-stable-diffusion-work/>

Stable Diffusion

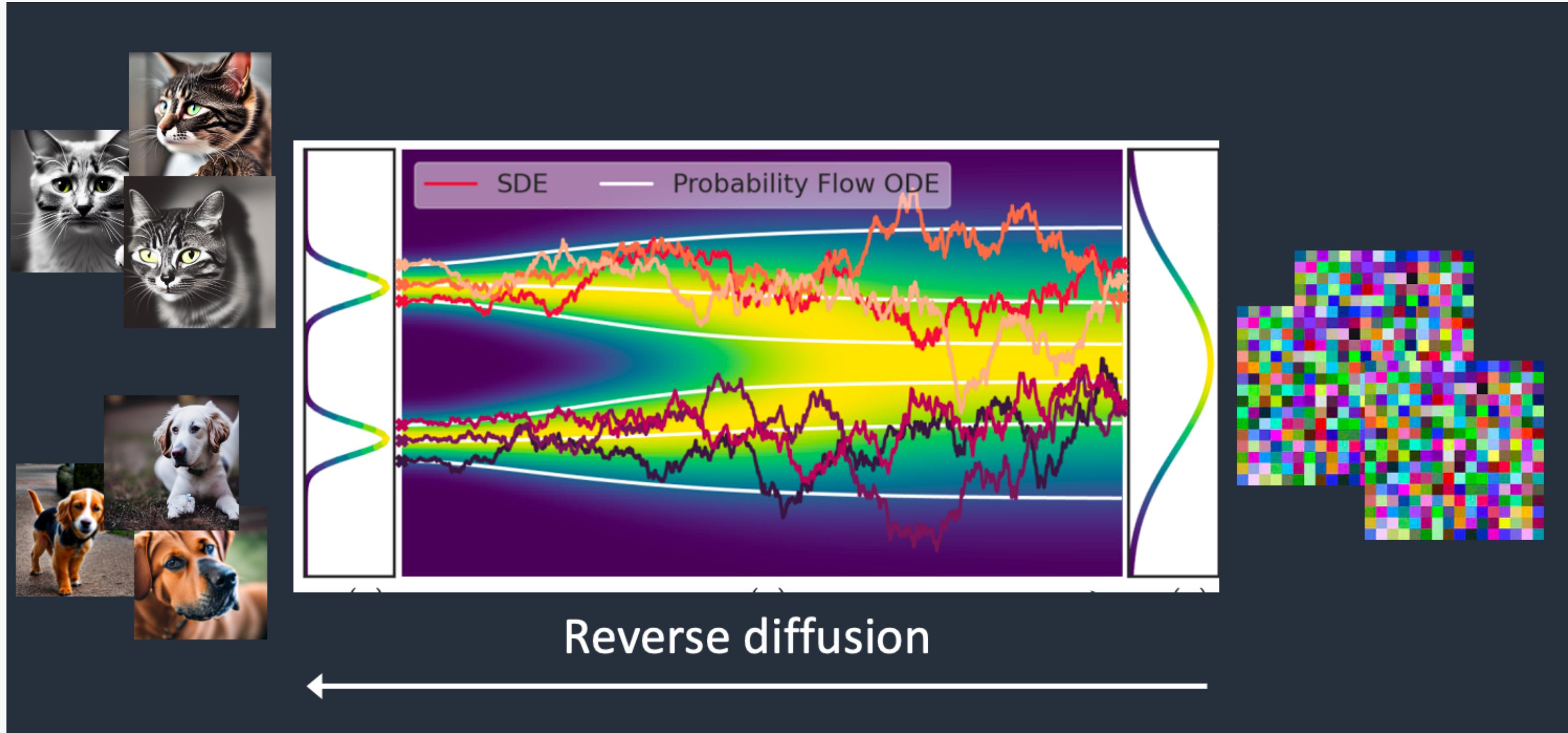


Forward diffusion



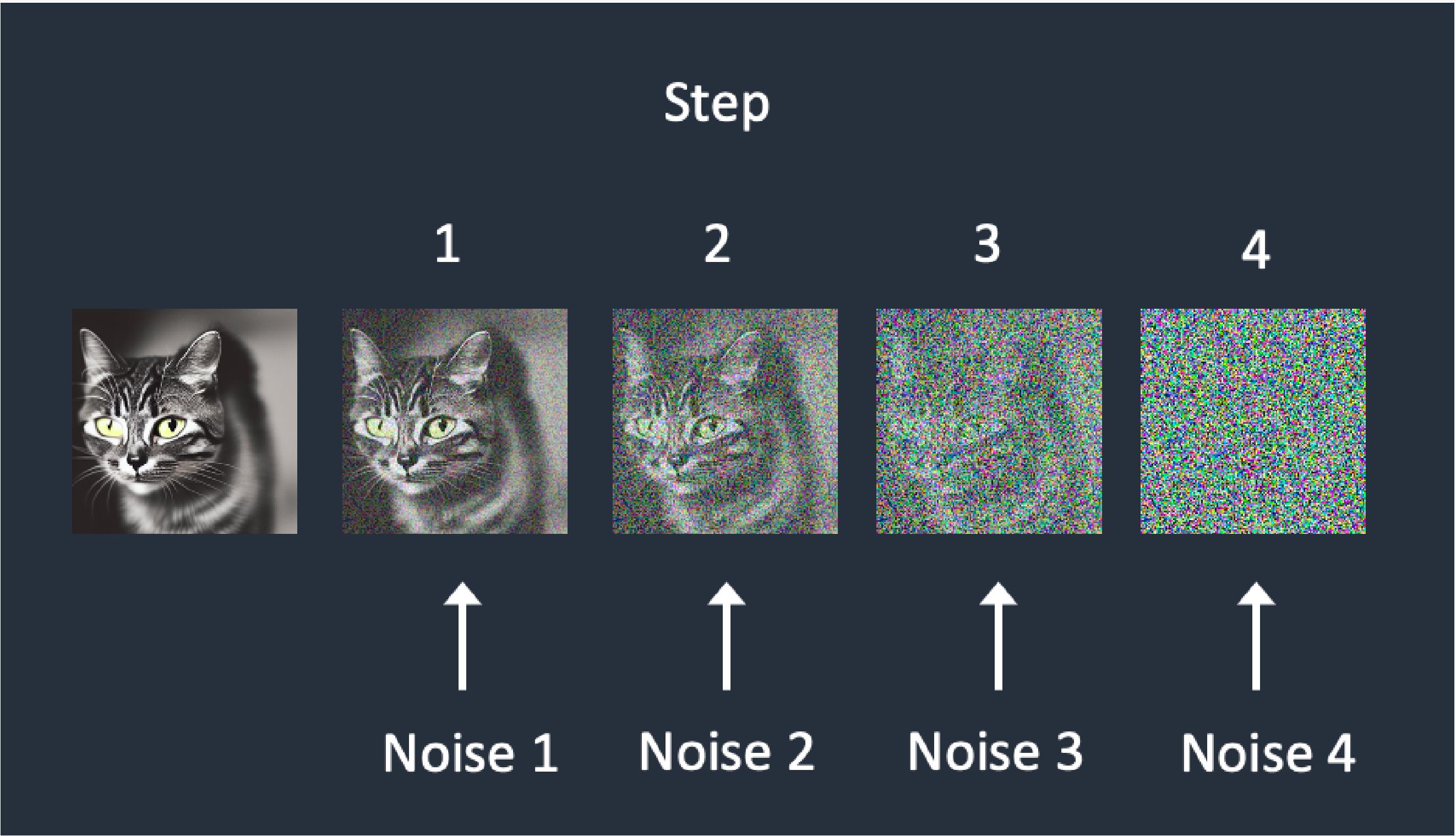
<https://stable-diffusion-art.com/how-stable-diffusion-work/>

Stable Diffusion



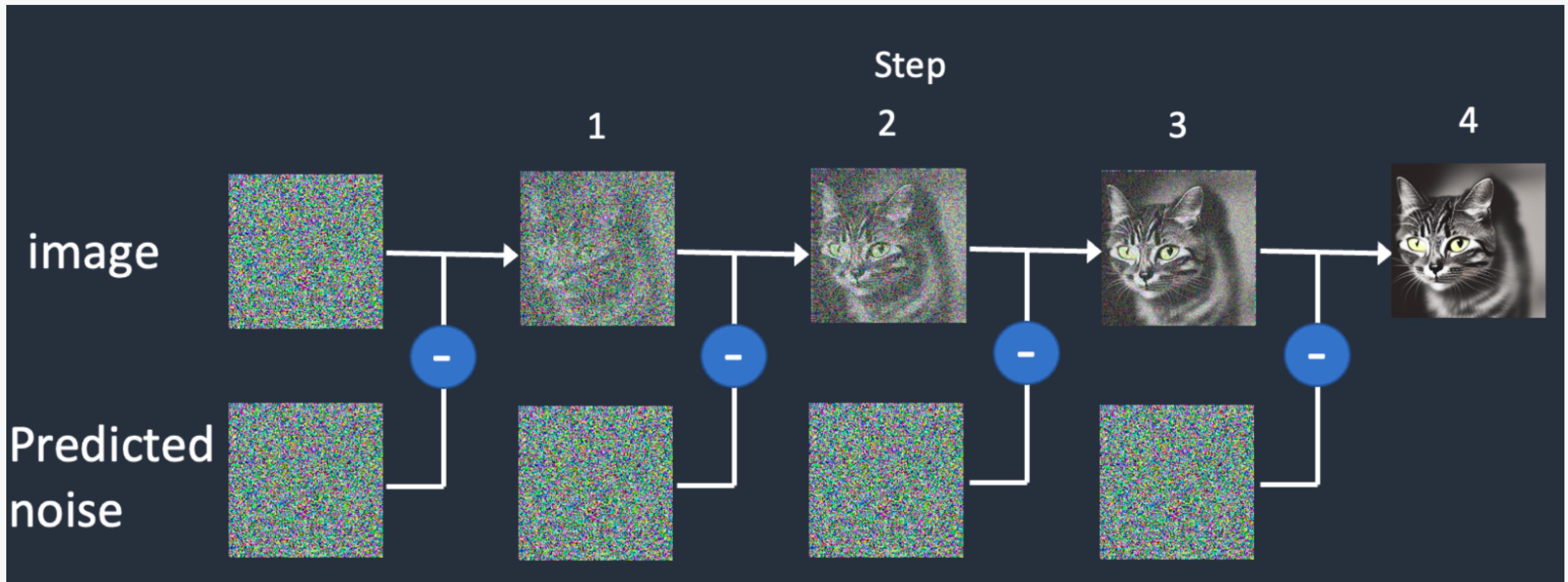
<https://stable-diffusion-art.com/how-stable-diffusion-work/>

Stable Diffusion



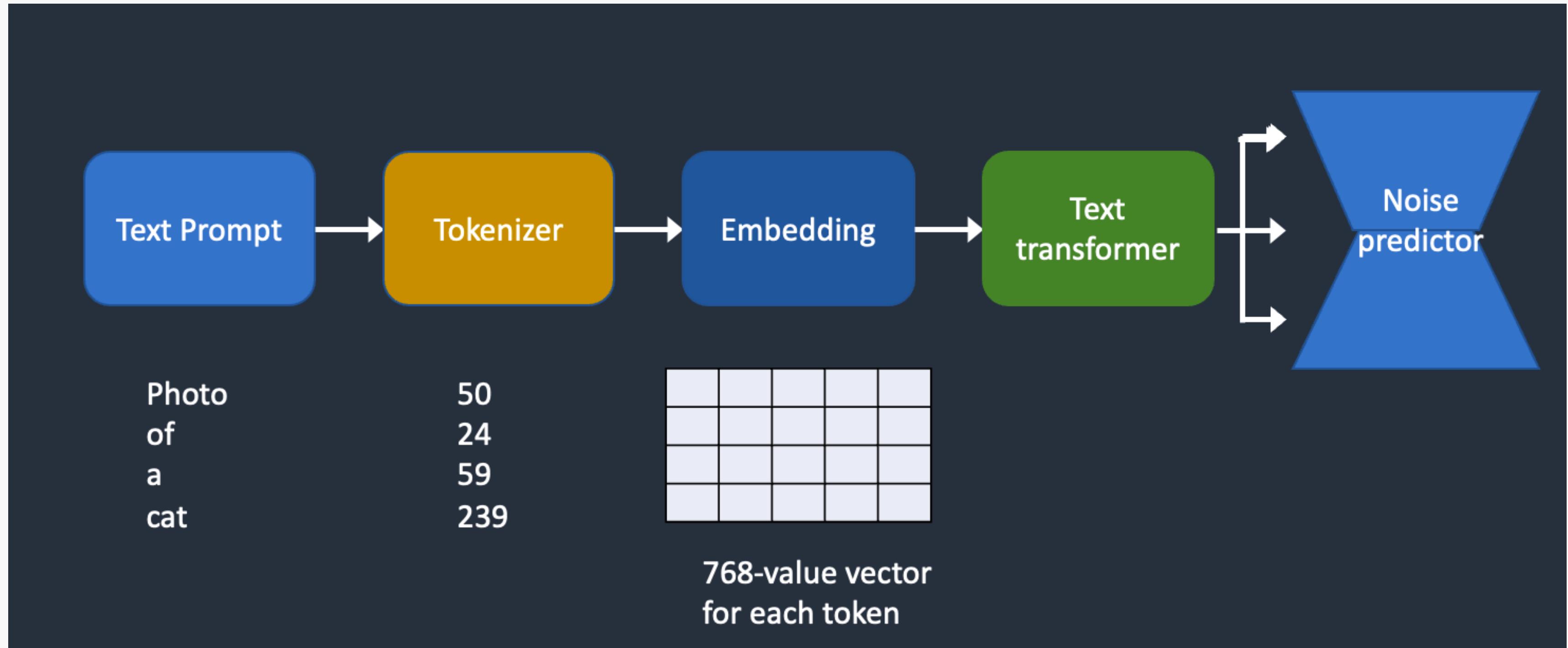
<https://stable-diffusion-art.com/how-stable-diffusion-work/>

Stable Diffusion

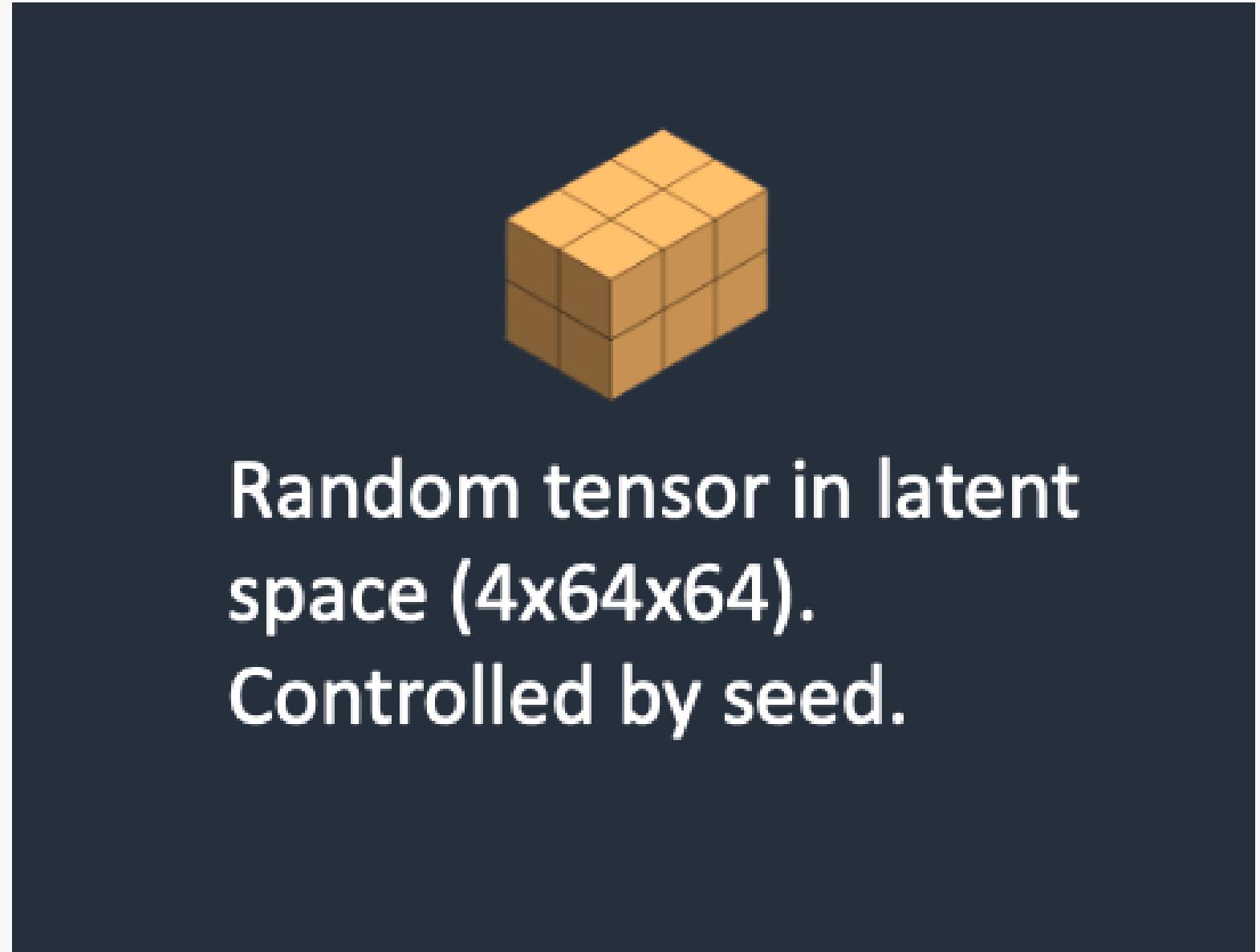


<https://stable-diffusion-art.com/how-stable-diffusion-work/>

Stable Diffusion

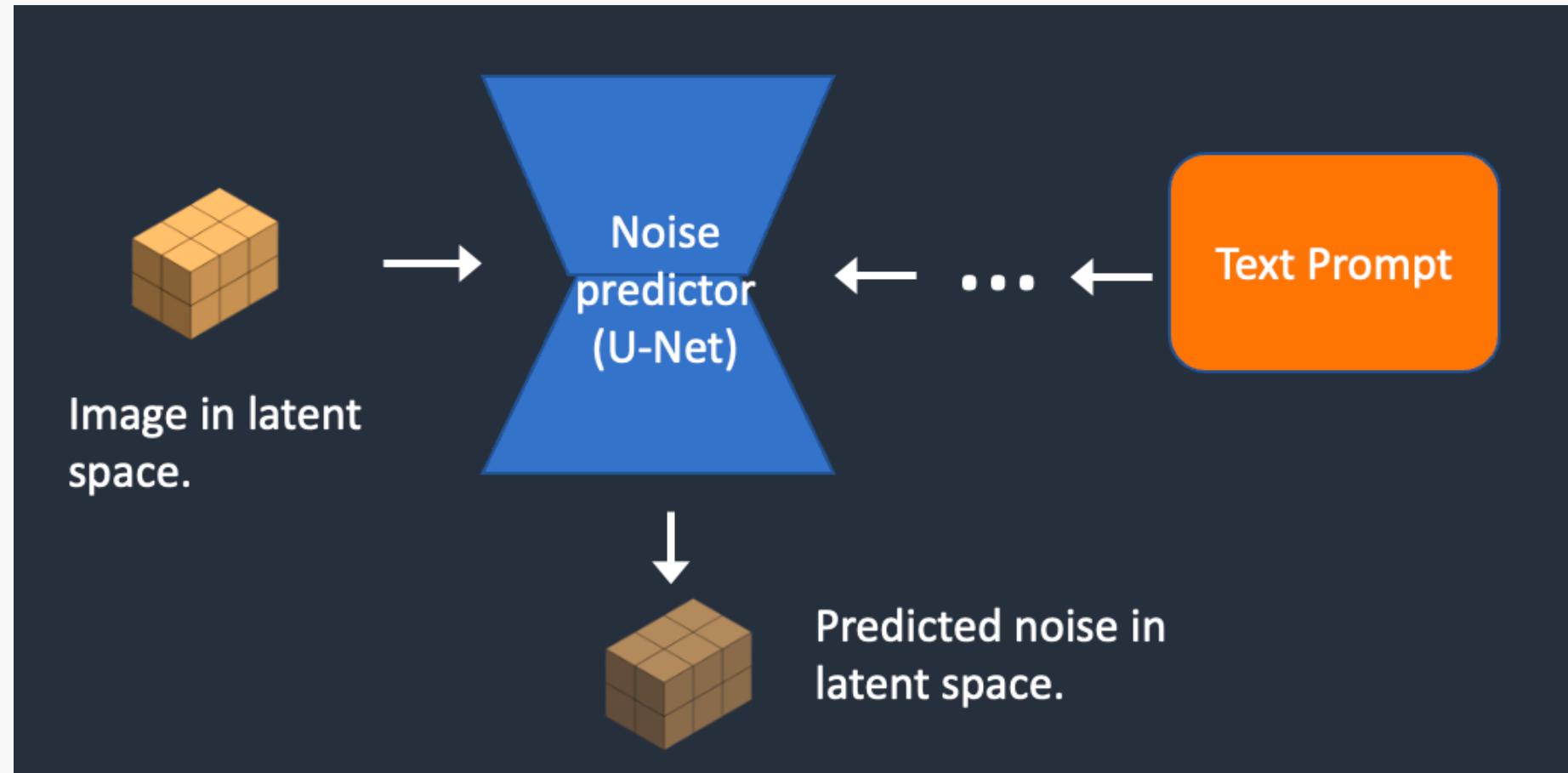


Stable Diffusion - Step by step



Step 1. สร้าง noise โดยการสุ่ม (จาก seed)

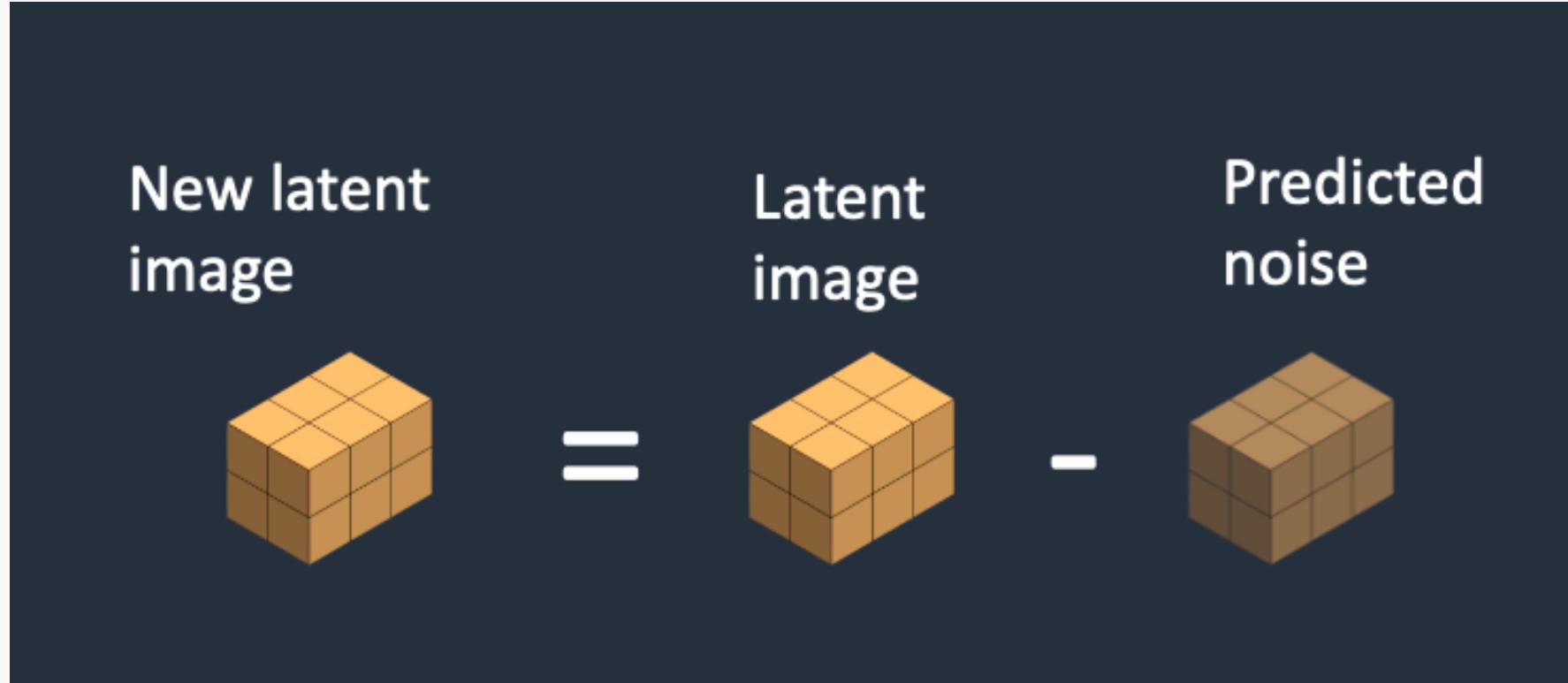
Stable Diffusion - Step by step



Step 1. สร้าง noise โดยการสุ่ม (จาก seed)

Step 2. กำนยา noise (จาก step 1)

Stable Diffusion - Step by step



Step 1. สร้าง noise โดยการสุ่ม (จาก seed)

Step 2. กำนวย noise (จาก step 1)

Step 3. ลบ noise จาก step 2 ออกจาก step 1

- ทำ 20-30 รอบ (sampling)

Stable Diffusion - Step by step



Step 1. สร้าง noise โดยการสุ่ม (จาก seed)

Step 2. กำนวย noise (จาก step 1)

Step 3. au noise จาก step 2 ออกจาก step 1

- กำ 20-30 รอบ (sampling)

Step 4. ได้รูปที่สวยงาน

