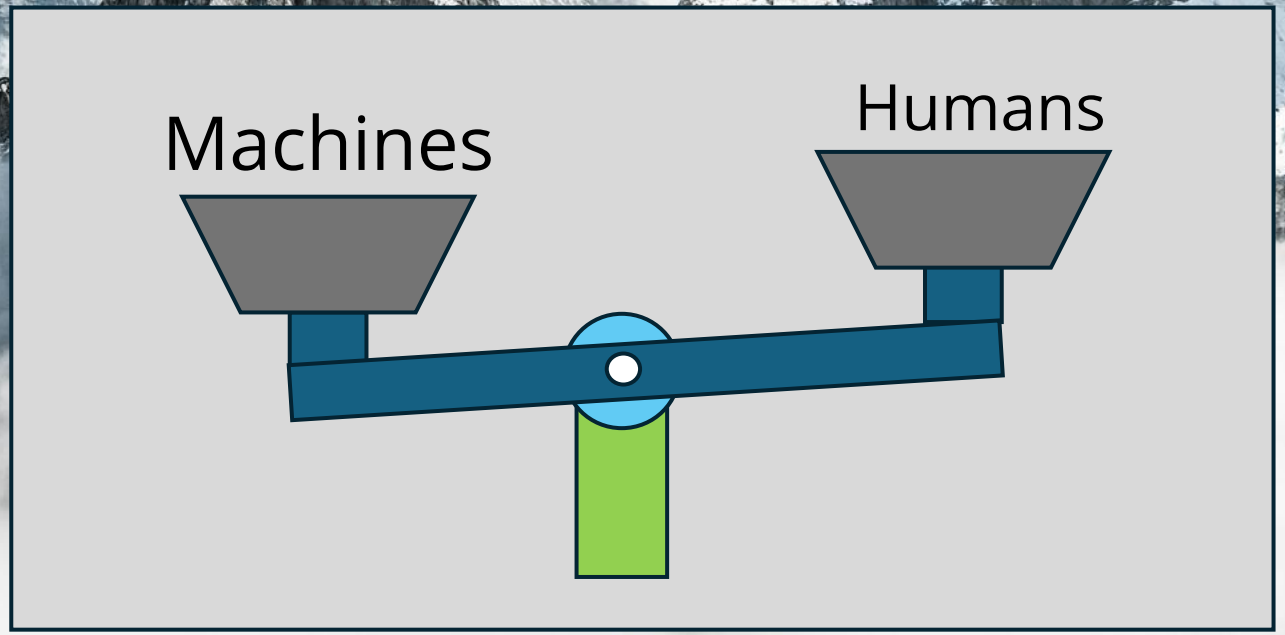
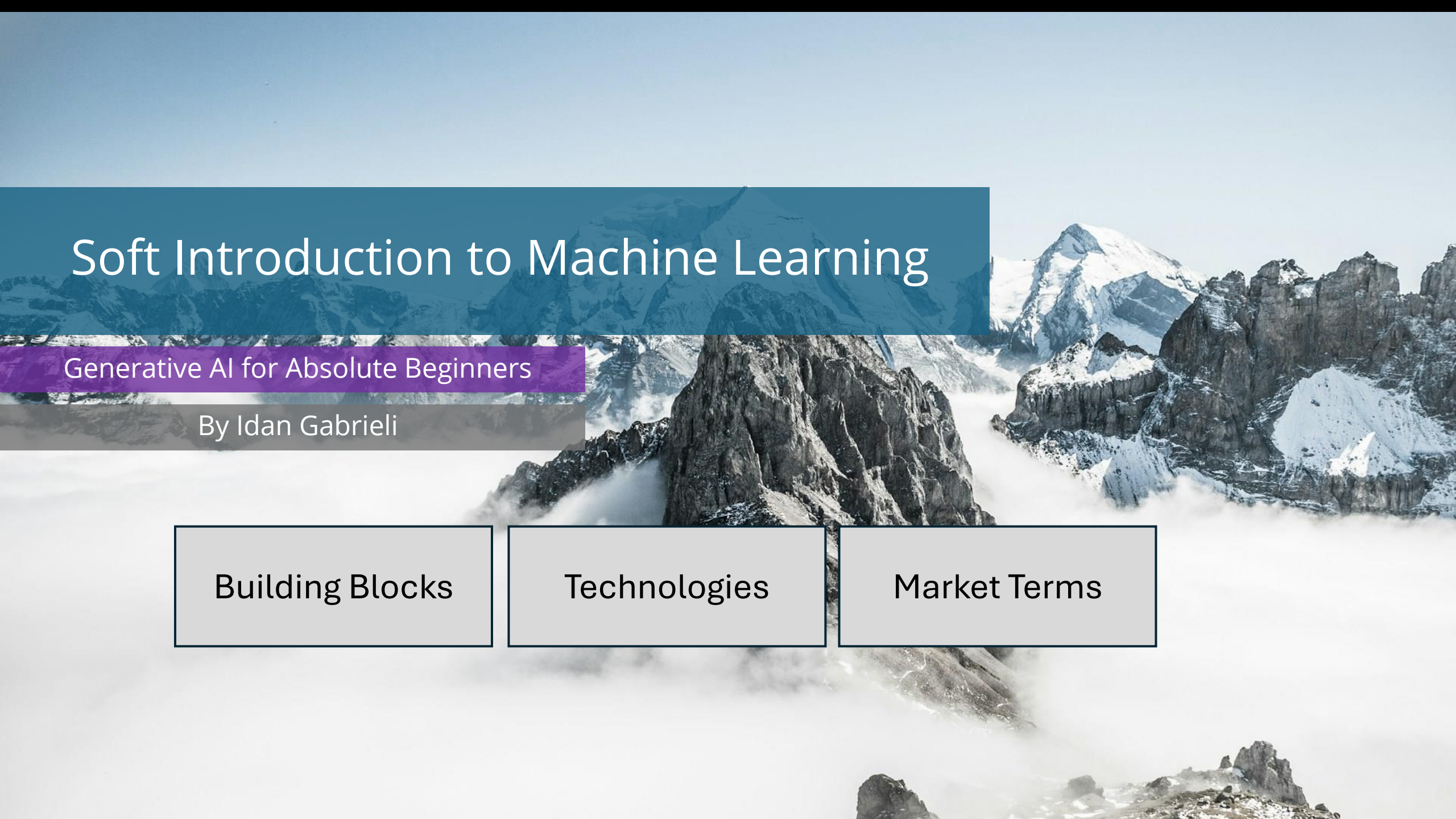
Soft Introduction to Machine Learning

Generative AI for Absolute Beginners

By Idan Gabrieli





Soft Introduction to Machine Learning

Generative AI for Absolute Beginners

By Idan Gabrieli

Building Blocks

Technologies

Market Terms

The ML Box (I/O)

The Box Analogy

INPUT → ML-Solution → OUTPUT

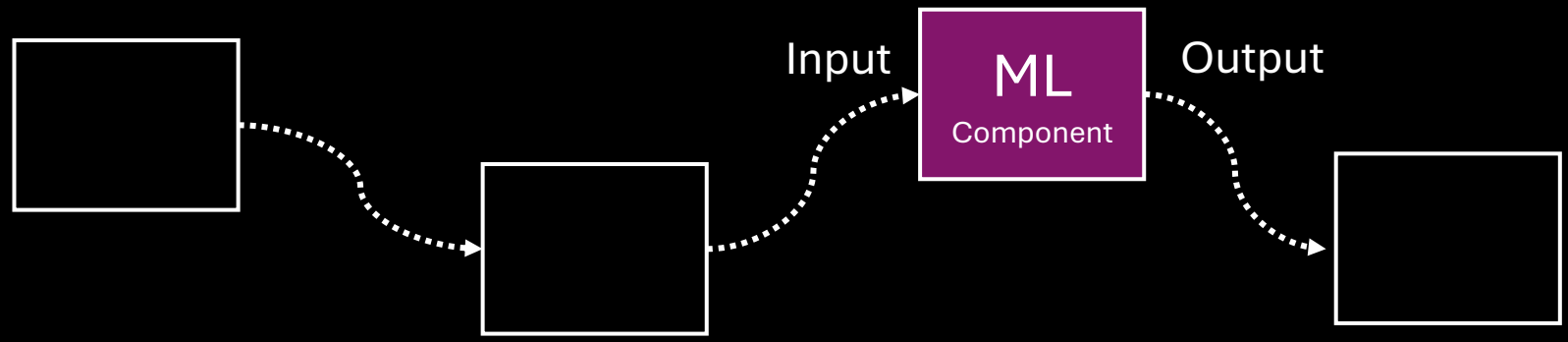
Product Picture → ML Box #1 → **Perfect\Defective**

Product Review (Text) → ML Box #2 → **Positive\Negative**

Story (Text) → ML Box #3 → **Video File**

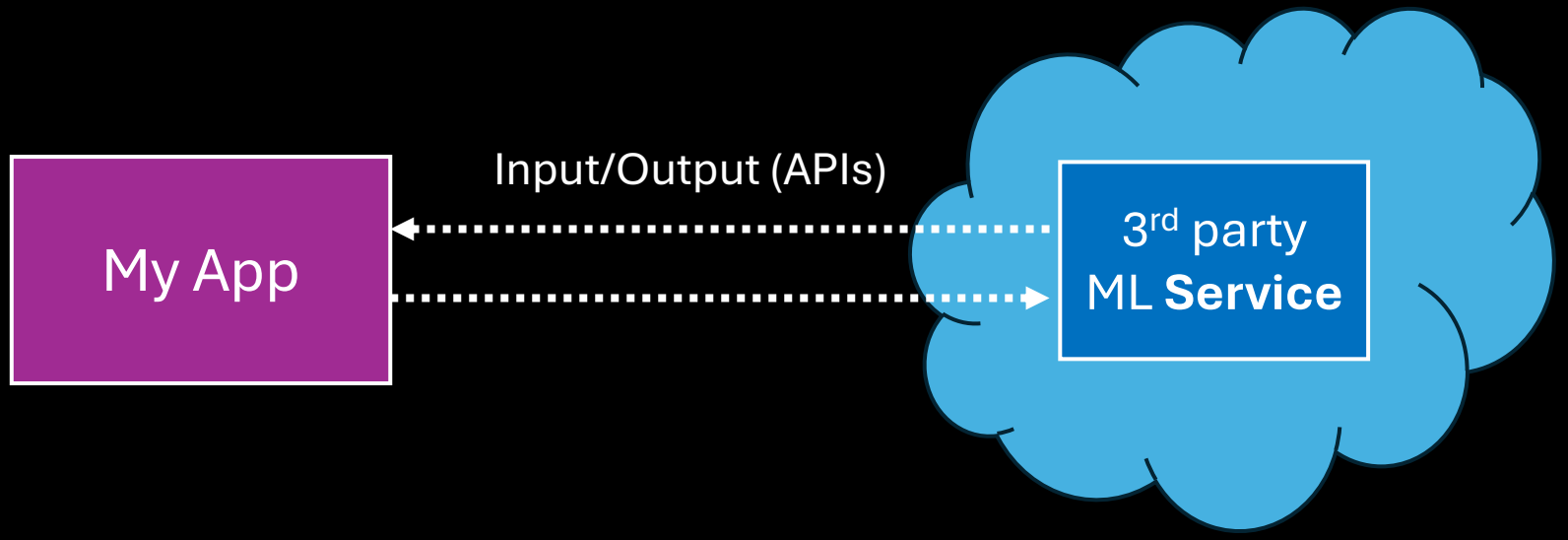
The ML Box (I/O)

The Box Analogy



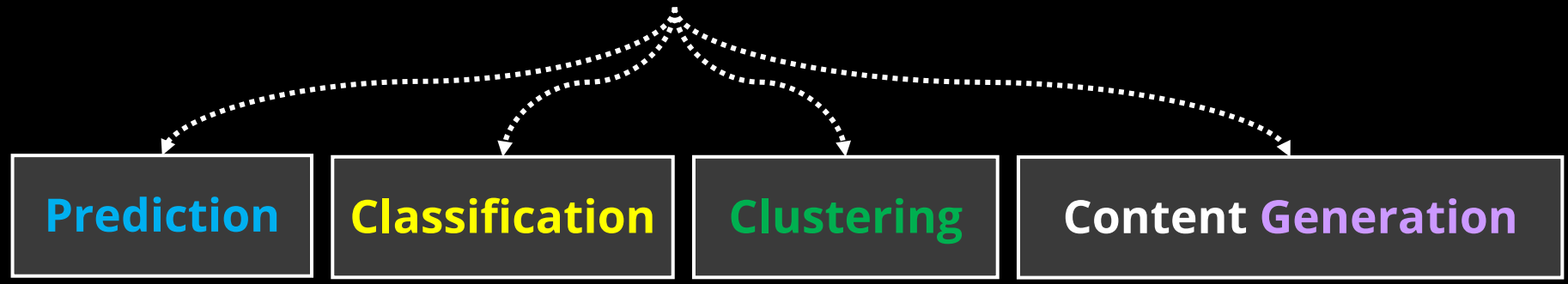
The ML Box (I/O)

The Box Analogy



The Box Analogy

Typical ML Tasks



Prediction

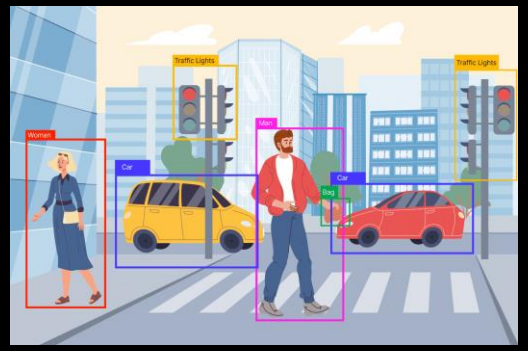
Input Data → ML Box → Value of a **Target Variable**

- ✓ **Stock Market Prediction**
- ✓ **Sales Forecasting**
- ✓ **Weather Forecasting**
- ✓ **Customer Churn Prediction**

Classification

Input Data → ML Box → **Labels** of **Categories**

- ✓ **Spam Detection**
- ✓ **Sentiment Analysis**
- ✓ **Image Classification**

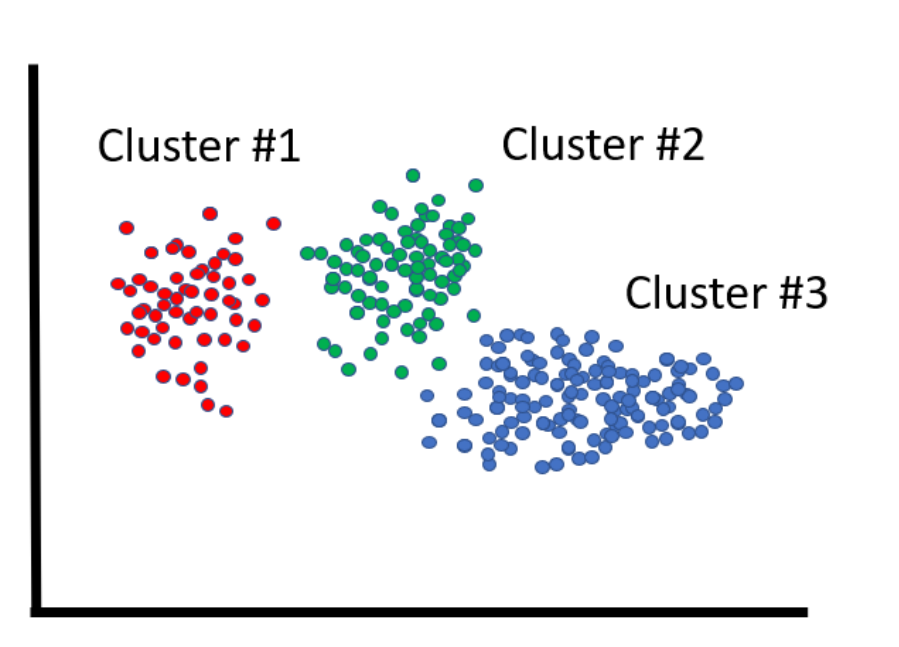


Clustering

Input Data → ML Box → **Clusters**

groups of similar data points

- ✓ **Personalized Product Recommendations**
- ✓ **Social Networks**
- ✓ **Knowledge Sharing**



Content Generation

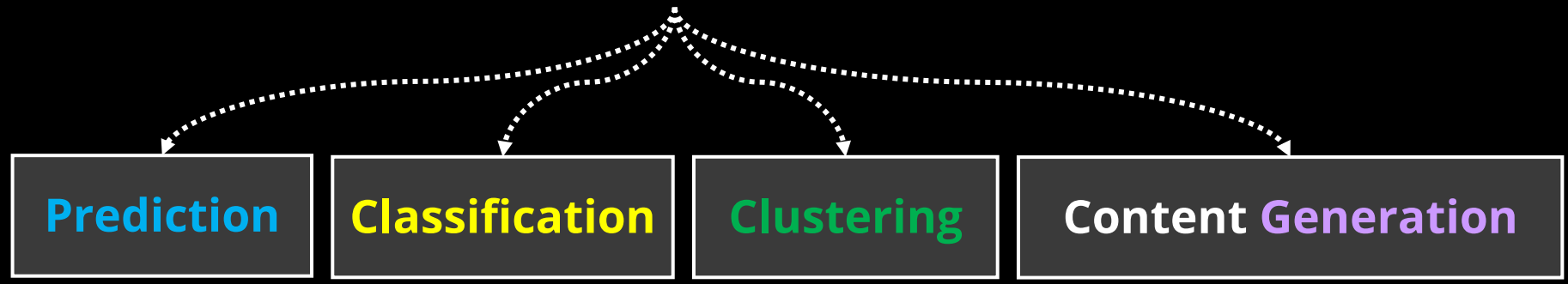
Text Description

ML Box

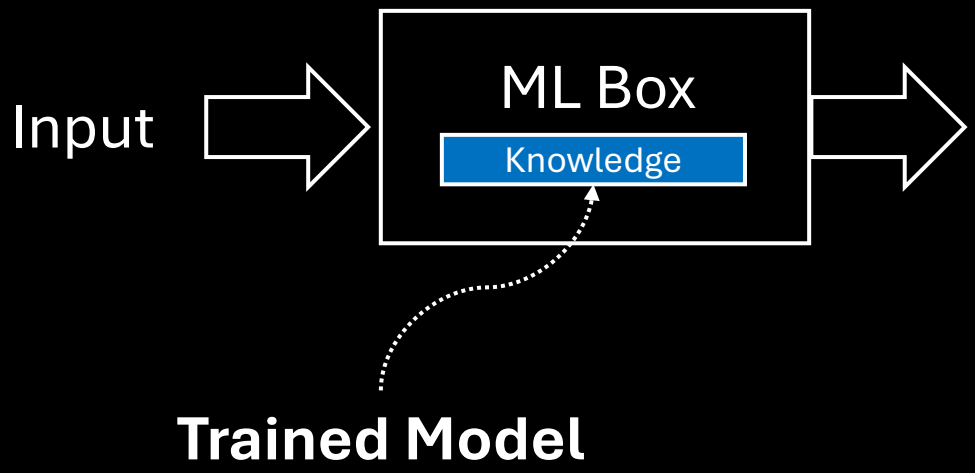
- Text-To-Text
- Text-To-Image
- Text-To-Video
- Text-To-Audio

The Box Analogy

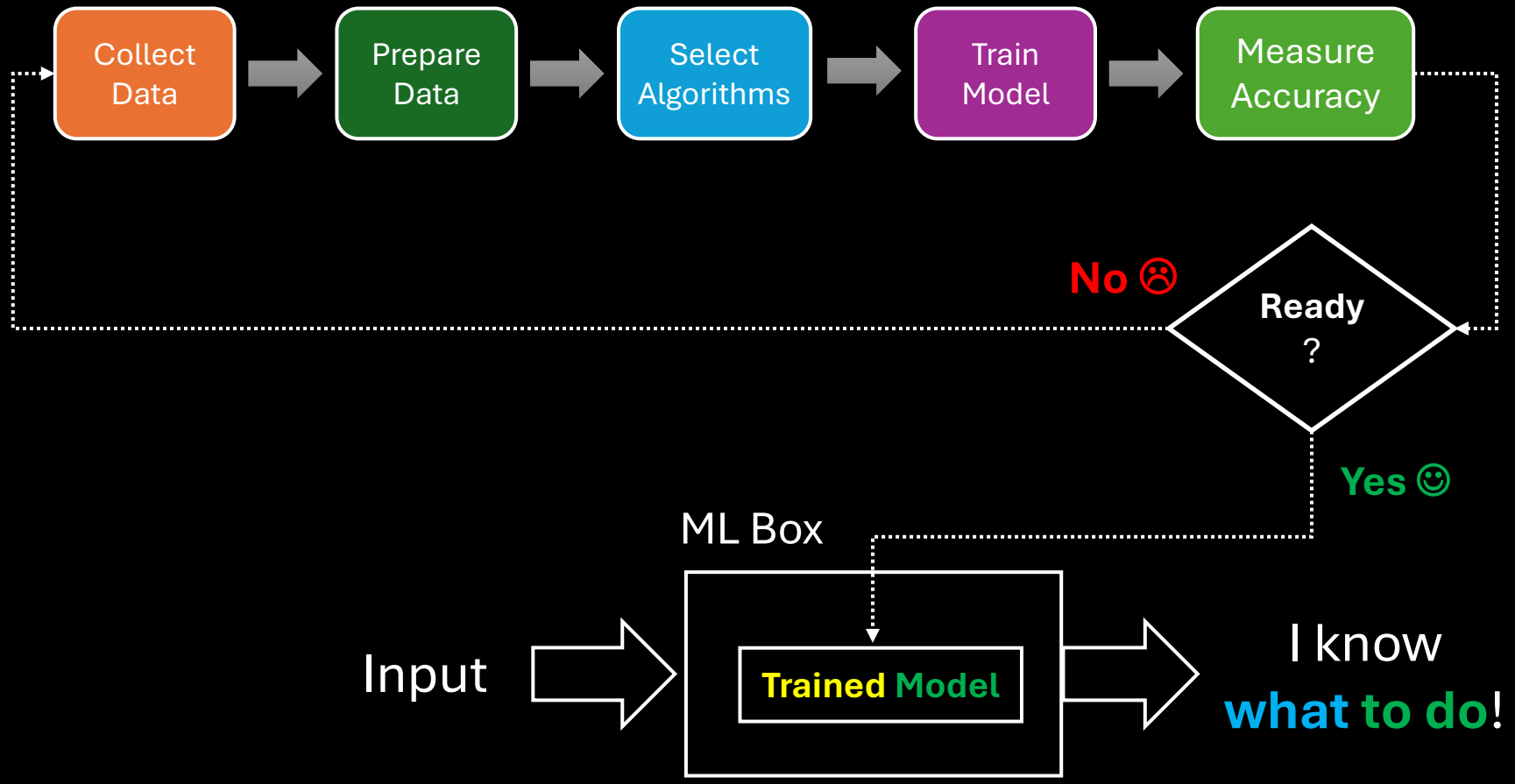
Typical ML Tasks



Training Phase

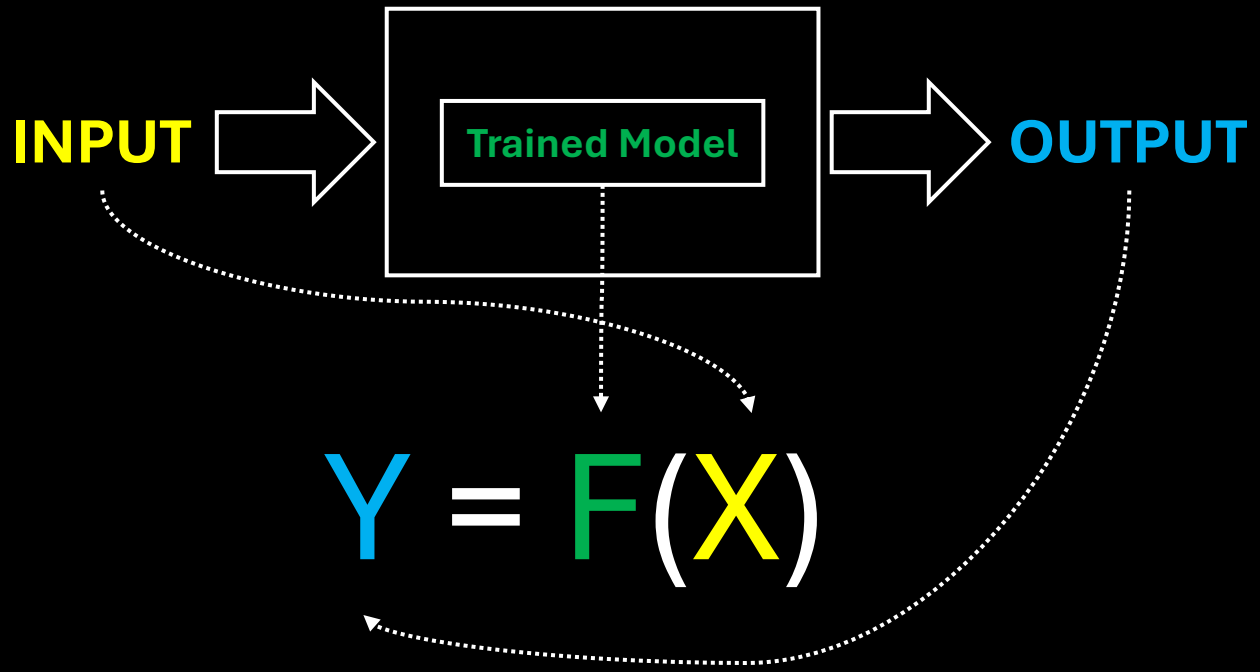


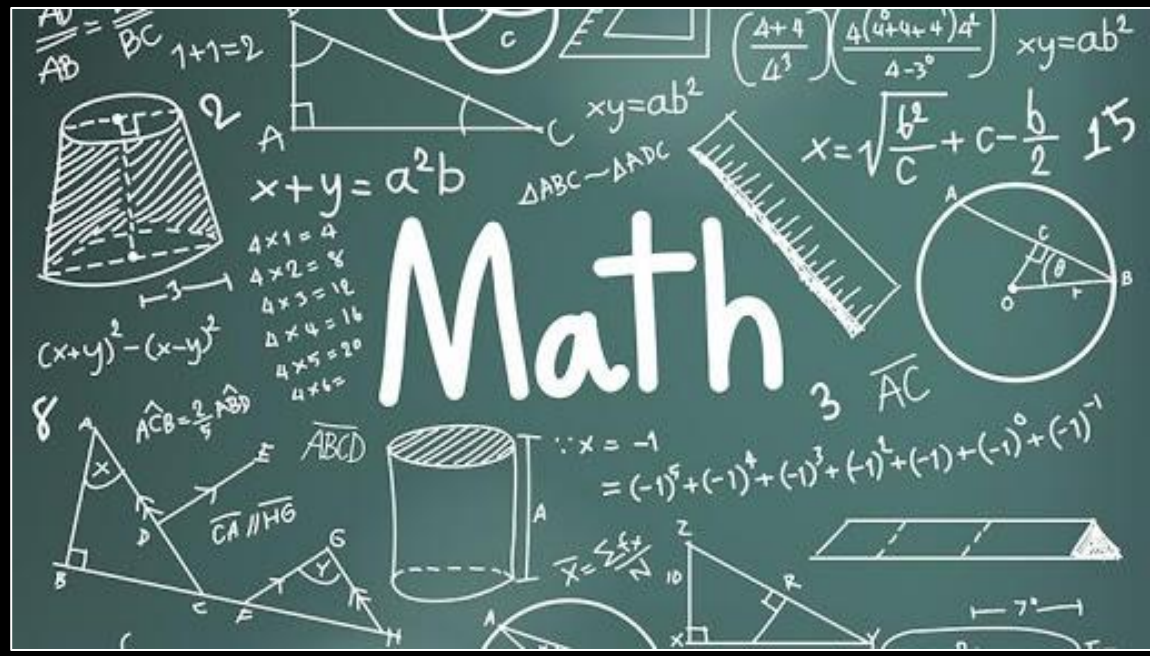
Creating a Trained Model



Y = F(X) – OMG!

The Box Analogy

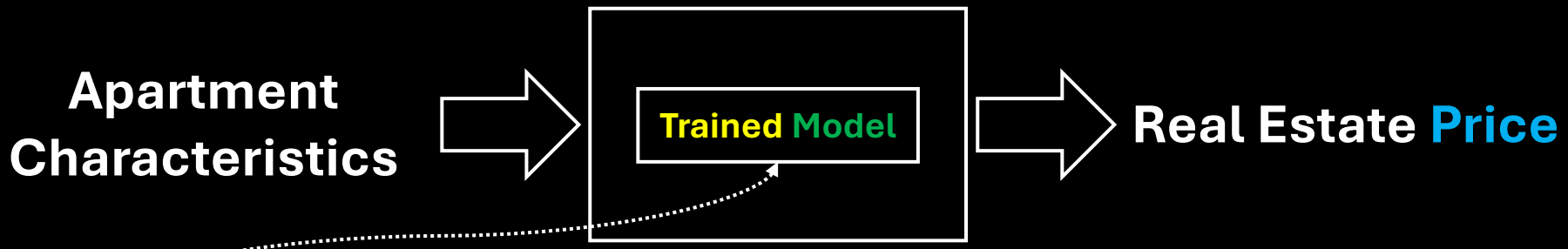




F - Mapping function

X → Y

Y = F(X) – OMG!

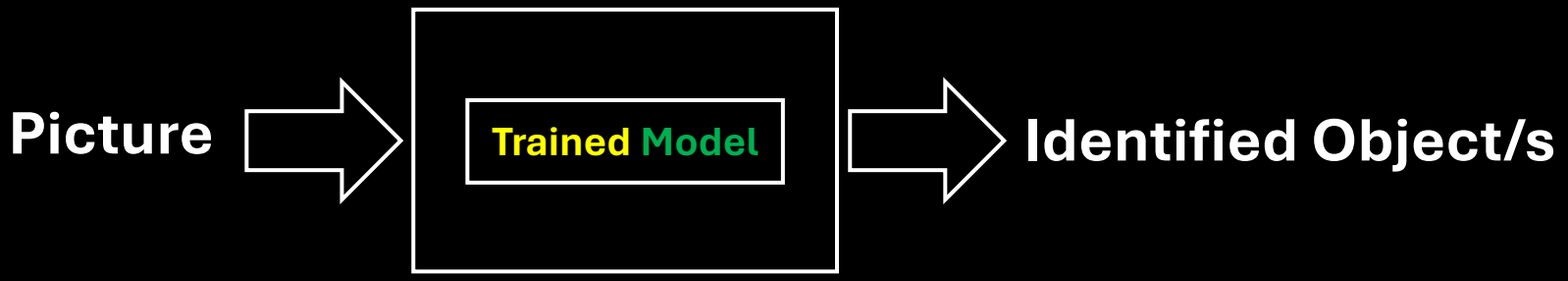


$$Y = F(X) = 0.1 + 3*X1 + 0.4*X2 + 1.6*X3 + 0.2*X4$$

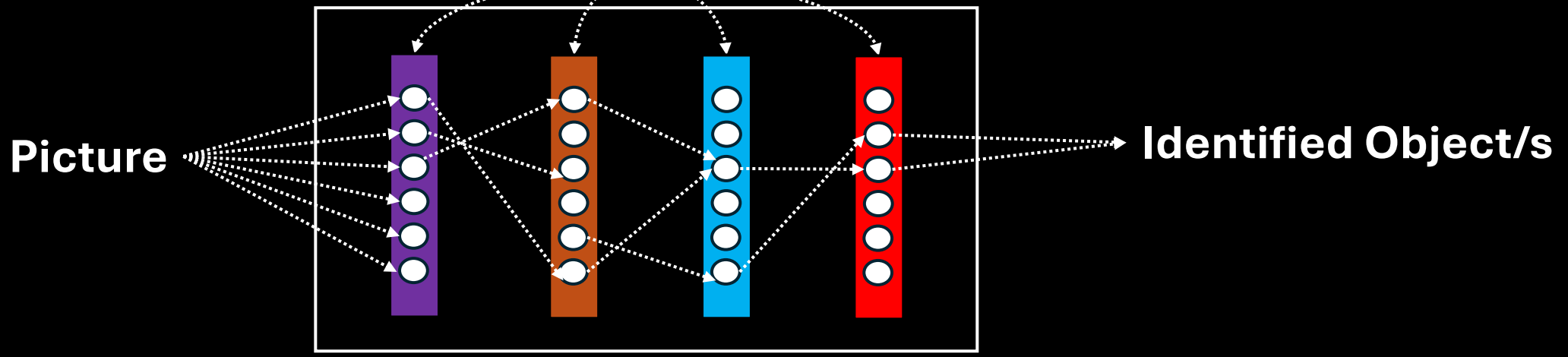
- X1 – number of rooms in the apartment
- X2 – the square size
- X3 – distance from nearby school
- X4 – distance from nearby hospital etc.

Model Weight/Parameter

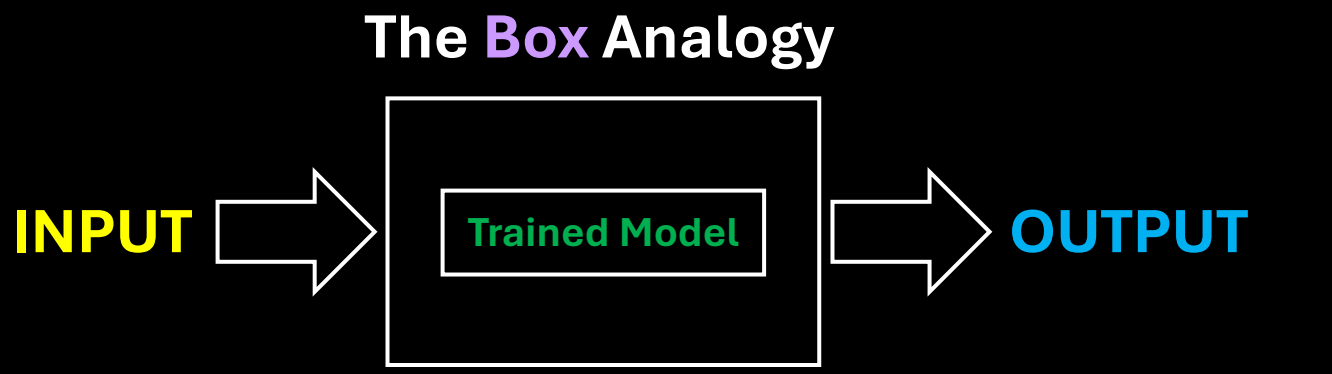
Y = F(X) – OMG!



$$Y = F(X) = F4\,(\,F3\,(\,F2\,(\,F1\,(X\,)\,)\,)\,)$$



4-layer neural network


The Box Analogy
INPUT
Trained Model
OUTPUT

#1 – Structured Data

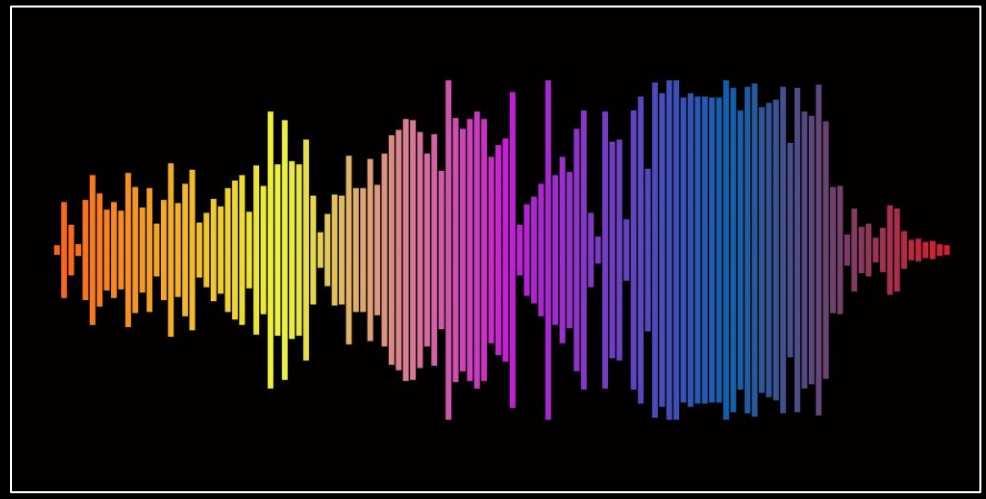
- ✓ Has a **defined format**
- ✓ highly organized and arranged in a predefined structure
- ✓ **Easily** searchable and accessible
- ✓ **Lower flexibility** for changing structure

List of Customers

Name	Age	Address	Phone	Email	ID
John	36	…	…	…	…
Michael	58				
Maya	29				

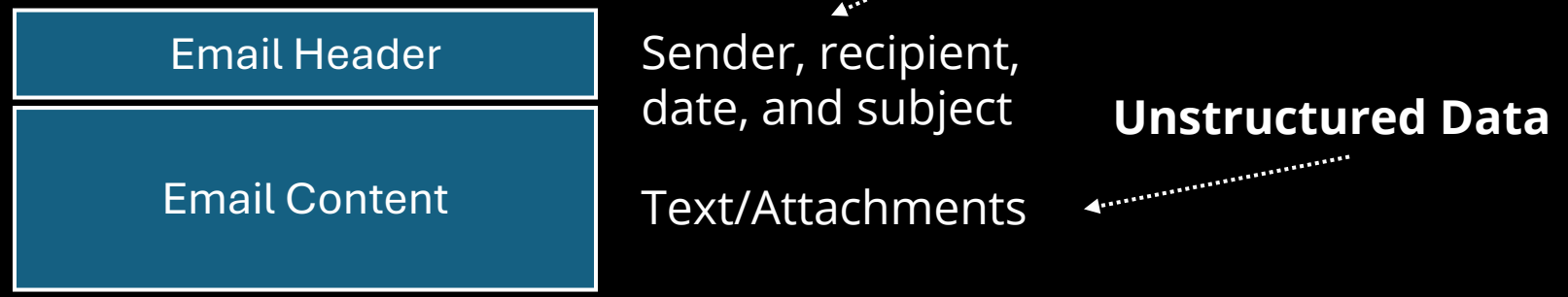
#2 – Unstructured Data

- ✓ **No pre-defined** data structure
- ✓ Examples
  - ✓ Textual data (emails, social media posts, documents…)
  - ✓ Multimedia (images, videos, audio files)
- ✓ More **complex** to process
  - ✓ Advanced machine learning methods

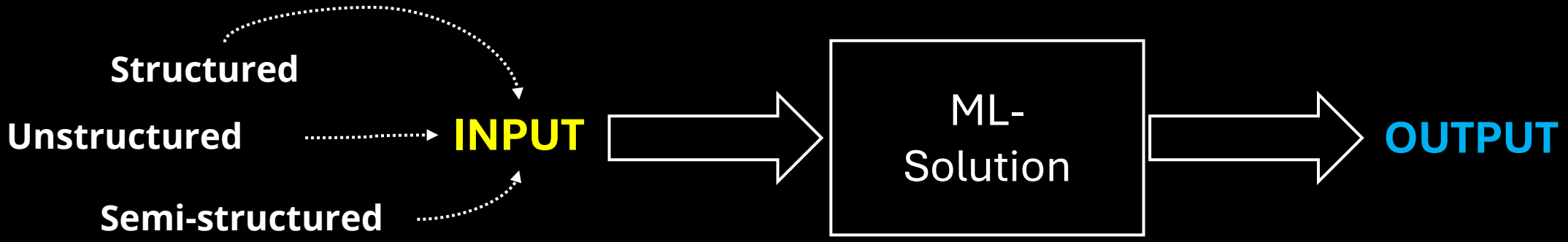


#3 – Semi-structured Data

- ✓ **Hybrid** type (between structured and unstructured)
- ✓ Has a **partial structure**

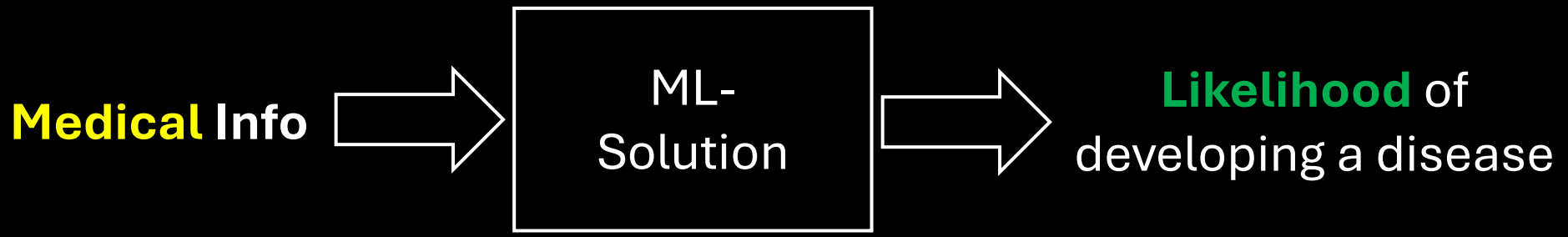


Log Message: 2024-08-25 12:34:56, HIGH, ERROR: Unable to connect to database



Features – Input Variables

- ✓ Numerical value (age)
- ✓ Categorical data (gender, color)
- ✓ Date and time
- ✓ Text
- ✓ Images
- ✓ Audio



Features – Input Variables

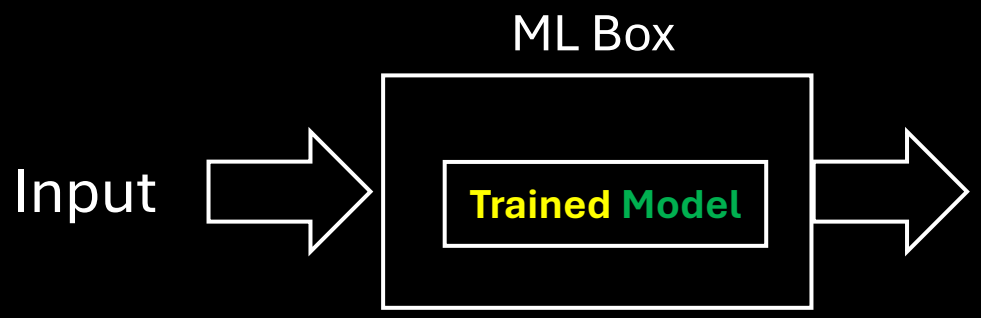
- ✓ Age
- ✓ Geo Location
- ✓ Gender
- ✓ Blood Pressure
- ✓ Smoking Status
- ✓ Health Condition

Feature Engineering

Age
(numerical number)

0-18
19-30
31-45
46-60
(categories)

Please Call the Supervisor!



- #1 – Supervised Learning
- #2 – Unsupervised Learning
- #3 – Semi-supervised Learning
- #4 – Reinforcement Learning

#1 – Supervised Learning

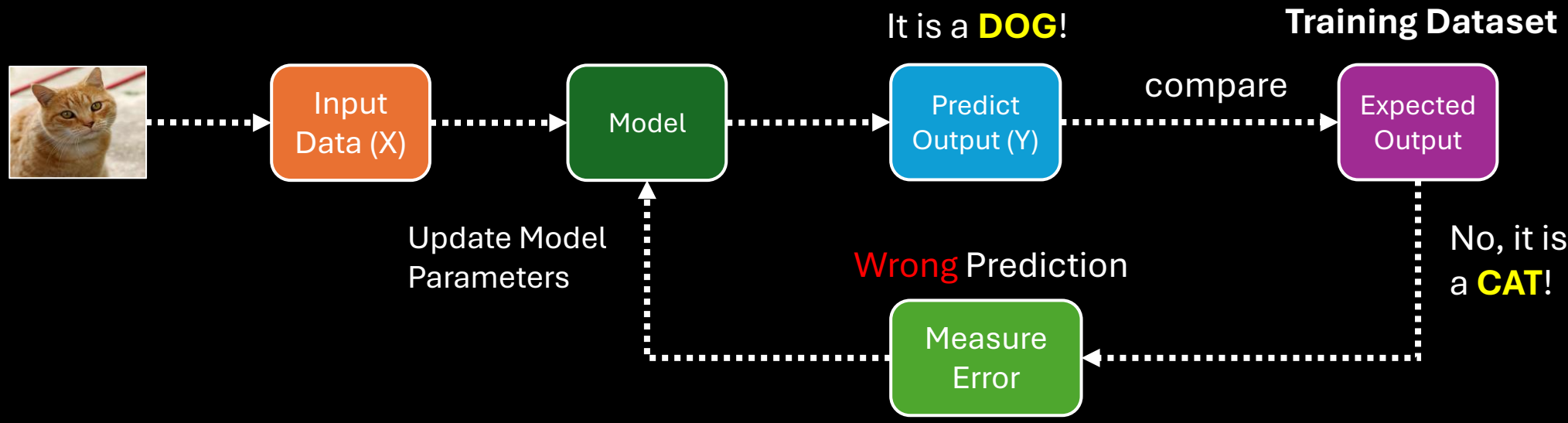
Training is guided by **labeled** data → **Training Dataset**

Labeled data = a collection of **many** data examples

Data example - Input + Output

#1 – Supervised Learning

Training Dataset



#2 – Unsupervised Learning

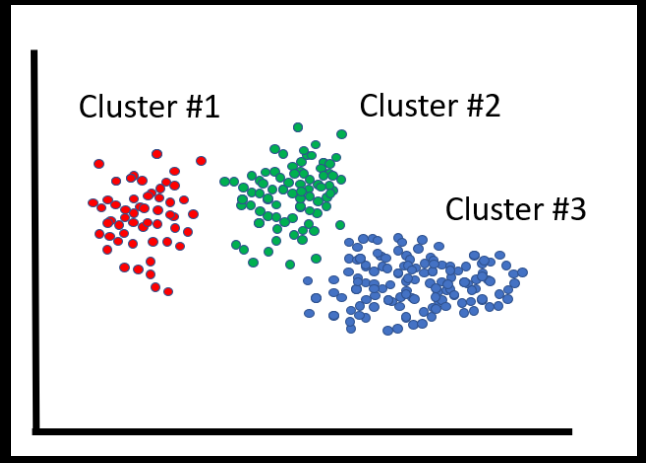
Training is guided by **unlabeled** data

Learns patterns directly from the data **without any guidance**

Main goal – **patterns discovery**

Discover **hidden patterns**, **structures**, or **relationships** within the data itself

Example - Clustering



#3 – Semi-supervised Learning

In some use cases, getting enough labeled data is a challenge

A **hybrid approach** – supervised + unsupervised

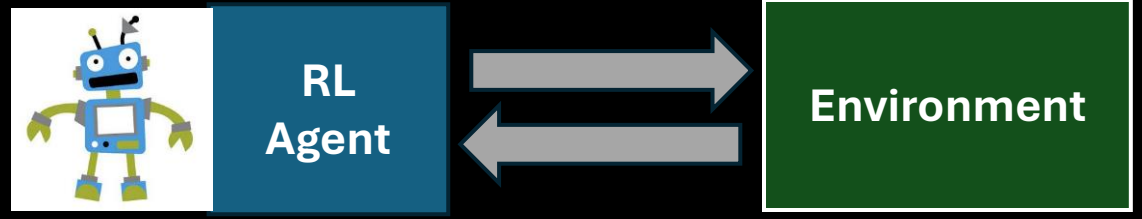
Training using:

- Small amount of **labeled** data
- Large amount of **unlabeled** data

#4 – Reinforcement Learning

Inspired by how humans and animals learn through **trial and error**

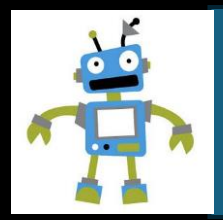
We **try things** and **get feedback**, and based on the feedback we are **learning**



Decision-making machine

Please Call the Supervisor!

#4 – Reinforcement Learning

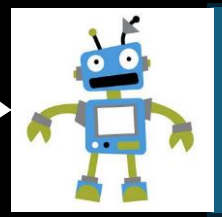


RL Agent ⇄ Environment

Winning a game → **Good Job!**

Losing a game → You should improve...

Game #1
Game #2
Game #3
Game #4



RL Agent ⇄ Policy/ Strategies

Trained Model