CSE 465
Lecture 1-2

Introduction

Course Name

- Course Name: Pattern Recognition & Neural Networks – (CSE465)
  - An undergraduate-level course – supposed to be more conceptual than theory
- 2 lectures every week (3 hours)

Workload and Grading Policy

- 2/3 class tests (theory + programming) worth 15%
  - Theory part: Why does it work?
  - Practical part: Usage of NN algos, analysis of results, metrics
  - Best 1/2 out of 2/3
- Project worth 20%-30%
  - Pick a project based on your works/idea – not from blogs/kaggle/etc.
  - Programming part: Implement/use NN algos, analysis of results
  - Must submit a power point presentation & a paper format report
  - Will also involve a viva
- Mid & Final exams worth 50%-60%
  - Mid & Final will be worth 25% each (total 50%)
  - They will have 3/4 questions (no choice – must answer all)
  - Questions will be part theory part practical concepts
- Attendance – 5%

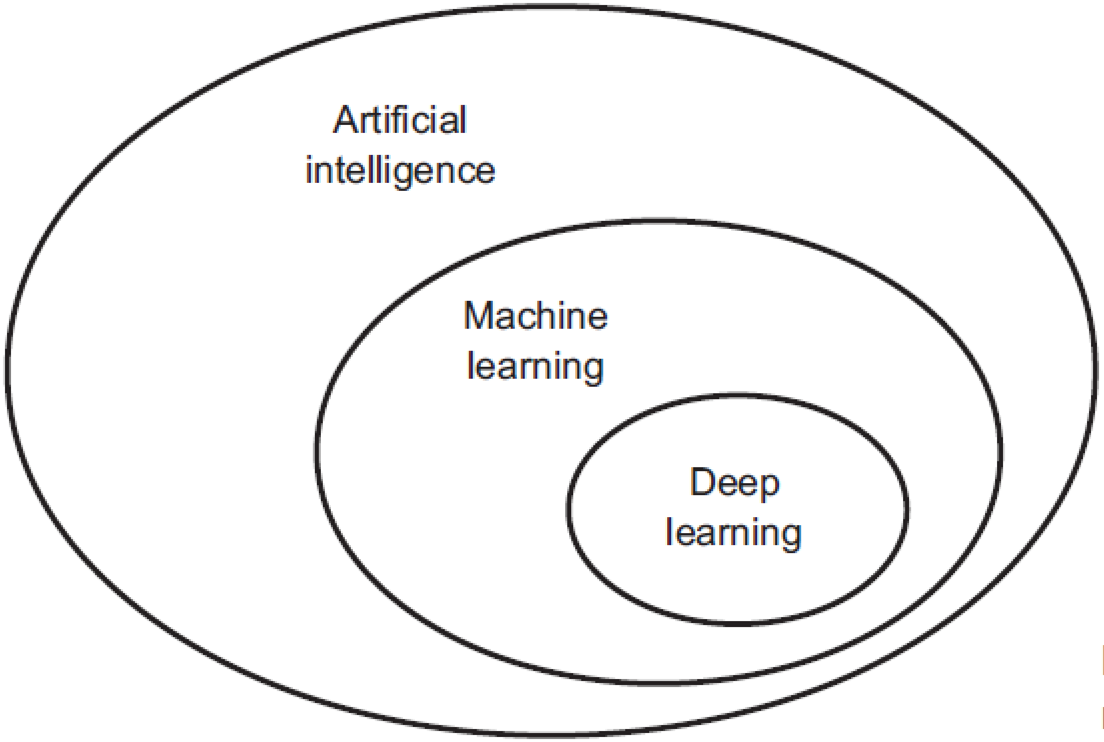
Reference materials

- Course slides and notebooks will be sourced from previous iterations of this course and relevant slides/materials from me/other researchers/blogs
- No fixed textbook
  - But you can read these free books
    - https://udlbook.github.io/udlbook/
    - Dive Into Deep Learning (https://d2l.ai/)
- Different books vary in terms of
  - Set of topics covered
  - Flavor (e.g., classical statistics, deep learning, probabilistic/Bayesian, theory)
  - Terminology and notation (beware of this especially)

Course Goals..

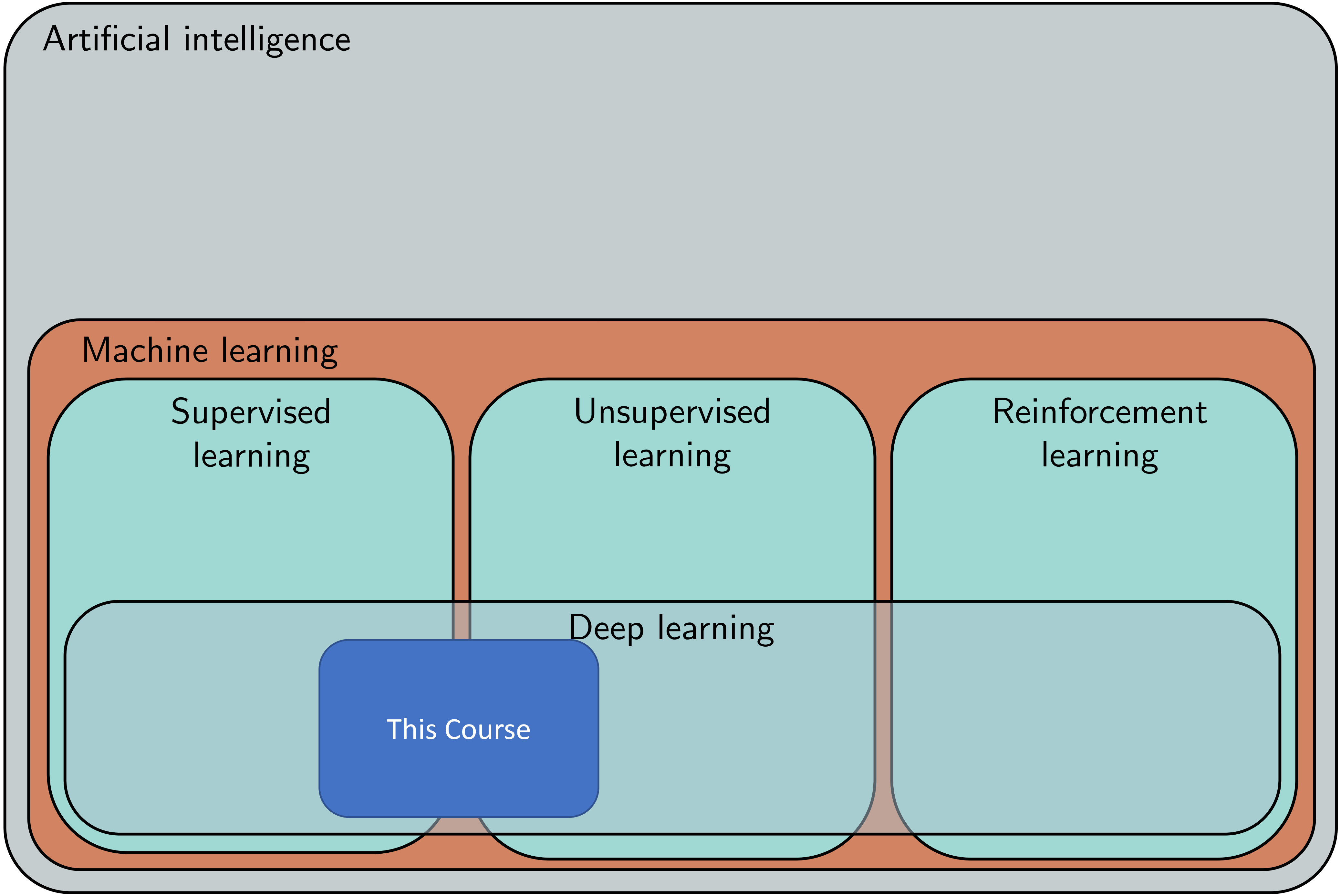
- Introduction to the foundations of Deep Learning
- Use unsupervised techniques to solve problems
- Focus on developing the ability to
  - Work on ML/DL projects at jobs/research
  - Be able to pick the right neural network for the job at hand
  - Being able to find computational complexities of different DL methods
  - Evaluate alternate solutions
  - Being able to understand the training complexities

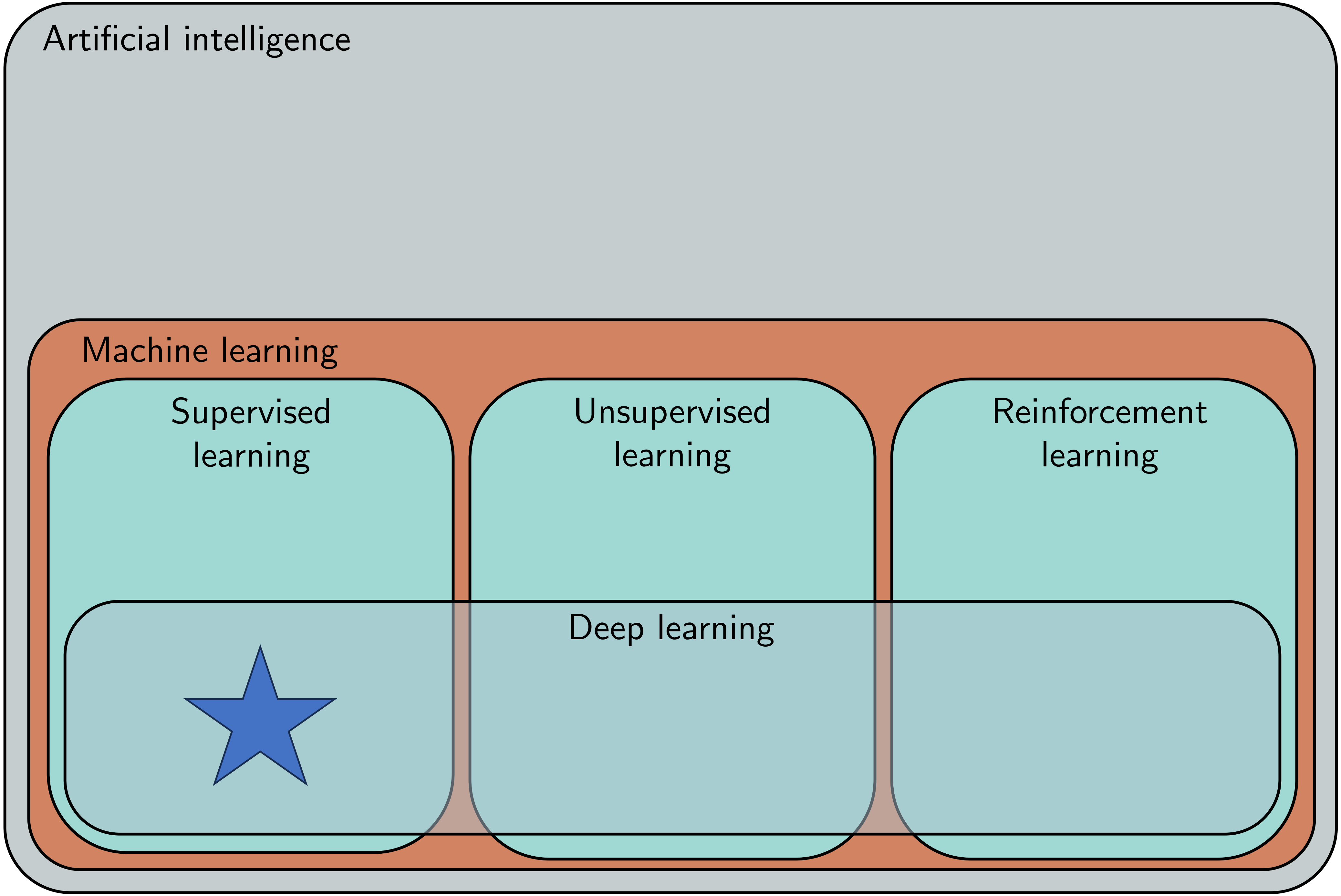
AI vs Machine Learning vs Deep Learning



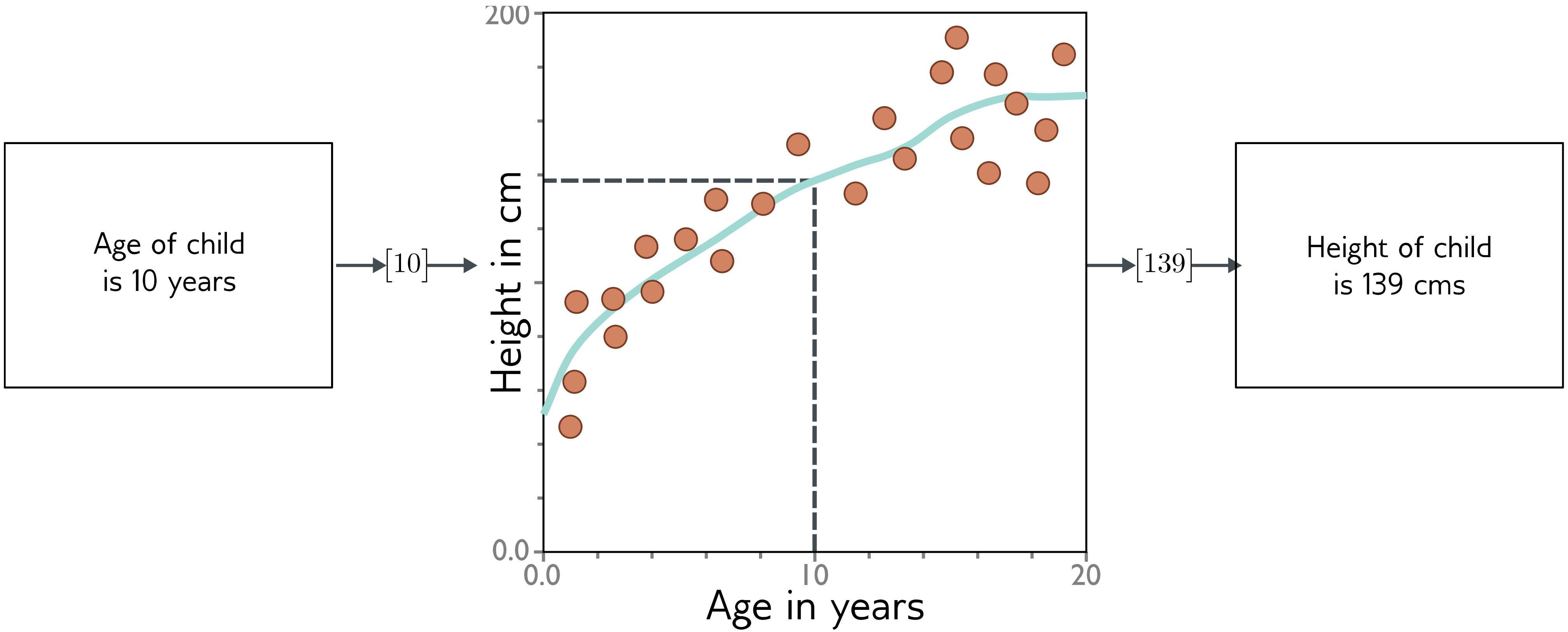
AI

- As the name suggests, artificial intelligence can be loosely interpreted as incorporating human intelligence to machines
- Artificial intelligence is the broader concept that consists of everything from good old-fashioned AI all the way to futuristic technologies such as deep learning
- Whenever a machine completes tasks based on a set of stipulated rules that solve problems – if that could be termed as “intelligent” behavior, then it is artificial intelligence


Artificial intelligence
Machine learning
Supervised learning
Unsupervised learning
Reinforcement learning
Deep learning
This Course

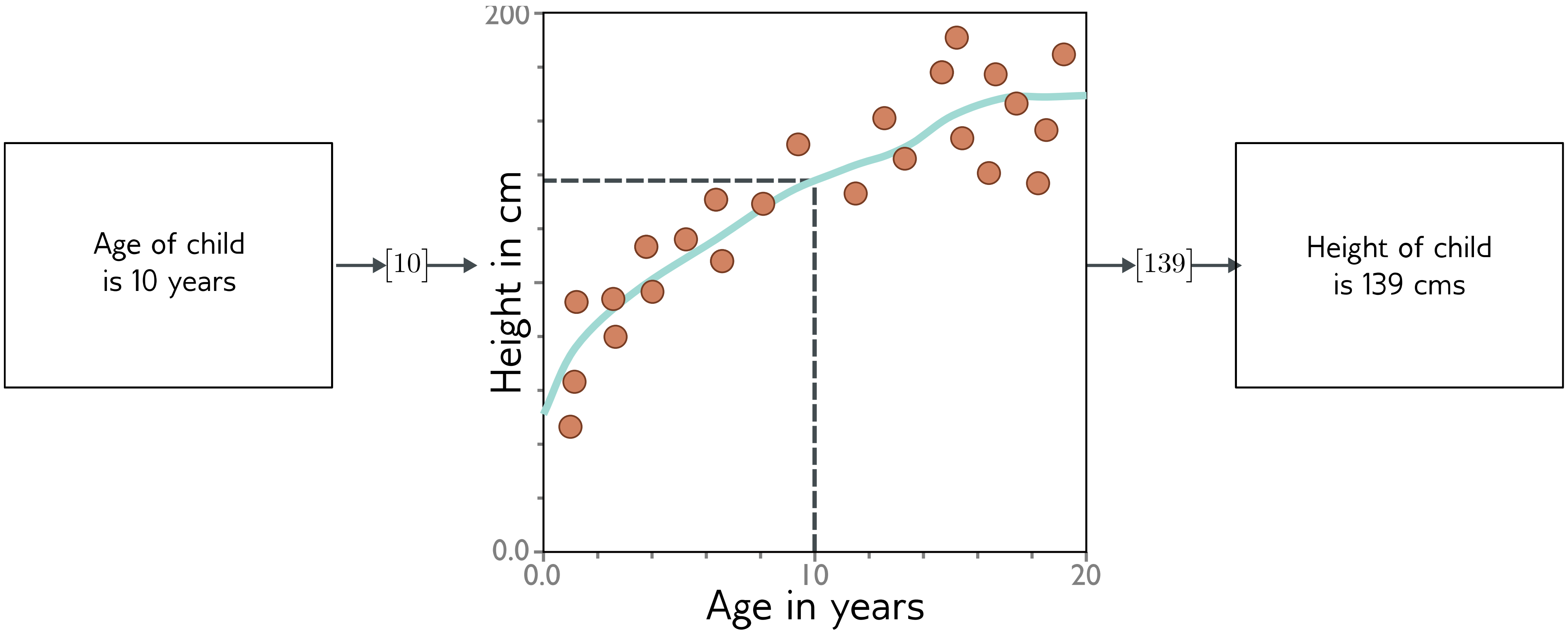

Artificial intelligence
Machine learning
Supervised learning
Unsupervised learning
Reinforcement learning
Deep learning

What is a supervised learning model?



- An equation relating input (age) to output (height)
- Search through family of possible equations to find one that fits training data well

What is a supervised learning model?



- Deep neural networks are just a very flexible family of equations
- Fitting deep neural networks = “Deep Learning”

Supervised learning

- Define a mapping from input to output
- Learn this mapping from paired input/output data examples
- In supervised learning, we aim to build a model that takes an *input x* and outputs a *prediction y*
- For simplicity, we assume that both the input x and output y are **vectors** of a predetermined and fixed size and that the elements of each vector are always ordered in the same way; this is termed structured or tabular data

Supervised Learning - Model

- To make the prediction, we need a model f[•] that takes input x and returns y, so: $\mathbf{y} = \mathrm{f}[\mathbf{x}]$.
  - When we compute the prediction y from the input x, we call this inference
- The model is just a mathematical equation/object with a fixed form
  - It represents a family of different relations between the input and the output
  - A neural network represents an equation with input (data) and outputs (labels)
  - The model also contains parameters ϕ.
- The choice of parameters determines the relation between input and output, so we should really write:

$$\mathbf{y} = \mathrm{f}[\mathbf{x}, \phi].$$

Supervised Learning - Training

- When we talk about learning or training a model, we mean that we attempt to find parameters ϕ that make sensible output predictions from the input
- We learn these parameters using a training dataset of *N pairs* of input and output examples $\{x_i , y_i\}$
- We aim to select parameters that map each training input to its associated output as closely as possible
  - We quantify the degree of mismatch in this mapping with the *loss L*
    - This is a scalar value that summarizes how poorly the model predicts the training outputs from their corresponding inputs for parameters ϕ

Supervised Learning - Loss

- We can treat the loss as a *function $L[\phi]$* of these ϕ parameters
- When we train the model, we are seeking parameters $\hat{\phi}$ that minimize this loss function:

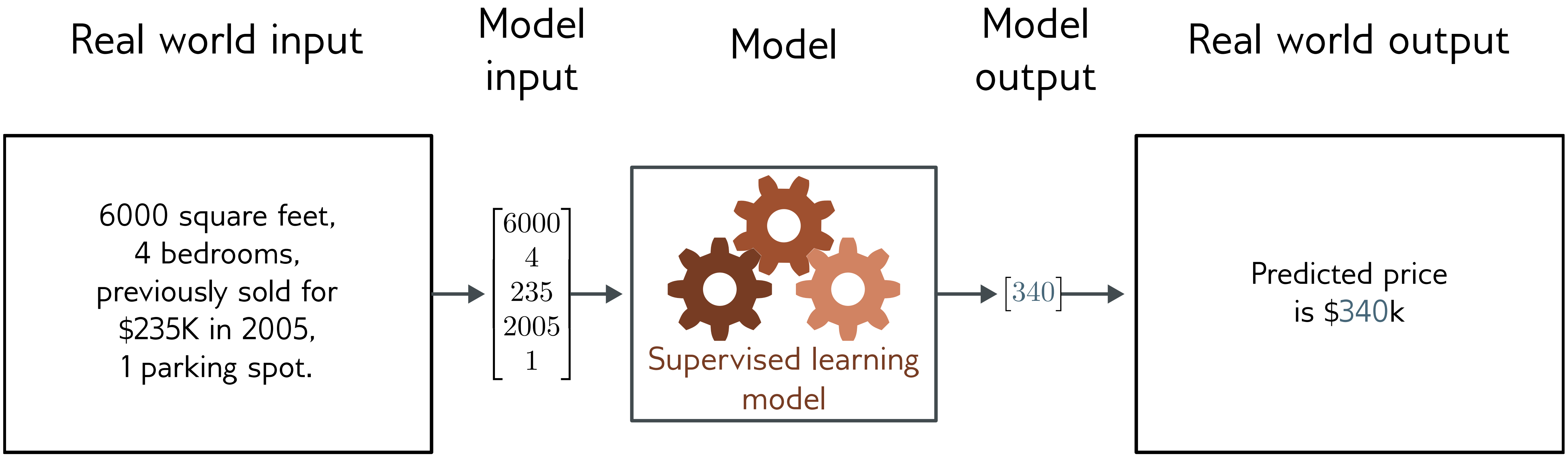
$$\hat{\phi} = \underset{\phi}{\operatorname{argmin}} \Big[L[\phi] \Big]$$

- If the loss is small after this minimization, we have found model parameters that accurately predict the training outputs y_i from the training inputs x_i
- After training a model, we must now assess its performance; we run the model on separate test data to see how well it generalizes to examples that it didn't observe during training
- If the performance is adequate, then we are ready to deploy the model

Terms

- Regression = continuous numbers as output
- Classification = discrete classes as output
- Two class and multiclass classification treated differently
- Univariate = one output
- Multivariate = more than one output

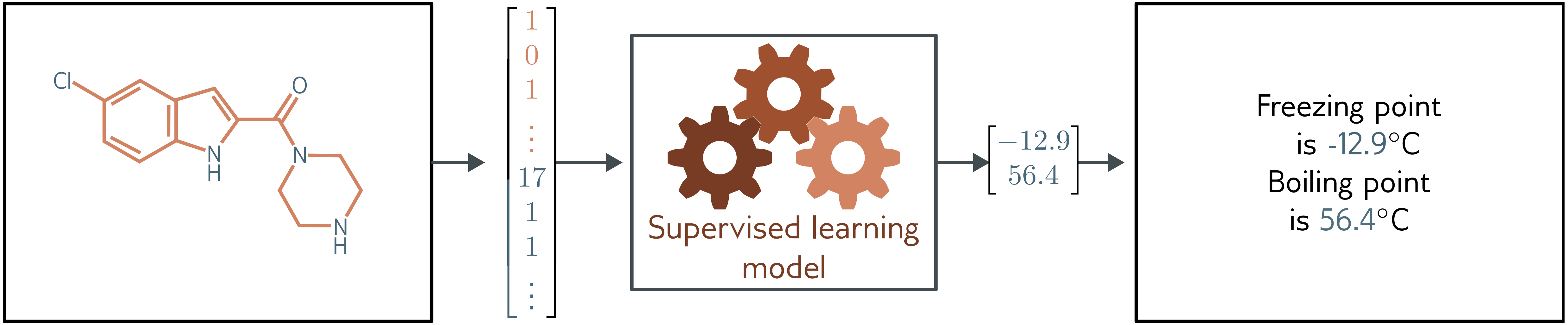
Regression



- Univariate regression problem (one output, real value)
- Fully connected network

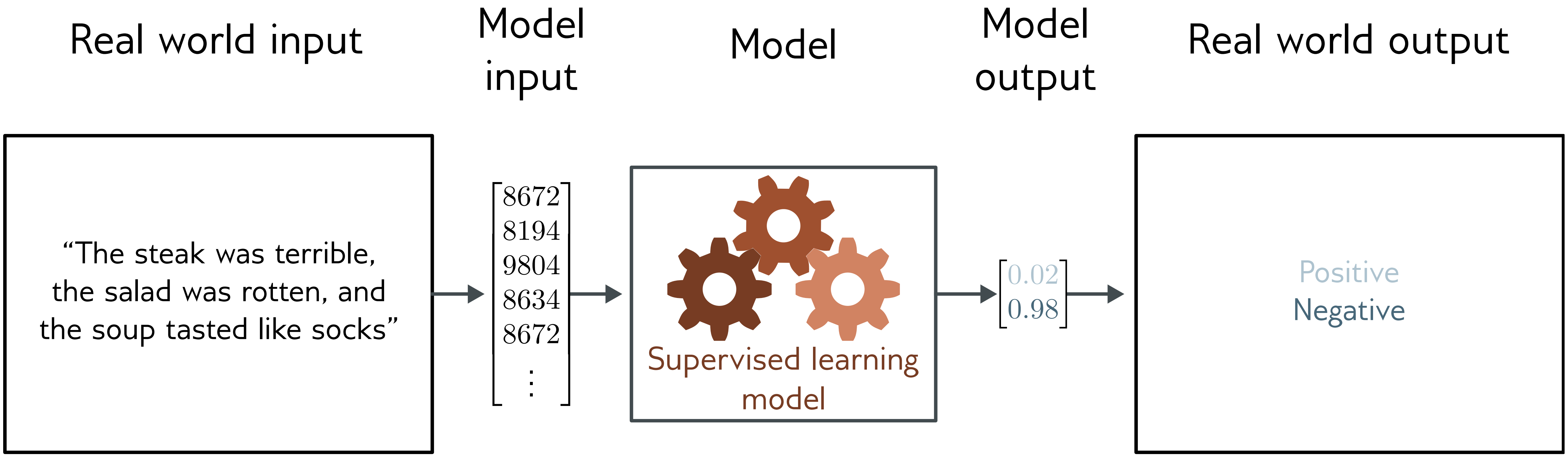
Graph regression

Real world input | Model input | Model | Model output | Real world output



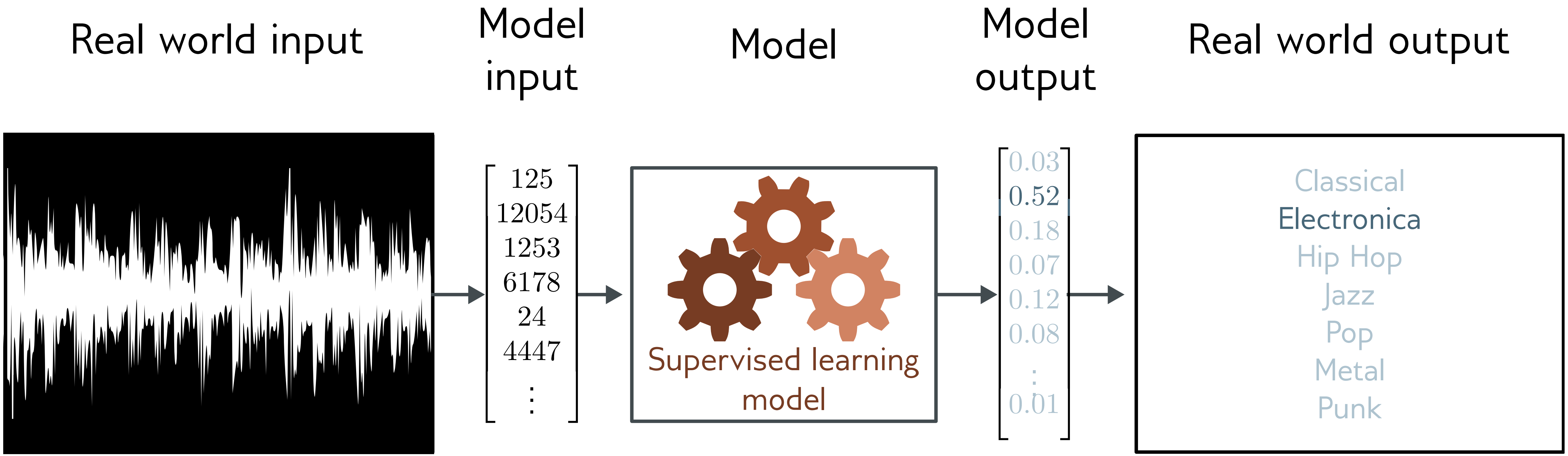
- Multivariate regression problem (>1 output, real value)
- Graph neural network

Text classification



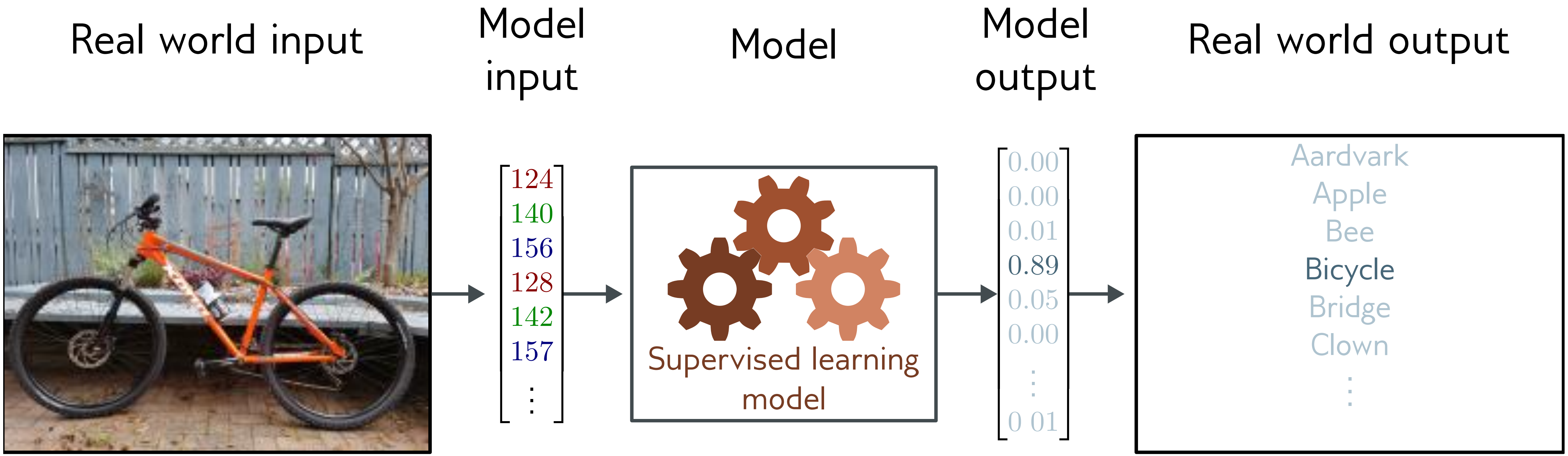
- Binary classification problem (two discrete classes)
- Transformer network

Music genre classification



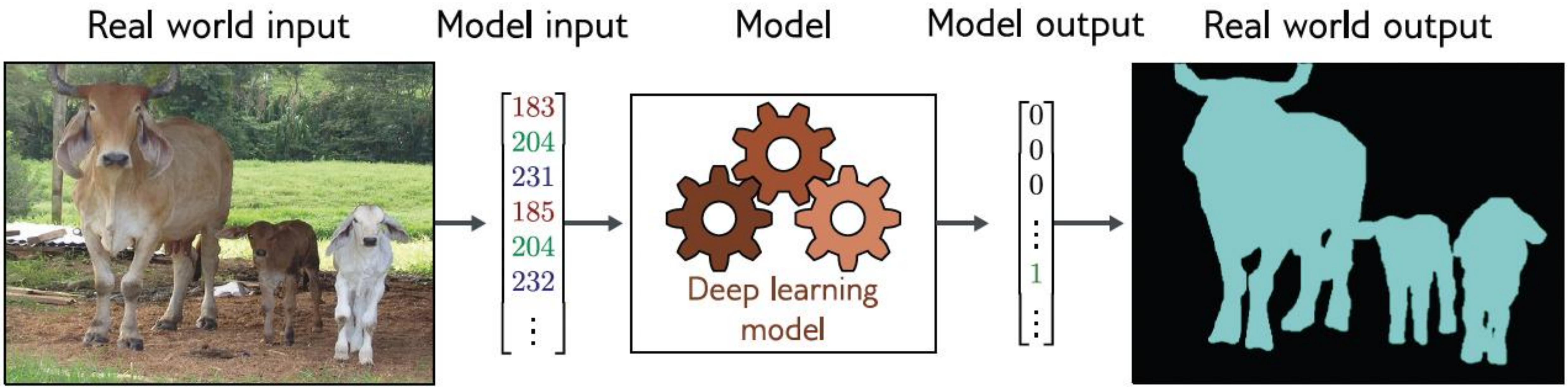
- Multiclass classification problem (discrete classes, >2 possible values)
- Recurrent neural network (RNN)

Image classification



- Multiclass classification problem (discrete classes, >2 possible classes)
- Convolutional network

Image segmentation

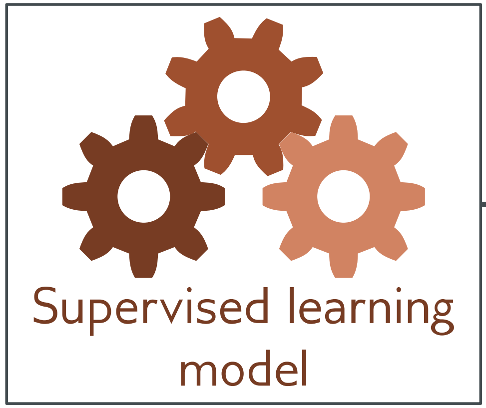


- Multivariate binary classification problem (many outputs, two discrete classes)
- Convolutional encoder-decoder network

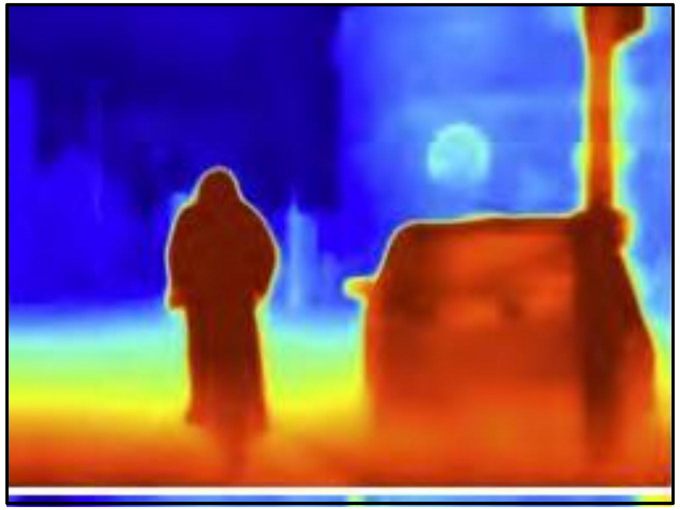
Depth estimation

Real world input	Model input	Model	Model output	Real world output

$$\begin{bmatrix} 255 \\ 254 \\ 255 \\ 254 \\ 254 \\ 255 \\ \vdots \end{bmatrix} \rightarrow$$



$$\rightarrow \begin{bmatrix} 0.001 \\ 0.002 \\ \vdots \\ 0.314 \\ 0.310 \\ \vdots \end{bmatrix} \rightarrow$$



- Multivariate regression problem (many outputs, continuous)
- Convolutional encoder-decoder network

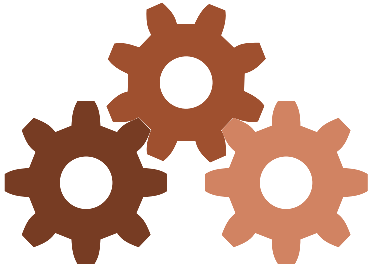
Pose estimation

Real world input

Model input

$$\begin{bmatrix} 3 \\ 5 \\ 4 \\ 3 \\ 5 \\ 5 \\ \vdots \end{bmatrix}$$

Model



Supervised learning model

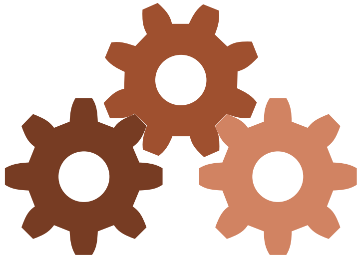
Model output

$$\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 3 \\ \vdots \end{bmatrix}$$

Real world output

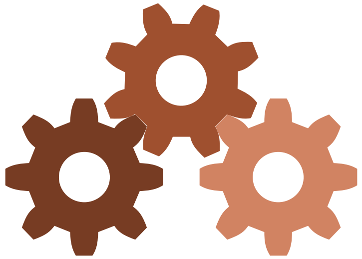
- Multivariate regression problem (many outputs, continuous)
- Convolutional encoder-decoder network

Translation

Real world input	Model input	Model	Model output	Real world output
"Skill without imagination is craftsmanship and gives us many useful objects such as wickerwork picnic baskets. Imagination without skill gives us modern art."	$\begin{bmatrix} 7800 \\ 9853 \\ 4520 \\ 4596 \\ 987 \\ \vdots \end{bmatrix}$	 Supervised learning model	$\begin{bmatrix} 6003 \\ 3689 \\ 4432 \\ 6003 \\ 2149 \\ \vdots \end{bmatrix}$	"L'habileté sans l'imagination est de l'artisanat et nous donne de nombreux objets utiles tels que des paniers de pique-nique en osier. L'imagination sans habileté nous donne l'art moderne."

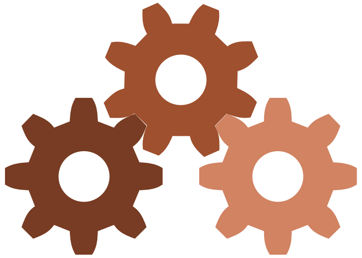
- Machine translation a sequence generation problem
- RNN/Transformer network

Image captioning

Real world input	Model input	Model	Model output	Real world output
	$\begin{bmatrix} 183 \\ 204 \\ 231 \\ 185 \\ 204 \\ 232 \\ \vdots \end{bmatrix}$	 Supervised learning model	$\begin{bmatrix} 1 \\ 5593 \\ 7532 \\ 7924 \\ 1 \\ \vdots \end{bmatrix}$	"A Kazakh man on a horse holding a bird of prey"

- Convolutional Neural Network for the embedding
- RNN/Transformer/CLIP network

Image generation from text

Real world input	Model input	Model	Model output	Real world output
"Teddy bears mixing sparkling chemicals as mad scientists, in a steampunk style."	$\begin{bmatrix} 8300 \\ 532 \\ 7676 \\ 7898 \\ 883 \\ \vdots \end{bmatrix}$	 Supervised learning model	$\begin{bmatrix} 2 \\ 2 \\ 2 \\ 2 \\ 2 \\ 1 \\ \vdots \end{bmatrix}$	

- Stable diffusion/GAN
- Generative models

What do these examples have in common?

- Very complex relationship between input and output of last three examples?
- Sometimes may be many possible valid answers
- But outputs (and sometimes inputs) obey rules

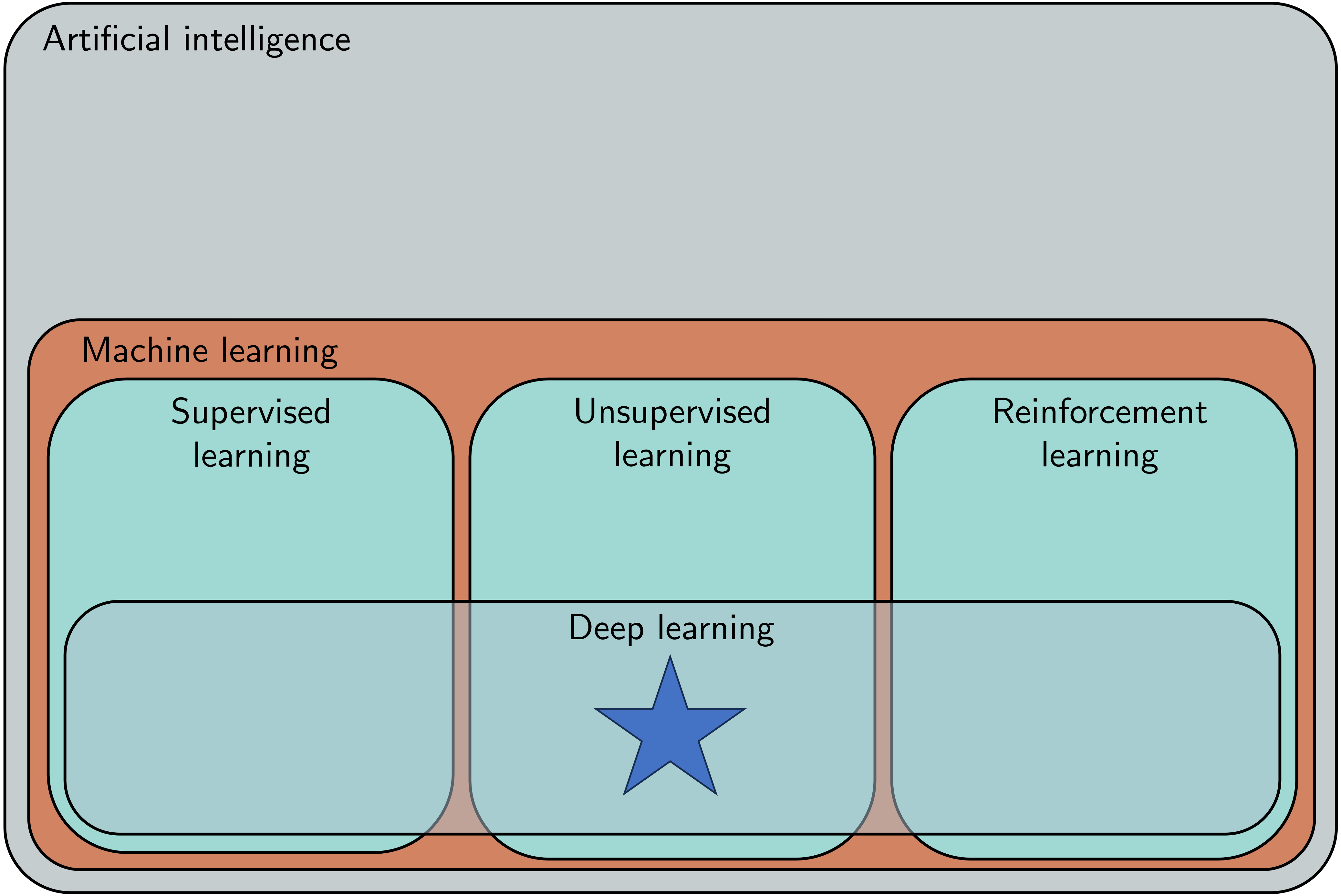
"A Kazakh man on a horse holding a bird of prey"

Language obeys grammatical rules

Natural images also have "rules"

How to incorporate external rules?

- Learn the “grammar” of the data from unlabeled examples
- Can use a gargantuan amount of data to do this (as unlabeled)
- Make the supervised learning task earlier by having a lot of knowledge of possible outputs

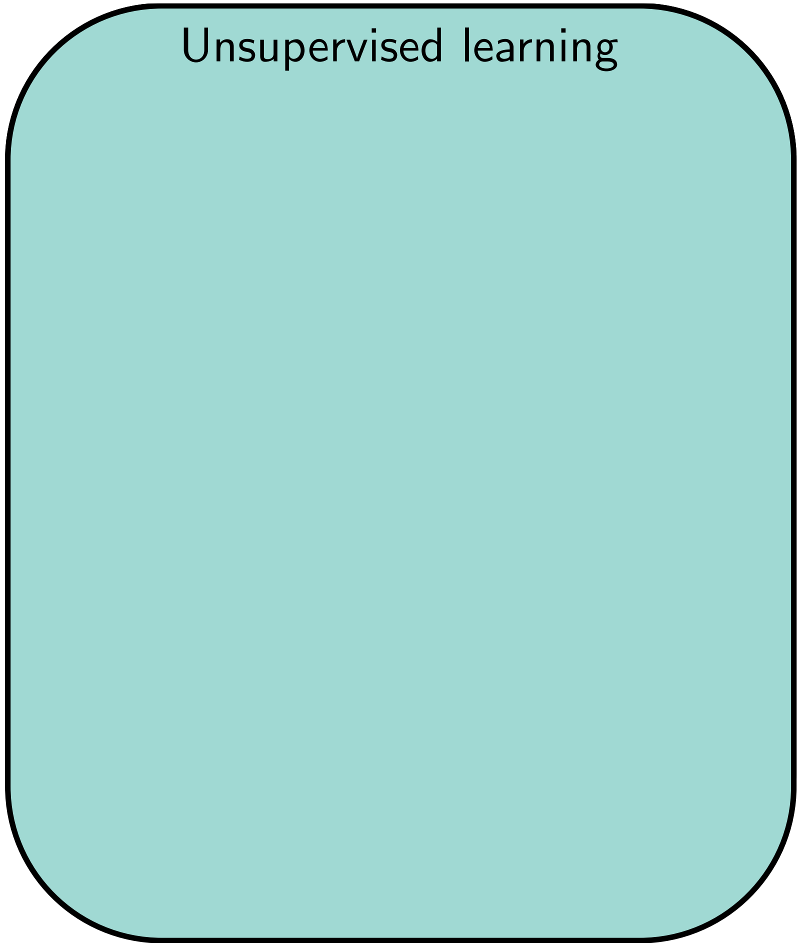


Artificial intelligence
Machine learning
Supervised learning
Unsupervised learning
Reinforcement learning
Deep learning

- Constructing a model from input data without corresponding output labels is termed unsupervised learning
  - Absence of output labels means there can be no “supervision”
    - Learns a mapping from input to output
    - Goal is to describe or understand the structure of the data
- There is no restrictions on the input data
  - It may be discrete or continuous, low-dimensional or high-dimensional, and of constant or variable length

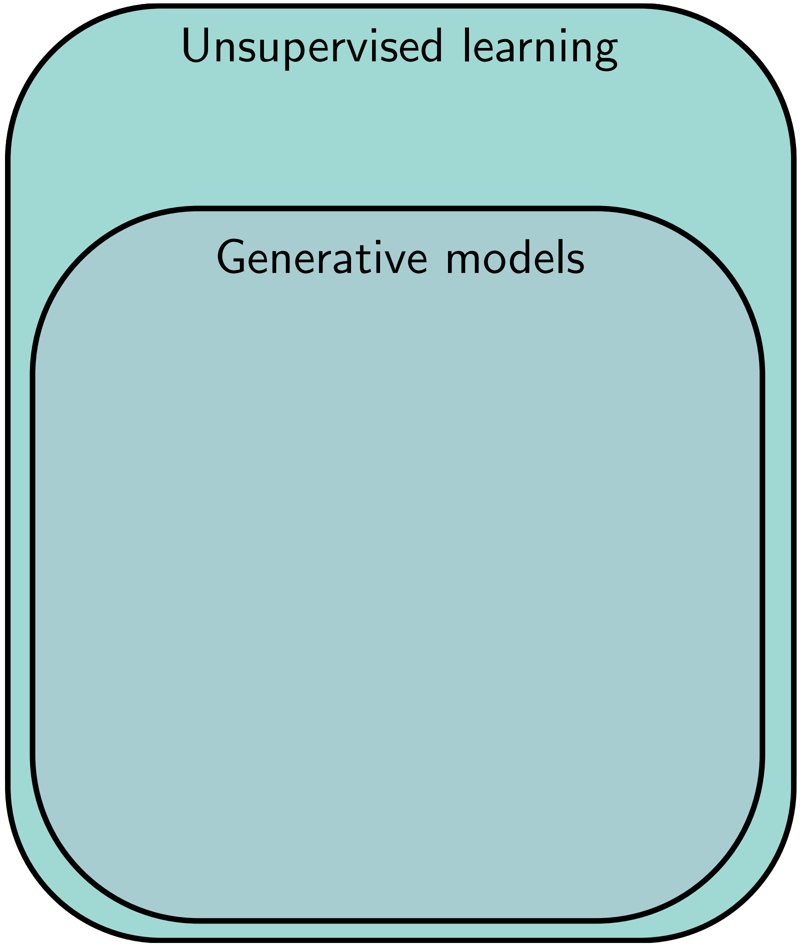
Unsupervised Learning

- Learning about a dataset without labels
  - Clustering
  - Finding outliers
  - Generating new examples
  - Filling in missing data



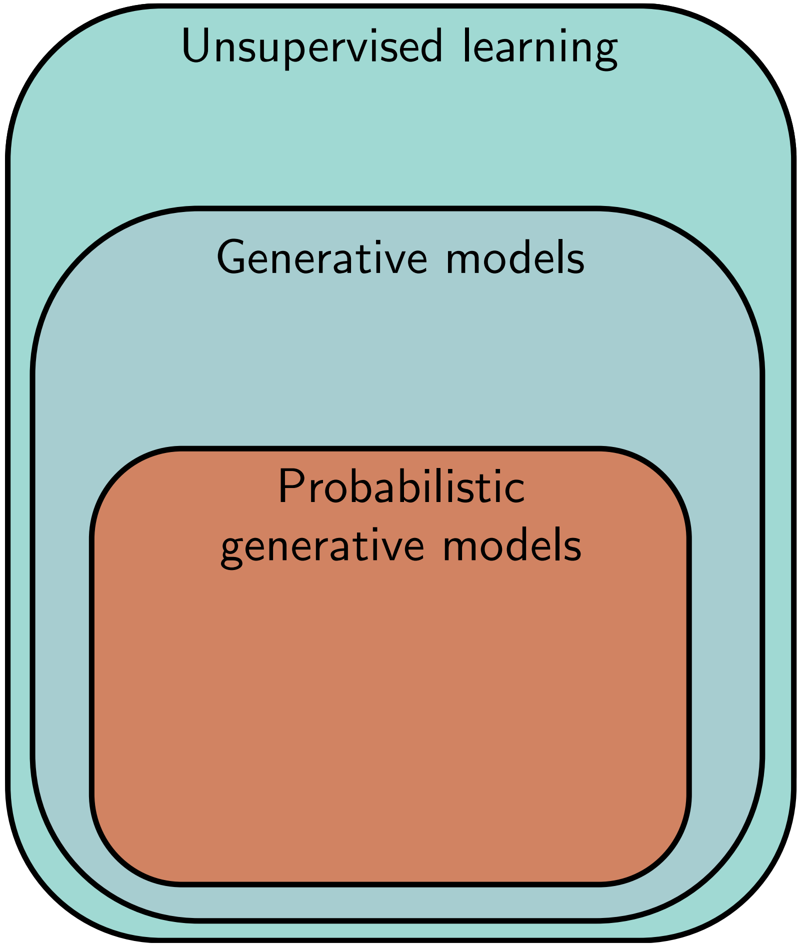
Unsupervised Learning

- Learning about a dataset without labels
  - e.g., clustering
- Generative models can create examples
  - e.g., generative adversarial networks

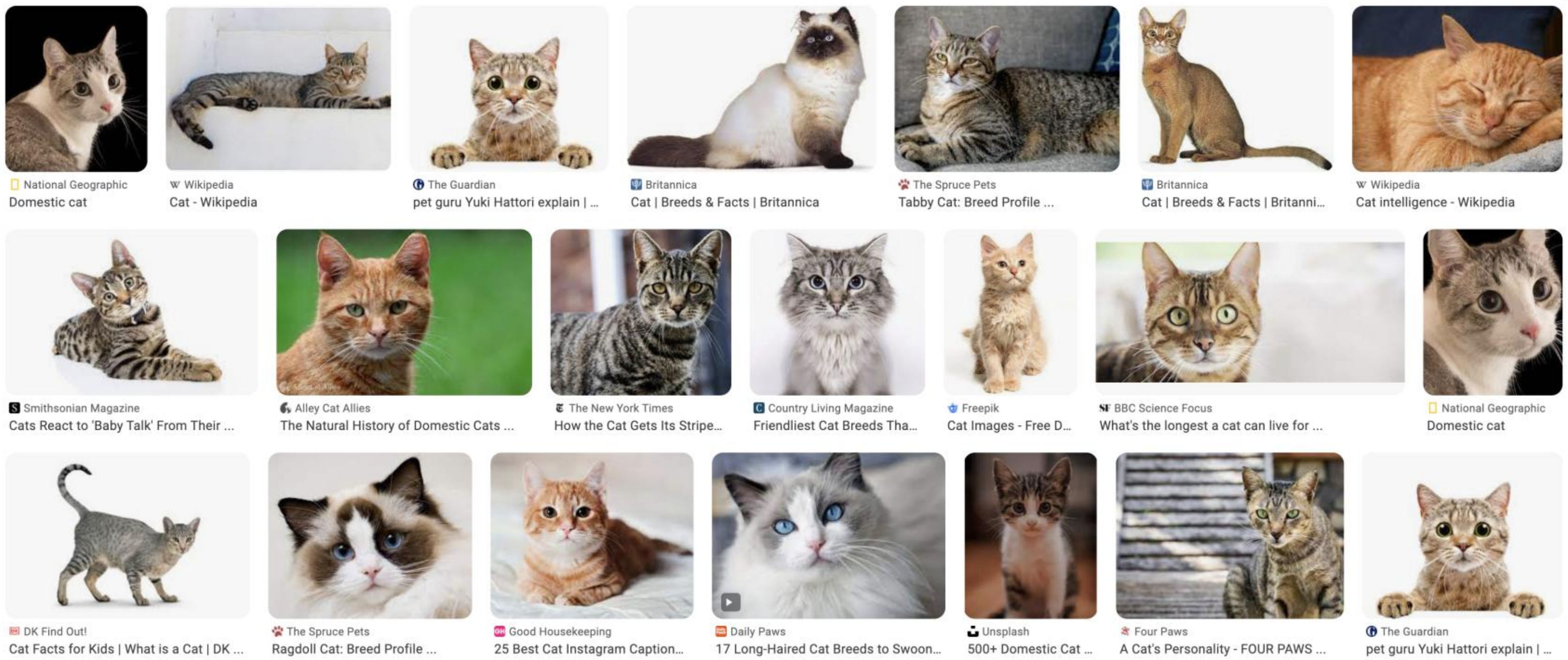


Unsupervised Learning

- Learning about a dataset without labels
  - e.g., clustering
- Generative models can create examples
  - e.g., generative adversarial networks
- PGMs learn distribution over data
  - e.g., variational autoencoders,
  - e.g., normalizing flows,
  - e.g., diffusion models

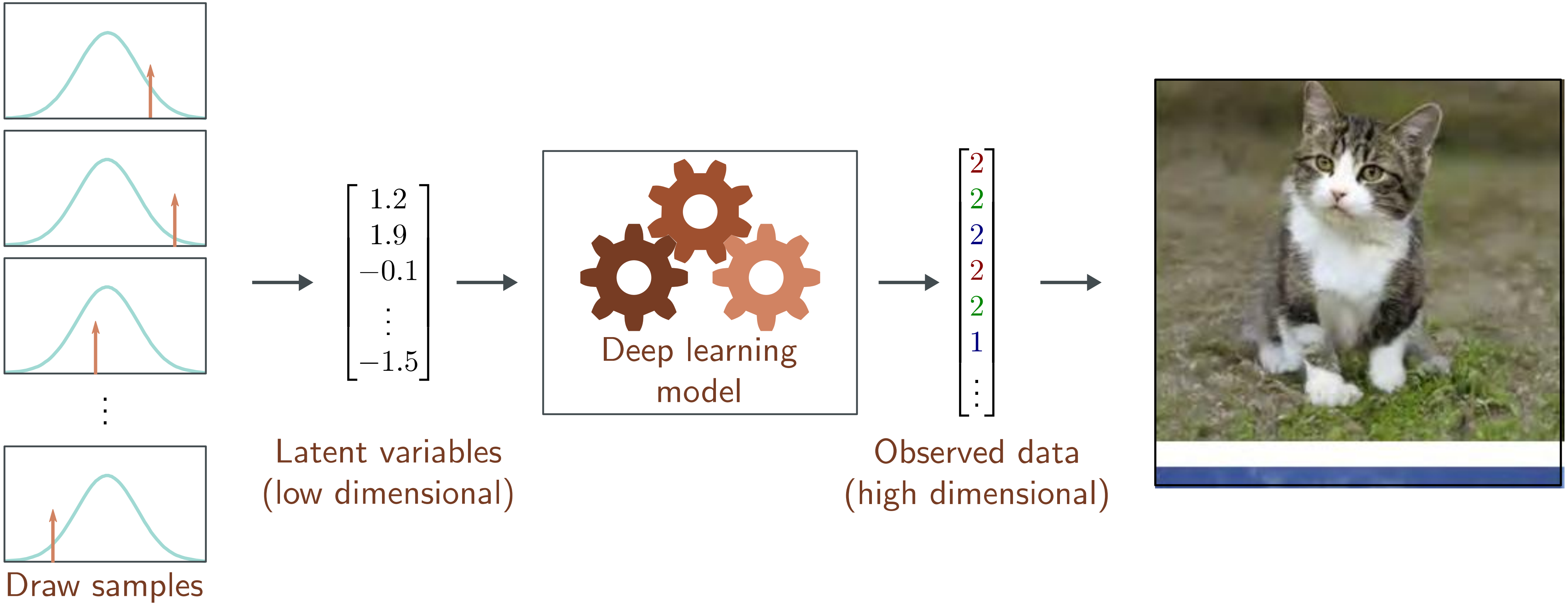


Generative models

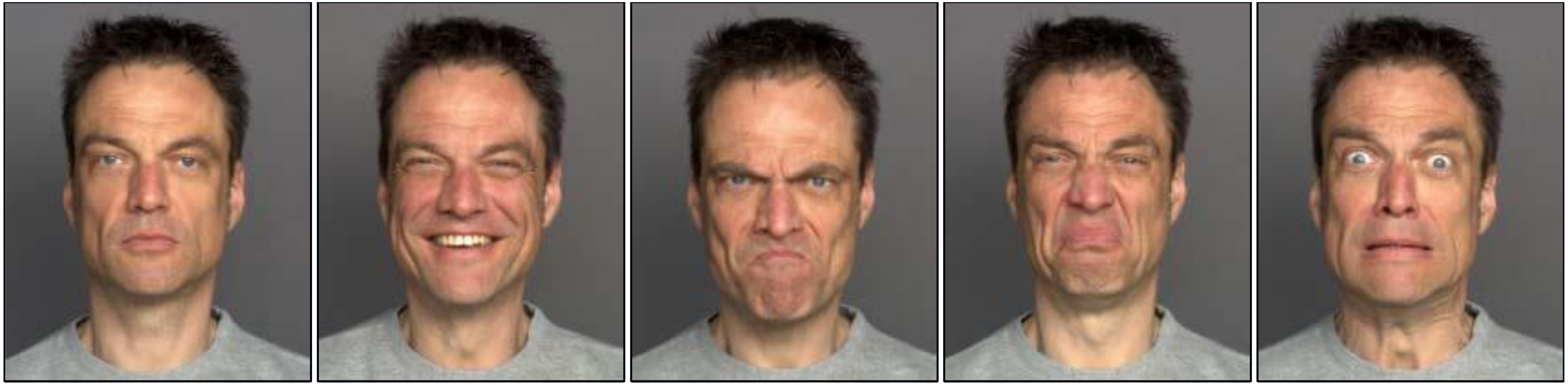


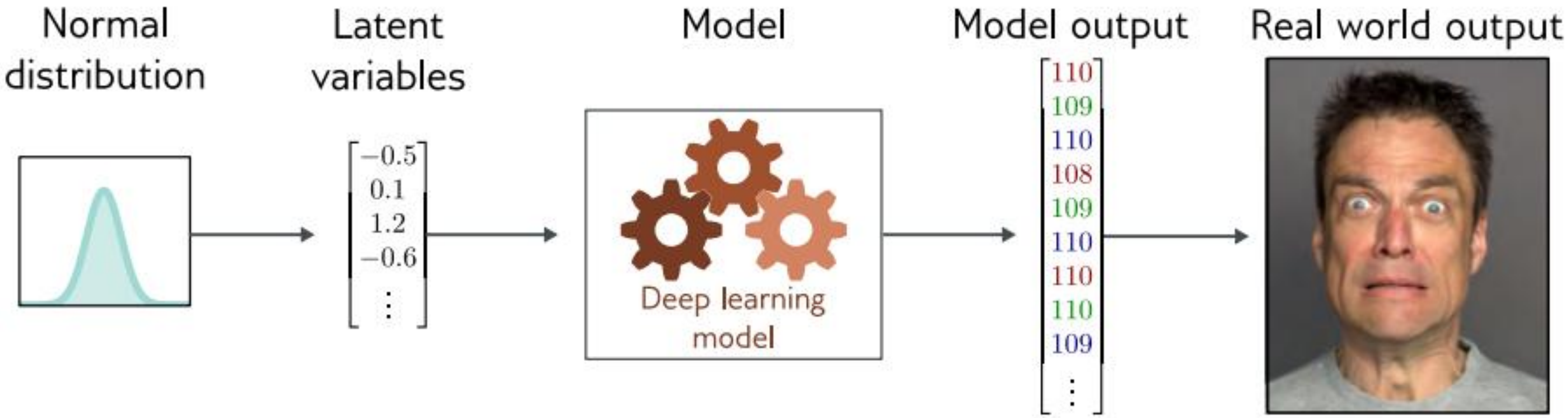
Generative models

Latent variables

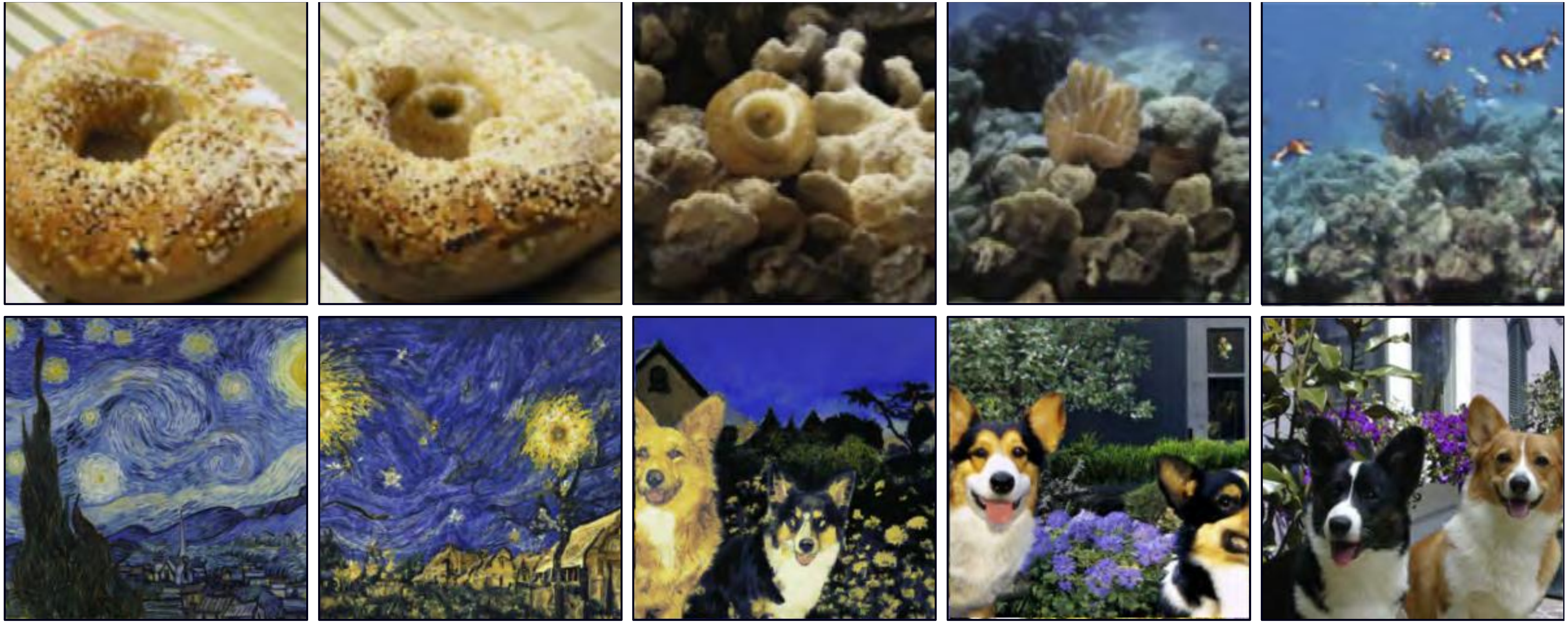


Why should this work?

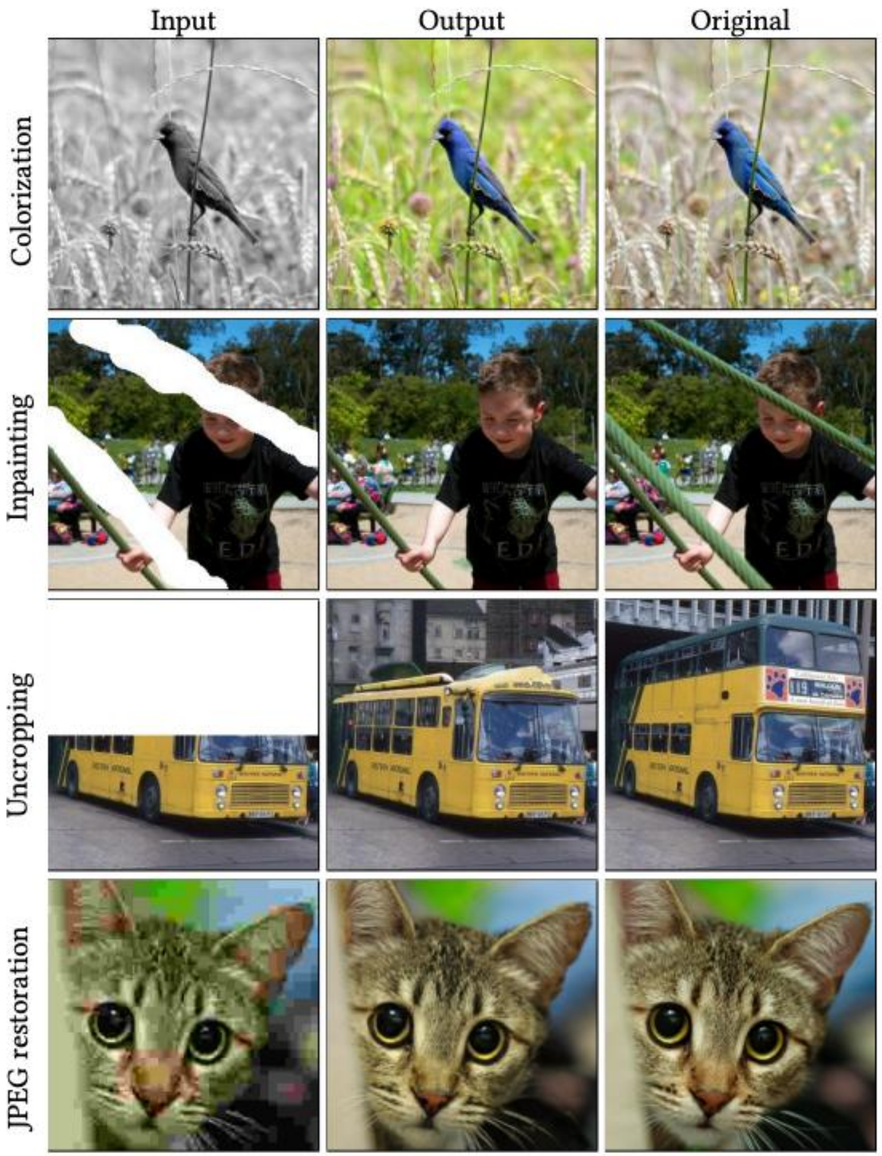




Interpolation



Conditional (diffusion) synthesis



NORTH SOUTH UNIVERSITY

I was a little nervous before my first lecture at the University of Bath. It seemed like there were hundreds of students and they looked intimidating. I stepped up to the lectern and was about to speak, when something bizarre happened.

Suddenly, the room was filled with a deafening noise, like a giant roar. It was so loud that I couldn’t hear anything else and I had to cover my ears. I could see the students looking around, confused and frightened. Then, as quickly as it had started, the noise stopped and the room was silent again.

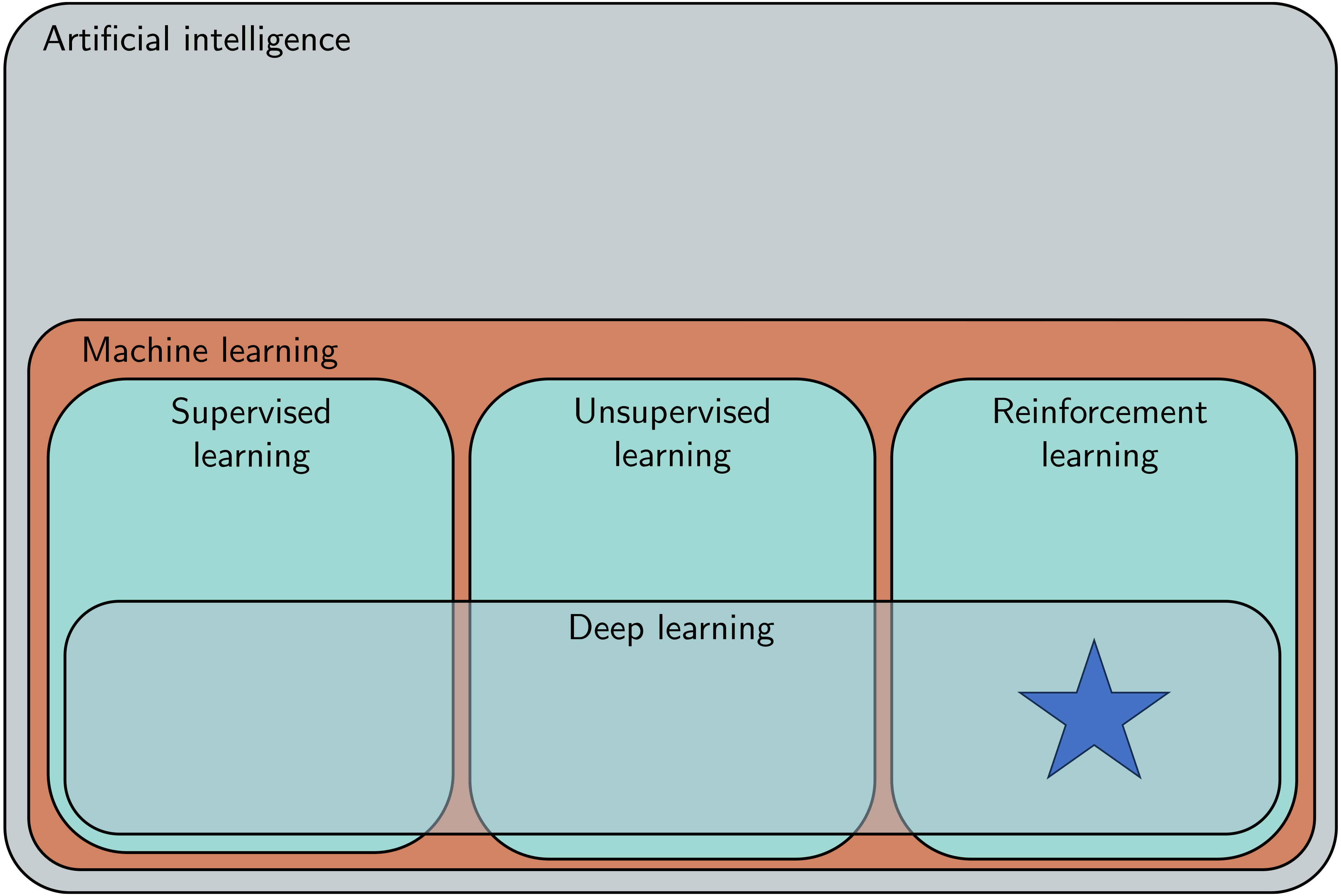
I stood there for a few moments, trying to make sense of what had just happened. Then I realized that the students were all staring at me, waiting for me to say something. I tried to think of something witty or clever to say, but my mind was blank. So I just said, ”Well, that was strange,” and then I started my lecture.

I was a little nervous before my first lecture at the University of Bath. It seemed like there were hundreds of students and they looked intimidating. I stepped up to the lectern and was about to speak, when something bizarre happened.

Suddenly, a giant rabbit ran into the lecture hall! The students started screaming and running around in panic. I was so shocked that I couldn’t move. The rabbit ran up to me and hopped onto the lectern. Then, in a booming voice, it said:

”I am the Easter Bunny! I have come to give you all a special gift!”

The students were so surprised that they stopped screaming and listened to the Easter Bunny. Then, the Easter Bunny started handing out chocolate eggs to everyone in the lecture hall. The students were so happy that they started cheering and clapping. I was so relieved that the Easter Bunny had saved my lecture! After that, I was able to continue and the students paid attention for the rest of the hour. It was a great success!


Artificial intelligence
Machine learning
Supervised learning
Unsupervised learning
Reinforcement learning
Deep learning

Reinforcement learning

- A set of states
- A set of actions
- A set of rewards

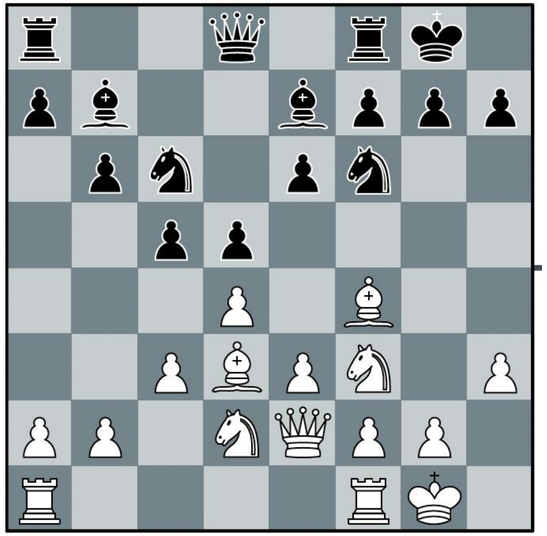
- Goal: take actions to change the state so that you receive rewards

- You don’t receive any data – you have to explore the environment yourself to gather data as you go

Example: chess

- States are valid states of the chess board
- Actions at a given time are valid possible moves
- Positive rewards for taking pieces, negative rewards for losing them

State

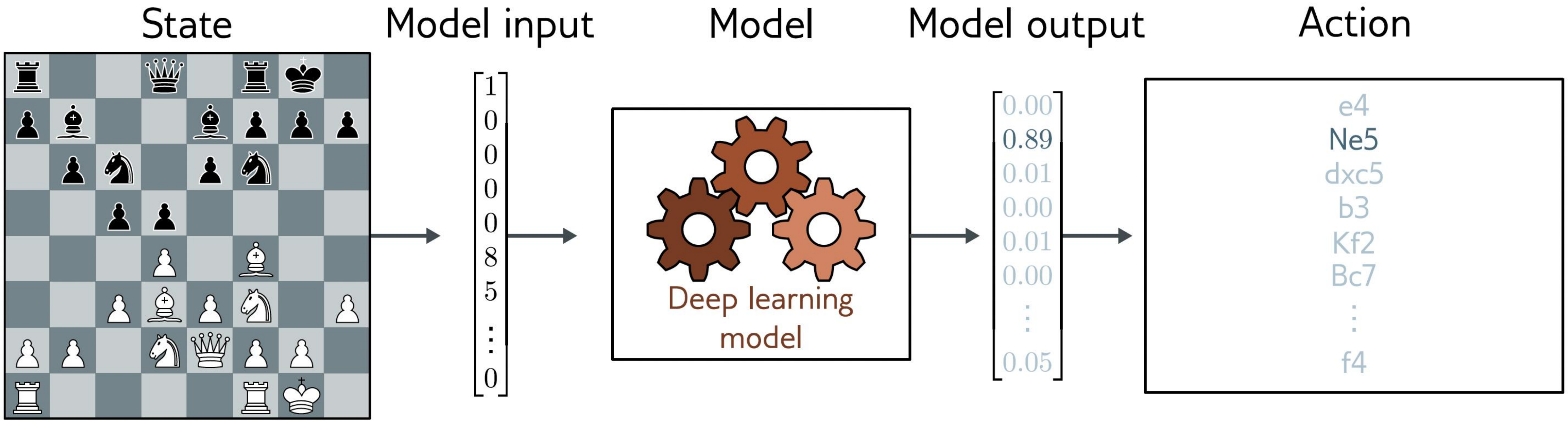


Action

e4
Ne5
dxc5
b3
Kf2
Bc7
⋮
f4

Example: chess

- States are valid states of the chess board
- Actions at a given time are valid possible moves
- Positive rewards for taking pieces, negative rewards for losing them



Why is this difficult?

- Stochastic
  - Make the same move twice, the opponent might not do the same thing
  - Rewards also stochastic (opponent does or doesn’t take your piece)
- Temporal credit assignment problem
  - Did we get the reward because of this move? Or because we made good tactical decisions somewhere in the past?
- Exploration-exploitation trade-off
  - If we found a good opening, should we use this?
  - Or should we try other things, hoping for something better?