

Machine Learning

1. Introduction

Book. Understanding Deep Learning

Simon J.D. Prince

May 29, 2025

Artificial intelligence

Artificial intelligence

Machine learning

Artificial intelligence

Machine learning

Supervised
learning

Unsupervised
learning

Reinforcement
learning

Artificial intelligence

Machine learning

Supervised
learning

Unsupervised
learning

Reinforcement
learning

Deep learning

Artificial intelligence

Machine learning

Supervised
learning

Unsupervised
learning

Reinforcement
learning

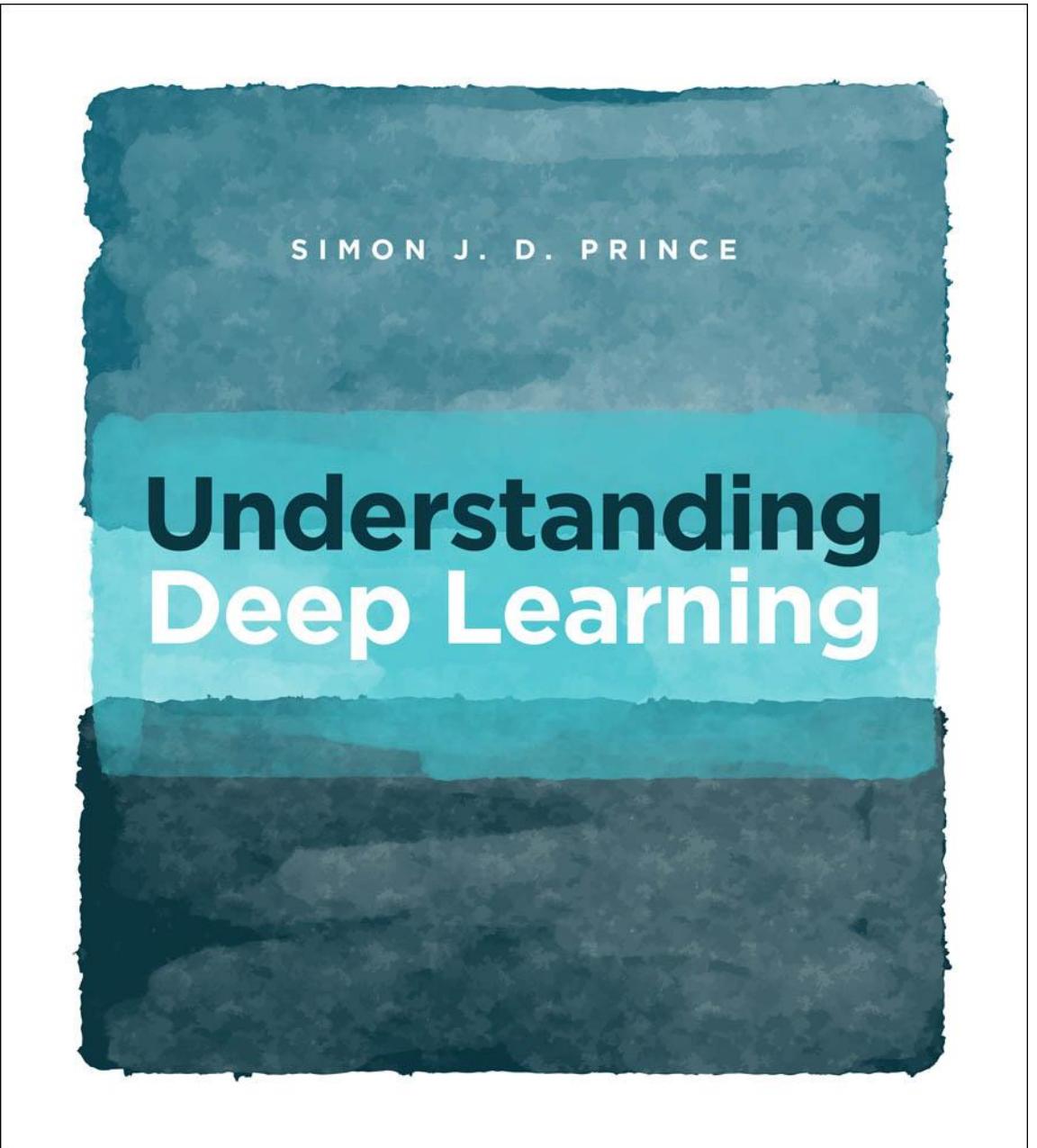
Deep learning

Lab sessions

- Python notebooks in CoLab
- You will need a Google account
 - Numpy
 - Matplotlib
 - PyTorch
- Problem sheets

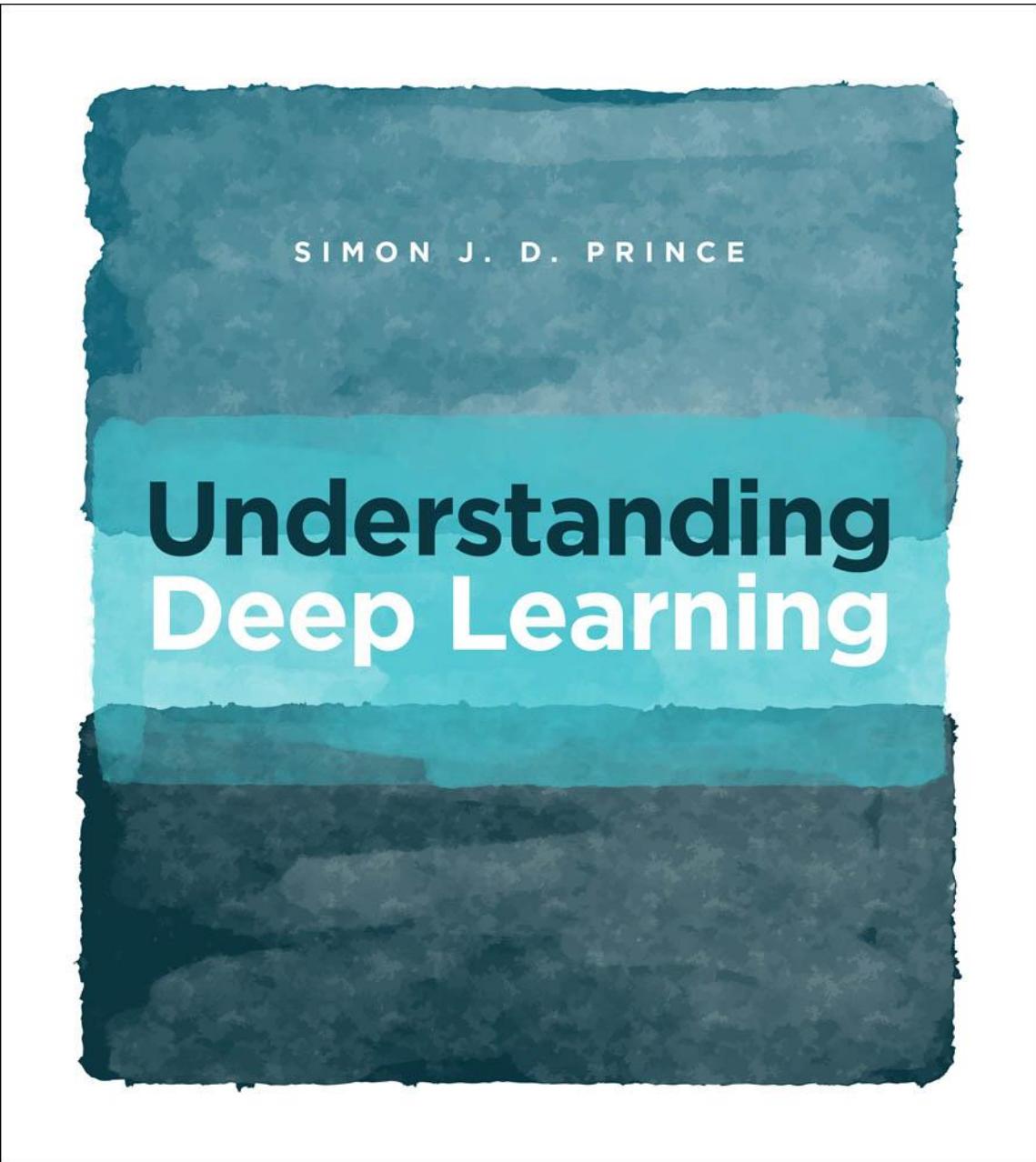
Book

Book



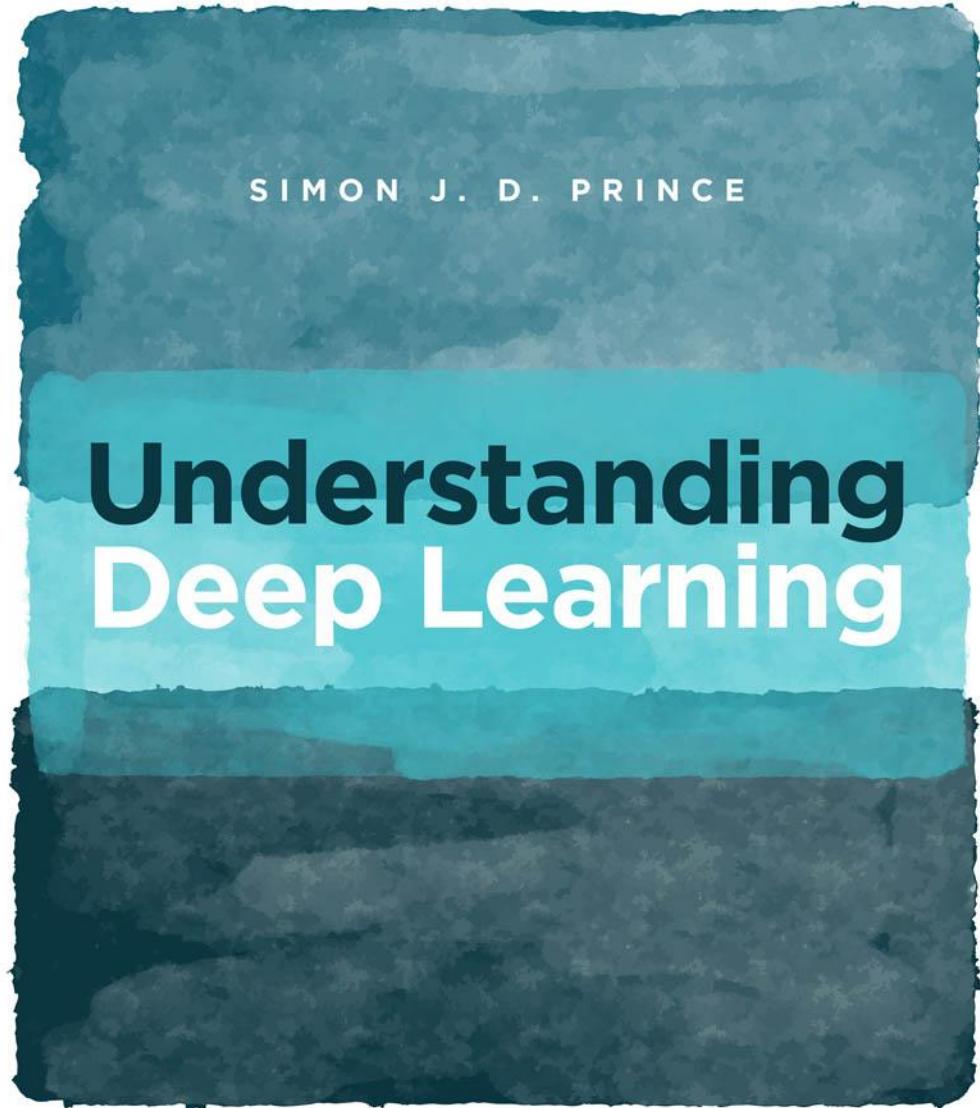
Book

Published 5th December 2023



Book

<http://udlbook.com>



Artificial intelligence

Machine learning

Supervised
learning

Unsupervised
learning

Reinforcement
learning

Deep learning



Supervised learning

- Define a mapping from input to output
- Learn this mapping from paired input/output data examples

Regression

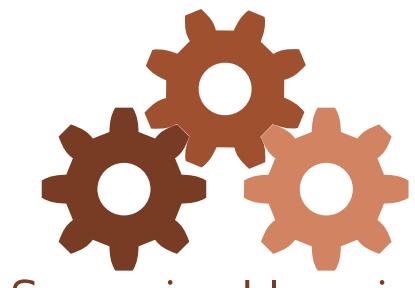
Real world input

6000 square feet,
4 bedrooms,
previously sold for
\$235K in 2005,
1 parking spot.

Model
input

$$\begin{bmatrix} 6000 \\ 4 \\ 235 \\ 2005 \\ 1 \end{bmatrix}$$

Model



Model
output

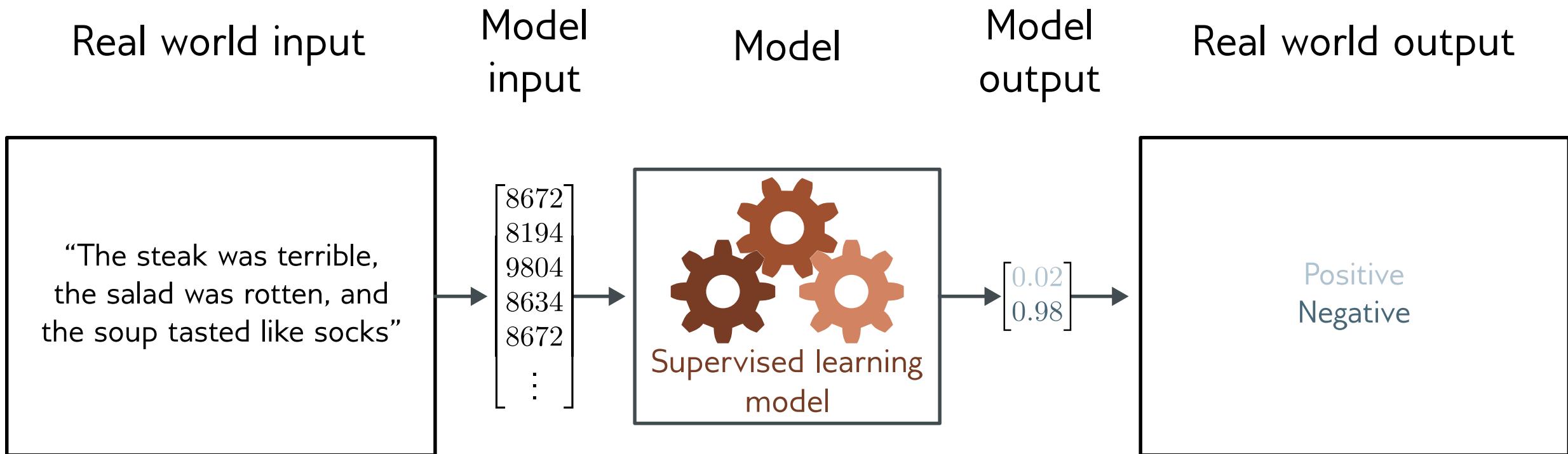
$$[340]$$

Real world output

Predicted price
is \$340k

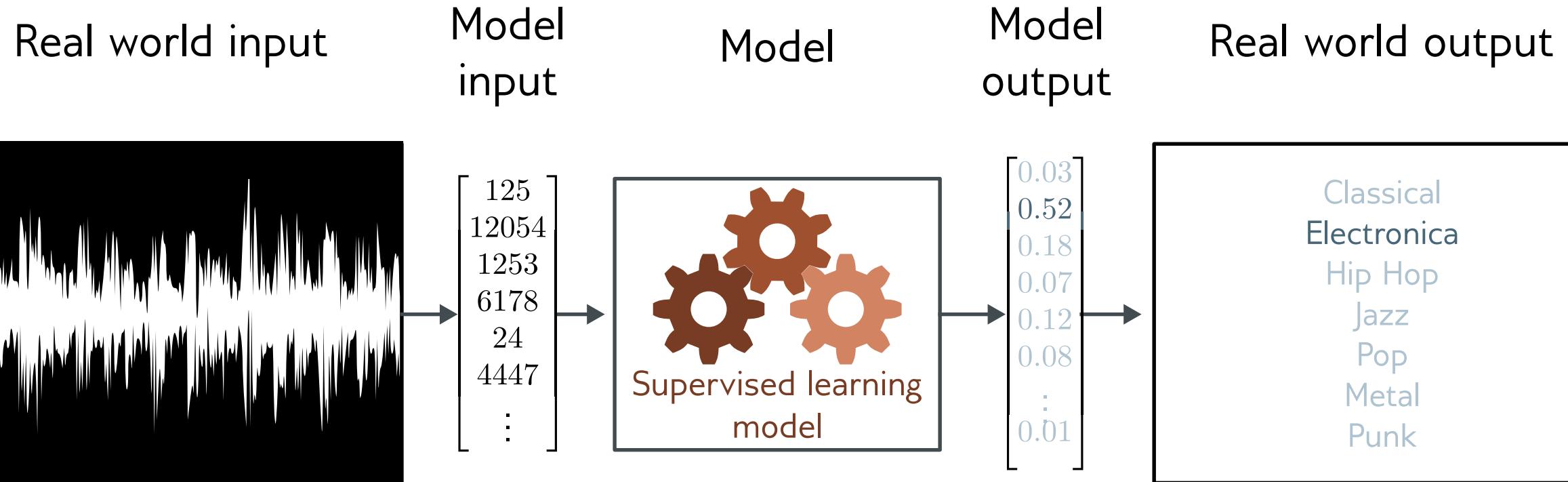
- Univariate regression problem (one output, real value)
- Fully connected network

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

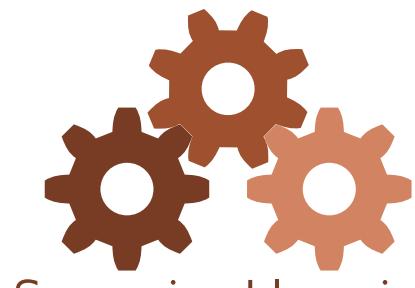
Real world input



Model
input

$$\begin{bmatrix} 124 \\ 140 \\ 156 \\ 128 \\ 142 \\ 157 \\ \vdots \end{bmatrix}$$

Model



Supervised learning
model

Model
output

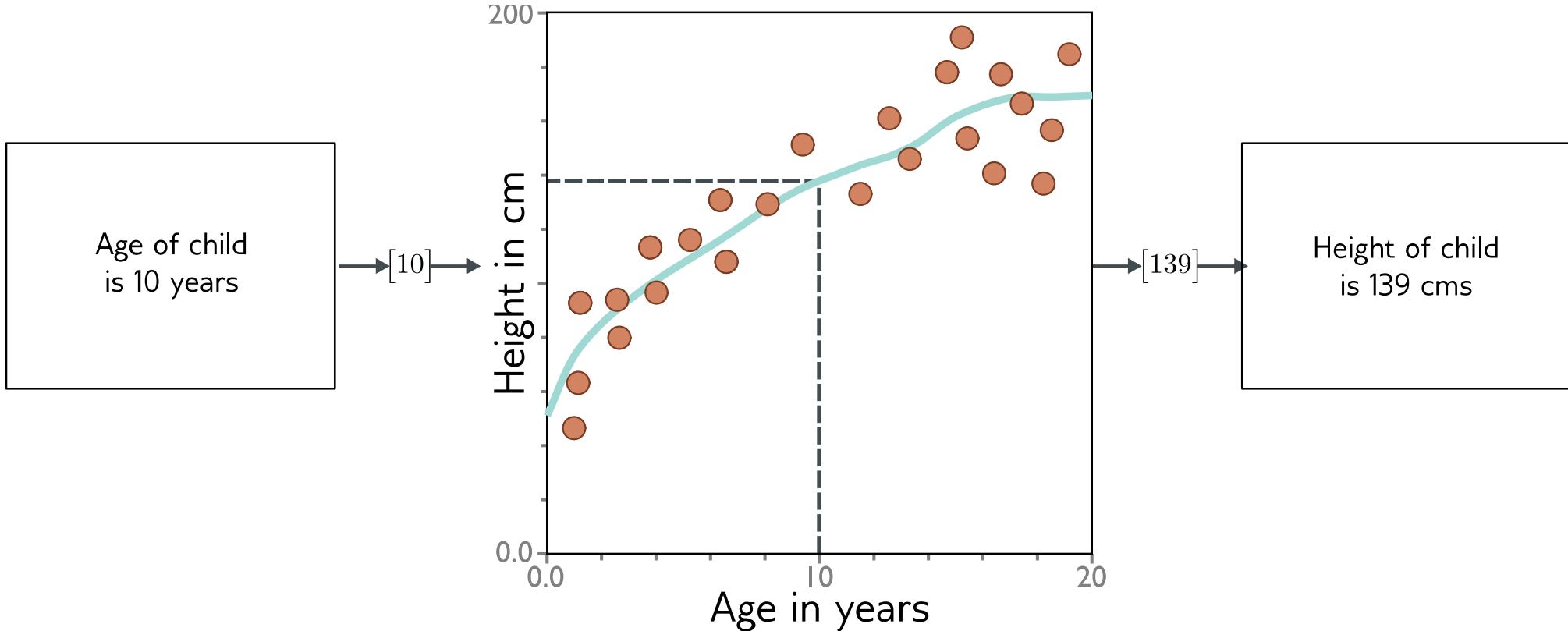
$$\begin{bmatrix} 0.00 \\ 0.00 \\ 0.01 \\ 0.89 \\ 0.05 \\ 0.00 \\ \vdots \\ 0.01 \end{bmatrix}$$

Real world output

Aardvark
Apple
Bee
Bicycle
Bridge
Clown
⋮

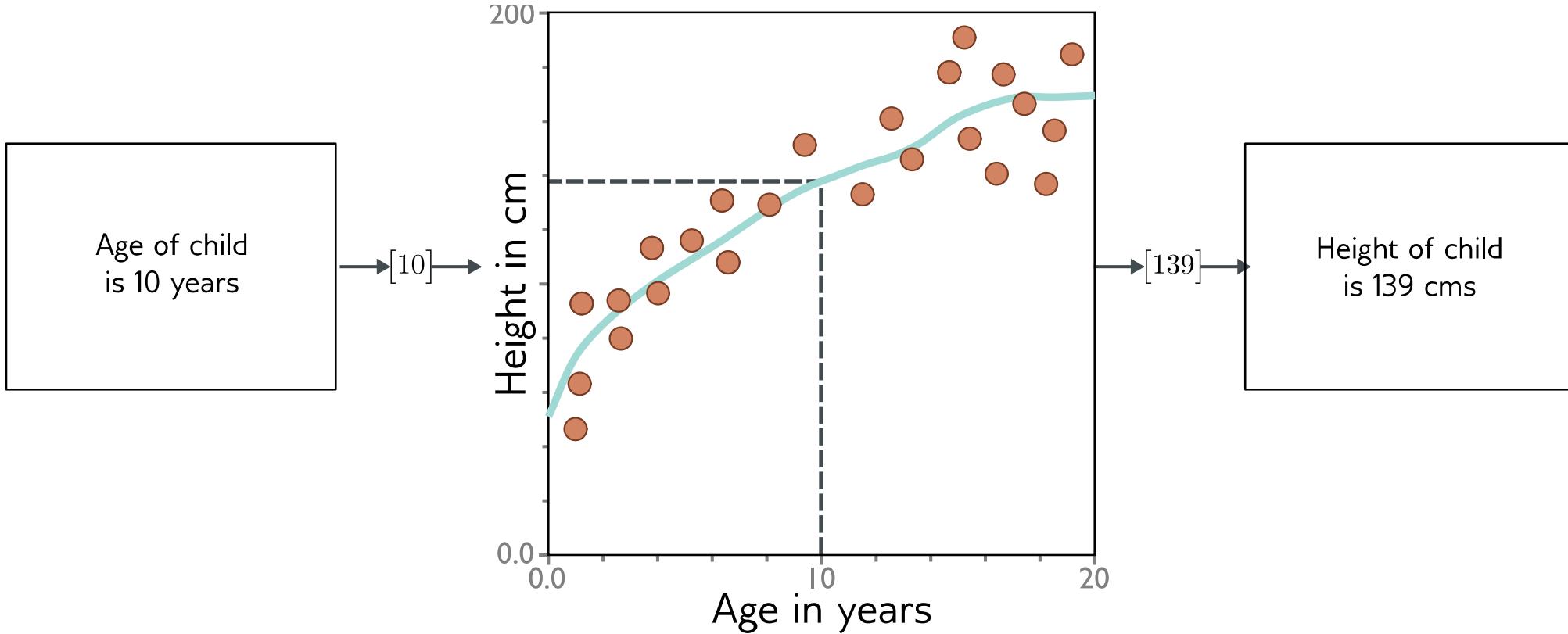
- Multiclass classification problem (discrete classes, >2 possible classes)
- Convolutional network

What is a supervised learning model?



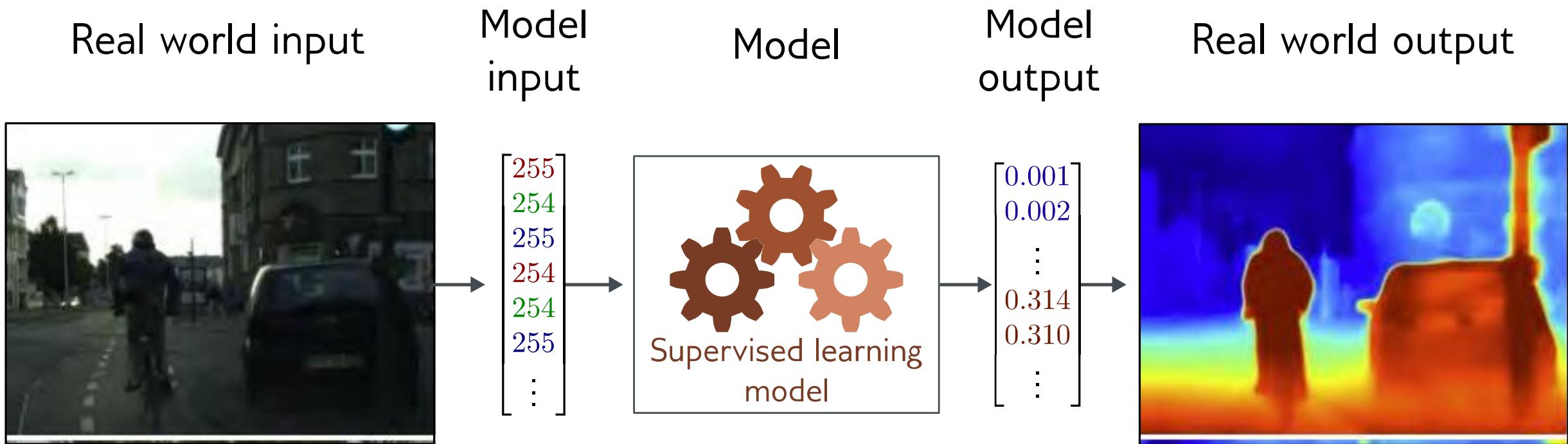
- An equation relating input (age) to output (height)
- Search through family of possible equations to find one that fits training data well

What is a supervised learning model?



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

Depth estimation



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

Pose estimation

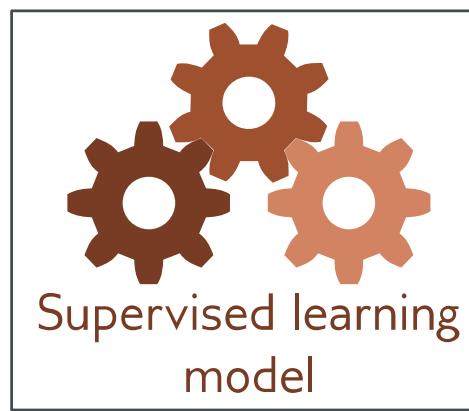
Real world input



Model
input

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

Model



Model
output

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

Real world output



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

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

Translation

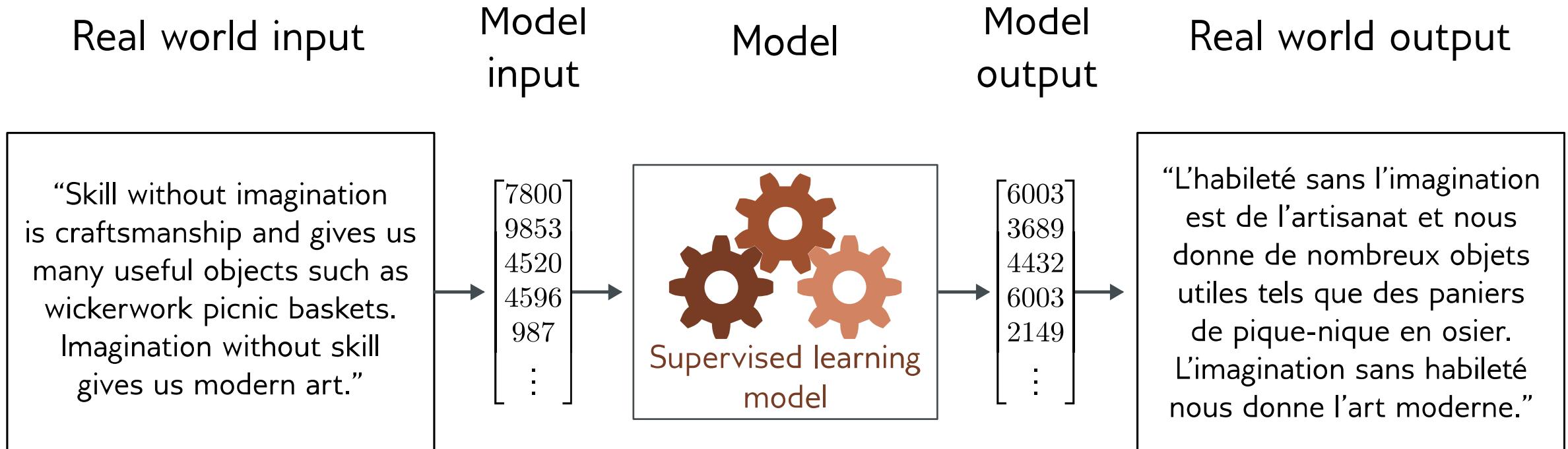


Image captioning

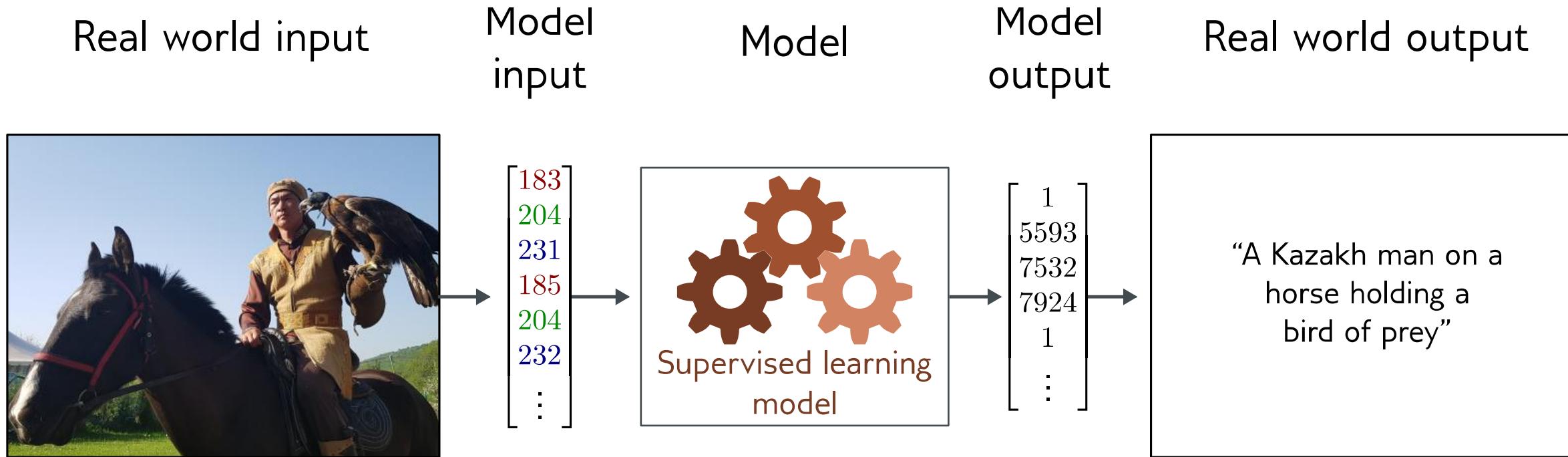
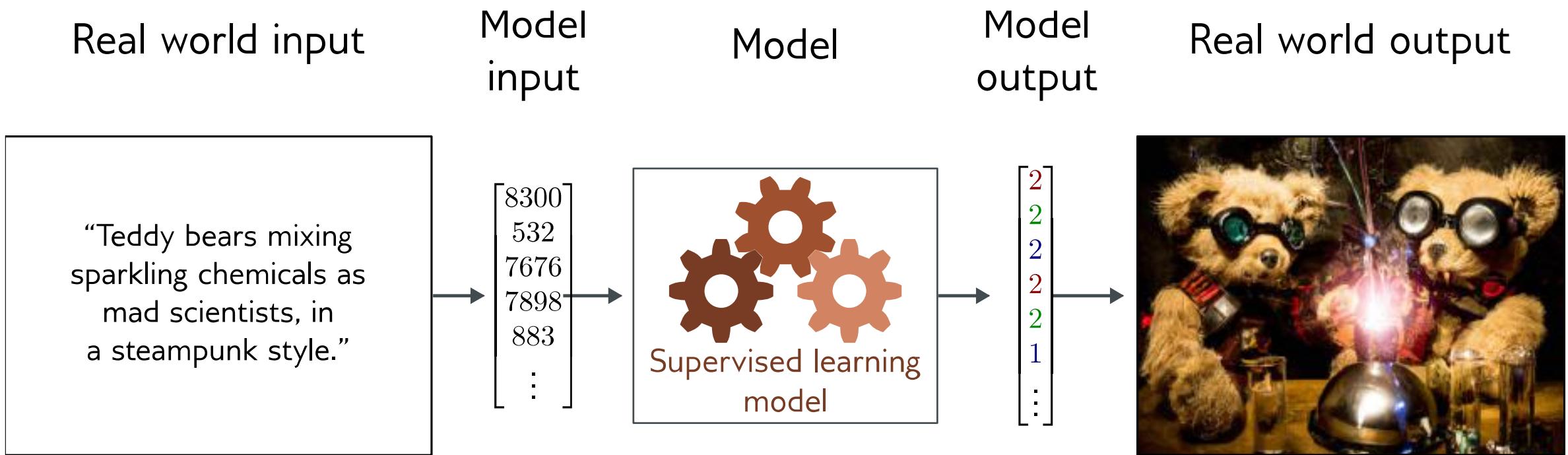


Image generation from text



What do these examples have in common?

- Very complex relationship between input and output
- 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”

Idea

- 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 easier by having a lot of knowledge of possible outputs

Artificial intelligence

Machine learning

Supervised
learning

Unsupervised
learning

Reinforcement
learning

Deep learning



Unsupervised Learning

Rather than learning a mapping from input to output, the goal is to describe or understand the structure of the data.

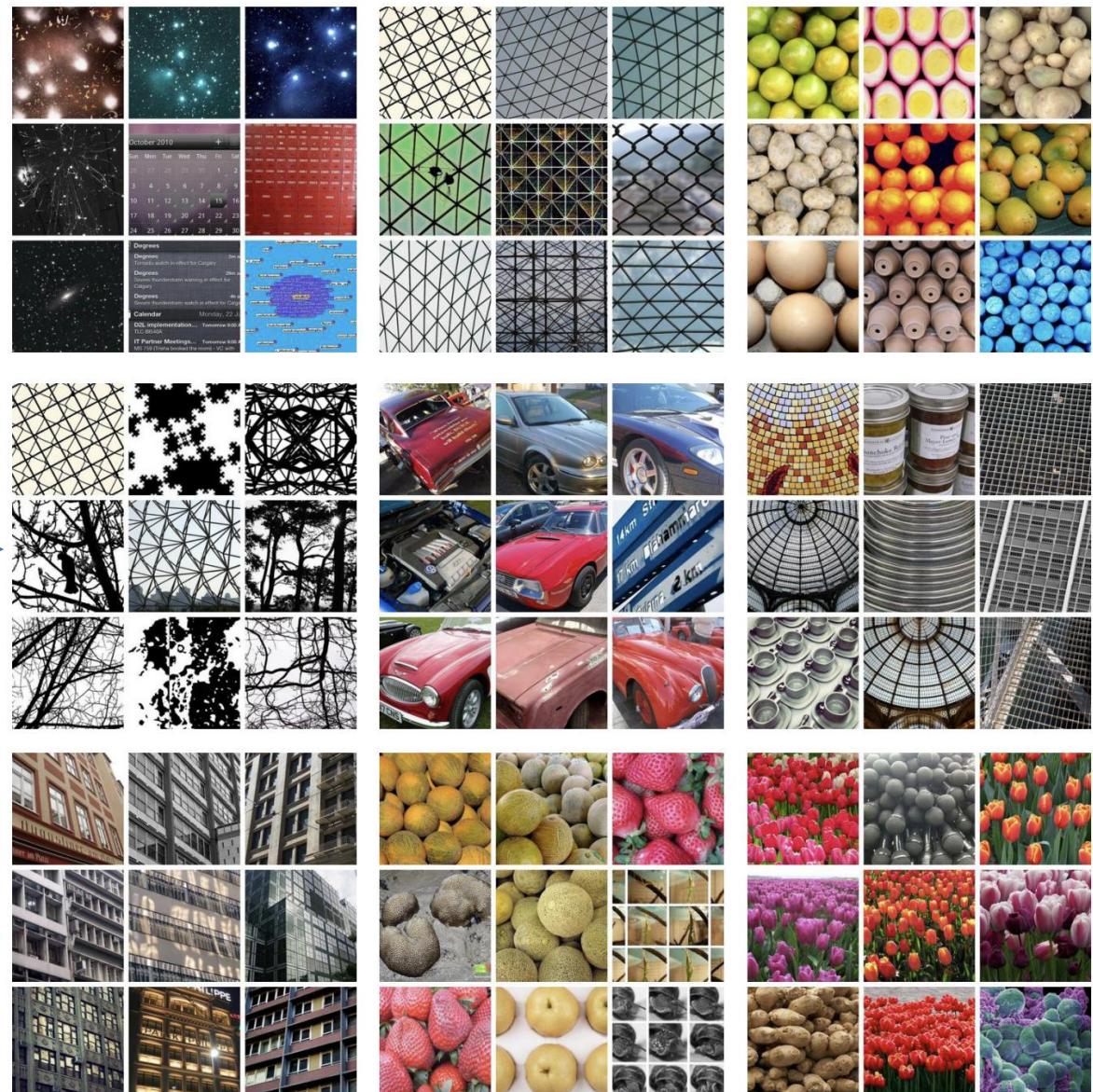
- Learning about a dataset without labels

- Clustering
- Finding outliers
- Generating new examples
- Filling in missing data

Unsupervised learning



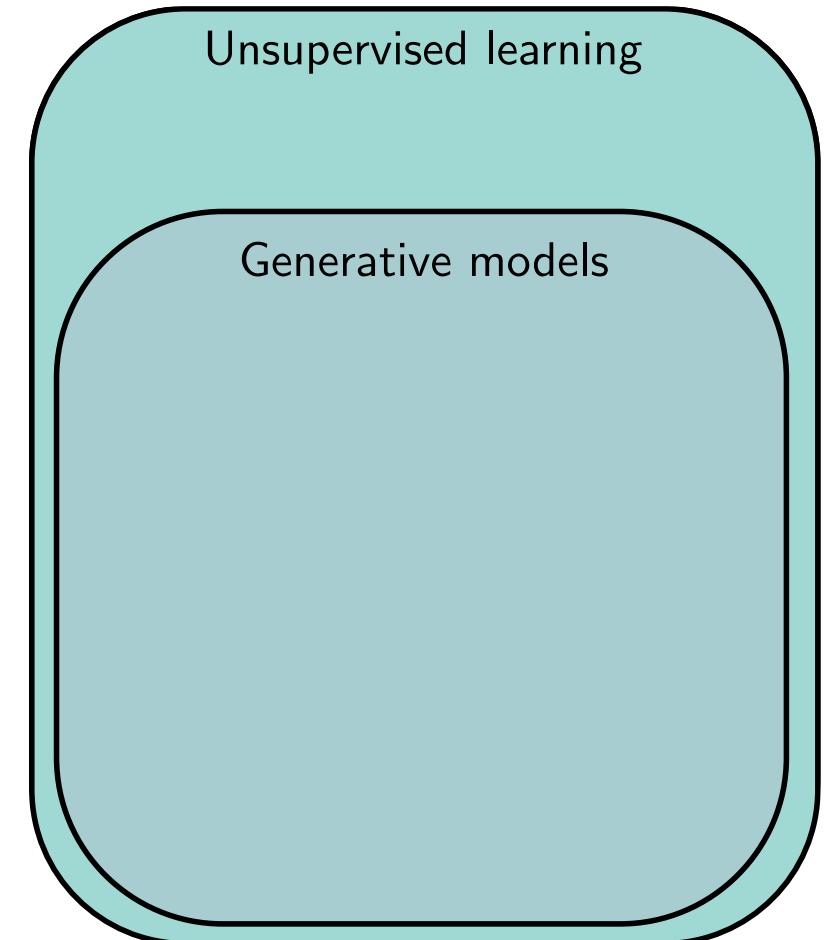
DeepCluster: Deep Clustering for Unsupervised Learning of Visual Features (Caron et al., 2018)



DeepCluster: Deep Clustering for Unsupervised Learning of Visual Features (Caron et al., 2018)

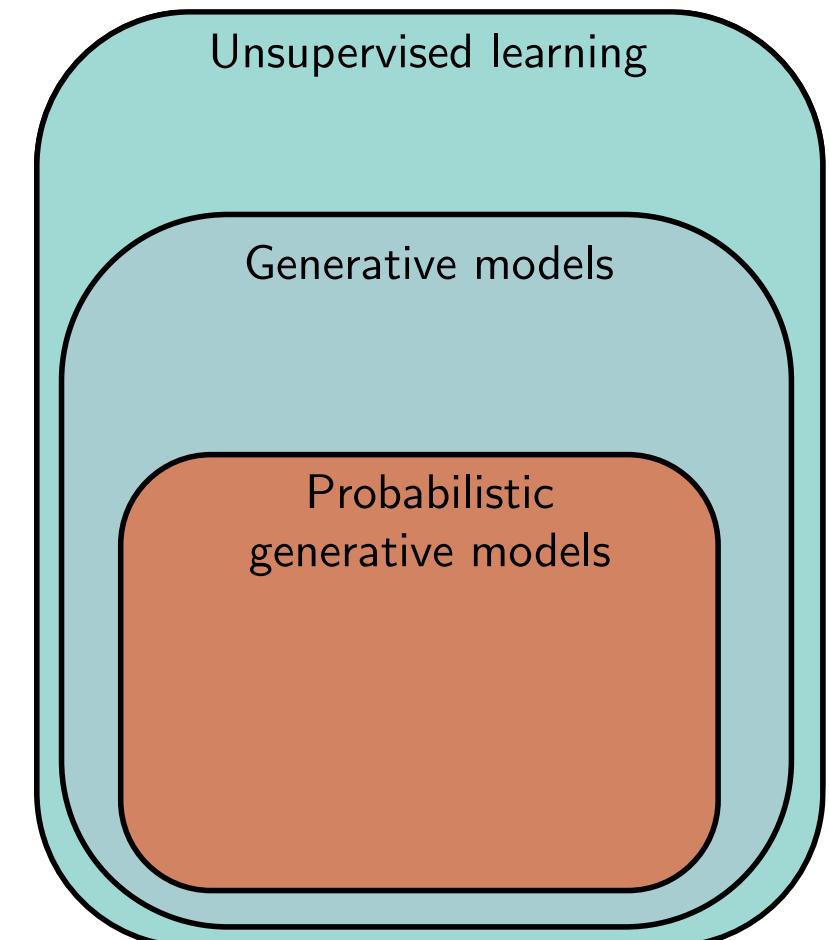
Unsupervised Learning

- 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
 - learn to synthesize new data examples that are statistically indistinguishable from the training data
- PGMs learn distribution over data
 - e.g., variational autoencoders,
 - e.g., normalizing flows,
 - e.g., diffusion models



Generative models

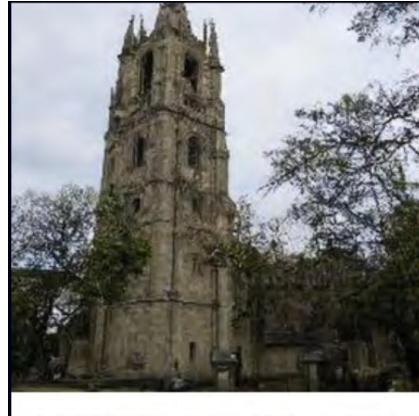
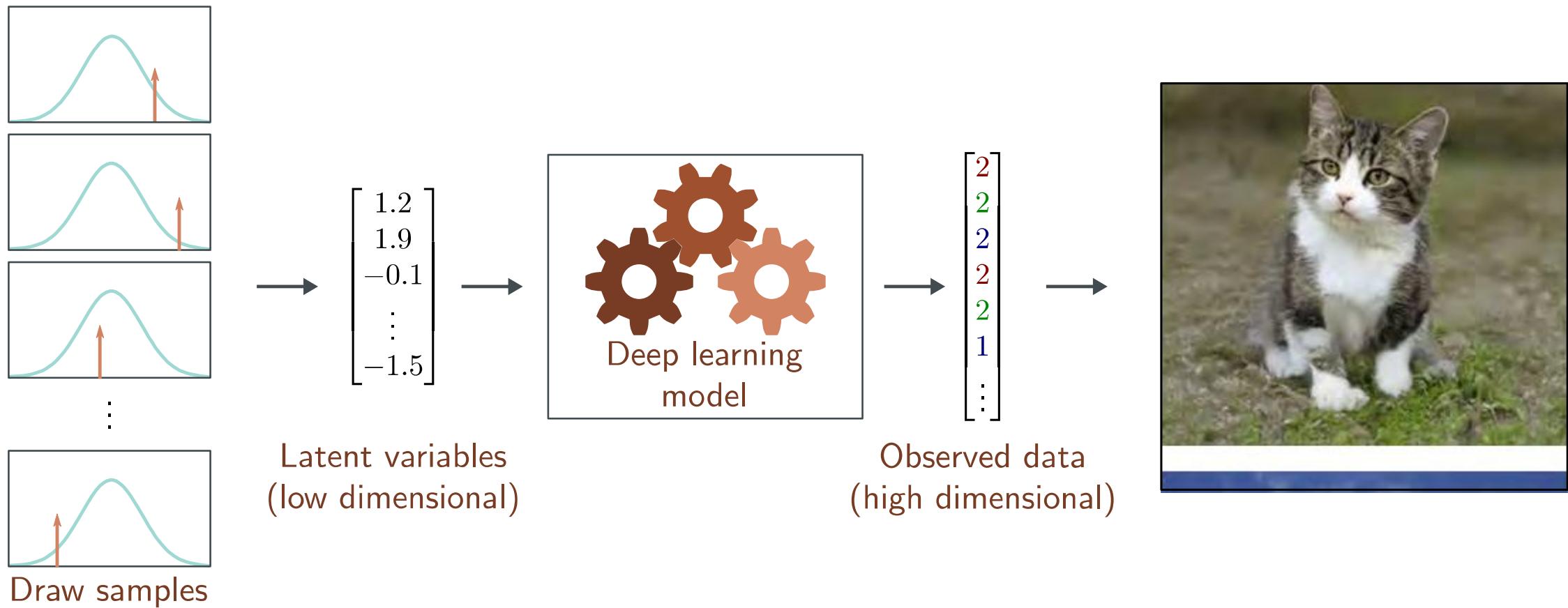


Figure 1.5 Generative models for images. Left: two images were generated from a model trained on pictures of cats. These are not real cats, but samples from a probability model. Right: two images generated from a model trained on images of buildings. Adapted from Karras et al. (2020b)

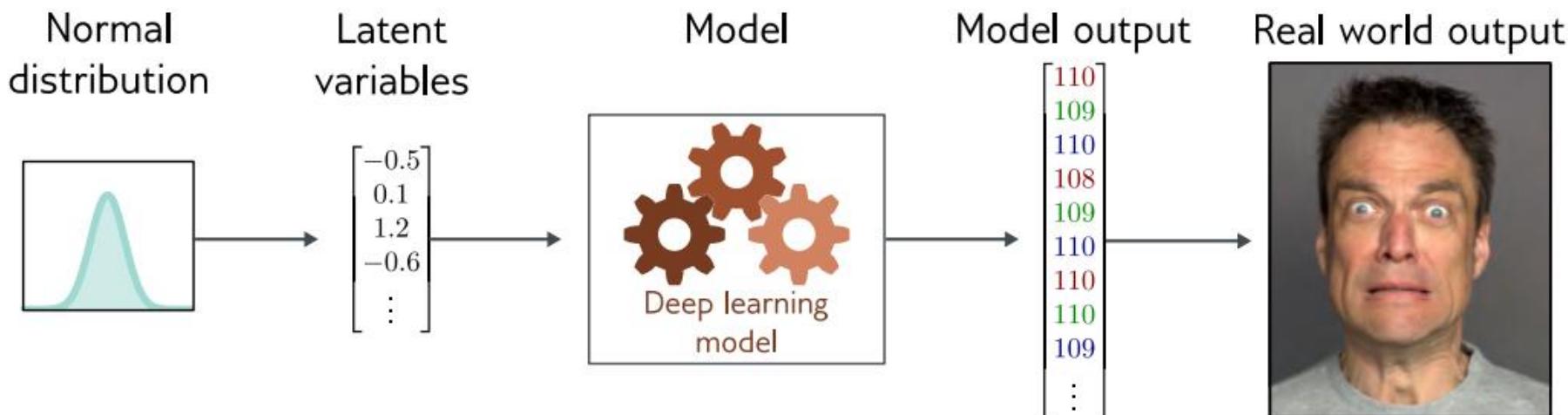
Latent variables



Why should this work?



Figure 1.9 Variation of the human face. The human face contains roughly 42 muscles, so it's possible to describe most of the variation in images of the same person in the same lighting with just 42 numbers



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!

Figure 1.8 Conditional text synthesis. Given an initial body of text (in black), generative models of text can continue the string plausibly by synthesizing the “missing” remaining part of the string. Generated by GPT3 (Brown et al., 2020).

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

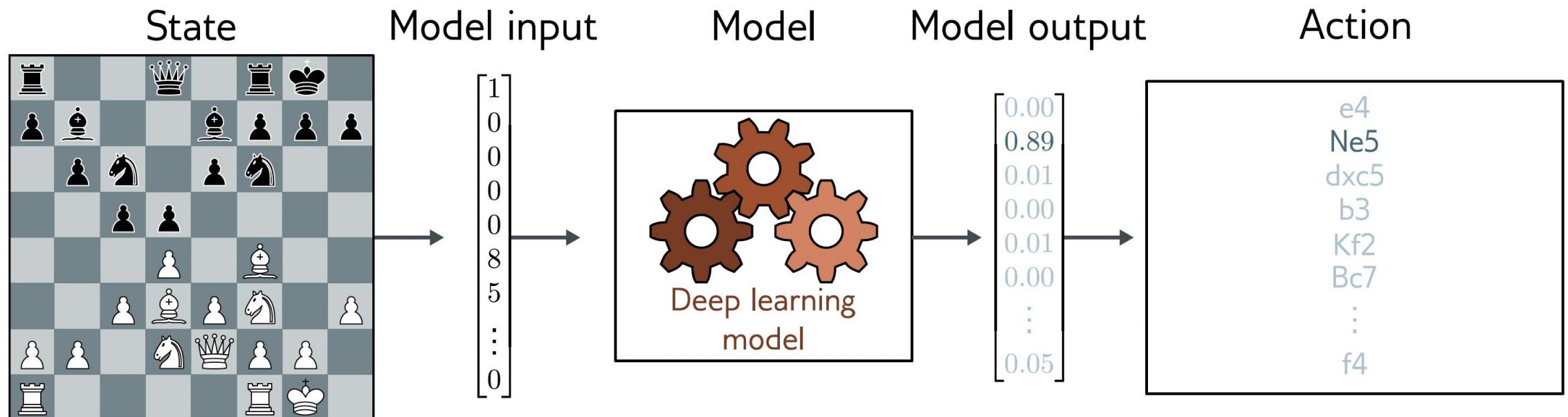


Action

:
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?

Landmarks in Deep Learning

- 1958 Perceptron (Simple ‘neural’ model)
- 1986 Backpropagation (Practical Deep Neural networks)
- 1989 Convolutional networks (Supervised learning)
- 2012 AlexNet Image classification (Supervised learning)
- 2014 Generative adversarial networks (Unsupervised learning)
- 2014 Deep Q-Learning -- Atari games (Reinforcement learning)
- 2016 AlphaGo (Reinforcement learning)
- 2017 Machine translation (Supervised learning)
- 2019 Language models ((Un)supervised learning)
- 2022 Dall-E2 Image synthesis from text prompts ((Un)supervised learning)
- 2022 ChatGPT ((Un)supervised learning)
- 2023 GPT4 Multimodal model ((Un)supervised learning)

2018 Turing award winners



Yoshua Bengio



Geoffrey Hinton



Yann LeCun

The book

- Chapter 1 - Introduction
 - Chapter 2 - Supervised learning
 - Chapter 3 - Shallow neural networks
 - Chapter 4 - Deep neural networks
 - Chapter 5 - Loss functions
 - Chapter 6 - Training models
 - Chapter 7 - Gradients and initialization
 - Chapter 8 - Measuring performance
 - Chapter 9 - Regularization
 - Chapter 10 - Convolutional networks
 - Chapter 11 - Residual networks
 - Chapter 12 - Transformers
 - Chapter 13 - Graph neural networks
 - Chapter 14 - Unsupervised learning
 - Chapter 15 - Generative adversarial networks
 - Chapter 16 - Normalizing flows
 - Chapter 17 - Variational autoencoders
 - Chapter 18 - Diffusion models
 - Chapter 19 - Deep reinforcement learning
 - Chapter 20 - Why does deep learning work?
 - Chapter 21 - Deep learning and ethics
- 
- Deep neural networks
How to train them
How to measure their performance
How to make that performance better

The book

- Chapter 1 - Introduction
 - Chapter 2 - Supervised learning
 - Chapter 3 - Shallow neural networks
 - Chapter 4 - Deep neural networks
 - Chapter 5 - Loss functions
 - Chapter 6 - Training models
 - Chapter 7 - Gradients and initialization
 - Chapter 8 - Measuring performance
 - Chapter 9 - Regularization
 - Chapter 10 - Convolutional networks
 - Chapter 11 - Residual networks
 - Chapter 12 - Transformers
 - Chapter 13 - Graph neural networks
 - Chapter 14 - Unsupervised learning
 - Chapter 15 - Generative adversarial networks
 - Chapter 16 - Normalizing flows
 - Chapter 17 - Variational autoencoders
 - Chapter 18 - Diffusion models
 - Chapter 19 - Deep reinforcement learning
 - Chapter 20 - Why does deep learning work?
 - Chapter 21 - Deep learning and ethics
- 
- Networks specialized to images
Image classification
Image segmentation
Pose estimation

The book

- Chapter 1 - Introduction
- Chapter 2 - Supervised learning
- Chapter 3 - Shallow neural networks
- Chapter 4 - Deep neural networks
- Chapter 5 - Loss functions
- Chapter 6 - Training models
- Chapter 7 - Gradients and initialization
- Chapter 8 - Measuring performance
- Chapter 9 - Regularization
- Chapter 10 - Convolutional networks
- Chapter 11 - Residual networks
- Chapter 12 - Transformers
- Chapter 13 - Graph neural networks
- Chapter 14 - Unsupervised learning
- Chapter 15 - Generative adversarial networks
- Chapter 16 - Normalizing flows
- Chapter 17 - Variational autoencoders
- Chapter 18 - Diffusion models
- Chapter 19 - Deep reinforcement learning
- Chapter 20 - Why does deep learning work?
- Chapter 21 - Deep learning and ethics



Networks specialized to text
Text generation
Automatic translation
ChatGPT

The book

- Chapter 1 - Introduction
- Chapter 2 - Supervised learning
- Chapter 3 - Shallow neural networks
- Chapter 4 - Deep neural networks
- Chapter 5 - Loss functions
- Chapter 6 - Training models
- Chapter 7 - Gradients and initialization
- Chapter 8 - Measuring performance
- Chapter 9 - Regularization
- Chapter 10 - Convolutional networks
- Chapter 11 - Residual networks
- Chapter 12 - Transformers
- Chapter 13 - Graph neural networks
- Chapter 14 - Unsupervised learning
- Chapter 15 - Generative adversarial networks
- Chapter 16 - Normalizing flows
- Chapter 17 - Variational autoencoders
- Chapter 18 - Diffusion models
- Chapter 19 - Deep reinforcement learning
- Chapter 20 - Why does deep learning work?
- Chapter 21 - Deep learning and ethics



Generative learning (unsupervised)
Generating random cats!