Learning Perspectives: From Discriminative to Generative

COMP6258

Let’s start with from a discriminative perspective

Task: given a classification problem, train a good model

Task: given a classification problem, train a good model

- You can now reason about choosing:
  - An architecture,
  - A loss function,
  - Some regularisers,
  - Data augmentation, etc.

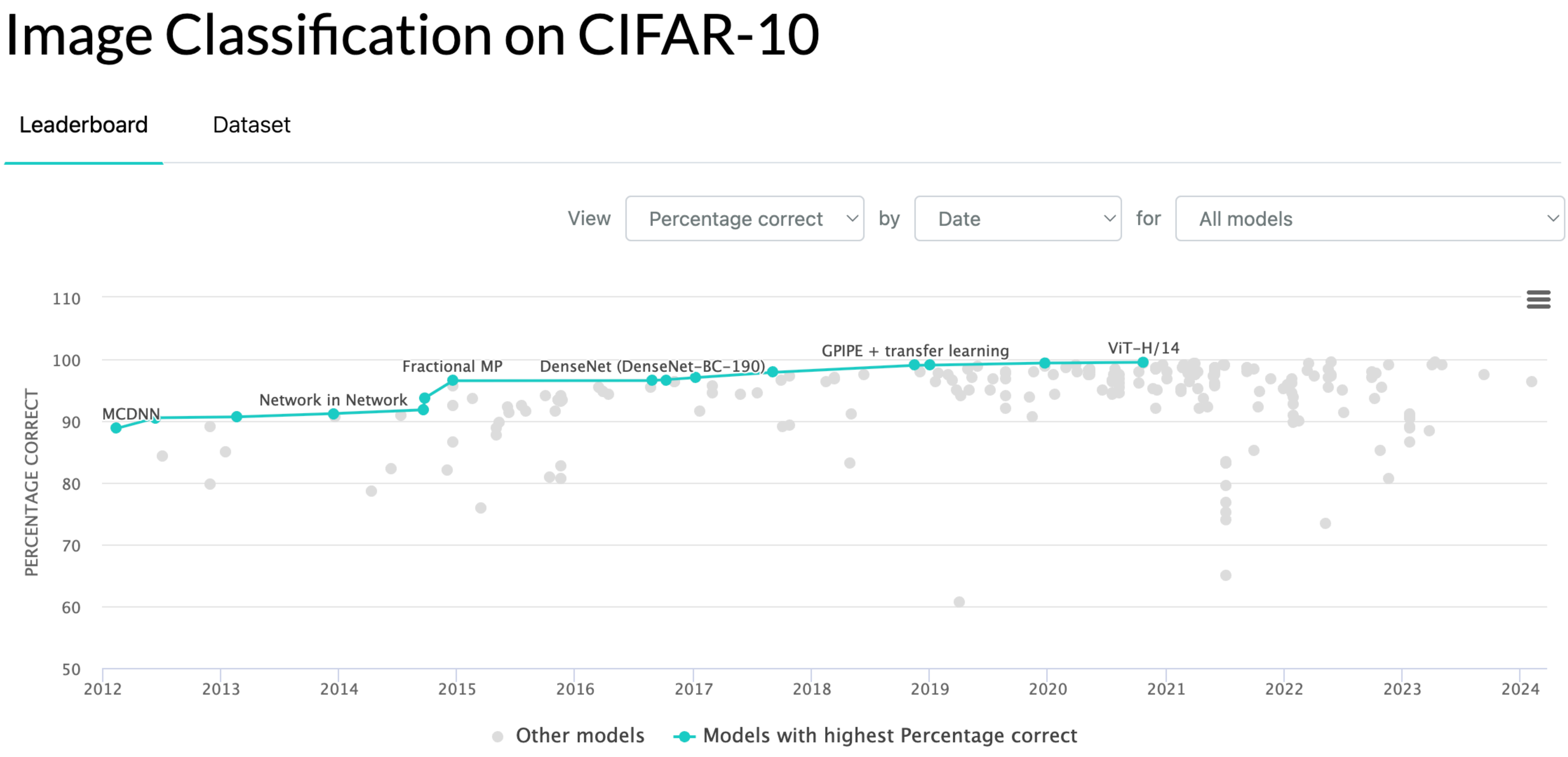
Task: given a classification problem, train a good model

- You can now reason about choosing:
    - An architecture,
    - A loss function,
    - Some regularisers,
    - Data augmentation, etc.
- How can you tell if the model you decide to use it’s good enough?

Test Accuracy Is Not Everything

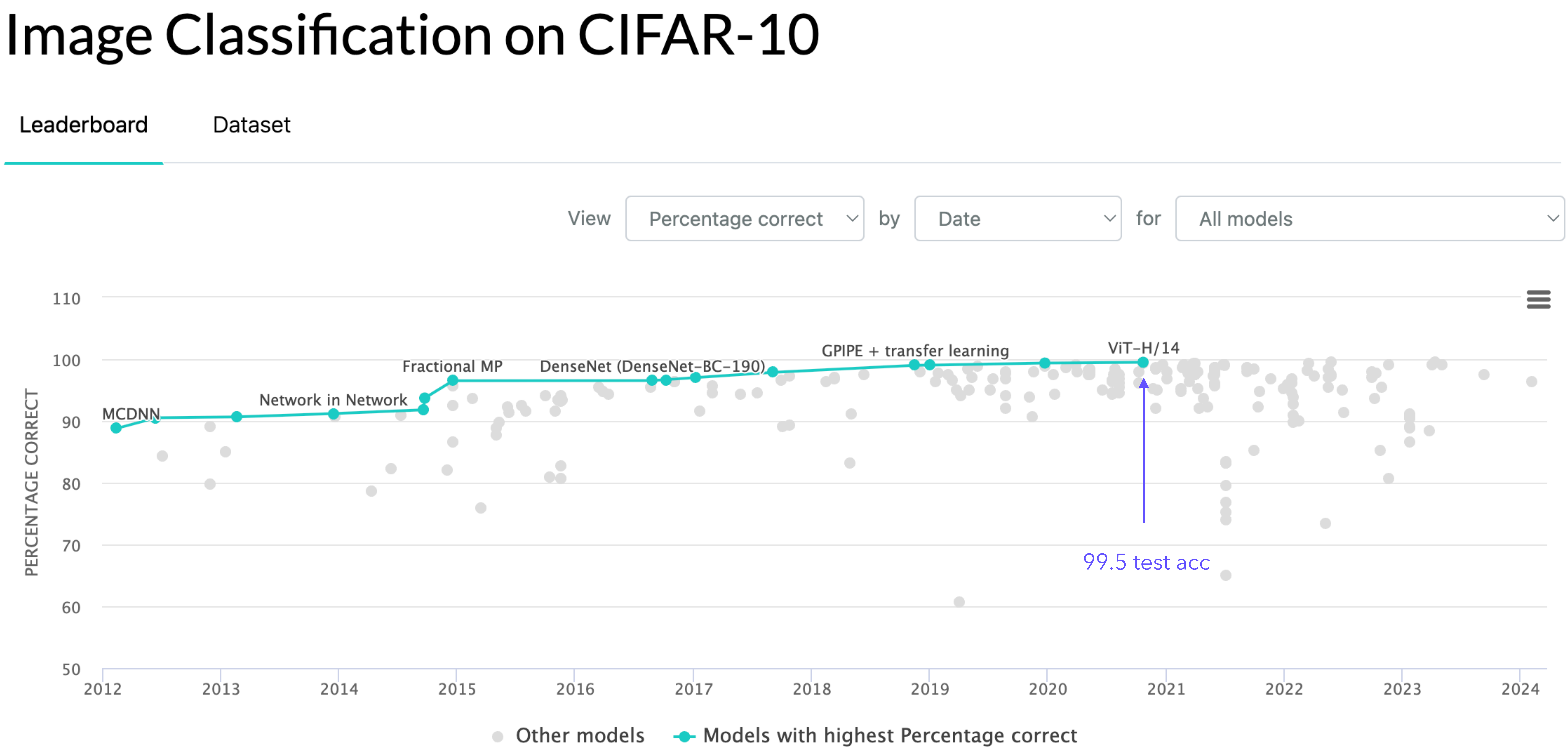
(Reminder)

CIFAR-10 state-of-the-art



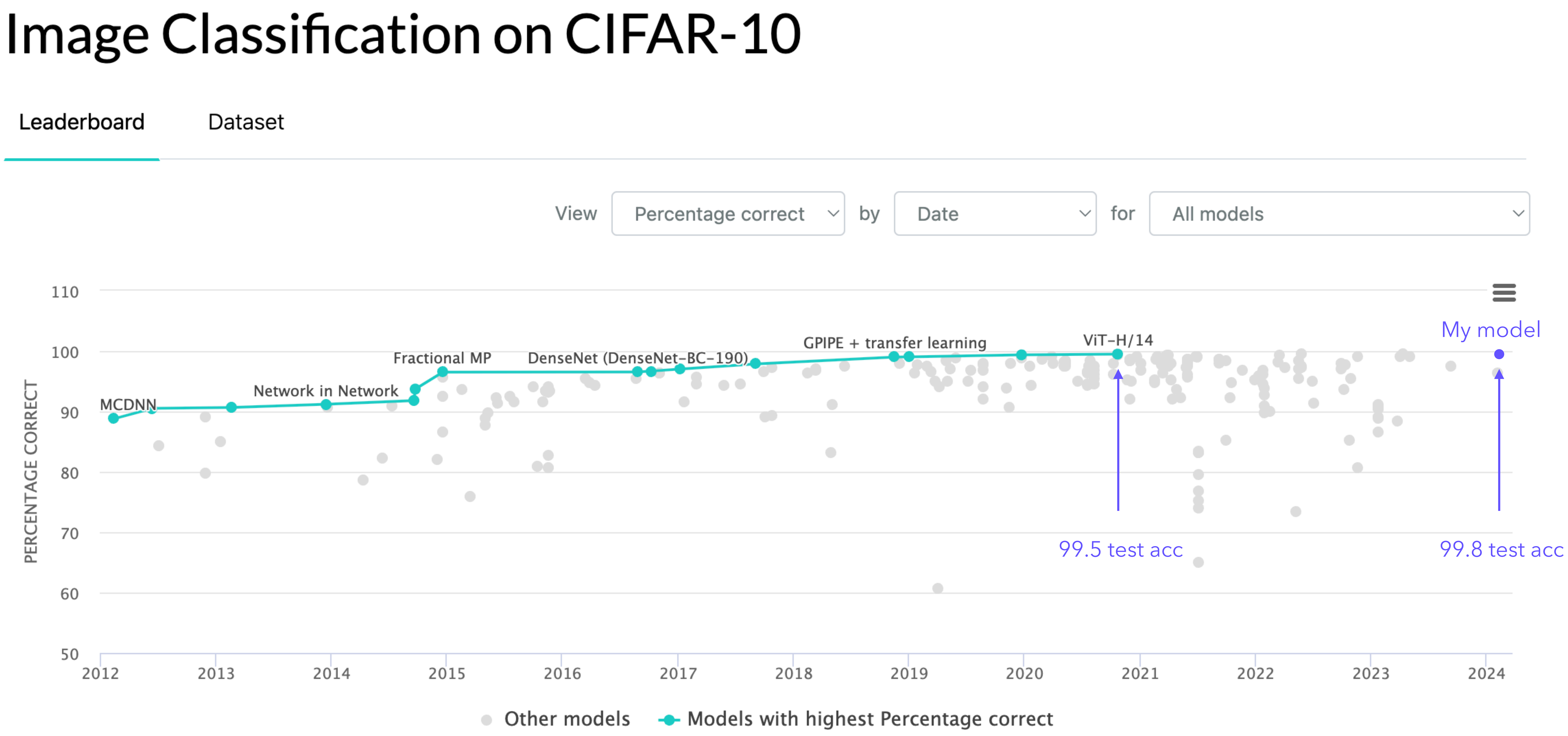
Graph stolen from Papers with code

CIFAR-10 state-of-the-art



Graph stolen from Papers with code

CIFAR-10 state-of-the-art



Graph stolen from Papers with code

Context

- You are building a fraud detection model for a financial institution
- You are building a model for new medical discovery

Context

- You are building a fraud detection model for a financial institution
- You are building a model for new medical discovery

Additional data might be expensive to acquire or even inexistent

Assume: no additional data

Assume: no additional data

- Reminder: How can you tell if the model you decide to use it’s good enough?

Assume: no additional data

- Reminder: How can you tell if the model you decide to use it's good enough?
- We don't have a way of knowing

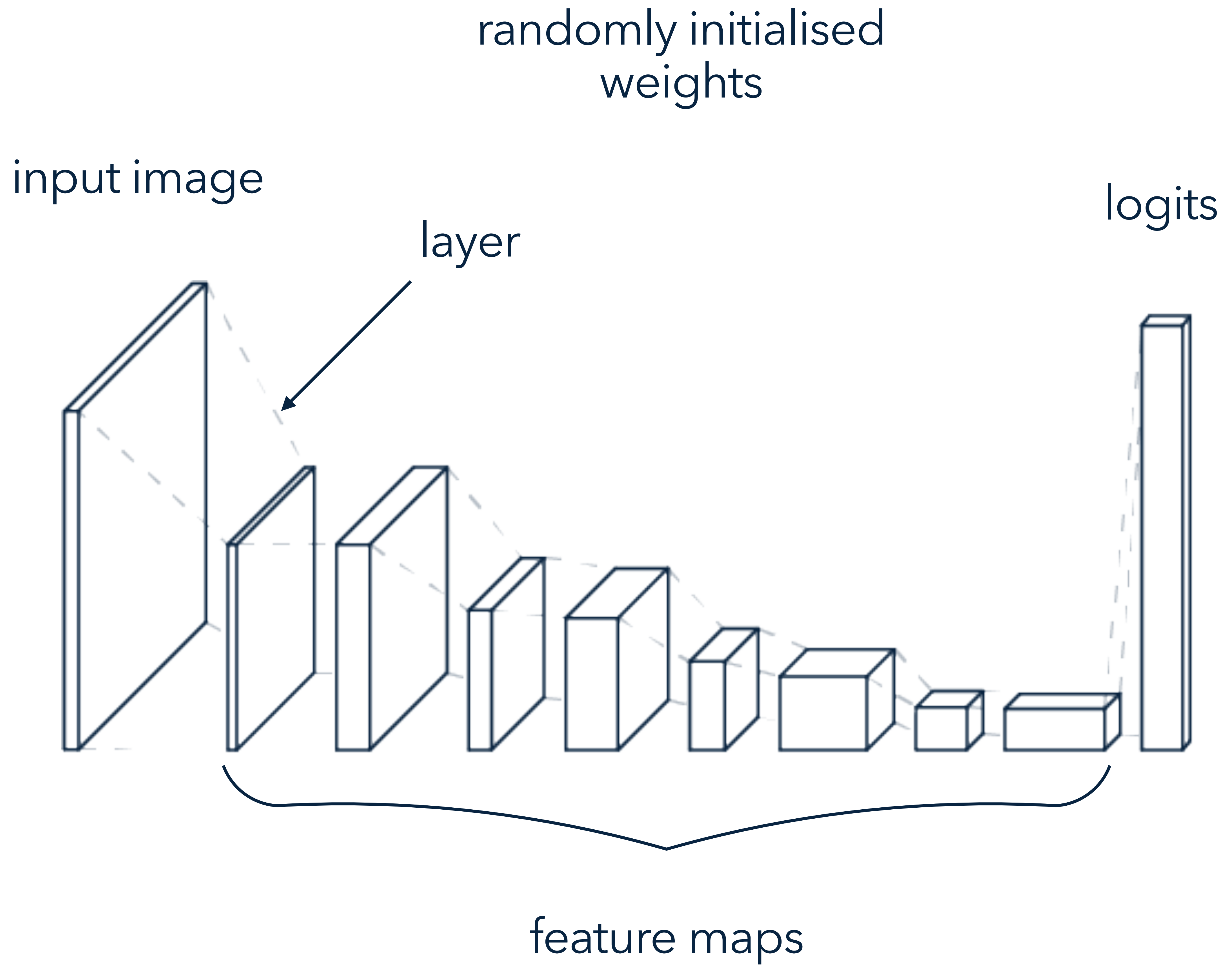
Assume: no additional data

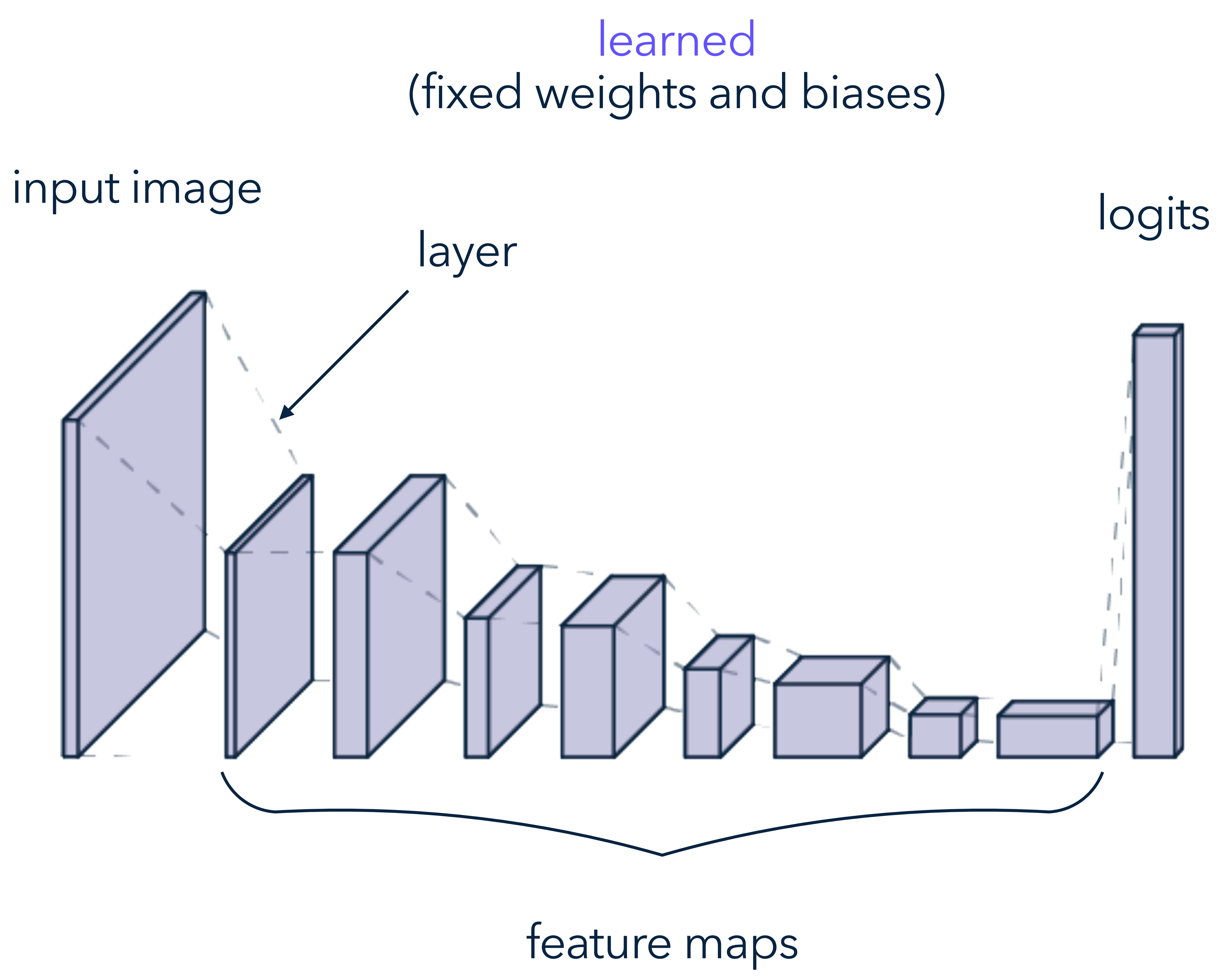
- Reminder: How can you tell if the model you decide to use it’s good enough?
- We don’t have a way of knowing but we can use different perspectives on learning to analyse models and understand what they are doing

Assume: no additional data

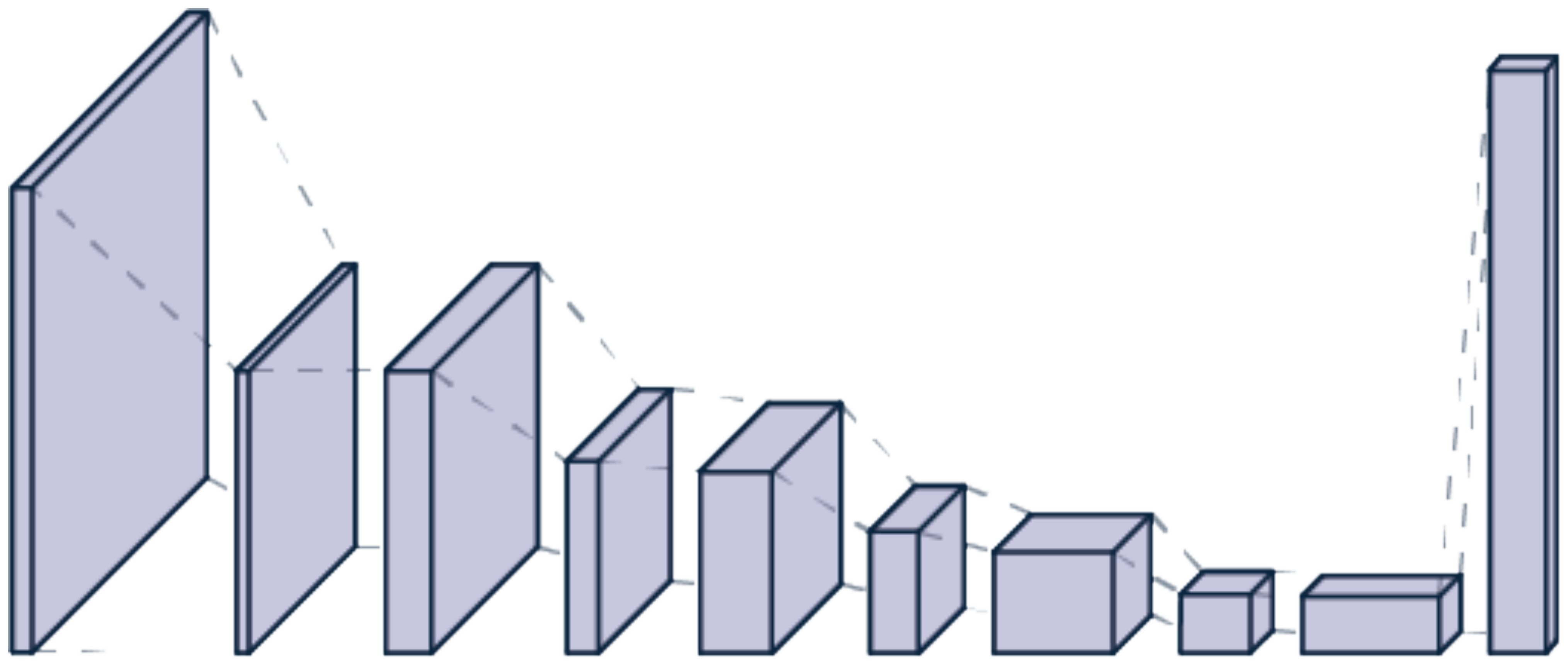
- Reminder: How can you tell if the model you decide to use it's good enough?
- We don't have a way of knowing but we can use different perspectives on learning to analyse models and understand what they are doing
- Let's first reason about what happens:
  - while training and
  - throughout the model

randomly initialised
weights


randomly initialised
weights
input image
layer
logits
feature maps

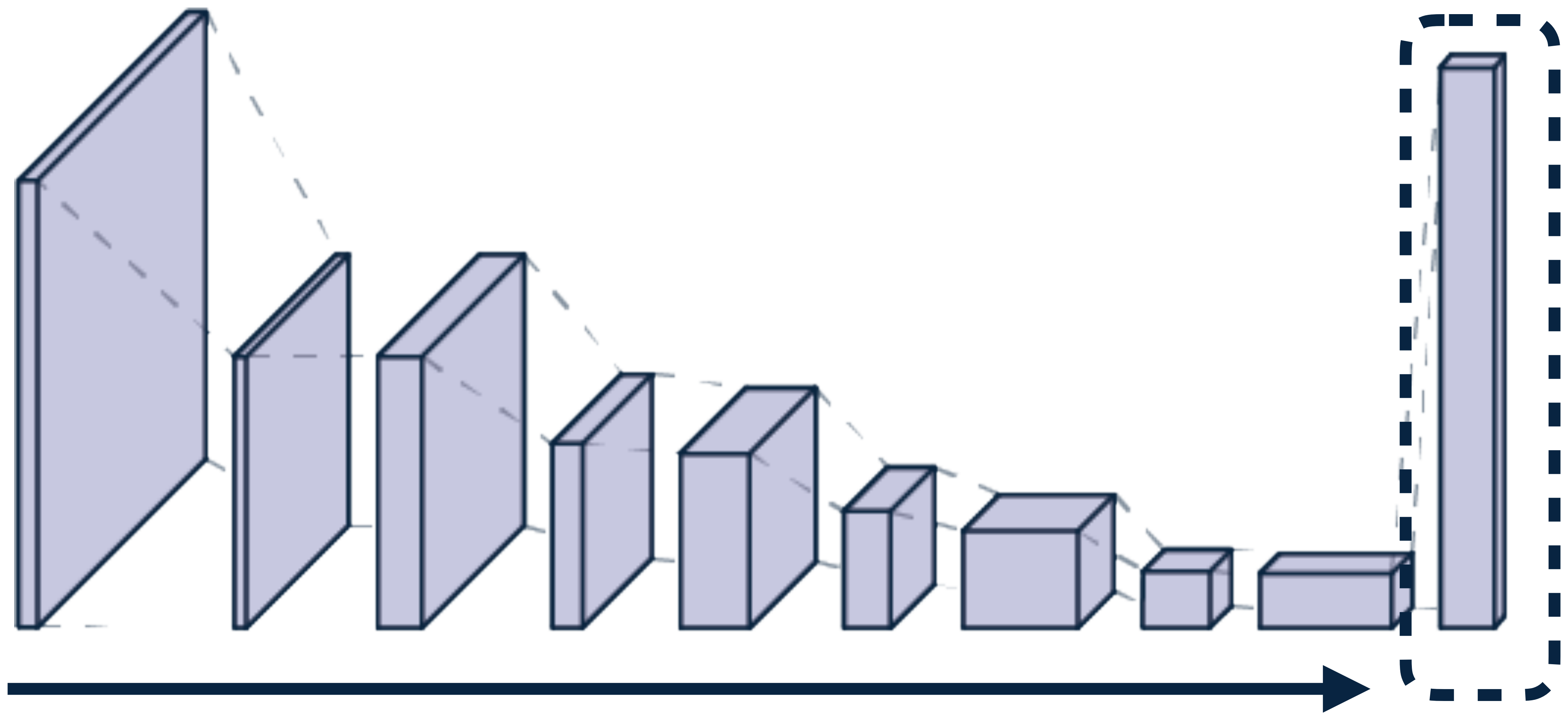

learned
(fixed weights and biases)
input image
layer
logits
feature maps

learned
(fixed weights and biases)



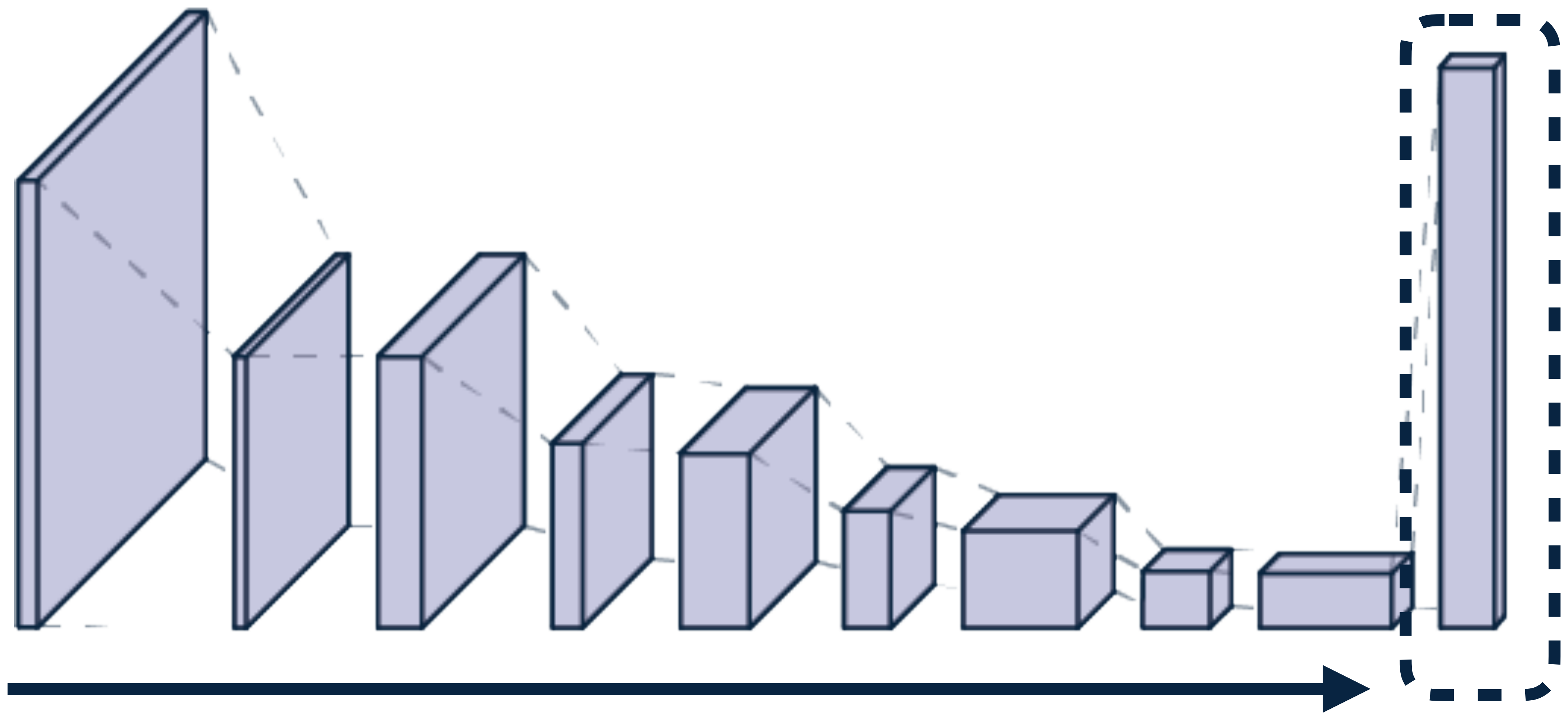
What is this model doing?

learned
(fixed weights and biases)



What is this model doing?

learned
(fixed weights and biases)

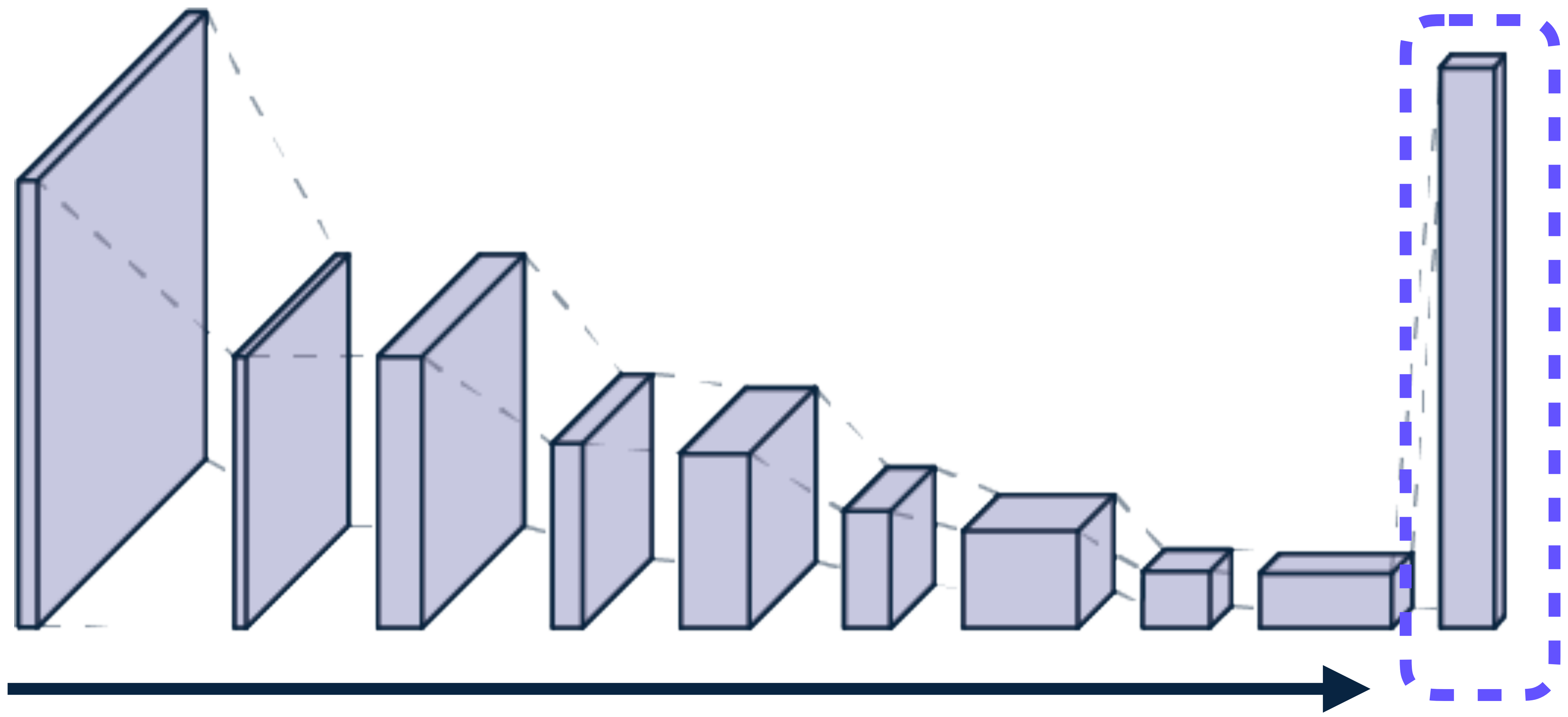


What is this model doing?

extracts and compresses information

uses info to make a decision

learned
(fixed weights and biases)



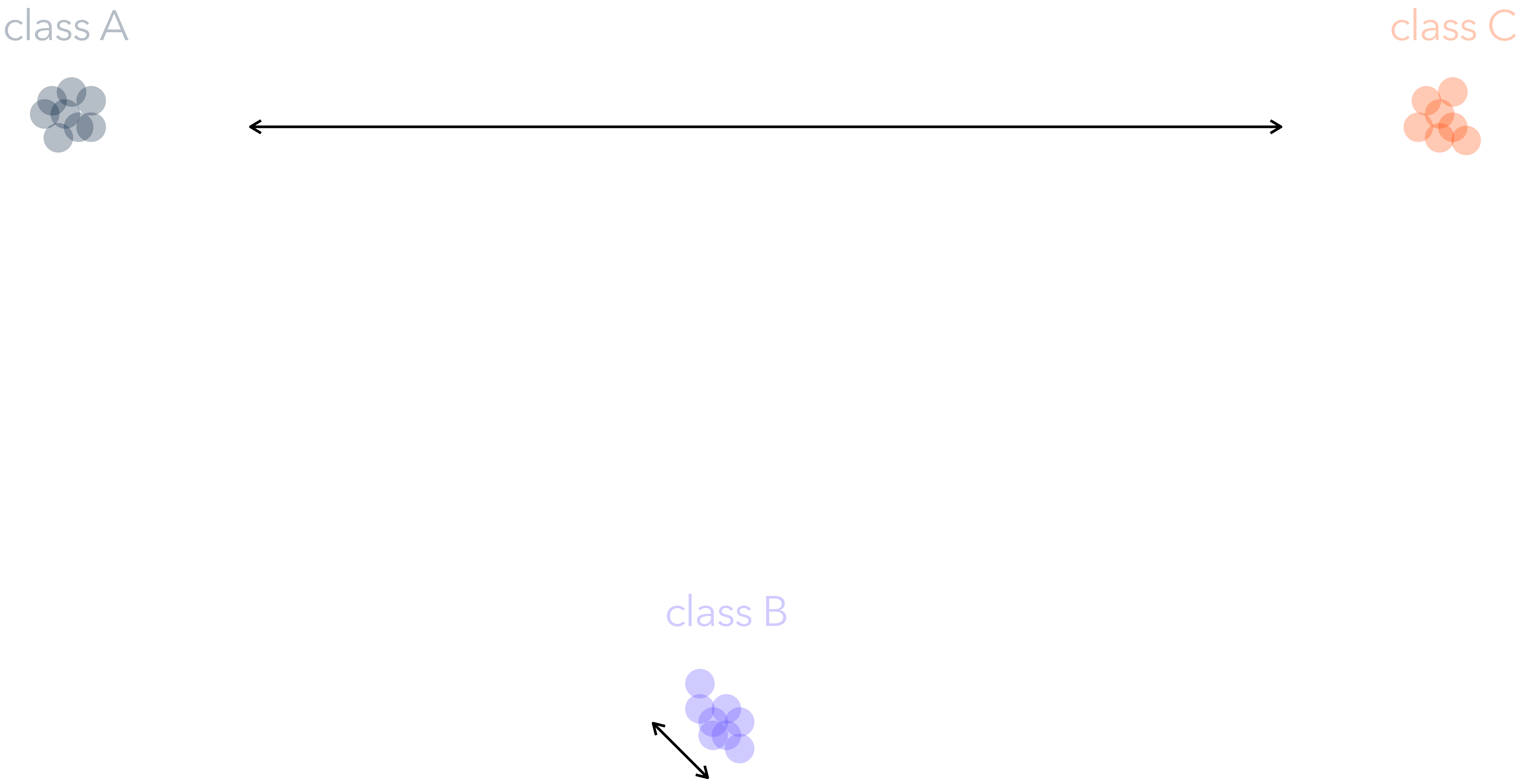
What is this model doing?

extracts and compresses information
so that samples become
linearly separable

uses info to make a decision

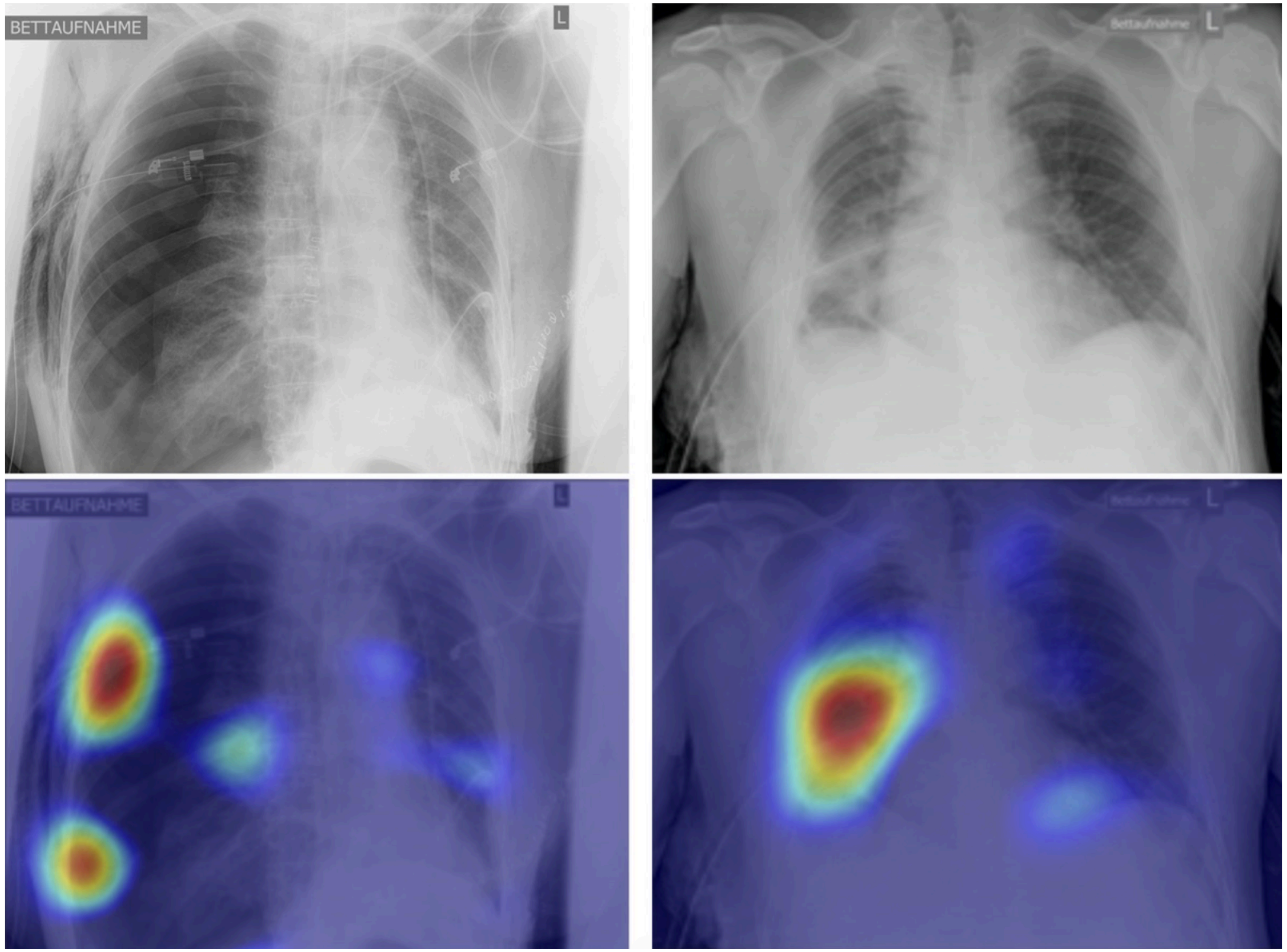
How can we achieve good separability?

(Think back to Foundations)


class A
class C
class B

What is the underlying assumption?

What can go wrong?


BETTAUFNAHME
L
Bettaufnahme
L

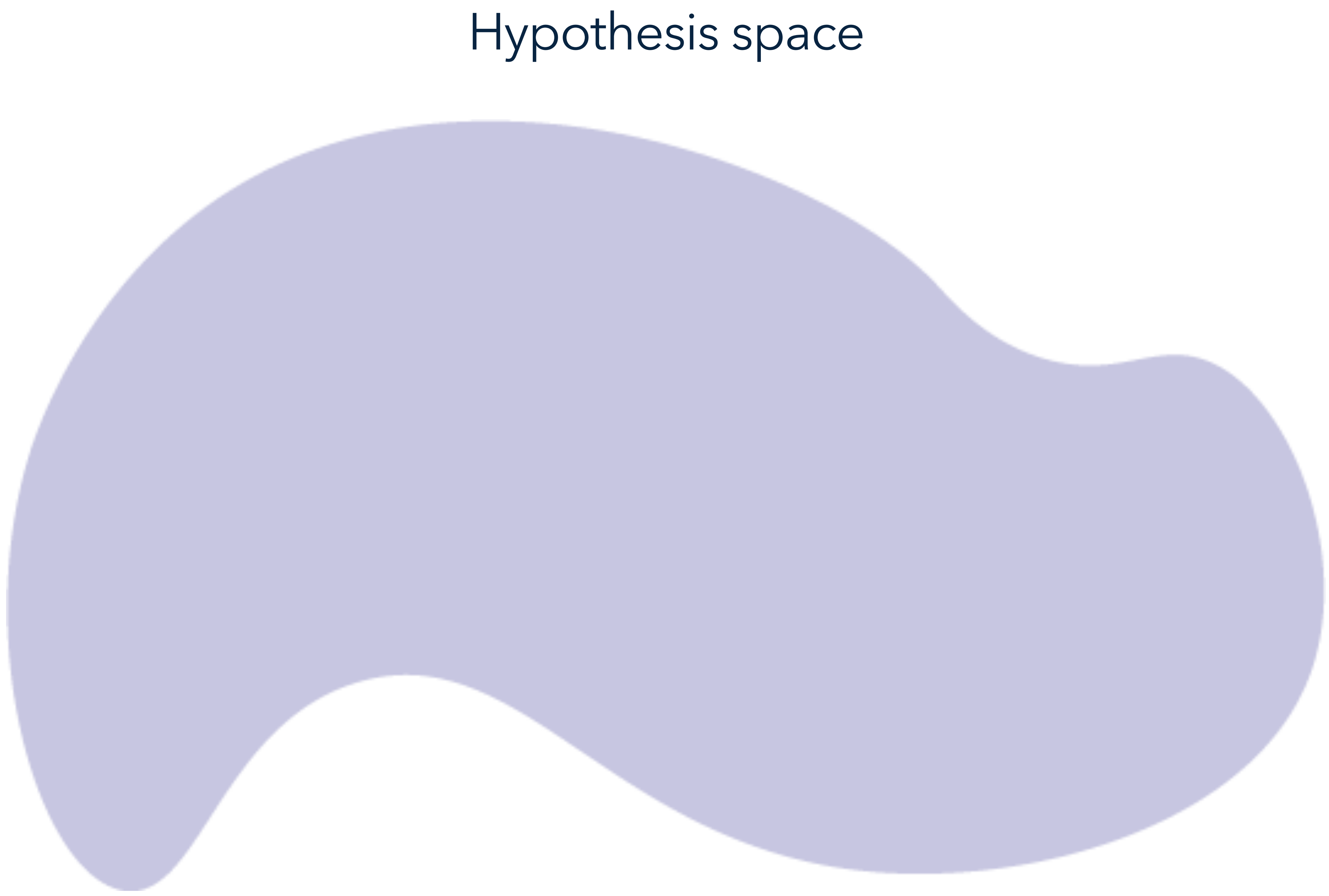
Are these anecdotes becoming DL myths?

"One pixel attack for fooling deep neural networks"

With this in mind, let’s talk about learning representations

(Discriminative and Generative)

The Hypothesis Space View



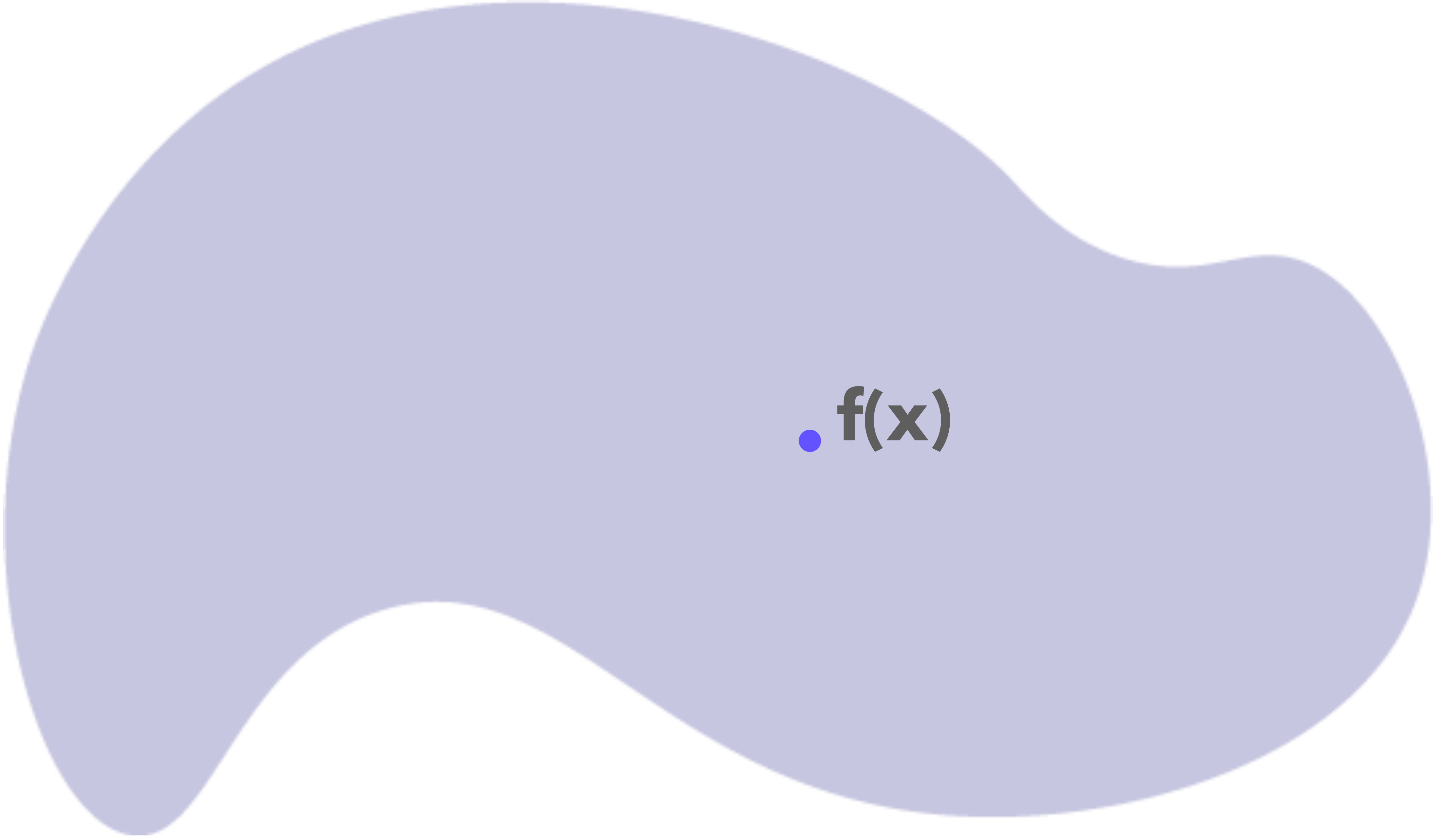
The Hypothesis Space View

Hypothesis space

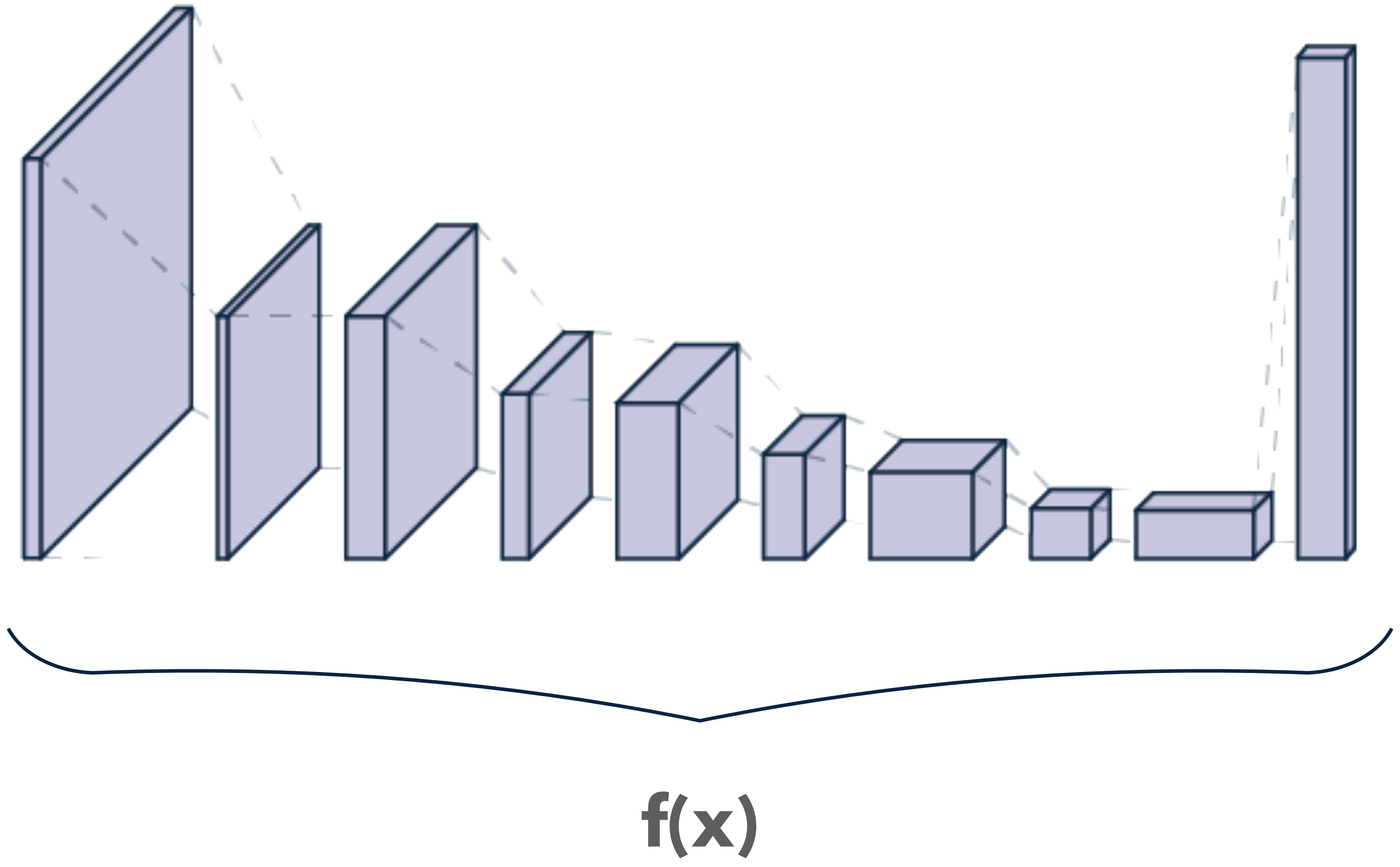
Deep Learning “equivalent”

The Hypothesis Space View

One hypothesis



Deep Learning "equivalent"



What happens during learning?

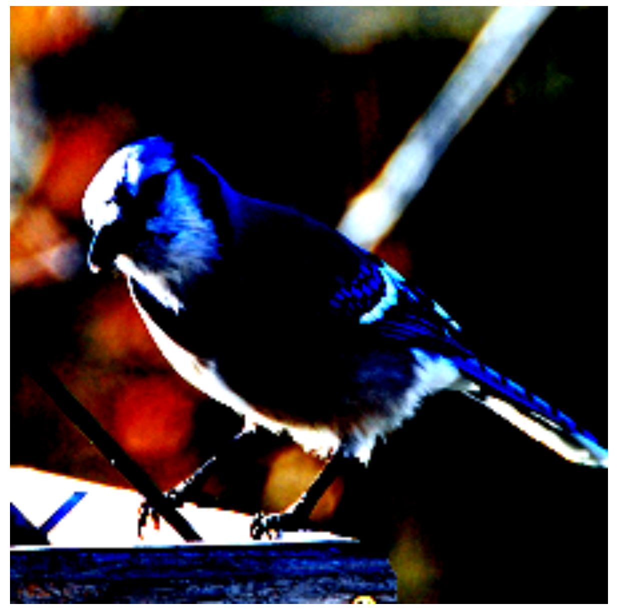
The Hypothesis Space View

Hypothesis Space

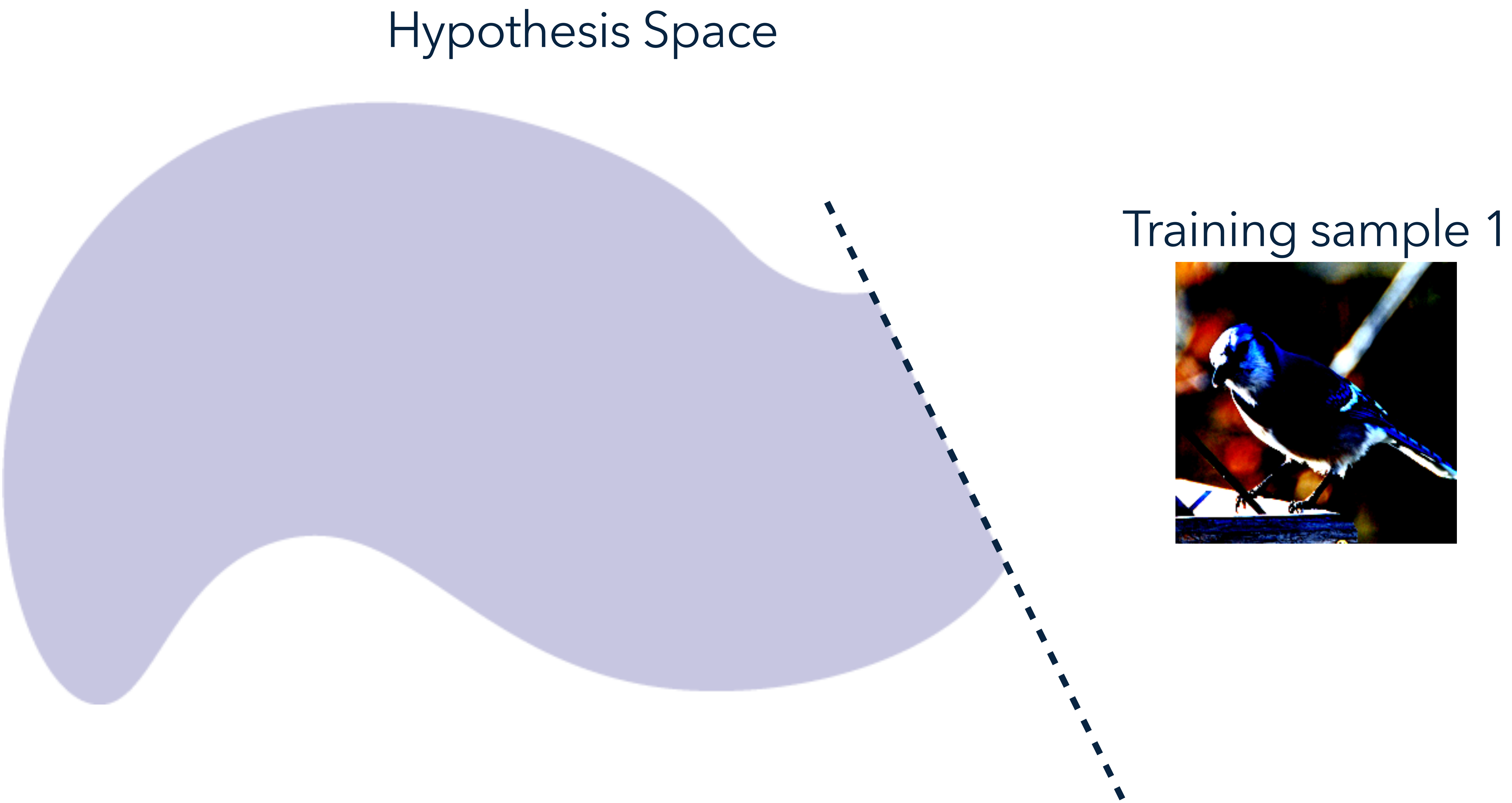
The Hypothesis Space View

Hypothesis Space

Training sample 1



The Hypothesis Space View



The Hypothesis Space View

Hypothesis Space

Training sample 2

The Hypothesis Space View

Hypothesis Space

Training sample 3

The Hypothesis Space View

Hypothesis Space

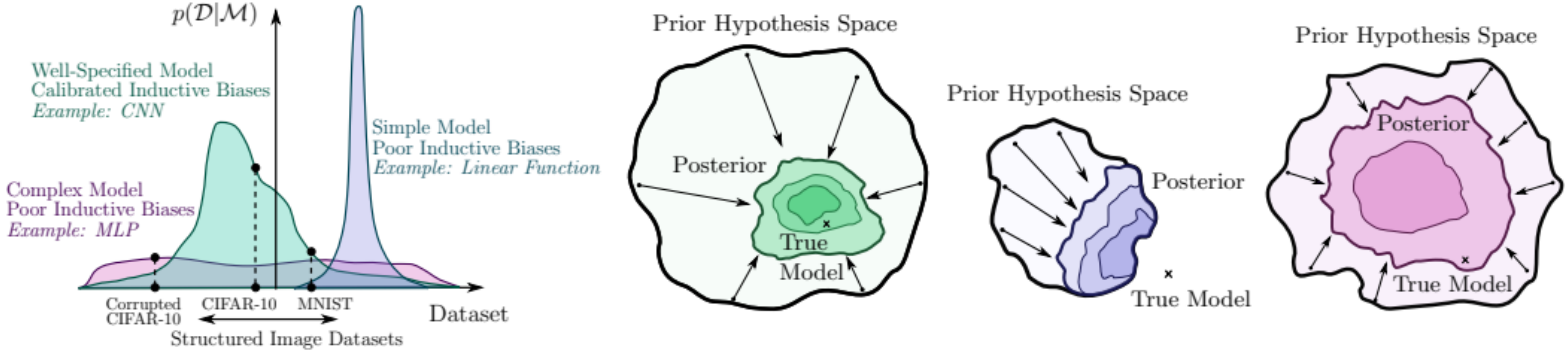
...

The Hypothesis Space View

Are we uniformly sampling from this space?

The Hypothesis Space View

Are we uniformly sampling from this space?



“Bayesian Deep Learning and a Probabilistic Perspective of Generalisation”

The Hypothesis Space View

- Captures only what happens throughout learning, not throughout network as well

The Hypothesis Space View

- Captures only what happens throughout learning, not throughout network as well
- Particularly useful for Bayesian Deep Nets

The Manifold View

- **Assumption:** "Data lies on a low-dimensional manifold"

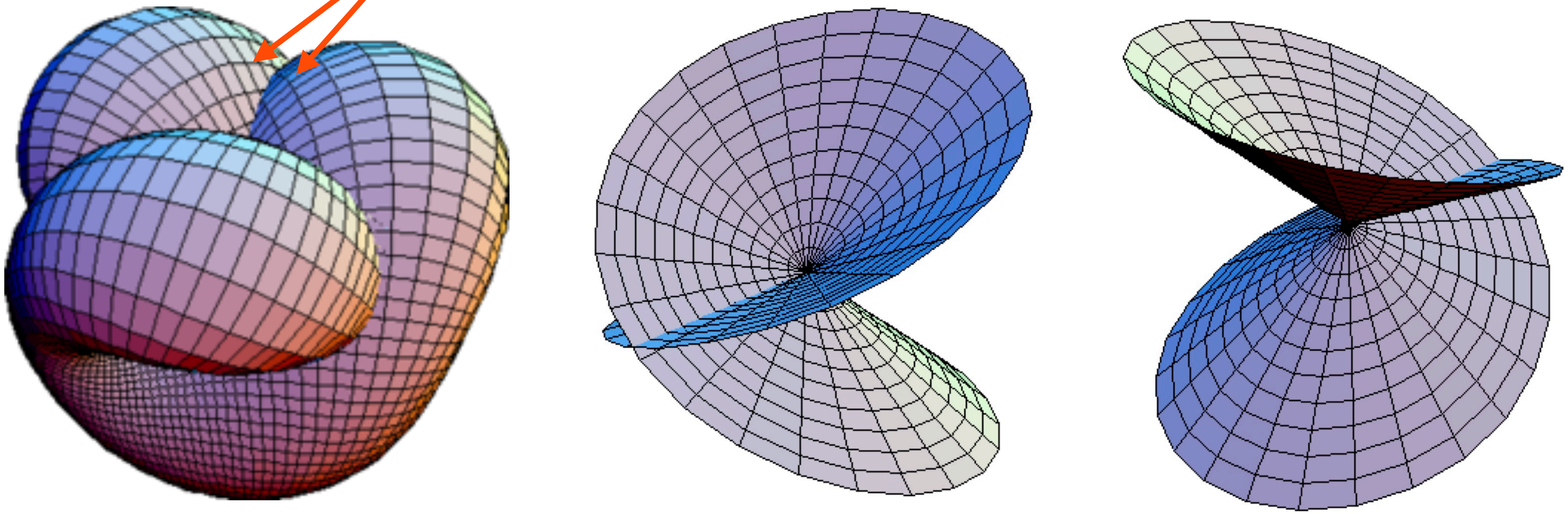
The Manifold View

- **Assumption:** "Data lies on a low-dimensional manifold"

The Manifold View

- **Assumption:** "Data lies on a low-dimensional manifold"

Close in space but far away on the manifold



Stolen from Wikipedia

The Manifold View

In Deep Learning we don’t actually work with the mathematical object

- **Assumption:** “Data lies on a low-dimensional manifold”

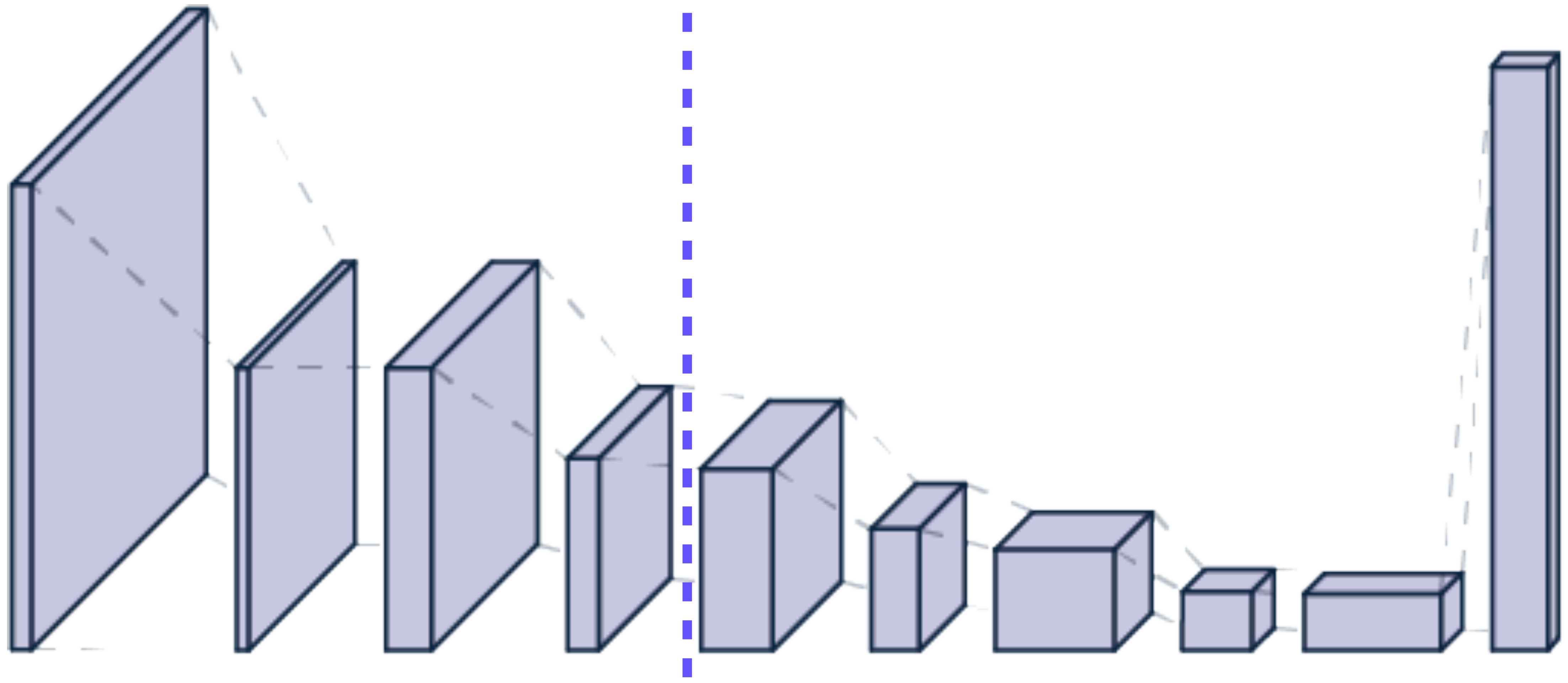
The Manifold View

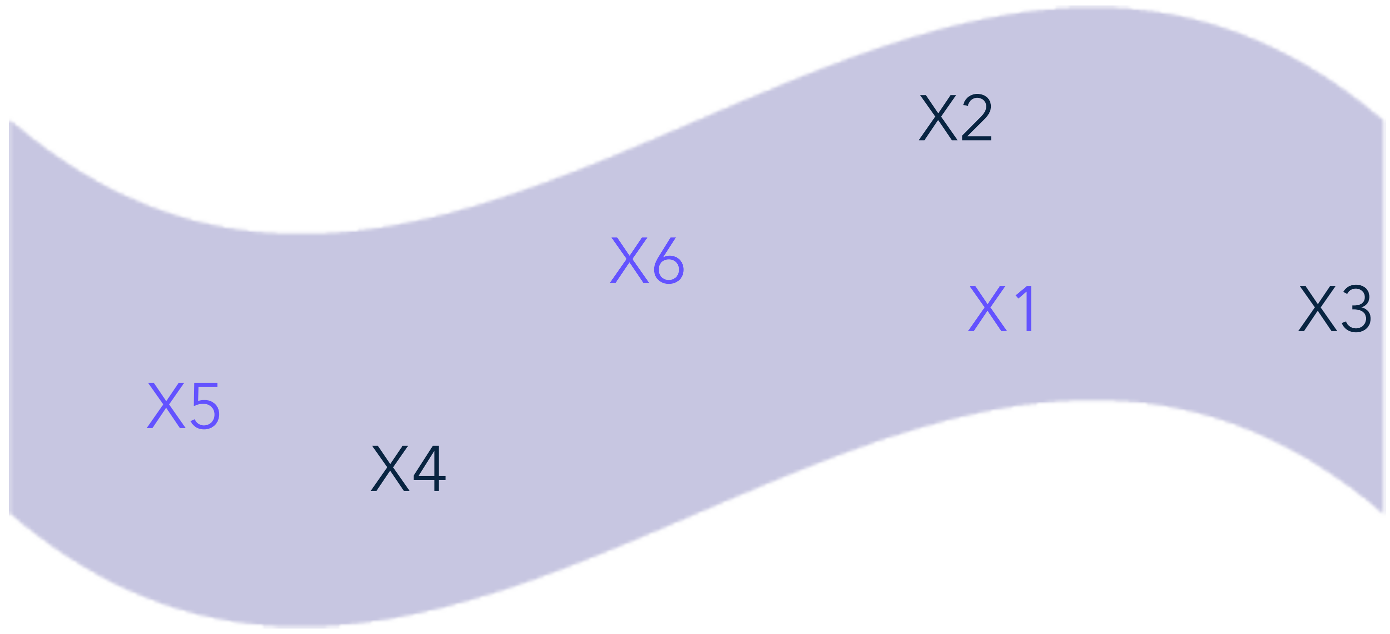
- **Assumption:** “Data lies on a low-dimensional manifold”

Let’s massively simplify things

The Manifold View

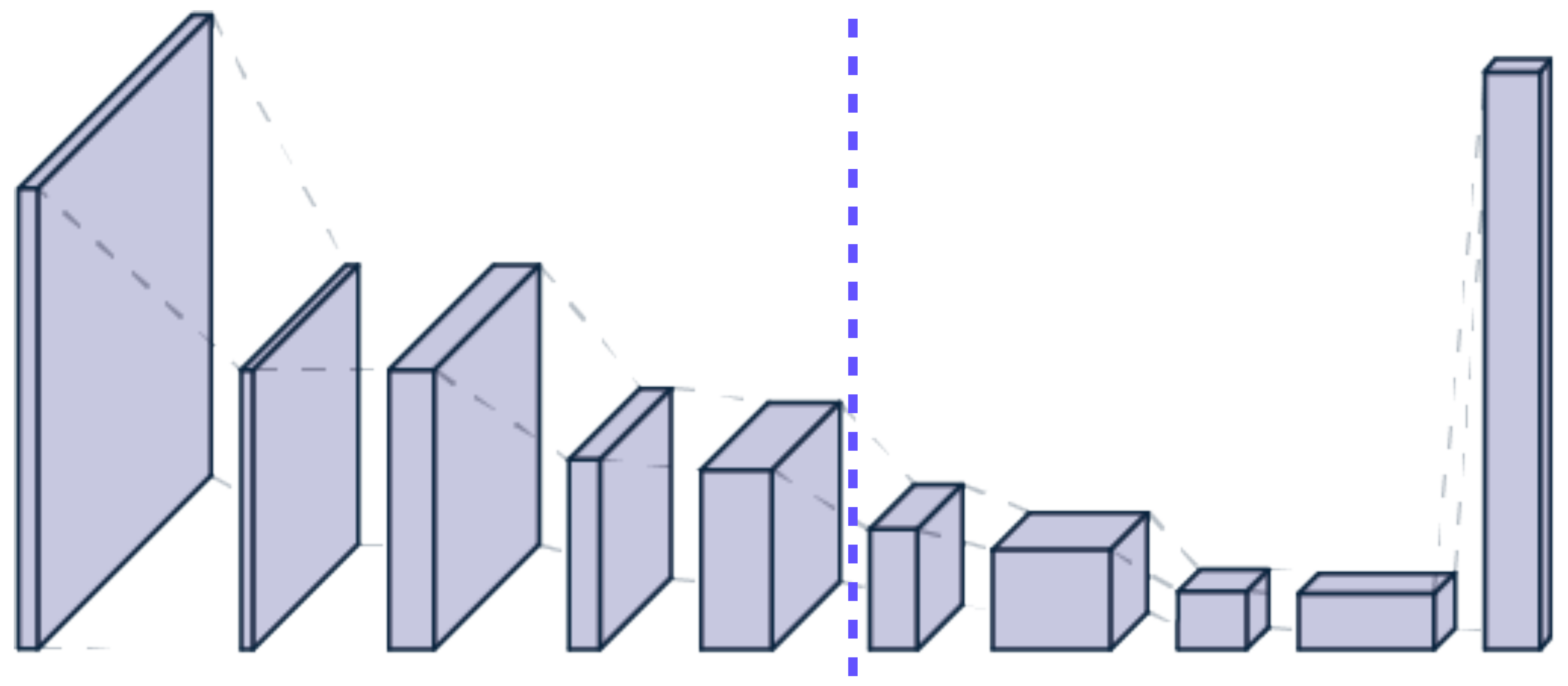
- **Assumption:** “Data lies on a low-dimensional manifold”

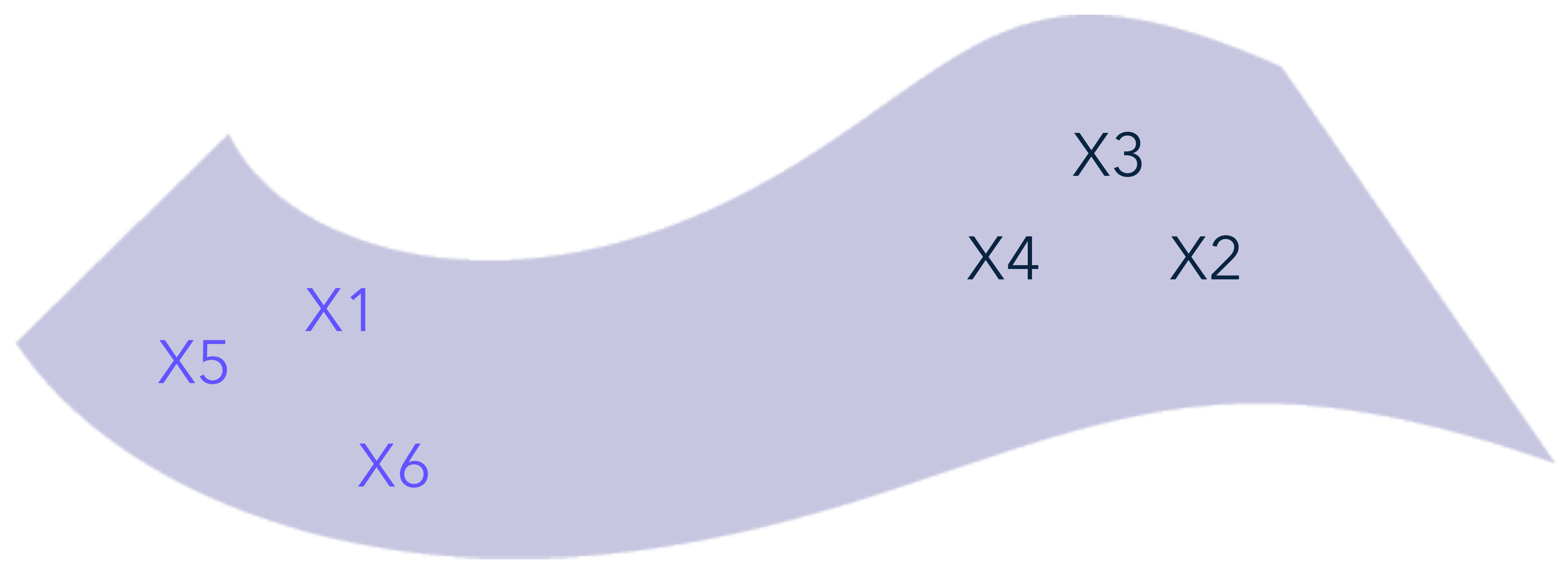




The Manifold View

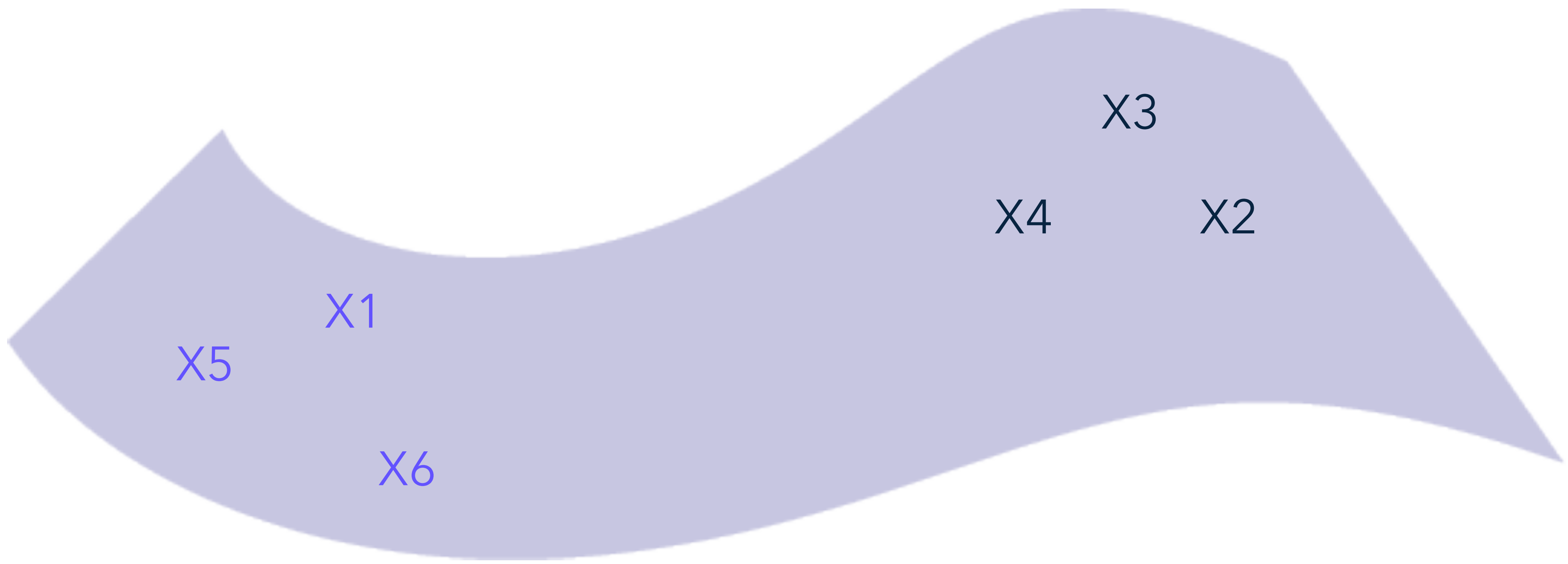
- **Assumption:** "Data lies on a low-dimensional manifold"



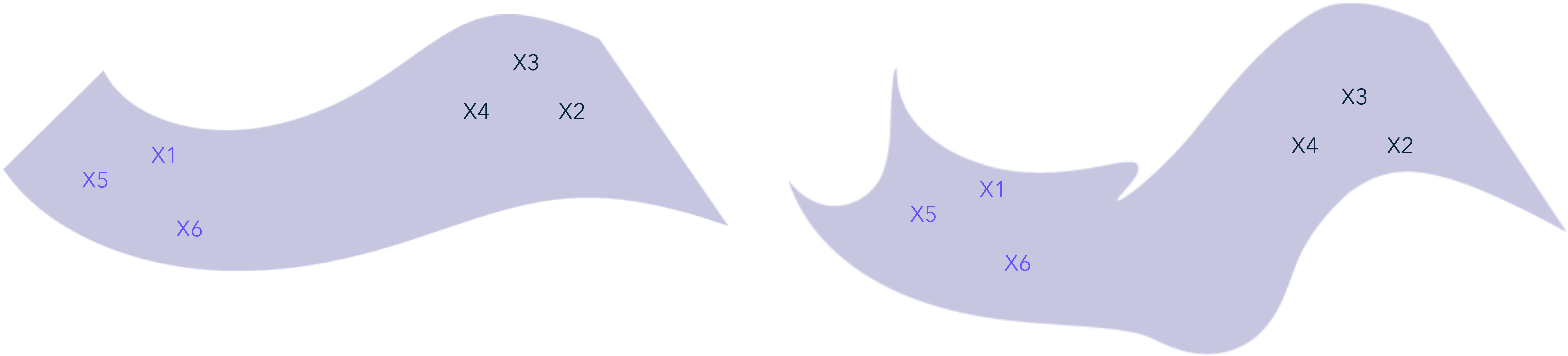


What’s the biggest challenge?

The Manifold View

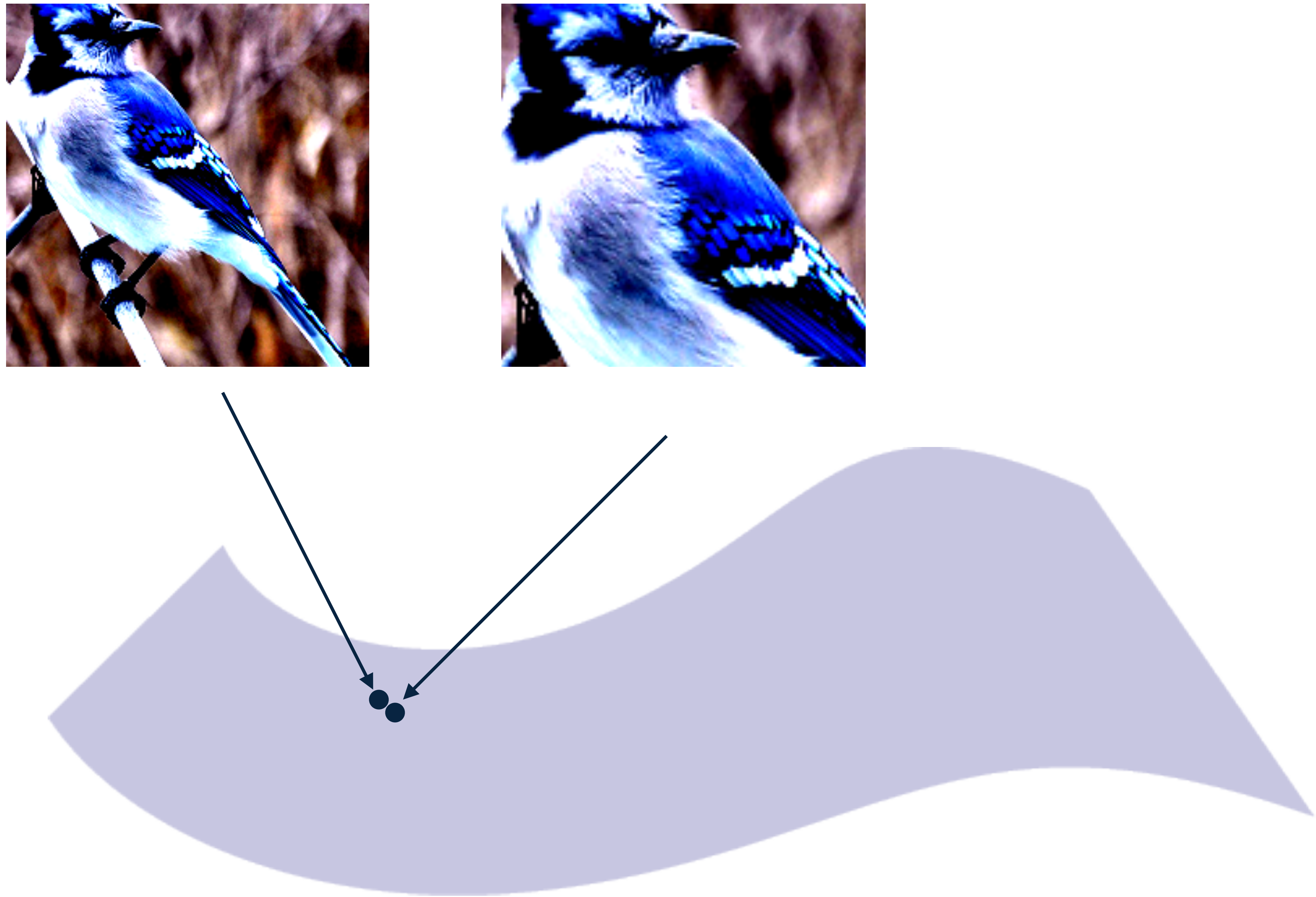


The Manifold View



What are classical augmentations trying to do?

The Manifold View



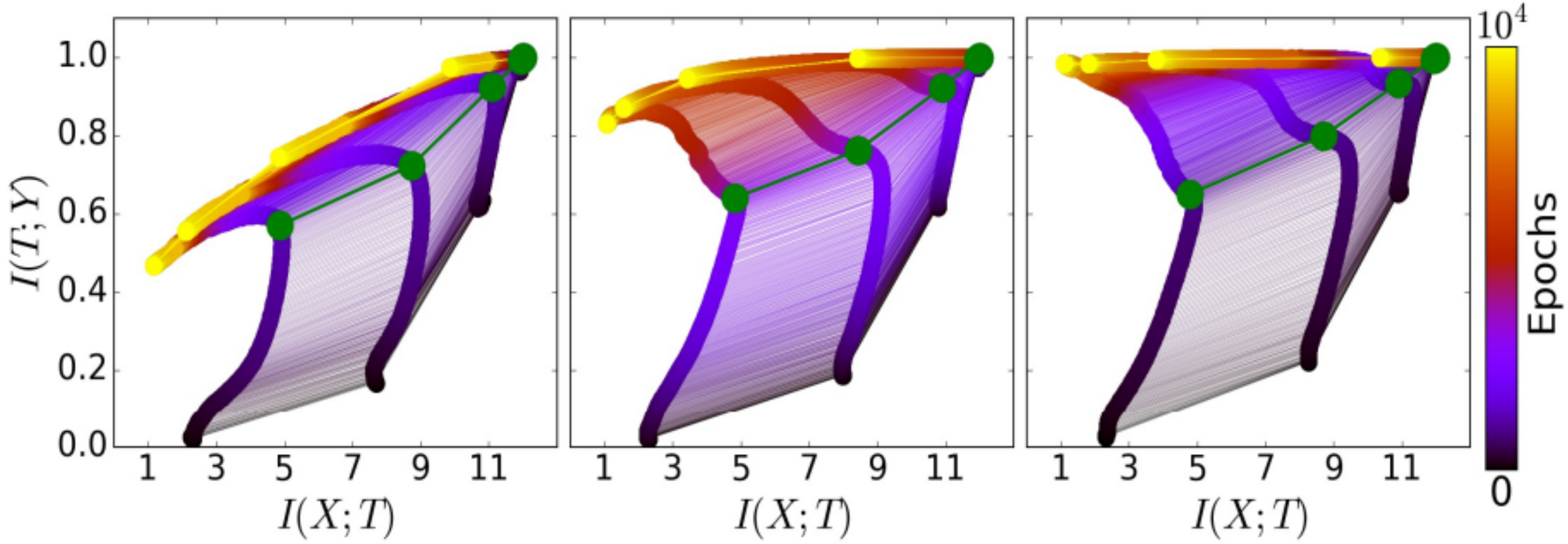
What about augmentations like Mixup?

What about augmentations like Mixup?

(The hypothesis perspective might be more useful here)

The Information View

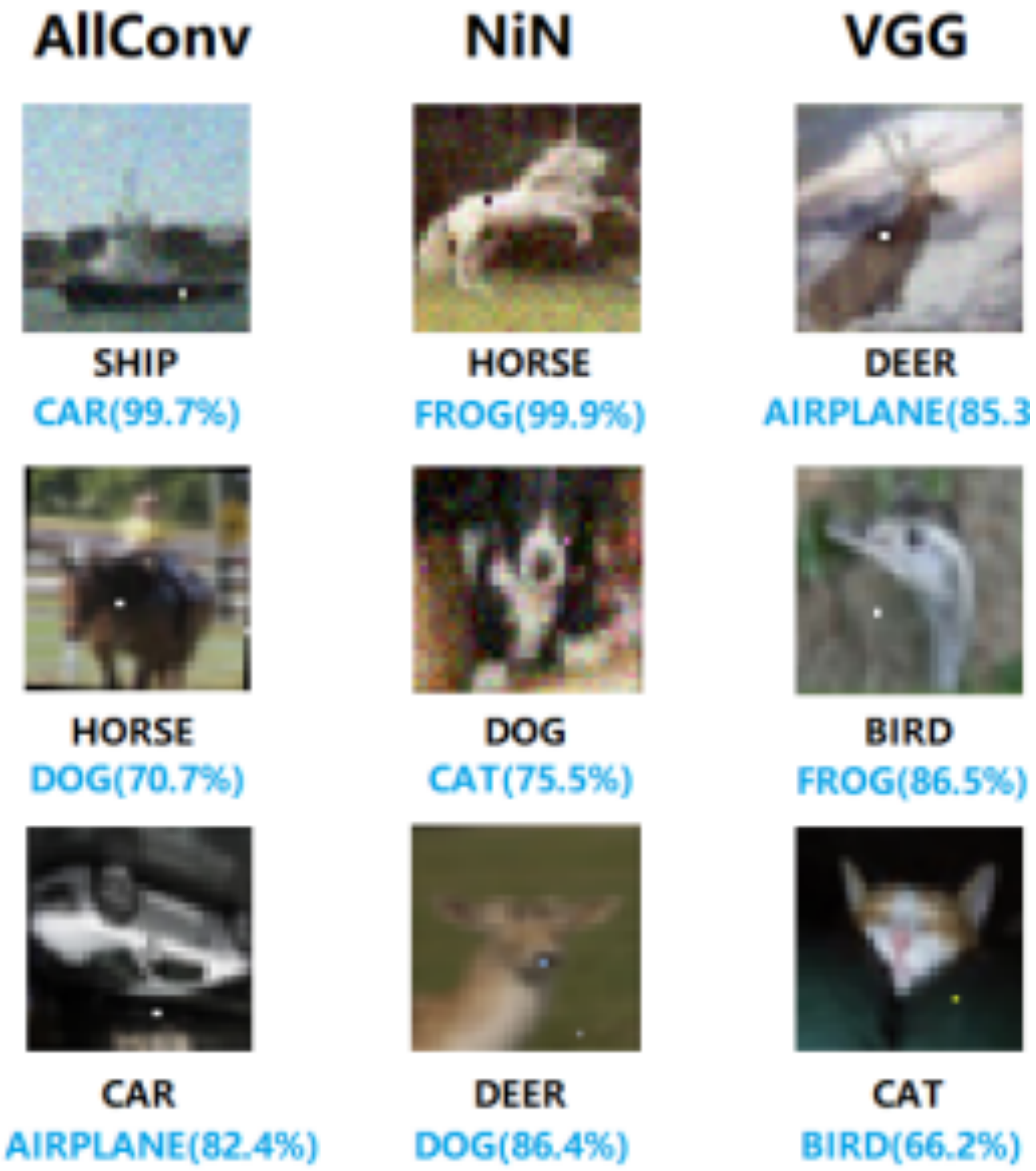
The Information Bottleneck Theory



"Opening the black box of Deep Neural Networks via Information"

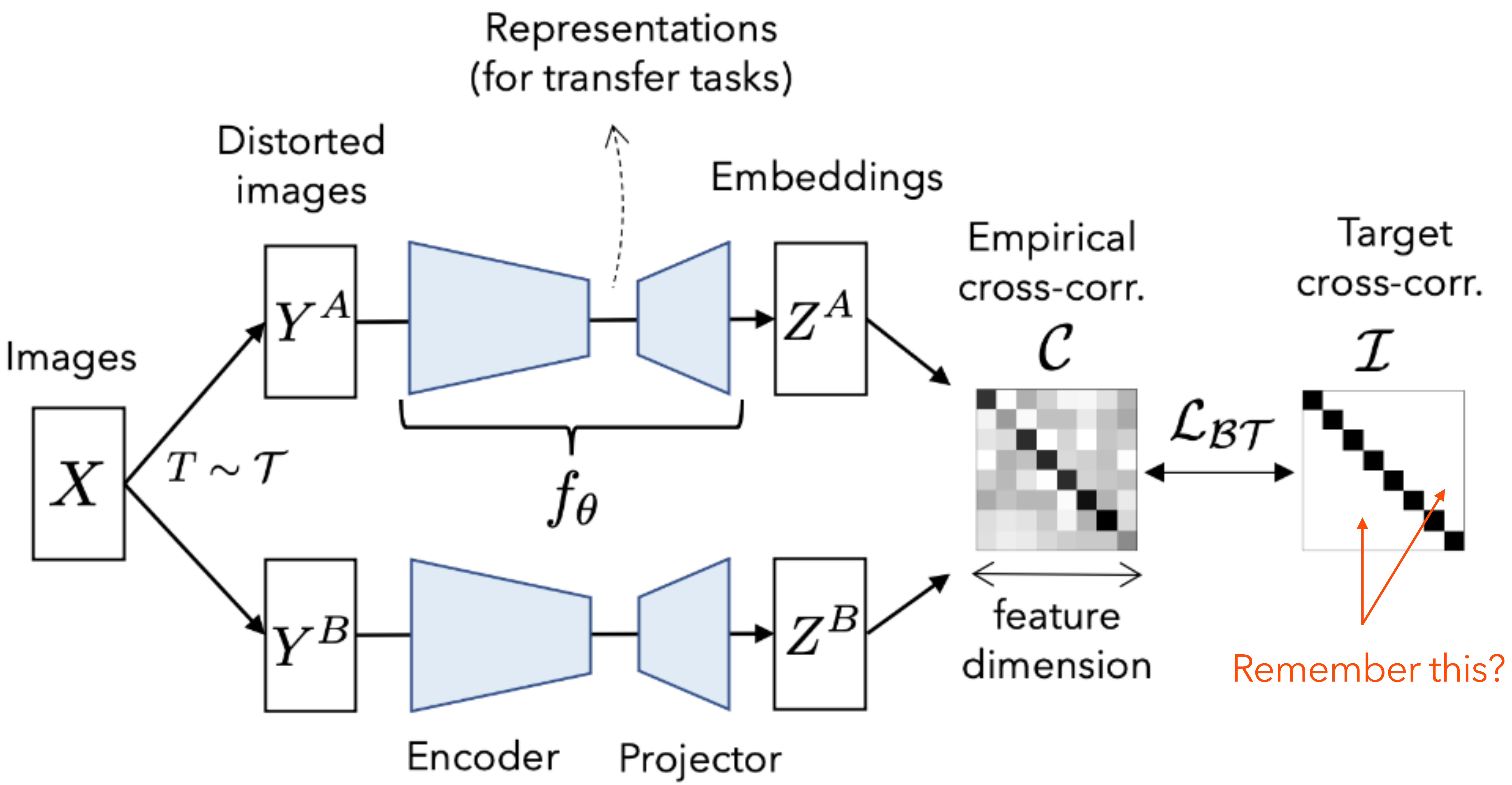
The Information Bottleneck View

- $I(X, Y) = H(X) - H(X|Y)$
- “Minimal sufficient statistics” perspective
- What’s the assumption and what can go wrong?



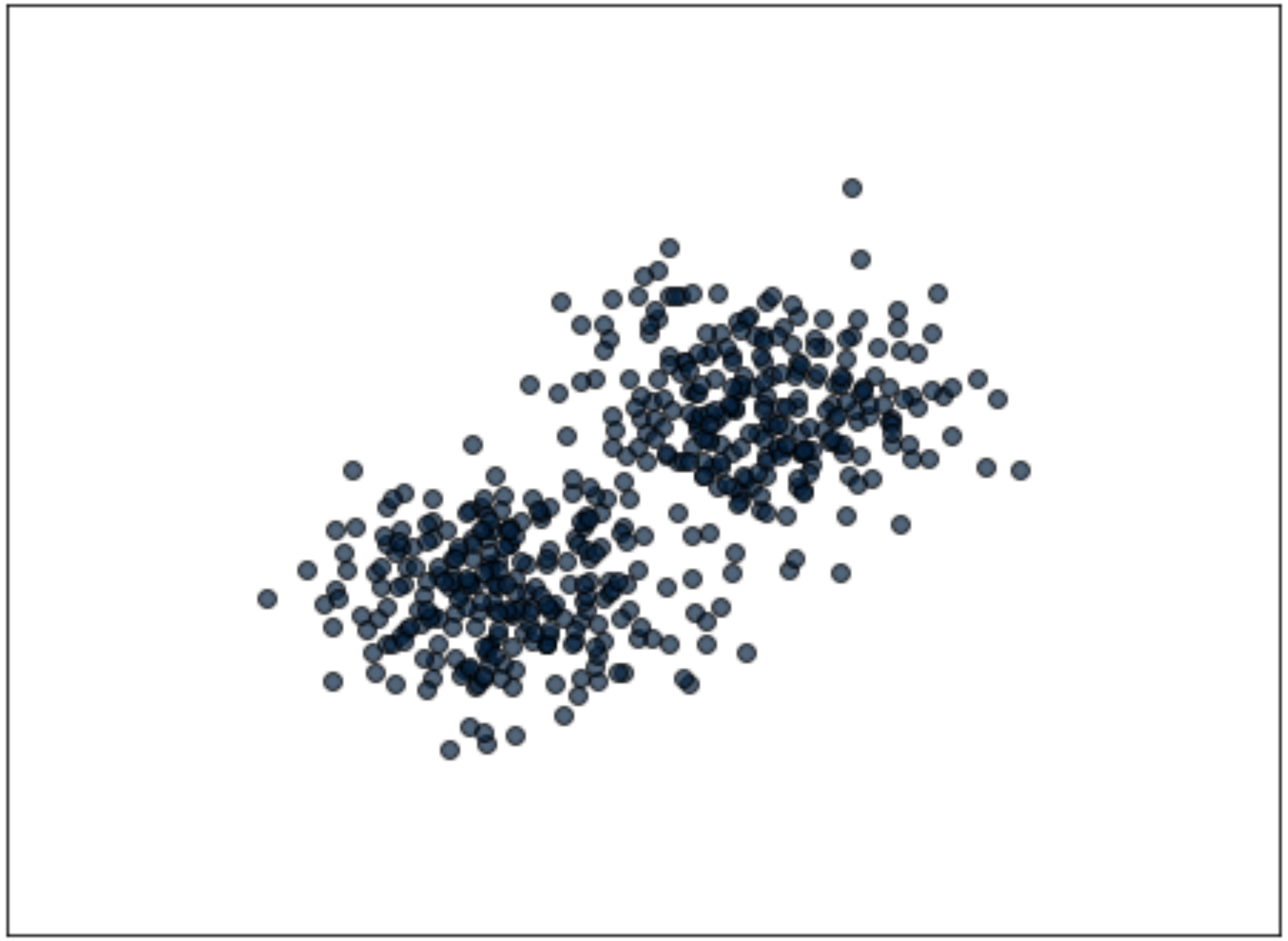
"One pixel attack for fooling deep neural networks"

Back to Barlow Twins



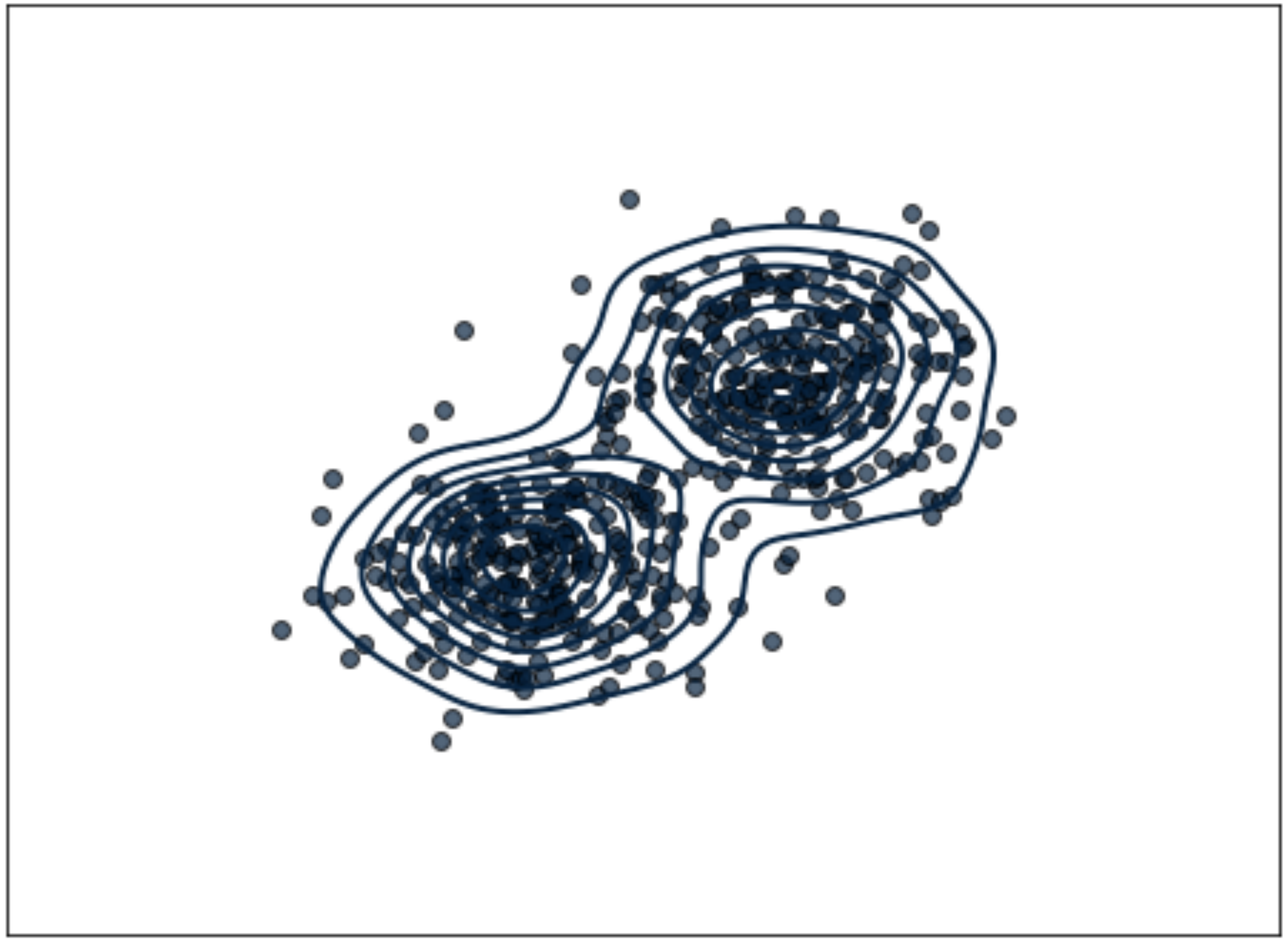
The Distribution Modelling View

The Distribution Modelling View

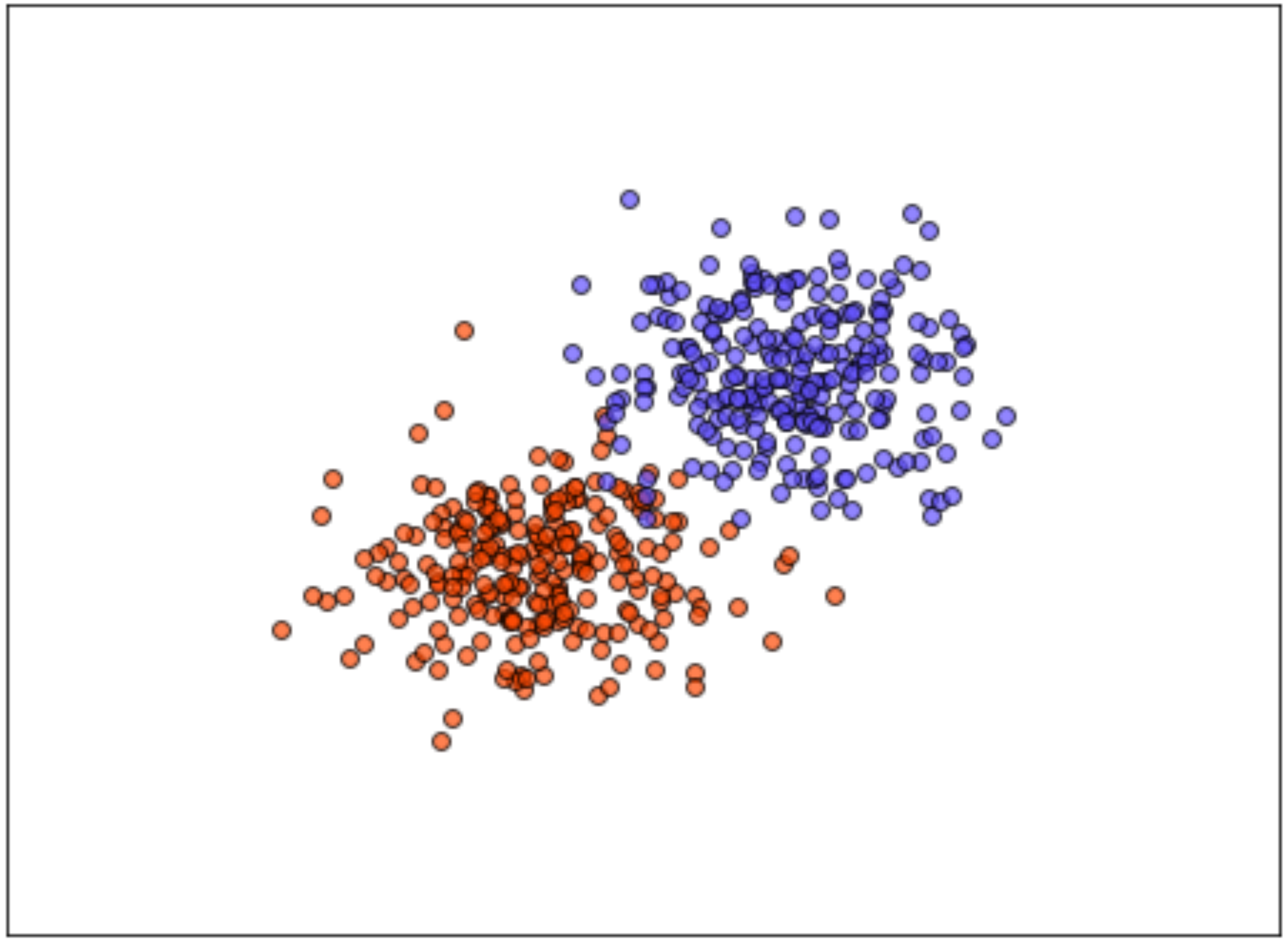


The Distribution Modelling View

P(X)

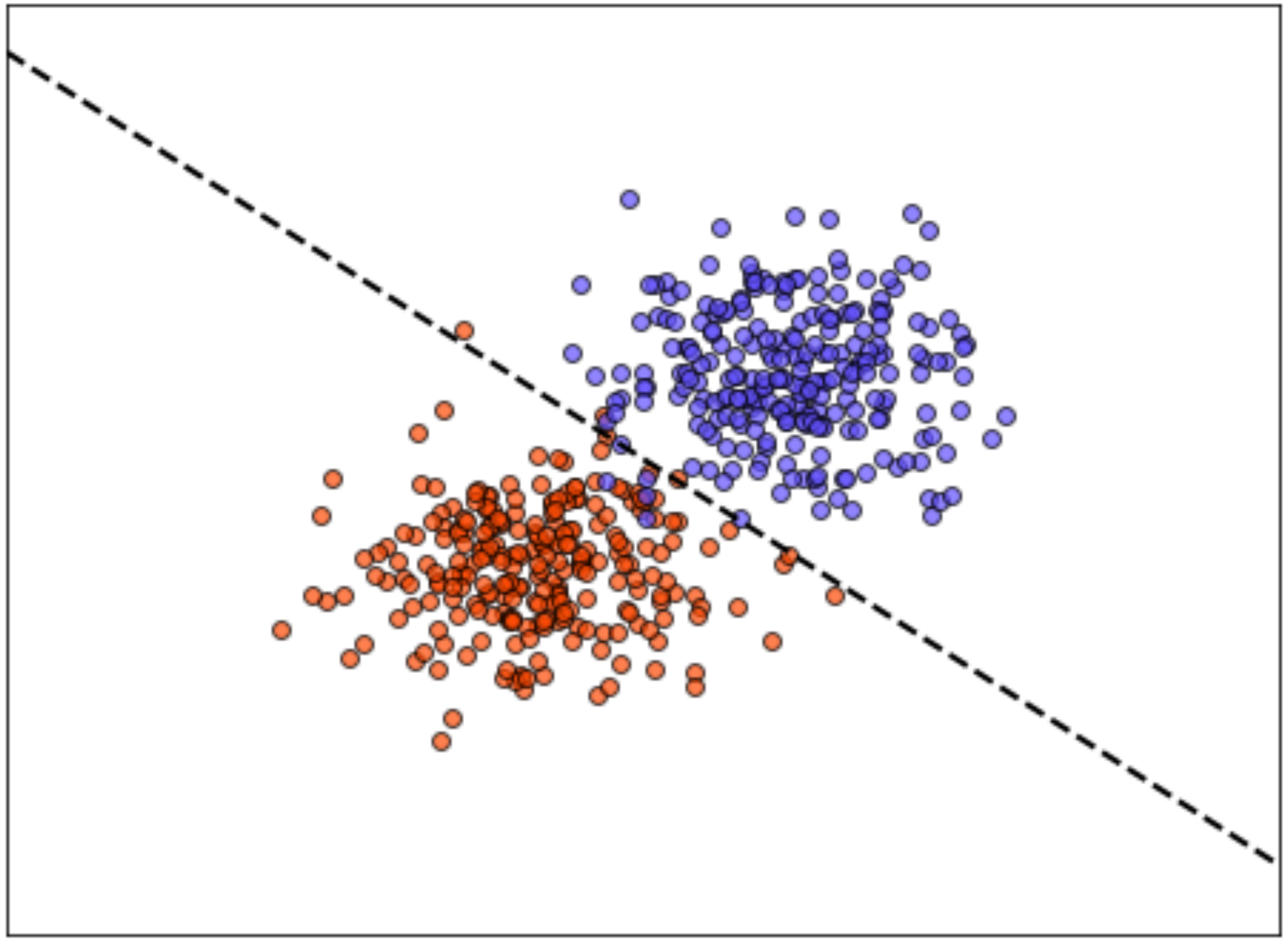


The Distribution Modelling View



The Distribution Modelling View

P(Y|X)



Take-aways

- To better analyse and evaluate models we first need to reflect on what happens throughout the model and during learning
- A number of different (sometimes complementary) perspectives exist

Take-aways

- To better analyse and evaluate models we first need to reflect on what happens throughout the model and during learning
- A number of different (sometimes complementary) perspectives exist each with its own strengths