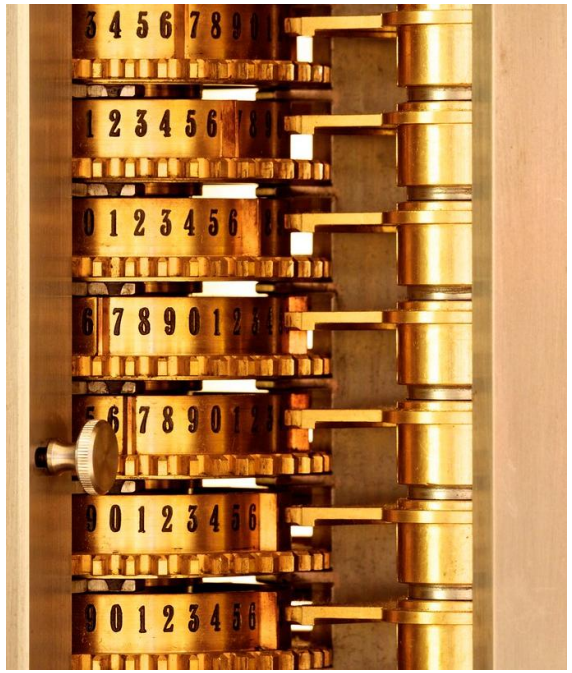
Computers, brains and AI

5: Deep learning





Difference engine by Marcin Wichary / CC BY
Ada Lovelace by Science Museum Group / CC0

1) Intro

2) Foundations
3) Silicon

4) Deep learning — Brains and neural networks
5) Deep learning — Learning

6) ChatGPT
7) ChatGPT
8) ChatGPT

9) The future of AI
10) The future of AI



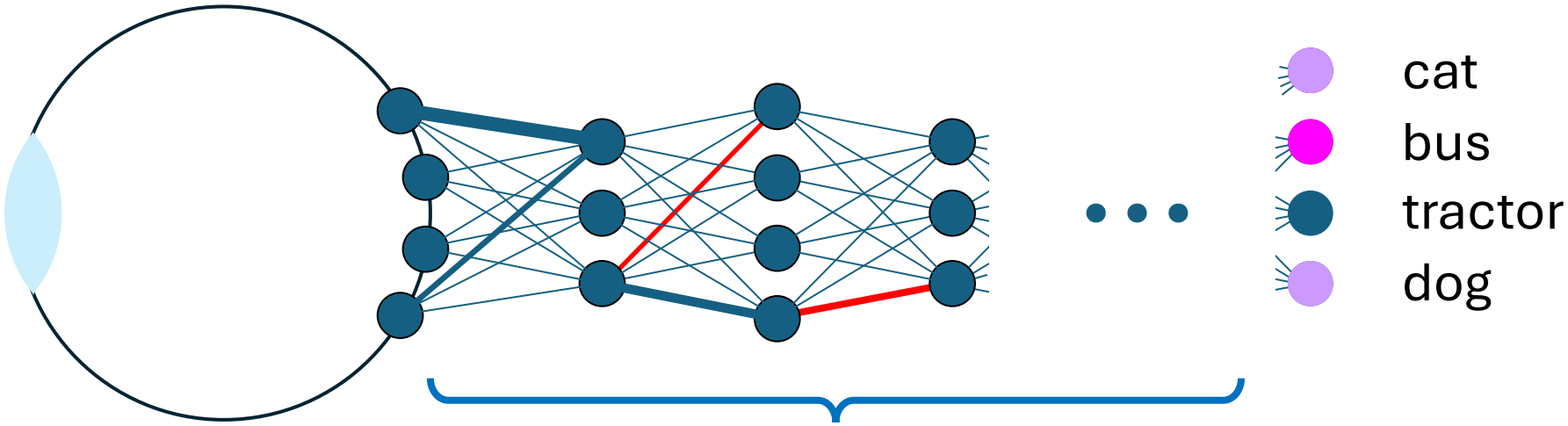
AlexNet

What AlexNet learns









Learning means choosing these numbers



DJ equipment by Paulo Guereta / CC BY 2.0



Training









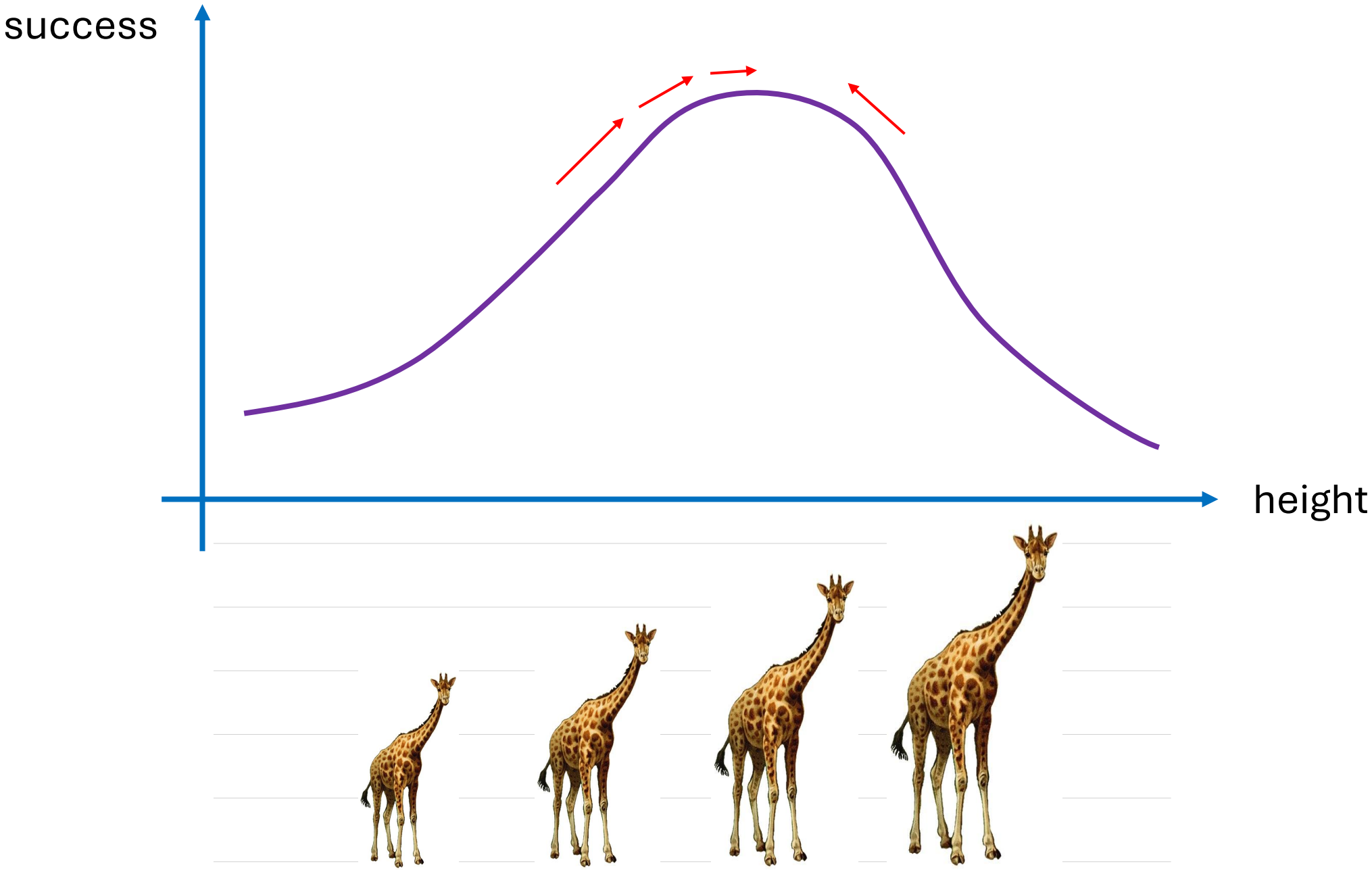
dog

bus

cat

Deep learning
vs
Evolution




success
height

badness ?

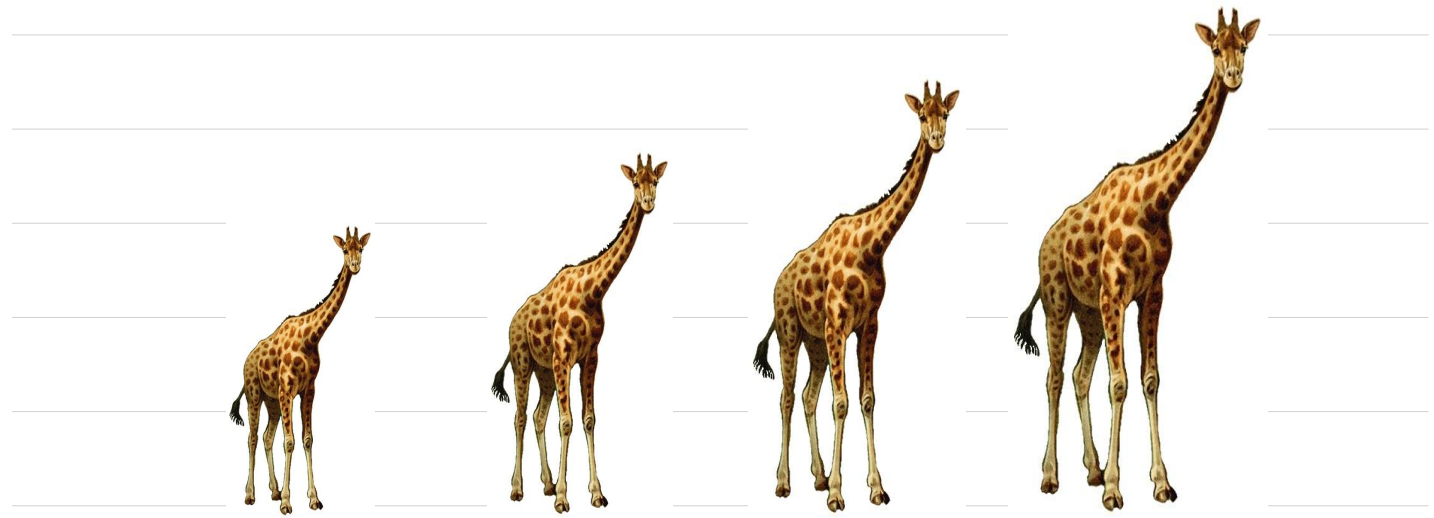
cost ?

loss

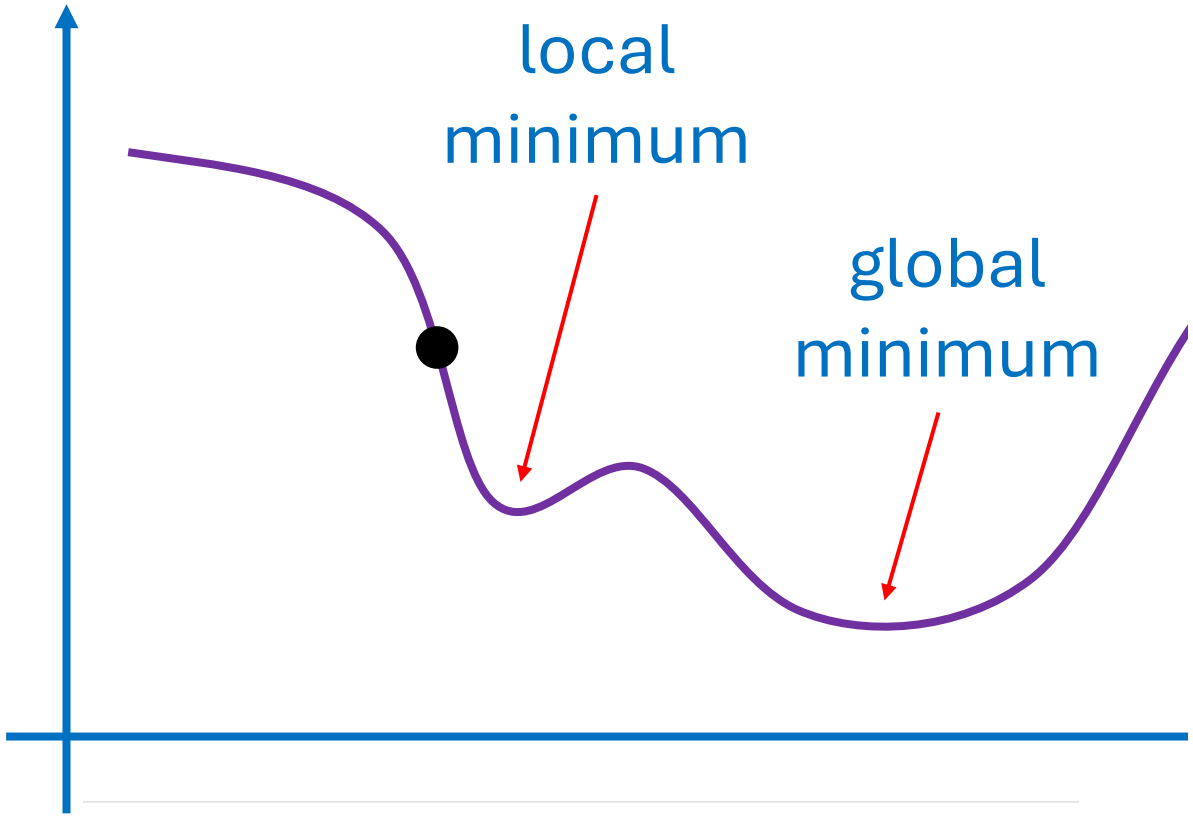
Gradient descent

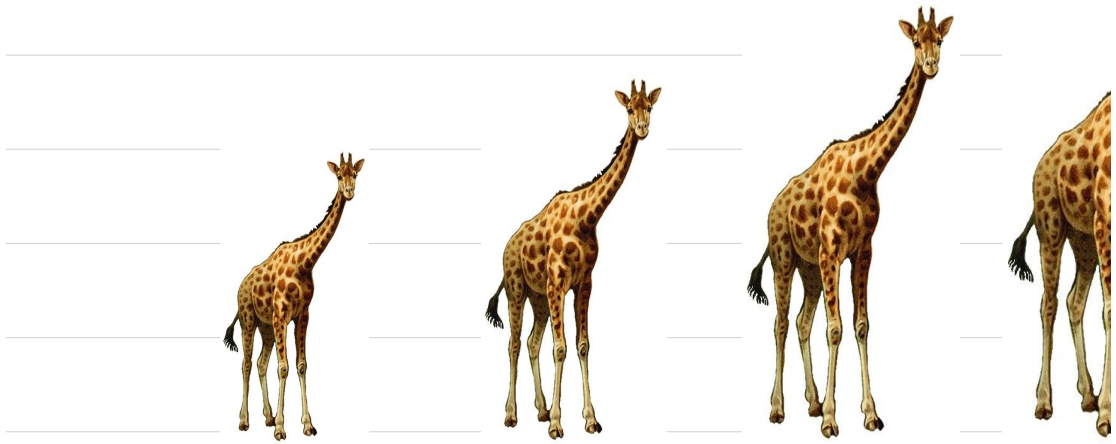
success

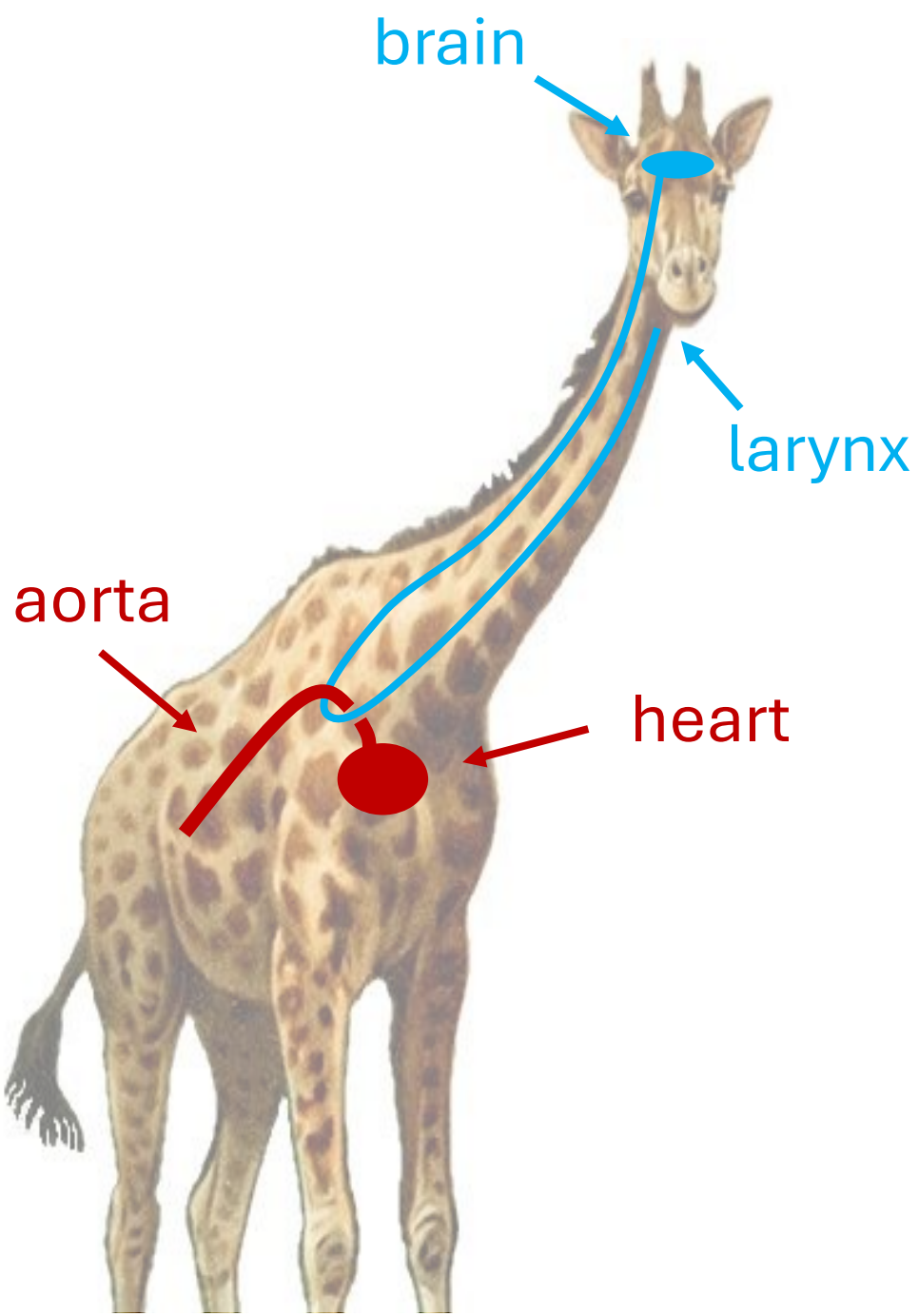
height



loss







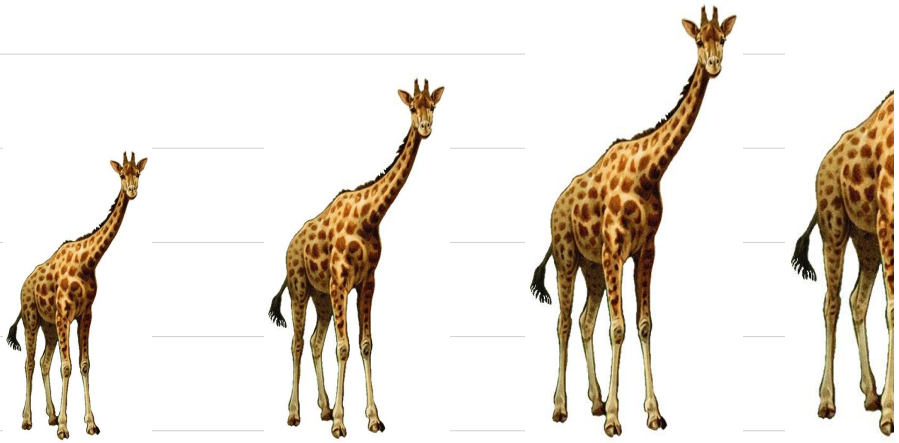


Giraffe / CC0

loss

local minimum

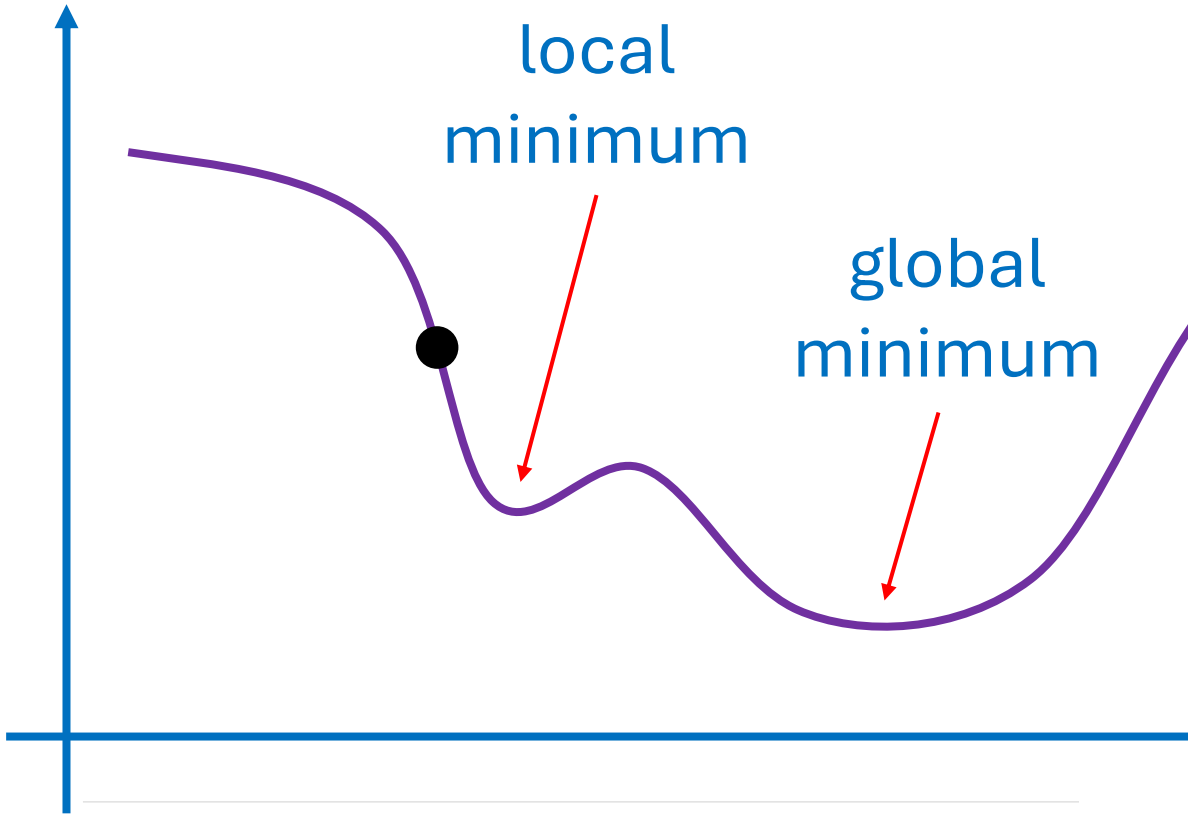
global minimum

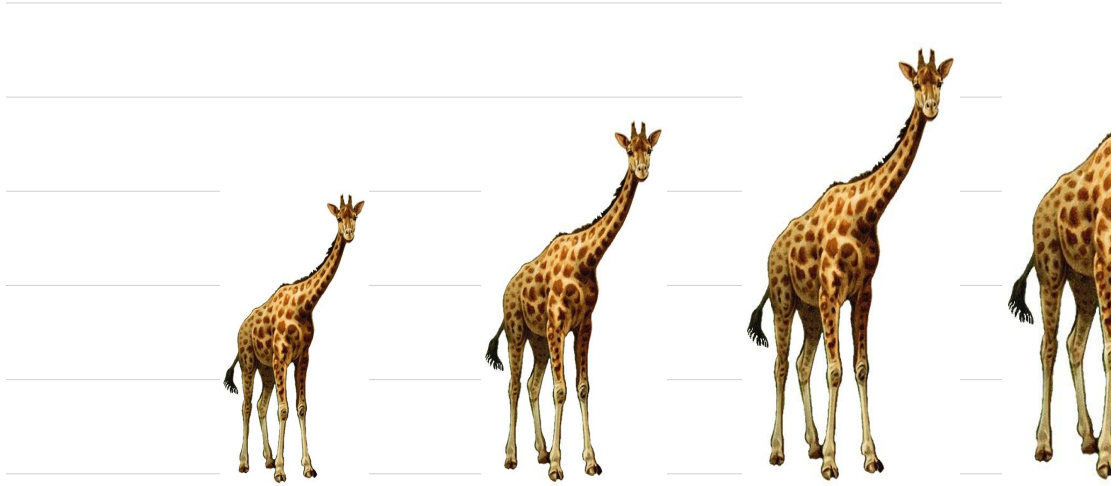


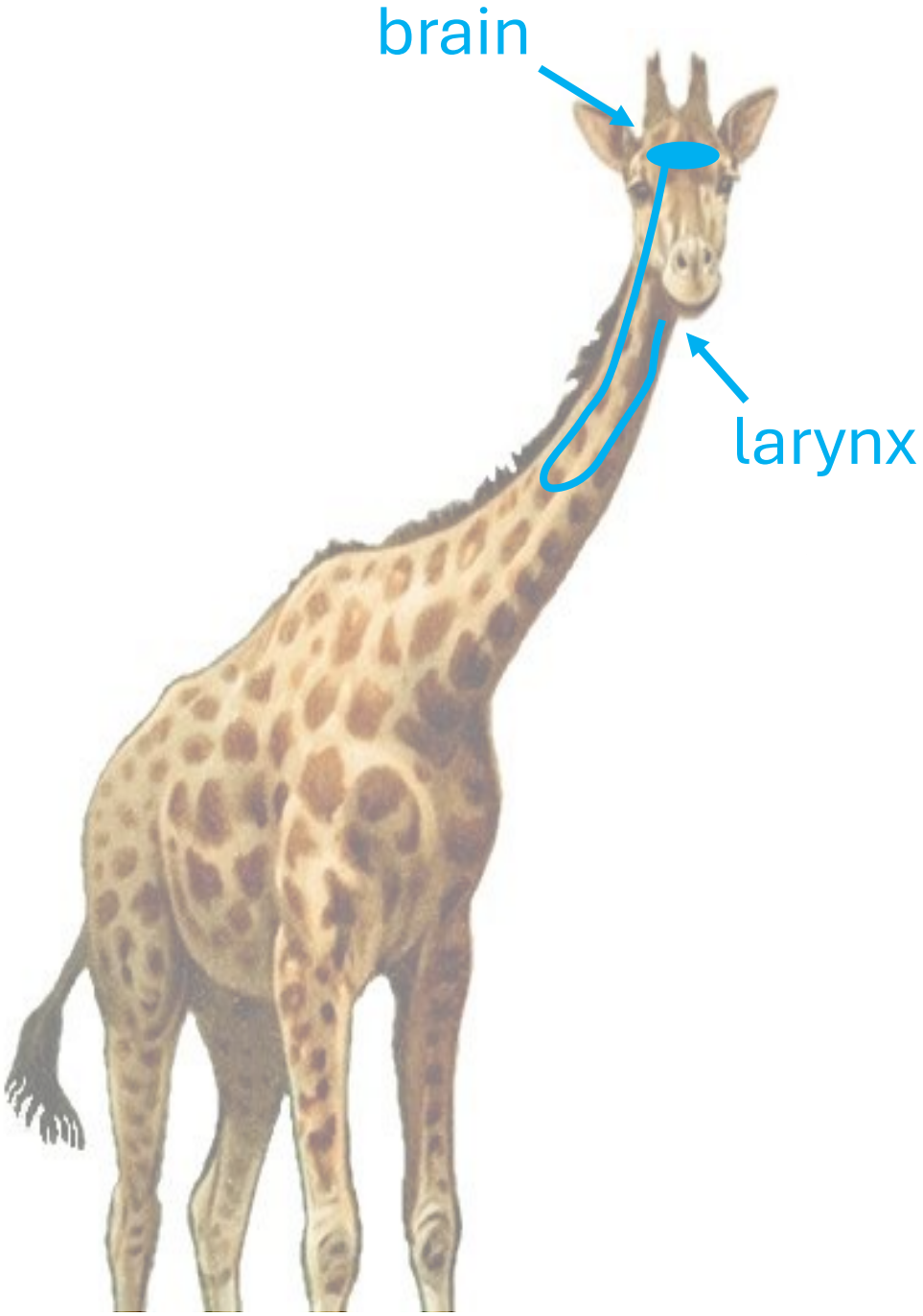
brain

larynx

loss











Giraffe / CC0



Marichini lakes (tarns) / CC0

How AlexNet learns

~~One photo at a time~~

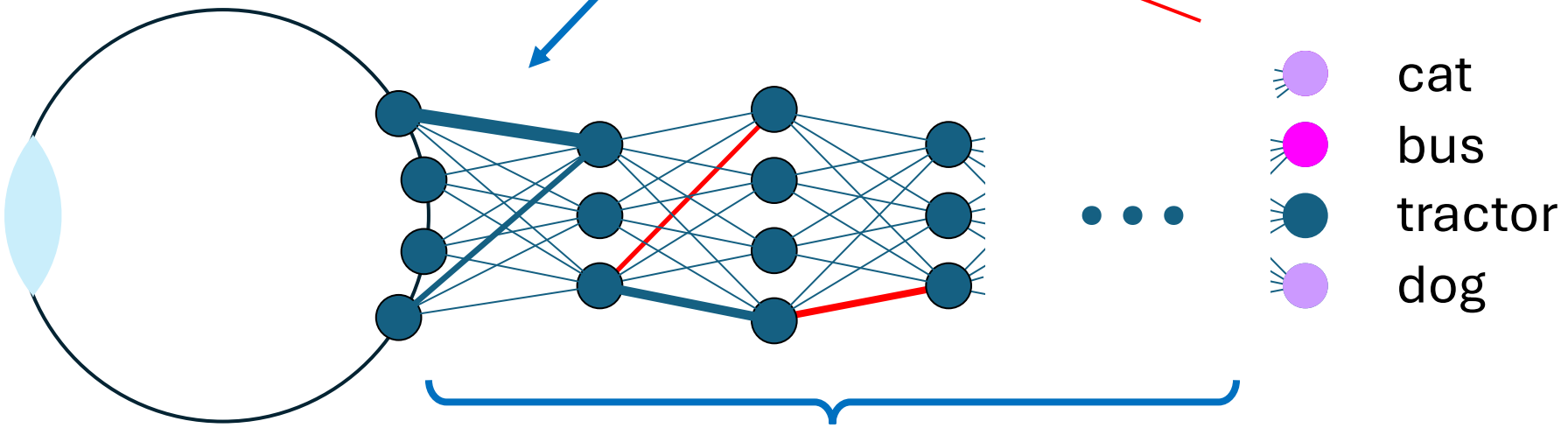
~~One weight at a time~~

~~Guessing ?!~~









choose these numbers!

?

10^{160} permutations of two packs of cards

10^{140}

10^{120}

10^{100} the observable universe as a computer

10^{80} atoms in the observable universe

permutations of a pack of cards

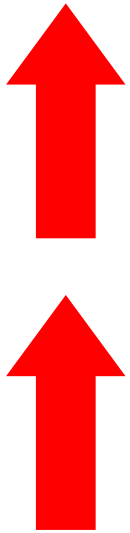
10^{60}



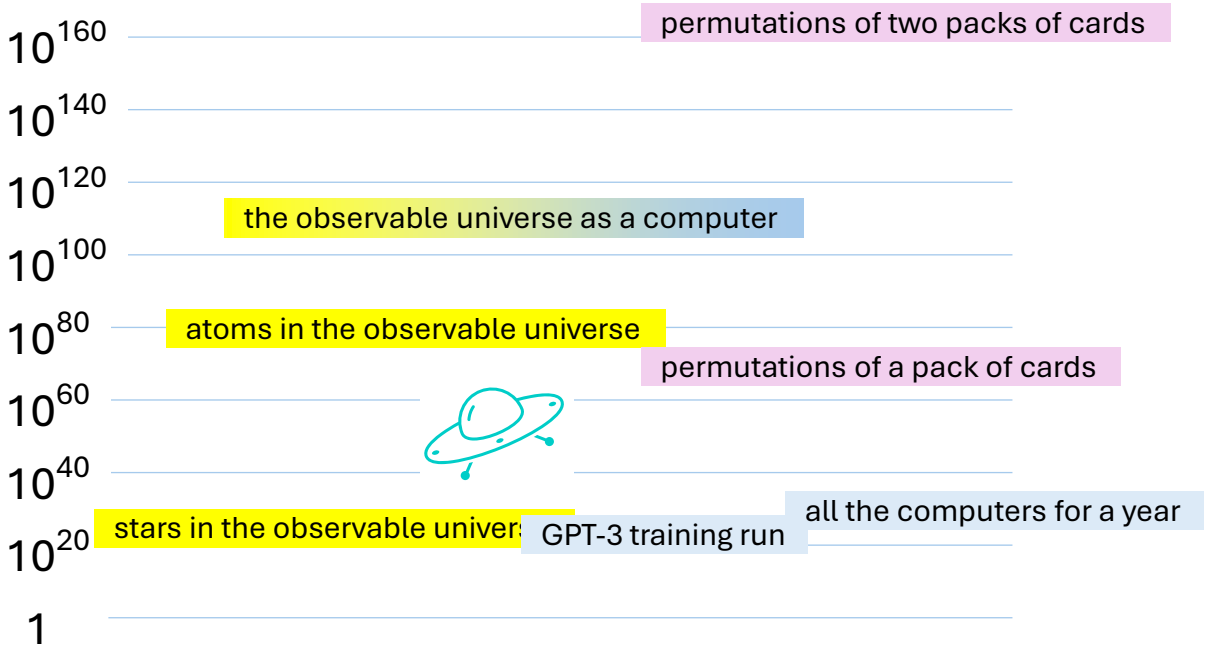
10^{40}

10^{20} stars in the observable universe GPT-3 training run all the computers for a year

1



Number of guesses



~~One photo at a time~~

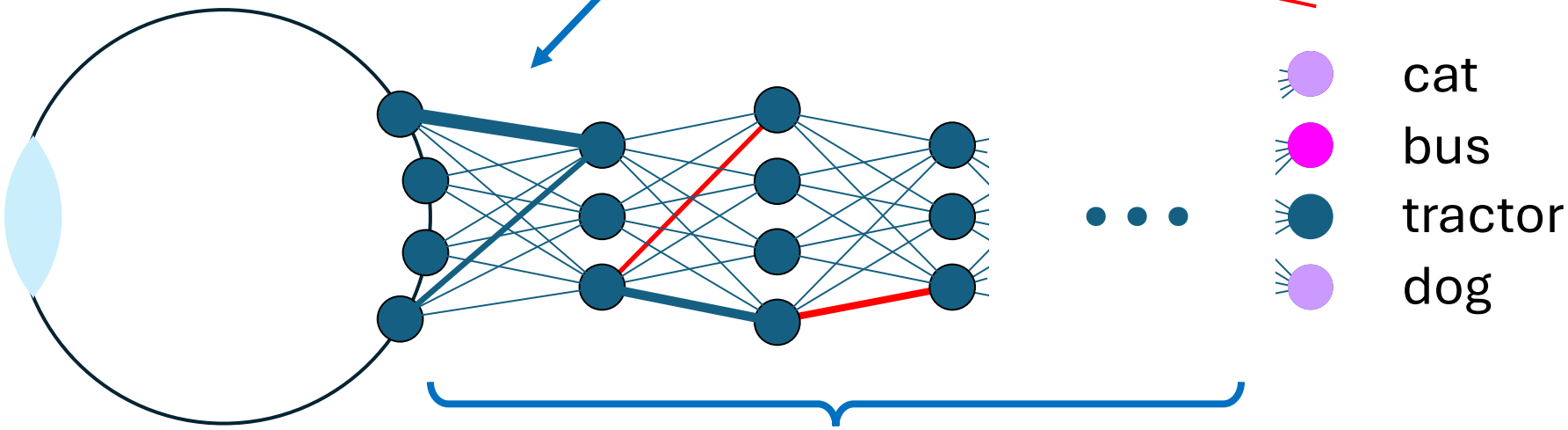
~~One weight at a time~~

~~Guessing ?!~~





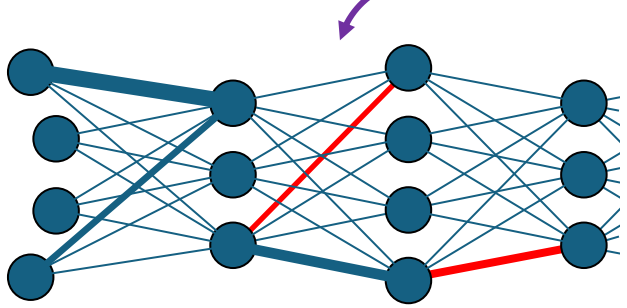




choose these numbers!

?

The algorithm





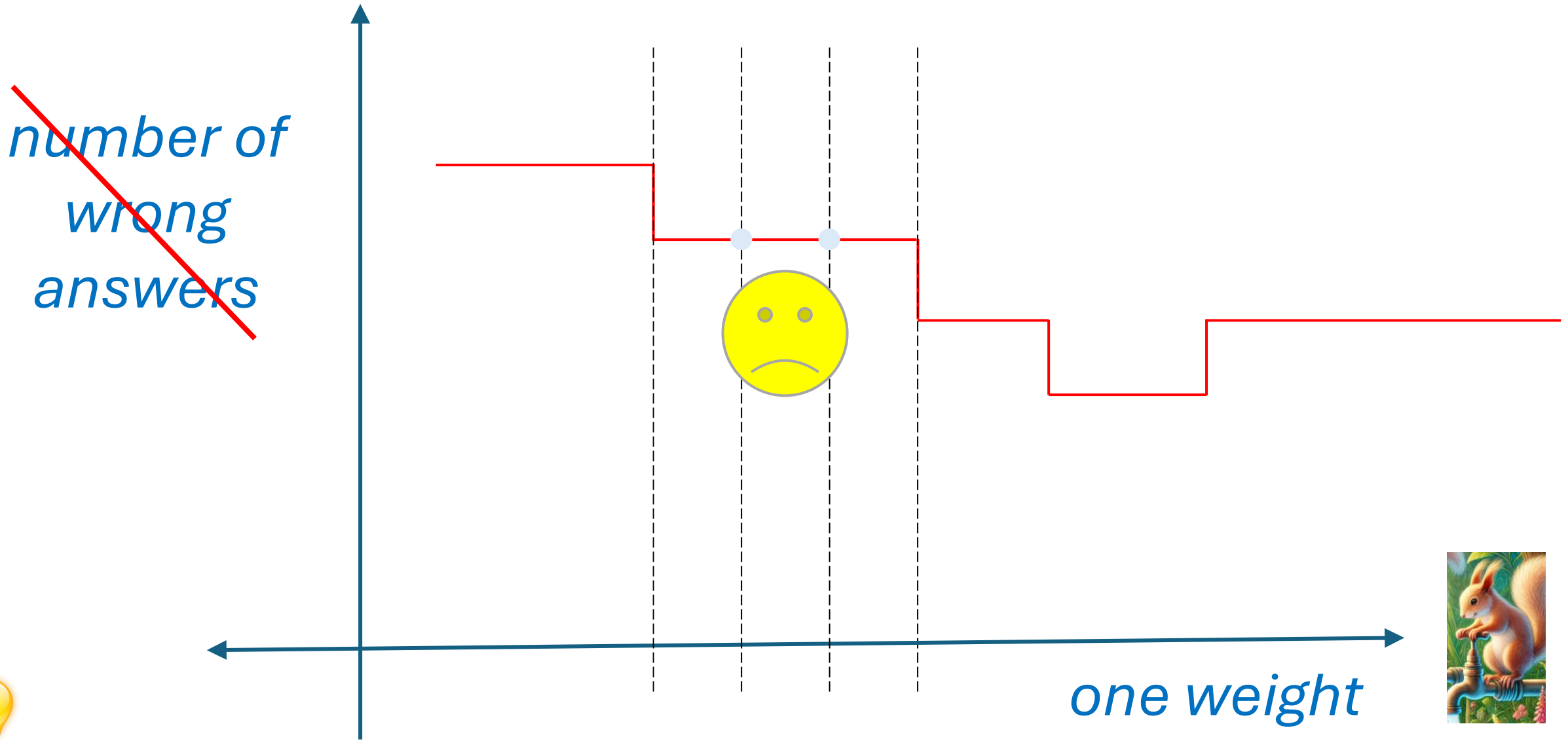
- Start with a newborn AlexNet
- Improve things a *tiny* step at a time



? - (by adjusting all the weights)

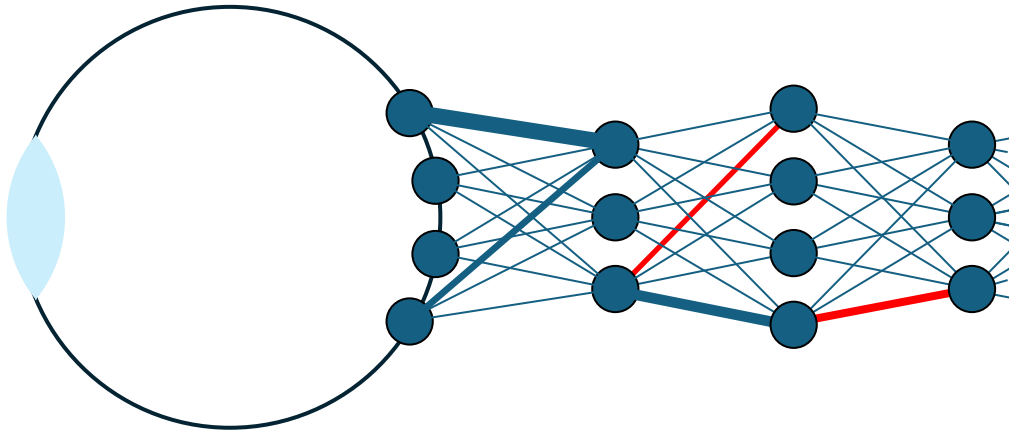
? ❖ (by adjusting all the weights)

? ❖ (by adjusting all the weights)



dog





50% cat
bus
tractor
20% dog



1%



cat



20% ⟶ 21% ✓

60% ⟶ 61% ✓

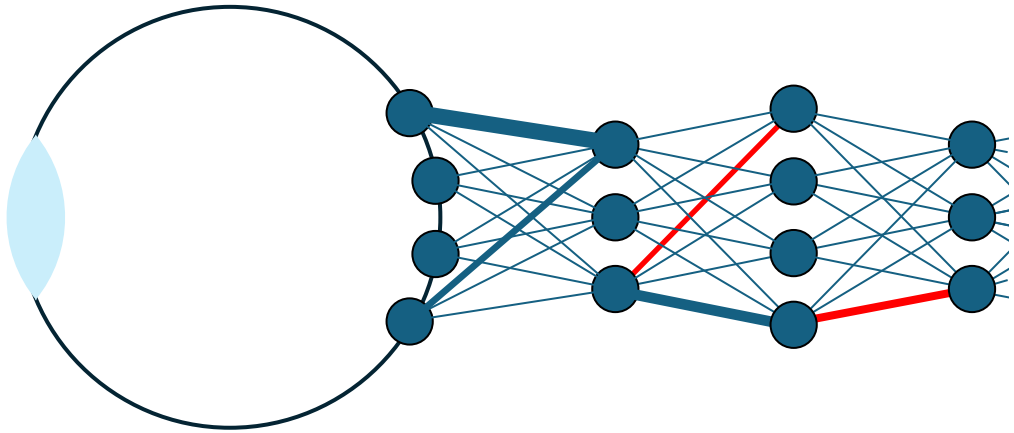
~~*number of wrong answers*~~

~~*tell it when it's wrong*~~



dog





50% cat
bus
tractor
20% dog

loss



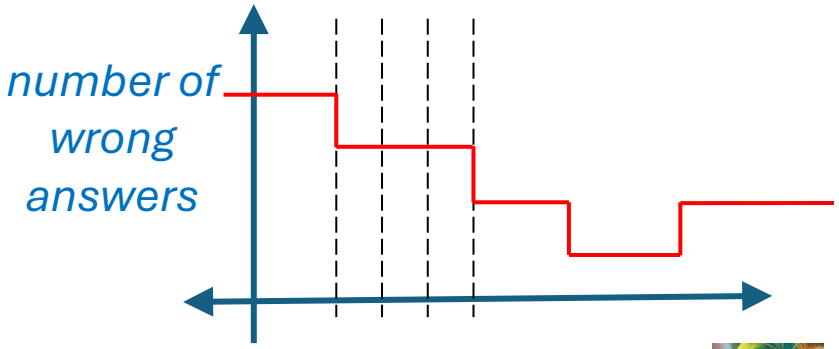
20% 80% 0.8



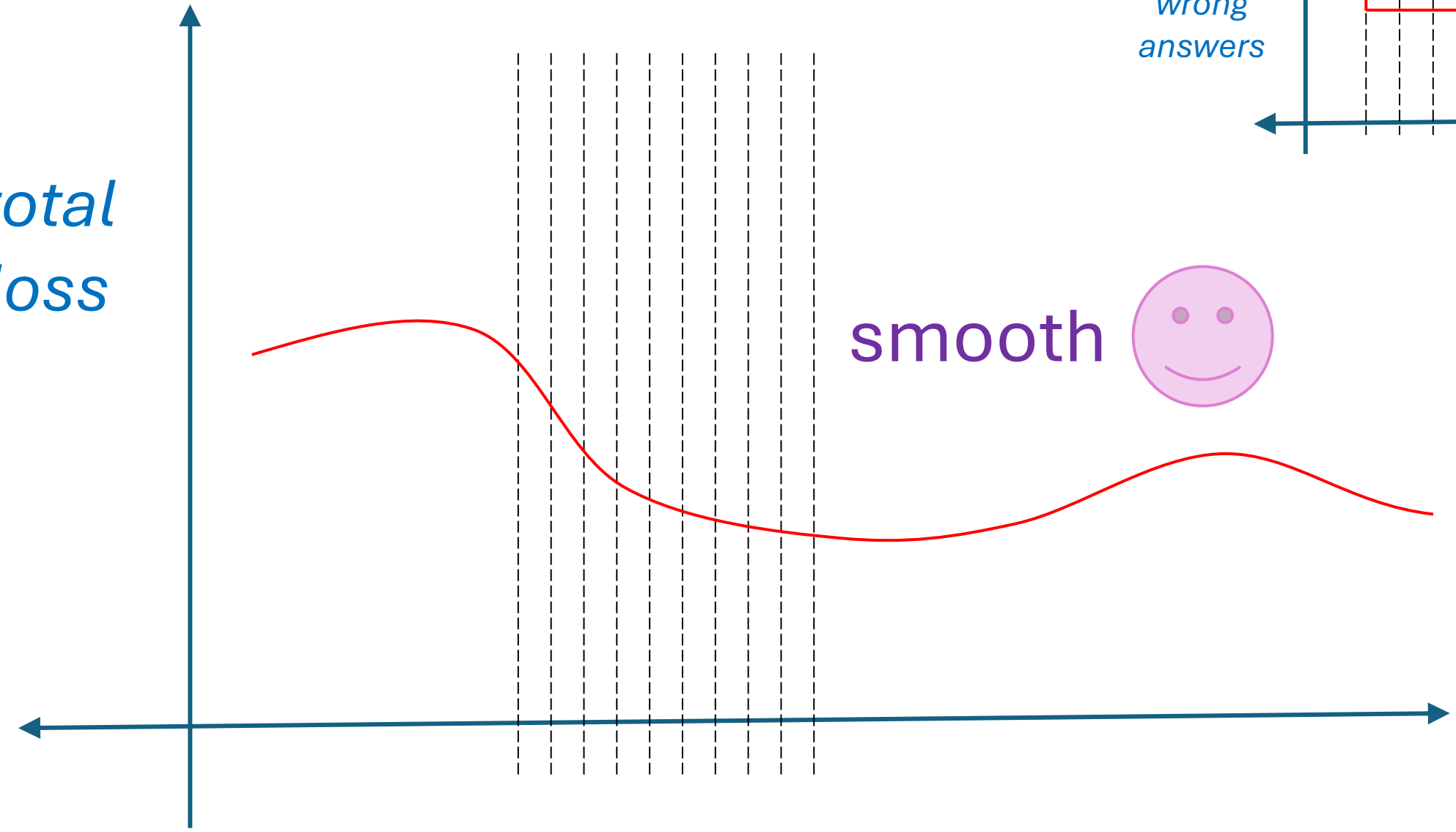
70% 30% 0.3

all the images

total loss


number of
wrong
answers

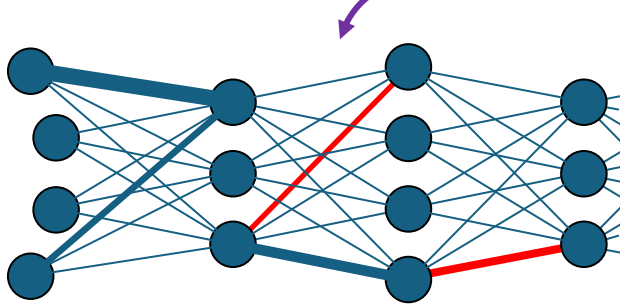



total
loss
smooth





The algorithm

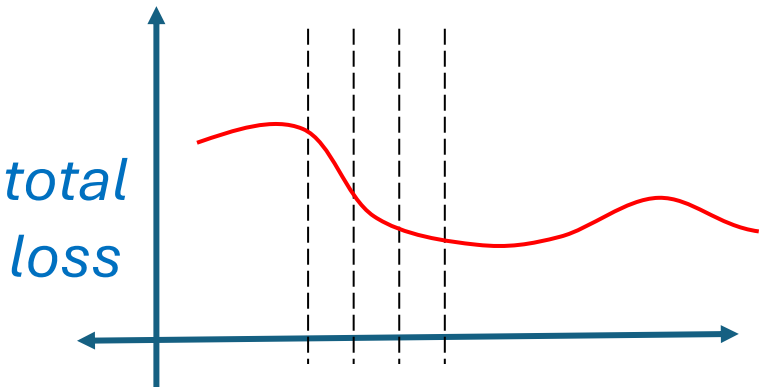




- Start with a newborn AlexNet
- Improve things a *tiny* step at a time



? - (by adjusting all the weights)



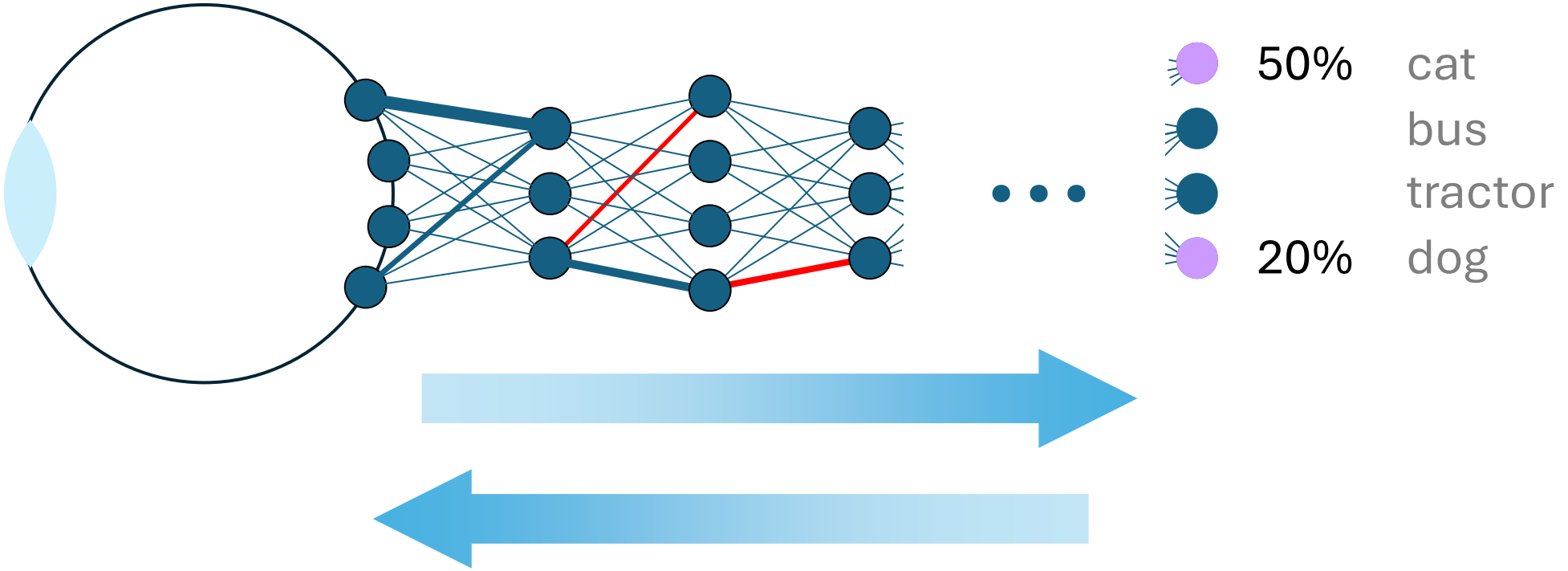






dog





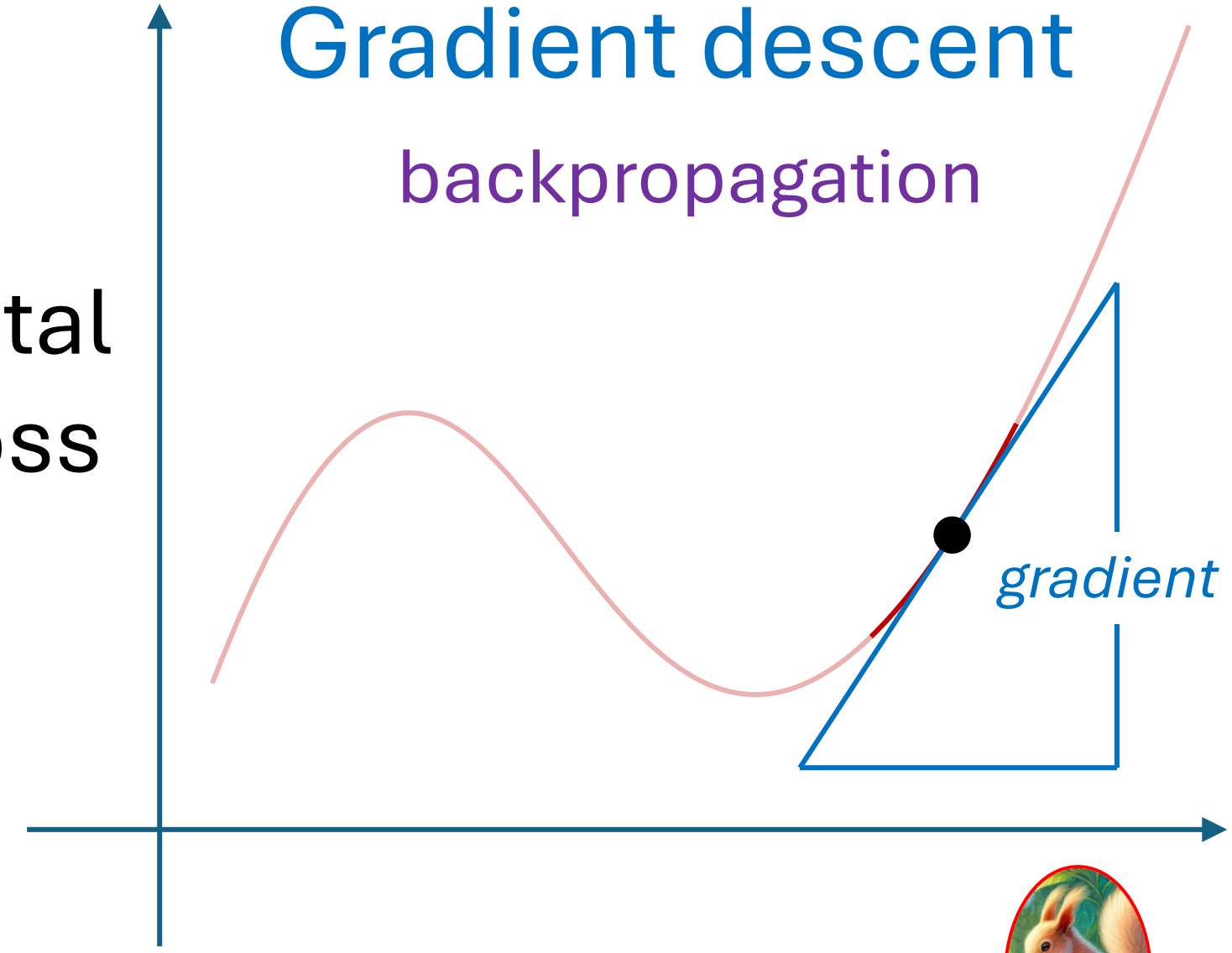
backpropagation



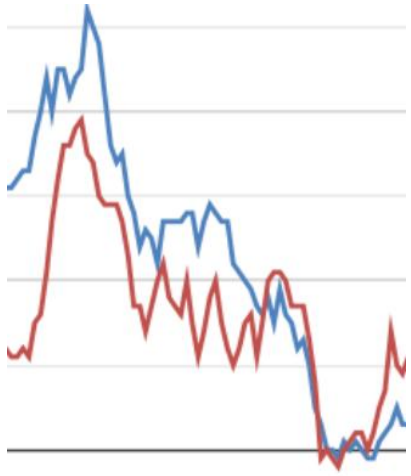
Gradient descent

backpropagation

total loss











Slope by Bouwe Brouwer / CC0

Adversarial machine learning



panda

58% confidence

+

carefully chosen tiny changes



=



gibbon

99% confidence



https://arxiv.org/pdf/1412.6572

Ian J. Goodfellow, Jonathon Shlens & Christian Szegedy Google Inc. (2015)