DECOMPOSING NEURAL NETWORKS

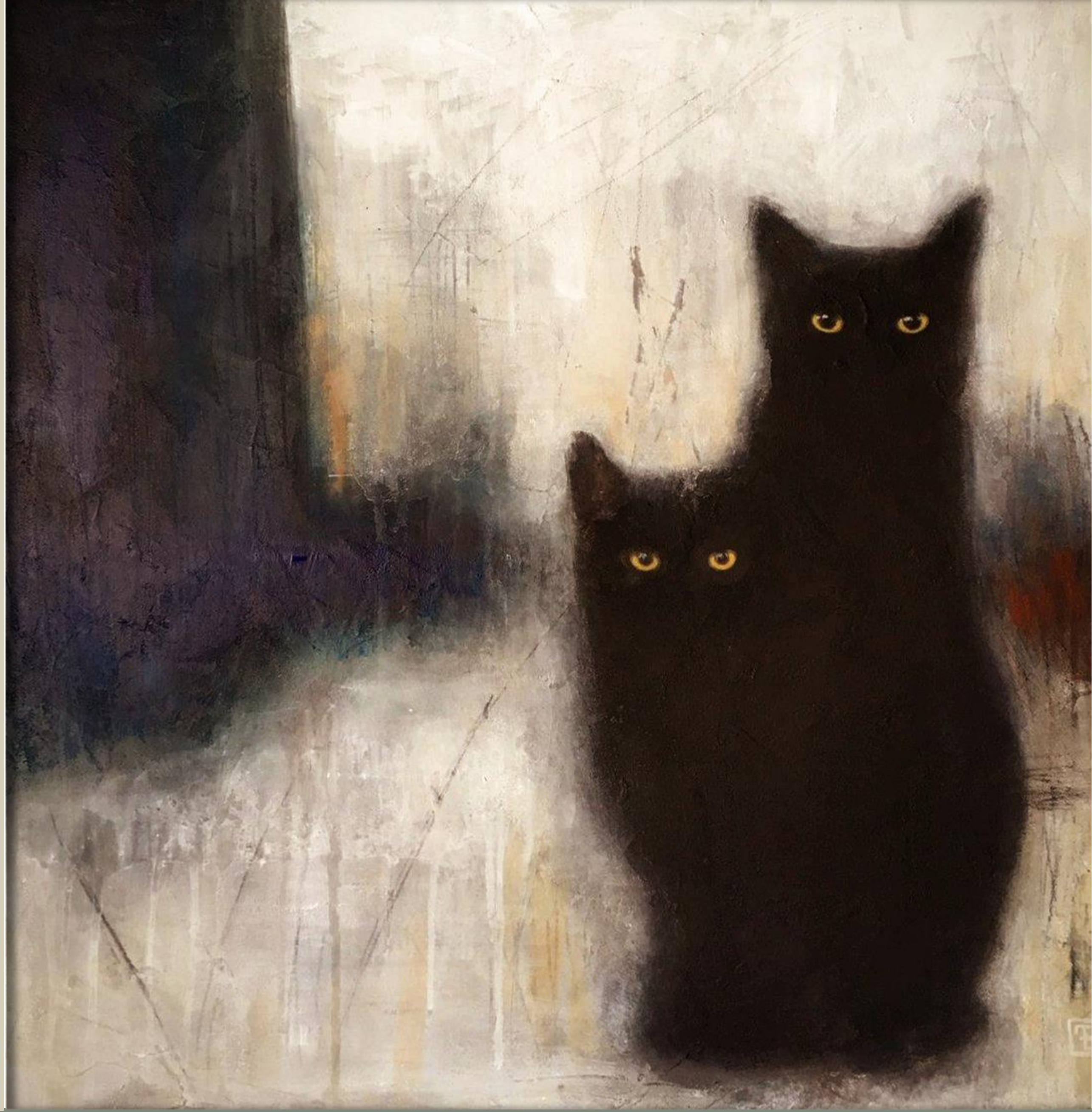
An applicant's guide to artificial learning
29.11.2022

→ JENNIFER MATTHIESEN & TINO PAULSEN | WINTERSEMESTER 2022

LEUPHANA
UNIVERSITÄT LÜNEBURG

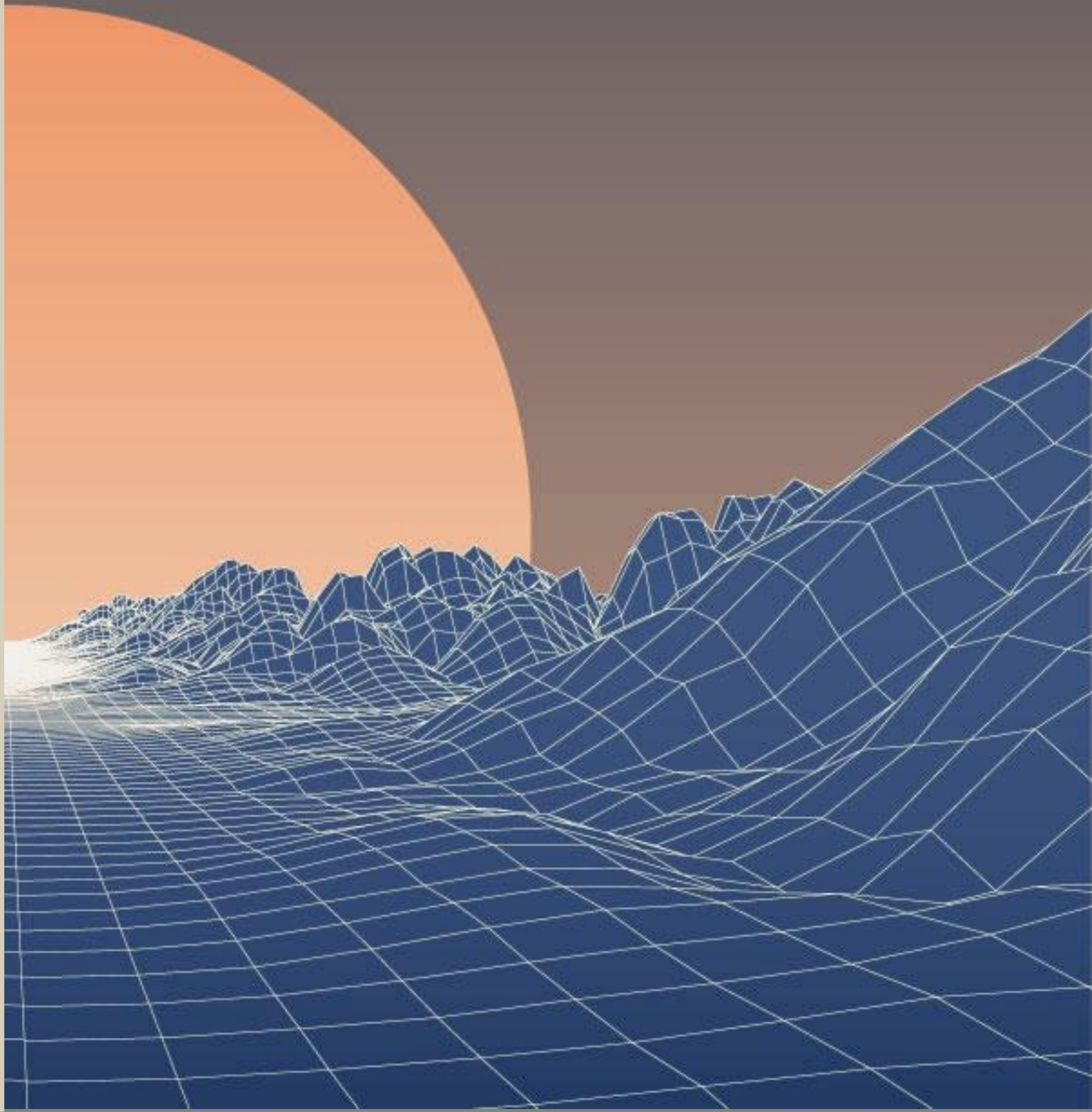
RECAP
ABOUT DATA, CATS & DOGS

- Vectors, Matrix, Tensor
- Classification & bias
- BIAS from data
- How NN's understand images
  - Convolutions & pooling
- Train your own network in a supervised manner

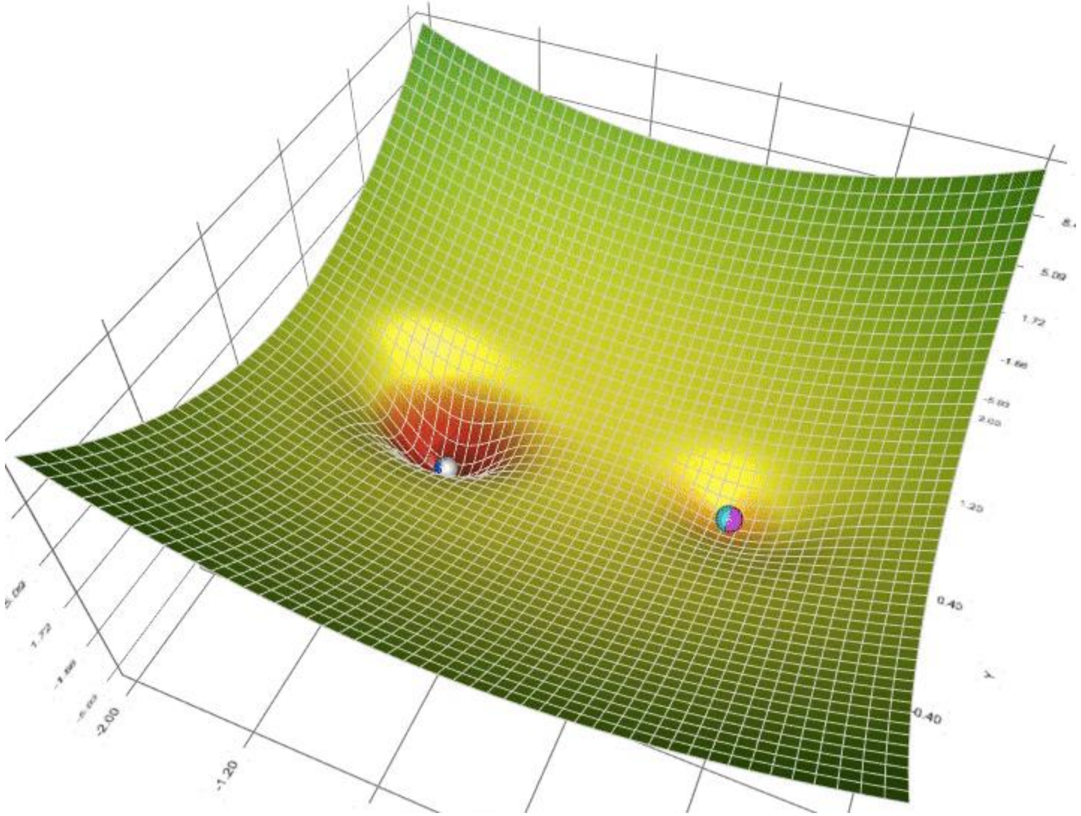


TODAY: OPTIMISATION

- Objective functions
- Numerical optimization
- How to train your network
- Optimizers?



What are we doing today?
Marbles!



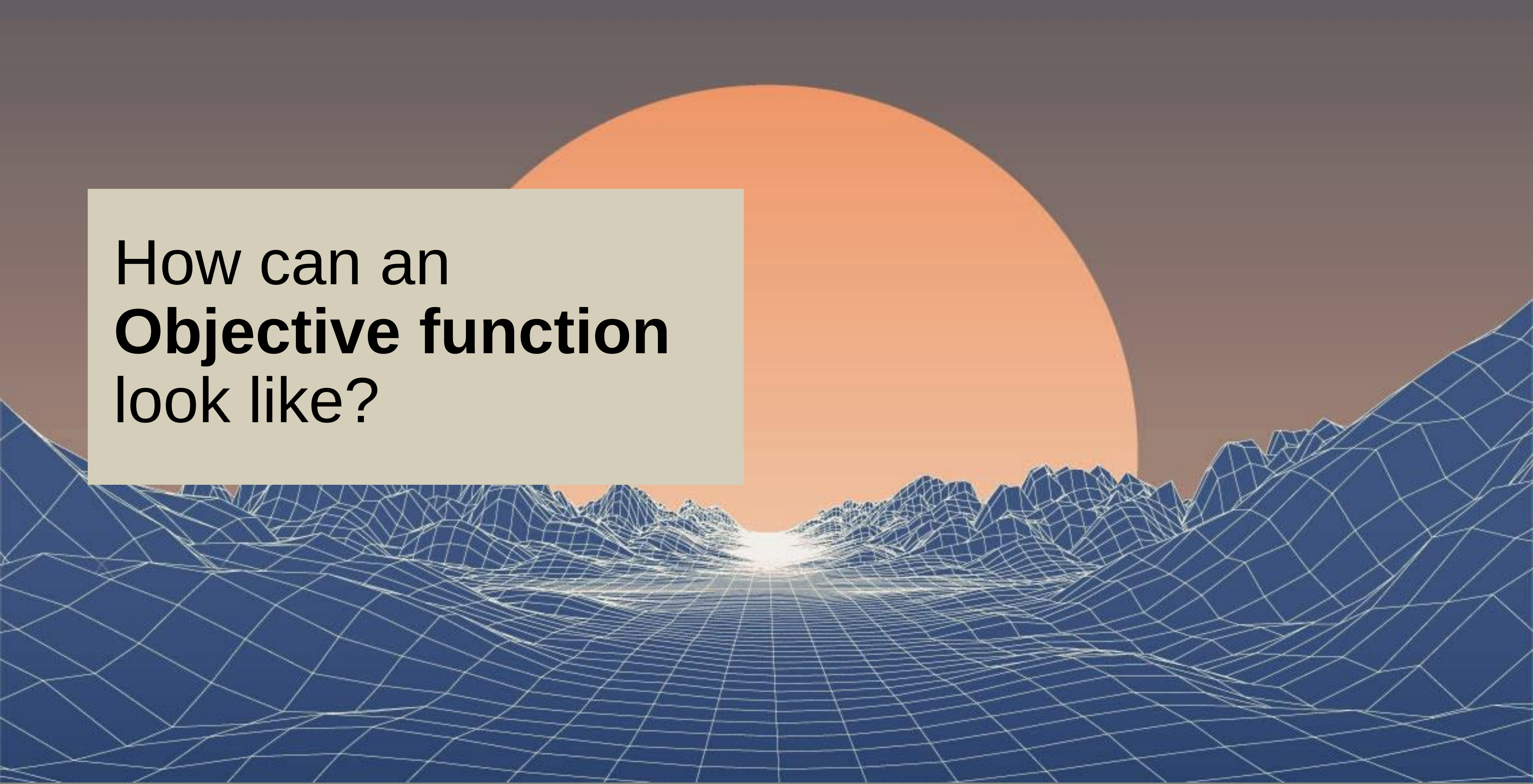
https://towardsdatascience.com/a-visual-explanation-of-gradient-descent-methods-momentum-adagrad-rmsprop-adam-f898b102325c

How does a neural network understand what to learn?

- A neural network is trained on data, but how?
- The “goal” must be understandable for the machine

  ➡ Quantifiable

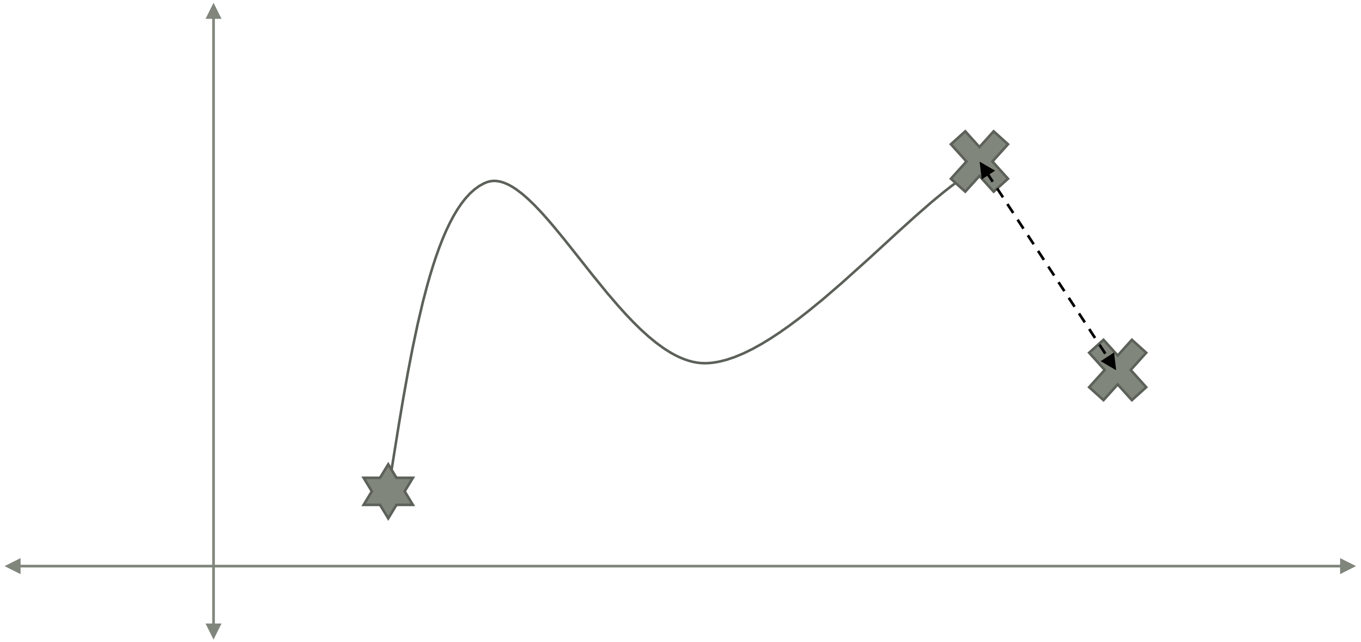
- Operationalizing “goals” into objective functions



How can an **Objective function** look like?

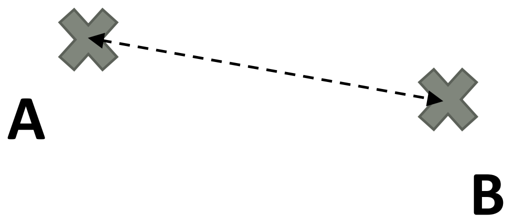
Objective functions
How do they need to be?

Imagine we have the telemetry data of a smartphone…



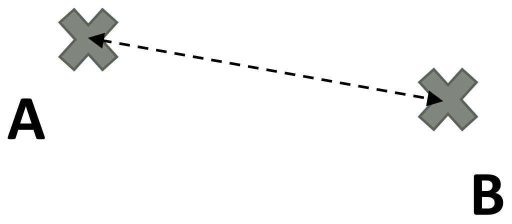
Objective functions
Properties

- An objective function is typically like a distance metric:
  - Symmetric: $dist(a,b) = dist(b,a)$



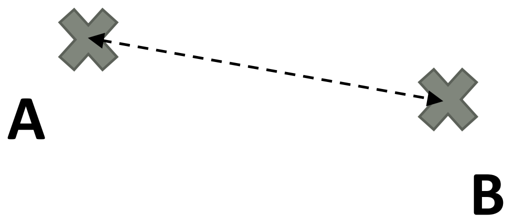
Objective functions
Properties

- An objective function is typically like a distance metric:
  - Positive: $dist(a, b) \geq 0$



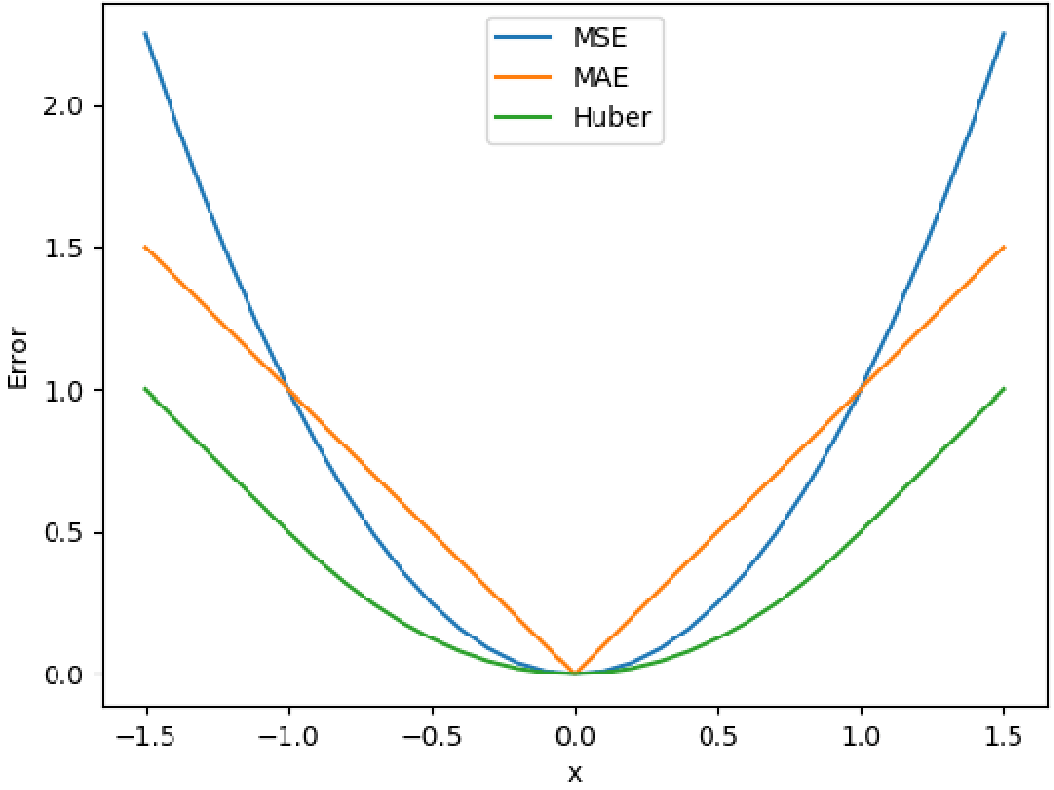
Objective functions
Properties

- An objective function is typically like a distance metric:
  - Symmetric: $dist(a, b) = dist(b, a)$
  - Positive: $dist(a, b) \geq 0$



  - Beware! This is in no way exhaustive!

Objective functions
Examples



https://www.researchgate.net/profile/Thibaut-Theate/publication/340644261/figure/fig3/AS:880423299186688@1586920680062/Comparison-of-the-MSE-MAE-and-Huber-losses.png

Objective functions
Examples

- Mean squared error

$$mse = \sum_{i=1}^{N}(x - \bar{x})^2$$

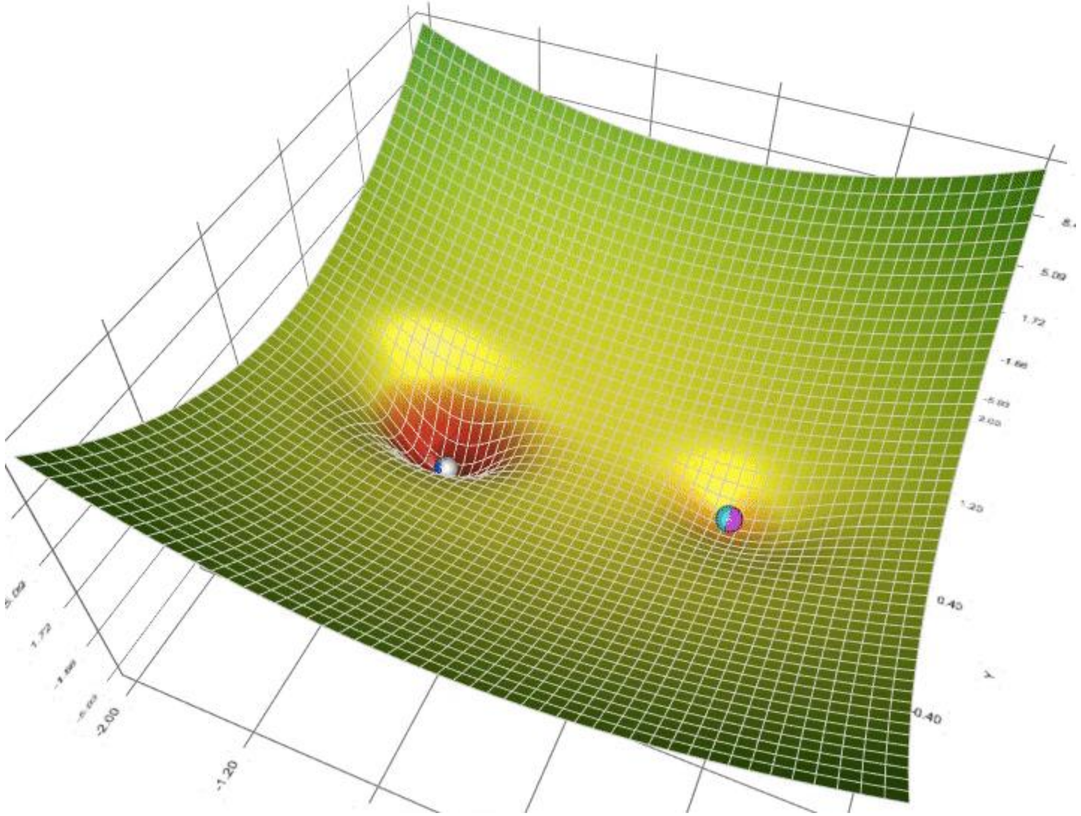
- Mean absolute error

$$mae = \sum_{i=1}^{N}|x - \bar{x}|$$

How can a function for **Classification** look like?

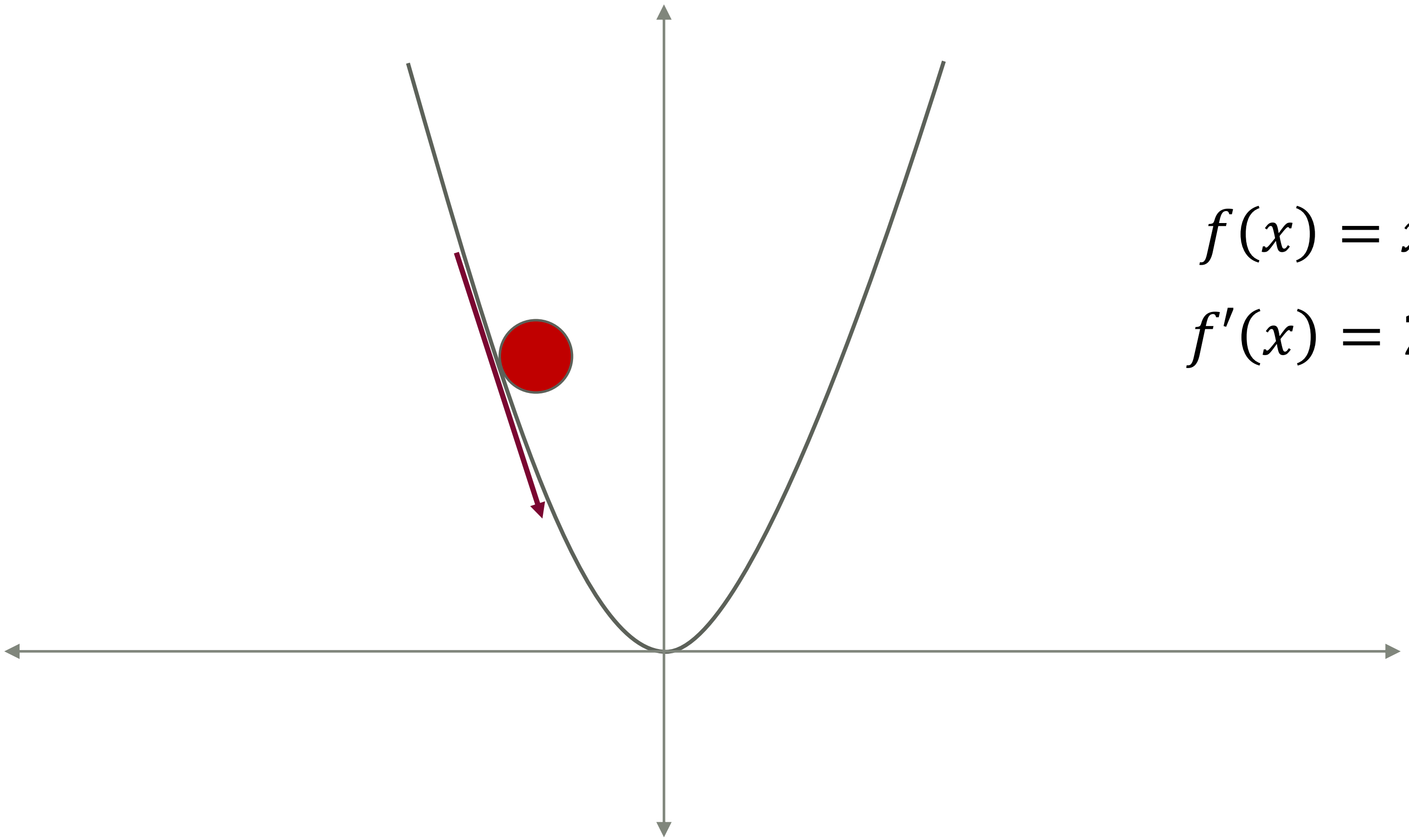
Create your own **Objective function!**
Desmos

What are we doing today?
Marbles!



https://towardsdatascience.com/a-visual-explanation-of-gradient-descent-methods-momentum-adagrad-rmsprop-adam-f898b102325c

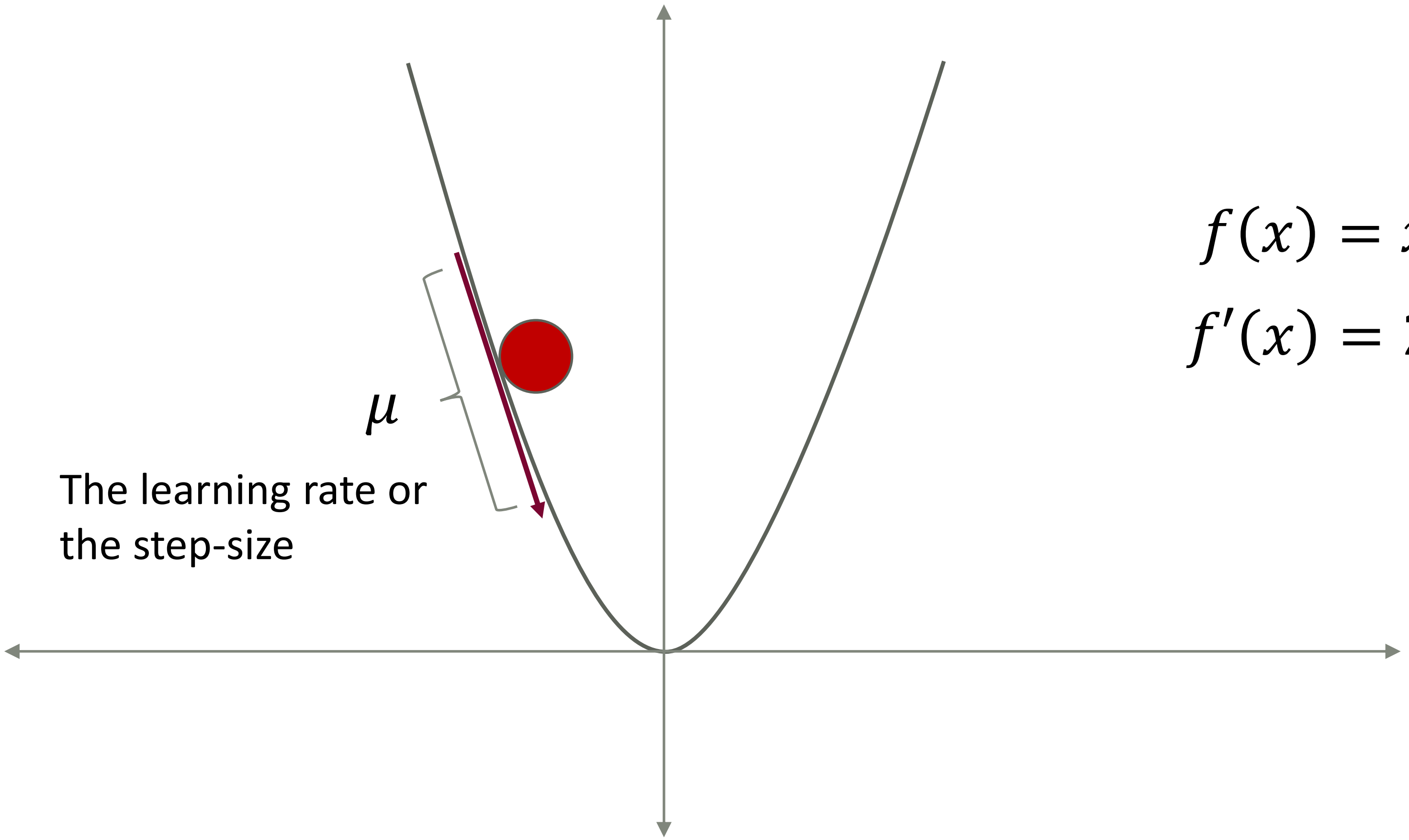
Numerical optimization
The bowl (in technical)



$$f(x) = x^2$$

$$f'(x) = 2x$$

Numerical optimization
The bowl (in technical)



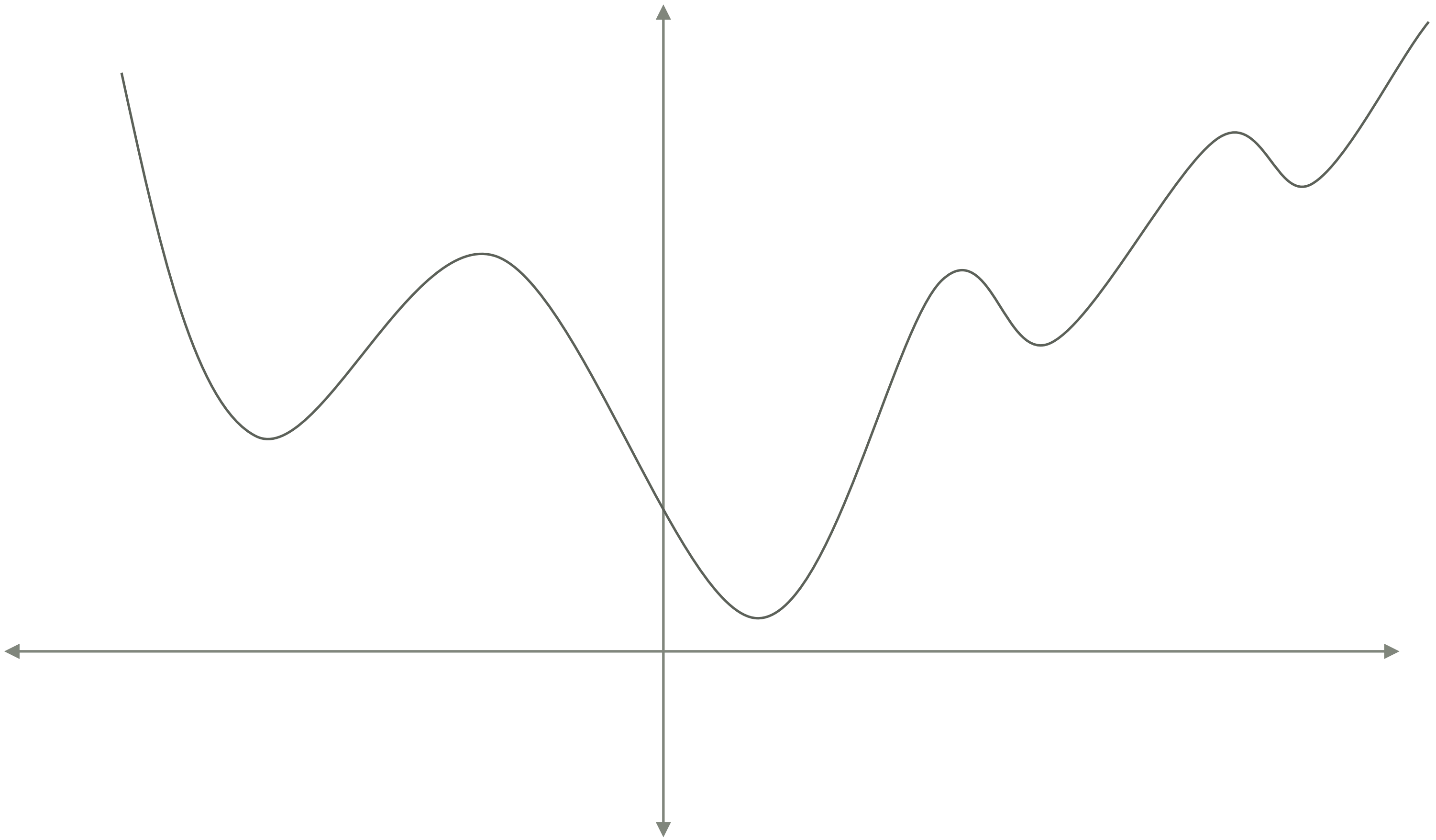
$$f(x) = x^2$$

$$f'(x) = 2x$$

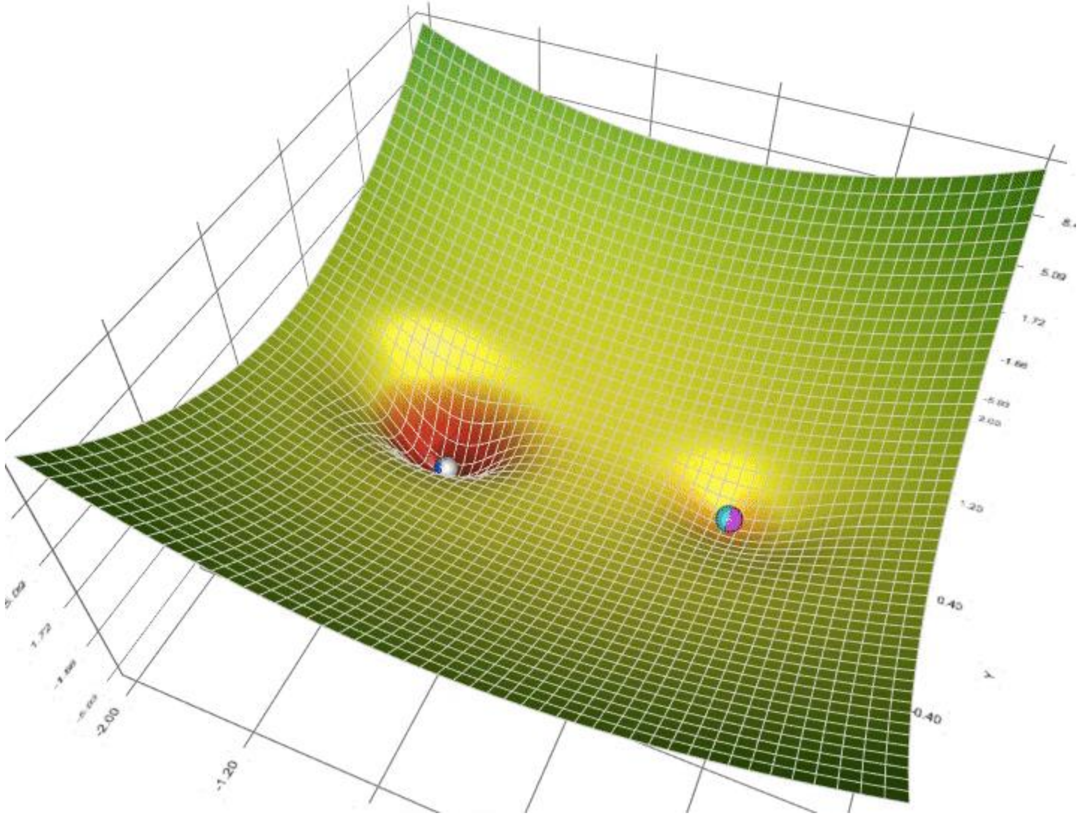
TIME FOR TRYING IT OUT!

https://www.benfrederickson.com/numerical-optimization/

Numerical optimization
The problems with local minima



What are we doing today?
Marbles!



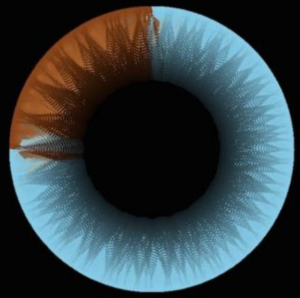
https://towardsdatascience.com/a-visual-explanation-of-gradient-descent-methods-momentum-adagrad-rmsprop-adam-f898b102325c

How to train your network
Combining knowledge

– Our objective function defines our measure of “goodness”, the resulting mathematical landscape is called the loss surface

– This loss surface is the “bowl” we optimize then, as it can be very complex, there can be many local minima

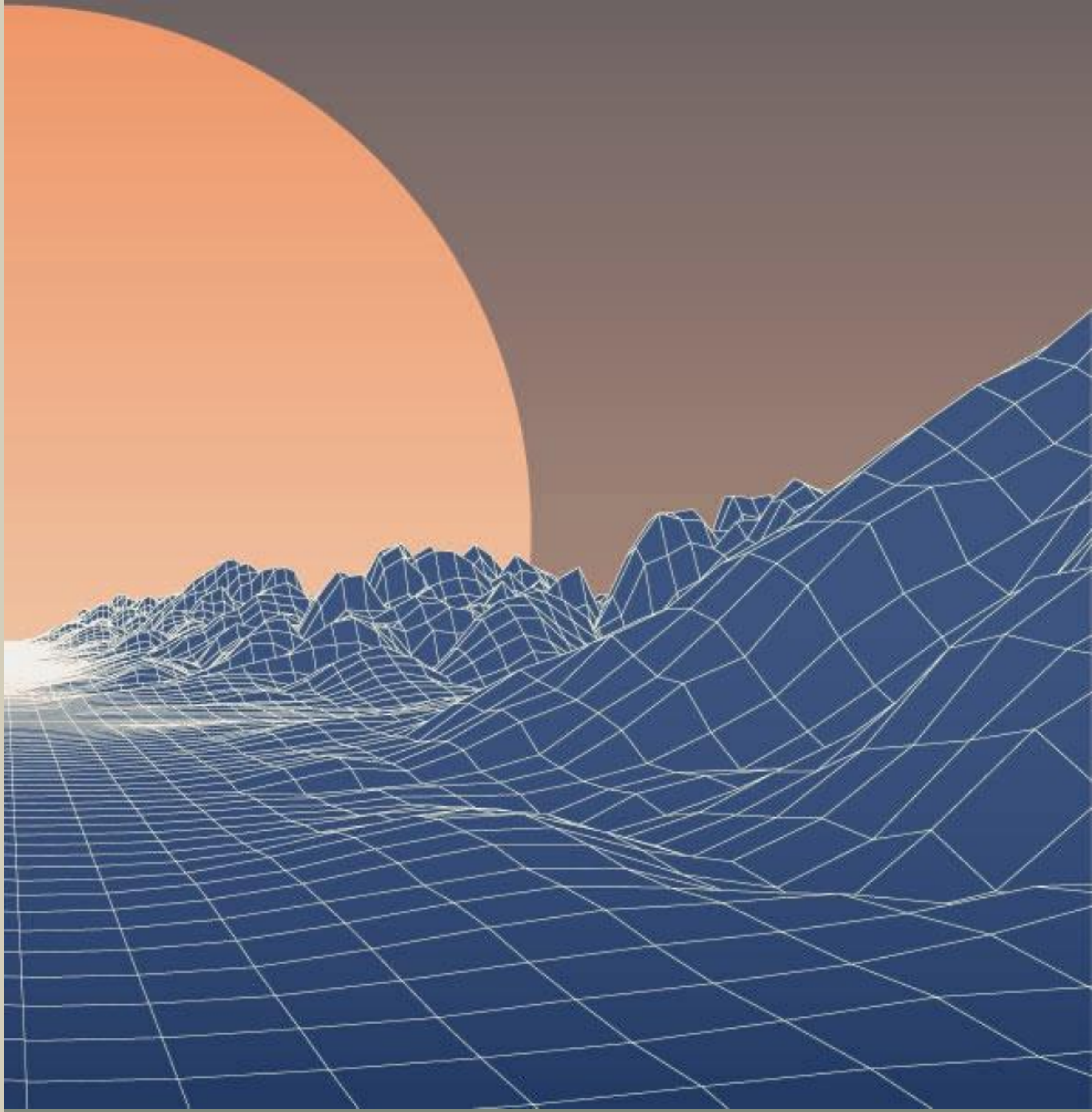
Backpropagation



3Blue1Brown Backpropagation

Practice: Plotting

- Train your network (again)
- Keep track of the losses
- Plot them (you can use the provided function, if you want)
- Try out different optimizers



SUMMARY OF TODAY
OPTIMISATION

- Objective functions
  - What they look like, what they do and why they limit a network
- Numerical optimization and local vs globa minima
- How to train your network
  - And plot it!
- Optimizers and a rough idea why they are powerful

If you want to repeat todays content in a guided manner