Machine Learning for Design

Lecture 3

Machine Learning for Images. *Part 1*

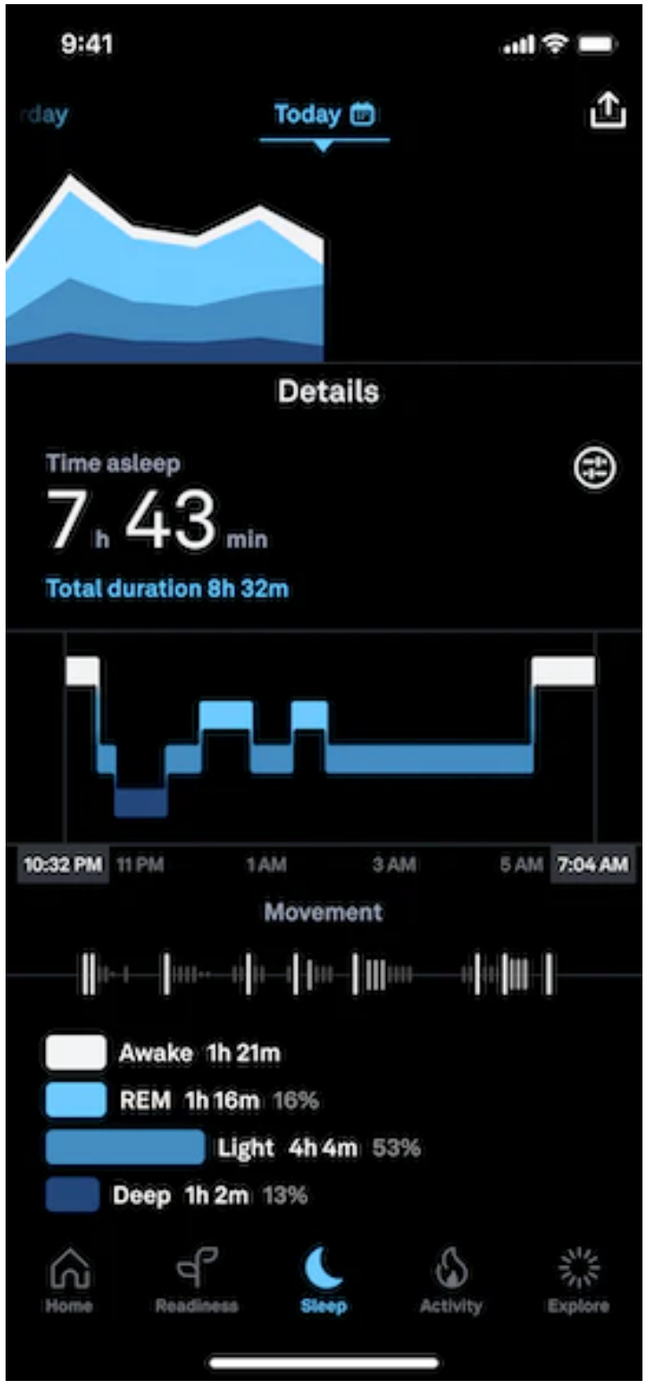
Administration

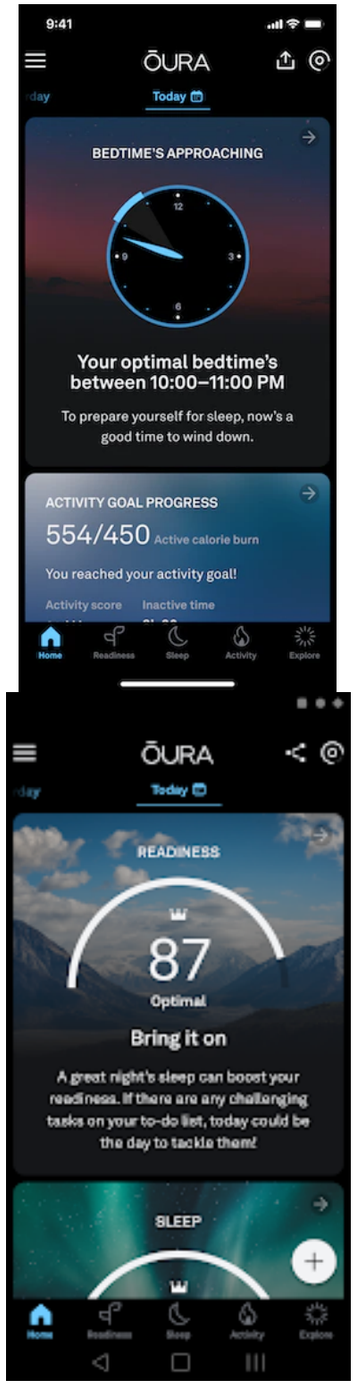
- Group formation
  - 48 students
  - 3 group of 3 → get together?
    - Or reach out to me, and I will re-distribute
- Forum in Brightspace
  - Collaboration > Discussions
  - For questions and discussions about lectures, assignments, etc.

(Reverse) Design Exercise

Oura Ring 3

https://ouraring.com/oura-experience


9:41
rday
Today
Details
Time asleep
7 h 43 min
Total duration 8h 32m
10:32 PM
11 PM
1 AM
3 AM
5 AM
7:04 AM
Movement
Awake 1h 21m
REM 1h 16m 16%
Light 4h 4m 53%
Deep 1h 2m 13%
Home
Readiness
Sleep
Activity
Explore


9:41
ŌURA
rday
Today
BEDTIME'S APPROACHING
Your optimal bedtime's between 10:00–11:00 PM
To prepare yourself for sleep, now's a good time to wind down.
ACTIVITY GOAL PROGRESS
554/450 Active calorie burn
You reached your activity goal!
Activity score
Inactive time
Home
Readiness
Sleep
Activity
Explore
ŌURA
Today
READINESS
87
Optimal
Bring it on
A great night's sleep can boost your readiness. If there are any challenging tasks on your to-do list, today could be the day to tackle them!
SLEEP
Home
Readiness
Sleep
Activity
Explore

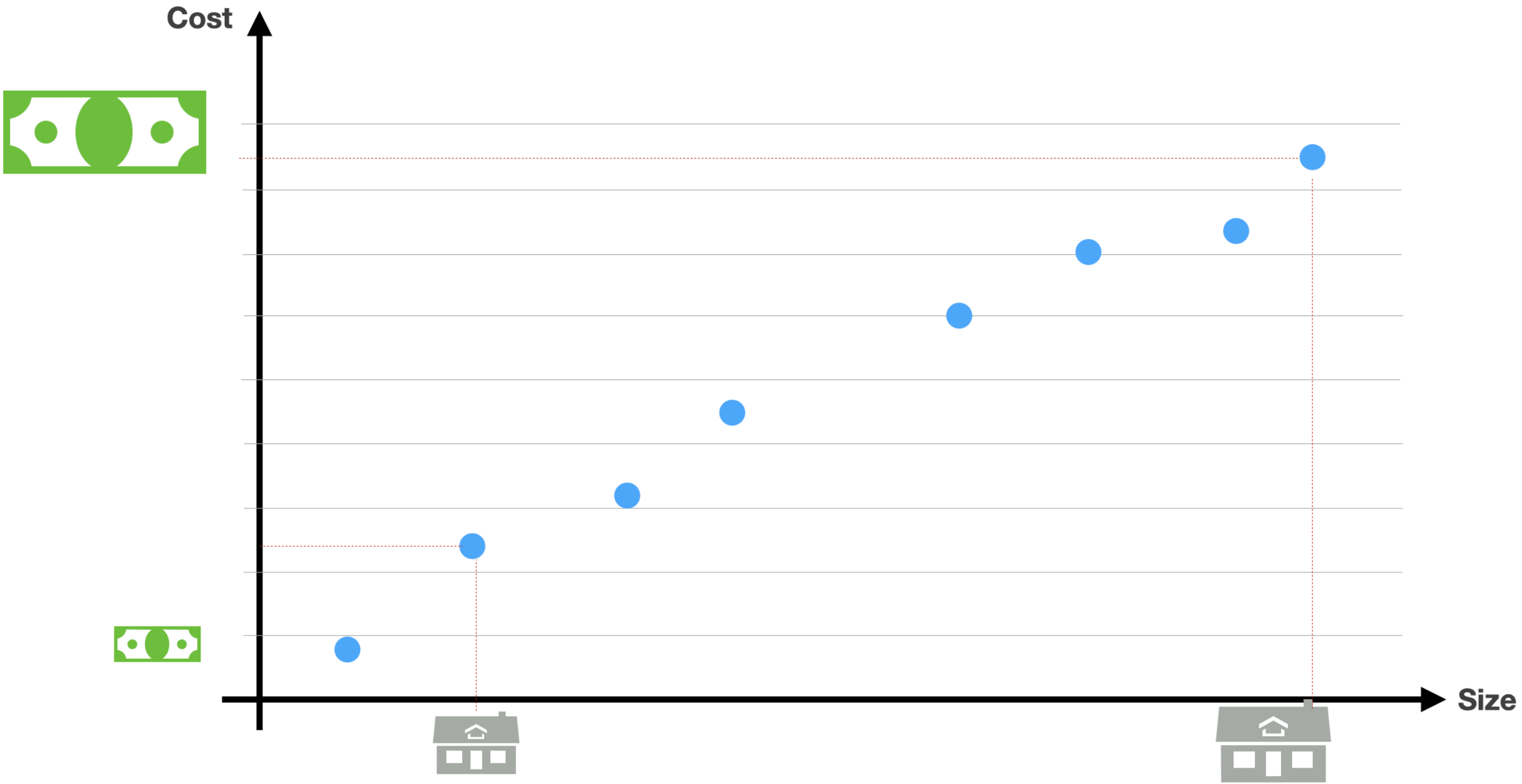
Formative Assignment

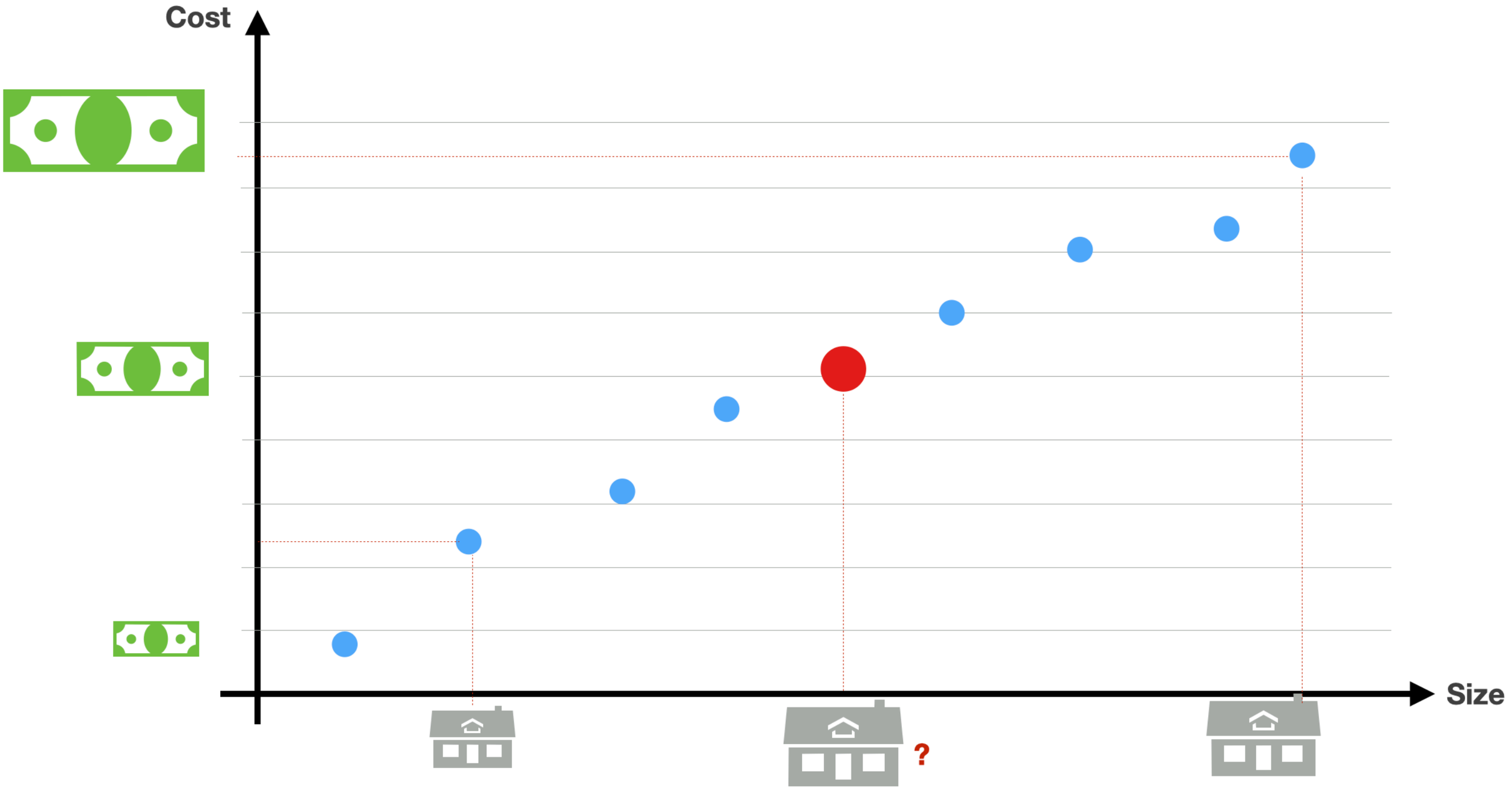
- Fill in the questionnaire in Bright Space
  - Materials > Week 3 > ML Design Questionnaire: OURA
- I will analyze your answers, and we will discuss them together next week

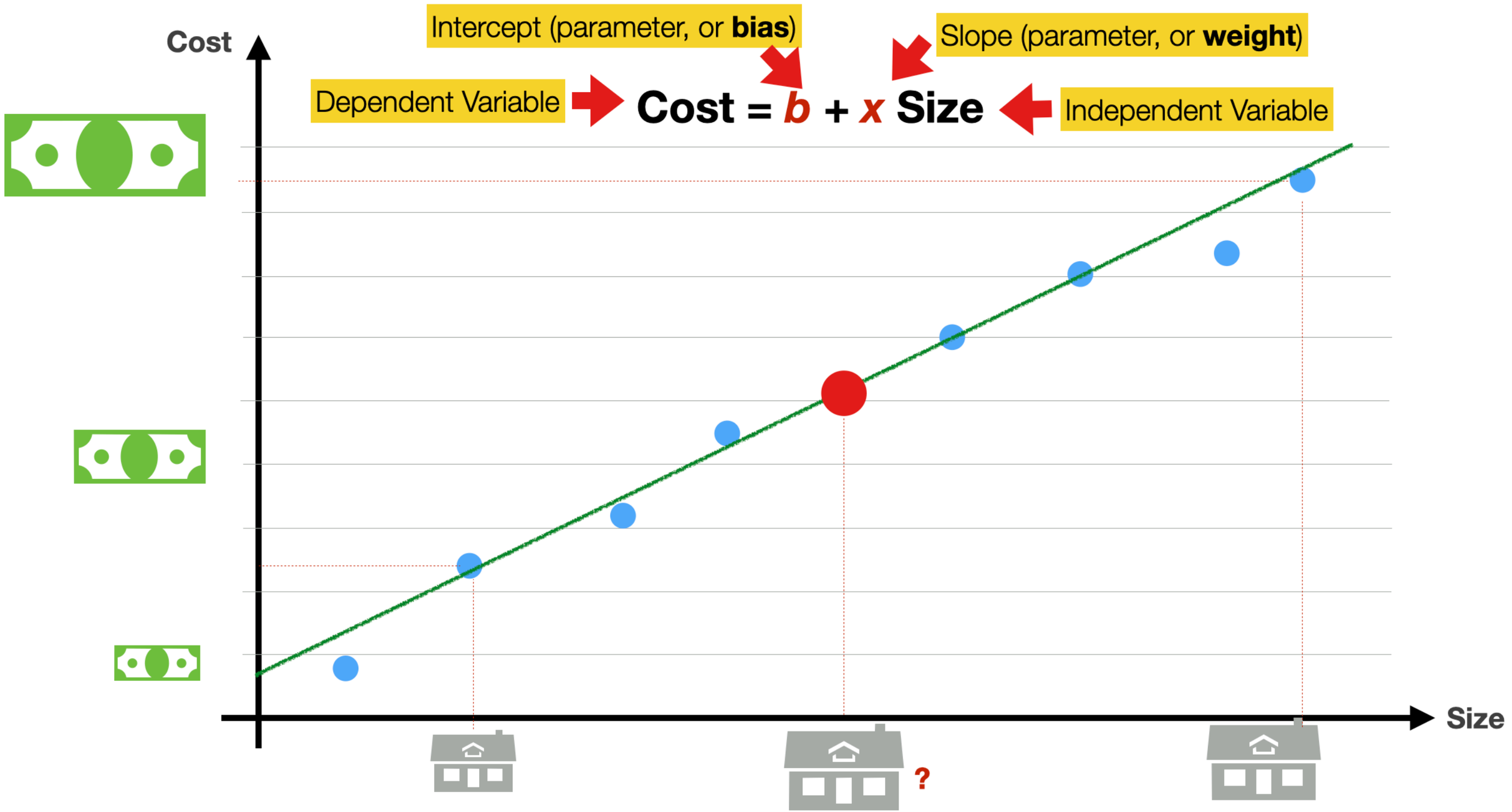
A bit more on regression and classification

And your very first contact with (deep) neural networks

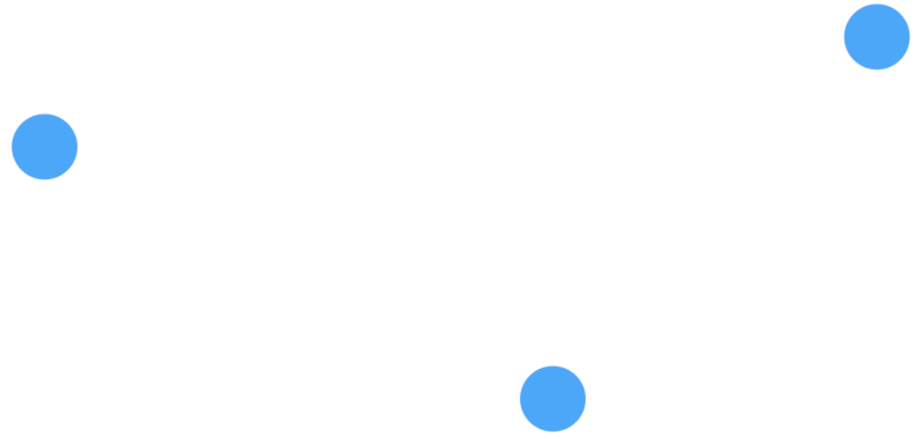
Linear Regression


Cost
Size

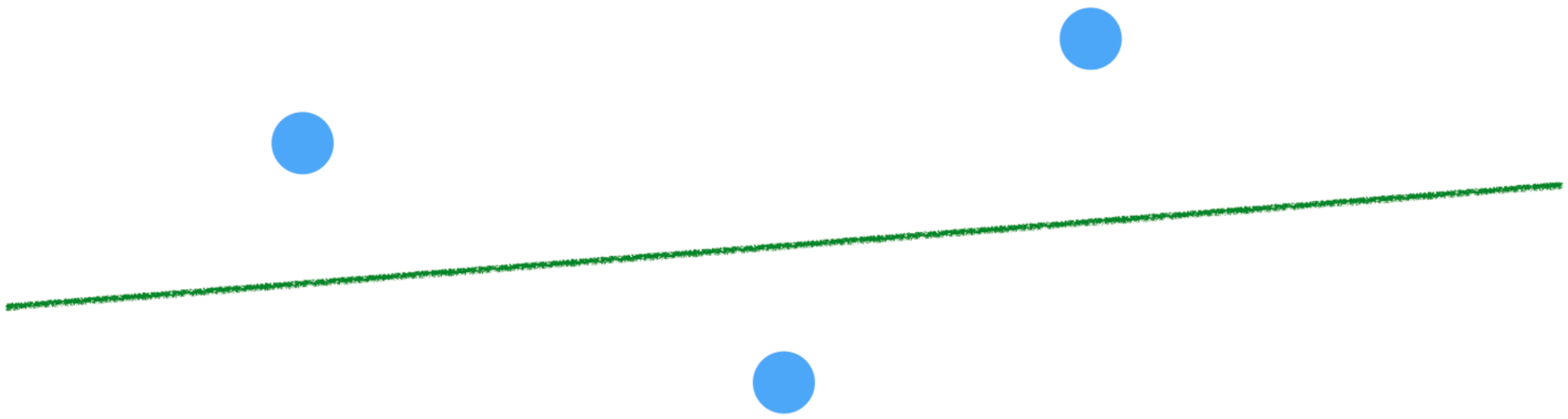

Cost
Size
?


Intercept (parameter, or **bias**)
Slope (parameter, or **weight**)
Dependent Variable
Cost = b + x Size
Independent Variable
Cost
Size
?

Cost = *x* Size



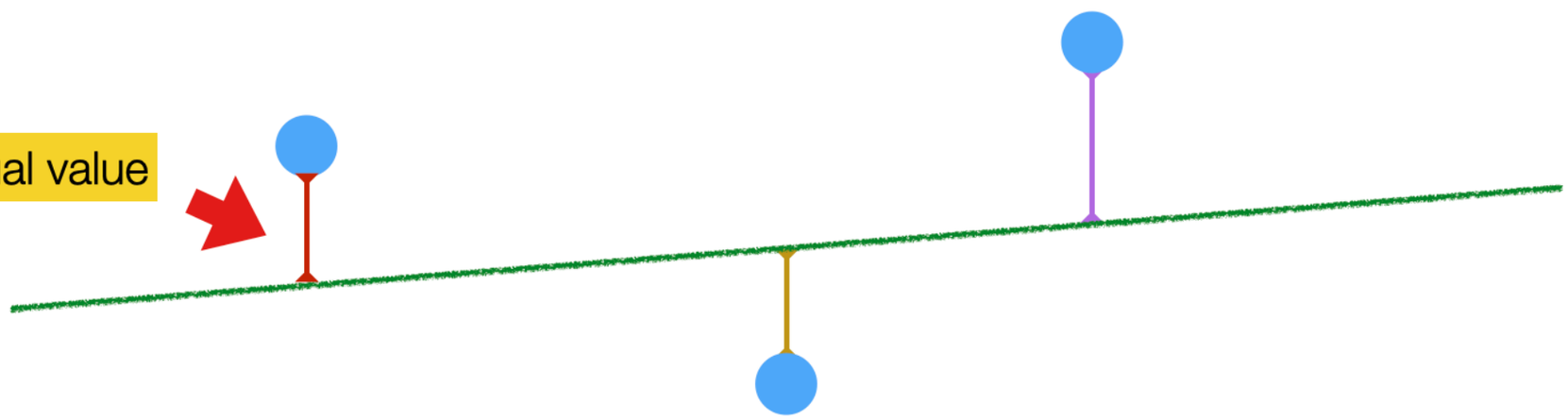
Cost = *x* Size



Cost = *x* Size

predicted value - actual value

Fit_1:

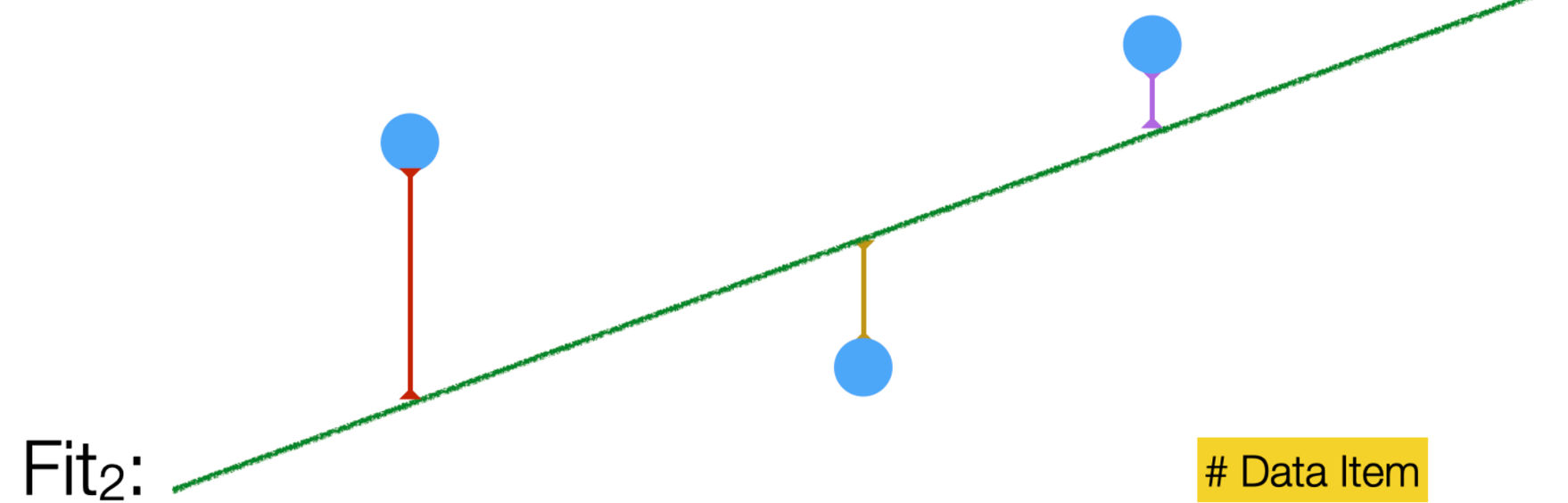


***x* = 0.05**

$Error_1$:

Cost = *x* Size

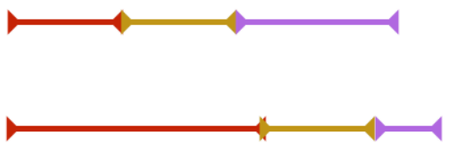
***x* = 0.2**



Fit$_2$:

Error$_1$:

Error$_2$:



Data Item

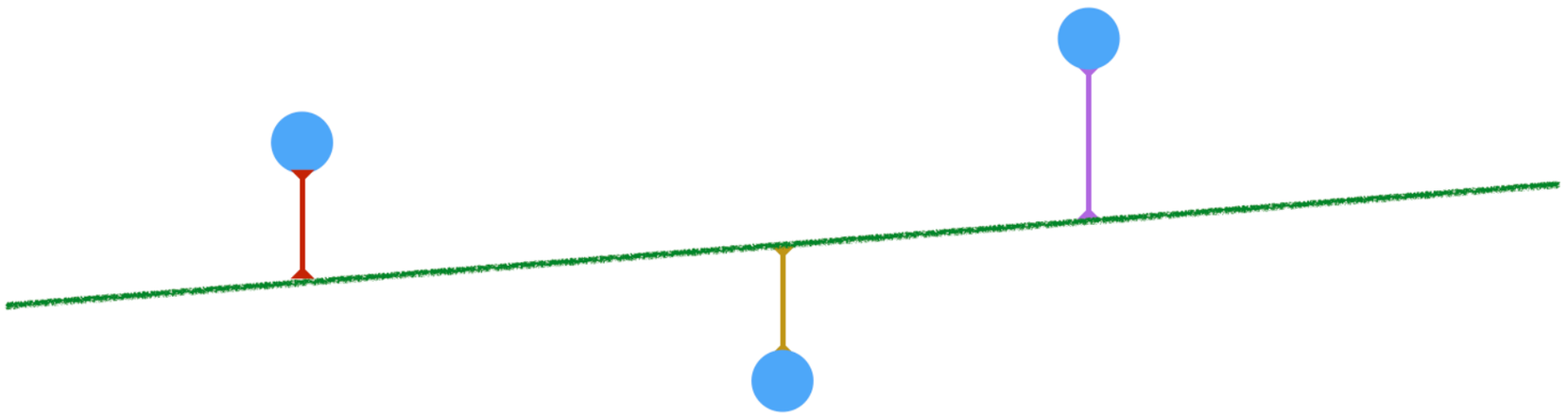
$$Error = \frac{1}{2d}\sum_{i=1}^{d}(Prediction_i - Value_i)^2$$

Mean Squared Error (MSE)

Current data item

Cost = *x* Size

Fit_1:



***b* = 0.05**

$Error_1$:

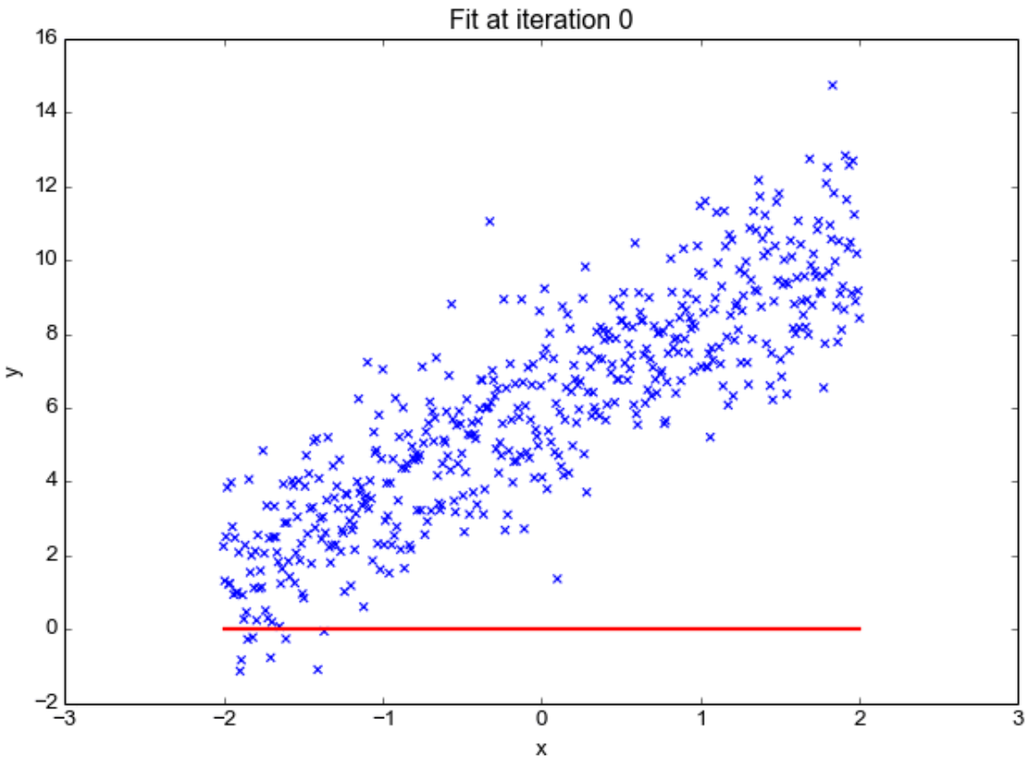
$Error_2$:

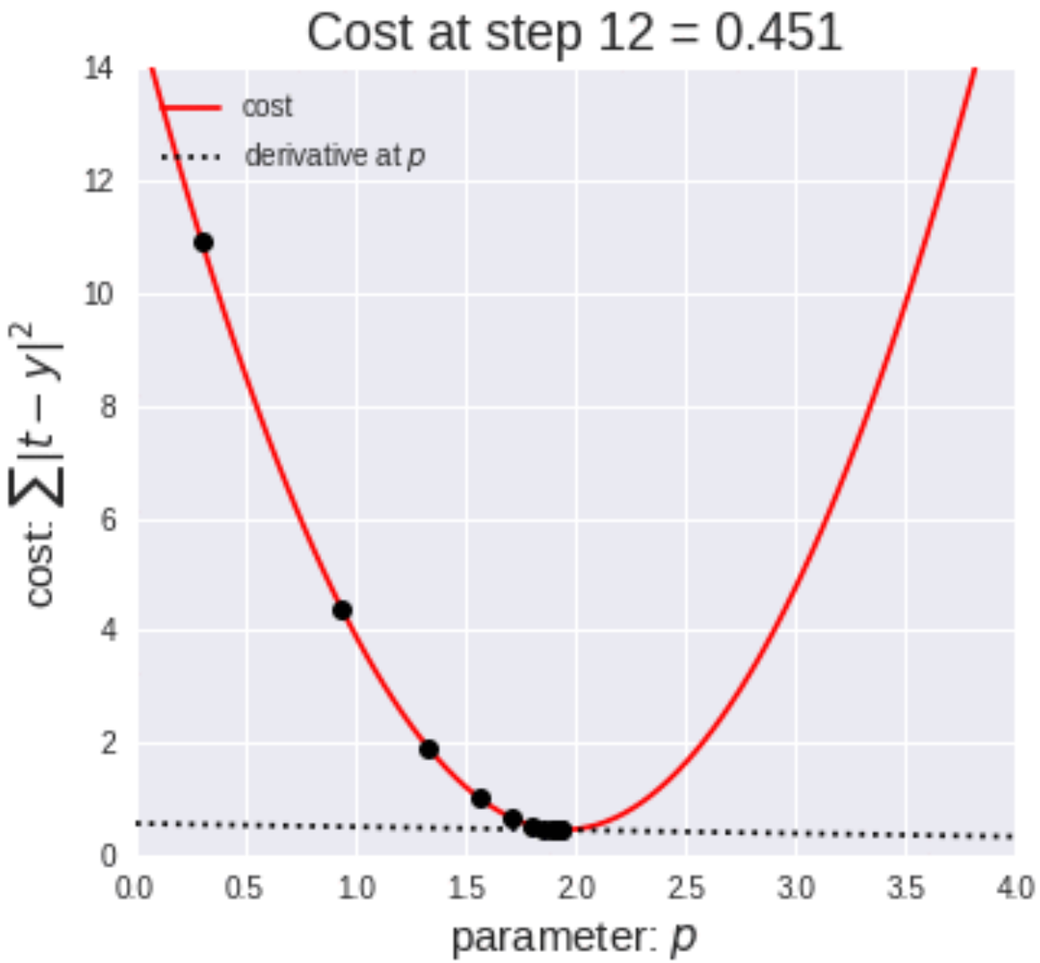
Fit_1 is a better fit on the training data than Fit_2

We select ***b* = 0.05**

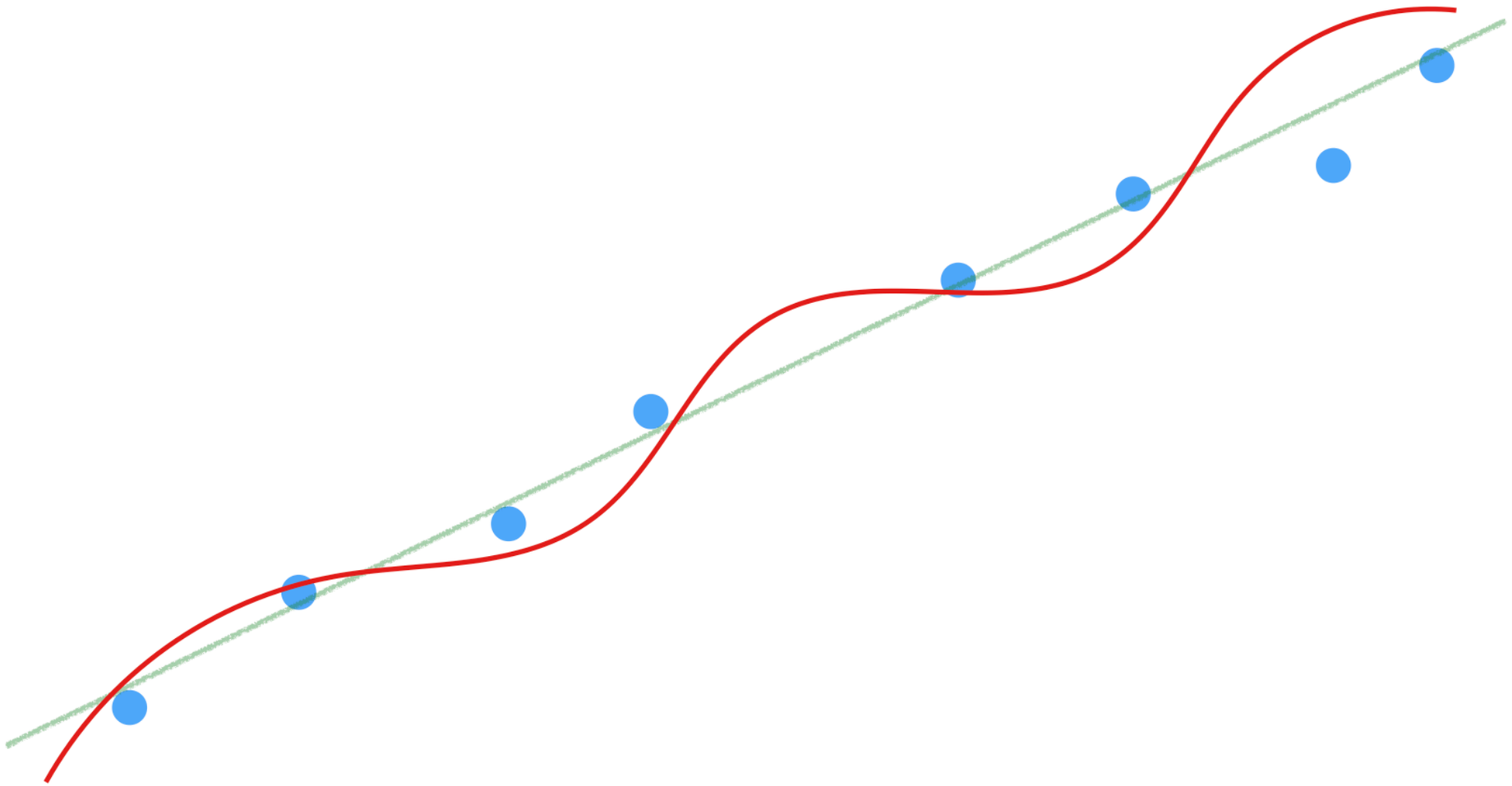
Finding the best parameter values

- Training the model
  - **Gradient Descent**: an algorithm to find the minimum point of a function
  - **Hyper parameters**: parameters of the Gradient Descent
    - *Learning Rate*: speed of descent
    - *Epochs*: max number of steps

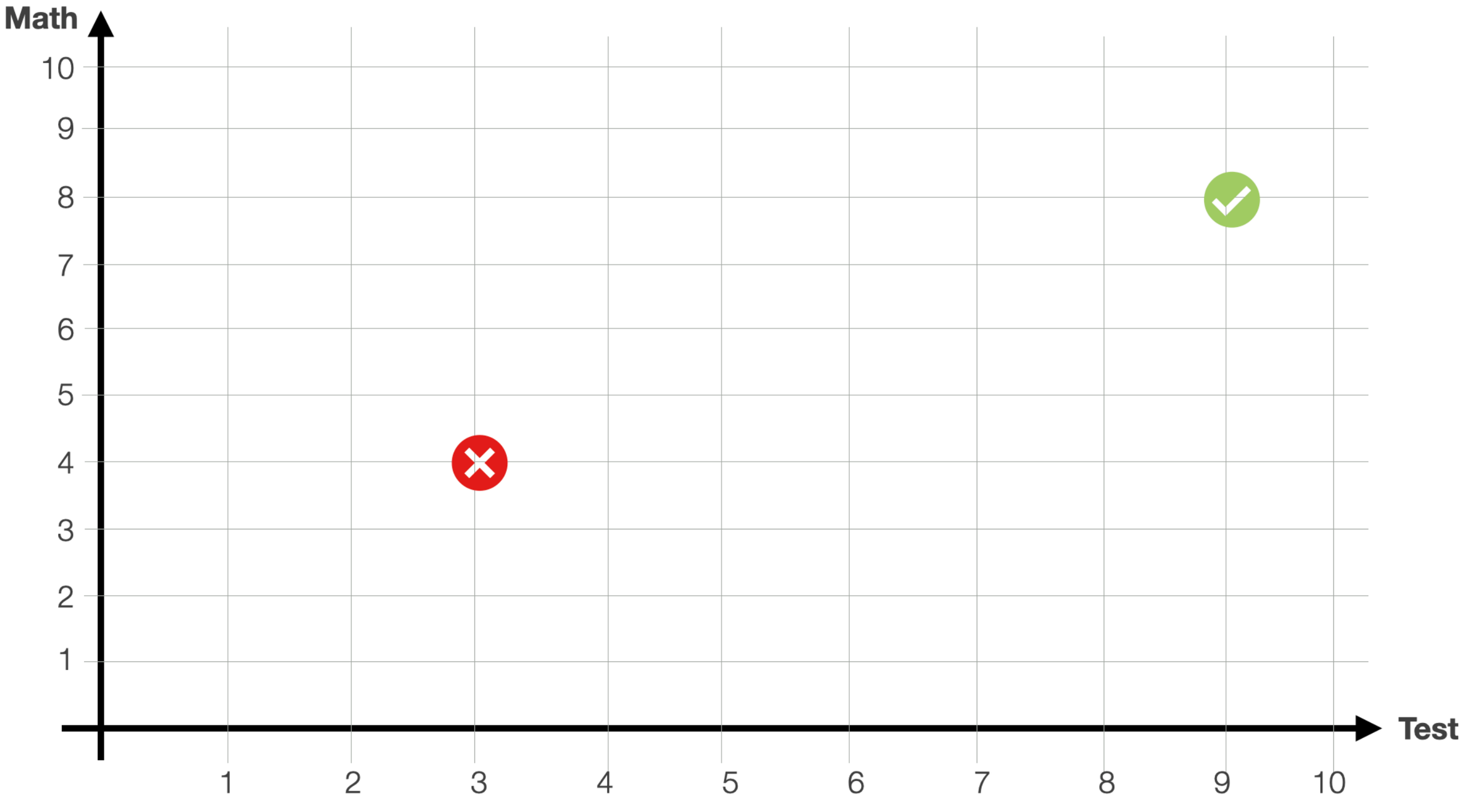

Fit at iteration 0
x
y

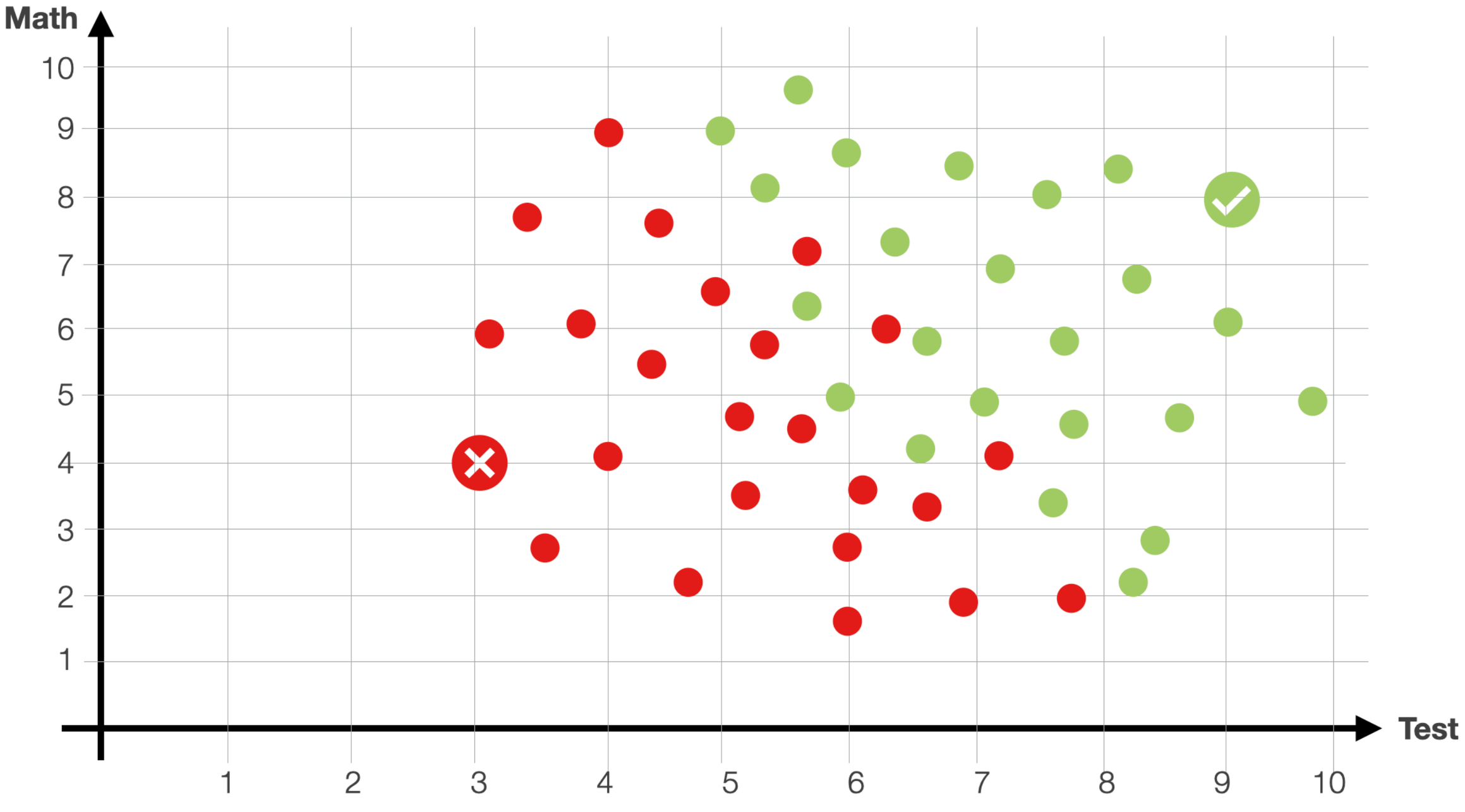

Cost at step 12 = 0.451
cost
derivative at p
cost: ∑|t − y|²
parameter: p

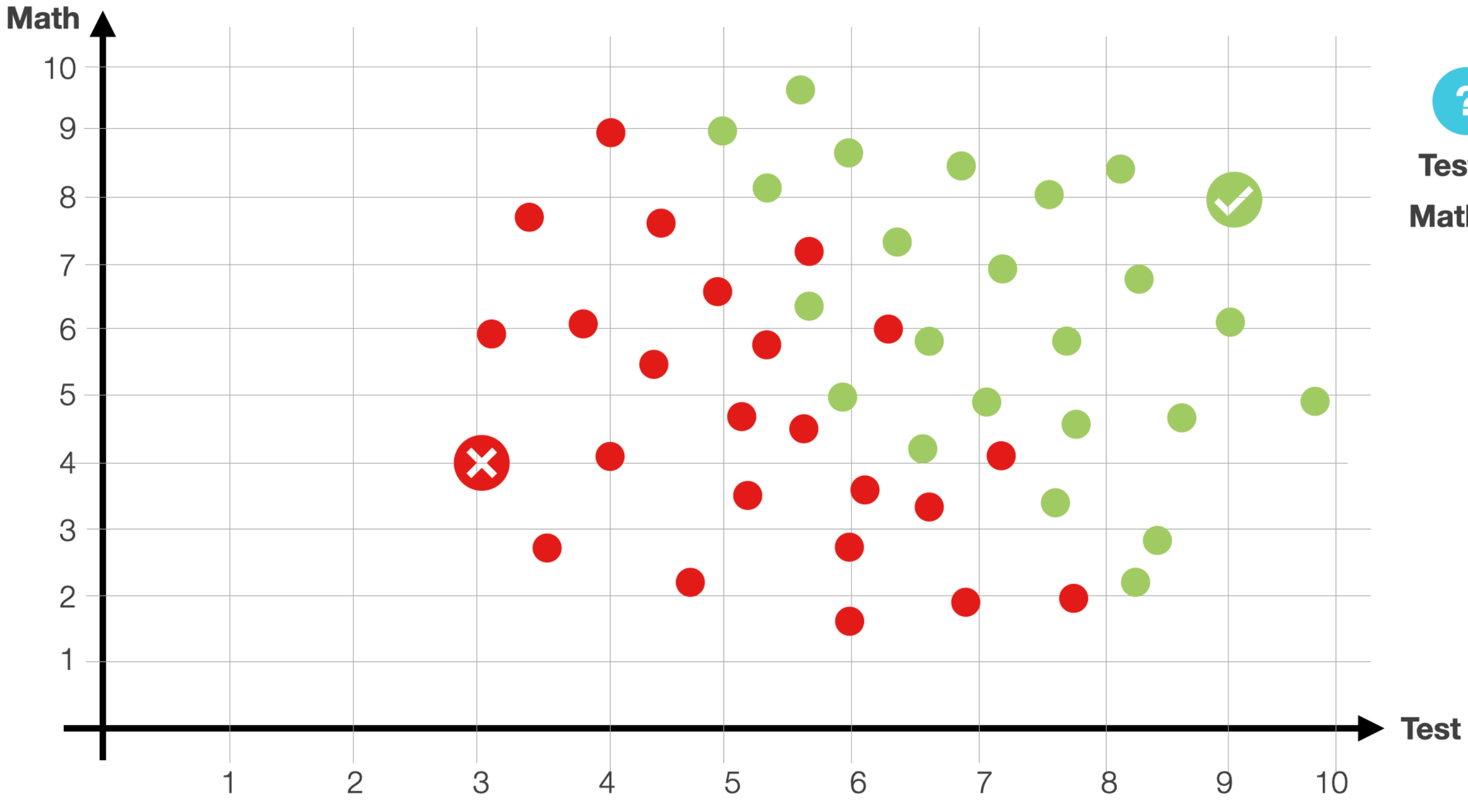
Nth degree polynomial ➡ **Cost = b + x_1 Size + x_2 Size^2 + ... + x_n Size^n**



Classification


Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10


Math
10
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
10
Test

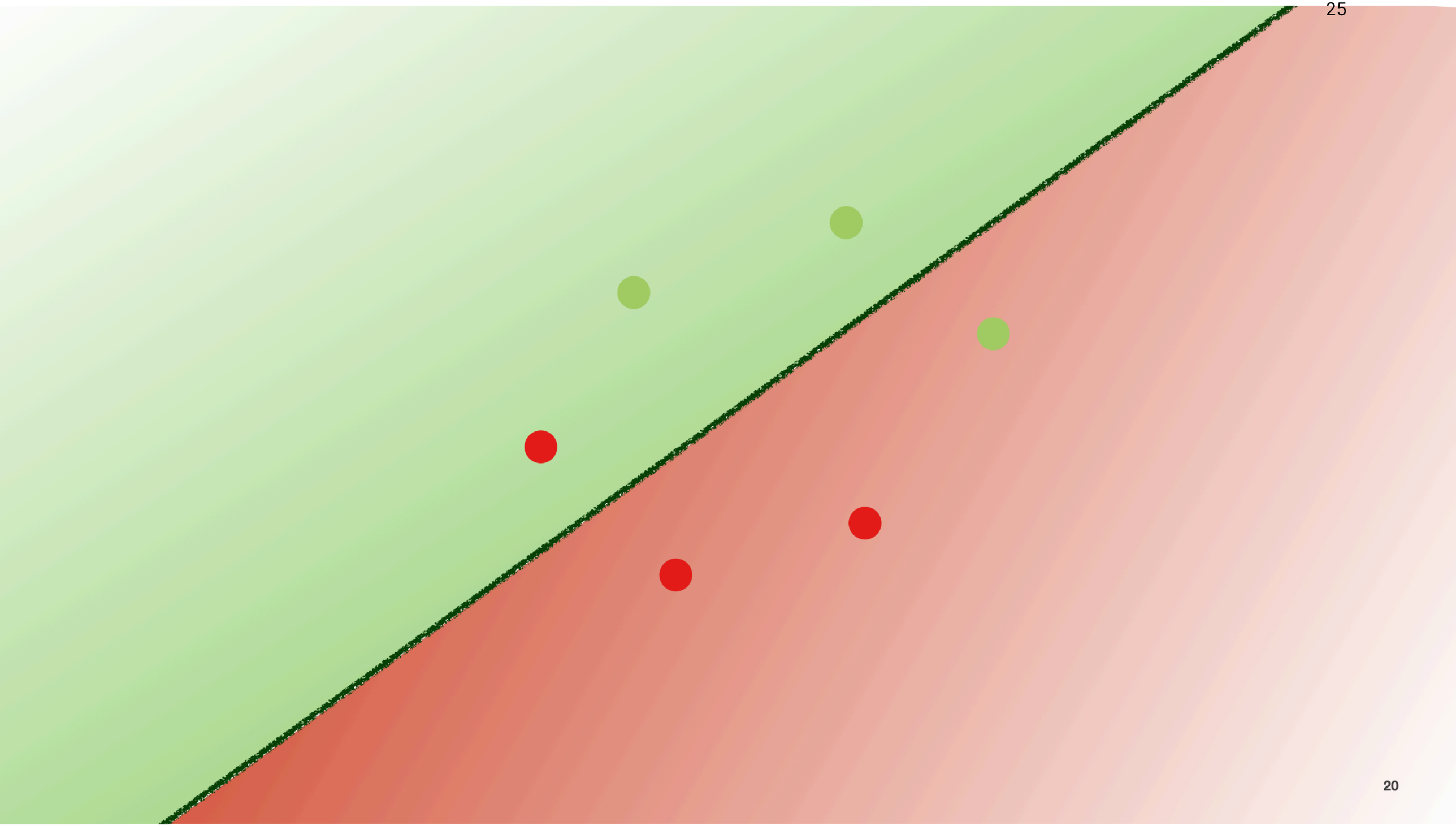


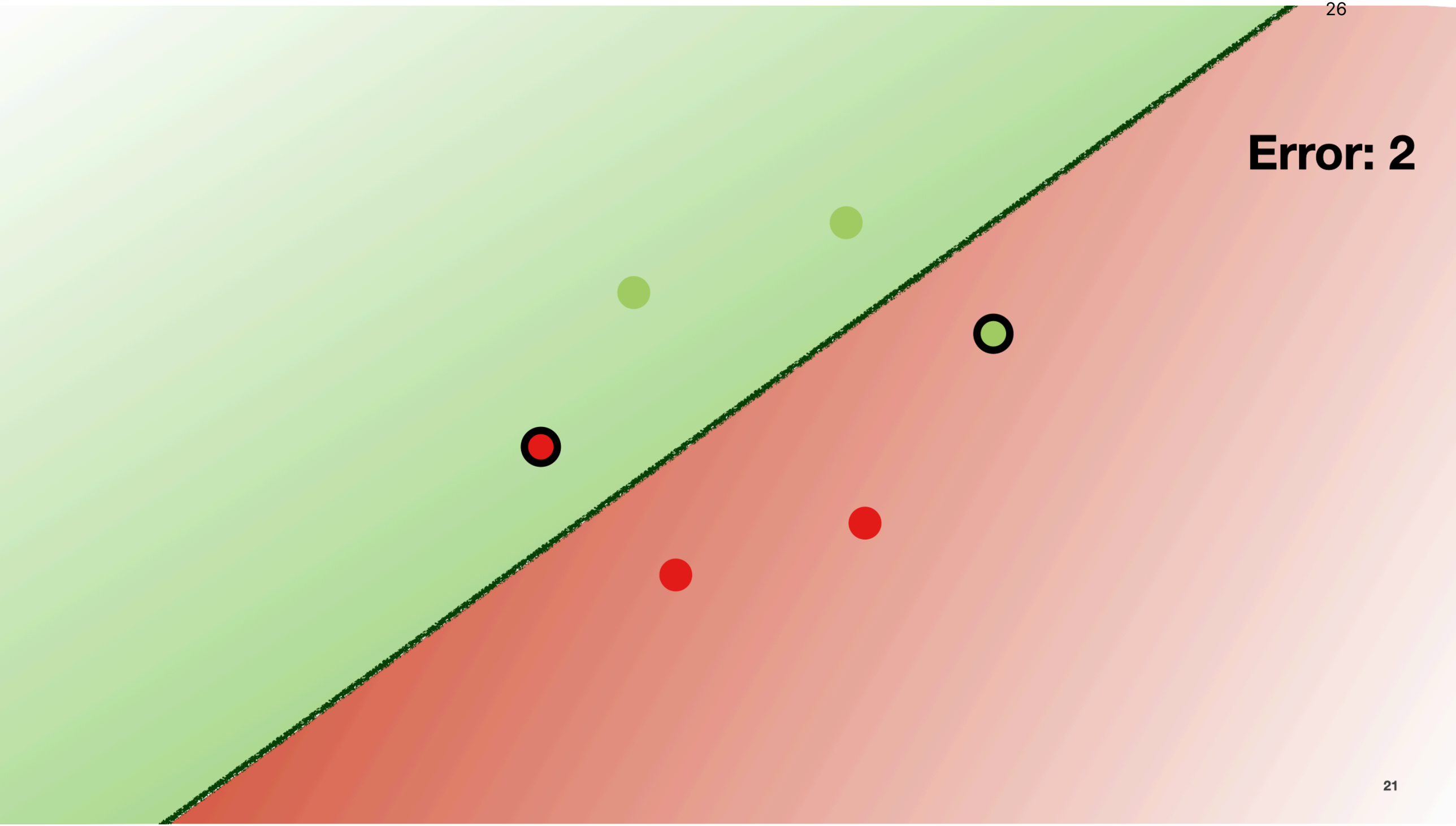
?

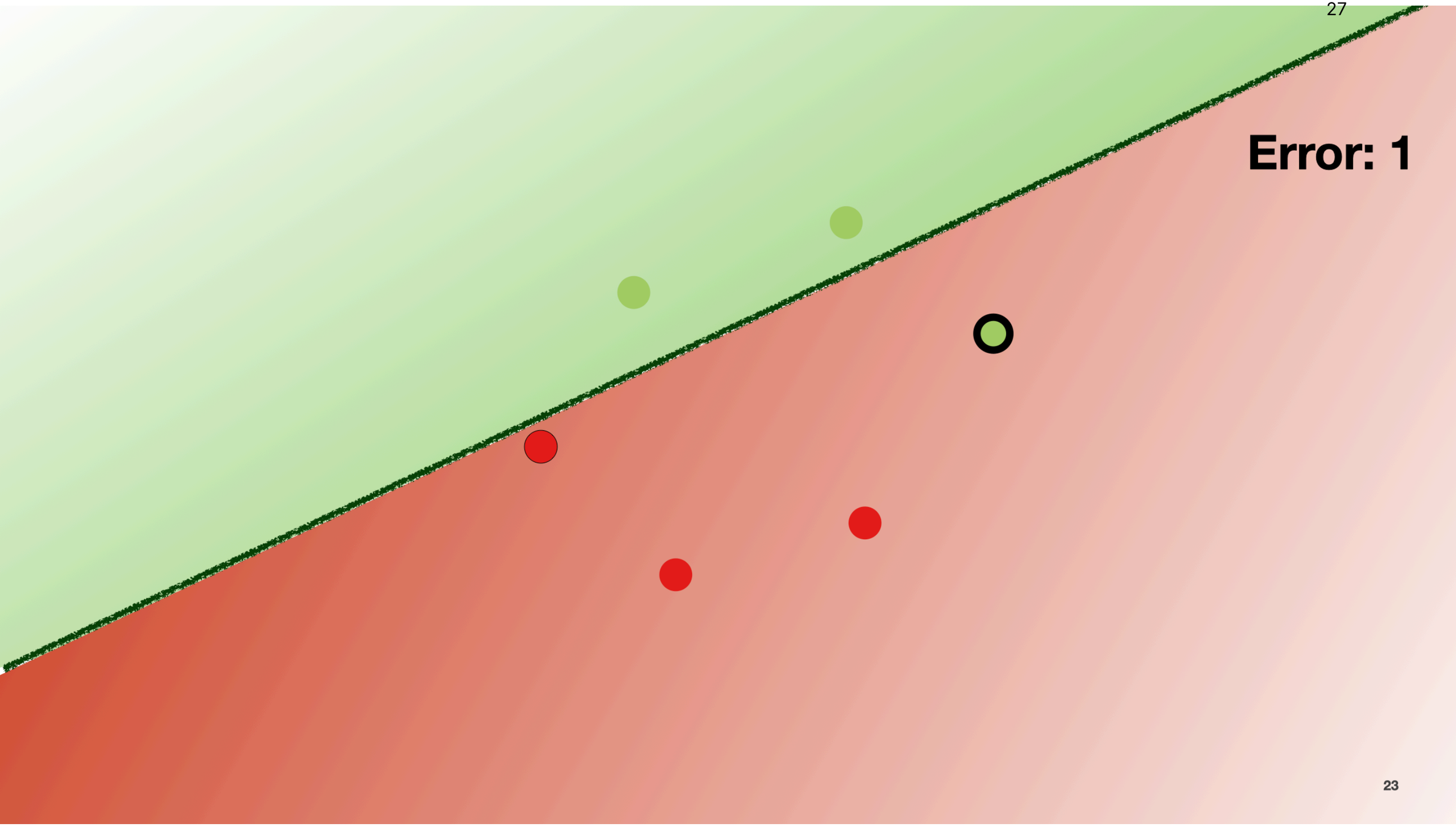
Test: 7

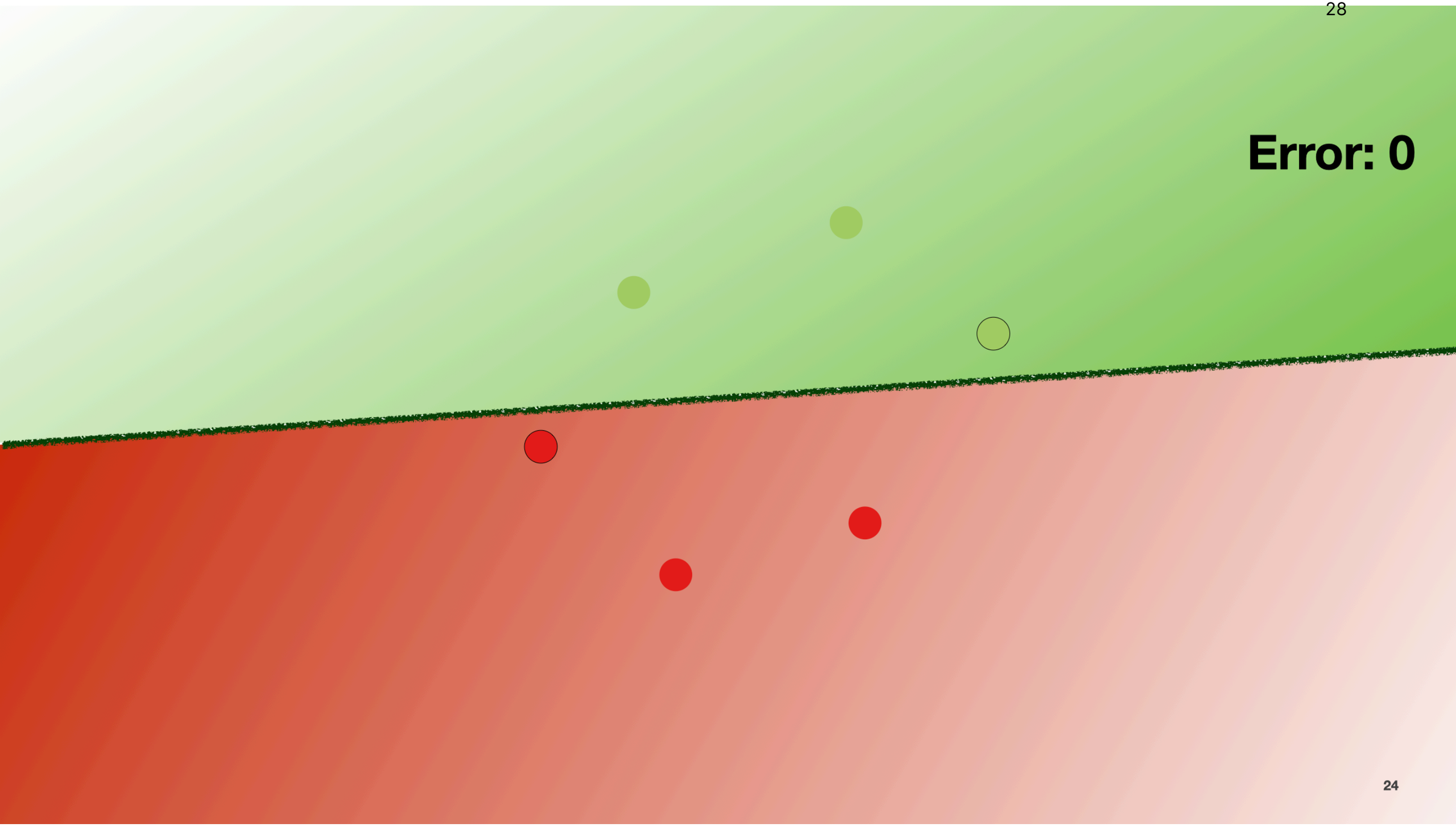
Math: 6

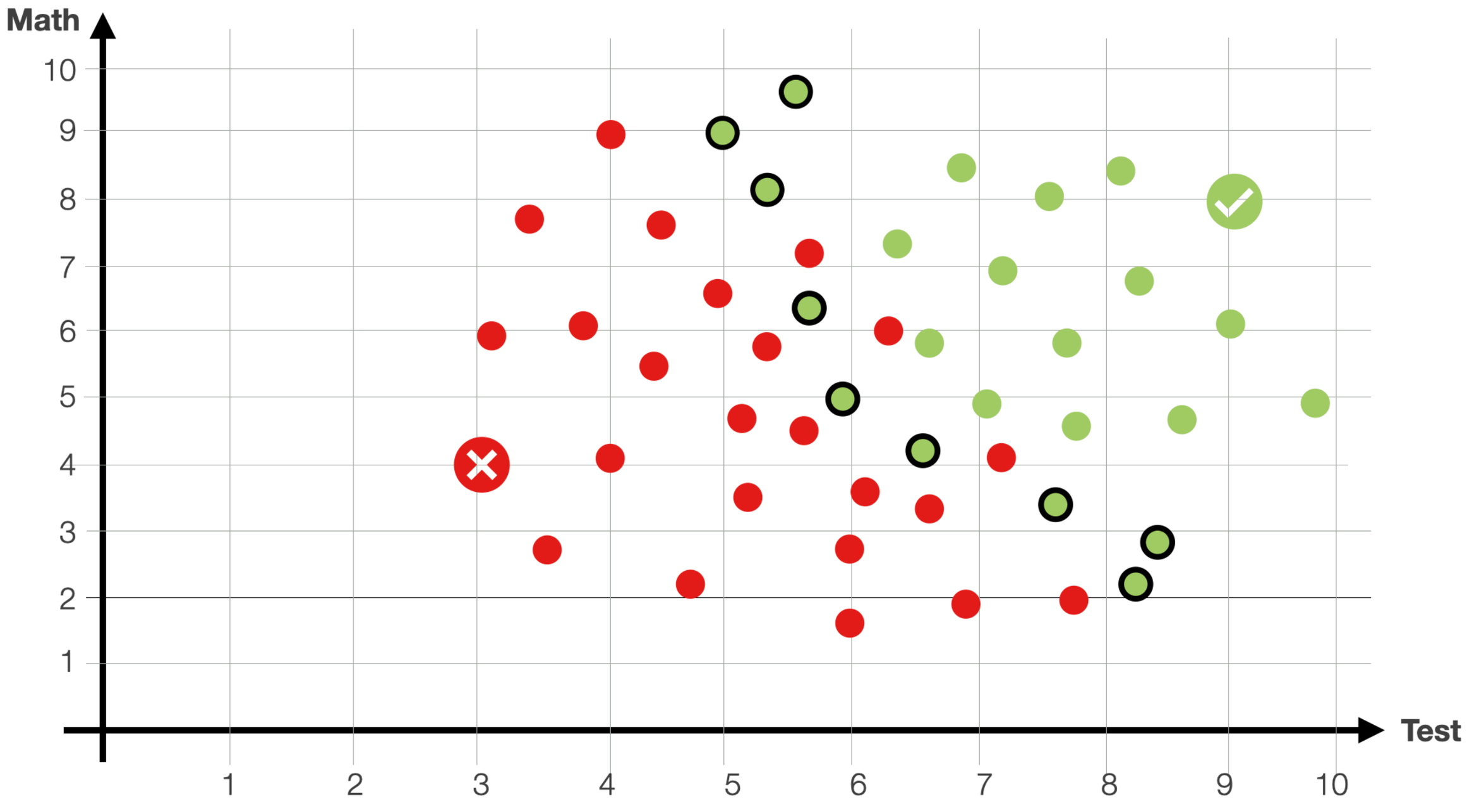

Logistic Regression
Math
10
9
8
7
6
5
4
3
2
1
?
Test
1
2
3
4
5
6
7
8
9
10

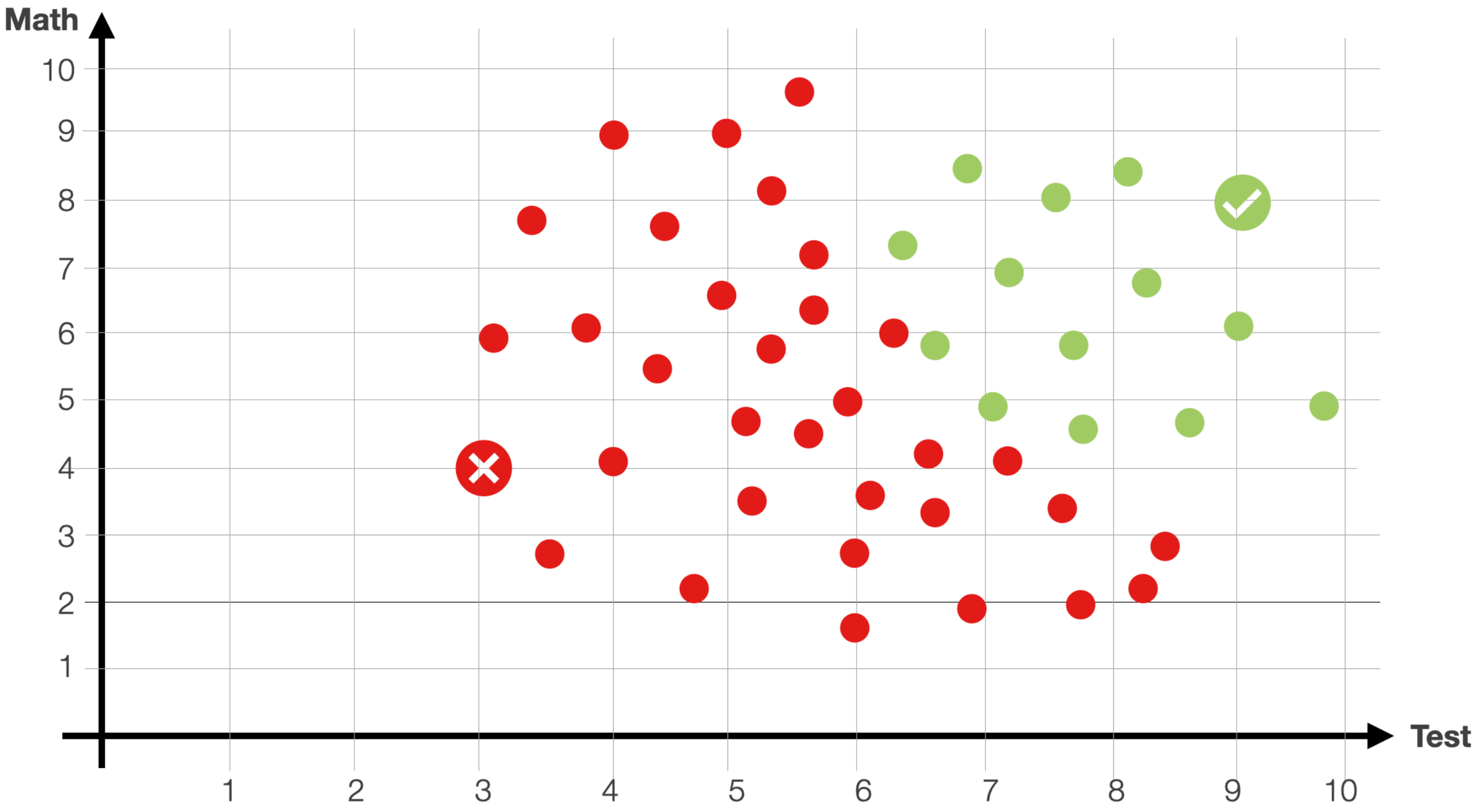


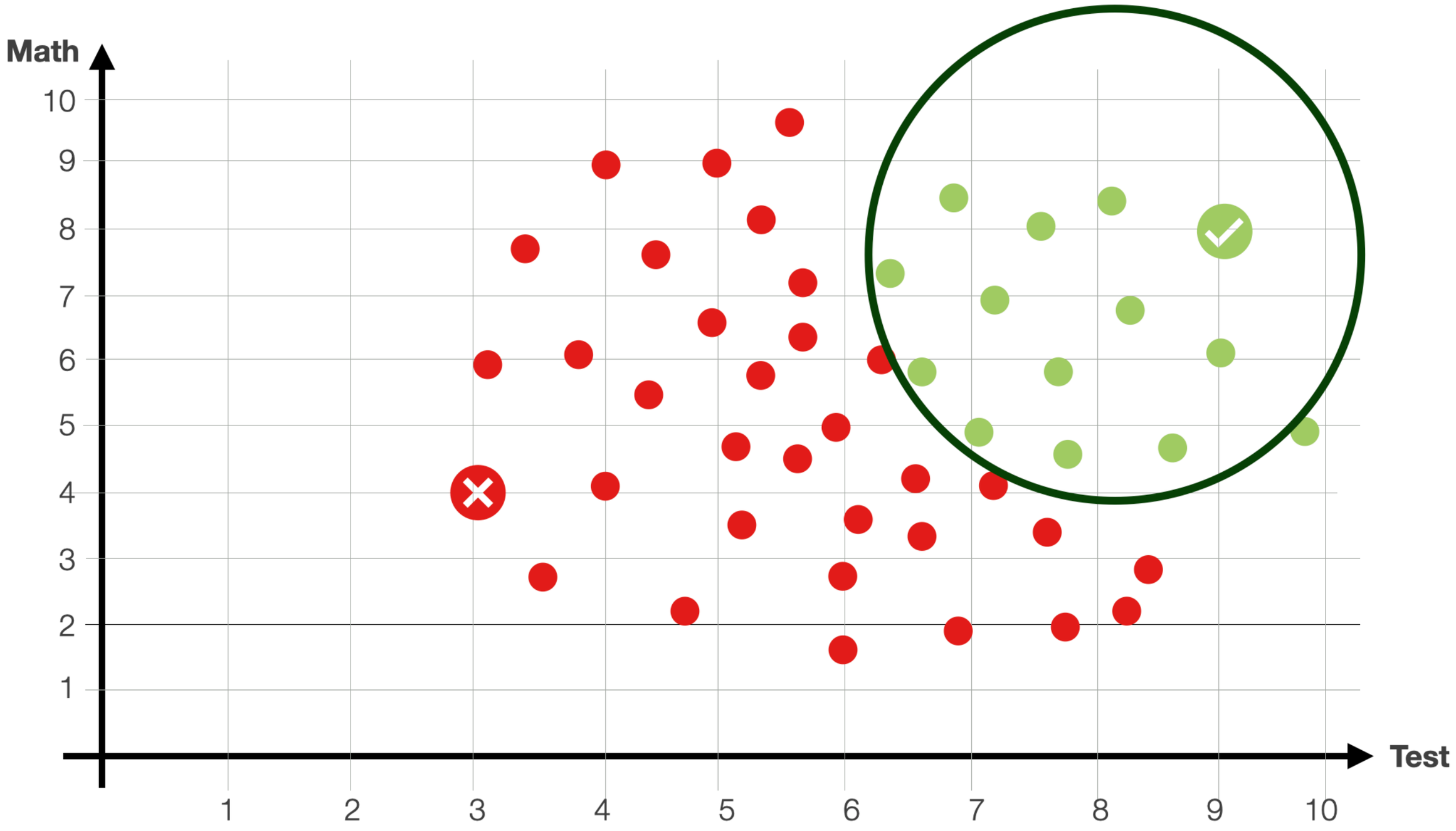

Error: 2

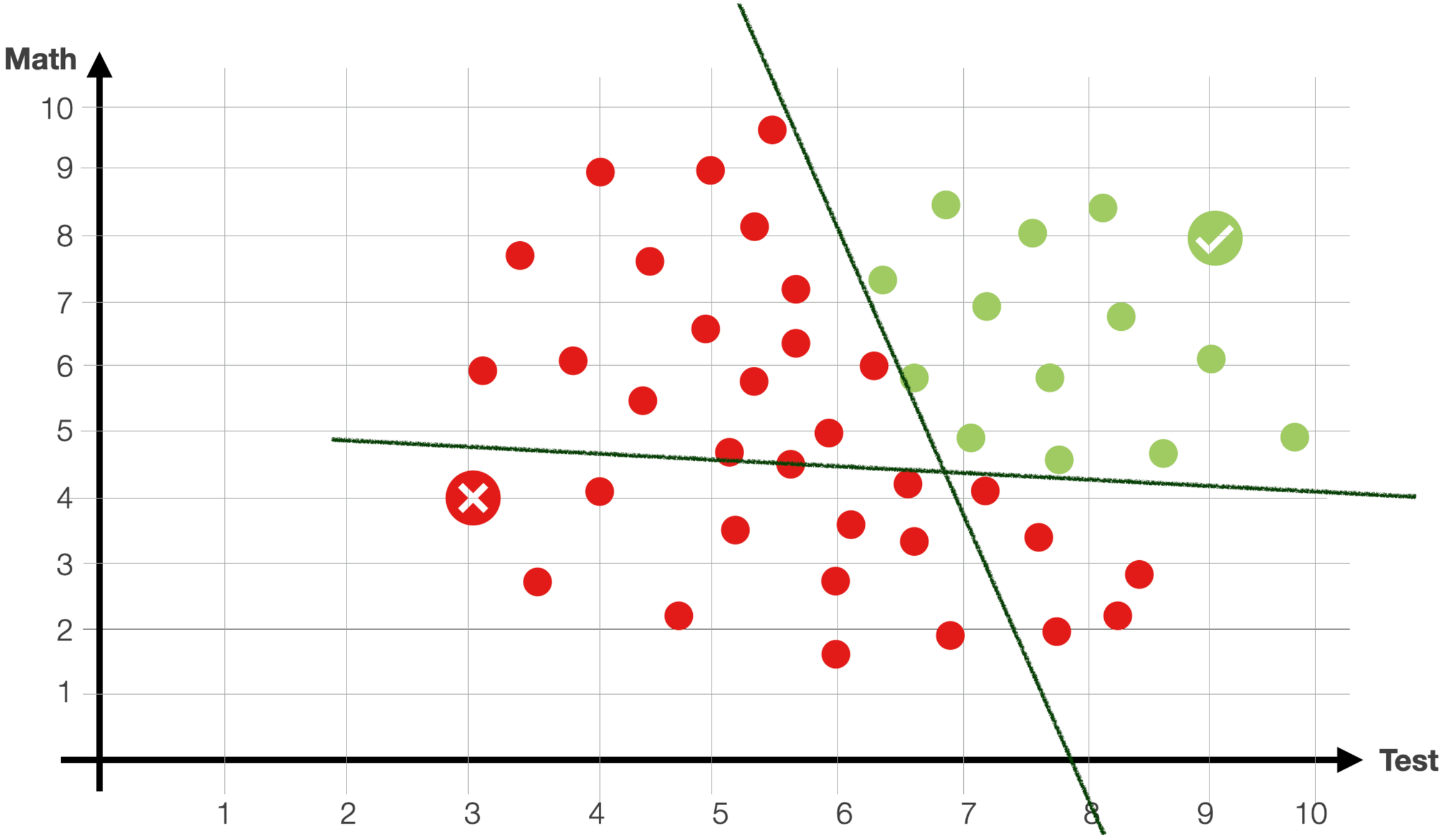

Error: 1

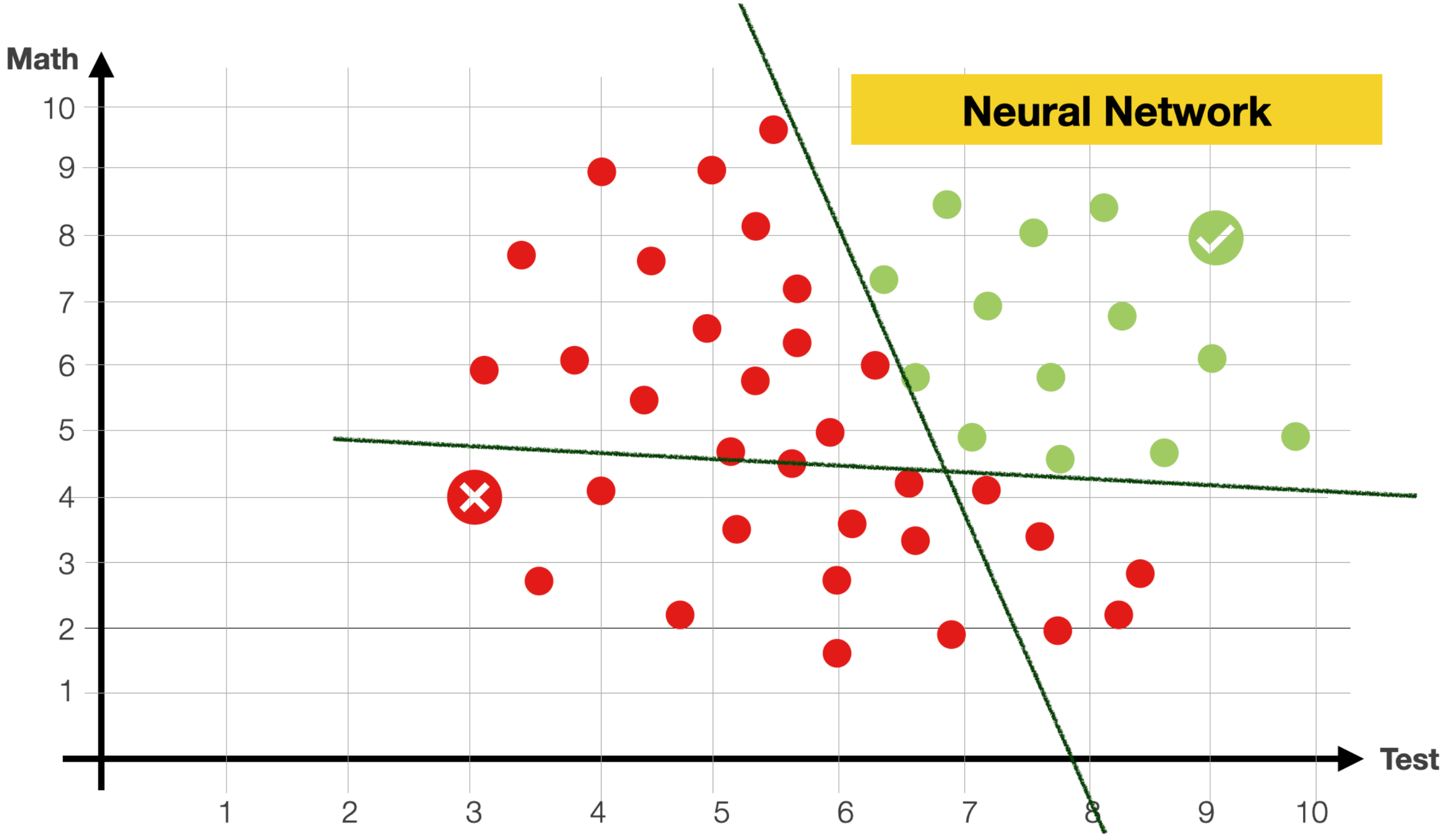

Error: 0

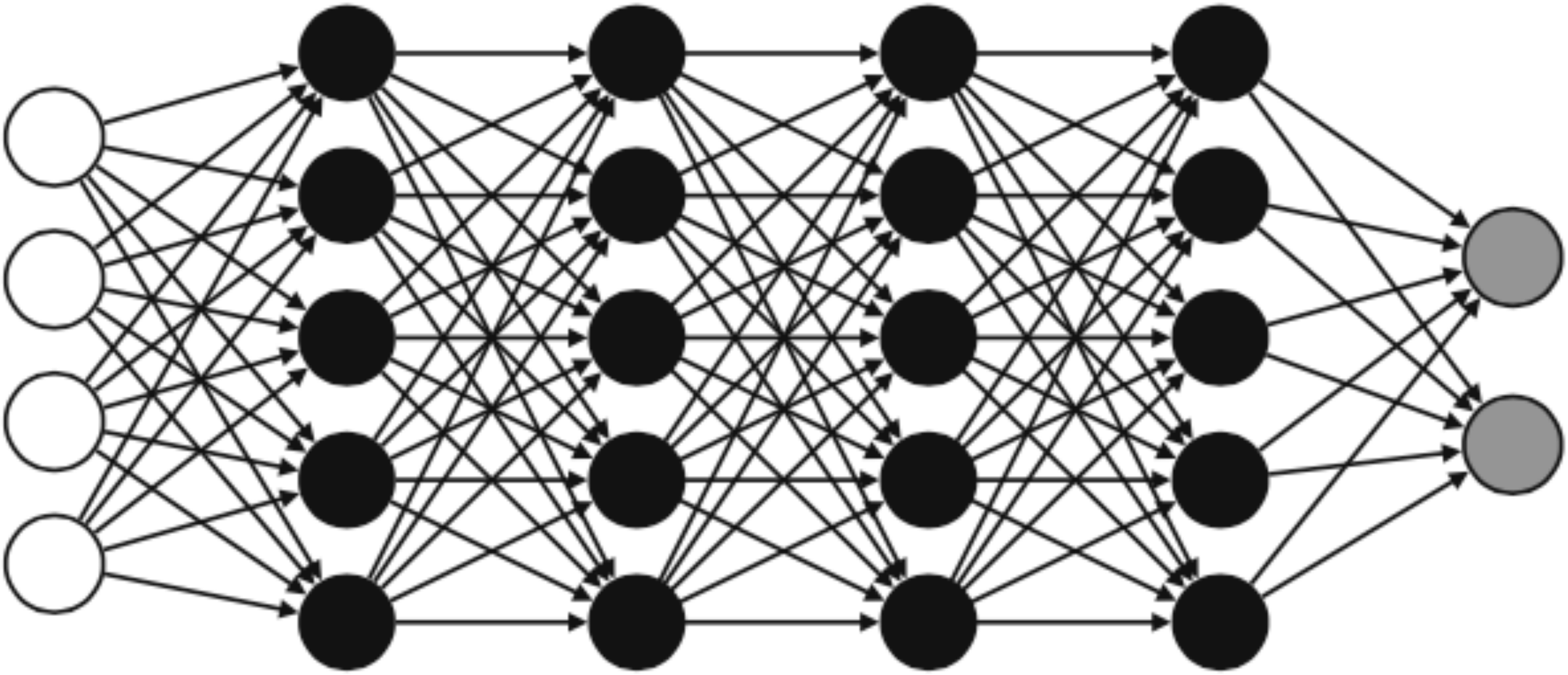

Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10

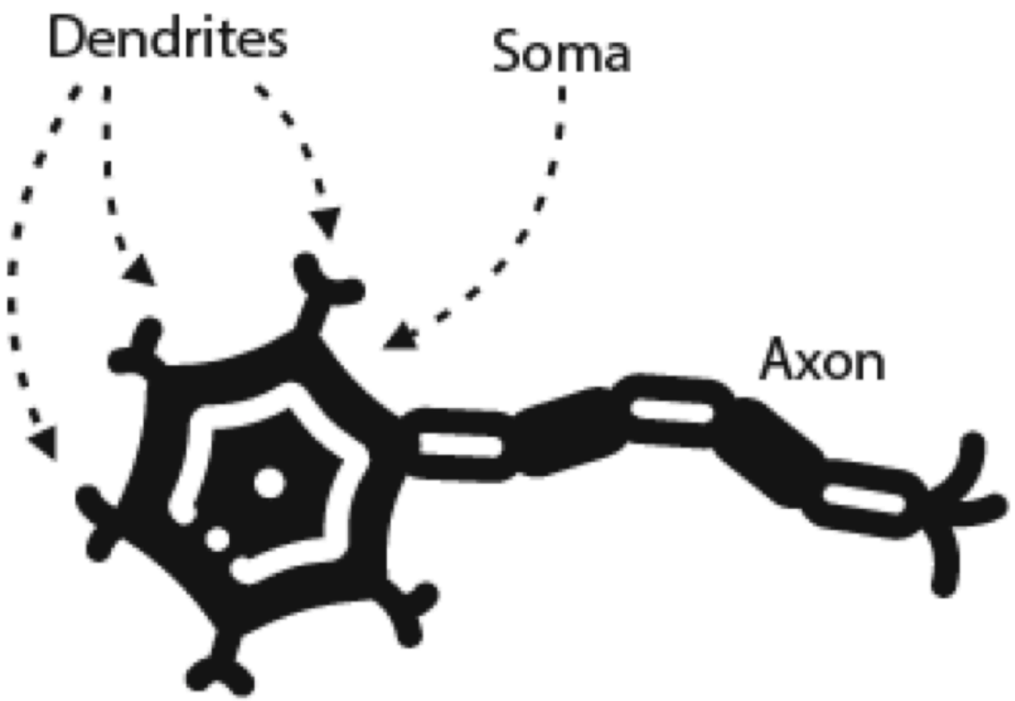

Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10


Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10

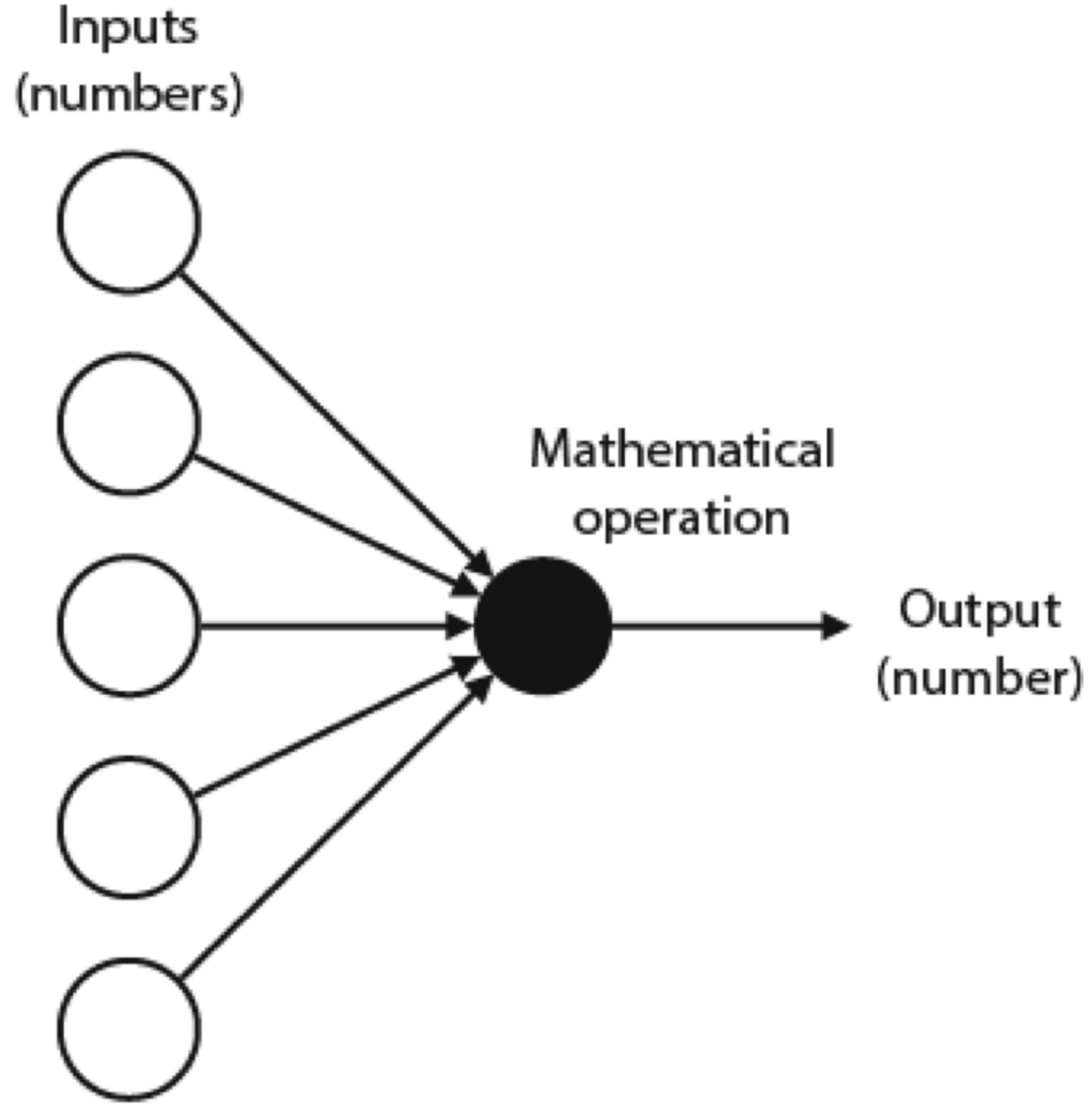

Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10


Neural Network
Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10

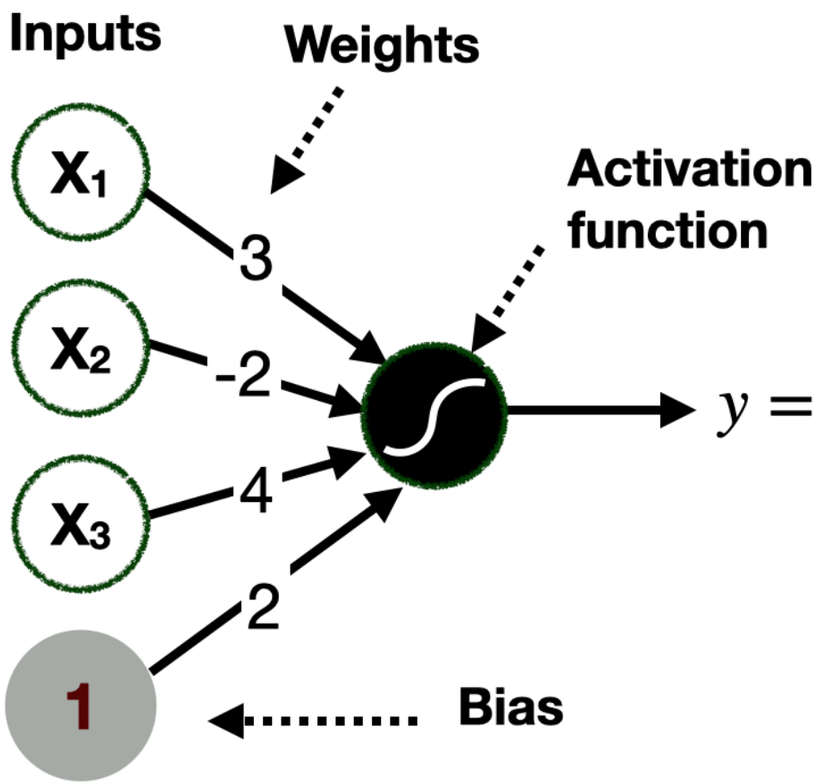




Neuron



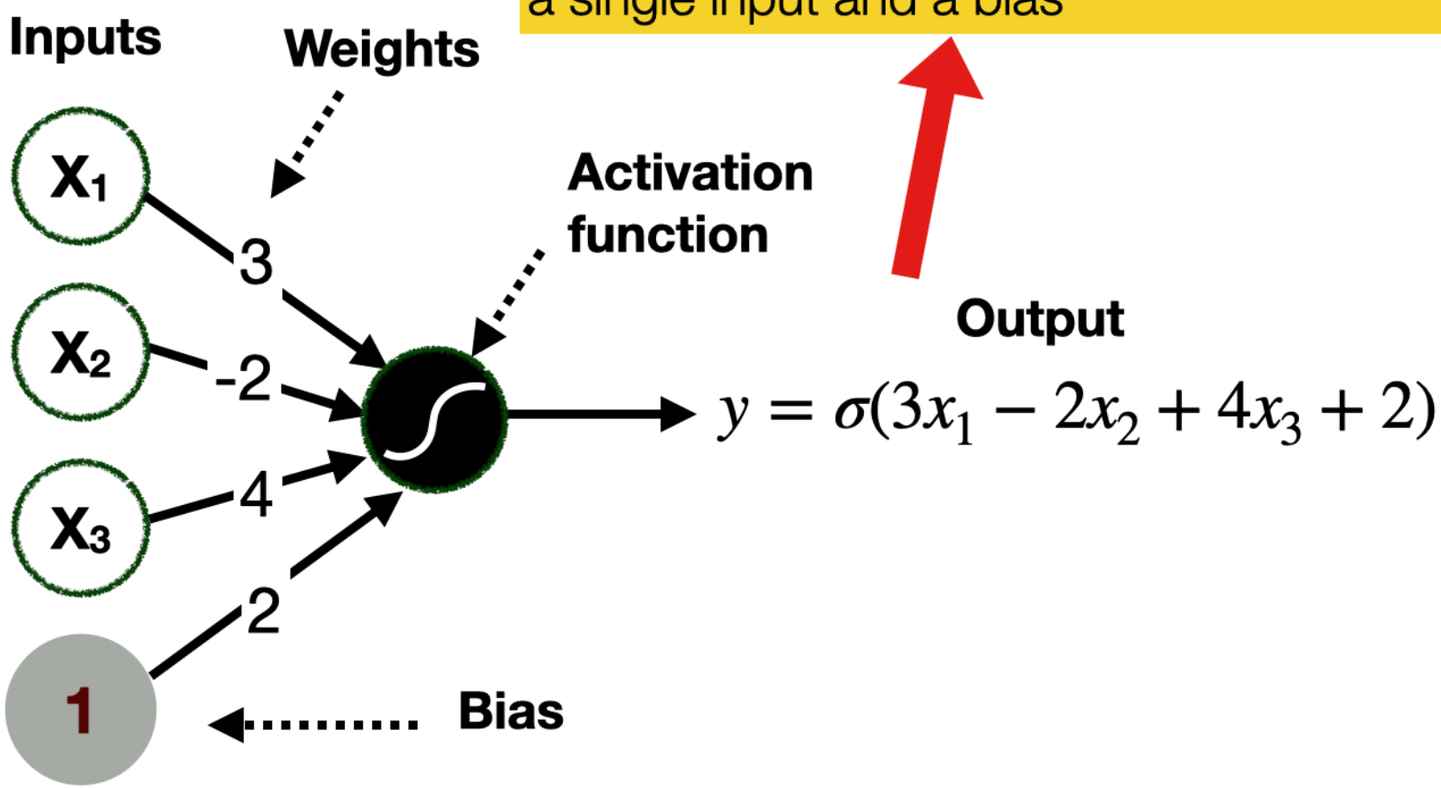
Perceptron



$$y = \sigma(3x_1 - 2x_2 + 4x_3 + 2)$$

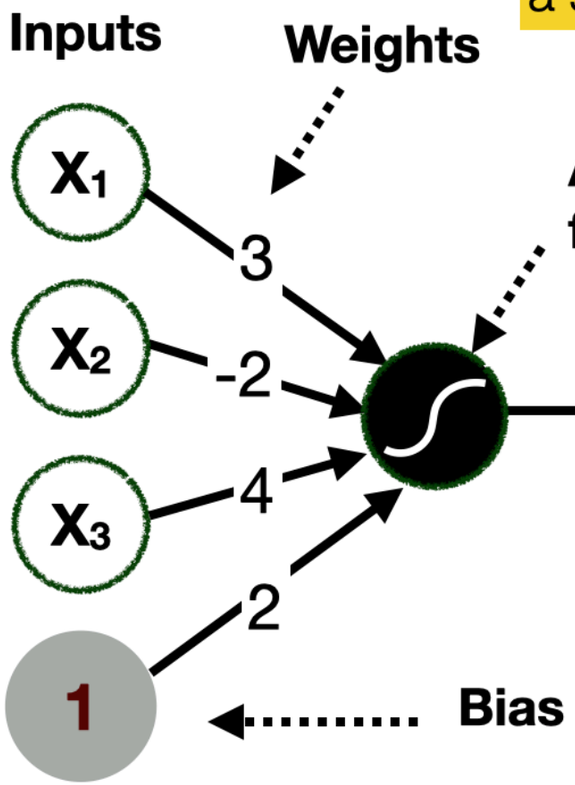
Cost = *a* + *b* Size

Remember our linear regression function?
It can be represented with a perceptron having a single input and a bias



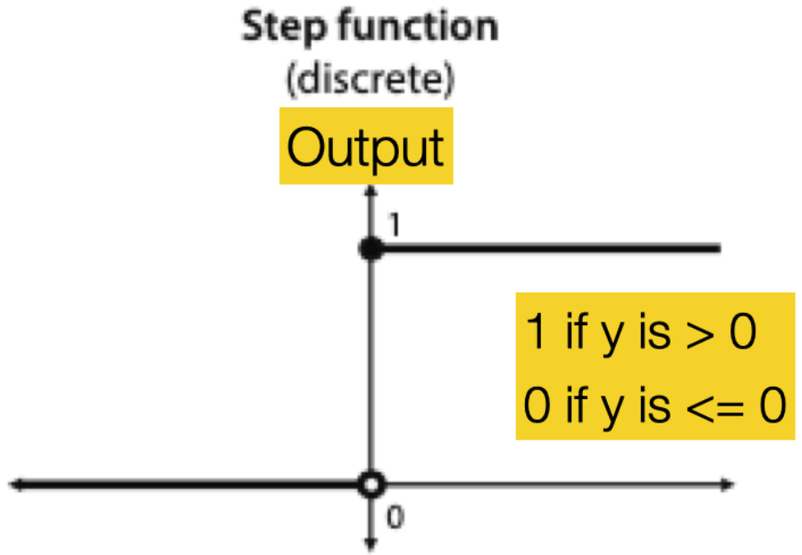
Cost = *a* + *b* Size

Remember our linear regression function?
It can be represented with a perceptron having a single input and a bias

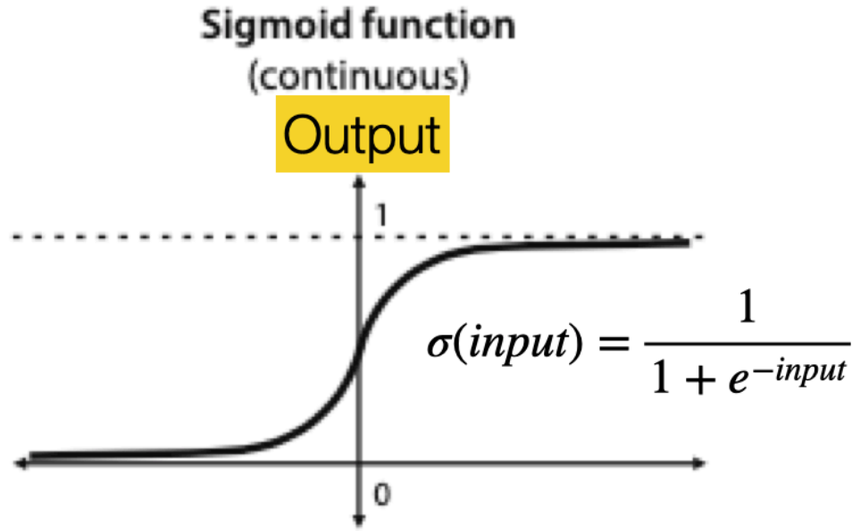


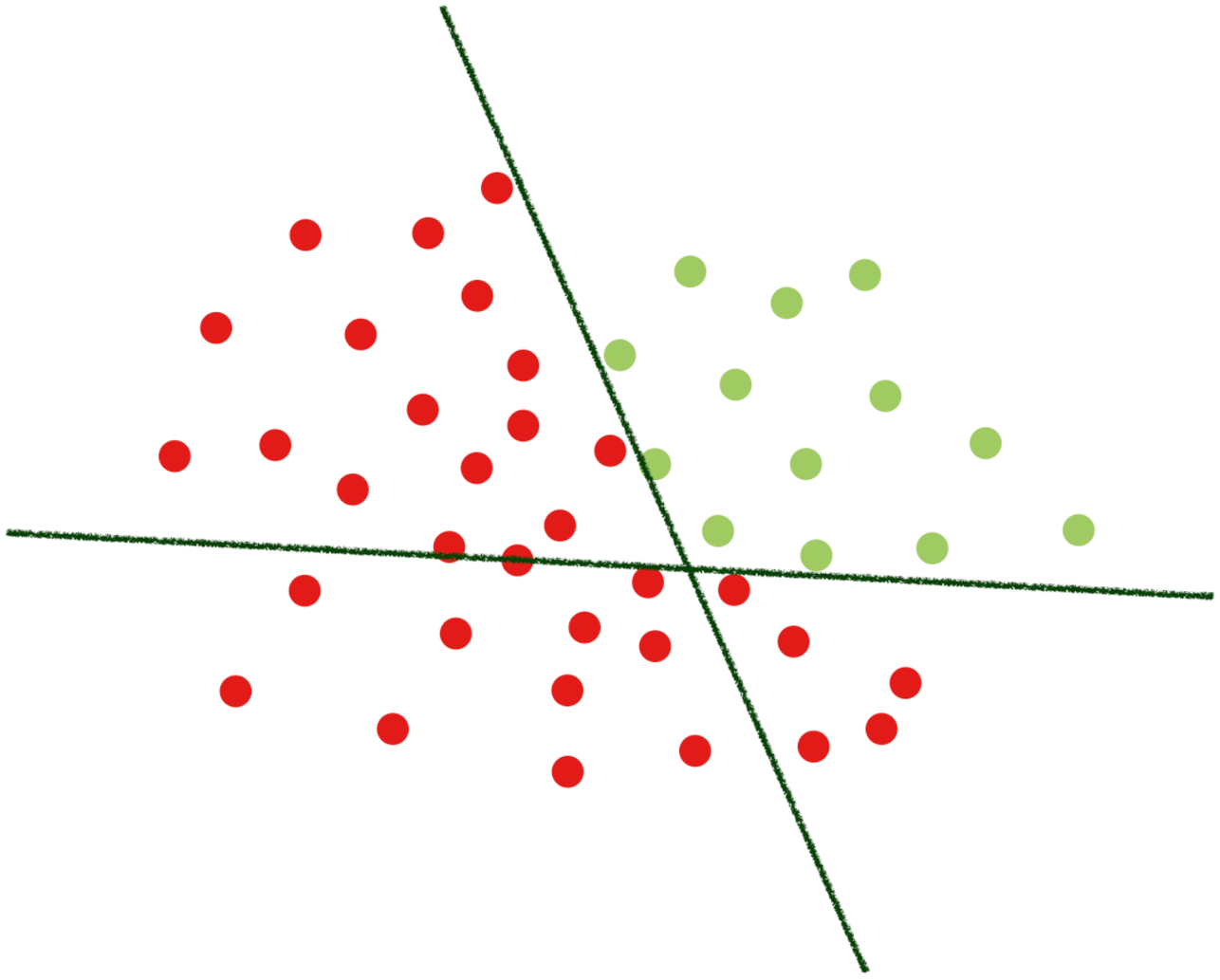
$$y = \sigma(3x_1 - 2x_2 + 4x_3 + 2)$$

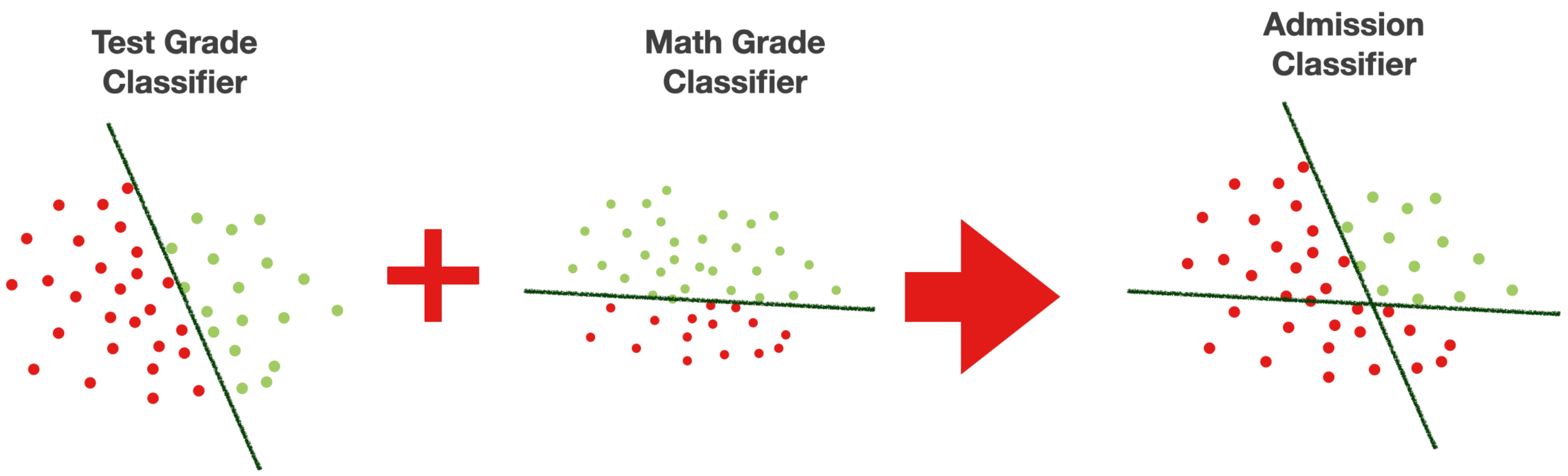
Input (weighted sum of values)

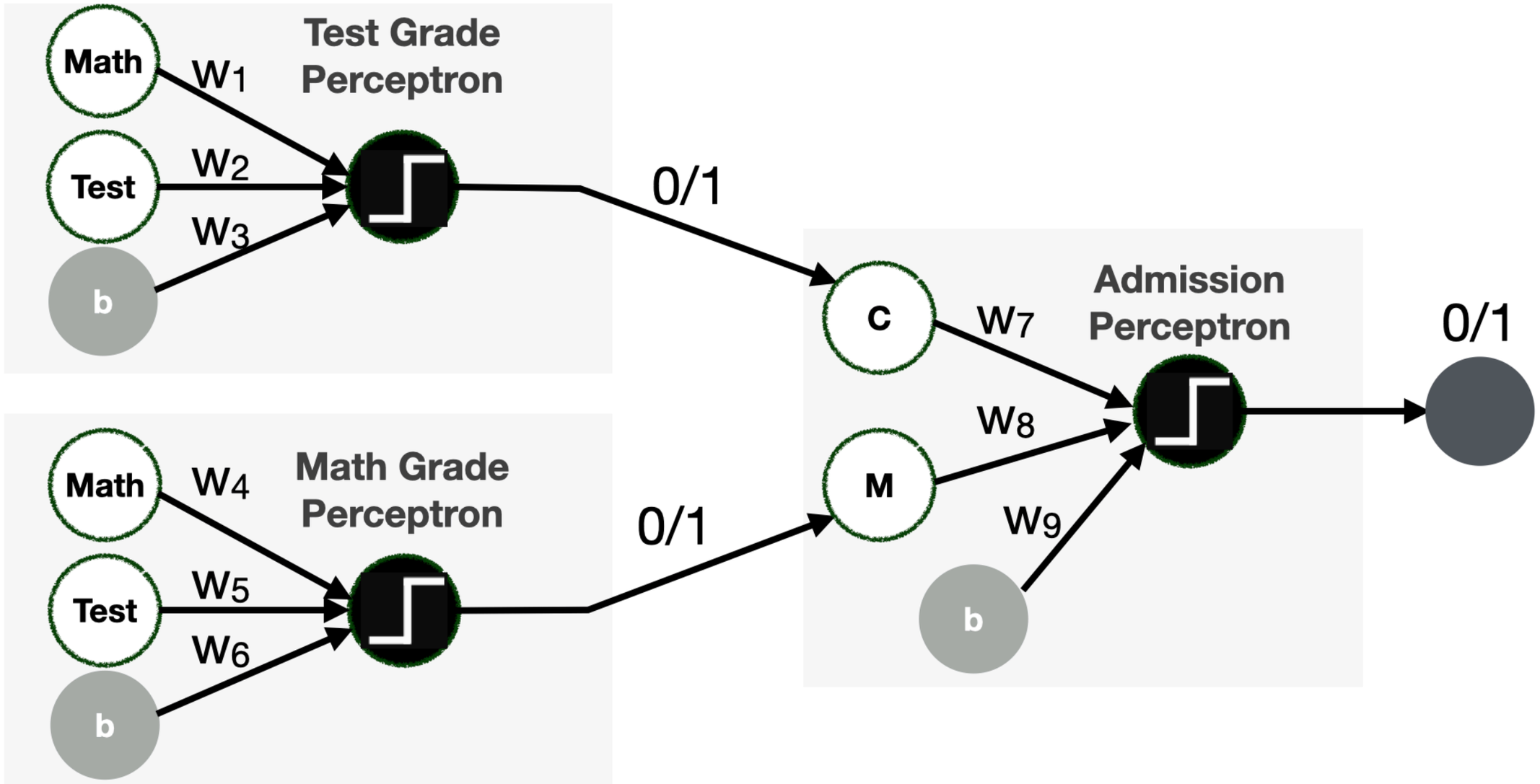


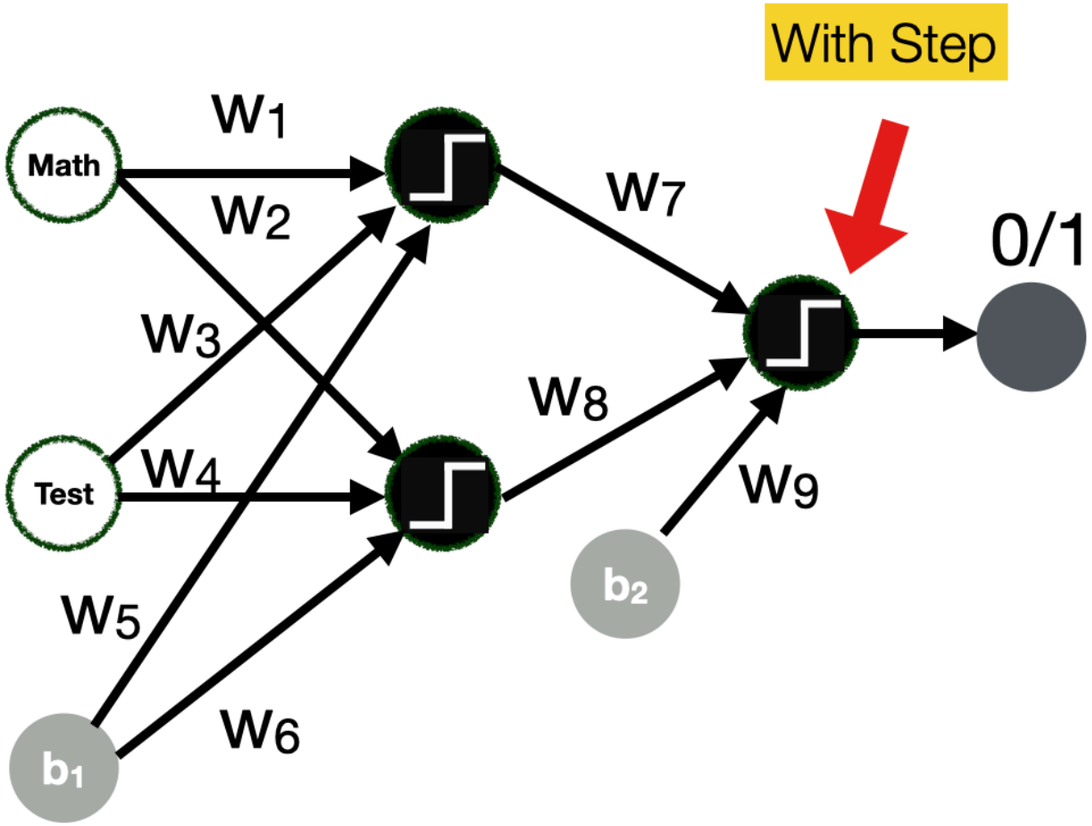
Input (weighted sum of values)

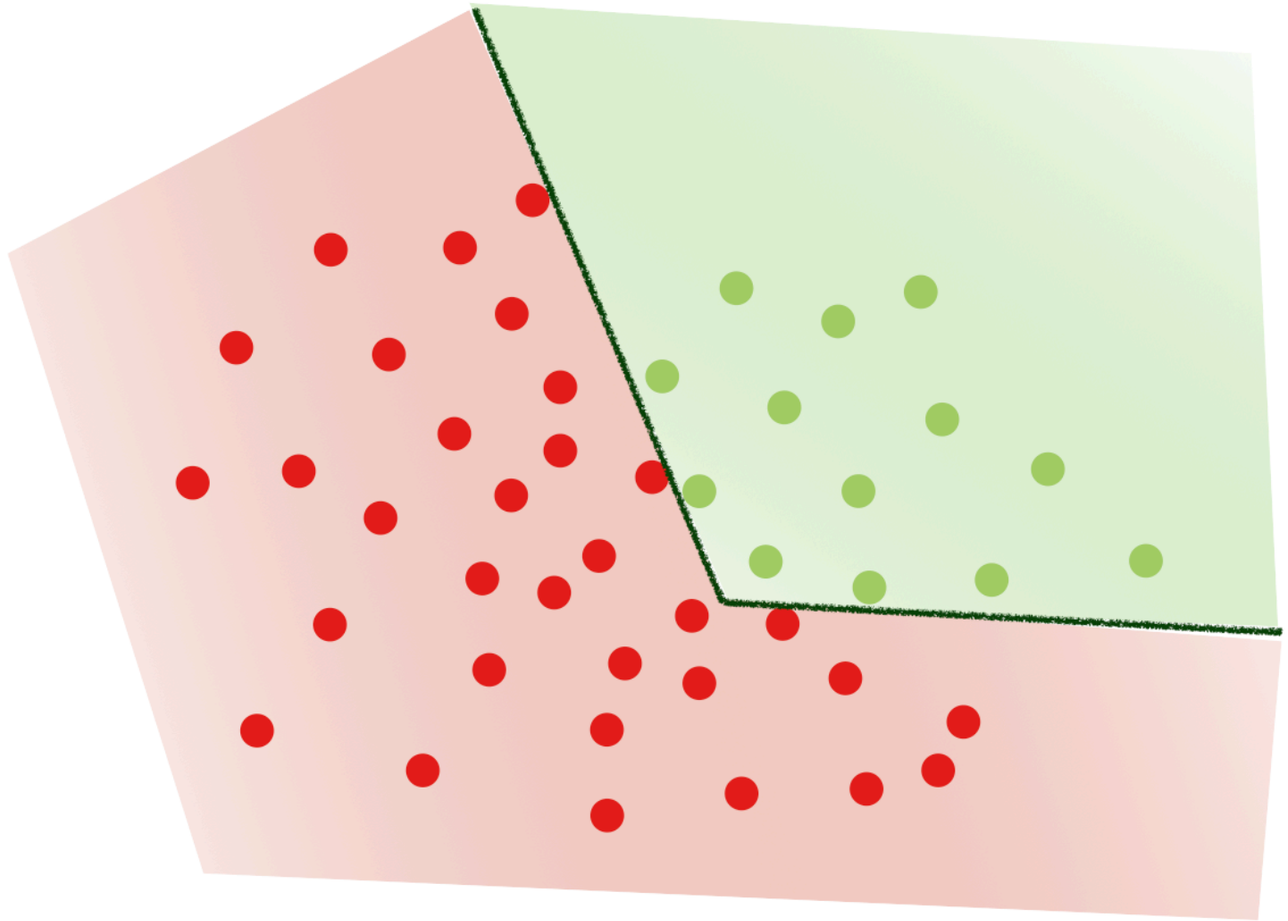


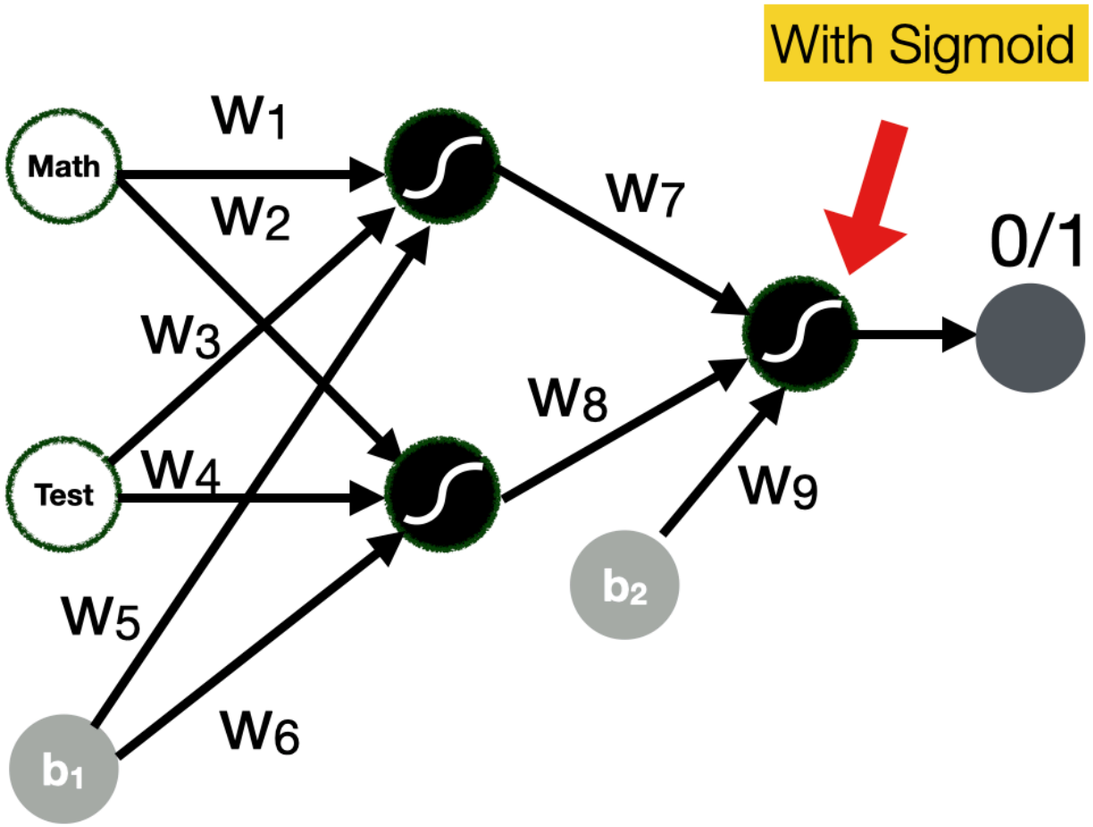


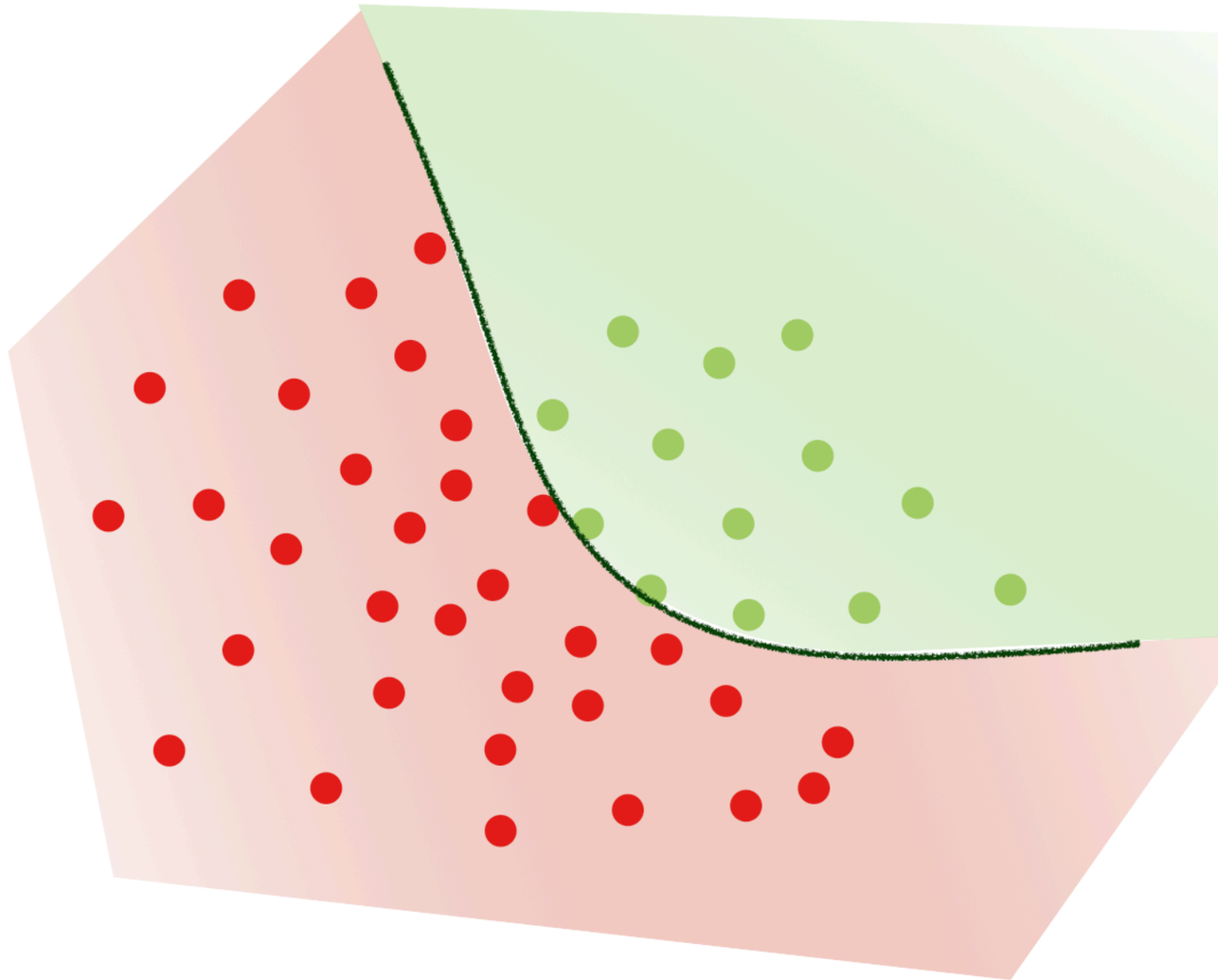

Test Grade
Classifier
Math Grade
Classifier
Admission
Classifier


Math
Test
b
W1
W2
W3
Test Grade Perceptron
0/1
C
W7
Admission Perceptron
0/1
W8
M
W9
b
Math
Test
b
W4
W5
W6
Math Grade Perceptron
0/1

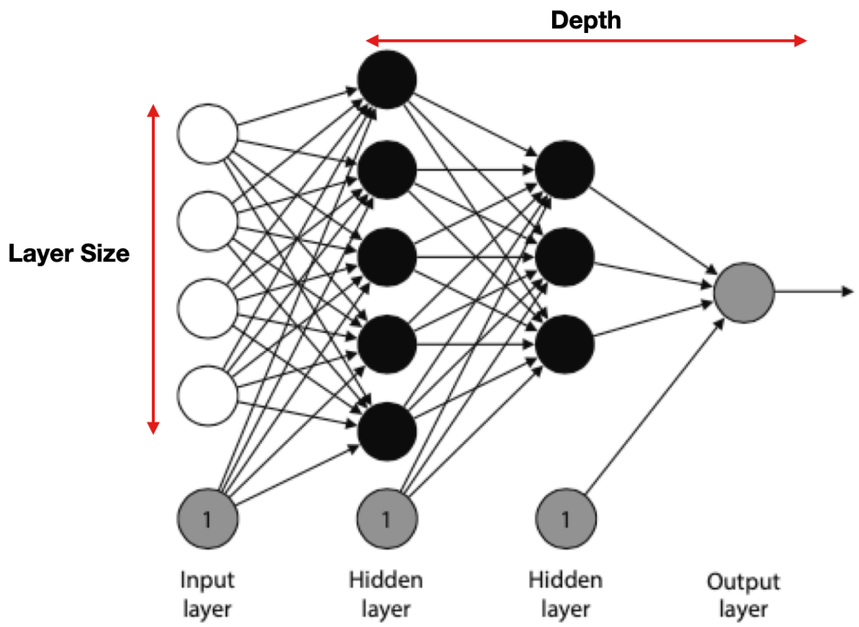

With Step
Math
Test
b_1
b_2
W_1
W_2
W_3
W_4
W_5
W_6
W_7
W_8
W_9
0/1




With Sigmoid
Math
Test
W1
W2
W3
W4
W5
W6
W7
W8
W9
b1
b2
0/1

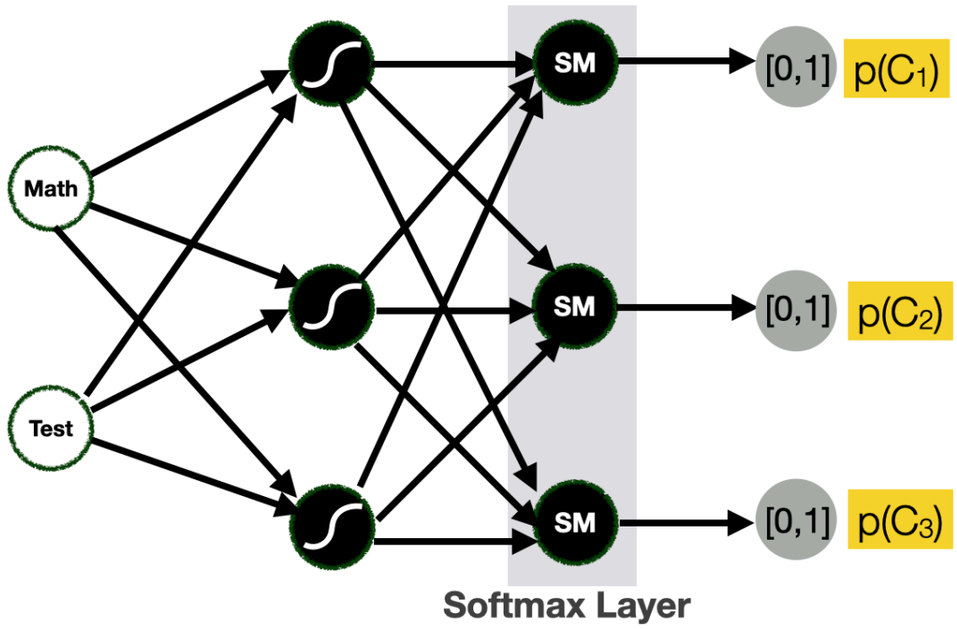


Fully connected Neural Network



- **Hyperparameters**
  - Learning rate
  - Number of epochs
  - Architecture
    - #layers, #nodes, activation functions
  - Batch vs. mini-batch vs. stochastic gradient descent
  - Regularization parameters:
    - Dropout probability p

Classifying into multiple classes - *Softmax* function



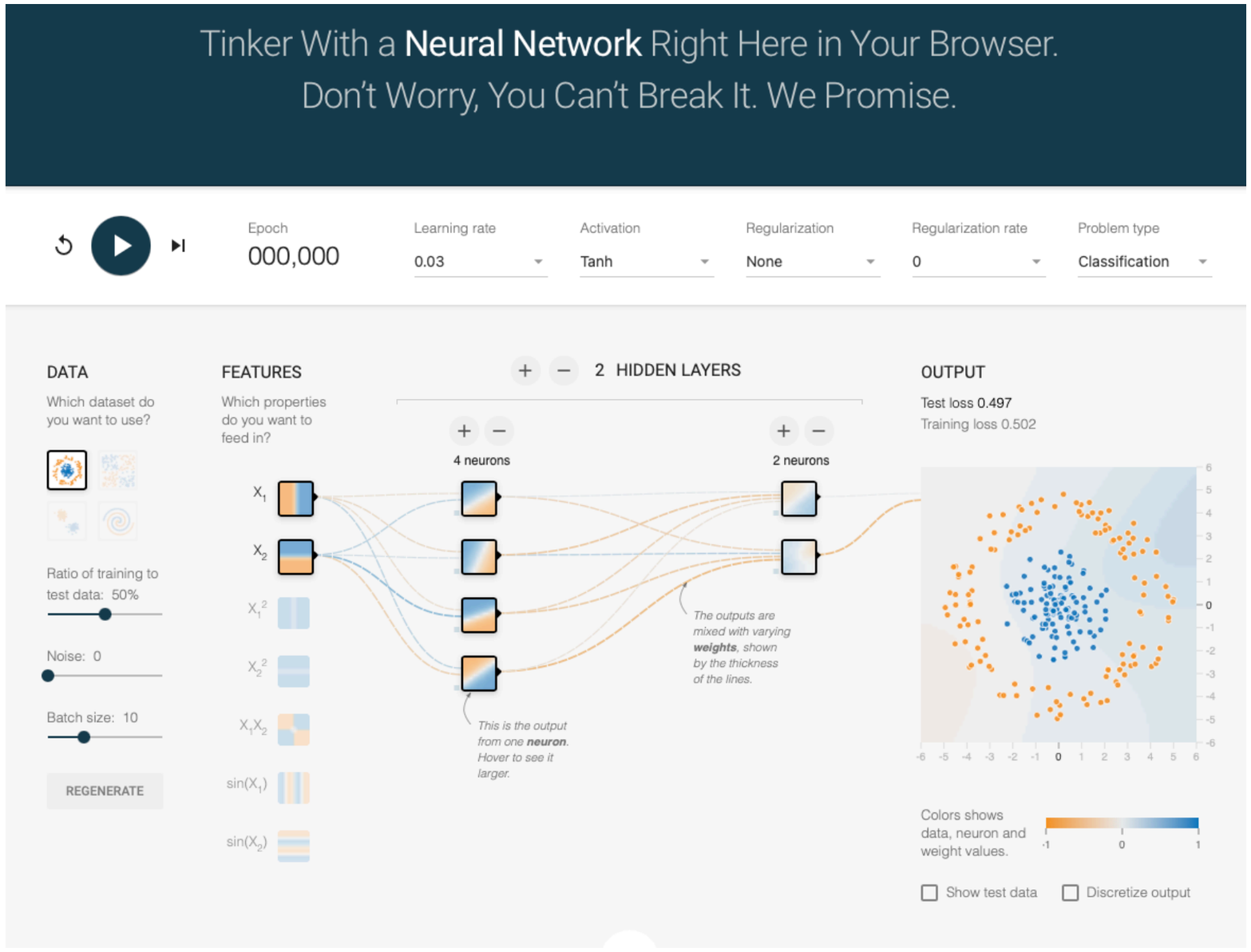
$$\text{Softmax}(x_i) = \frac{e^{(x_i)}}{\sum_j^K e^{(x_j)}}$$

Value of class *i*

Normalisation term on K classes

- Return a probability for each class
  - example C1= ADMITTED, C2 = NOT ADMITTED, C3 = NEW TEST
  - p(C1) = 0.37, p(C2) = 0,21, p(C3) = 0,42
- We use the *Softmax* activation function for the output layer

Tensorflow Playground

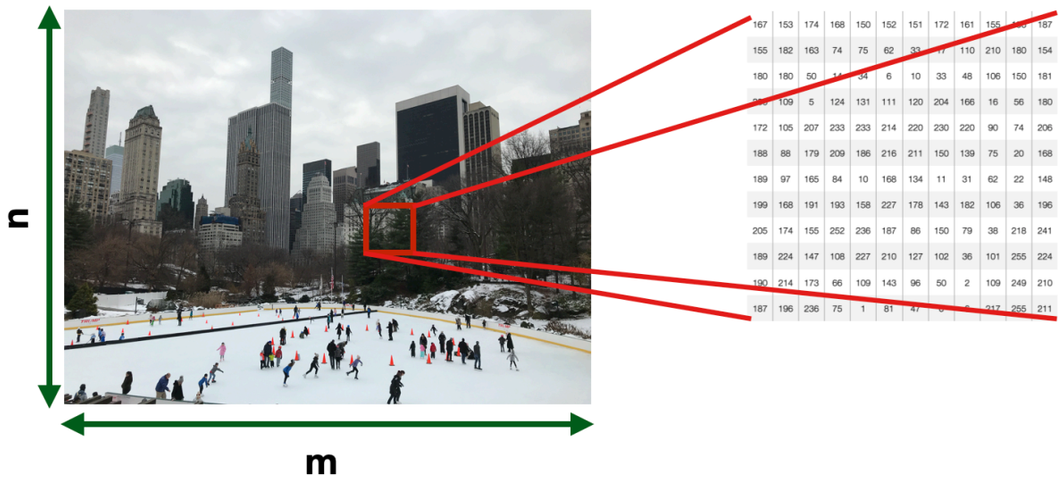


Machine Learning and Images

What do you see?

This is what a computer “sees”

Images



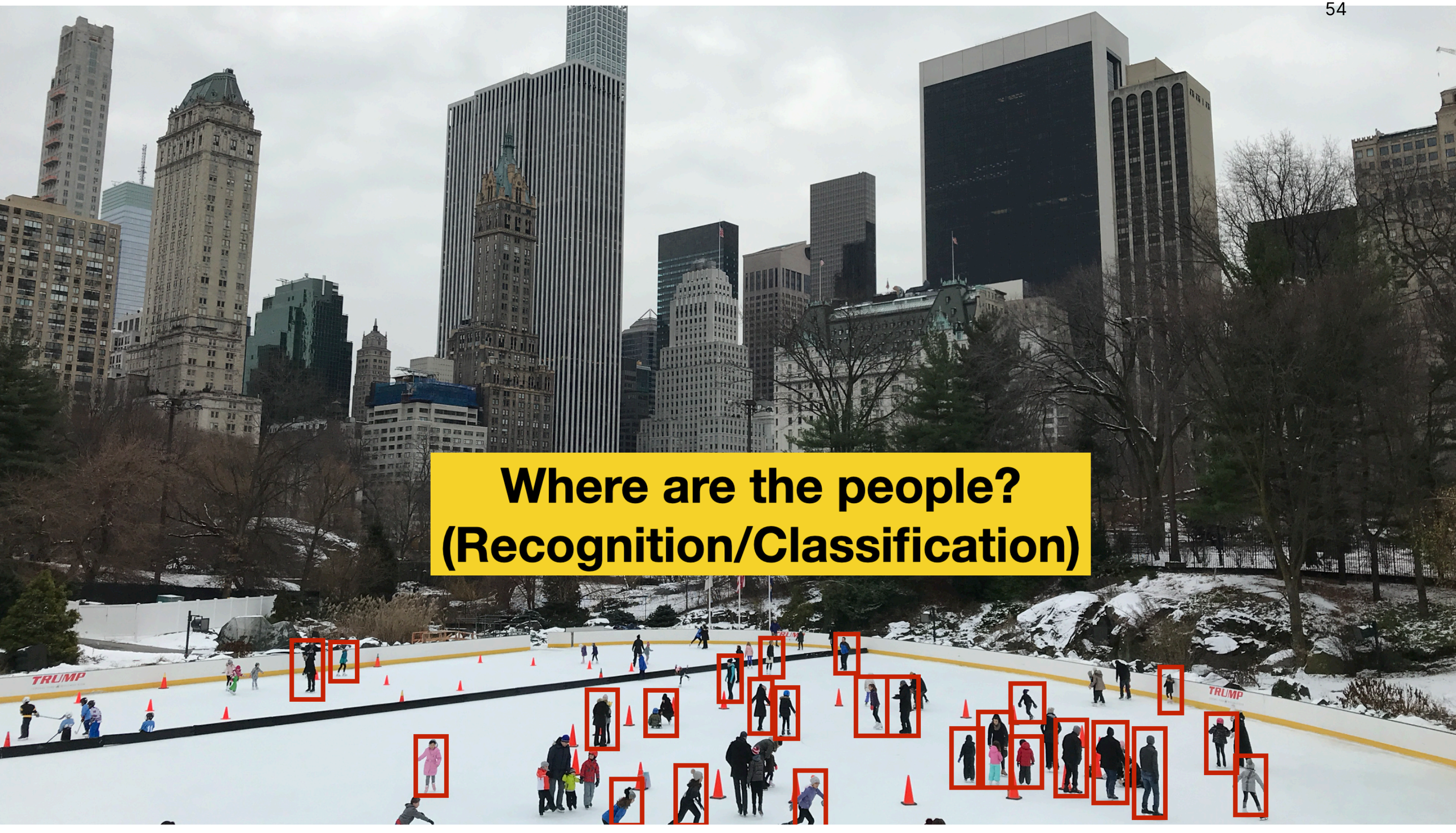
- Each pixel in an image is a *feature*
  - numerical
    - 0 or 1 for *Black and White*
    - Between 0 and 255 for *greyscale*
    - 16M values for *RGB*
- Dimensionality → *n x m*

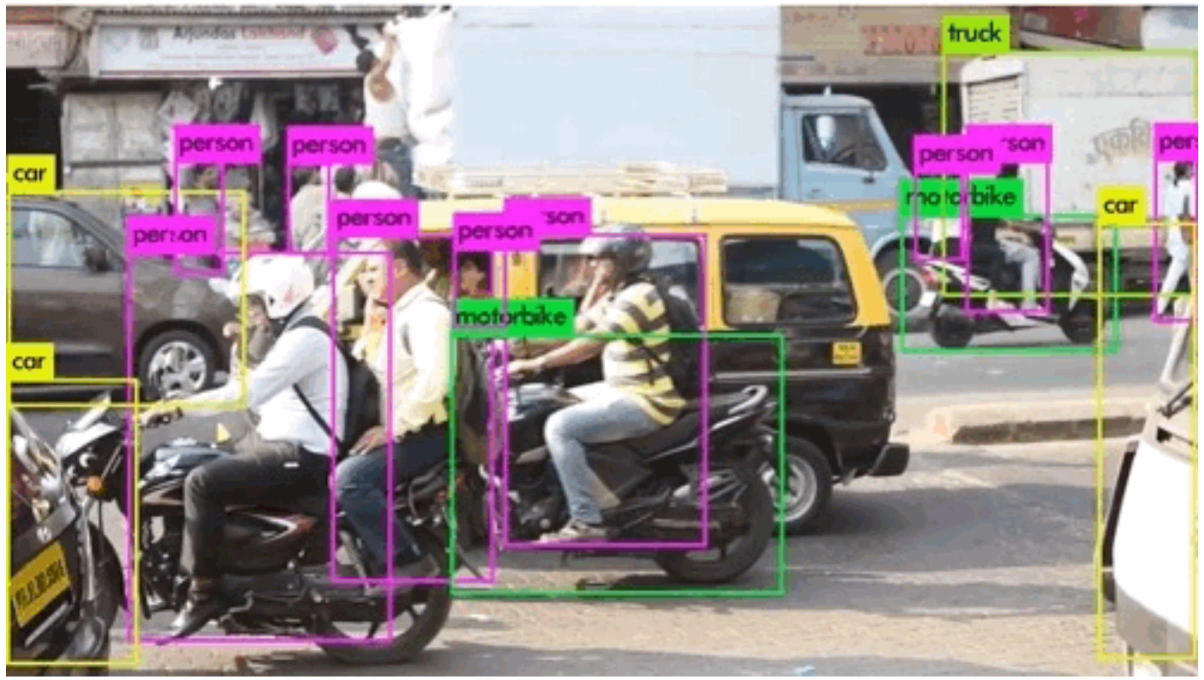
Computer Vision

- Building algorithms that can “understand” the content of images and use it for other applications
- It is a “Strong AI” problem
  - signal-to-symbol conversion
  - The **semantic gap**
- A general-purpose vision system **requires**
  - Flexible, robust visual representation
    - Updated and maintained
  - Reasoning
  - Interfacing with attention, goals, and plans

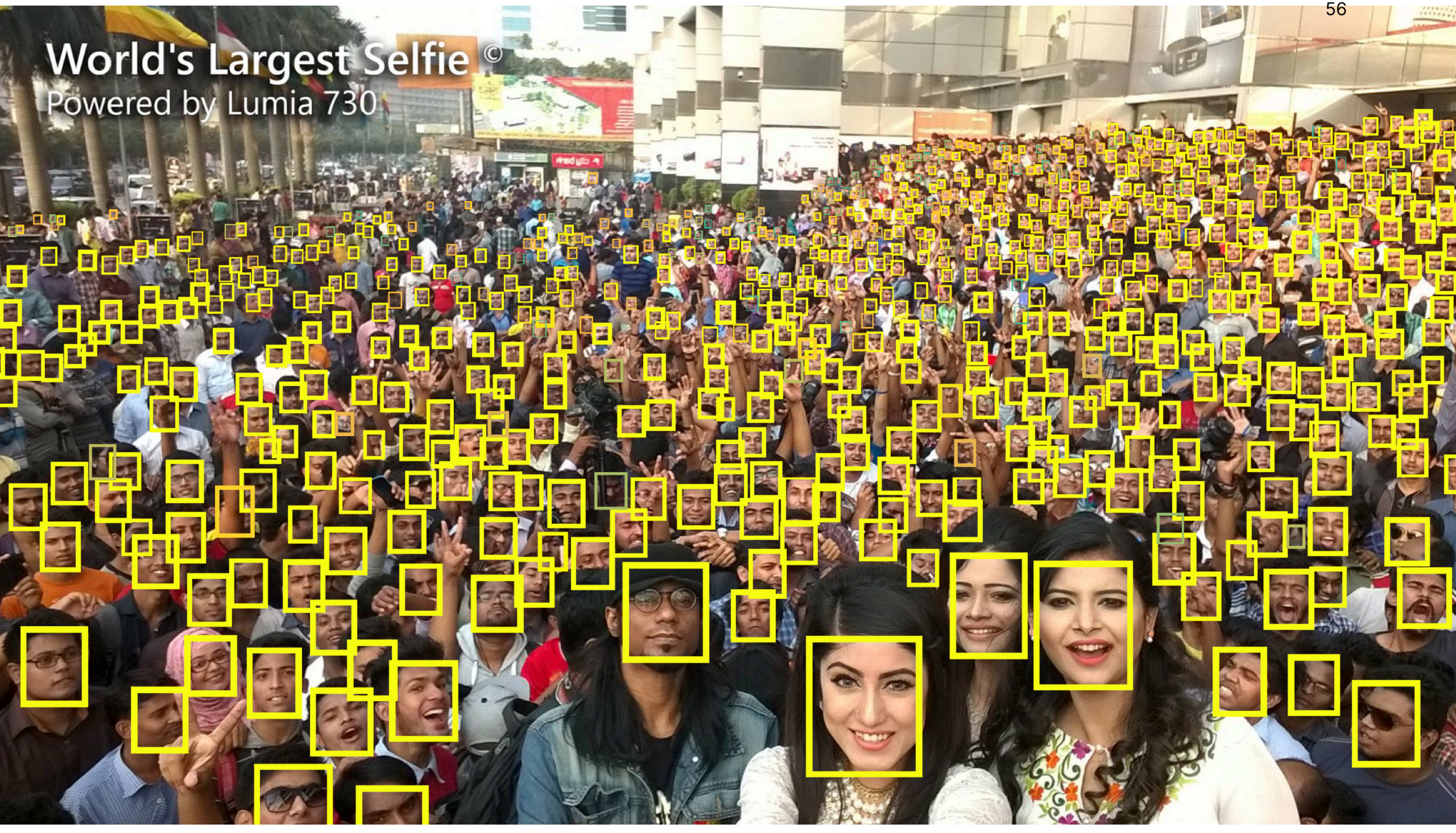
What specific tasks can we train a CV system to perform?


Is this a flag?
(Recognition / Classification)
TRUMP
TRUMP

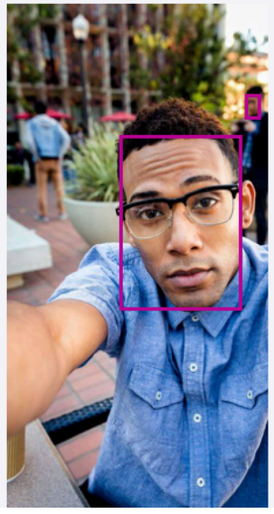

Where are the people?
(Recognition/Classification)
TRUMP
TRUMP

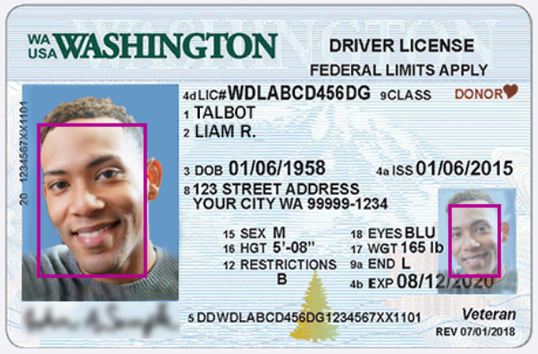

truck
person
car
motorbike


person


World's Largest Selfie ©
Powered by Lumia 730


Is this Jeff?
(Identification)
TRUMP
TRUMP



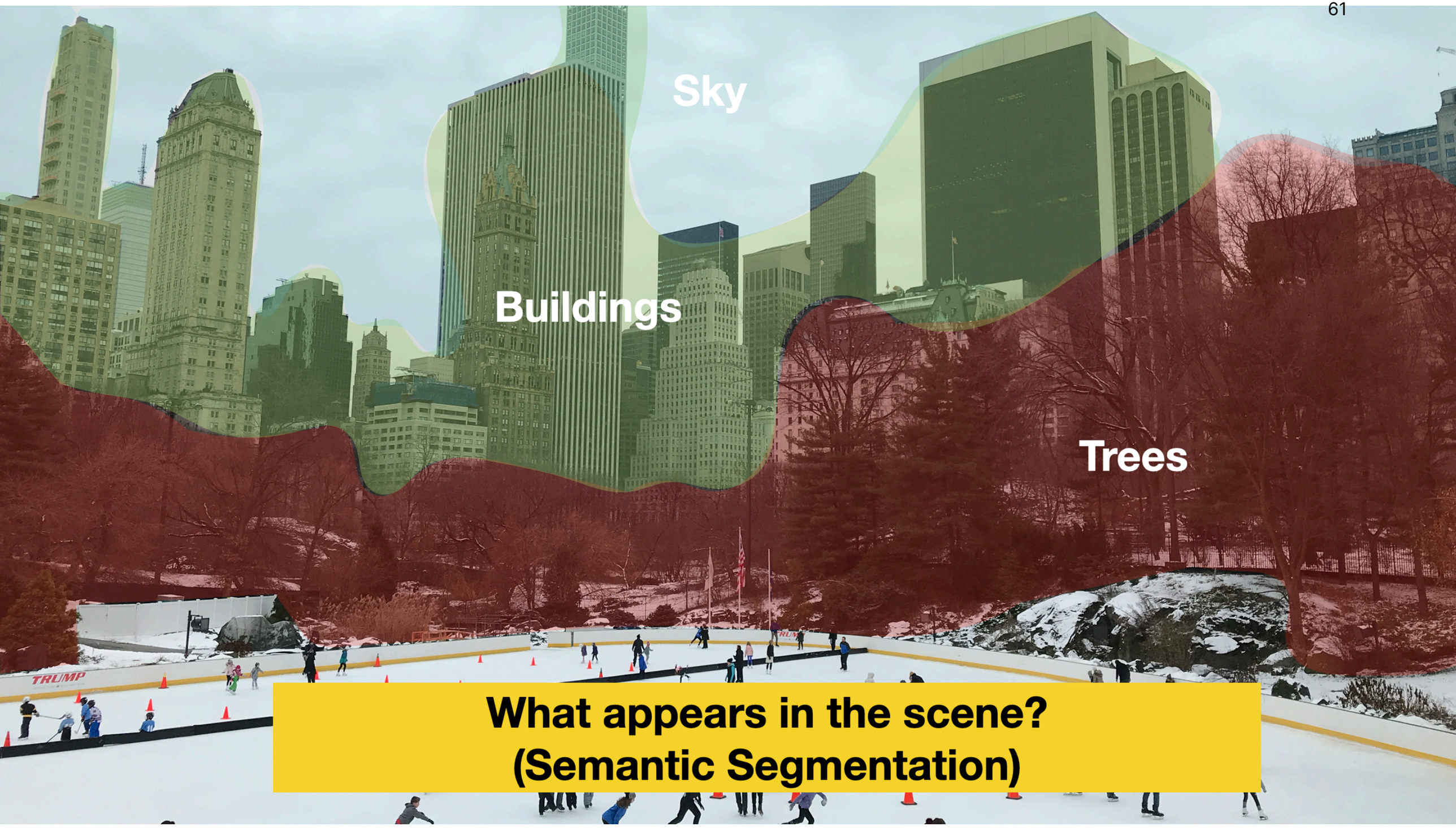

WA
USA WASHINGTON
DRIVER LICENSE
FEDERAL LIMITS APPLY
4d LIC# WDLABCD456DG 9 CLASS DONOR
1 TALBOT
2 LIAM R.
3 DOB 01/06/1958 4a ISS 01/06/2015
8 123 STREET ADDRESS
YOUR CITY WA 99999-1234
15 SEX M
16 HGT 5'-08"
12 RESTRICTIONS B
18 EYES BLU
17 WGT 165 lb
9a END L
4b EXP 08/12/2020
5 DDWDLABCD456DG1234567XX1101
Veteran
REV 07/01/2018


対象人物
対象人物
対象人物
対象人物

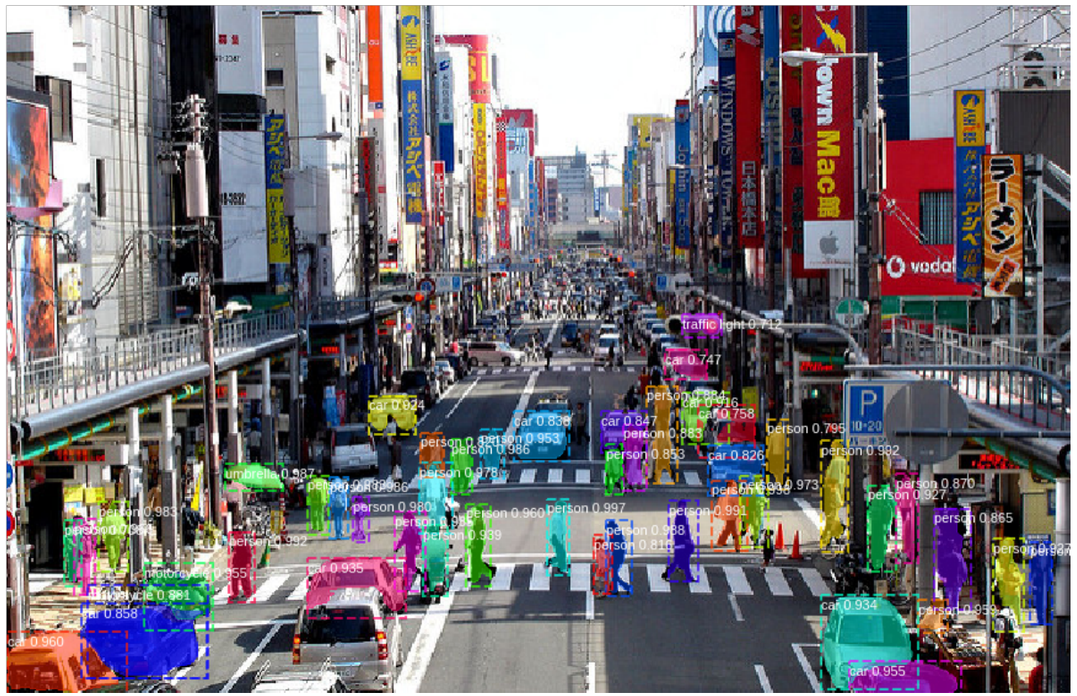

Is this the Wollman Rink?
(Identification)
TRUMP

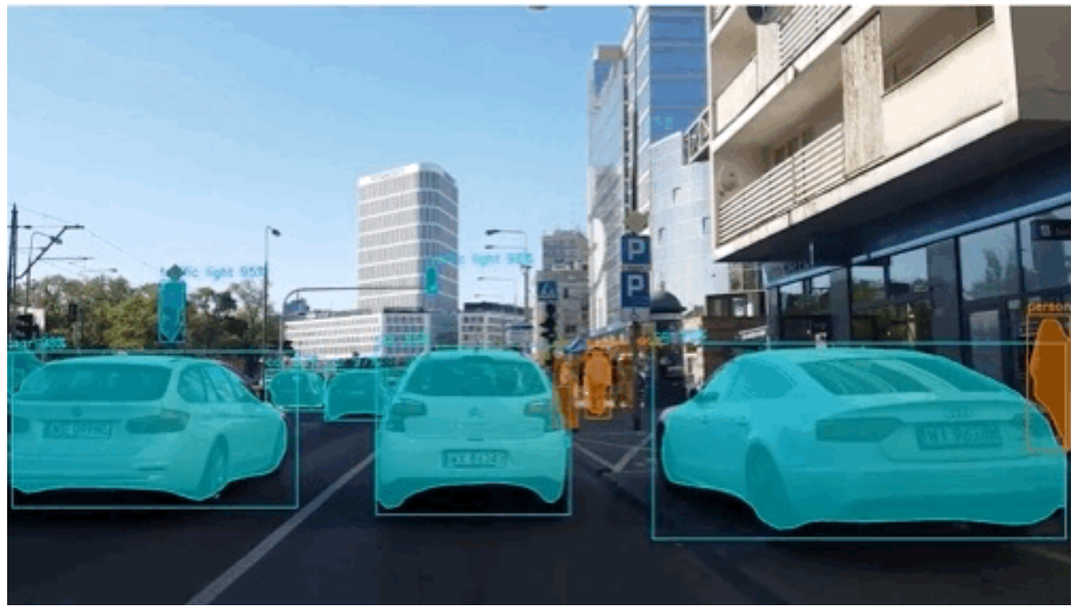

8:00
Google Lens
Coit Tower
Text selection
Coit Tower
4.4 2287 Reviews


Google Lens
DELFT
UNIVERSITY OF TECHNOLOGY / FACULTY OF INDUSTRIAL DESIGN ENGINEERING
DESIGN
GUIDE
PERSPECTIVES - MODELS - APPROACHES - MET
BISPUBLISHERS
Translate
Text
Search
Homework
Shoppi
Delft Design Guide: Design Strategies and Methods
Book
Search

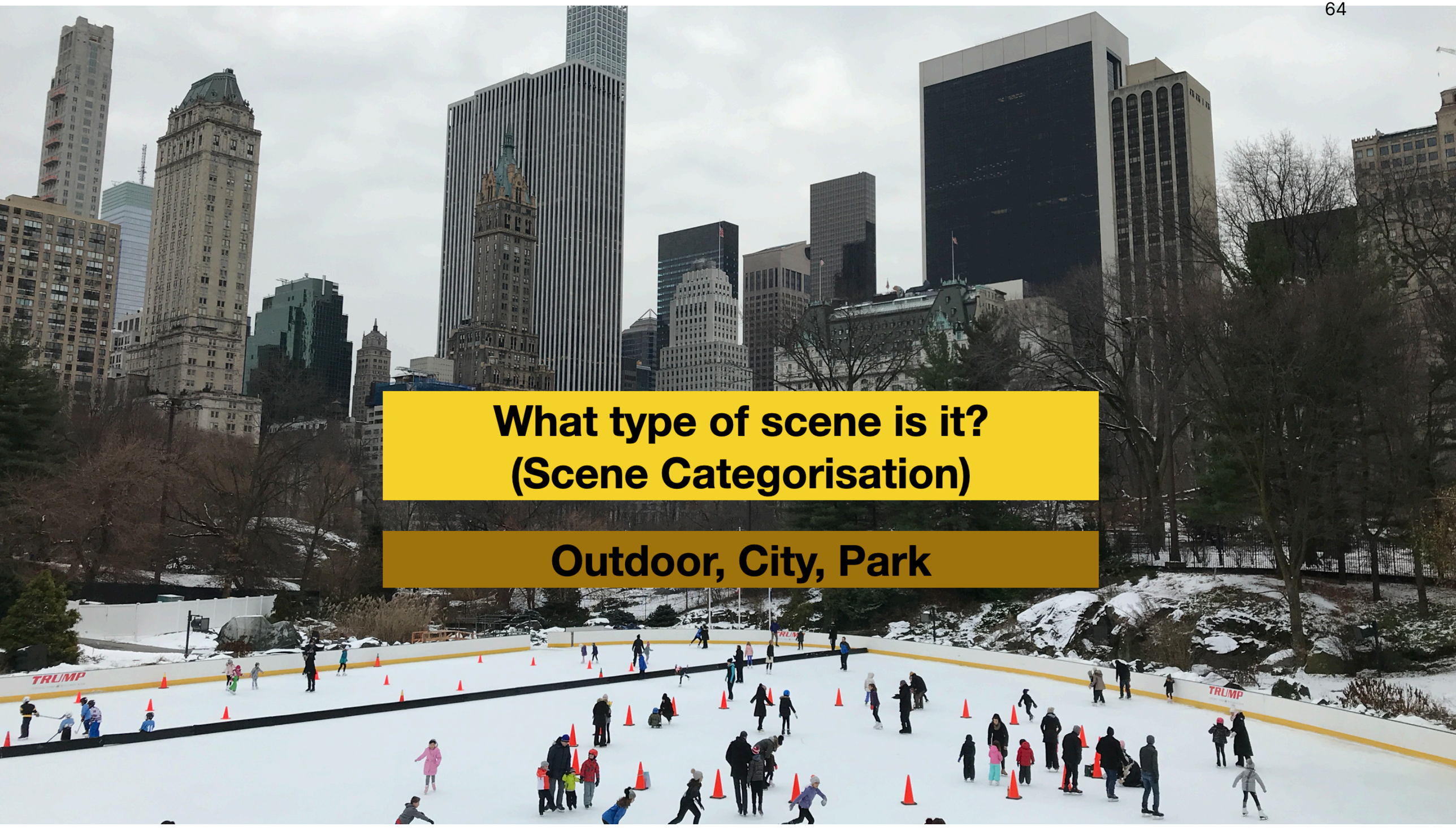


What appears in the scene?
(Semantic Segmentation)


traffic light 0.712
car 0.747
car 0.924
car 0.838
car 0.847
car 0.816
car 0.758
person 0.795
person 0.992
person 0.953
person 0.986
person 0.883
person 0.853
car 0.826
person 0.978
umbrella 0.987
person 0.973
person 0.870
person 0.927
person 0.865
person 0.983
person 0.980
person 0.960
person 0.997
person 0.991
person 0.939
person 0.988
person 0.811
car 0.935
motorcycle 0.881
car 0.858
car 0.960
car 0.934
car 0.955

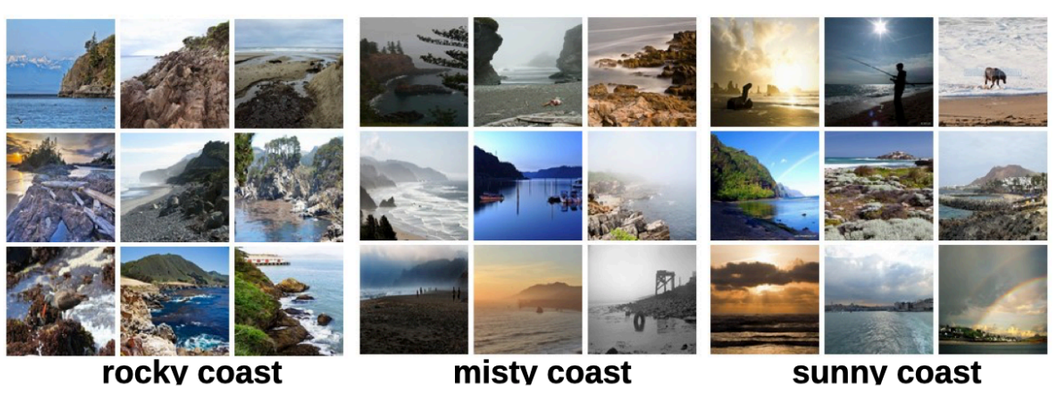


Project Sunroof



What type of scene is it? (Scene Categorisation)

Outdoor, City, Park



Predictions:

- **Type of environment:** outdoor
- **Scene categories:** skyscraper (0.704), downtown (0.211)
- **Scene attributes:** man-made, vertical components, open area, natural light, clouds, no horizon, metal, glass, sunny
- **Informative region for predicting the category *skyscraper* is:**

Hugging Face | Search models, datasets, users... | Models | Datasets | Spaces | Docs | Solutions | Pricing

Spaces: OFA-Sys / OFA-Image_Caption | like 12 | Running

App | Files and versions

OFA-Image_Caption

Gradio Demo for OFA-Image_Caption. Upload your own image or click any one of the examples, and click "Submit" and then wait for the generated caption.

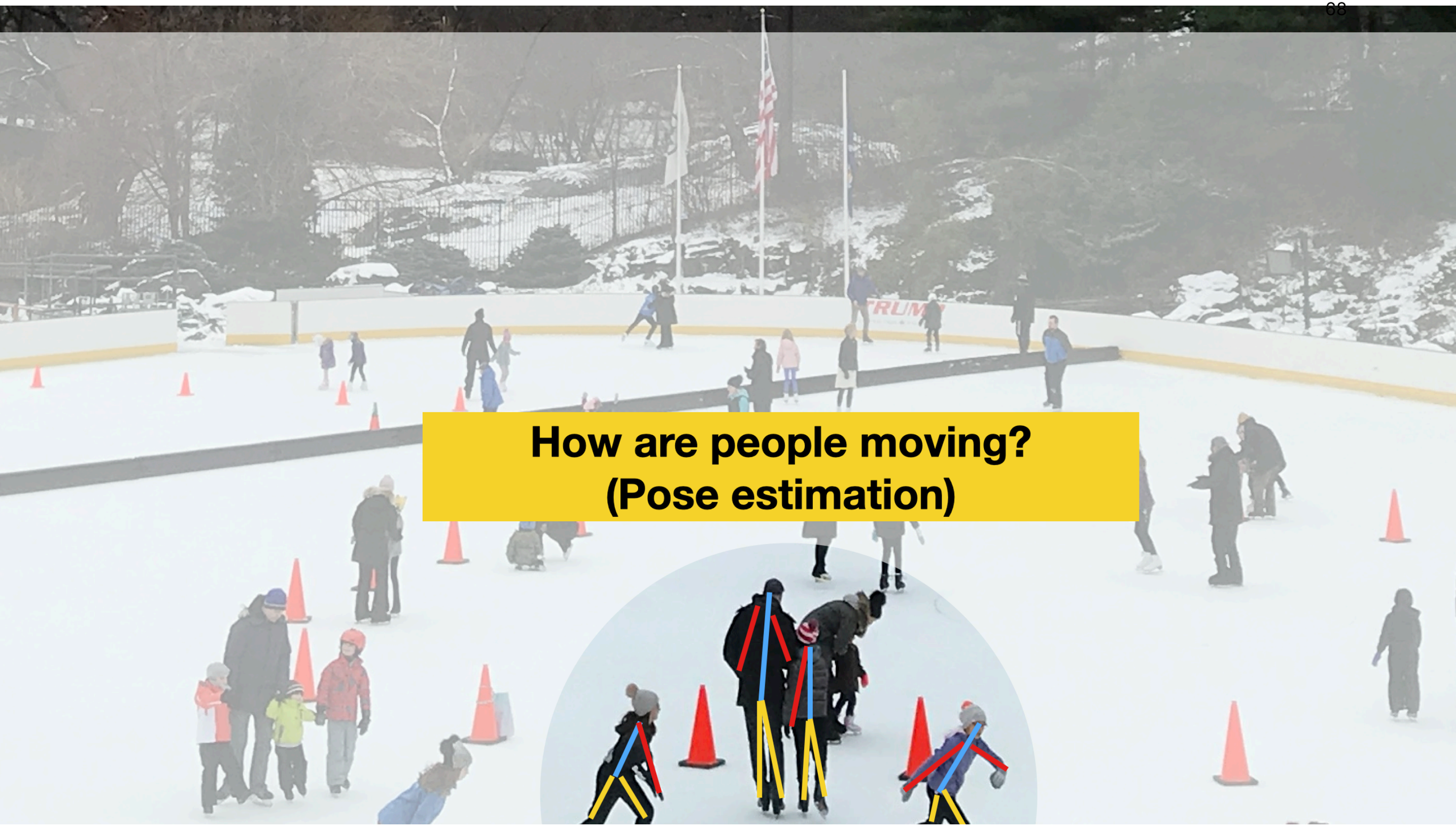
Image

Caption 8.28s

people skating on a rink in a city park with skyscrapers

Clear | Submit


What are these people doing?
(Activity / event recognition)
Skating
TRUMP
TRUMP

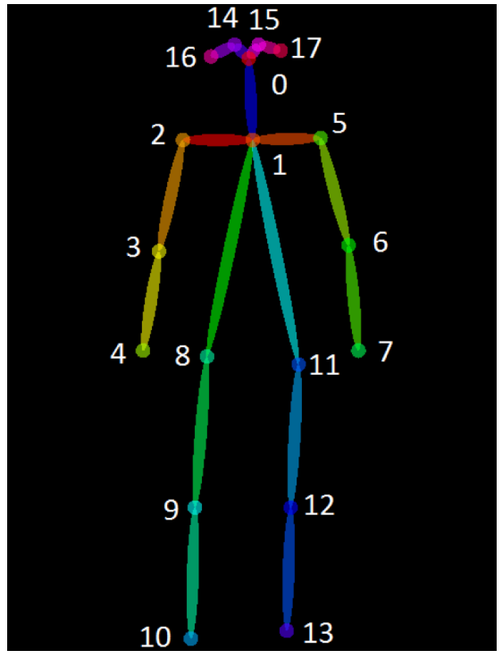

How are people moving?
(Pose estimation)

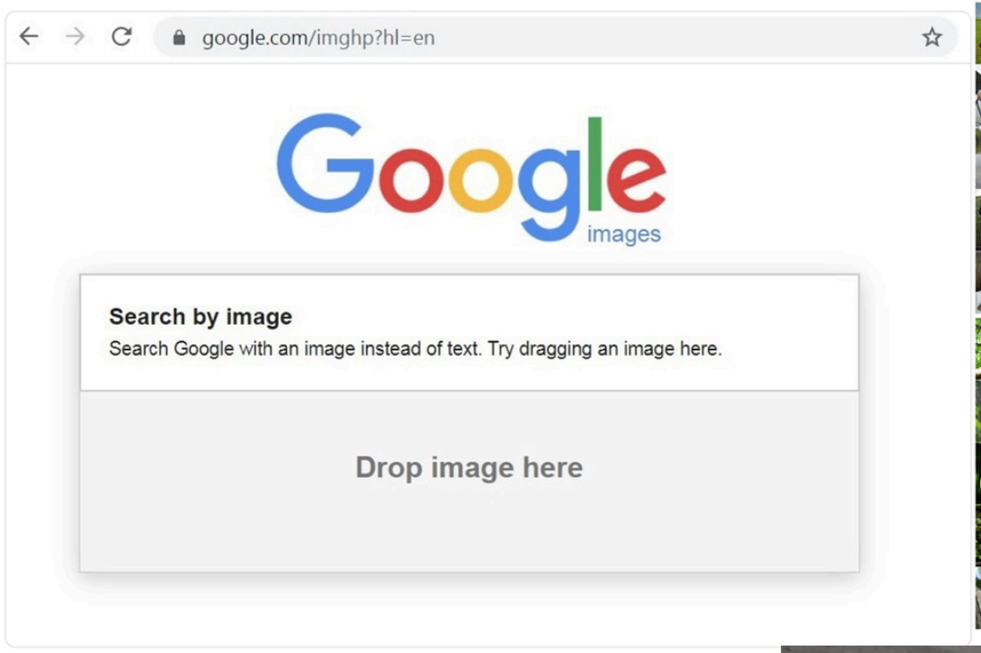
Stereolabs ZED Camera

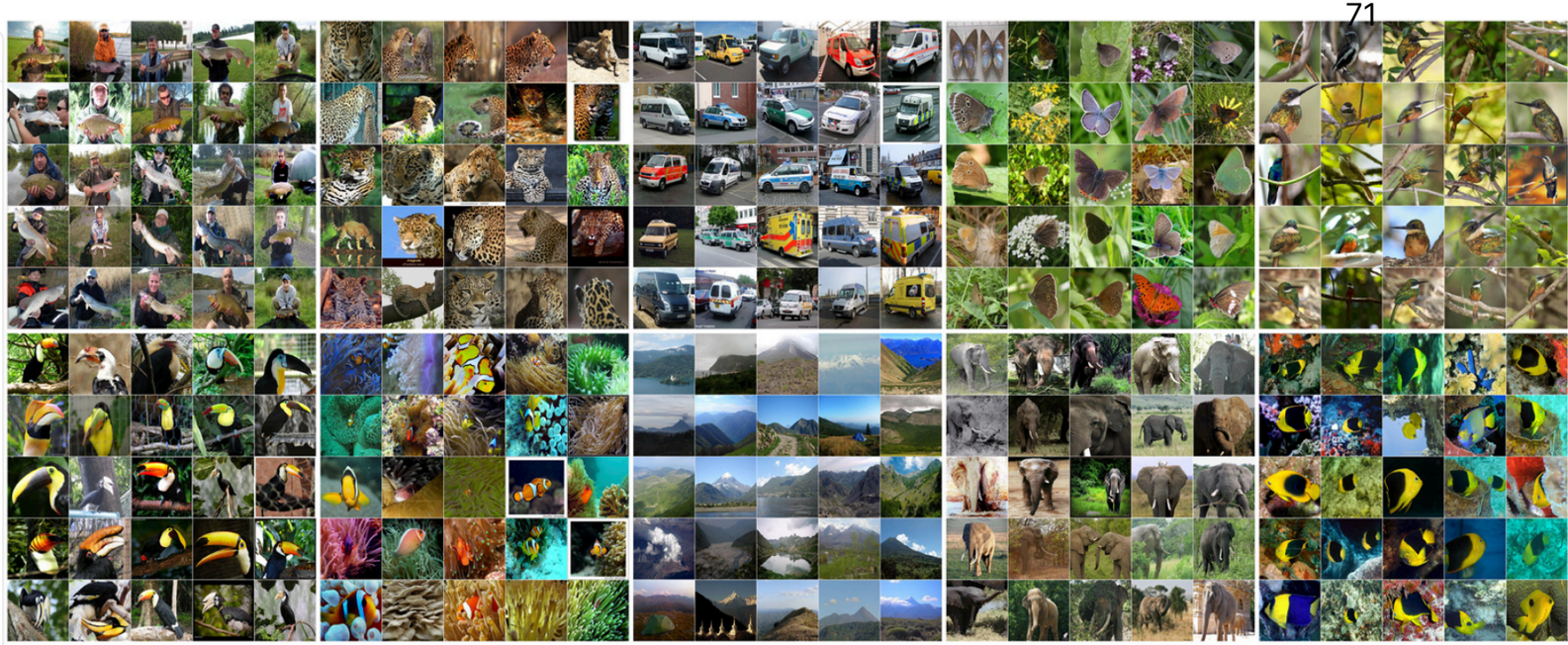
3D Object Detection

Body tracking

Positional tracking




google.com/imghp?hl=en
Google
images
Search by image
Search Google with an image instead of text. Try dragging an image here.
Drop image here



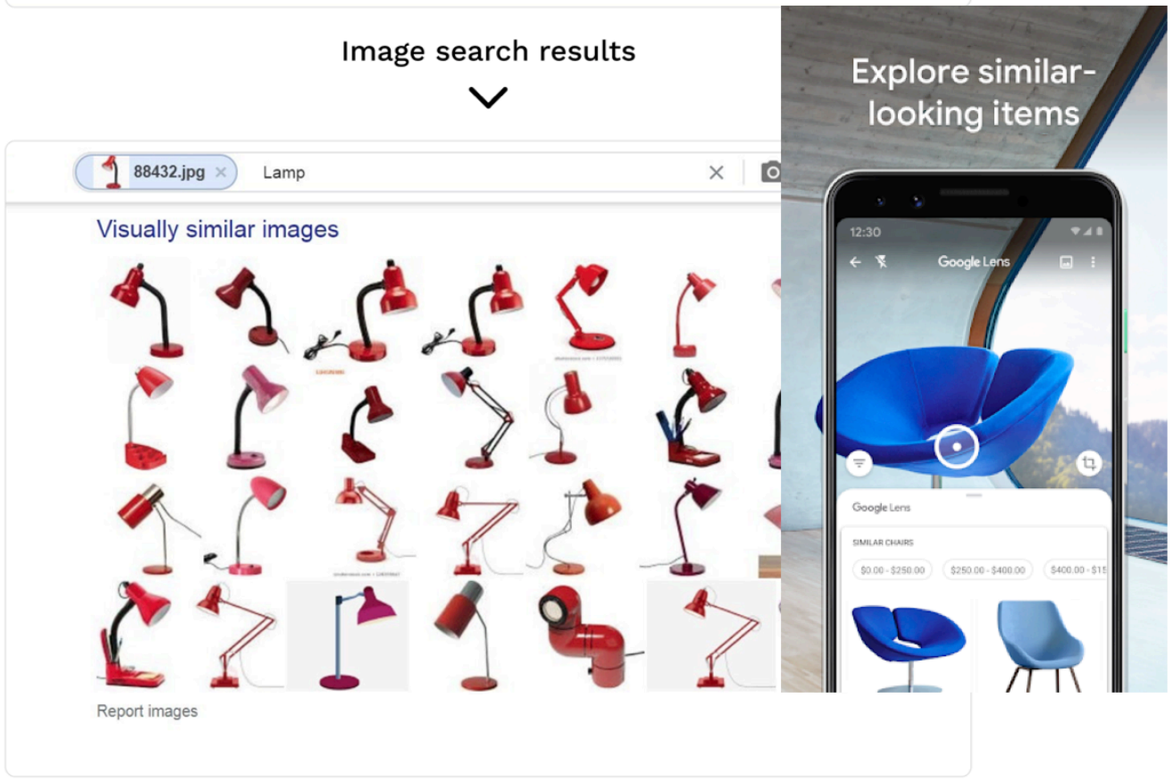
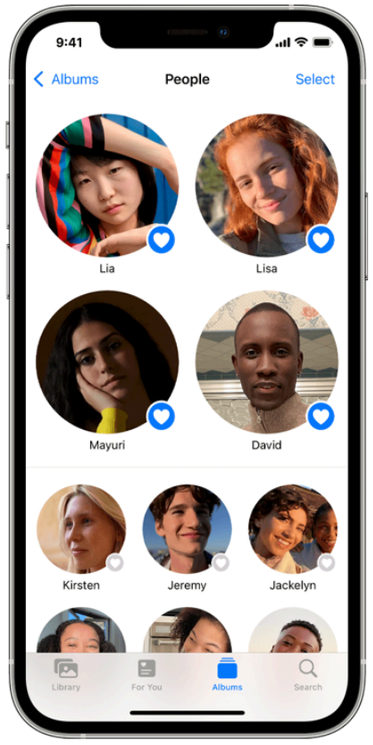
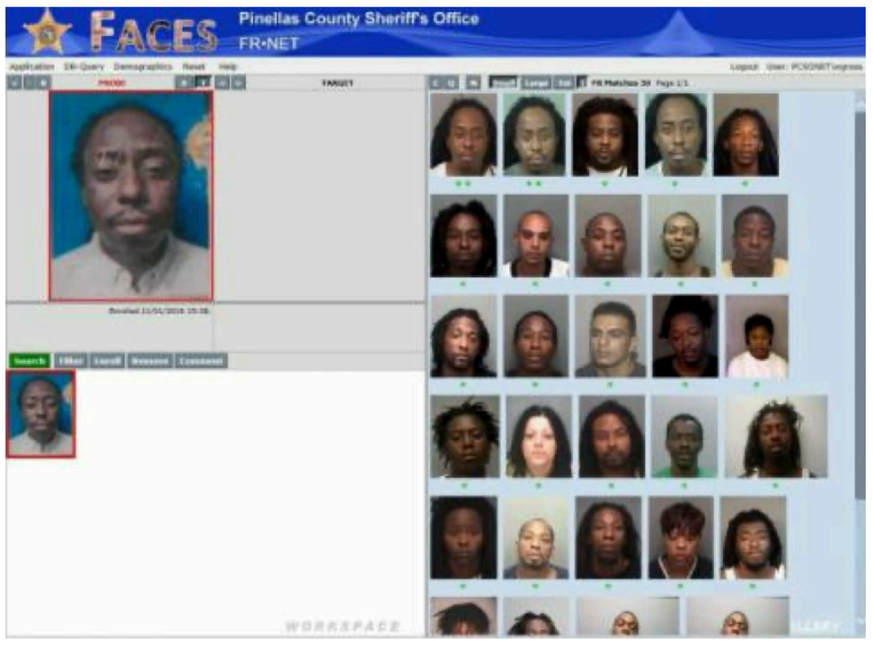


Image search results
88432.jpg
Lamp
Visually similar images
Report images

Explore similar-looking items
12:30
Google Lens
Google Lens
SIMILAR CHAIRS
$0.00 - $250.00
$250.00 - $400.00
$400.00 - $15


9:41
Albums
People
Select
Lia
Lisa
Mayuri
David
Kirsten
Jeremy
Jackelyn
Library
For You
Albums
Search


FACES
Pinellas County Sheriff's Office
FR•NET
WORKSPACE

Machine Learning for Design

Lecture 3

Machine Learning for Images. *Part 1*

Credits

CMU Computer Vision course - Matthew O’Toole.

Grokking Machine Learning. Luis G. Serrano. Manning, 2021