

Advanced Machine Learning For Design

Lecture 4 - Machine Learning
for Image Processing (part 1)

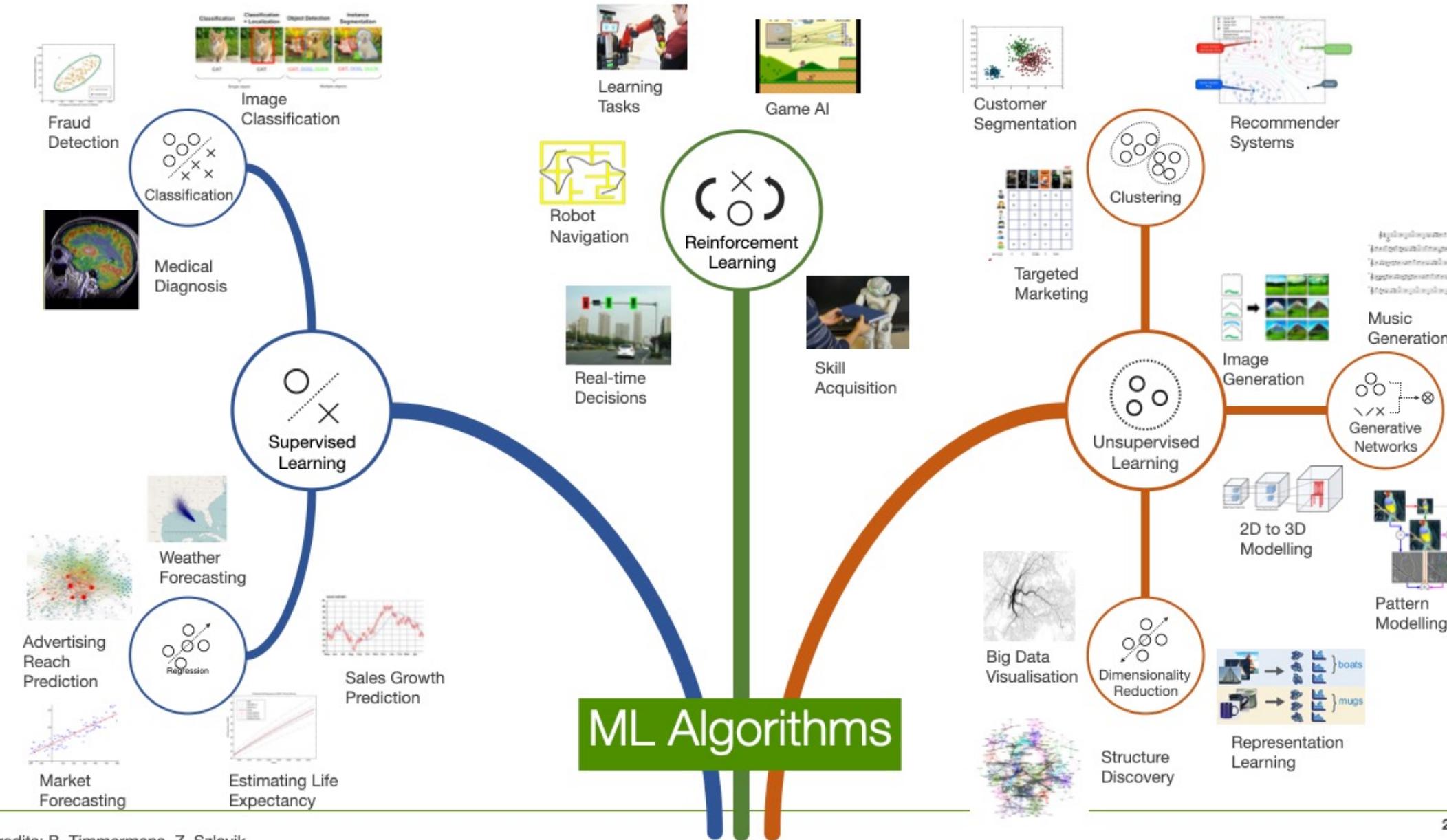
Module 2

Evangelos Niforatos
12/10/2022

aml4d-ide@tudelft.nl
<https://aml4design.github.io/>

Recap on ML/AI.

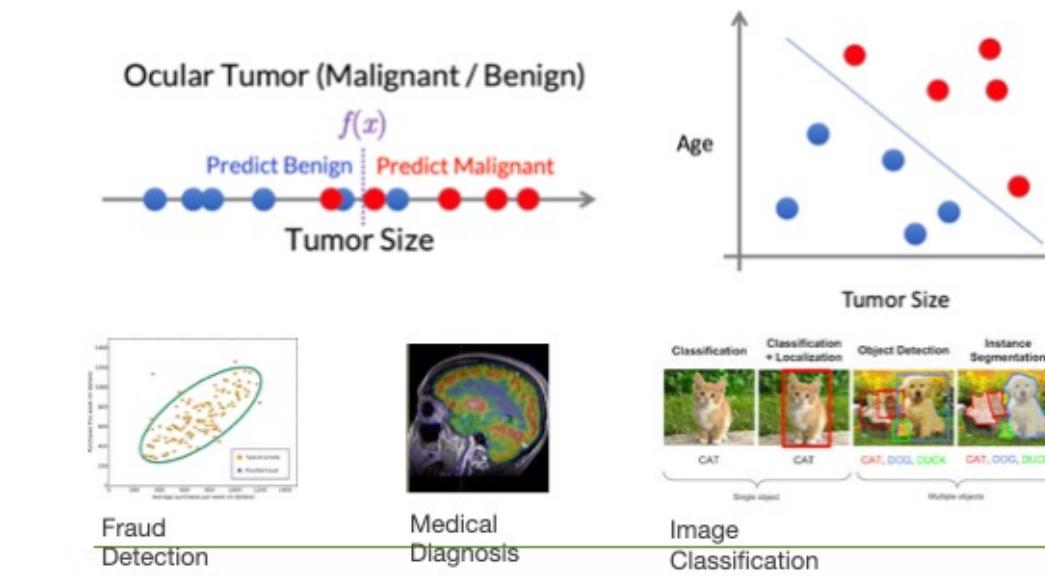




Classification / Regression

Classification

- Learn to output a **category** label
 - Binary (e.g. Spam / not Spam, Cat / not cat)
 - Multi-class (e.g. cat, dog, bird)



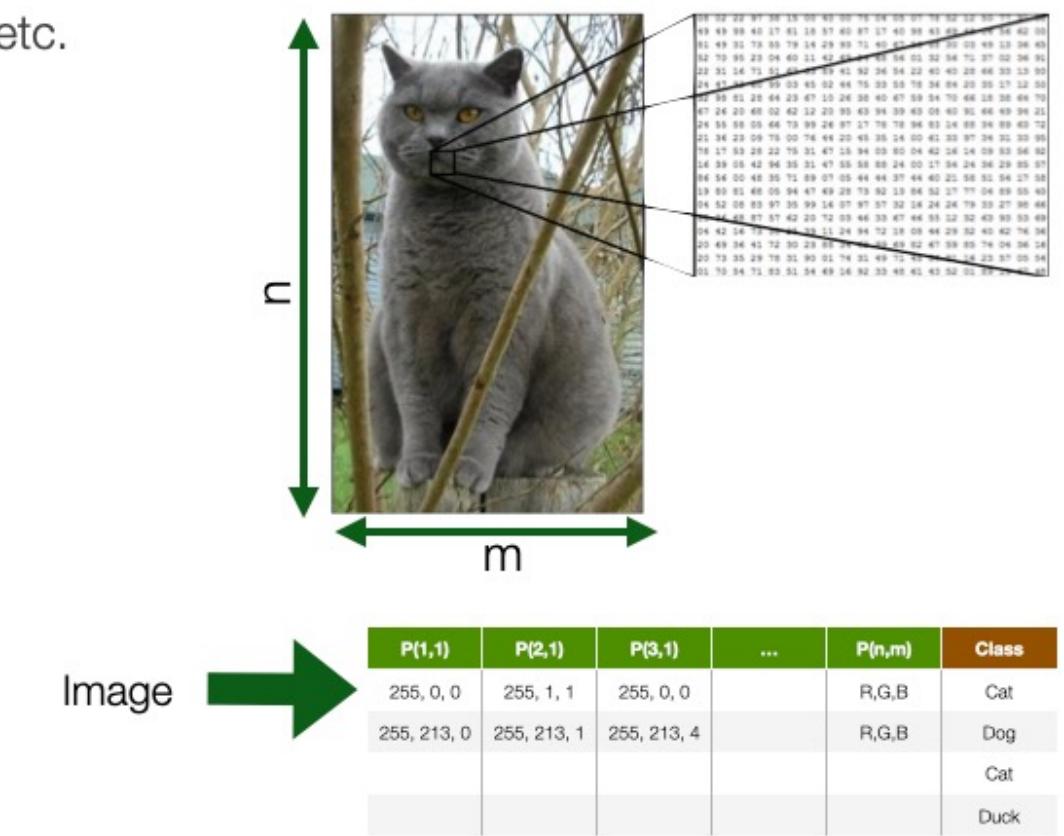
Regression (prediction)

- Learn to guess one or more numbers
 - e.g. value of a share, number of stars in a review



Images

- Visual content acquired through cameras, scanners, etc.
- Each pixel in an image is a feature
 - But spatially and geometrically organised
 - e.g. edges, corners
- Feature values are numerical values across channels
 - e.g. R,G,B
- Dimensionality $\rightarrow n \times m$

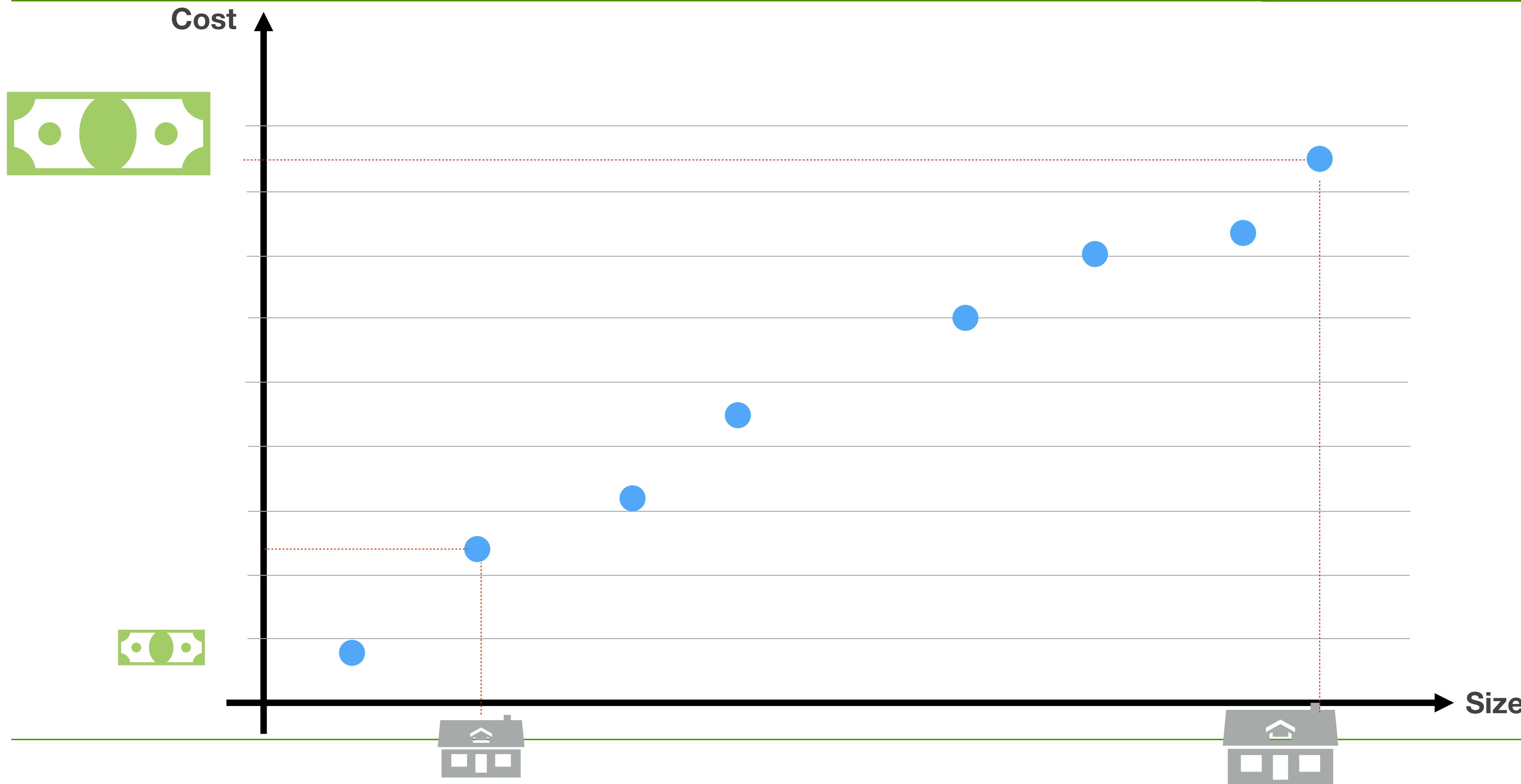


More in Module 2

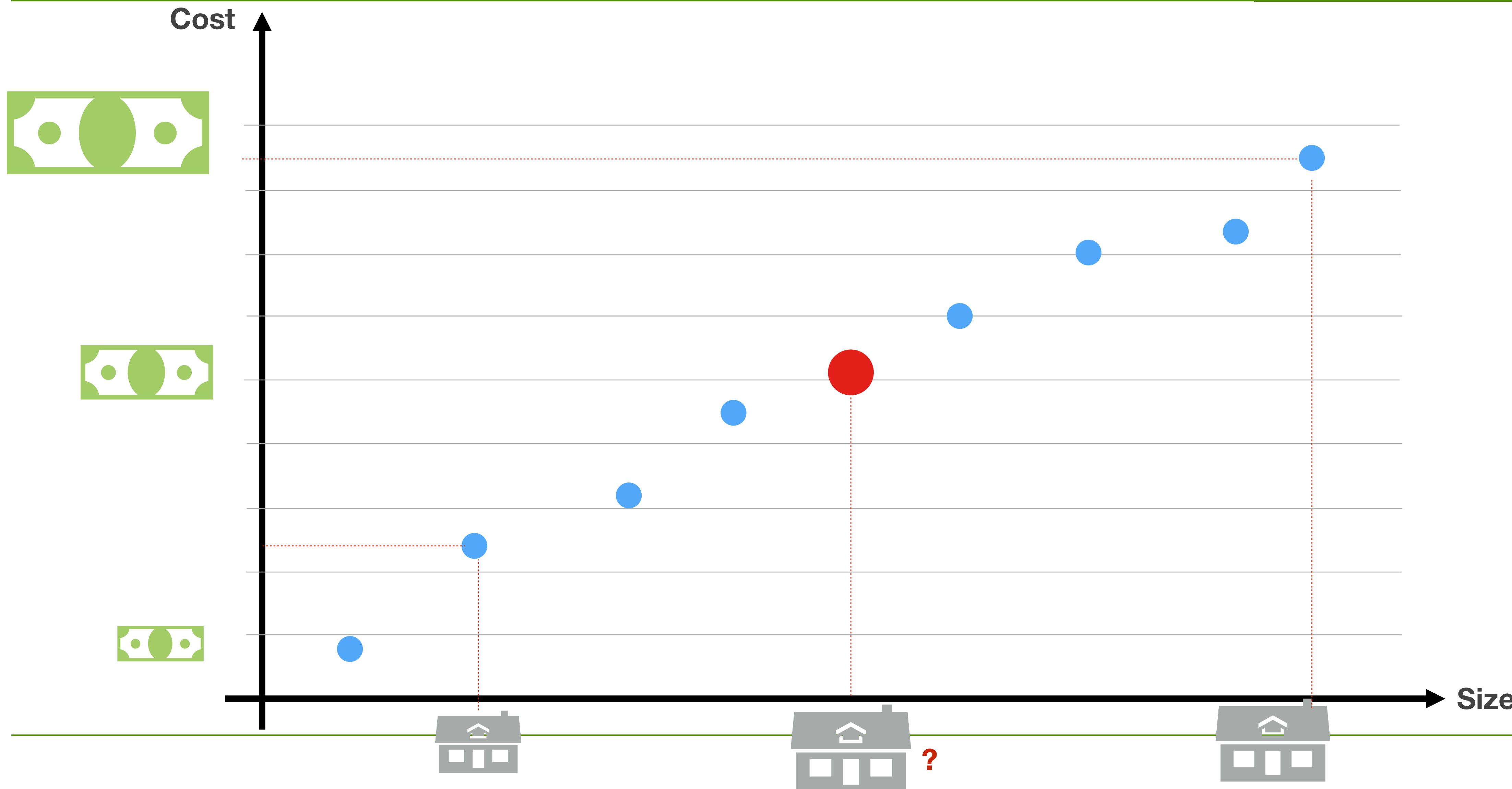
A bit more on regression and classification

And your 2nd contact with (deep) neural networks

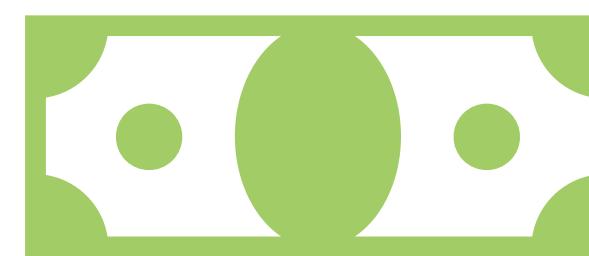
Linear Regression /1



Linear Regression /2



Linear Regression



Cost

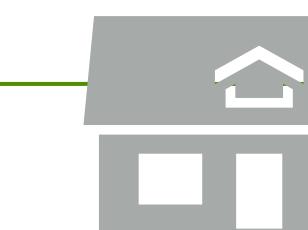
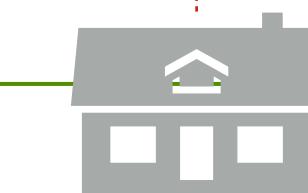
Dependent Variable

Intercept (parameter, or **bias**)

Cost = $b + w^*Size$

Slope (parameter, or **weight**)

Independent Variable

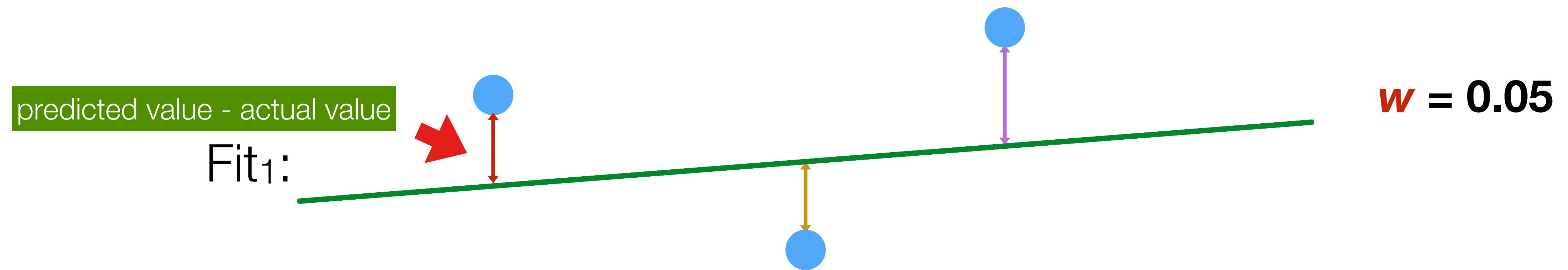


Size

Cost = w^*Size

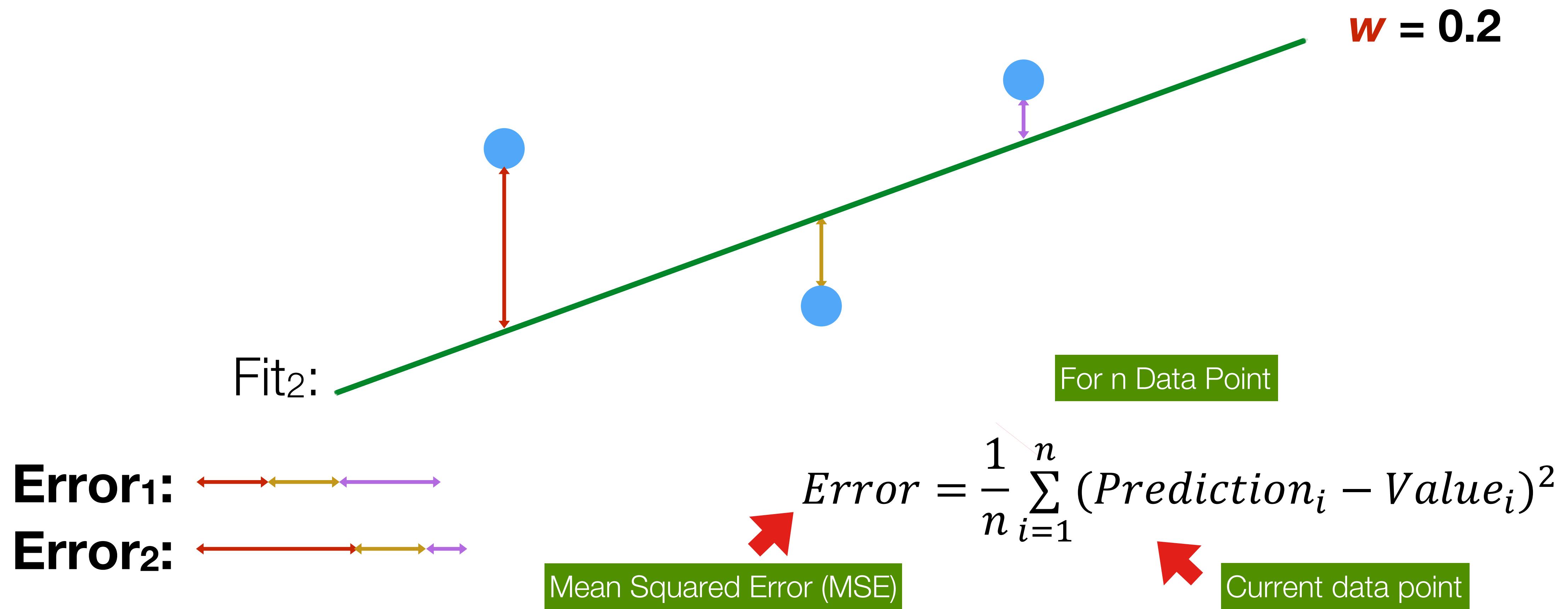


$$\text{Cost} = \mathbf{w}^* \mathbf{Size}$$

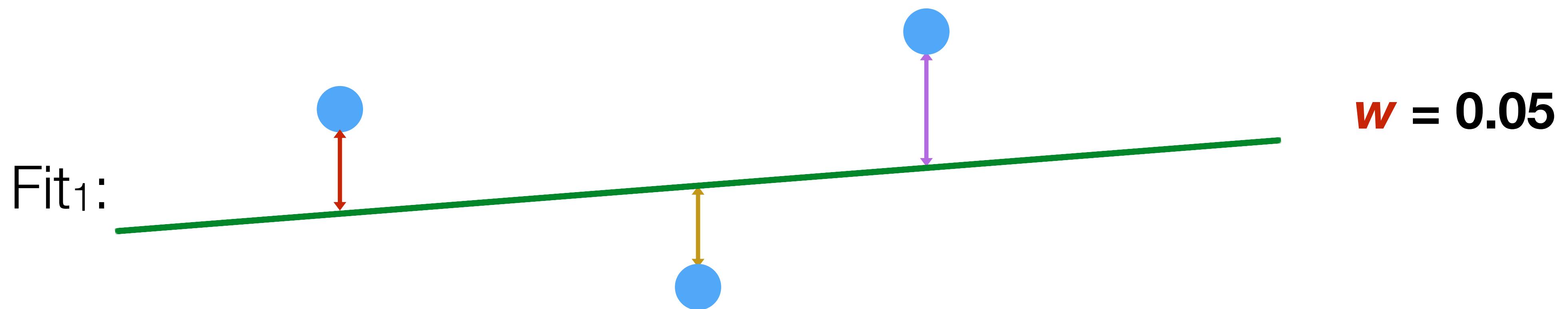


Error₁: ←→

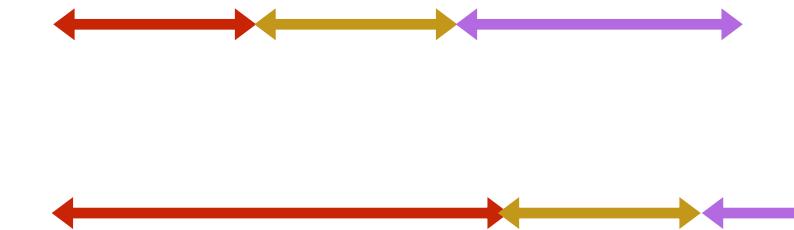
Cost = w^* Size



$$\text{Cost} = \mathbf{w}^* \mathbf{Size}$$



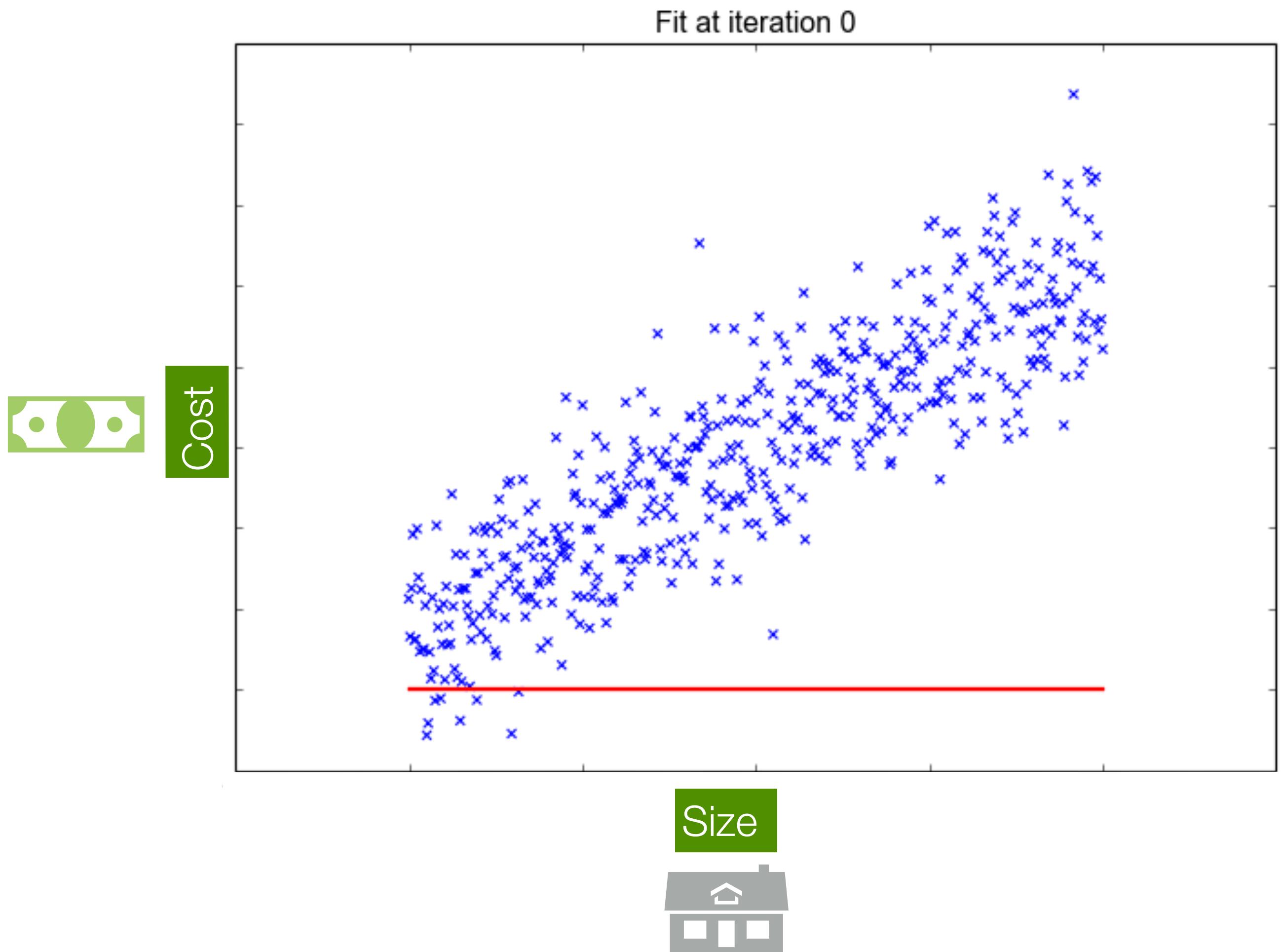
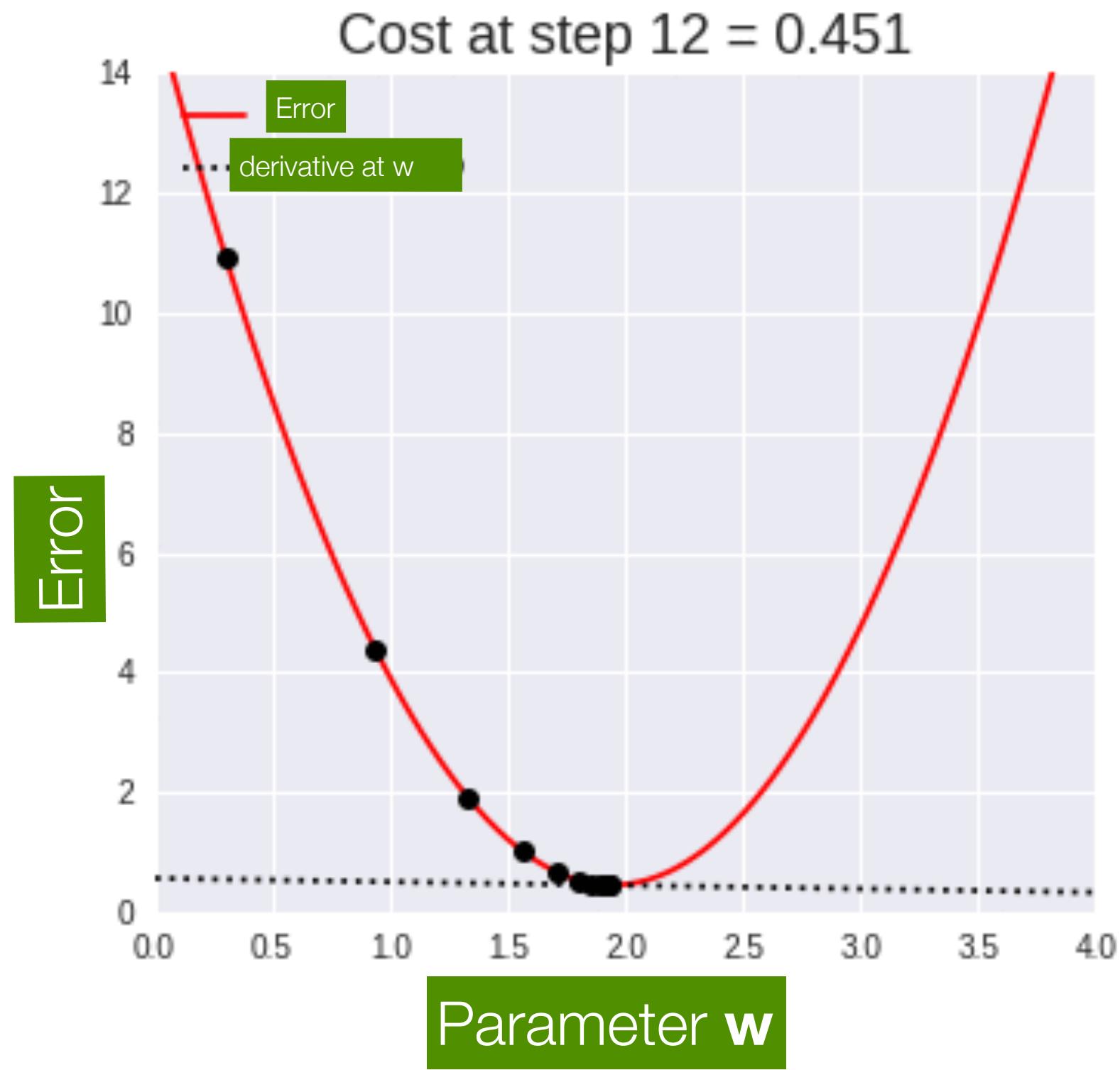
Error₁:
Error₂:



Fit₁ is a better fit on the training data than Fit₂
We select $w = 0.05$

Finding the best parameter values (training the model)

■ Gradient descent



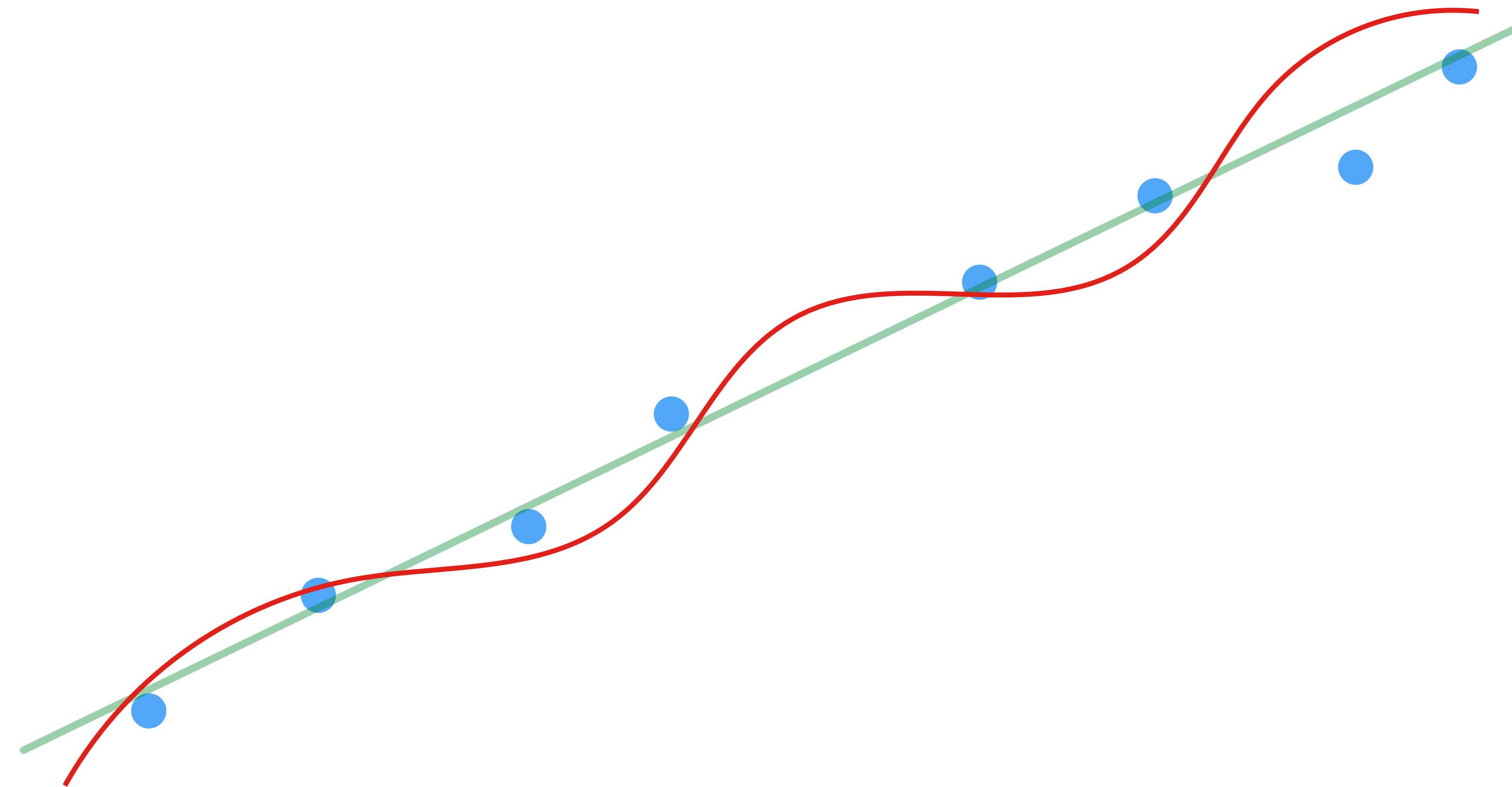
■ Hyperparameters

- **Learning Rate:** “speed” of descent
- **Epochs:** max number of steps

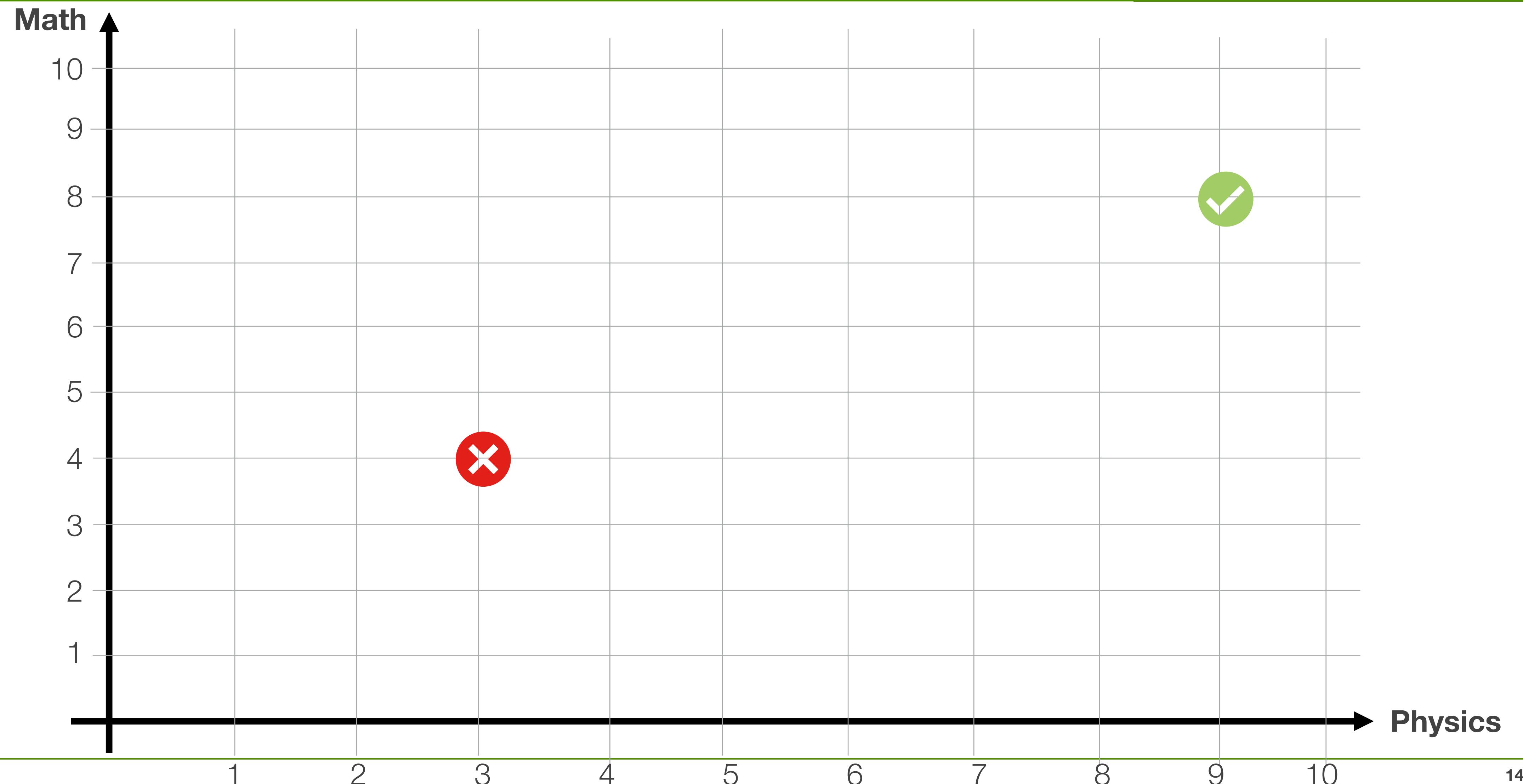
Polynomial Regression

Nth degree polynomial

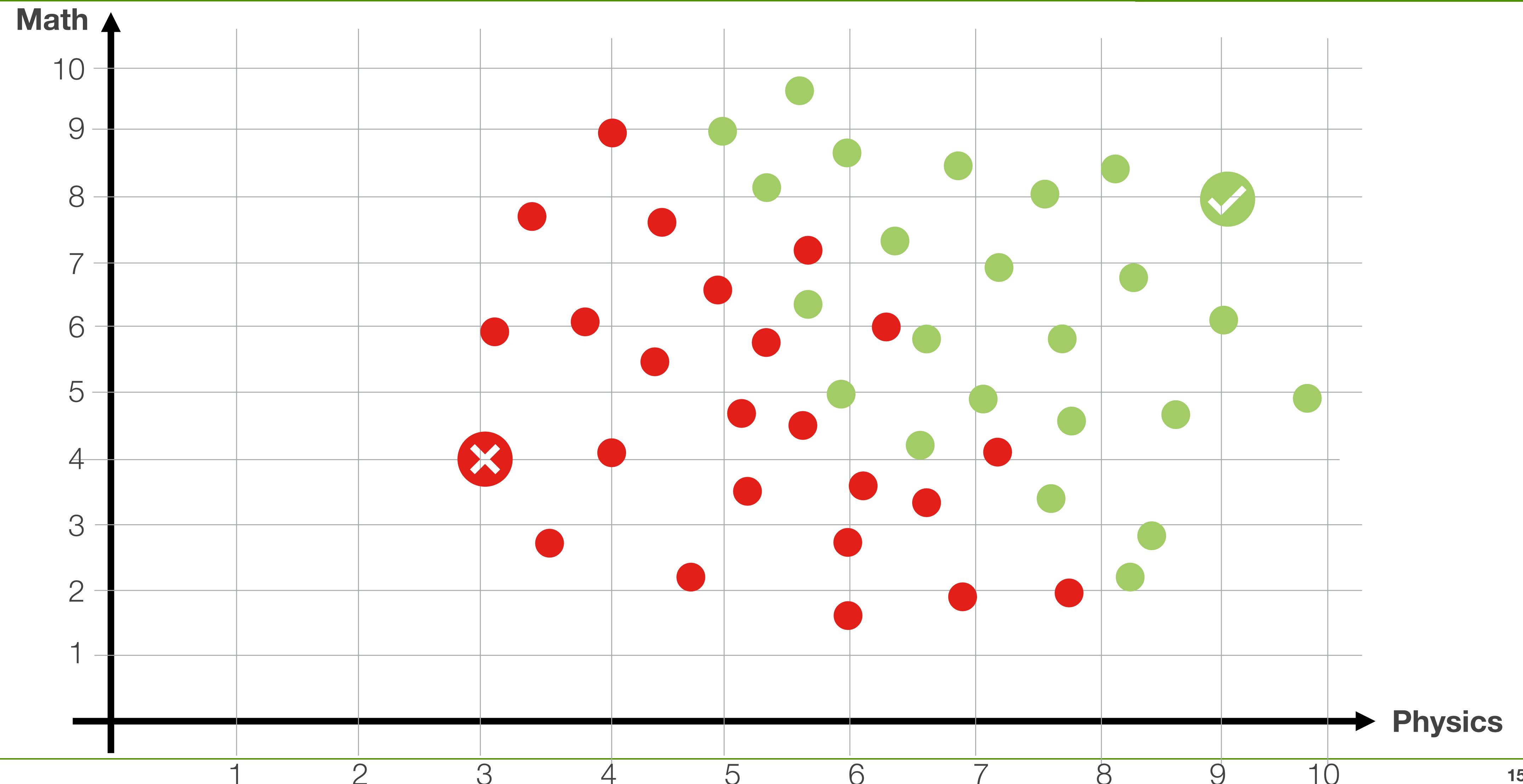
$$\rightarrow \text{Cost} = b + w_1 \text{ Size} + w_2 \text{ Size}^2 + \dots + w_n \text{ Size}^n$$



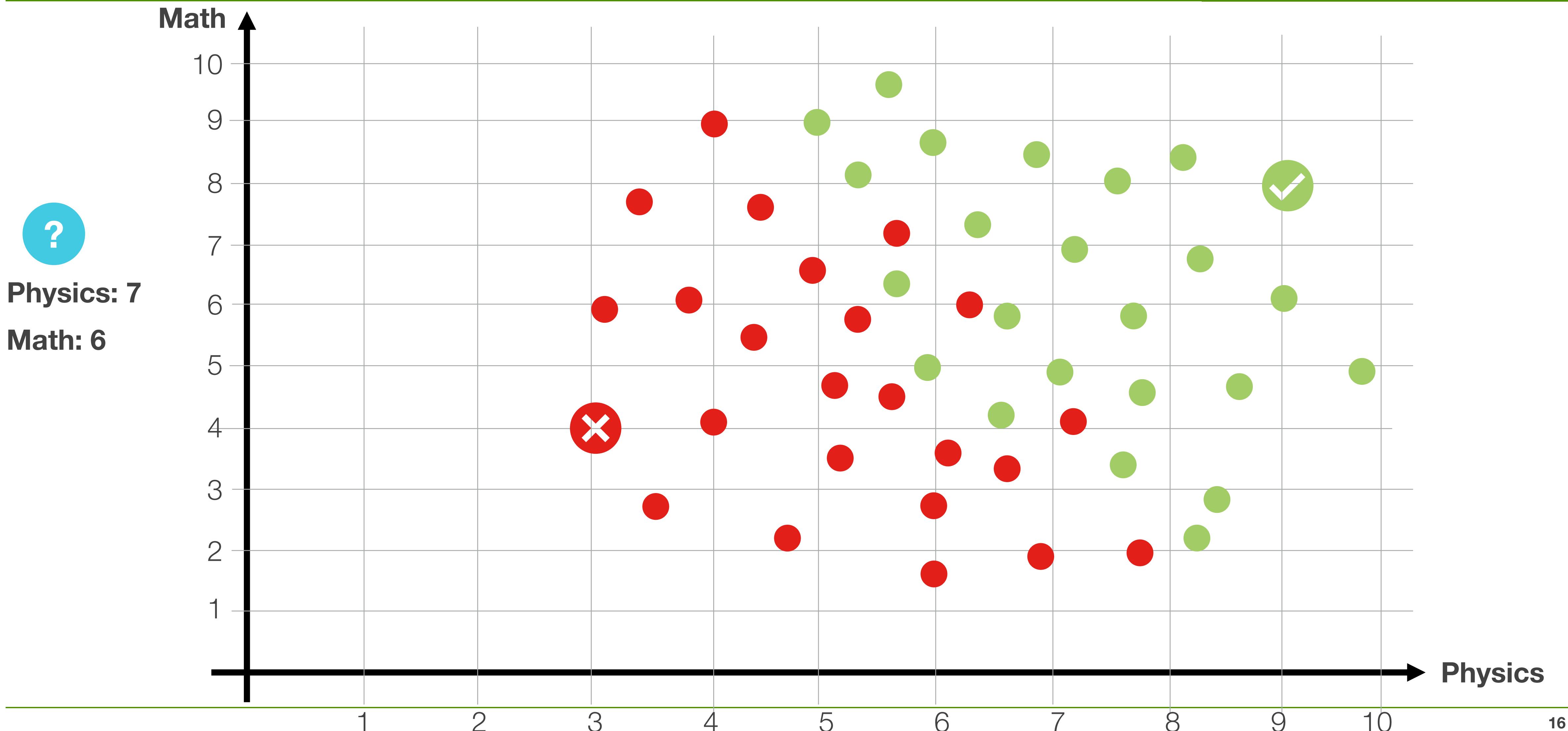
Classification



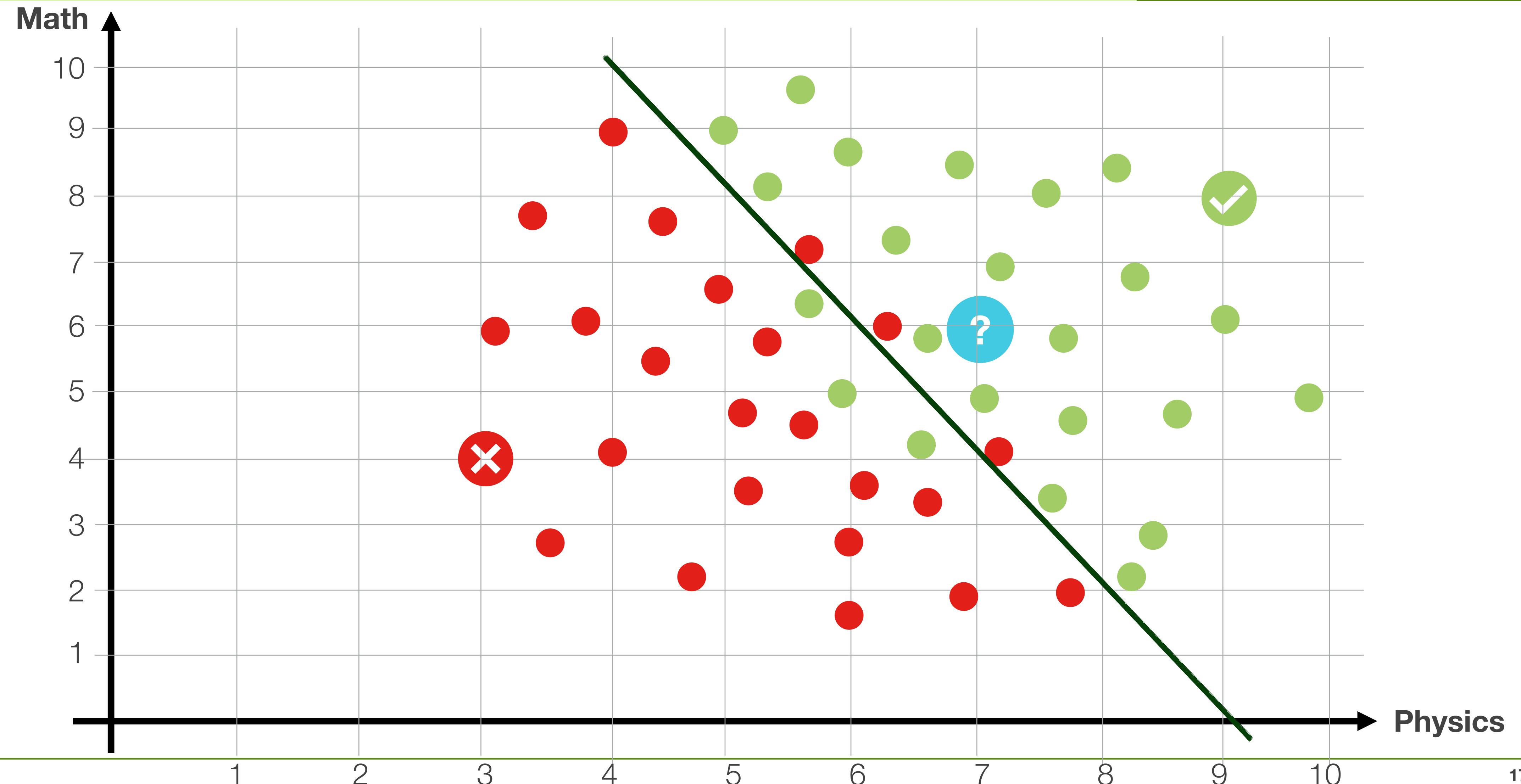
Classification

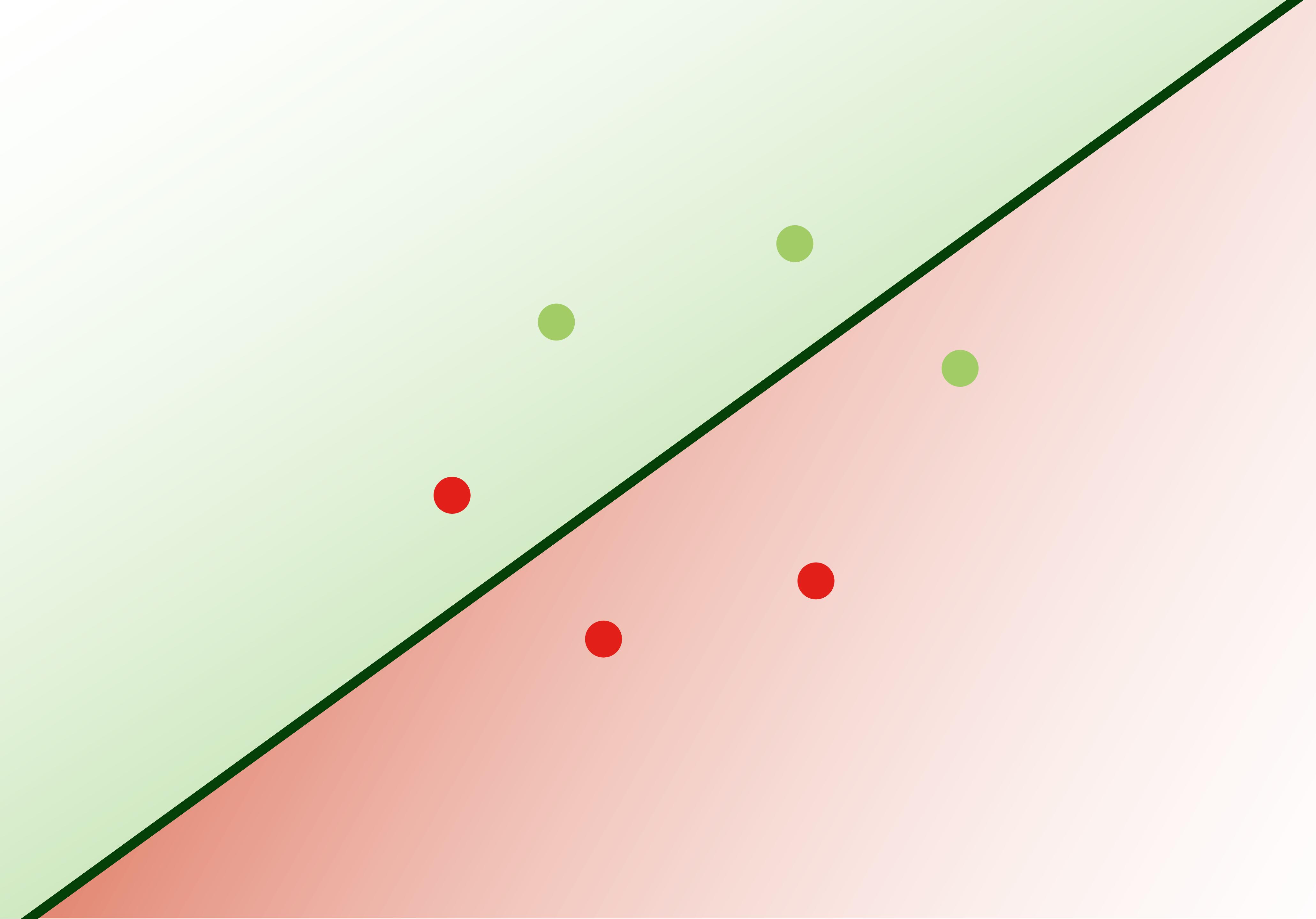


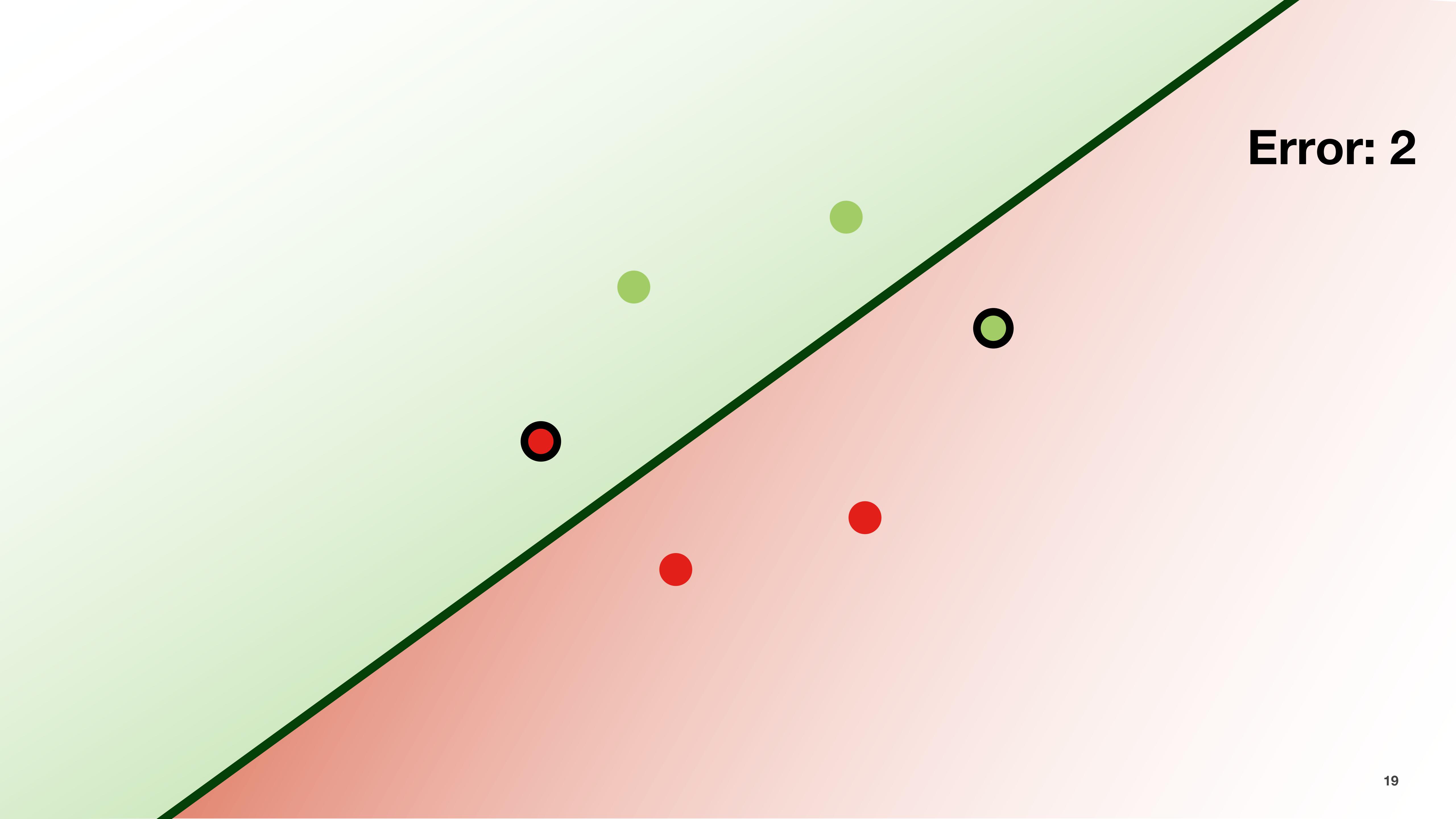
Classification



Logistic Regression

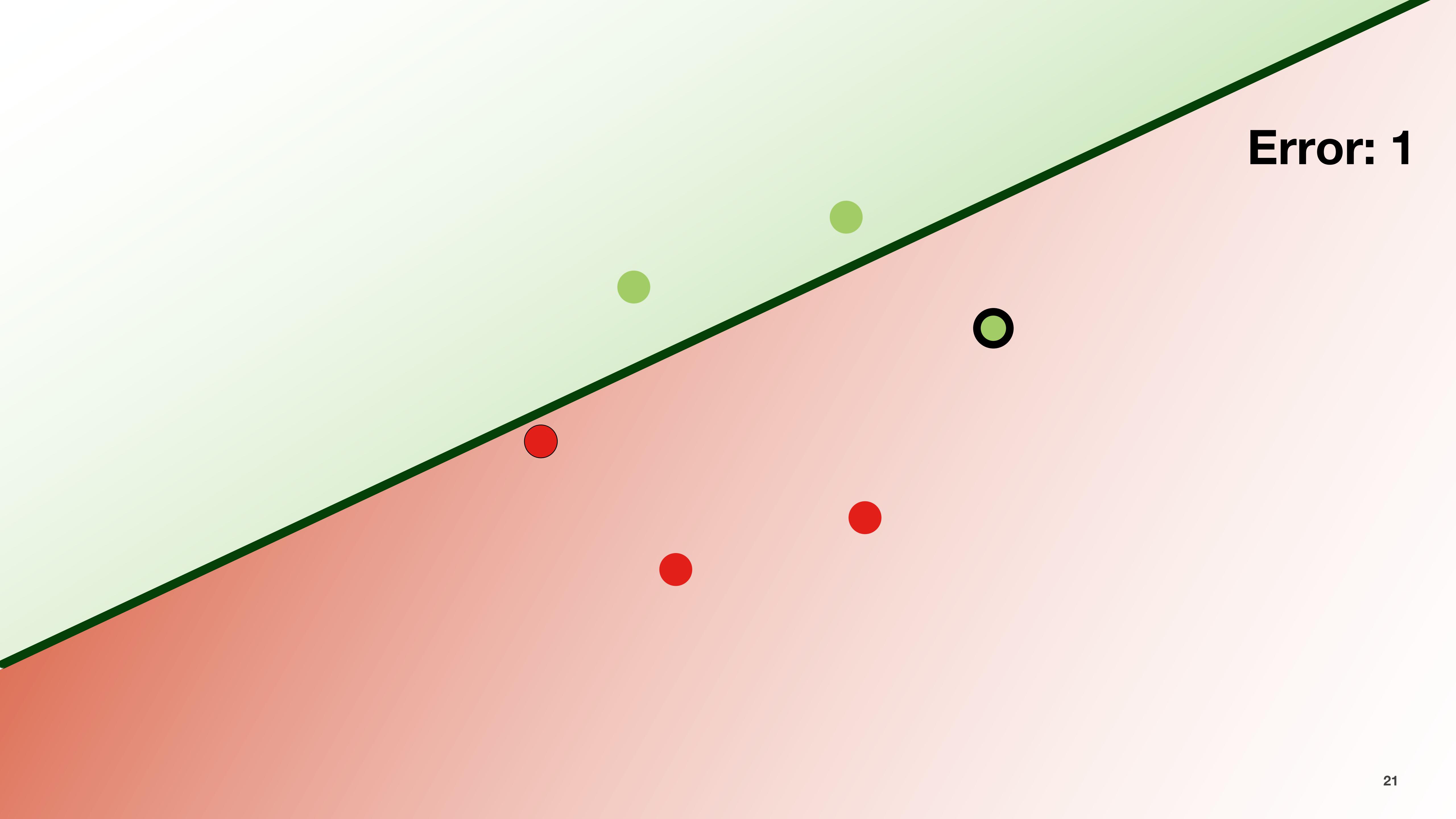






Error: 2

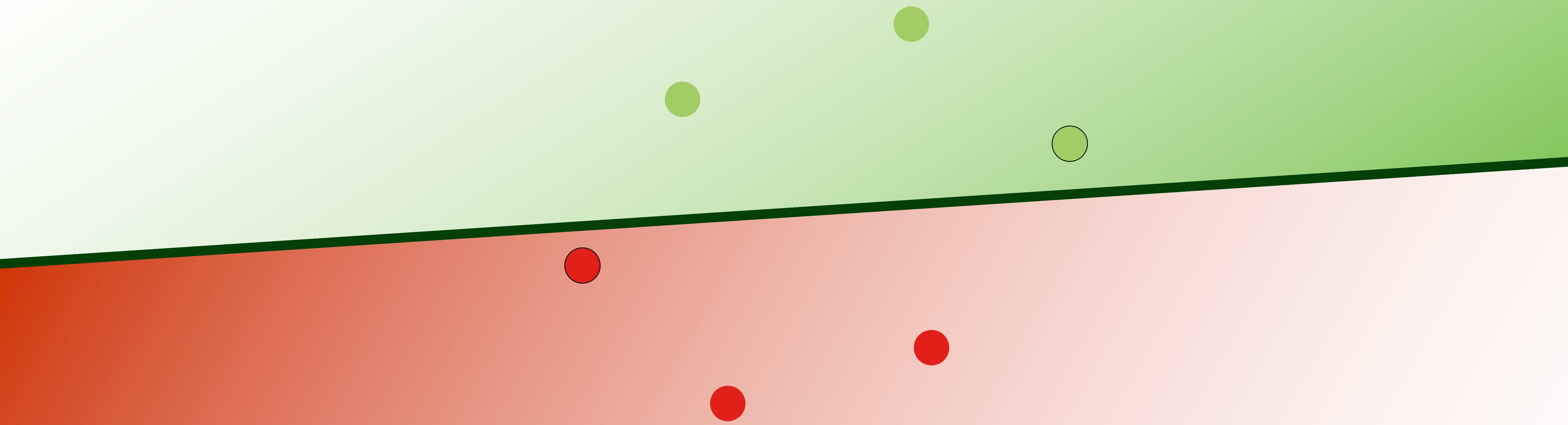
Error: 2



A scatter plot illustrating a linear classification model. The background is divided into two regions by a diagonal green line with a black border. The upper-right region is light green, while the lower-left region is light orange. There are six data points: one green point in the green region, three red points in the orange region, and two black-outlined green points. The text "Error: 1" is located in the top right corner.

Error: 1

Error: 0



Math

10

9

8

7

6

5

4

3

2

1

Physics

1

2

3

4

5

6

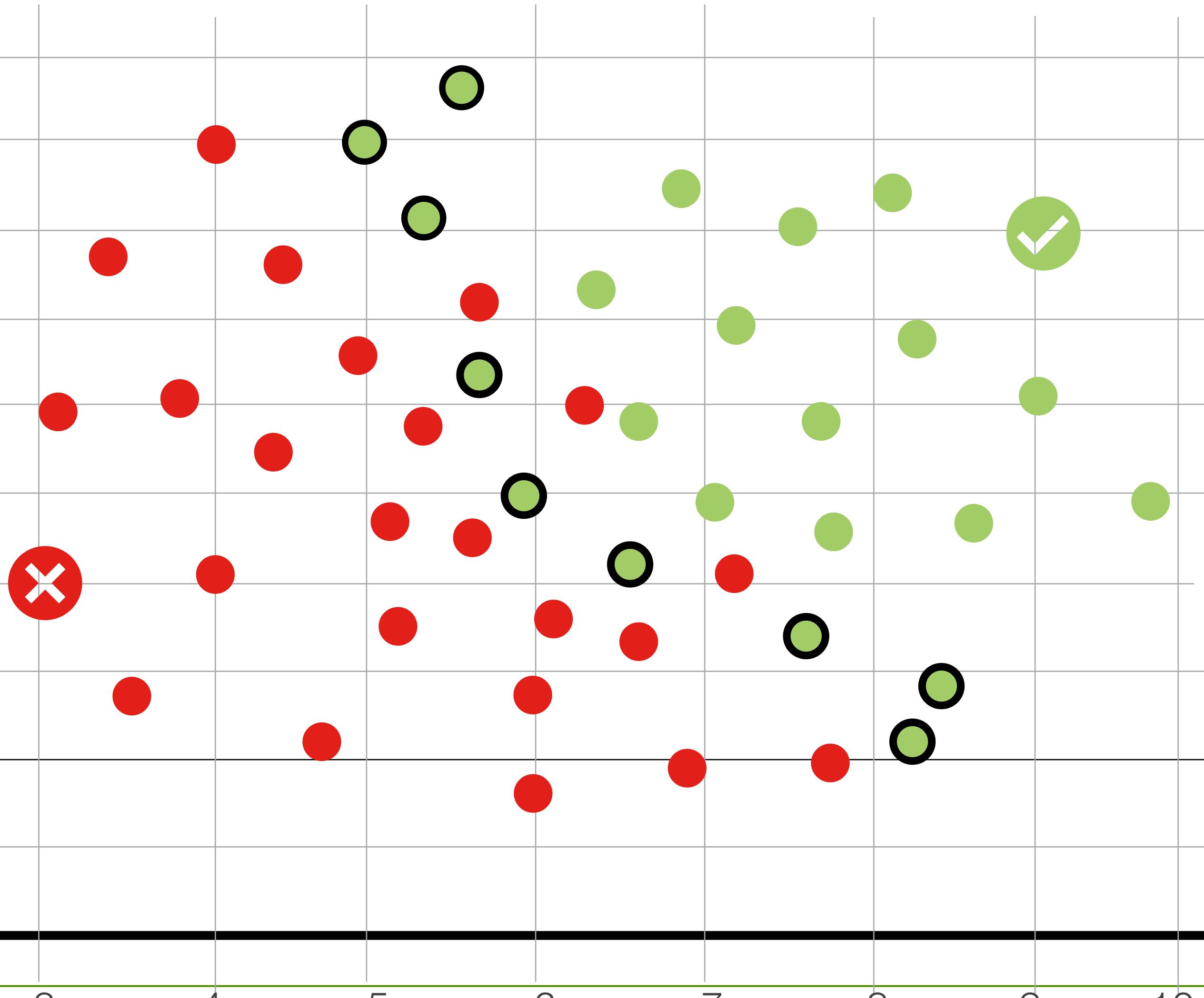
7

8

9

10

23



Math

10

9

8

7

6

5

4

3

2

1

Test

1

2

3

4

5

6

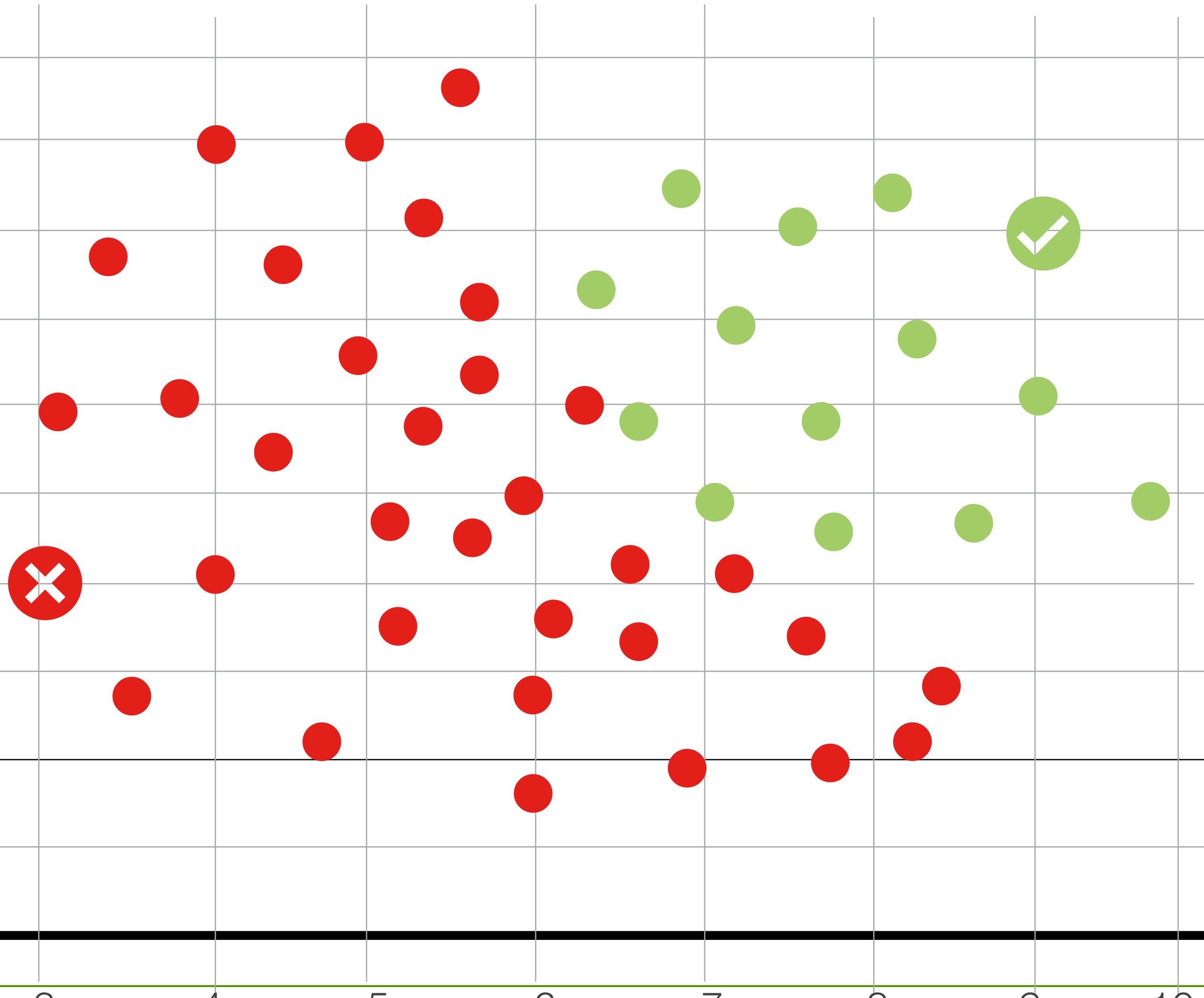
7

8

9

10

24



Math

10

9

8

7

6

5

4

3

2

1

Test

1

2

3

4

5

6

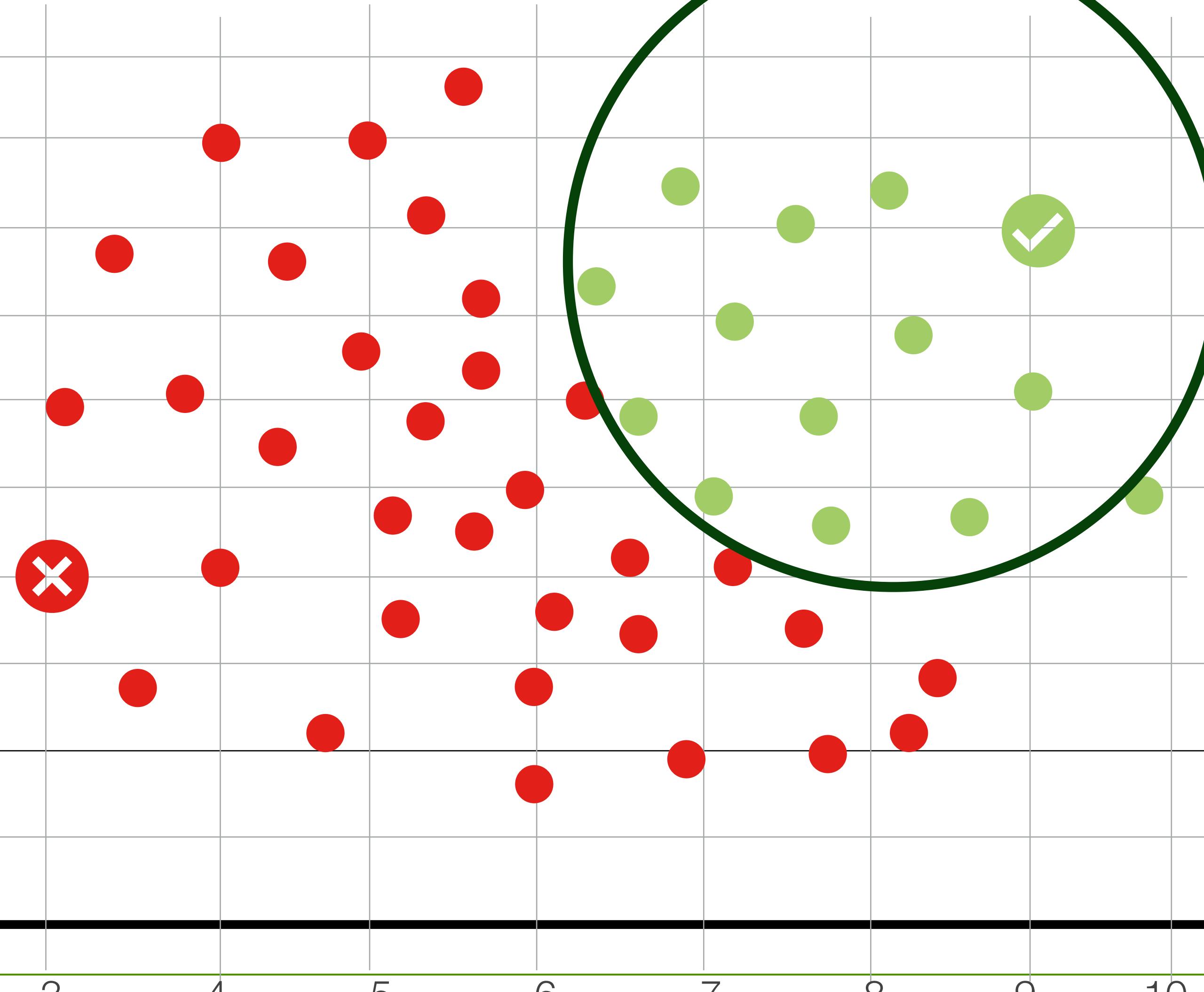
7

8

9

10

25



Math

10

9

8

7

6

5

4

3

2

1

1

2

3

4

5

6

7

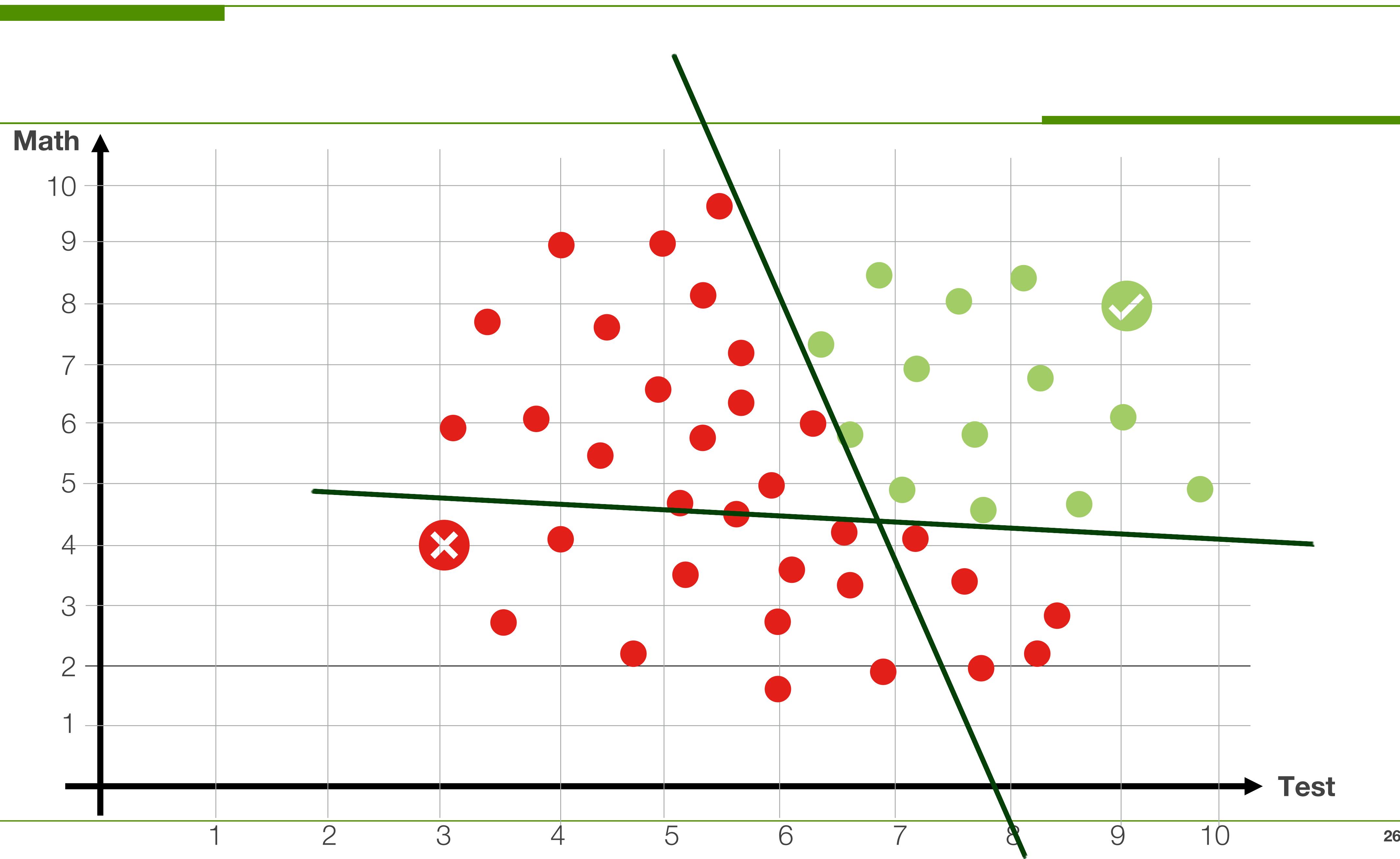
8

9

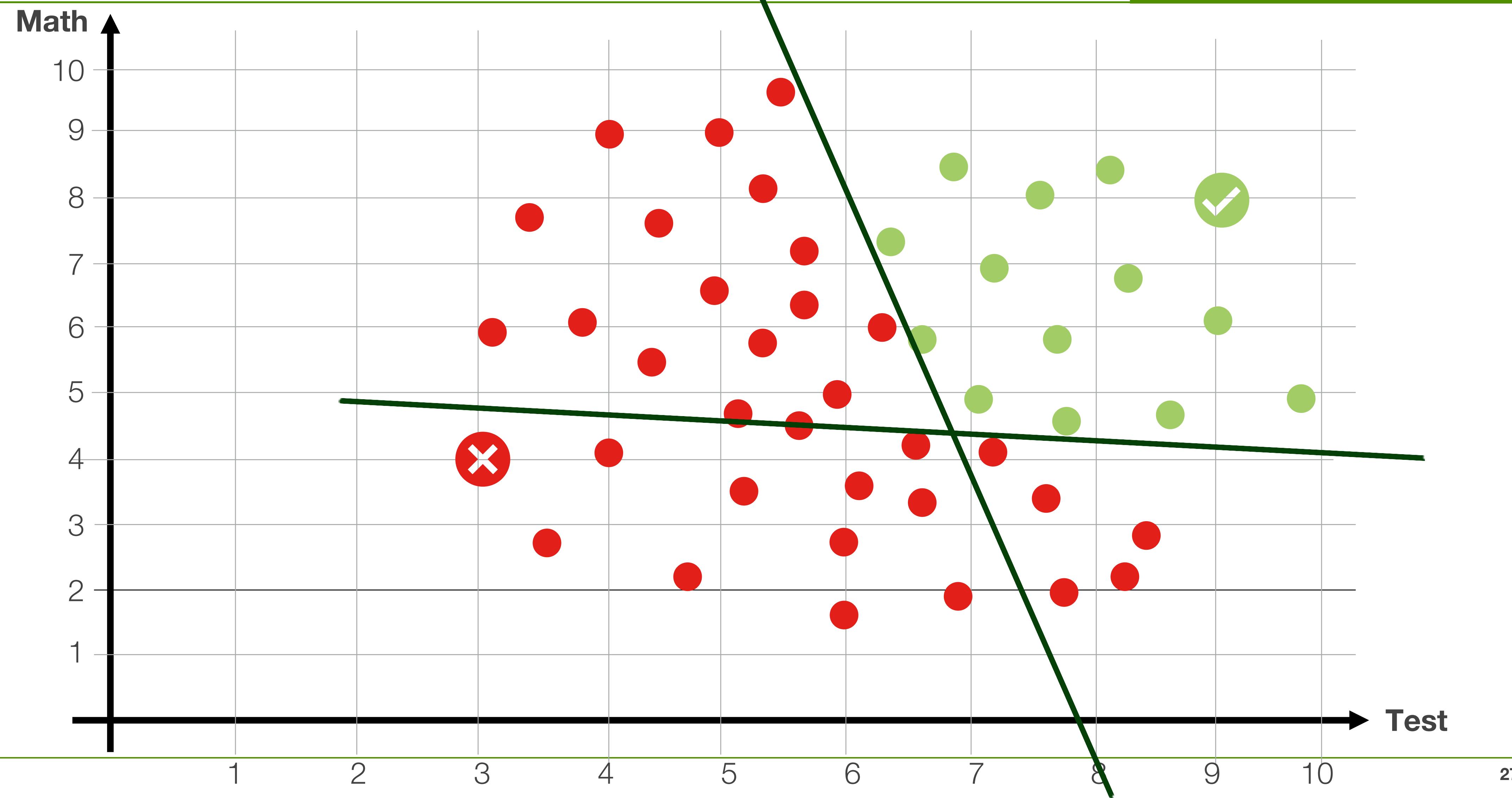
10

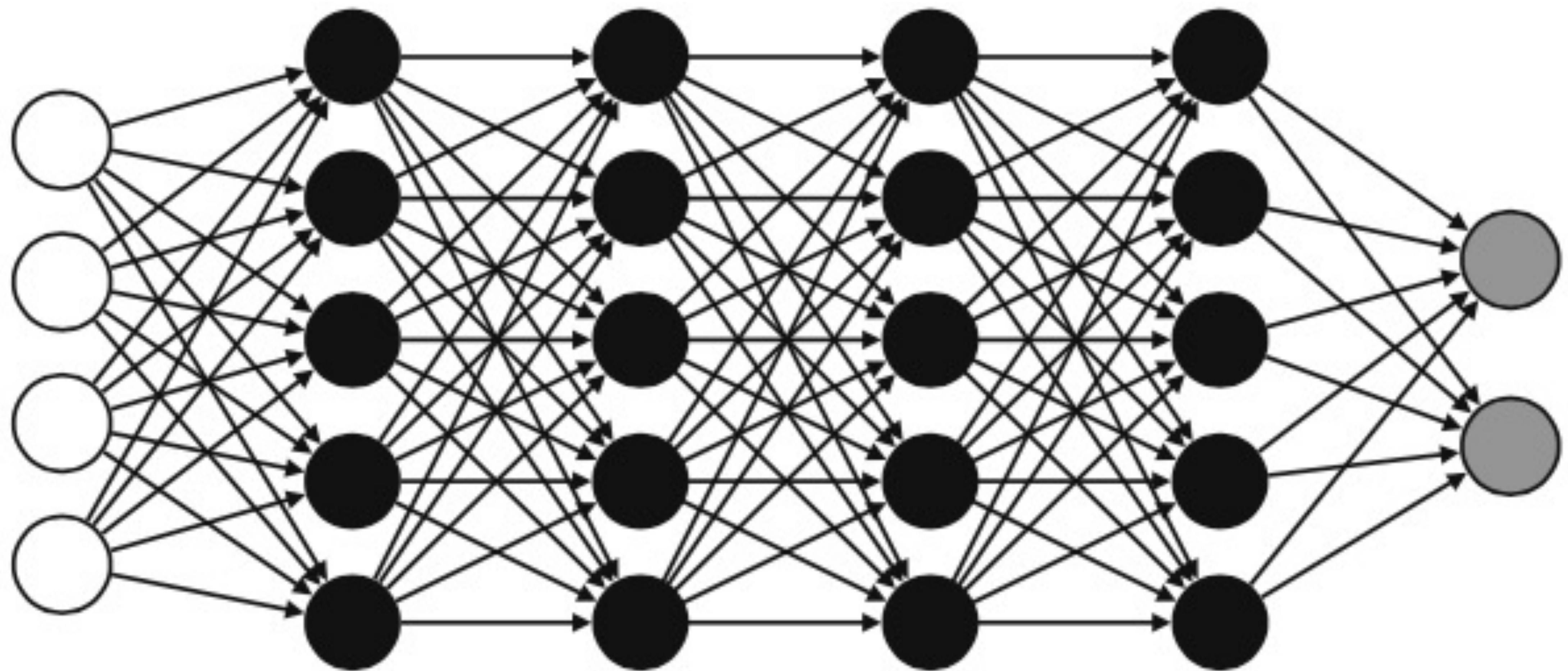
26

Test

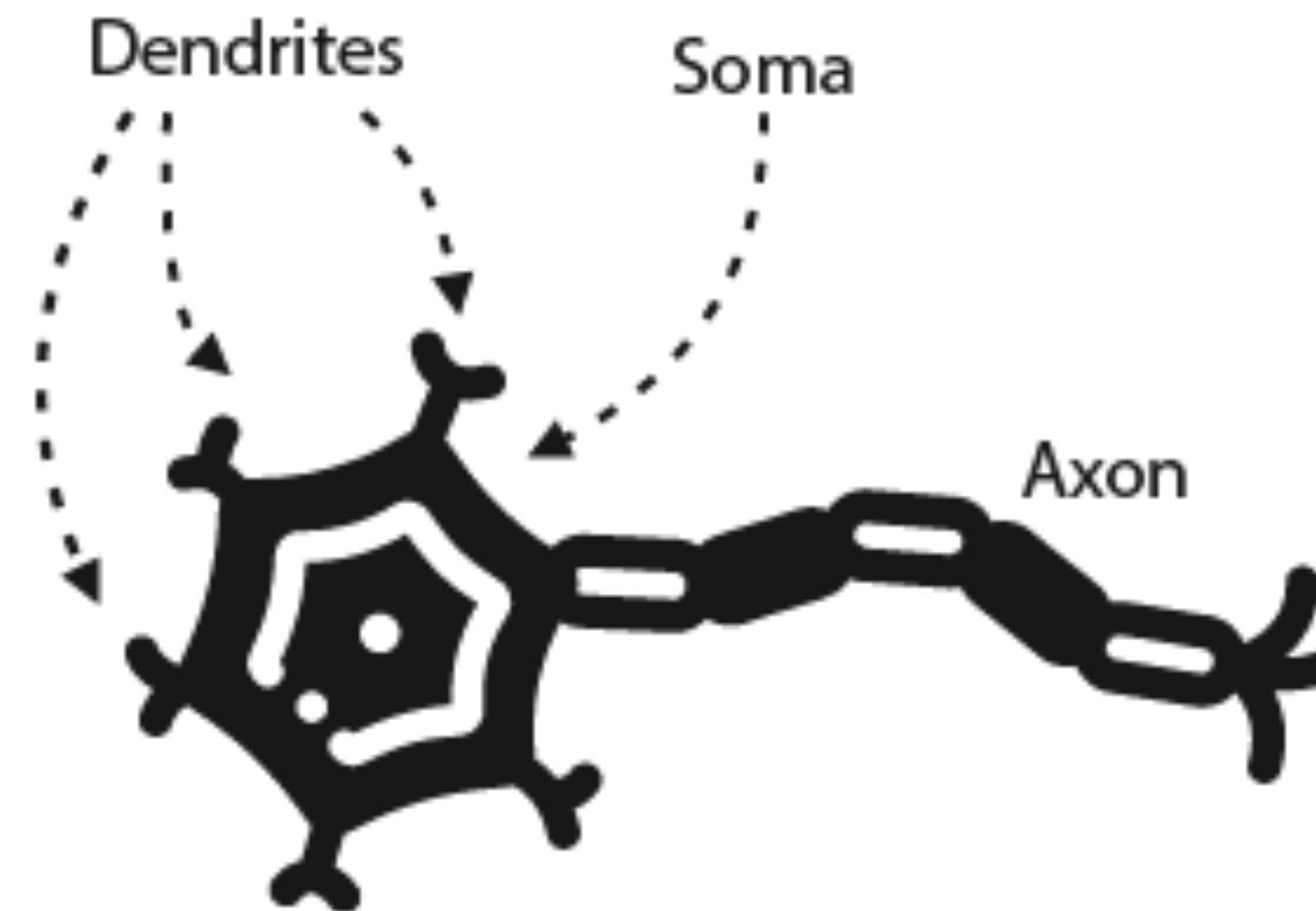


Neural Network

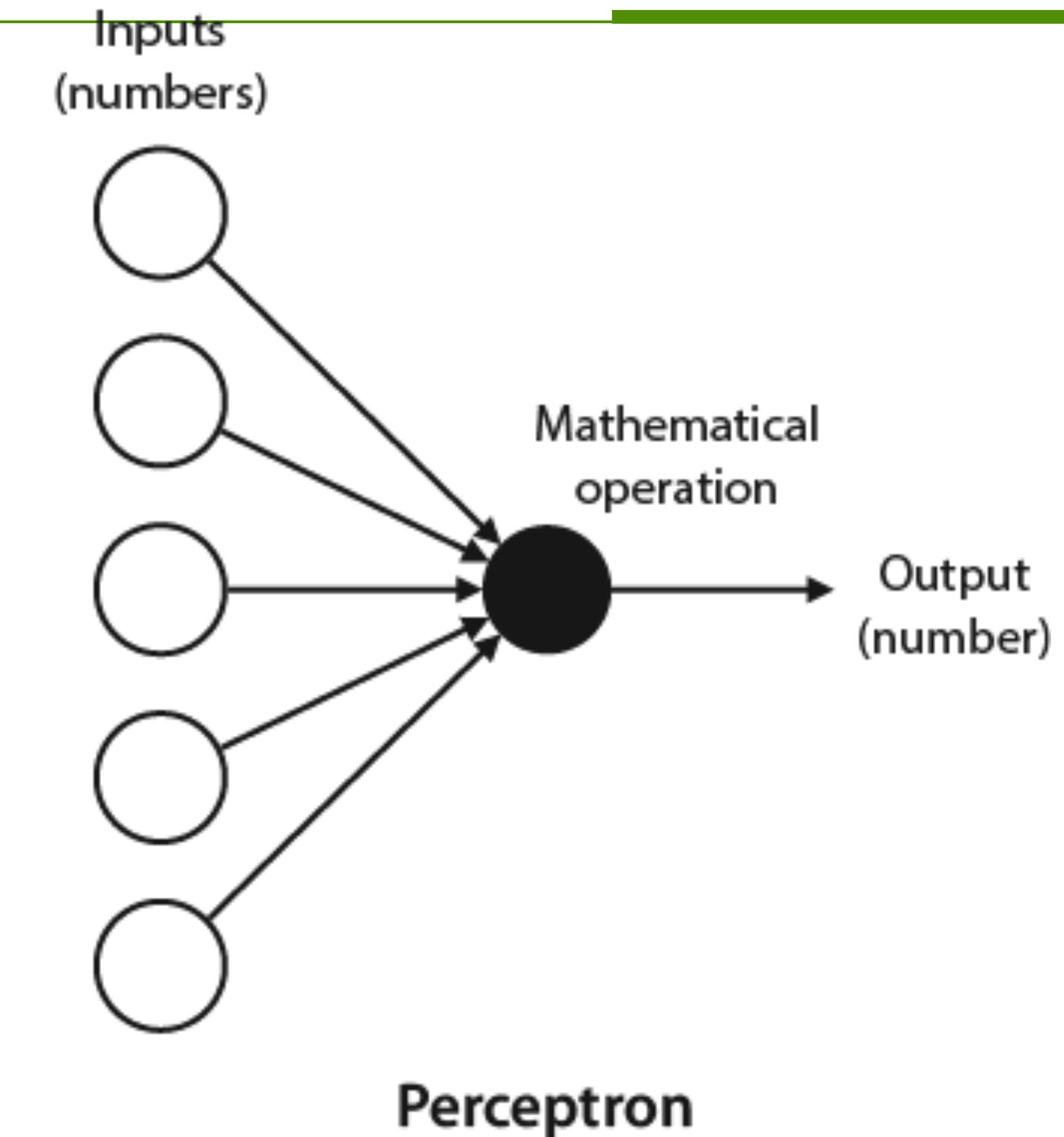




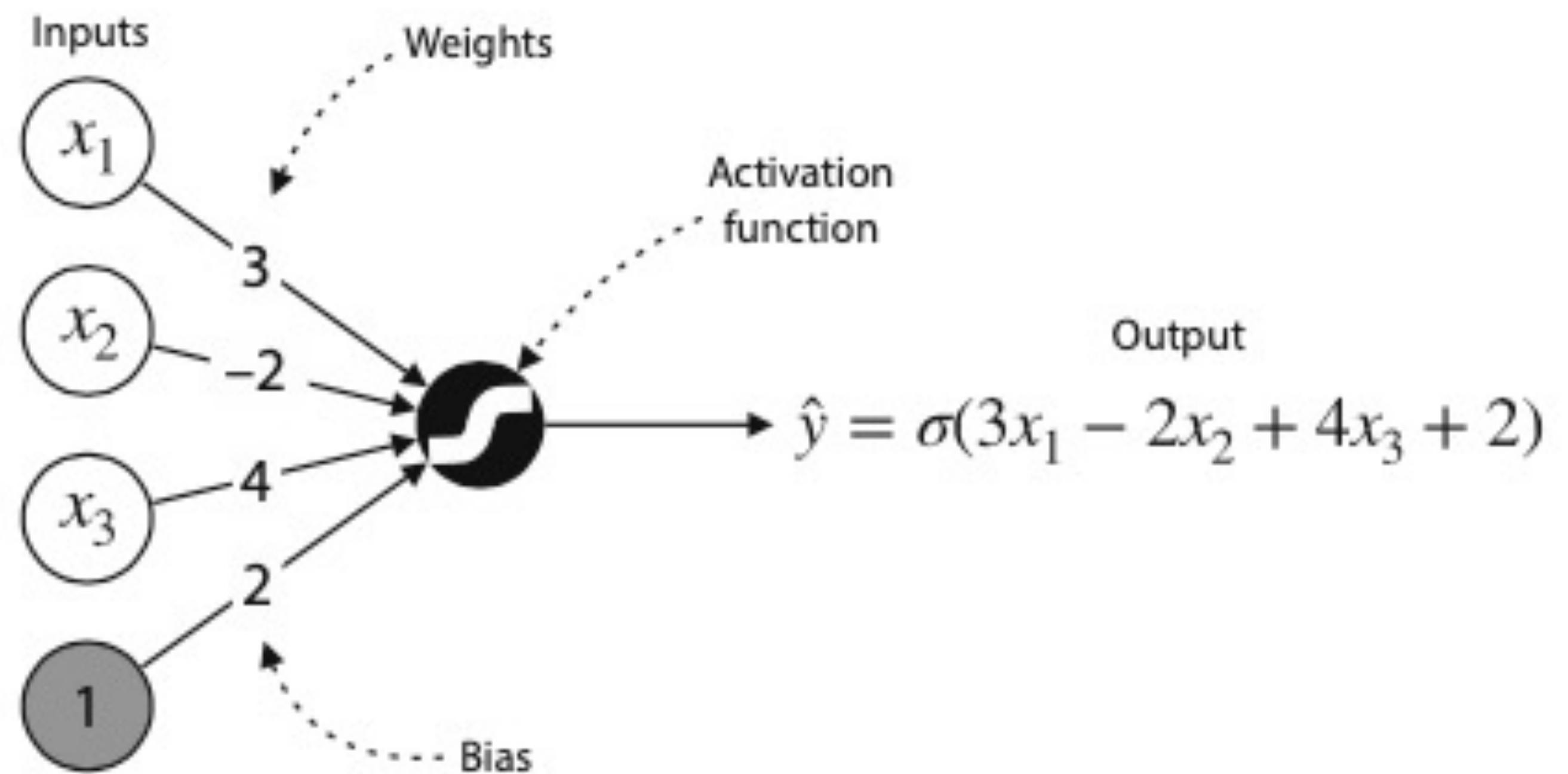
Neural Network



Neuron



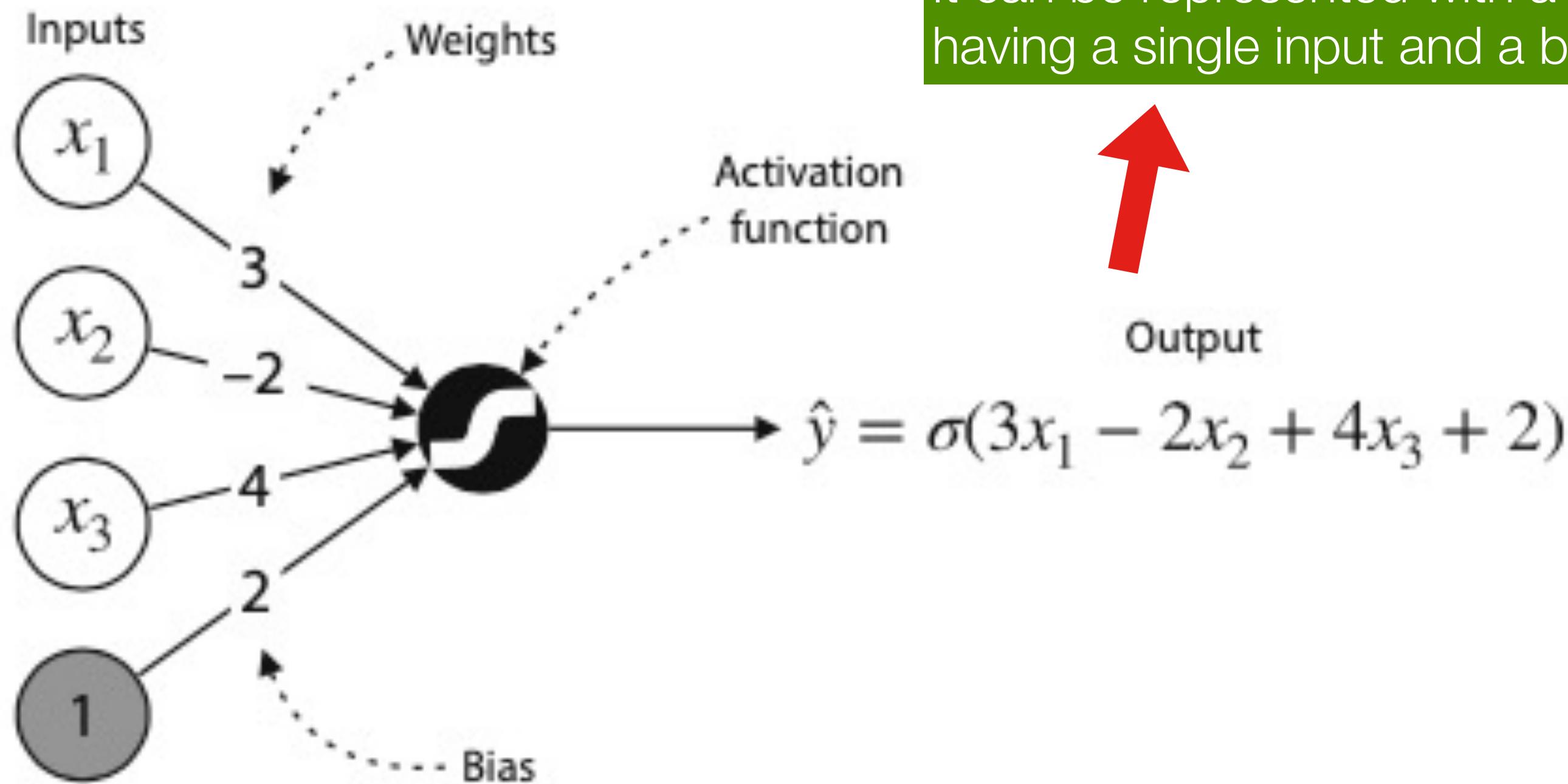
Perceptron



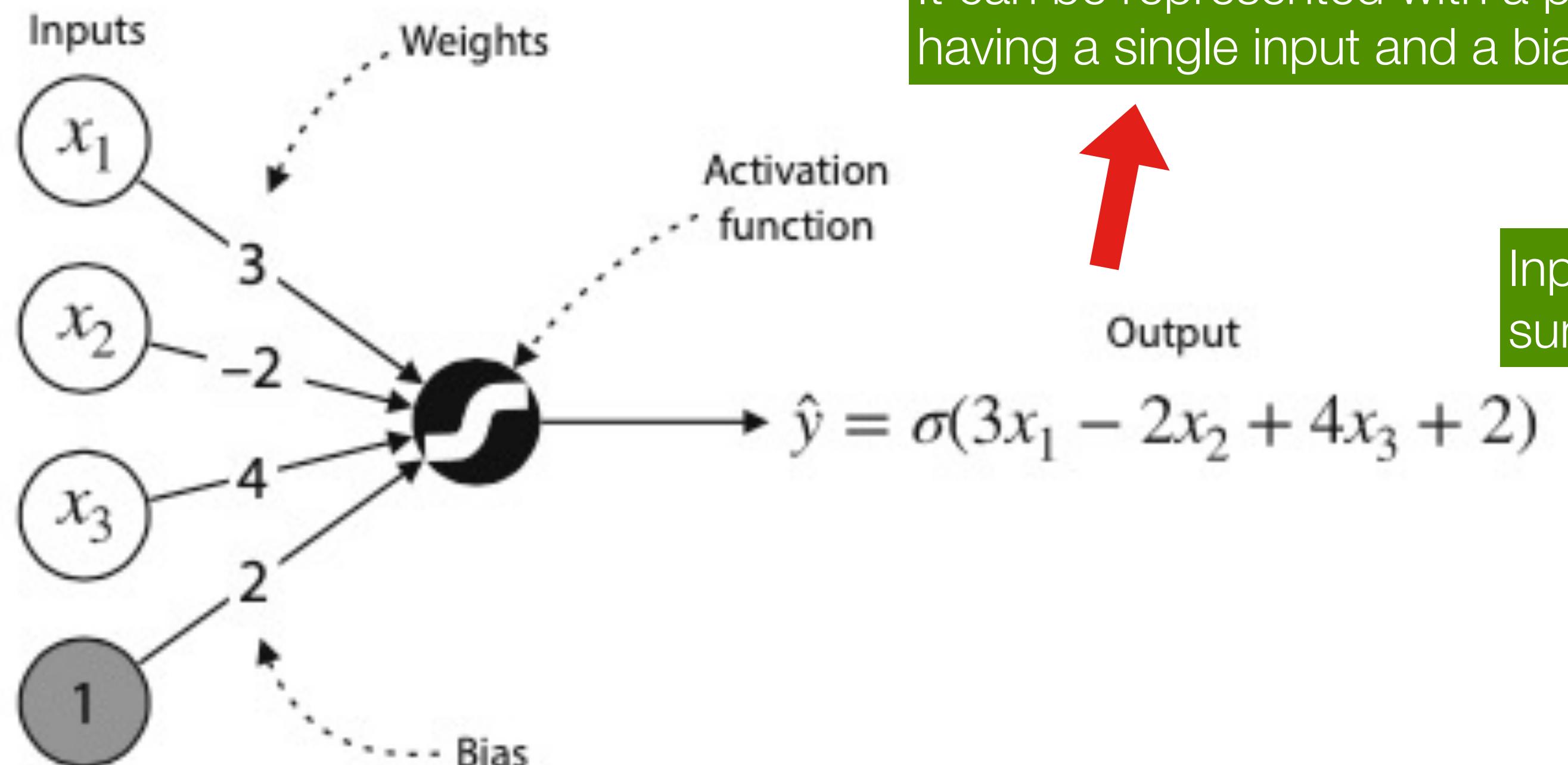
Perceptron

$$\text{Cost} = \mathbf{b} + \mathbf{w}^* \mathbf{Size}$$

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



Perceptron



$$\text{Cost} = a + w^* \text{Size}$$

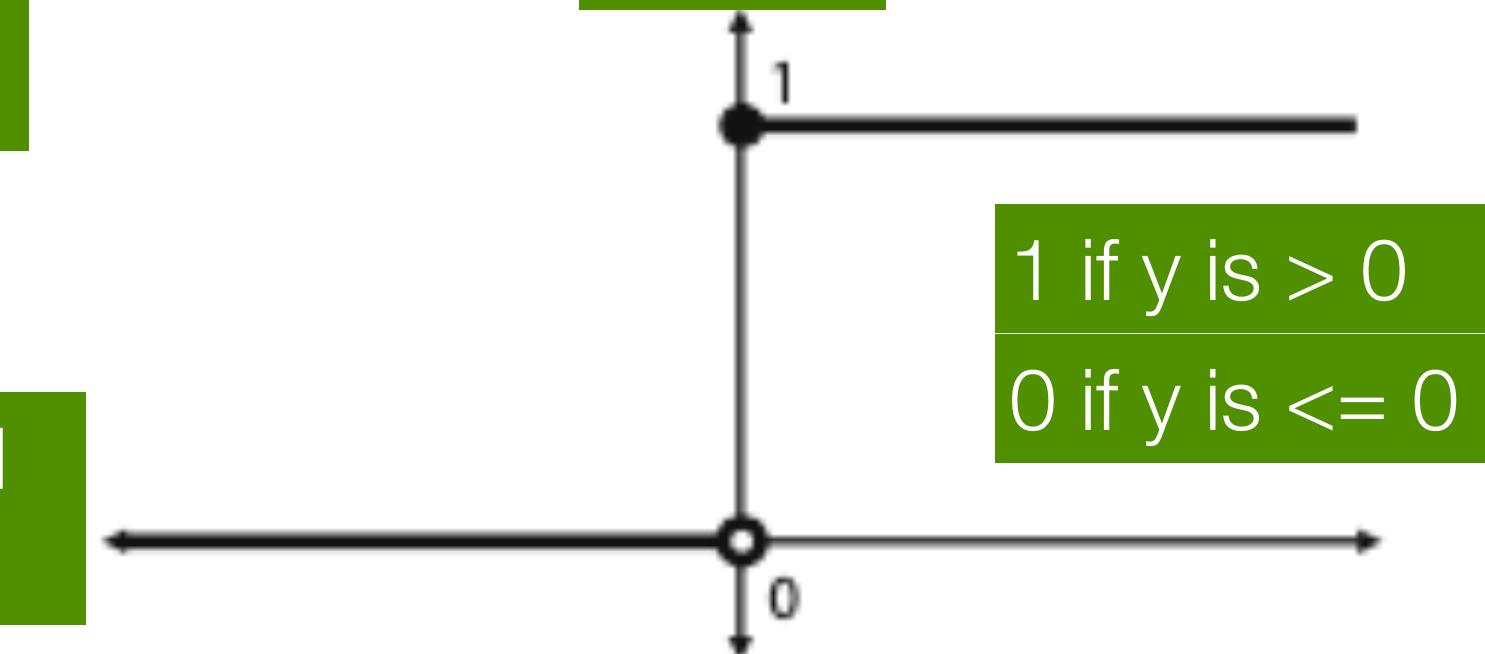
Remember our linear regression function?
It can be represented with a perceptron
having a single input and a bias

Input (weighted
sum of values)

Input (weighted
sum of values)

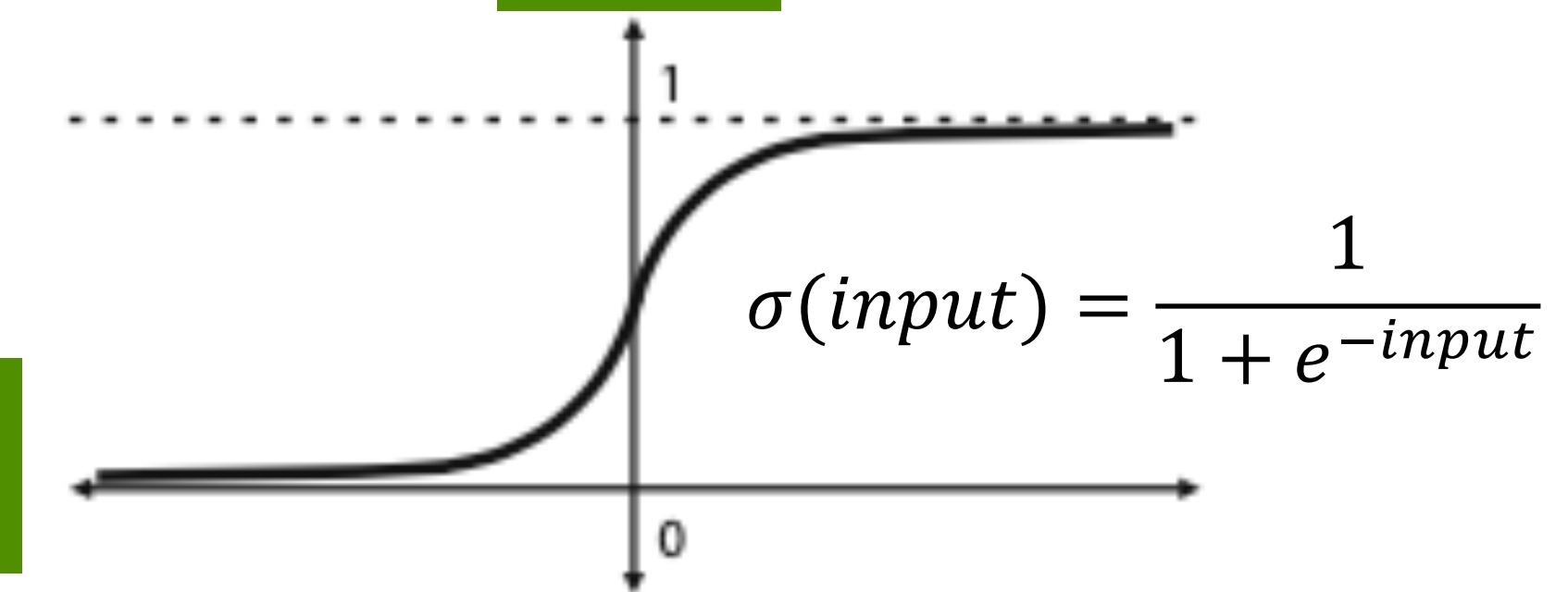
Step function
(discrete)

Output

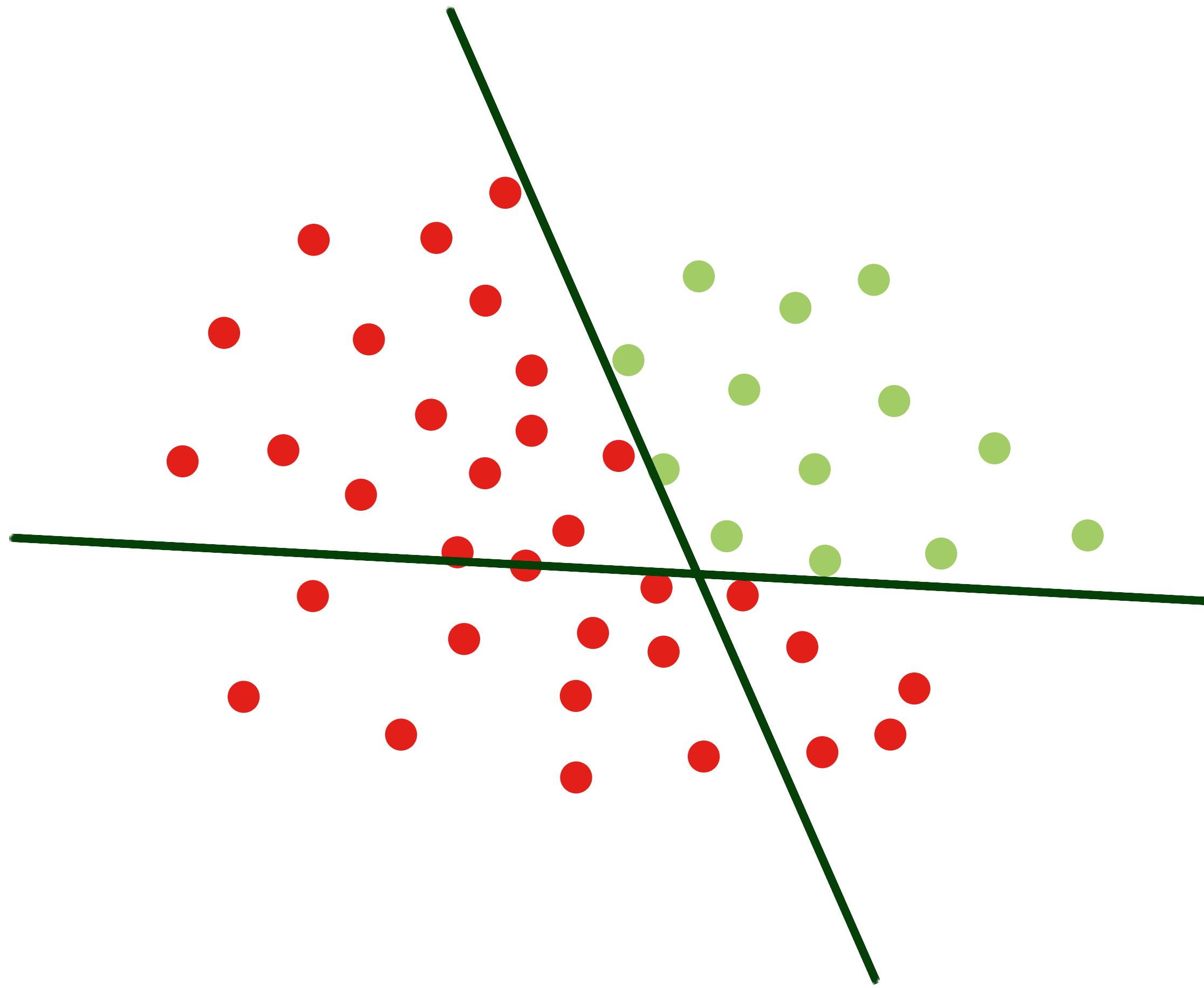


Sigmoid function
(continuous)

Output

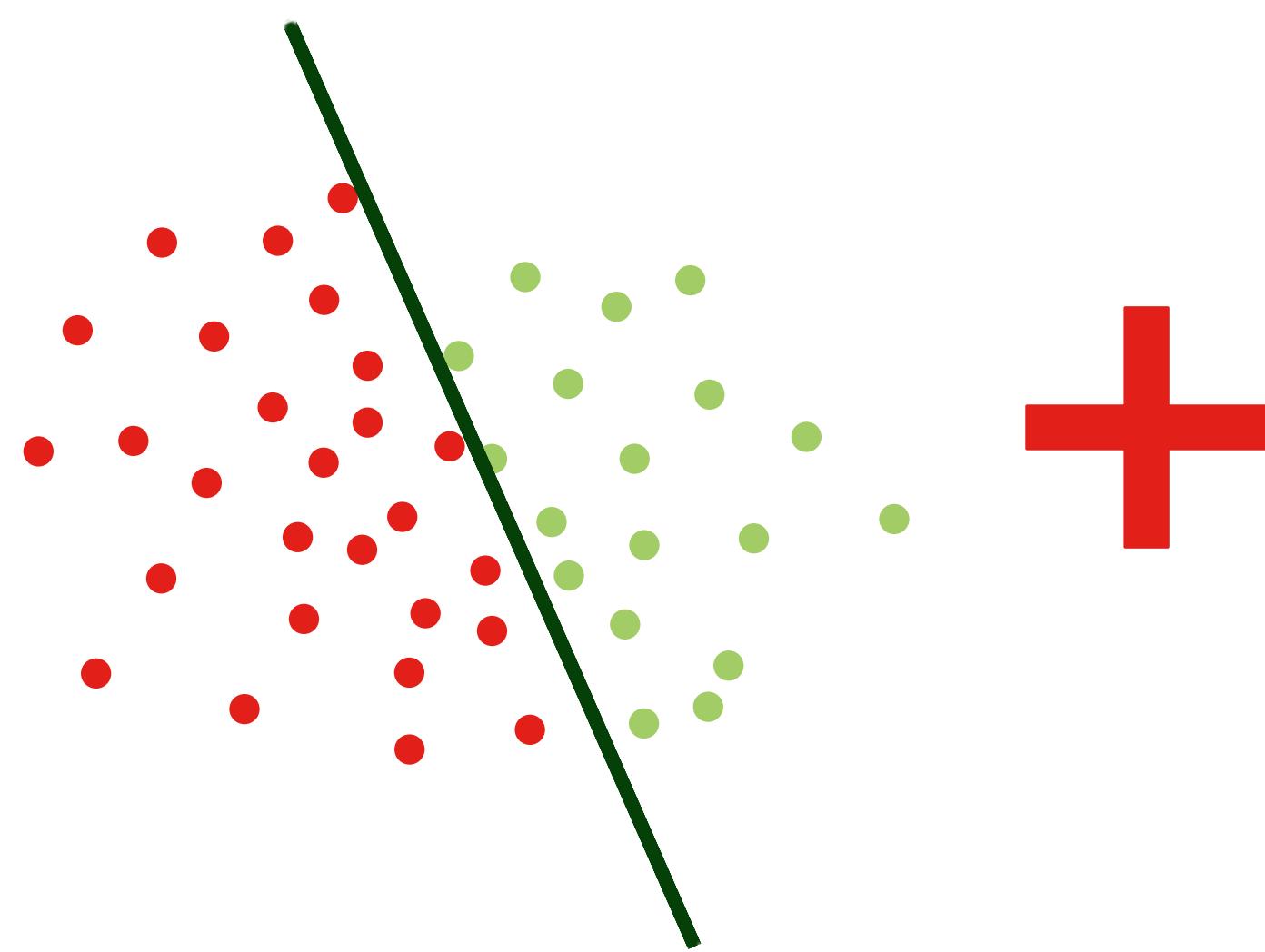


Neural Network

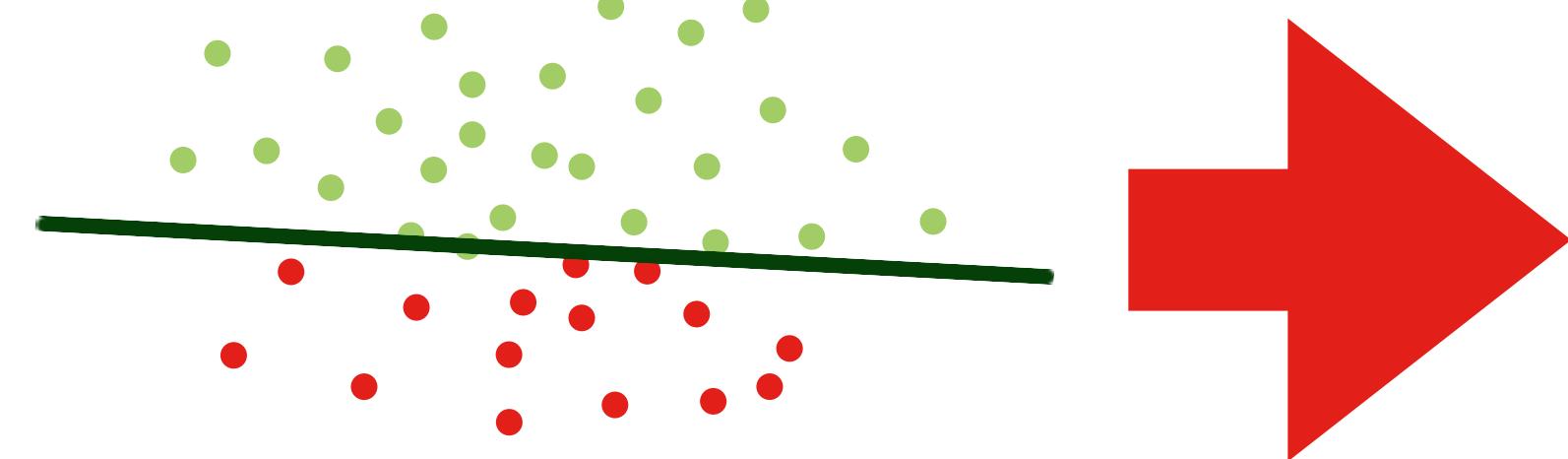


Neural Network

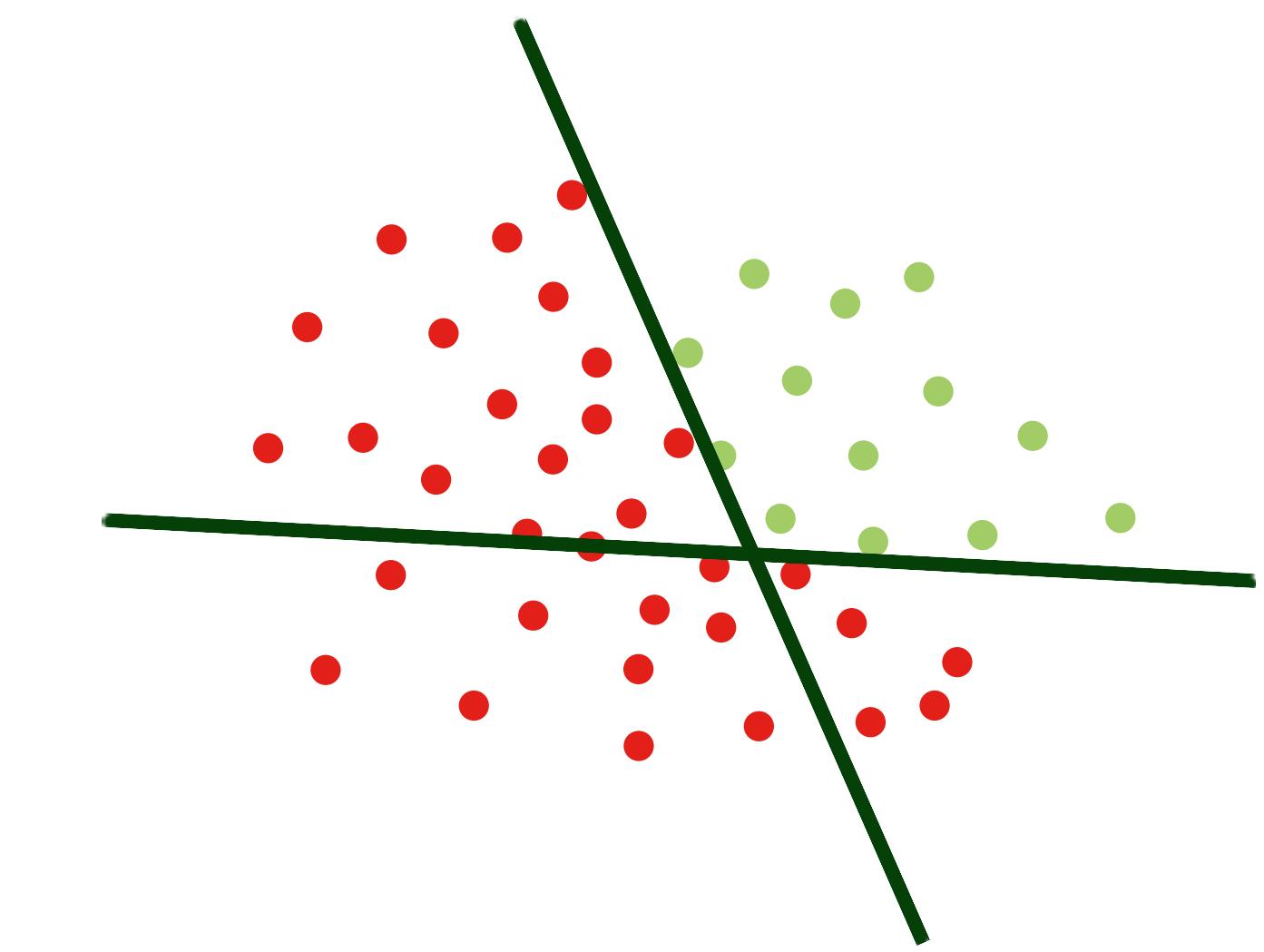
Physics Grade
Classifier



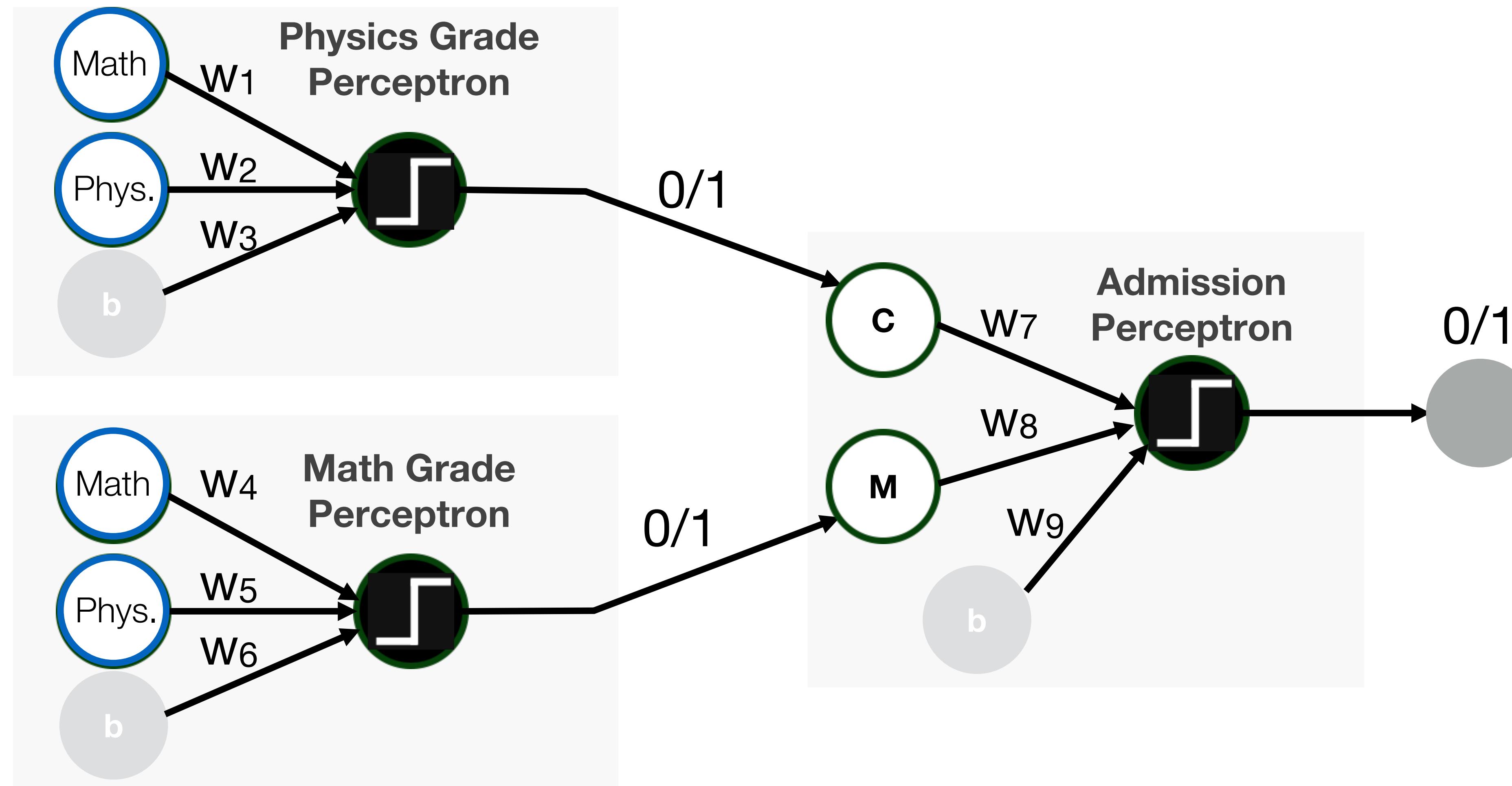
Math Grade
Classifier



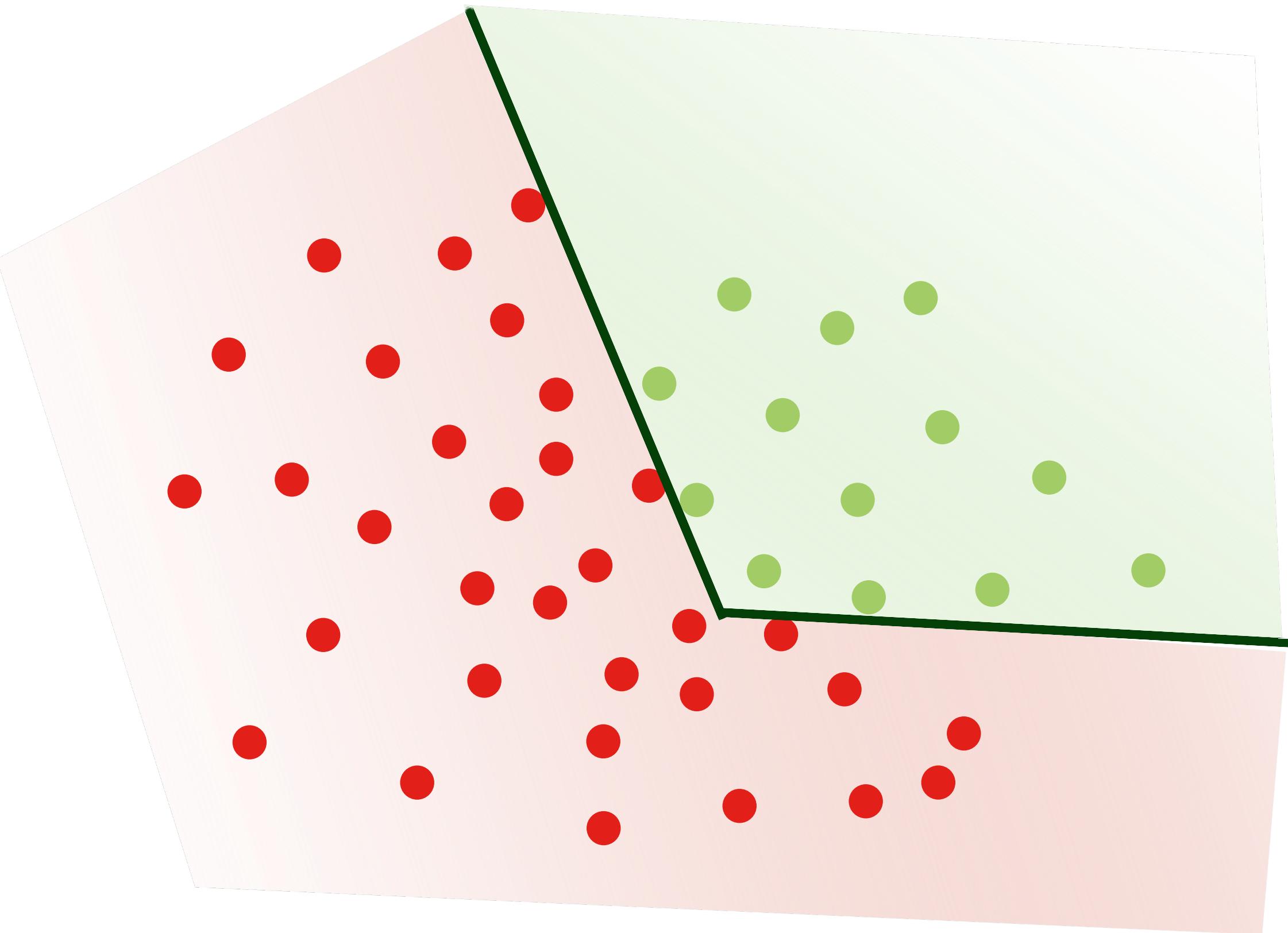
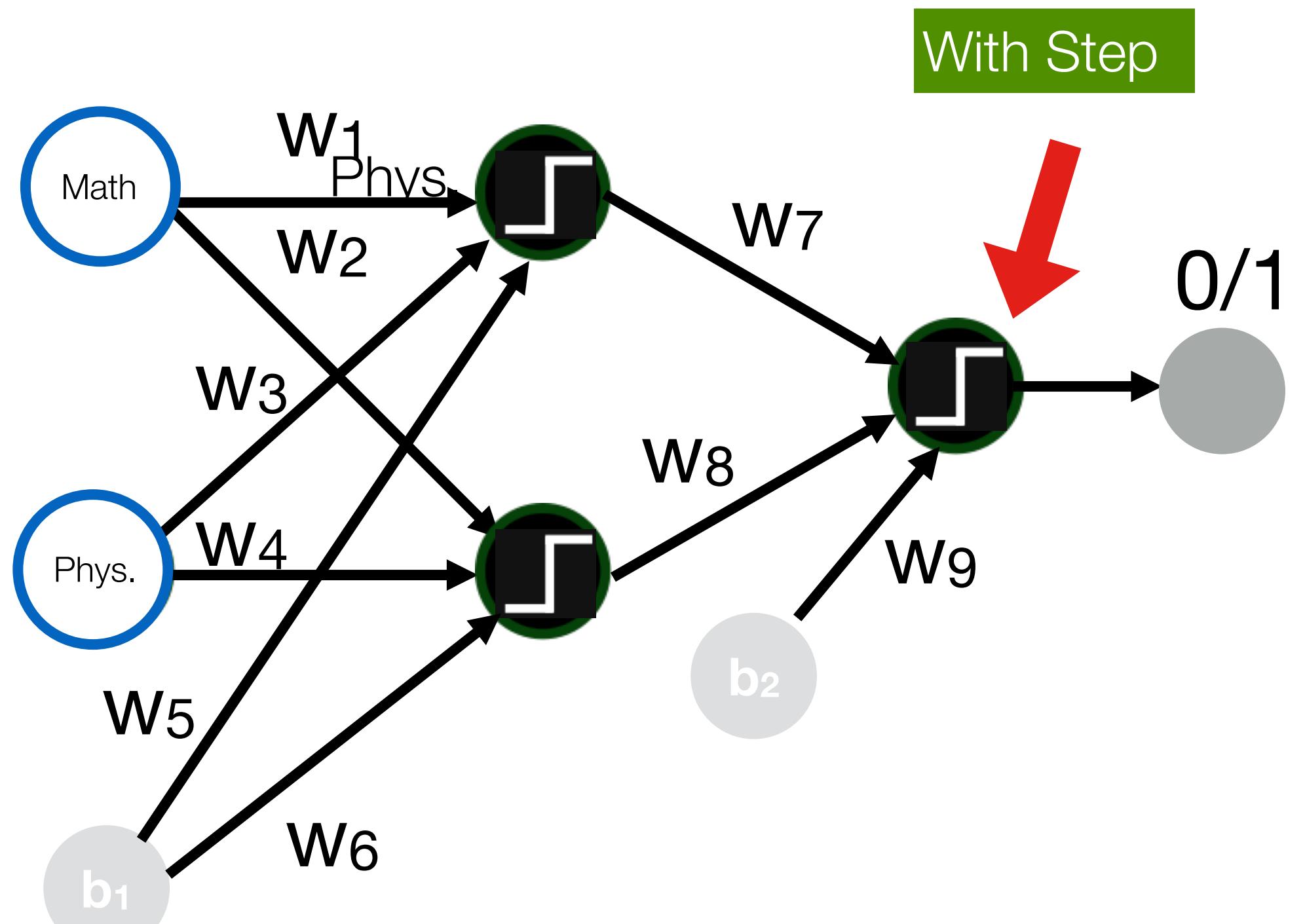
Admission
Classifier



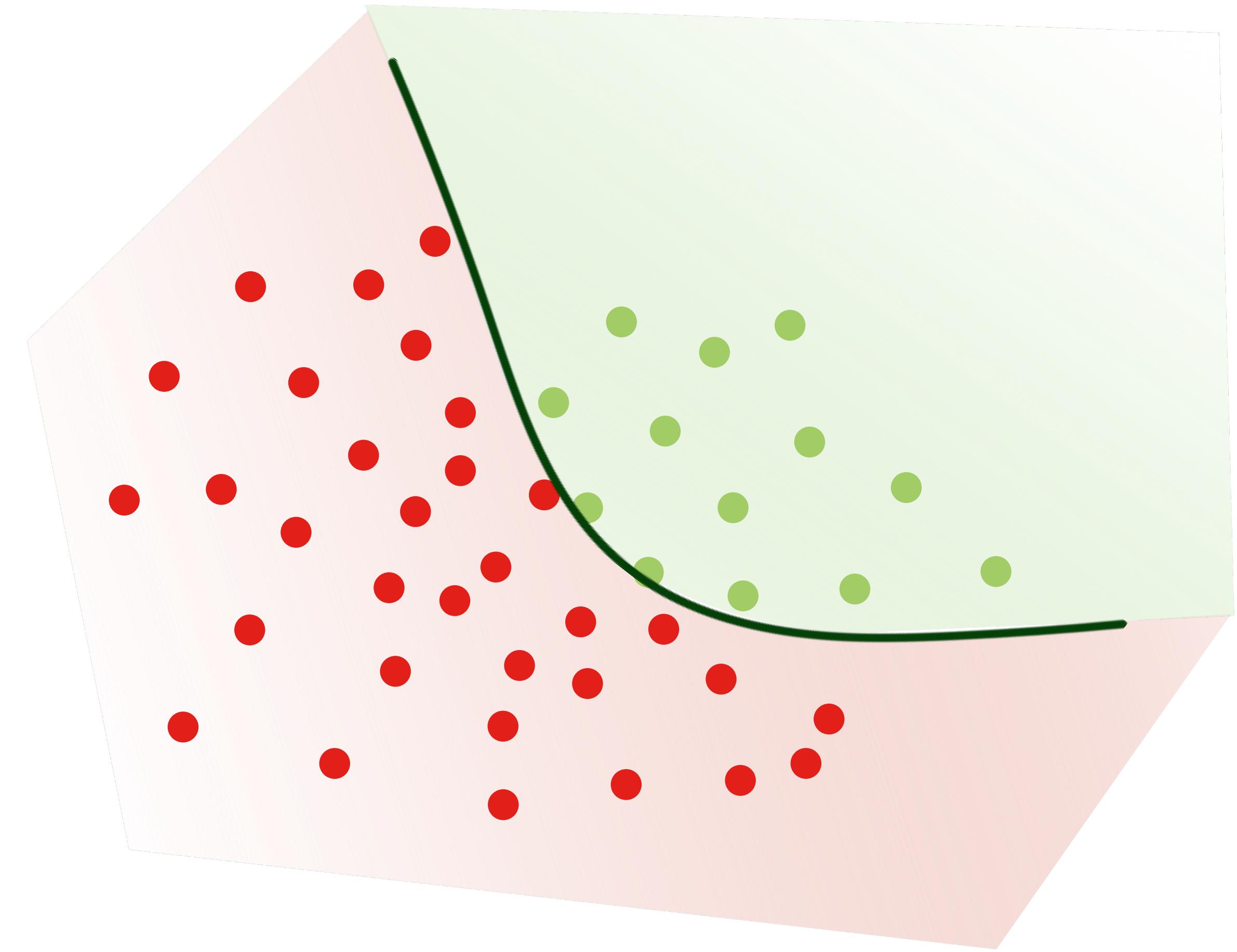
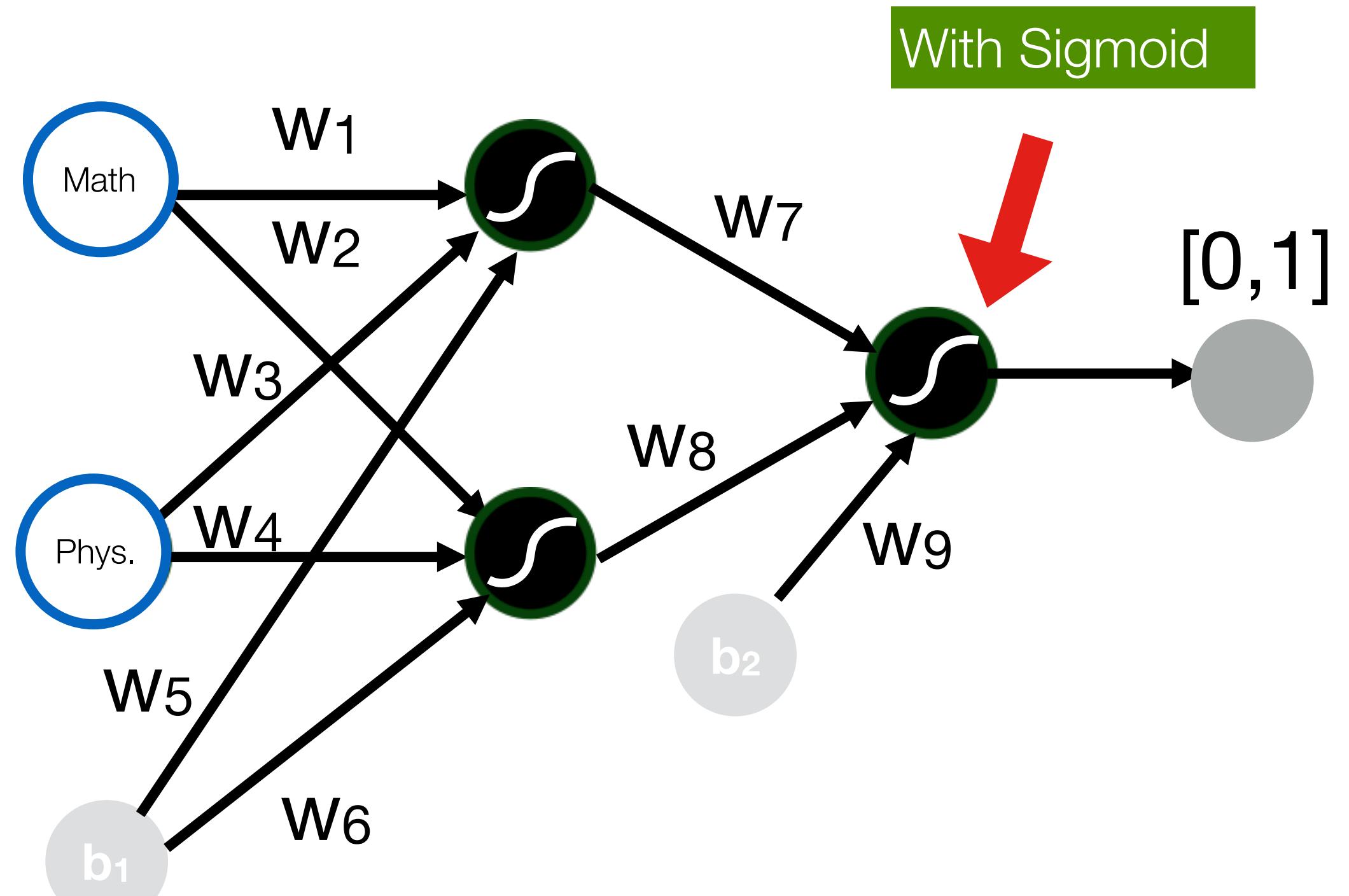
Neural Network



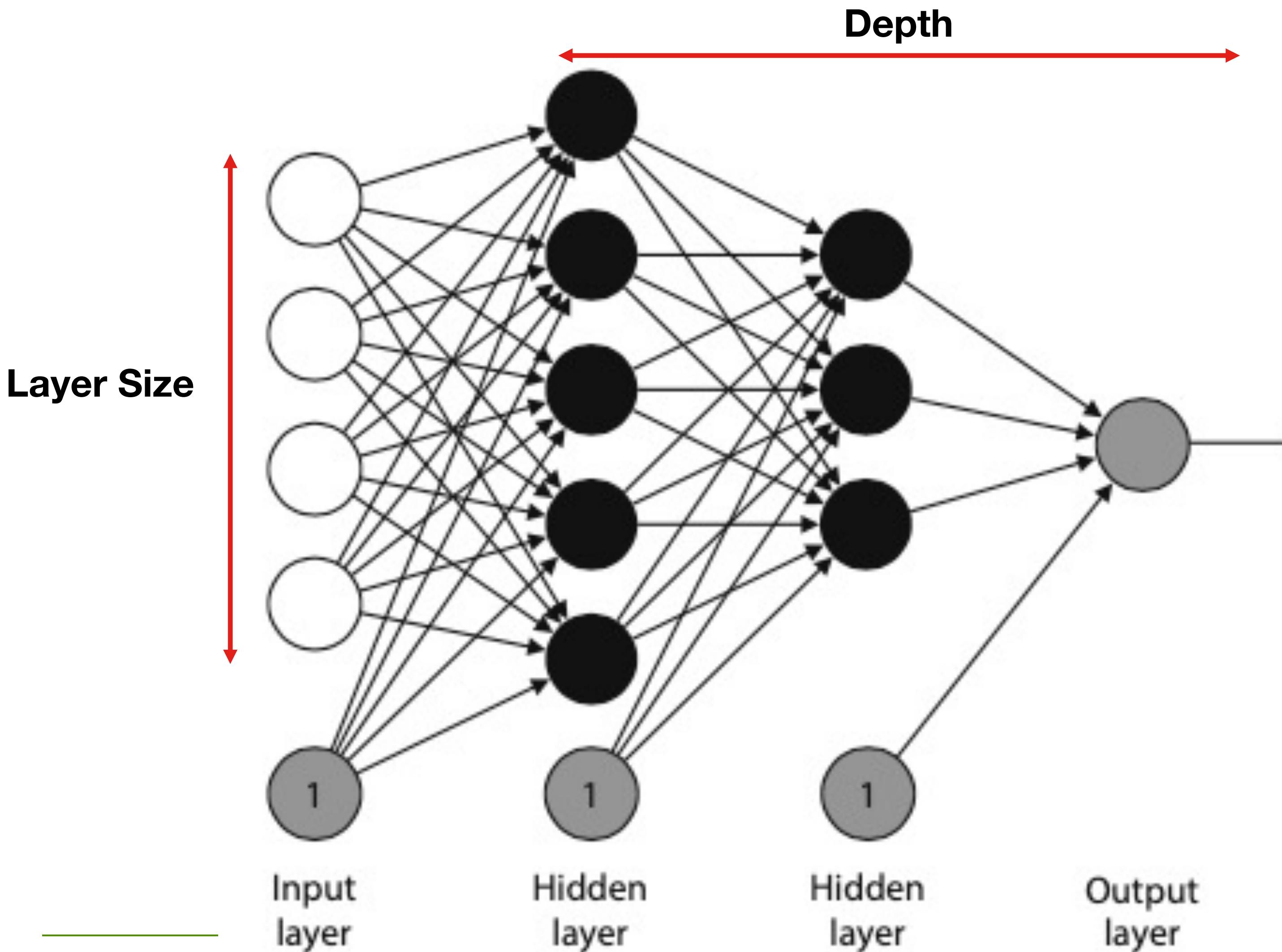
Neural Network



Neural Network



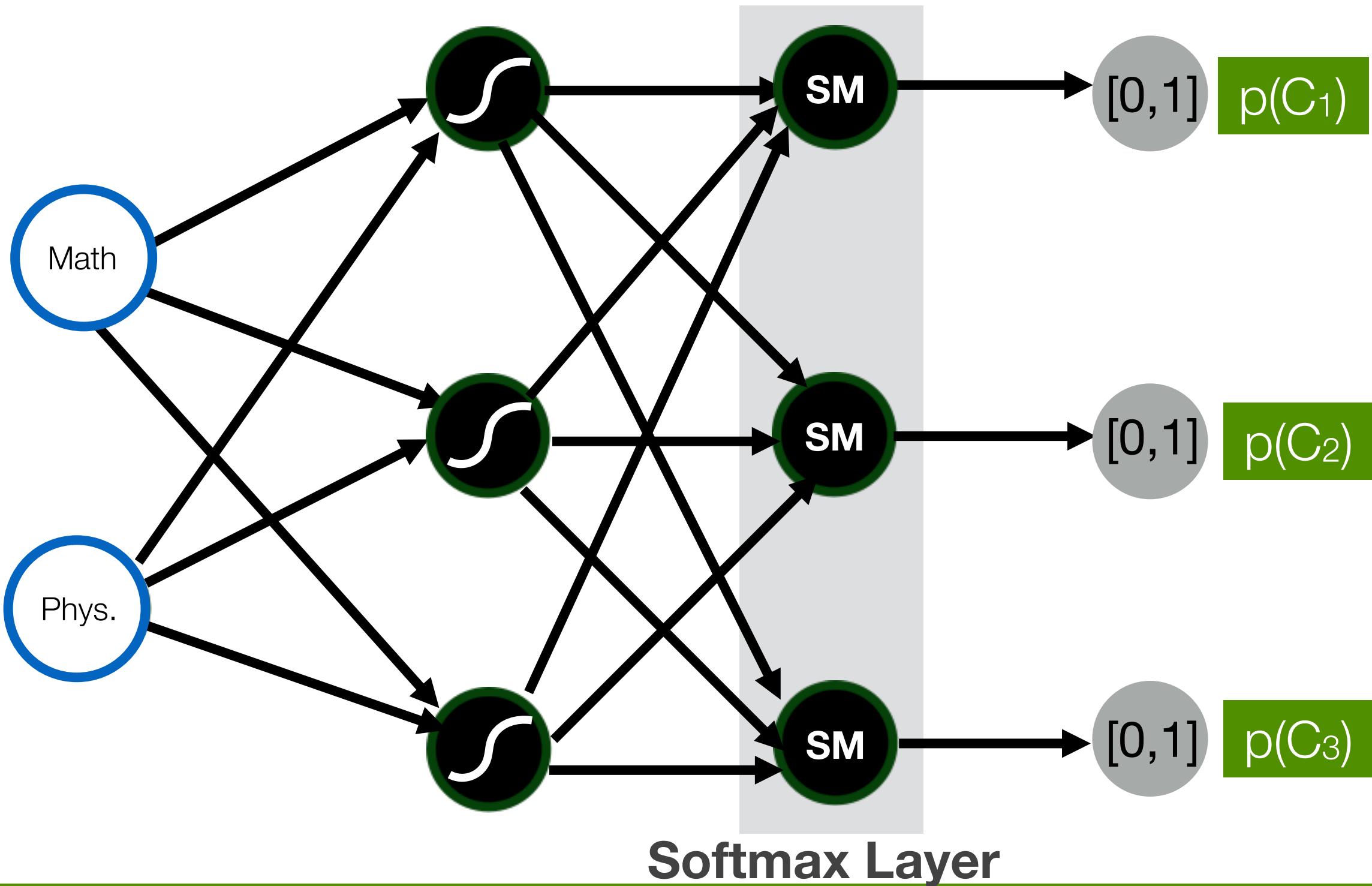
Fully Connected Neural Network



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

Classifying into multiple classes - softmax function

- Return a probability for **each class**
 - Imagine example $C_1 = \text{ADMITTED}$, $C_2 = \text{NOT ADMITTED}$, $C_3 = \text{TAKE TEST}$
 - $p(C_1) = 0.37$, $p(C_2) = 0.21$, $p(C_3) = 0.42$
- We use the Softmax activation function for the output layer

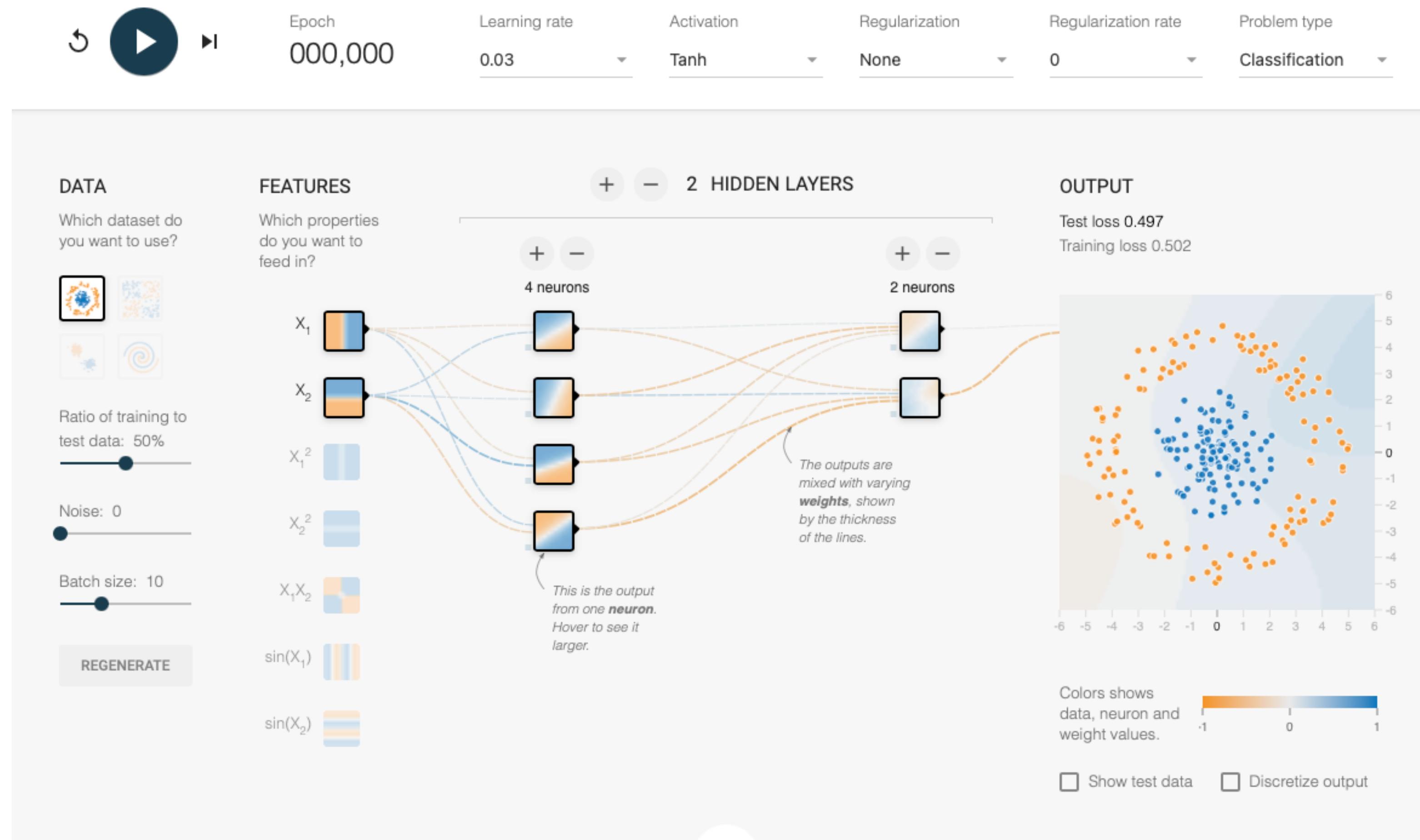


$$\text{Softmax}(x_i) = \frac{e^{(x_i)}}{\sum_j^K e^{(x_j)}}$$

Value of class i

Normalisation term on K classes

Tinker With a **Neural Network** Right Here in Your Browser. Don't Worry, You Can't Break It. We Promise.



<https://playground.tensorflow.org/>

Machine Learning and Images

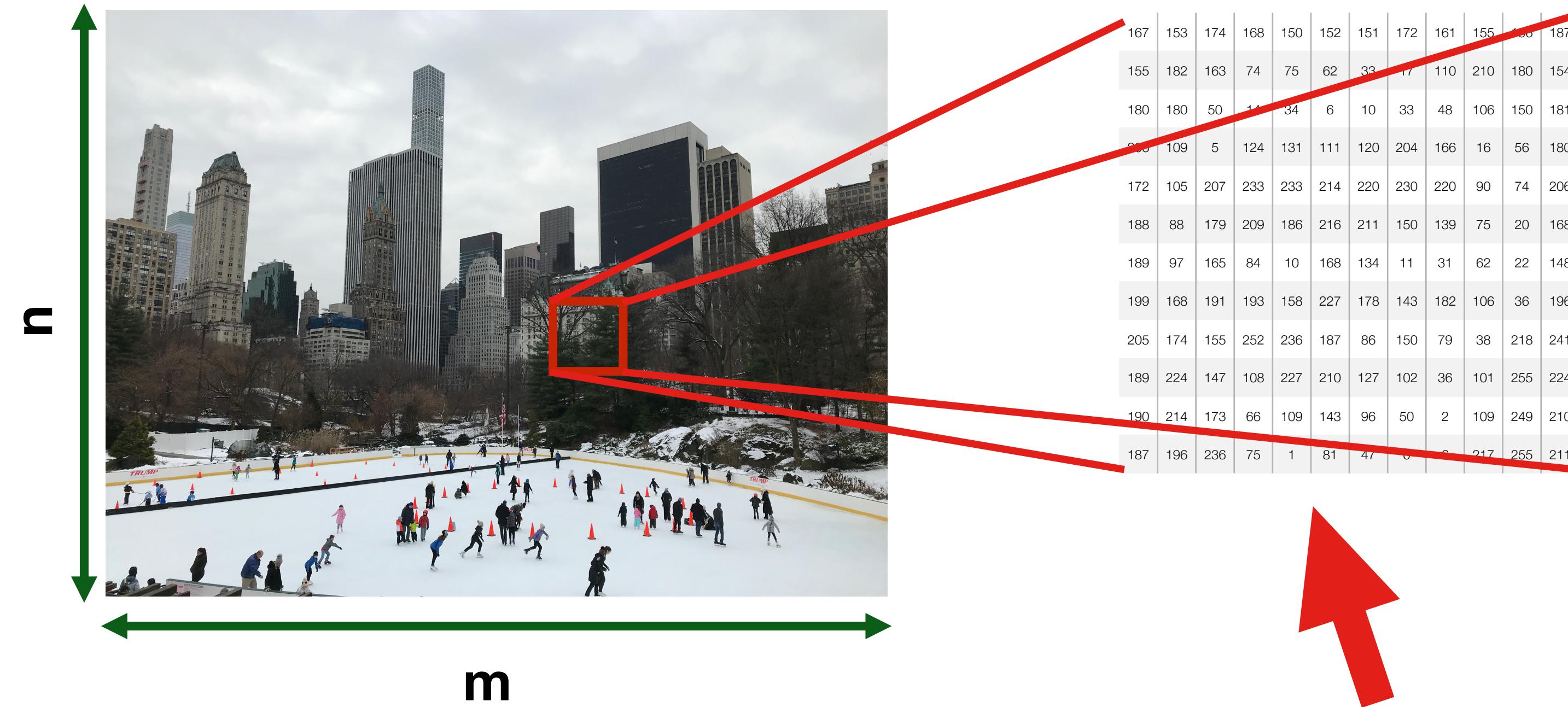
What do you see?



This is what a computer “sees”

167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187						
155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	187				
180	180	50	14	34	6	10	33	180	180	50	14	34	6	10	33	48	106	150	180	180	50	14	34	6	10	33	48	106	150	180	181	180	181	180	181	180	181	180	181	180	181
206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	180	180	180			
172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	206	206	206	206		
188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	168	168	168	168		
189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	148	148	148	148		
199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	196	196	196	196		
205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	241	241	241			
189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	224	224	224			
190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	210	210	210			
187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211	211	211	211			
167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	187	187	187			
155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	187	187	187	187	
180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181	181	180	181			
206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	180	180	180			
172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	206	206	206	206		
188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	168	168	168	168		
189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	148	148	148	148		
199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	196	196	196	196		
205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	241	241	241			
189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	224	224	224			
190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	210	210	210			
187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81</																								

Images



- Each pixel in an image is a feature
- Dimensionality
 - $n \times m$
- Feature (pixel) values are numerical
 - 0 or 1 for Black and White
 - Between 0 and 255 for greyscale
 - 16M values for RGB

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?

Strong vs. Weak Artificial Intelligence

- **Strong AI**
 - Artificial General Intelligence (AGI), human-level, general
 - The AI we see in movies
 - AI that can do everything we humans can do, and possibly much more
- **Weak AI**
 - Narrow AI
 - AI specialised in well-defined tasks
 - e.g. speech recognition, chess-playing, autonomous driving
- No AI program has been created yet that could be called intelligent in any general (Strong AI) sense
 - "A pile of narrow intelligence will never add up to a general intelligence. General intelligence isn't about the number of abilities, but about the integration between those abilities?"
- Superintelligence doesn't really mean anything - a basic calculator far exceeds any human benchmark for performing basic arithmetic

“Easy problems are hard”

Marvin Minsky

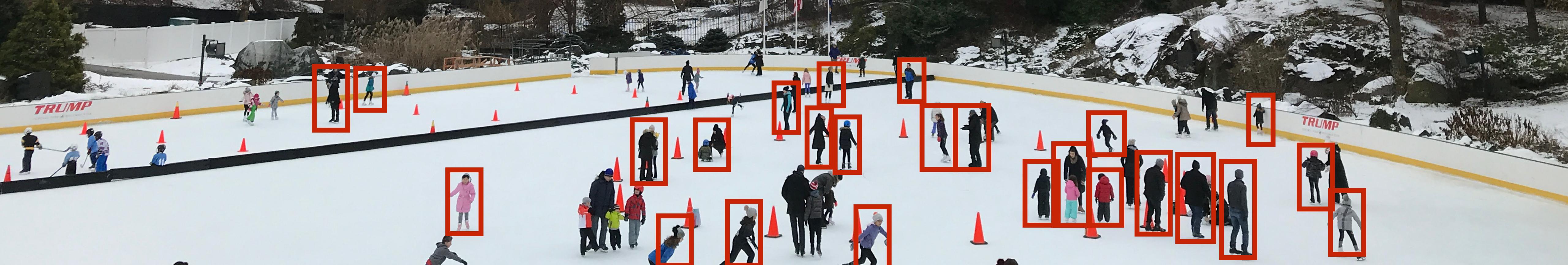


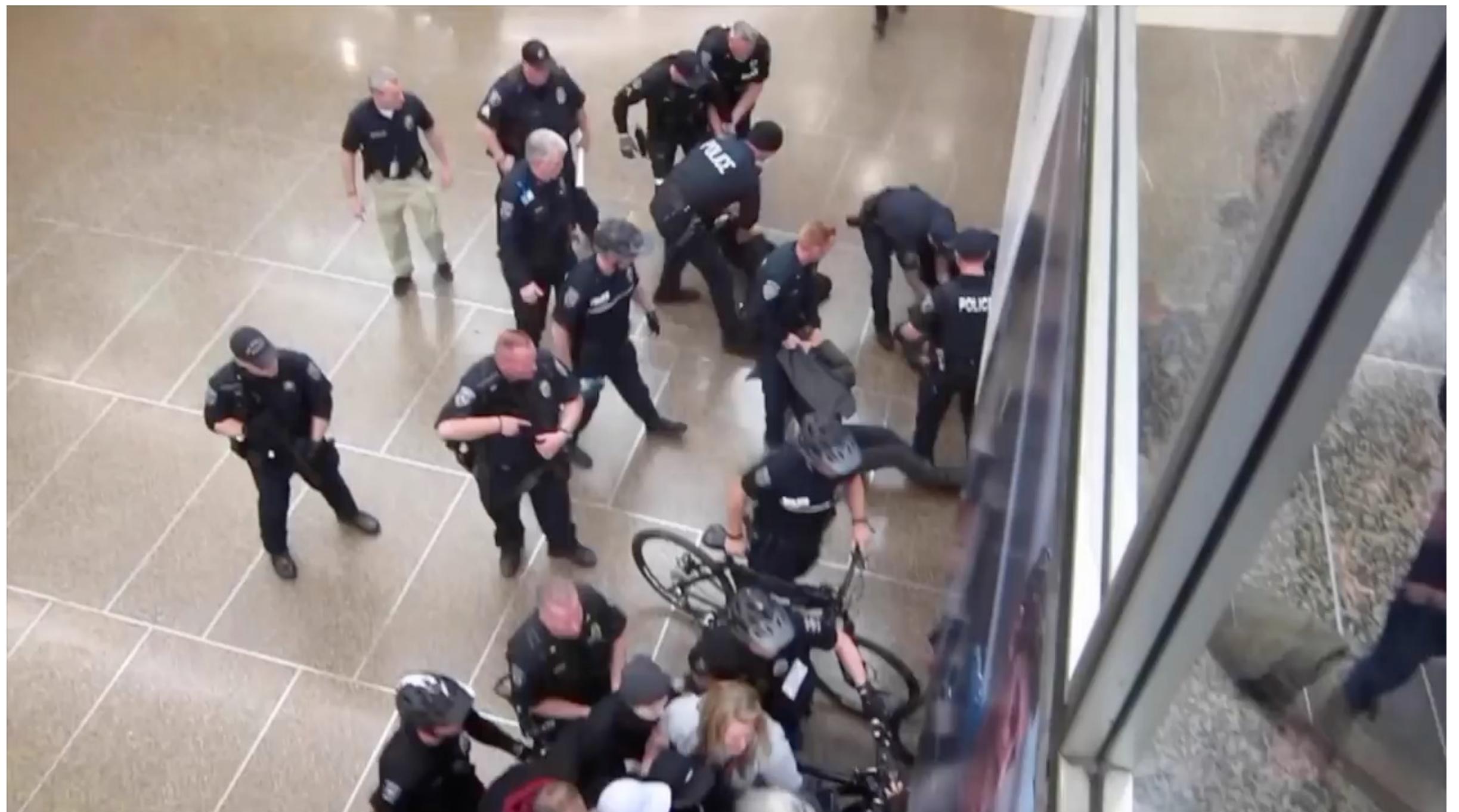
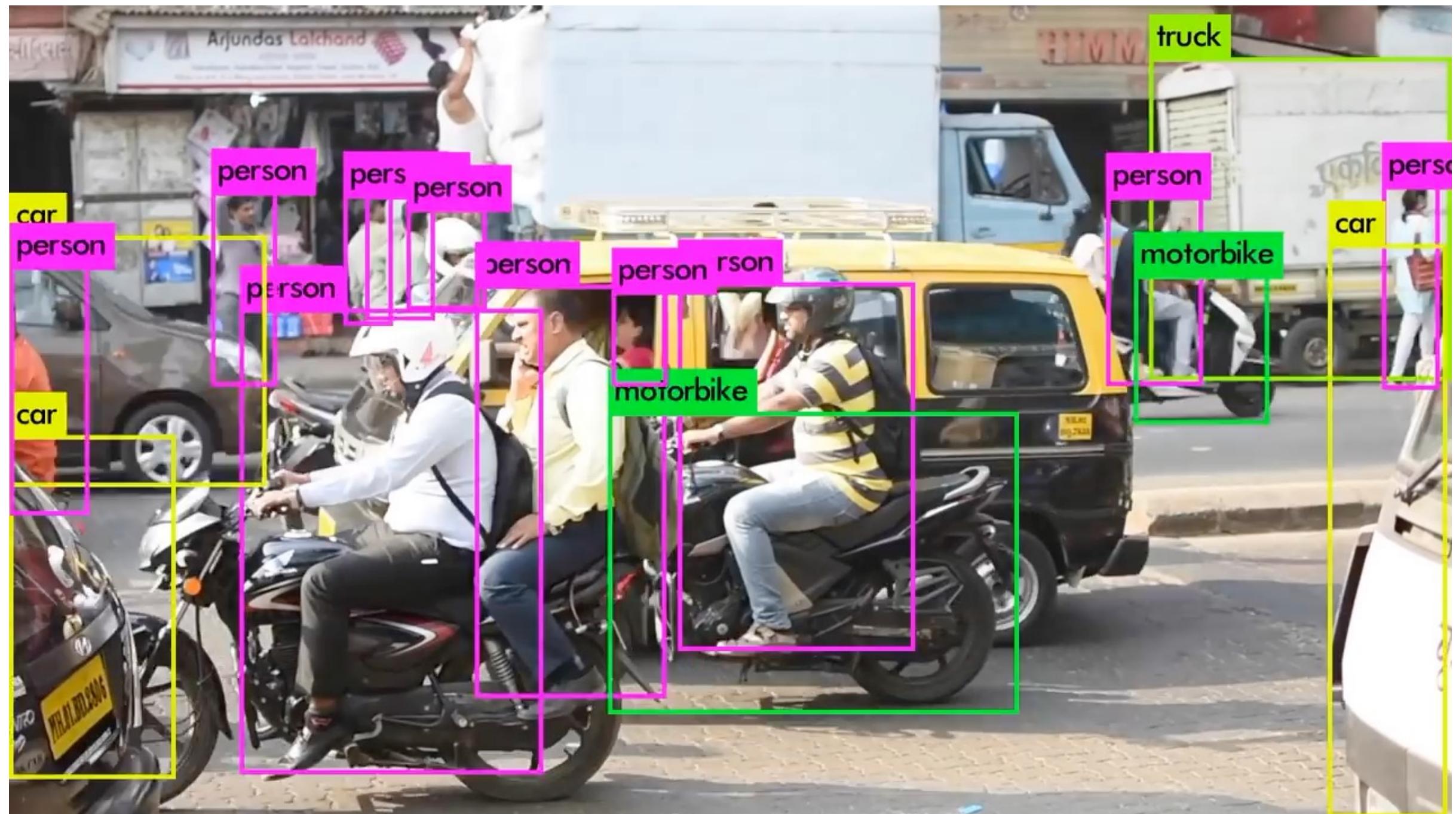
Is this a flag?
(Recognition / Classification)





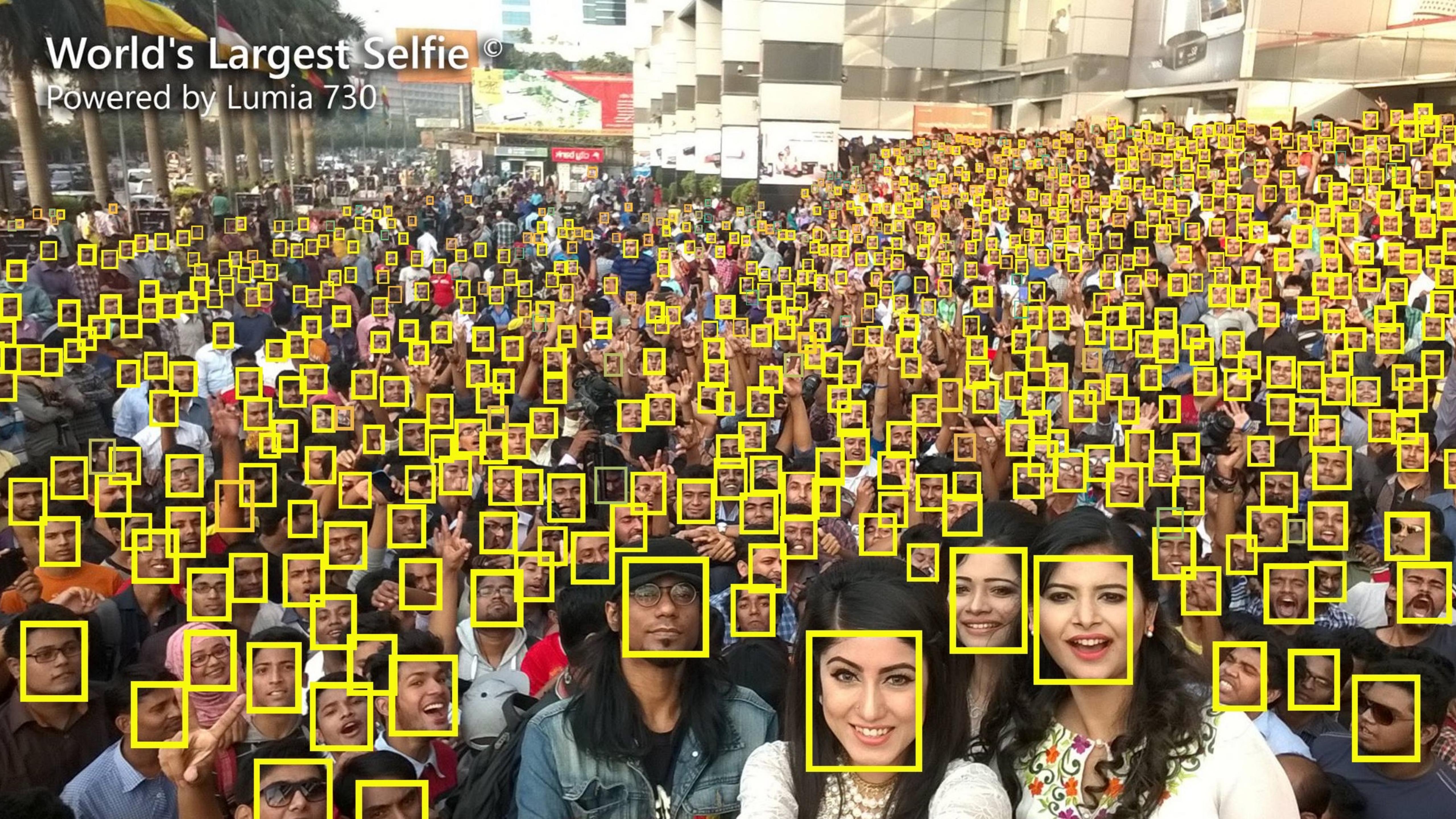
Where are the people?
(Recognition/Classification)





World's Largest Selfie ©

Powered by Lumia 730



MORPHCAST



AFFECT

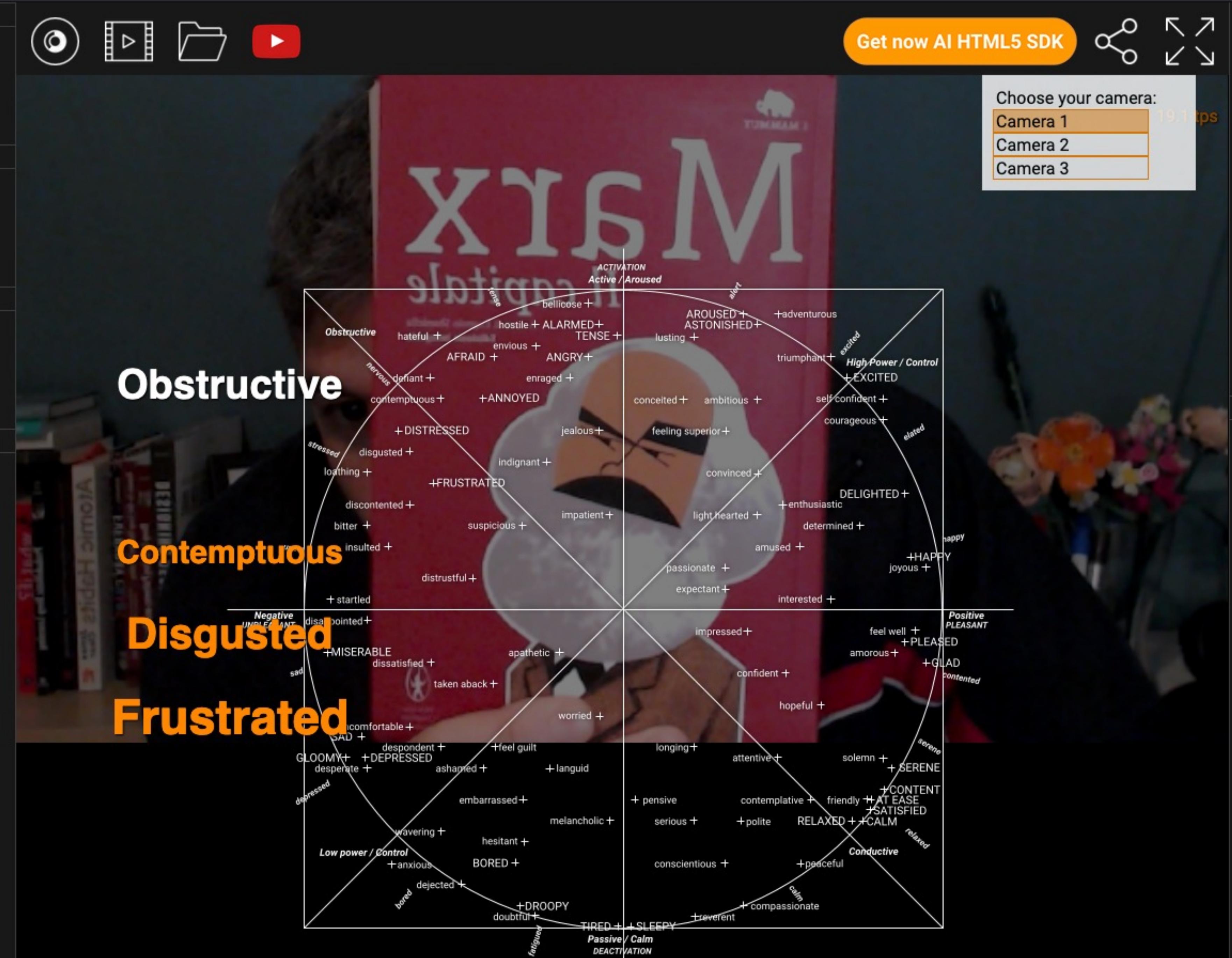


EMOTION



LIKELY AGE

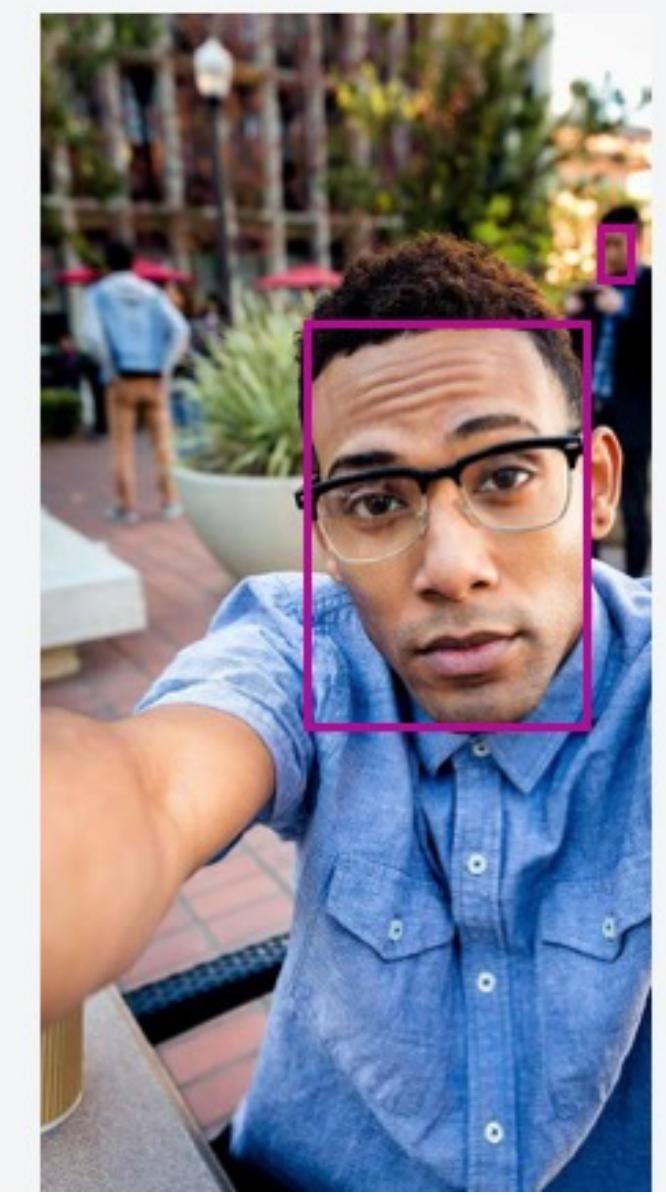
-100
- 90
- 80
- 70
- 60
- 50
- 40
- 30
- 20
- 10
- 0



Fully protects
your privacy
No recording
No storing of
biometric data
All images
overwritten
in 100 MS



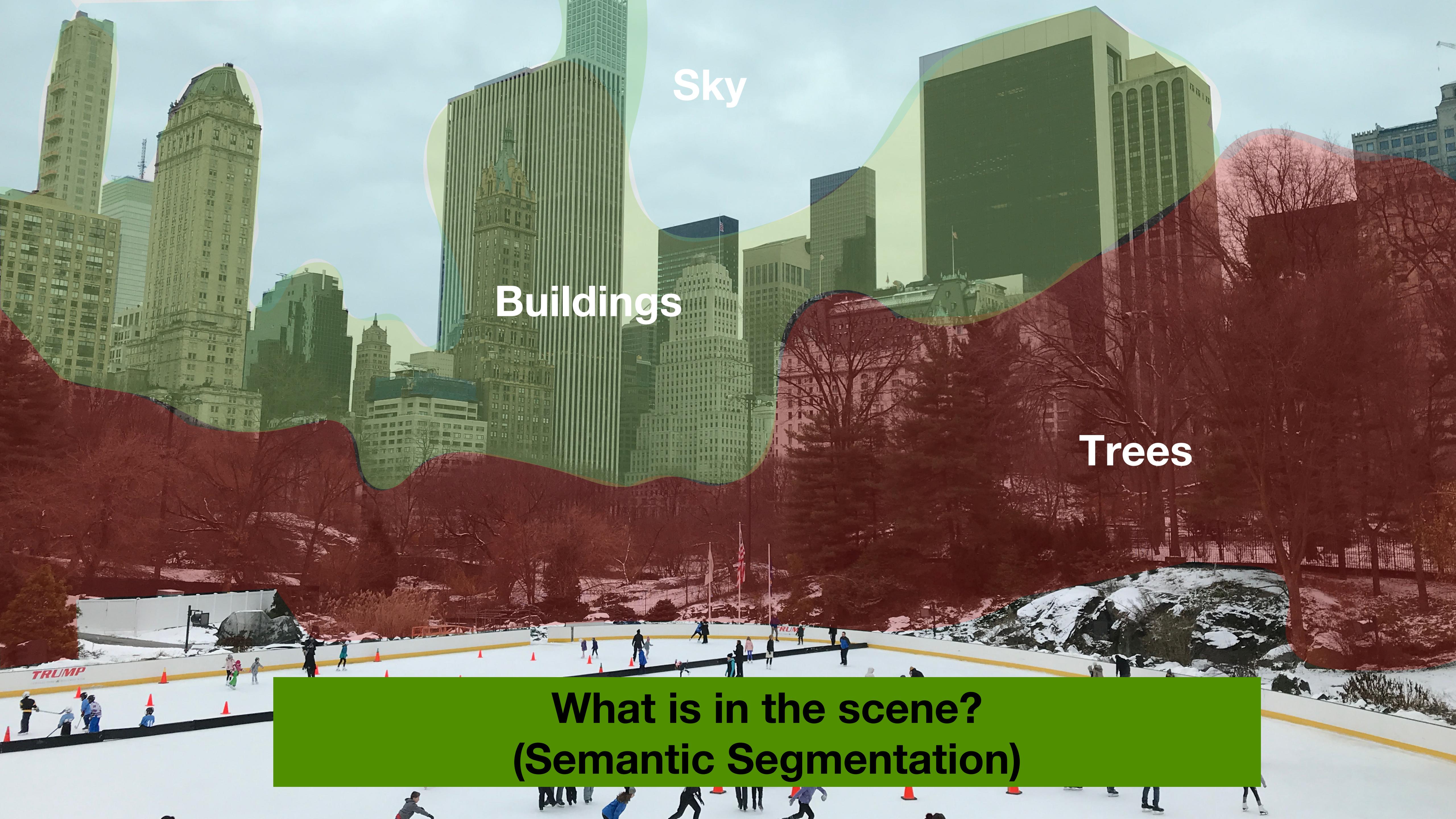
**Is this Jeff?
(Identification)**



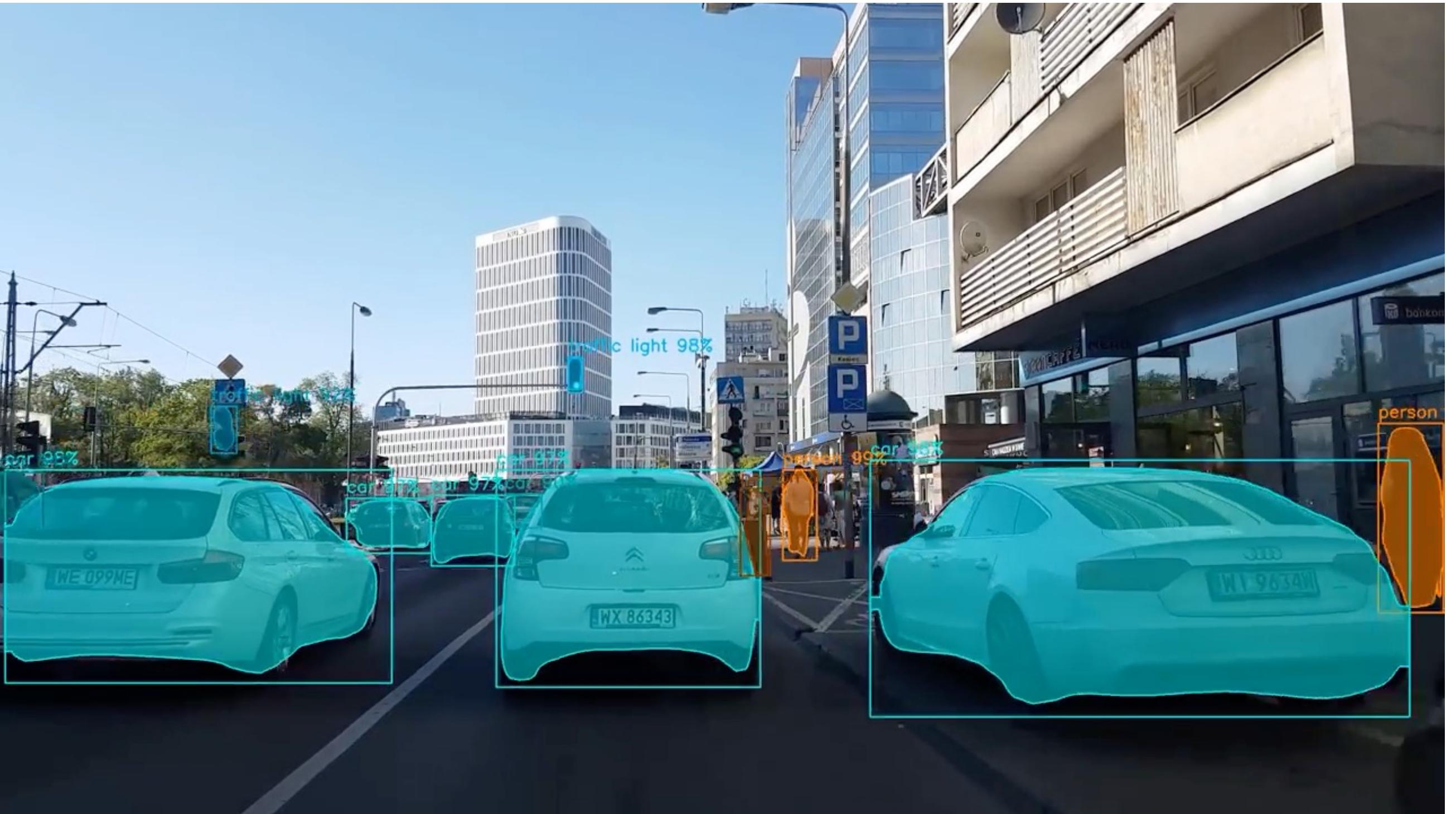


**Is this the Wollman Rink?
(Identification)**





**What is in the scene?
(Semantic Segmentation)**



https://github.com/matterport/Mask_RCNN

Project Sunroof

≡ Google Project Sunroof

Savings estimator

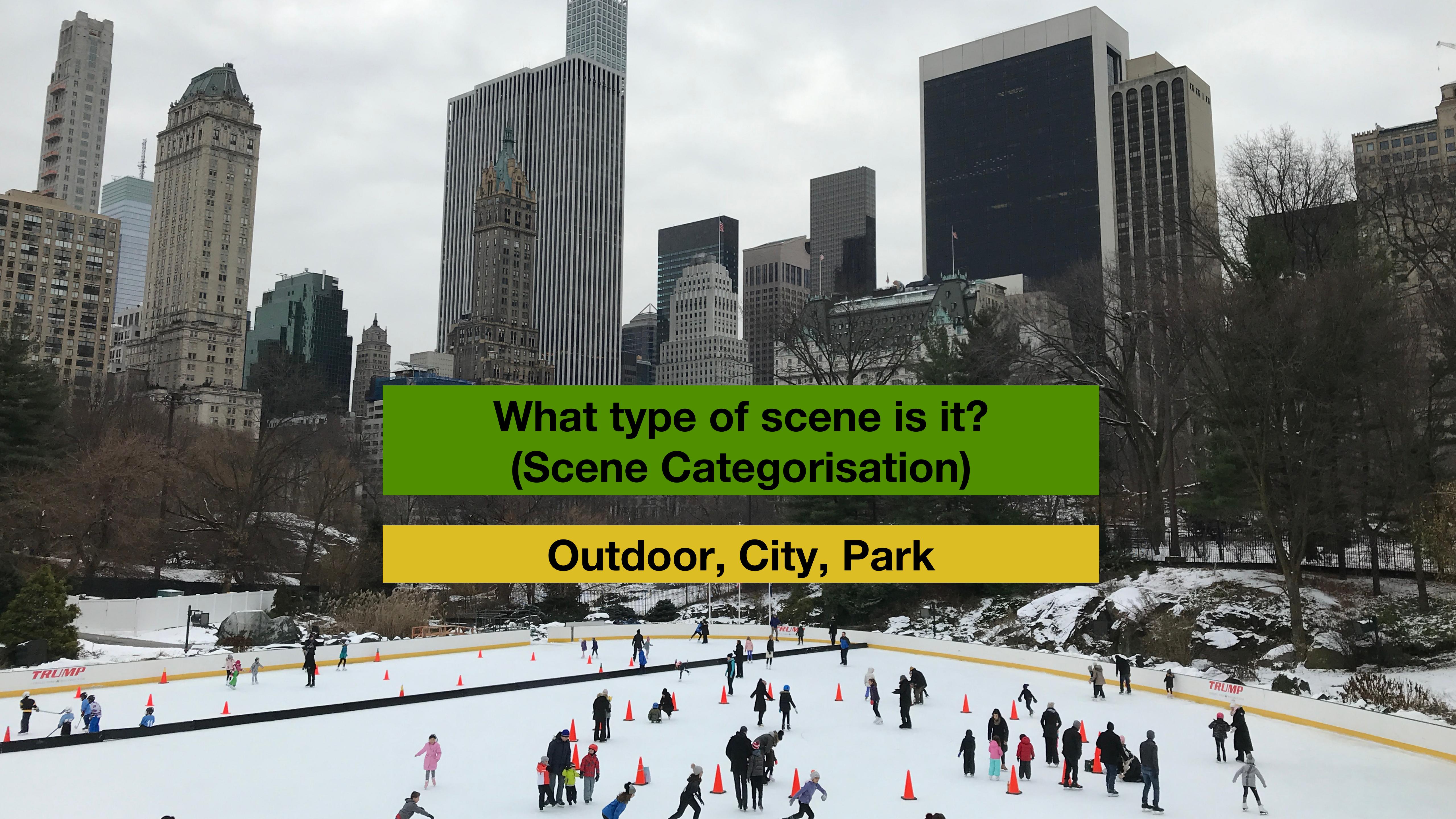
Data explorer

Solar 101

FAQ

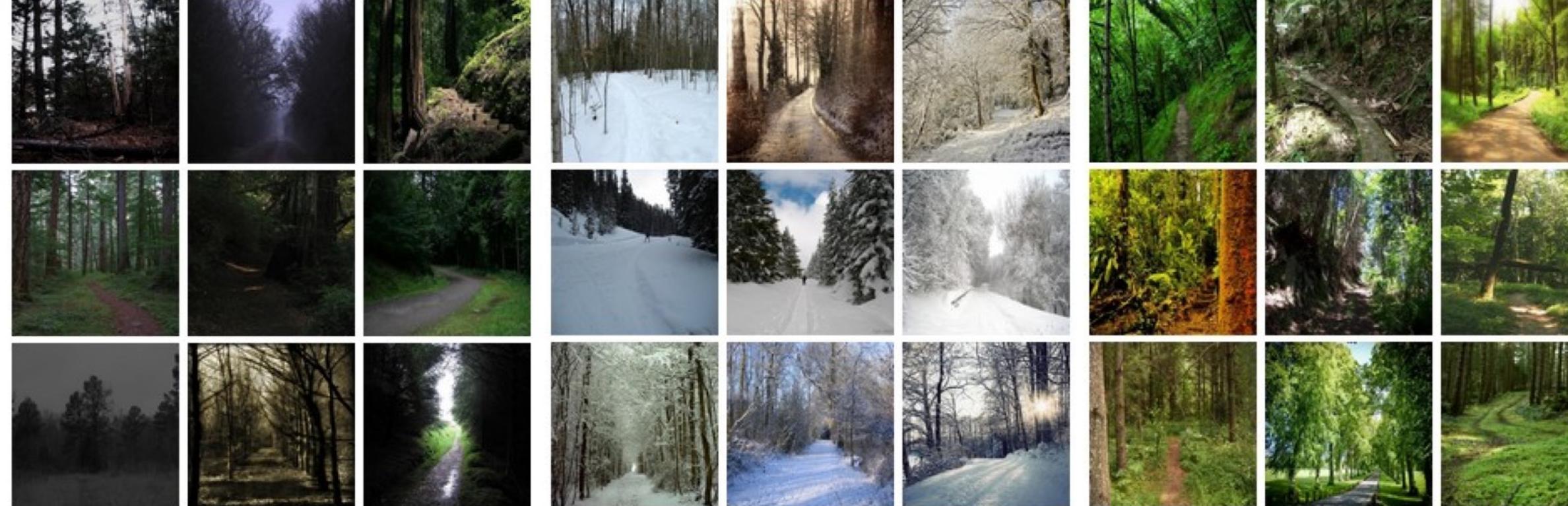


<https://www.google.com/get/sunroof>

The background image shows a wide outdoor ice skating rink in a park. In the foreground, many people are skating on the ice. The rink is surrounded by a white fence. In the background, there is a dense cluster of tall, modern skyscrapers under a clear blue sky.

**What type of scene is it?
(Scene Categorisation)**

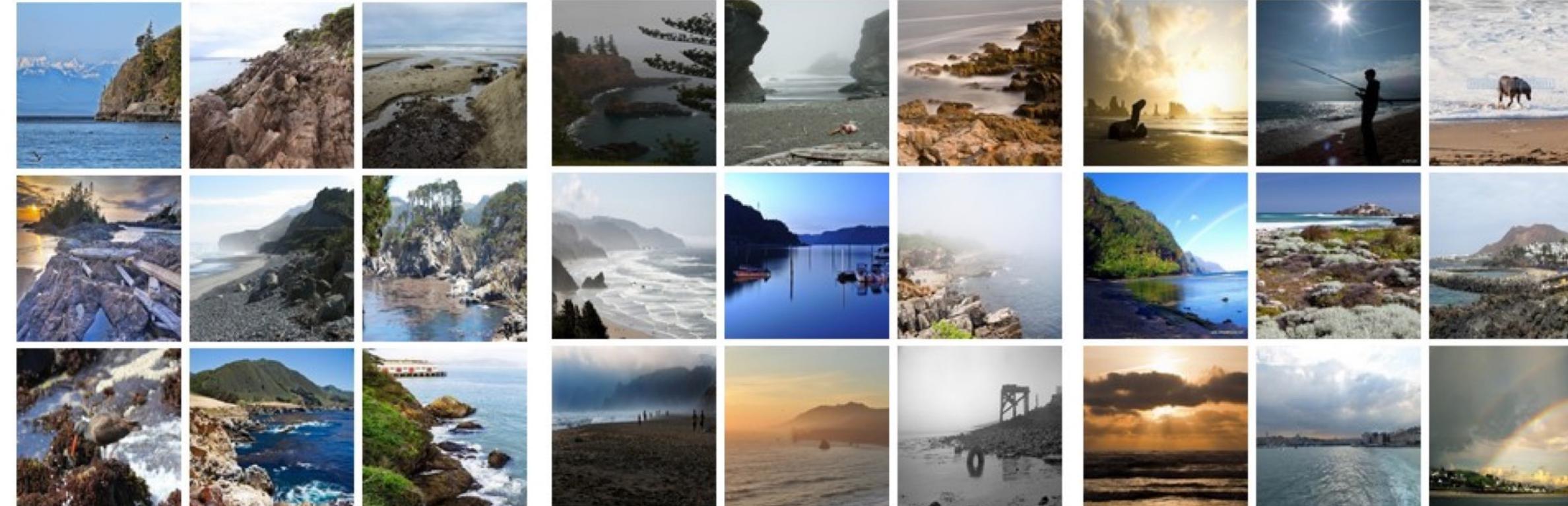
Outdoor, City, Park



darkest forest path

wintering forest path

greener forest path



rocky coast

misty coast

sunny coast



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:**



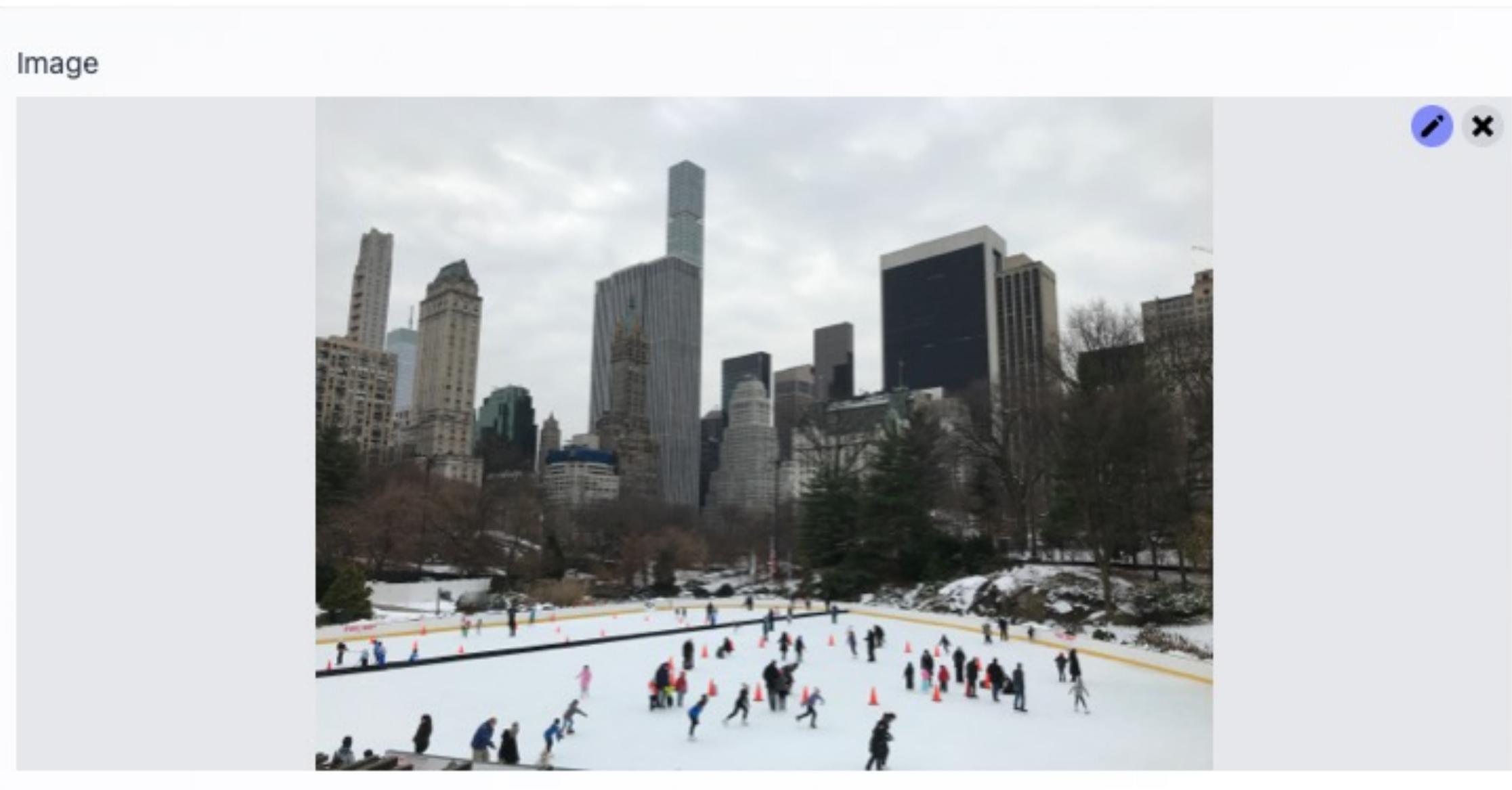


Spaces: OFA-Sys/OFA-Image_Caption □ like 12 • Running

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.

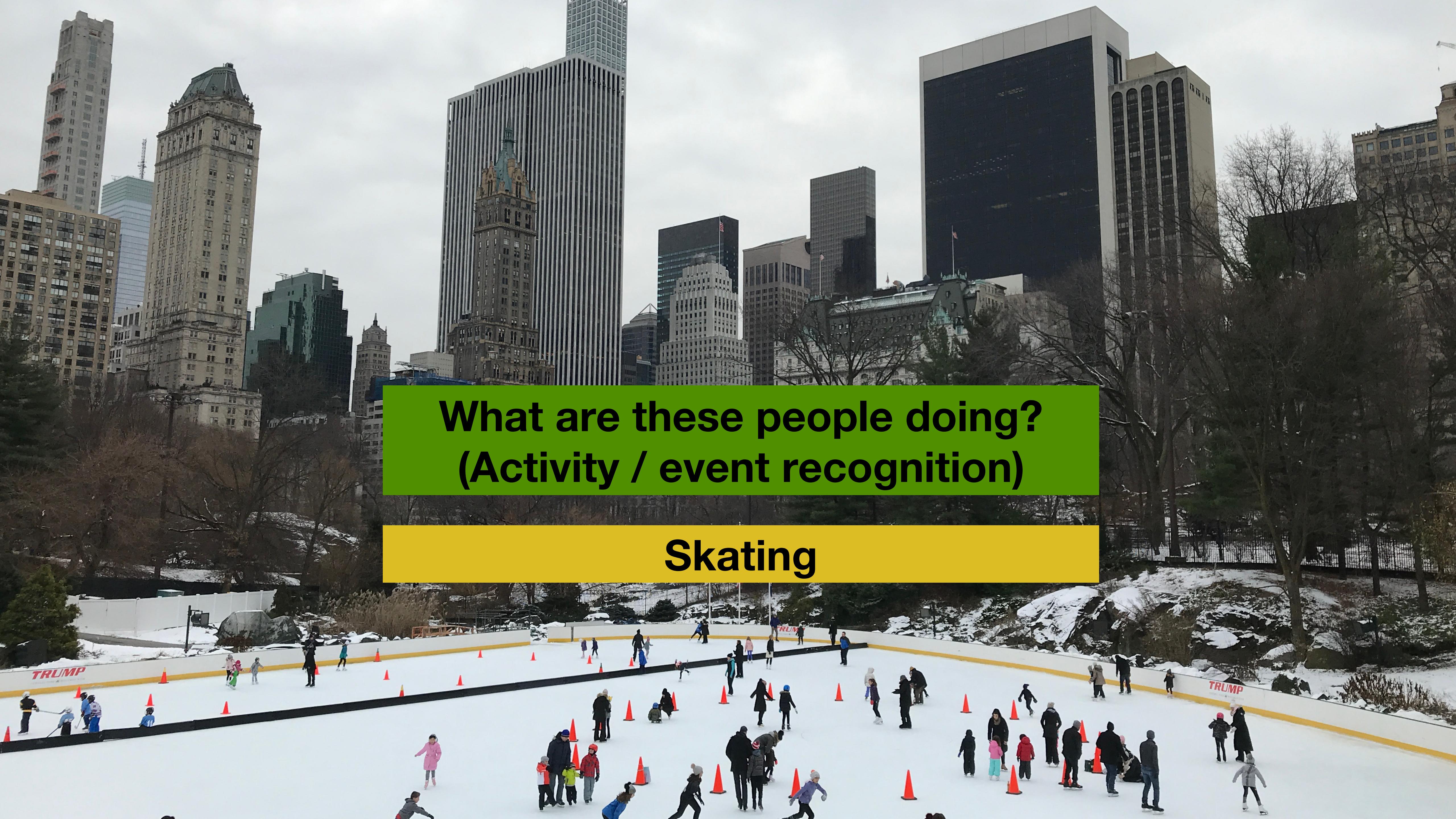


Caption 8.28s

people skating on a rink in a city park with skyscrapers

Clear

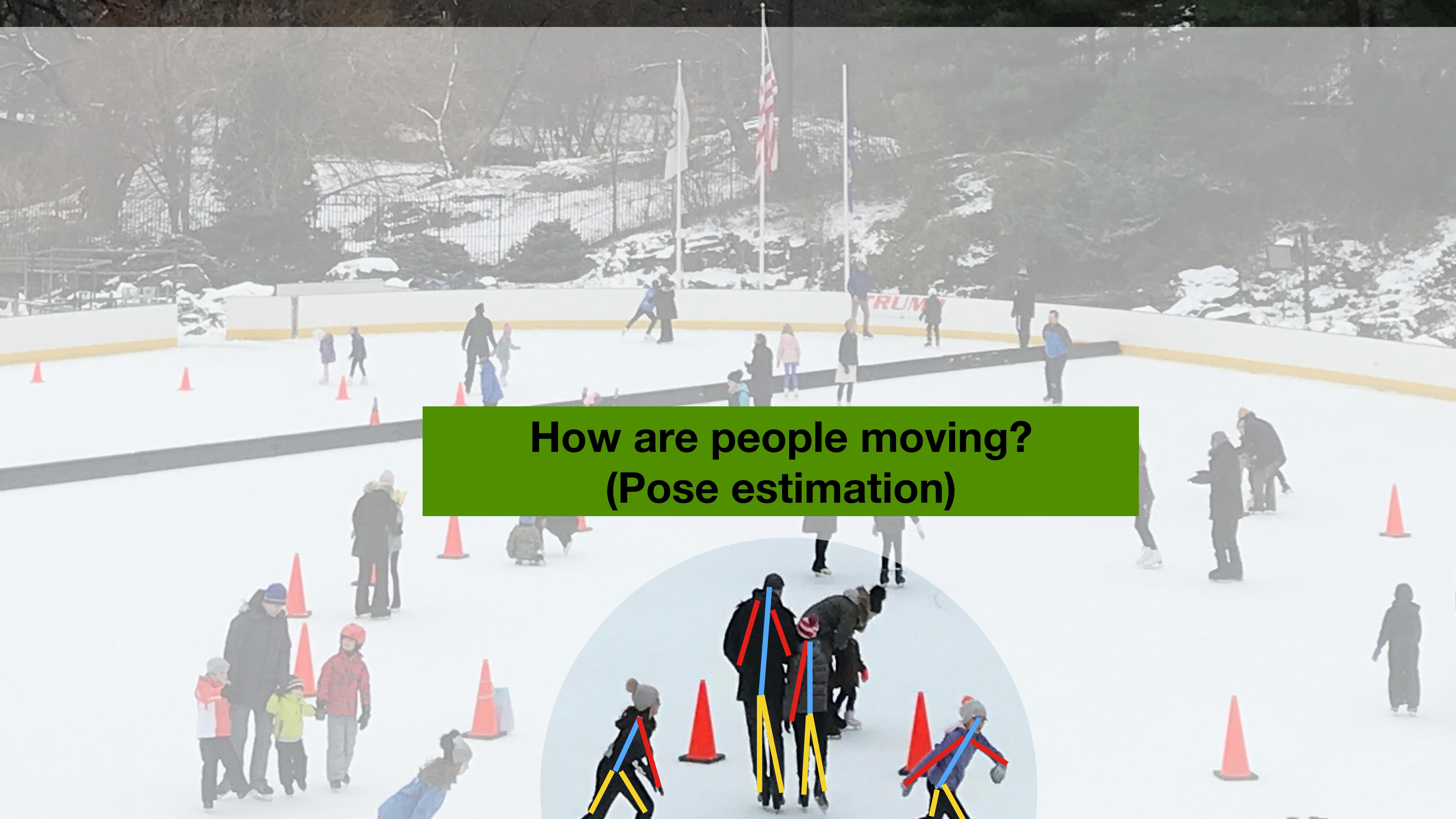
Submit



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

Skating





**How are people moving?
(Pose estimation)**



X: -1.11702489838885962072754
Z: -5.3908080944445801
Work station

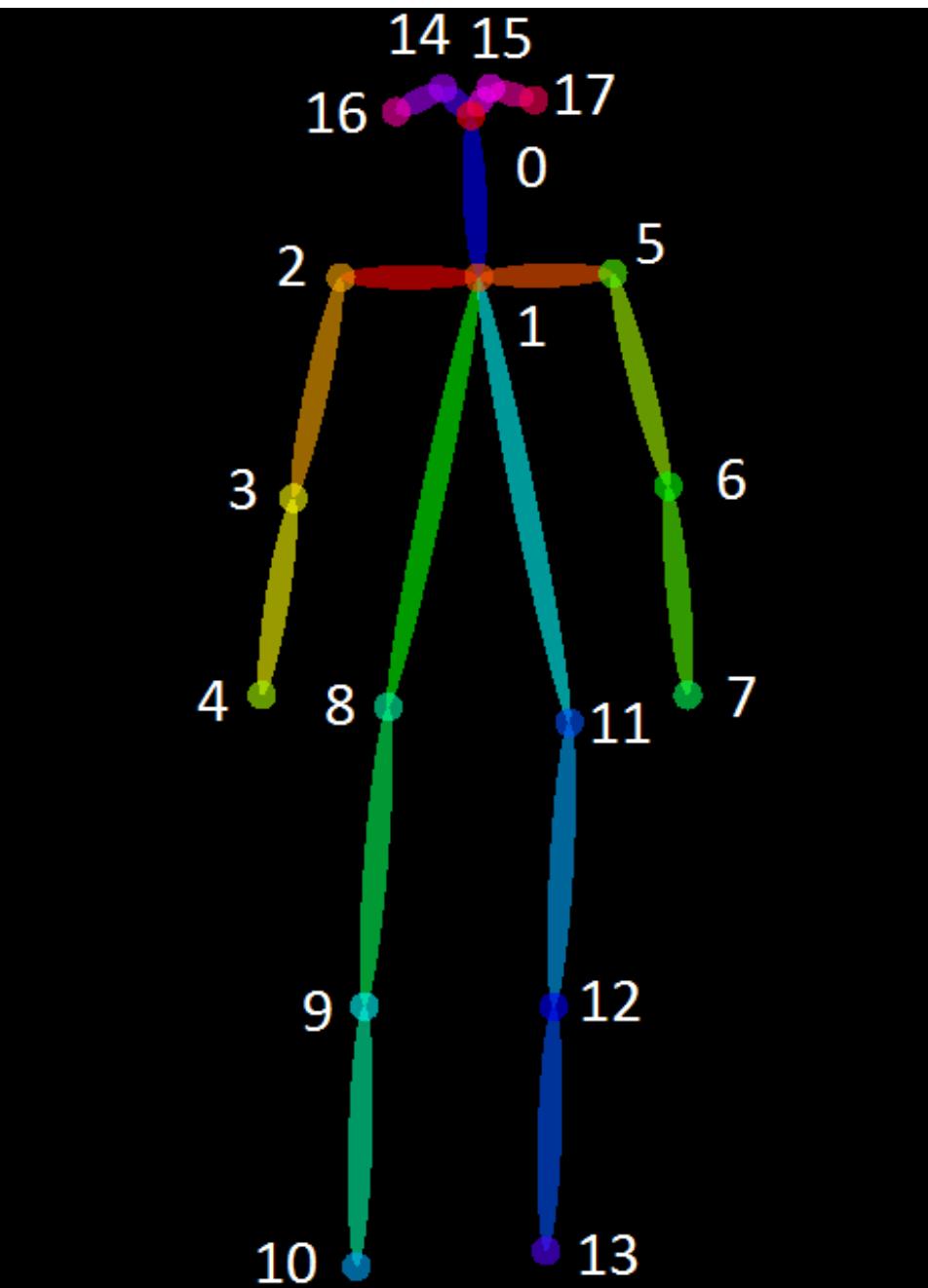


Ethical task tracking of operators in agile manufacturing

<http://resolver.tudelft.nl/uuid:3408e8c3-809b-436d-94eb-efb4f0532b17>

Stereolabs ZED Camera used in COALA

- 3D Object Detection
- Body tracking
- Positional tracking



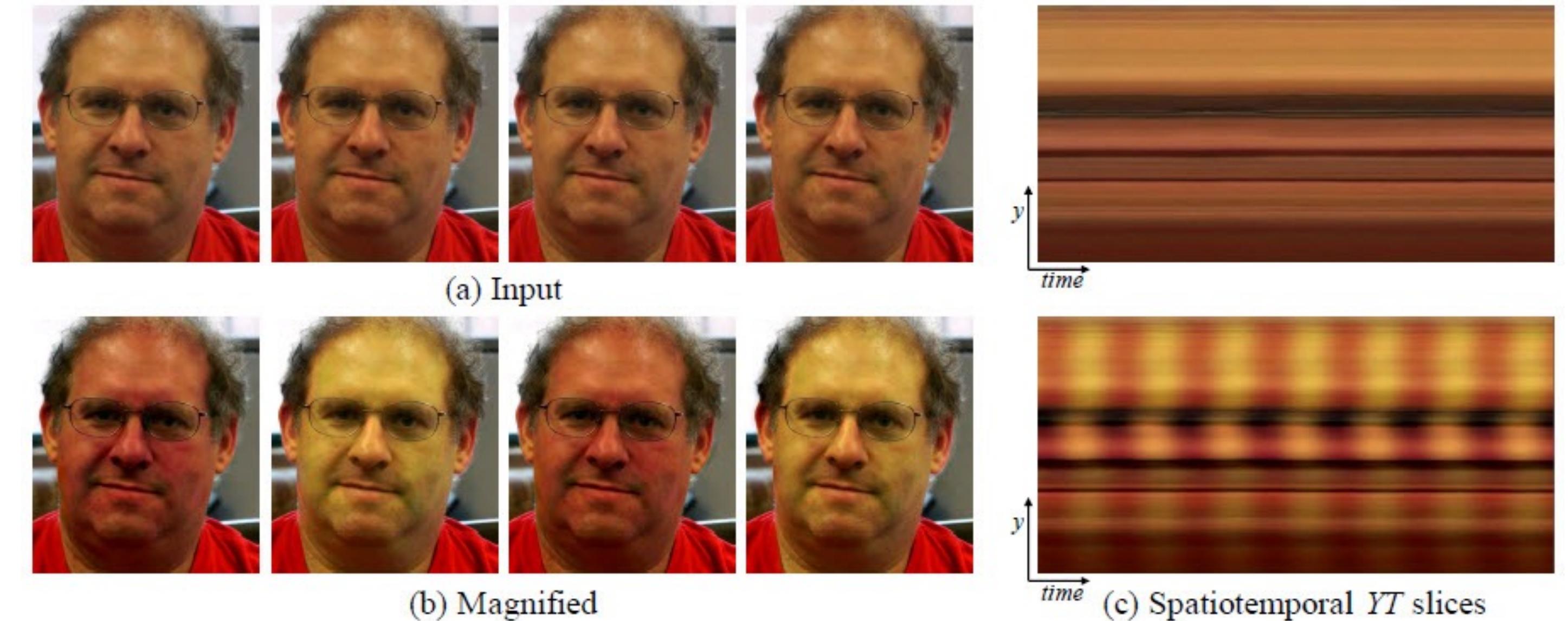
[https://www.stereolabs.com/
docs/object-detection/](https://www.stereolabs.com/docs/object-detection/)

[https://www.stereolabs.com/d
ocs/body-tracking/](https://www.stereolabs.com/docs/body-tracking/)



Eulerean Video Magnification

- Measure heart rate by magnifying and measuring the subtle changes in skin colour owed to blood vessels
- <https://people.csail.mit.edu/mrub/vidmag/>



Wu, H. Y., Rubinstein, M., Shih, E., Guttag, J., Durand, F., & Freeman, W. (2012). Eulerian video magnification for revealing subtle changes in the world. ACM transactions on graphics (TOG), 31(4), 1-8.

Google images

Search by image
Search Google with an image instead of text. Try dragging an image here.

Drop image here

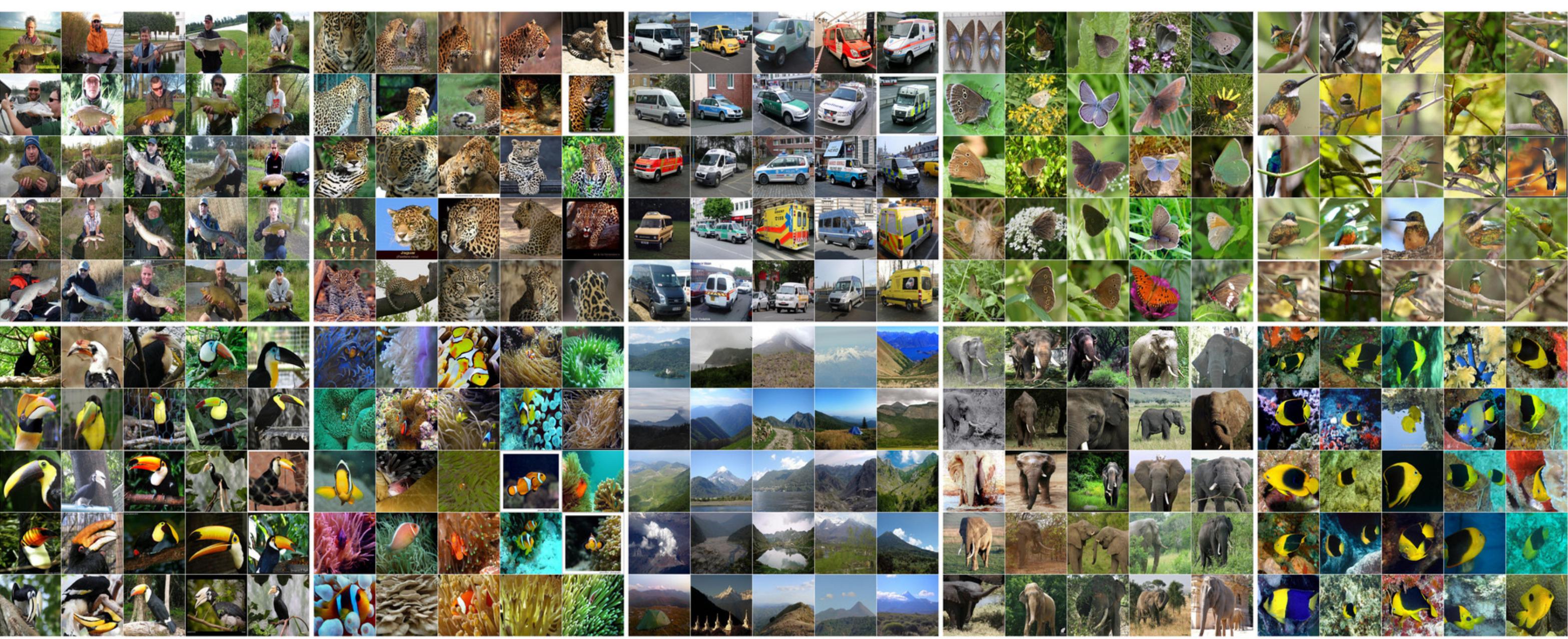


Image search results

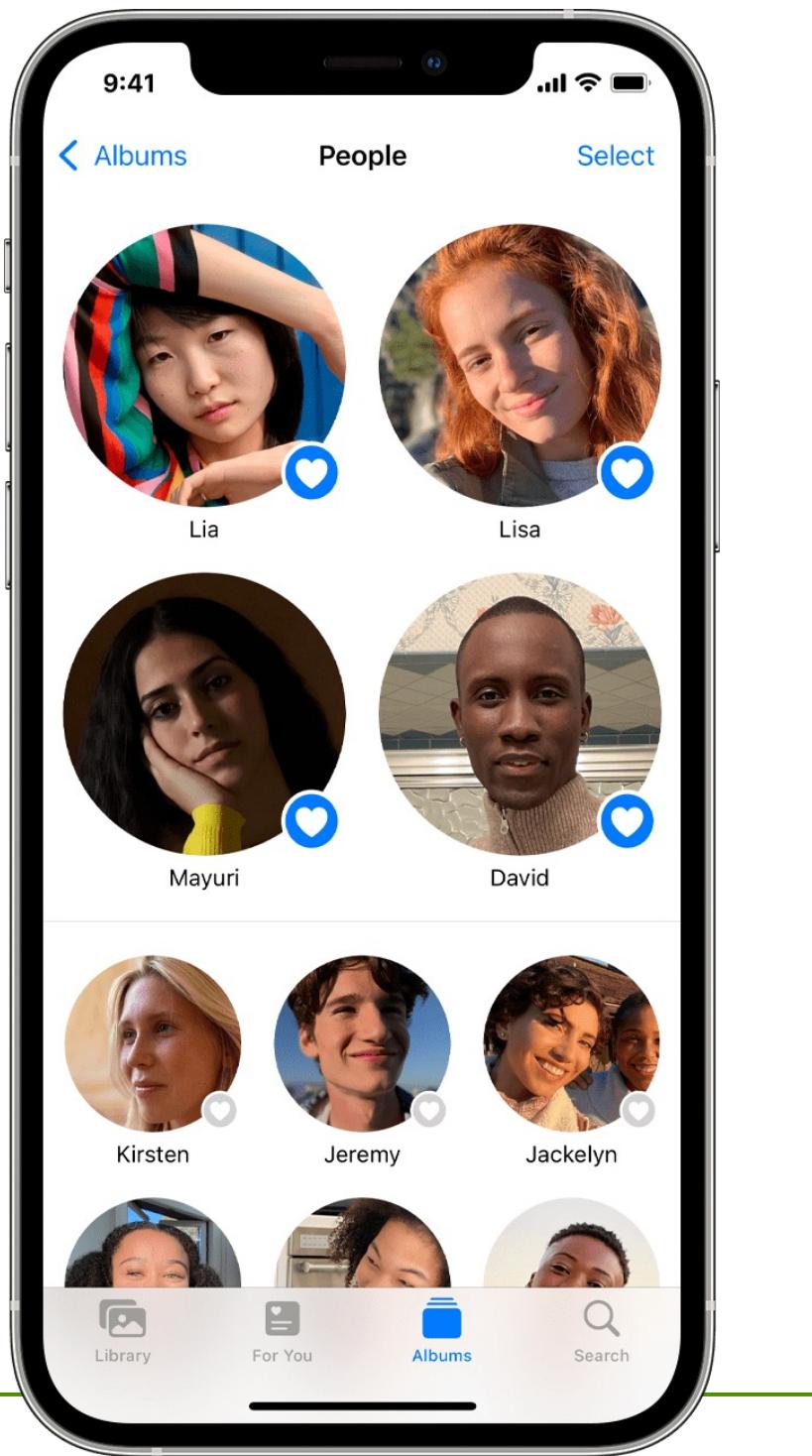
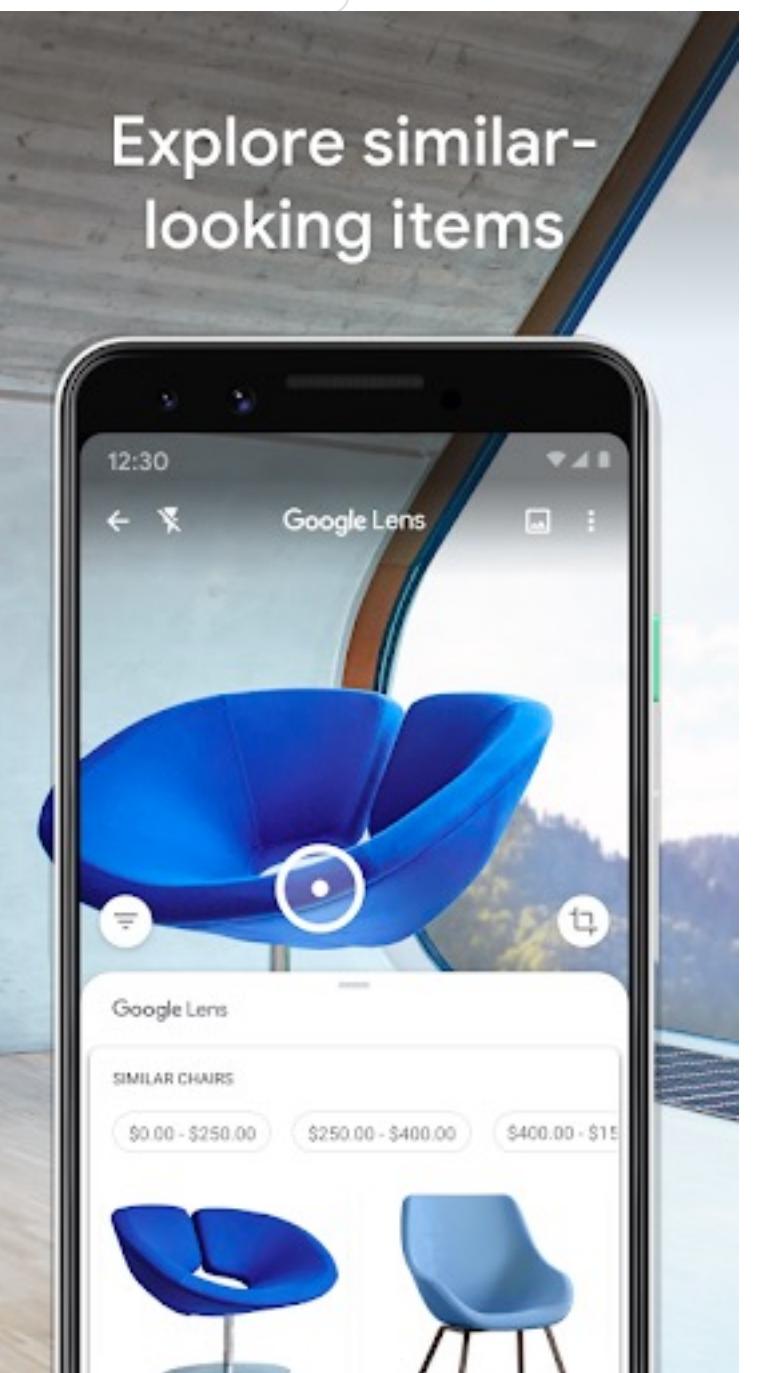


88432.jpg X Lamp

Visually similar images



Report images



Advanced Machine Learning For Design

Lecture 4 - Machine Learning
for Image Processing (part 1)

Module 2

Evangelos Niforatos
12/10/2022

aml4d-ide@tudelft.nl
<https://aml4design.github.io/>

Sources

- CMU Computer Vision course - Matthew O'Toole. <http://16385.courses.cs.cmu.edu/spring2022/>
- Grokking Machine Learning. Luis G. Serrano. Manning, 2021
- [What is back propagation really doing?](#)
 - YouTube walk-through
- “[Gradient descent, how neural networks learn](#)”
 - YouTube walk-through
- “[Neural networks and back-propagation explained in a simple way](#)”
 - Article on medium.com
- “[The MNIST DATABASE](#)” of handwritten digits
- [“Neural Networks from Scratch”](#)
 - YouTube tutorial
- [“Neural Networks from Scratch in Python”](#)
 - New book/e-book

