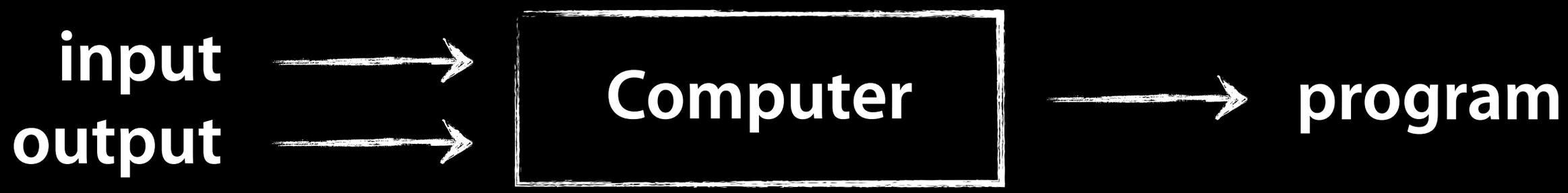
machine learning



“ Field of study that gives computers the ability to learn without being explicitly programmed. ”

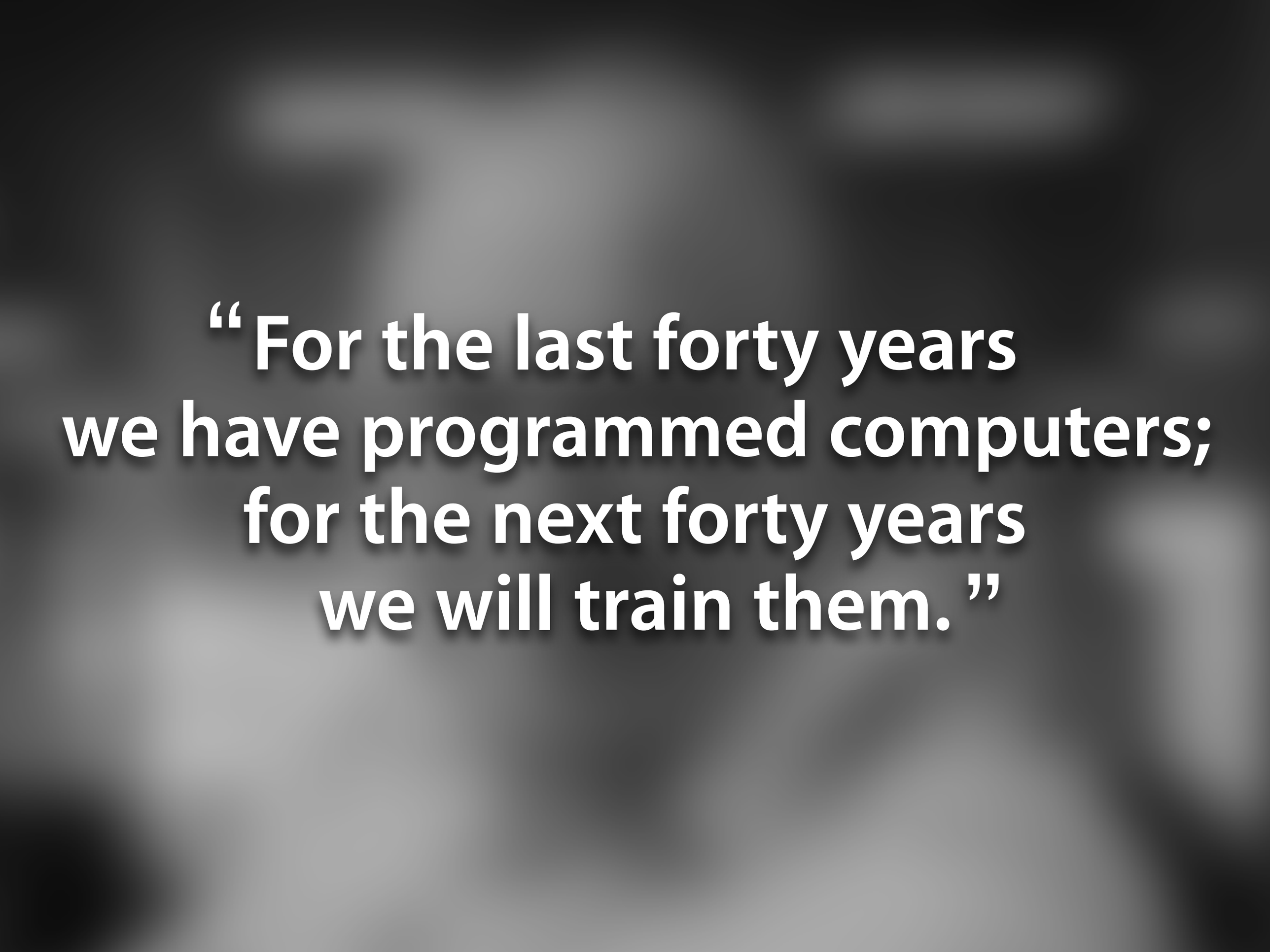
— Arthur Samuel

“ A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P if its performance at tasks in T, as measured by P, improves with experience E. ”

— Tom M. Mitchell

Christopher Bishop

Director of the Microsoft Research lab in Cambridge


“For the last forty years
we have programmed computers;
for the next forty years
we will train them.”

$$\mathbf{x} \in \mathbb{R}^N$$

$m \times n$

$m \times n$

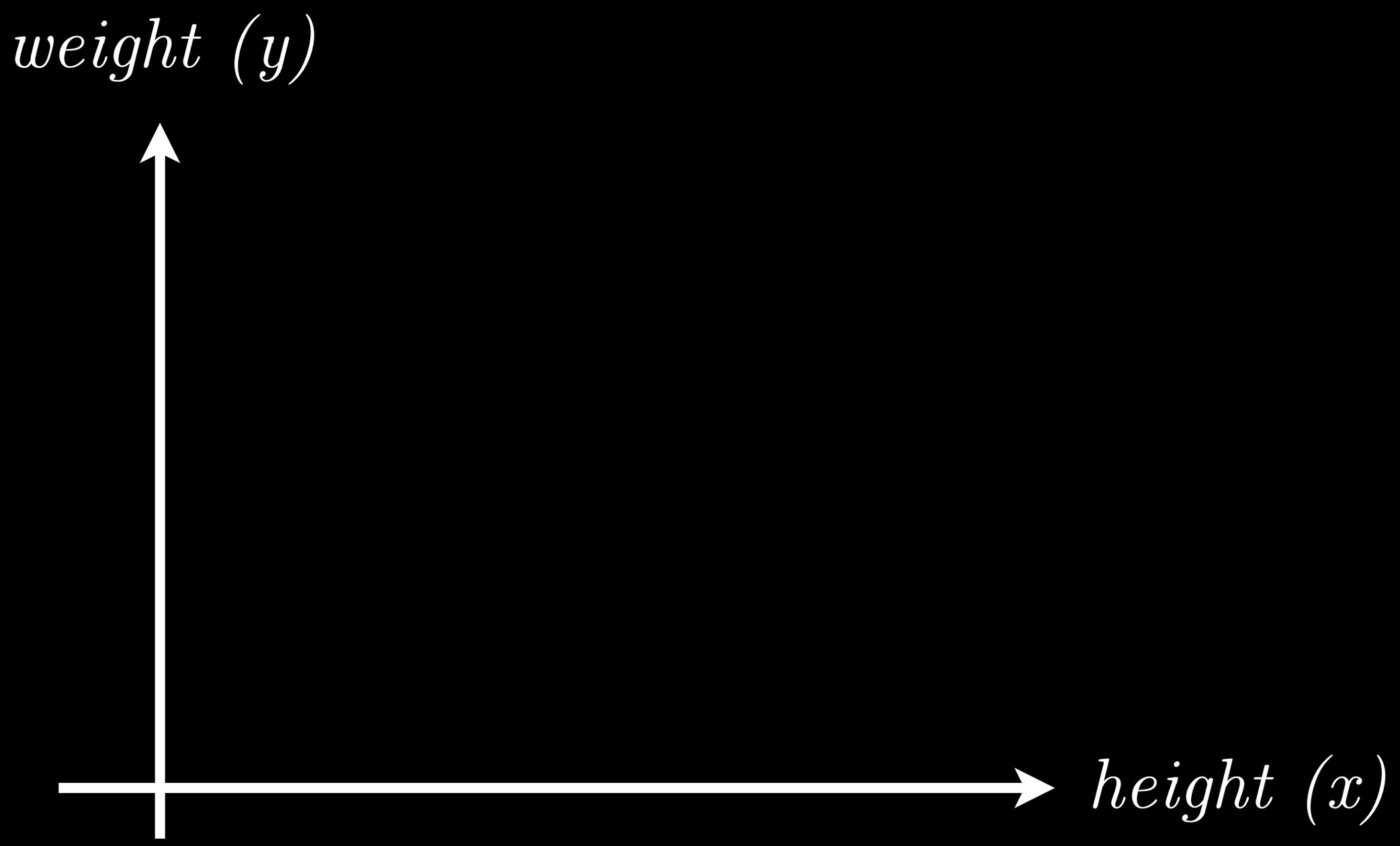
$3mn \times 1$

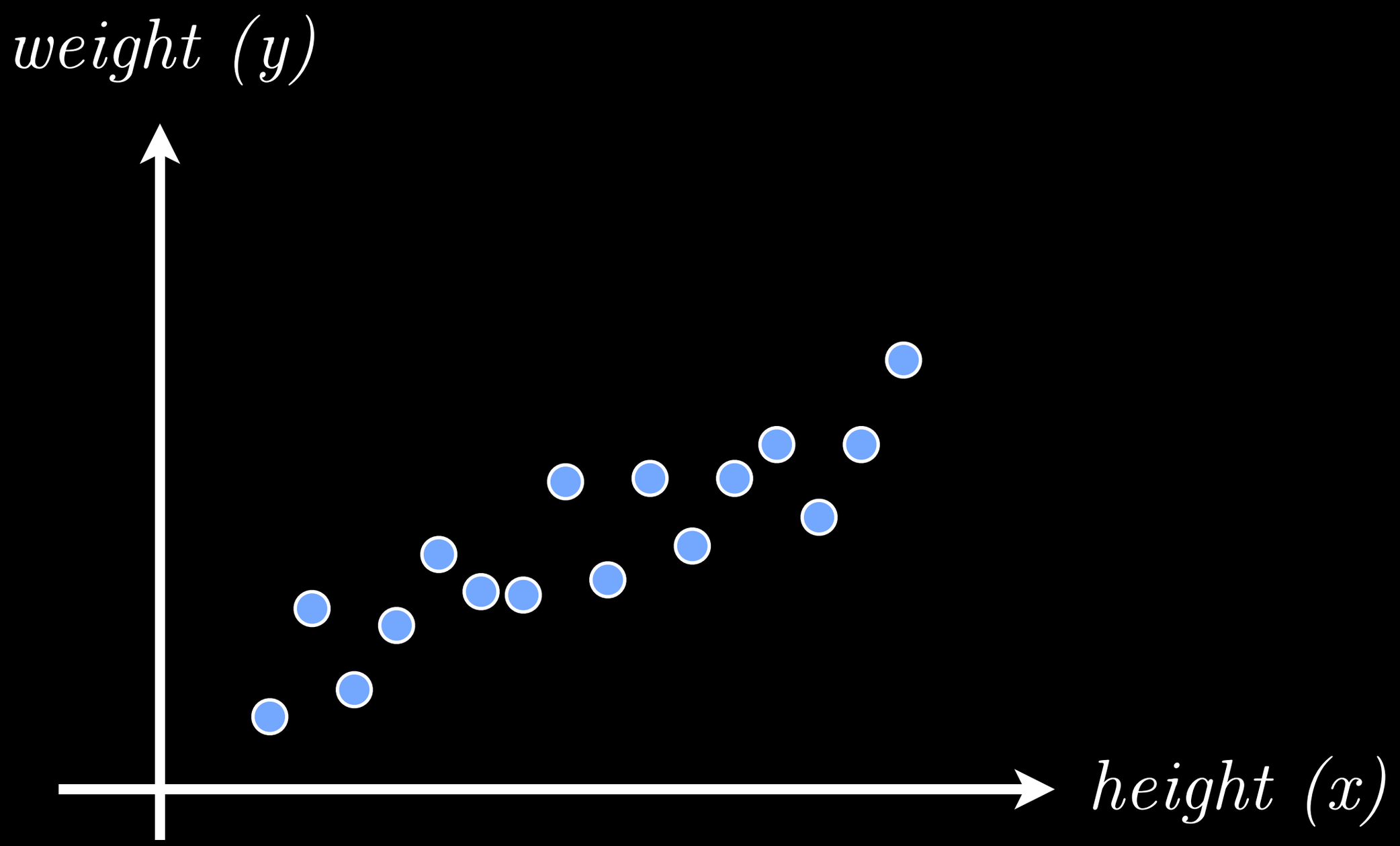
5

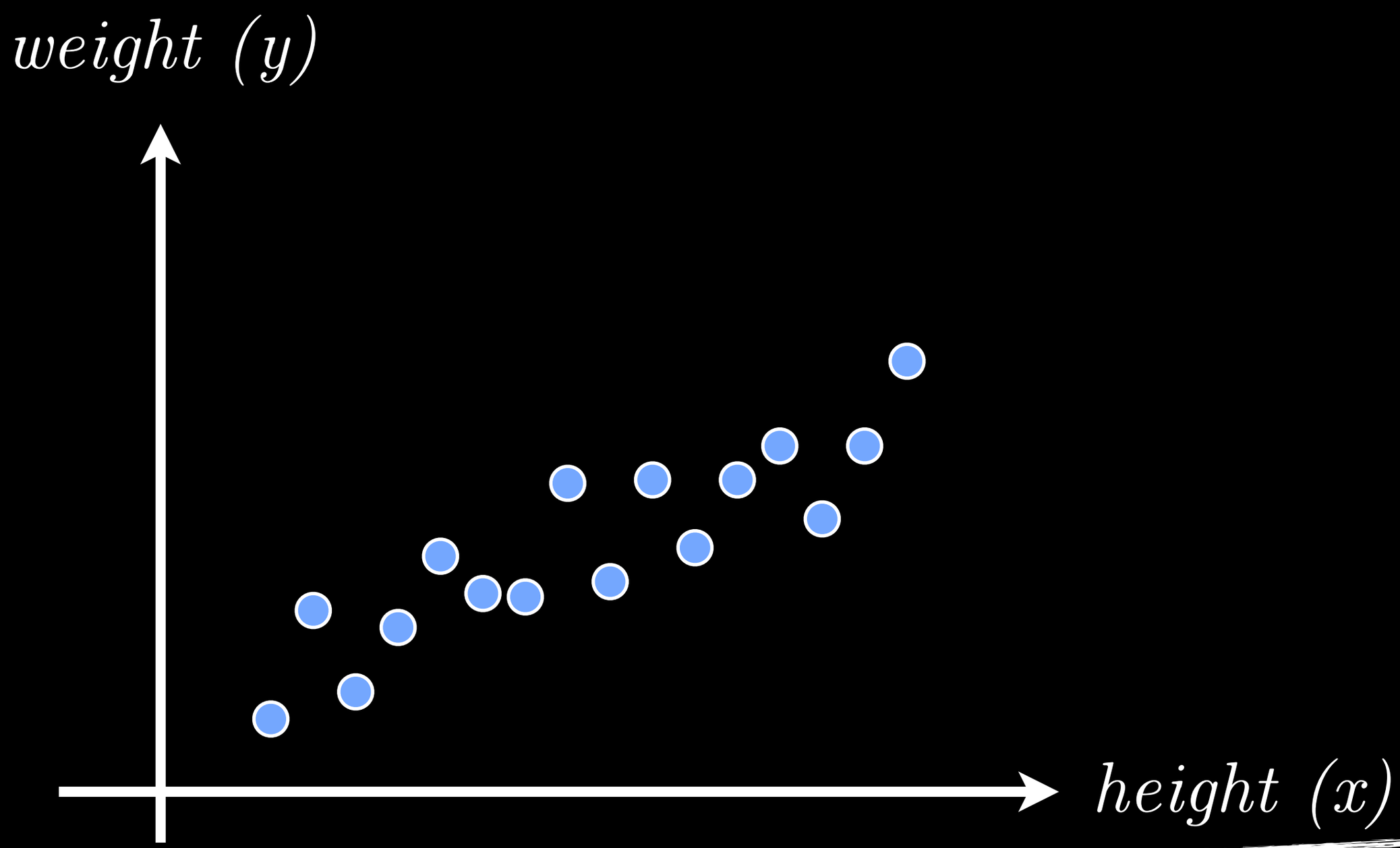
canonical learning problems

1

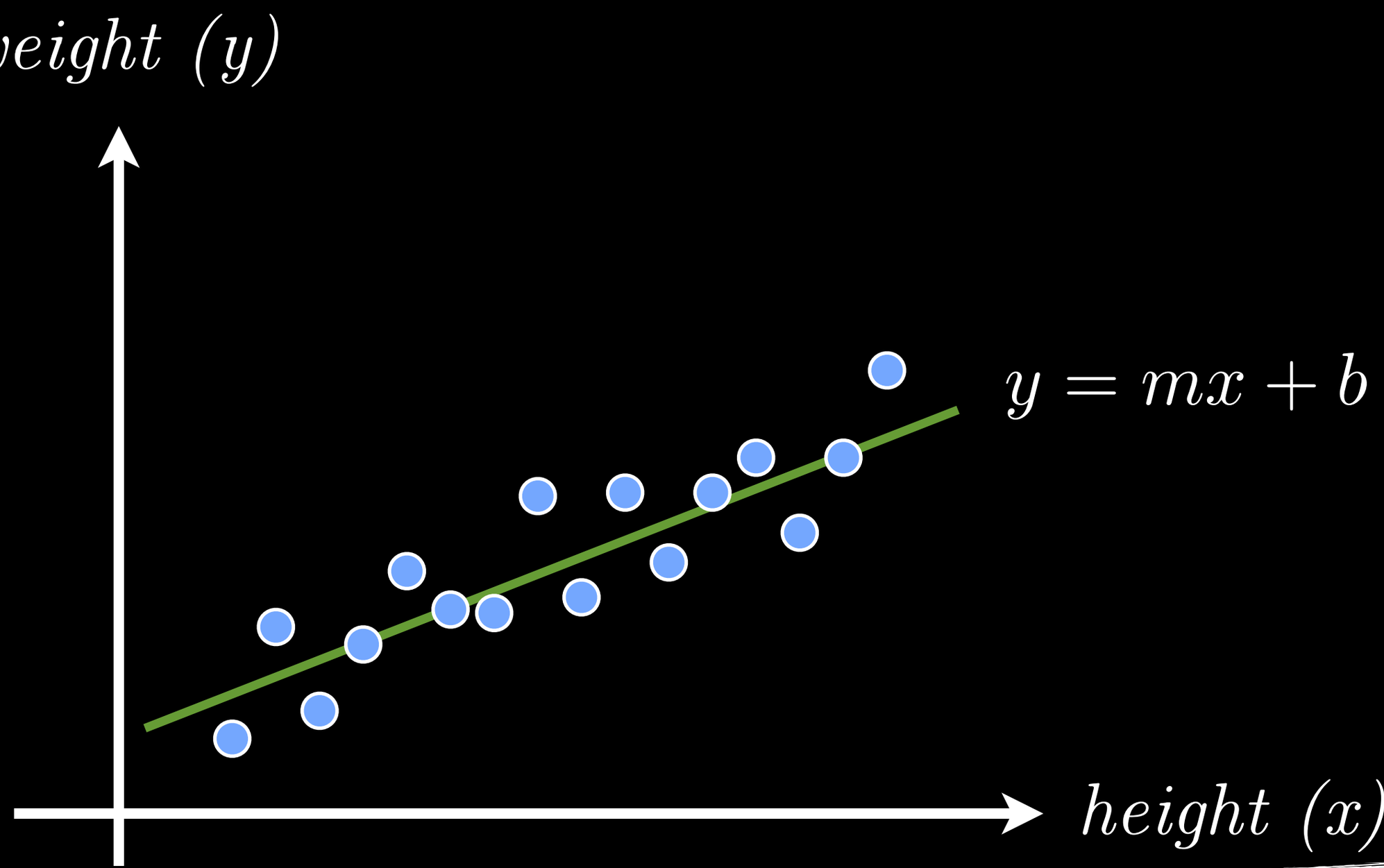
regression


weight (y)
height (x)


$weight\ (y)$
$height\ (x)$



estimate the parameters of the "best" fitting function



estimate the parameters of the "best" fitting function

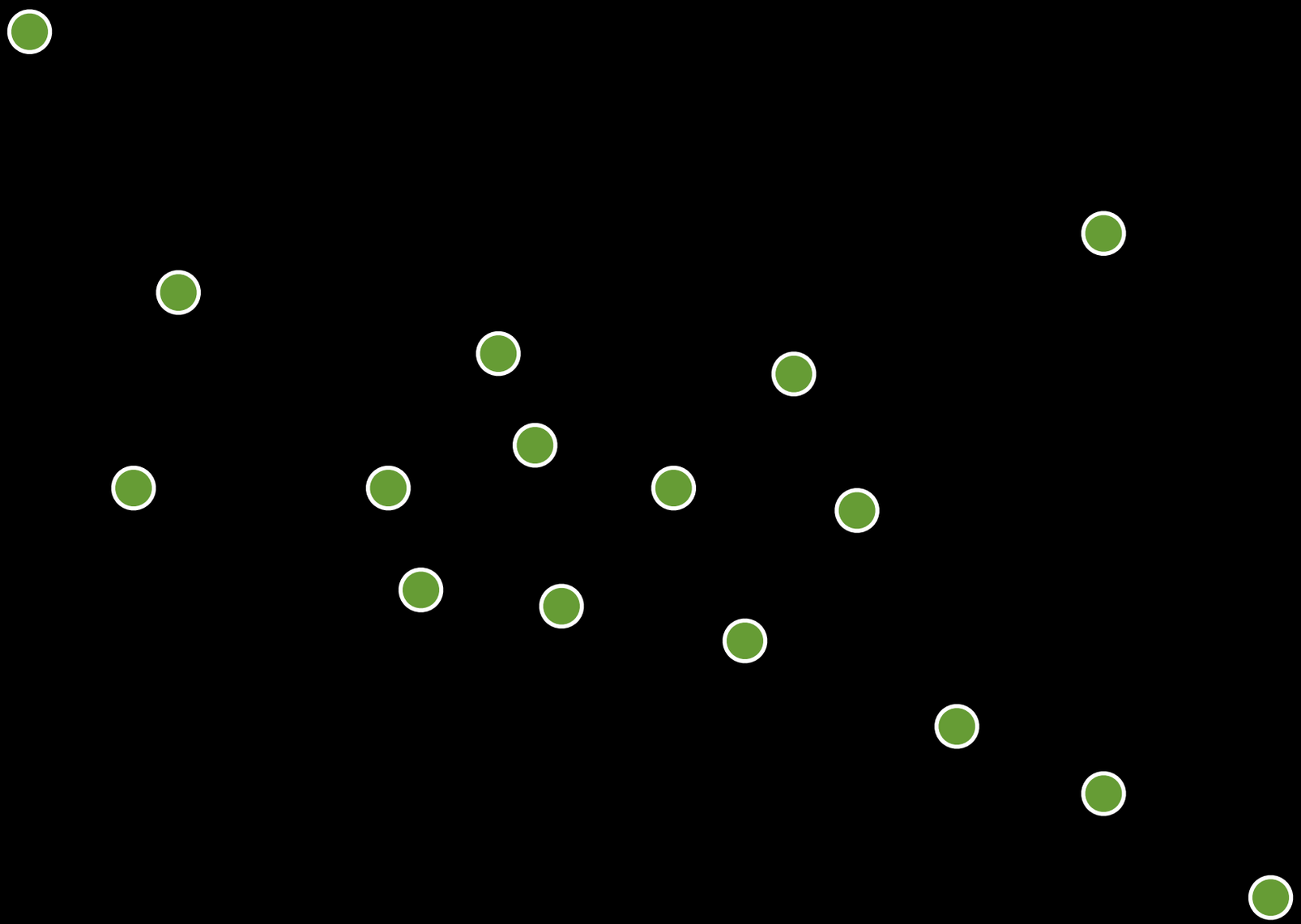
2

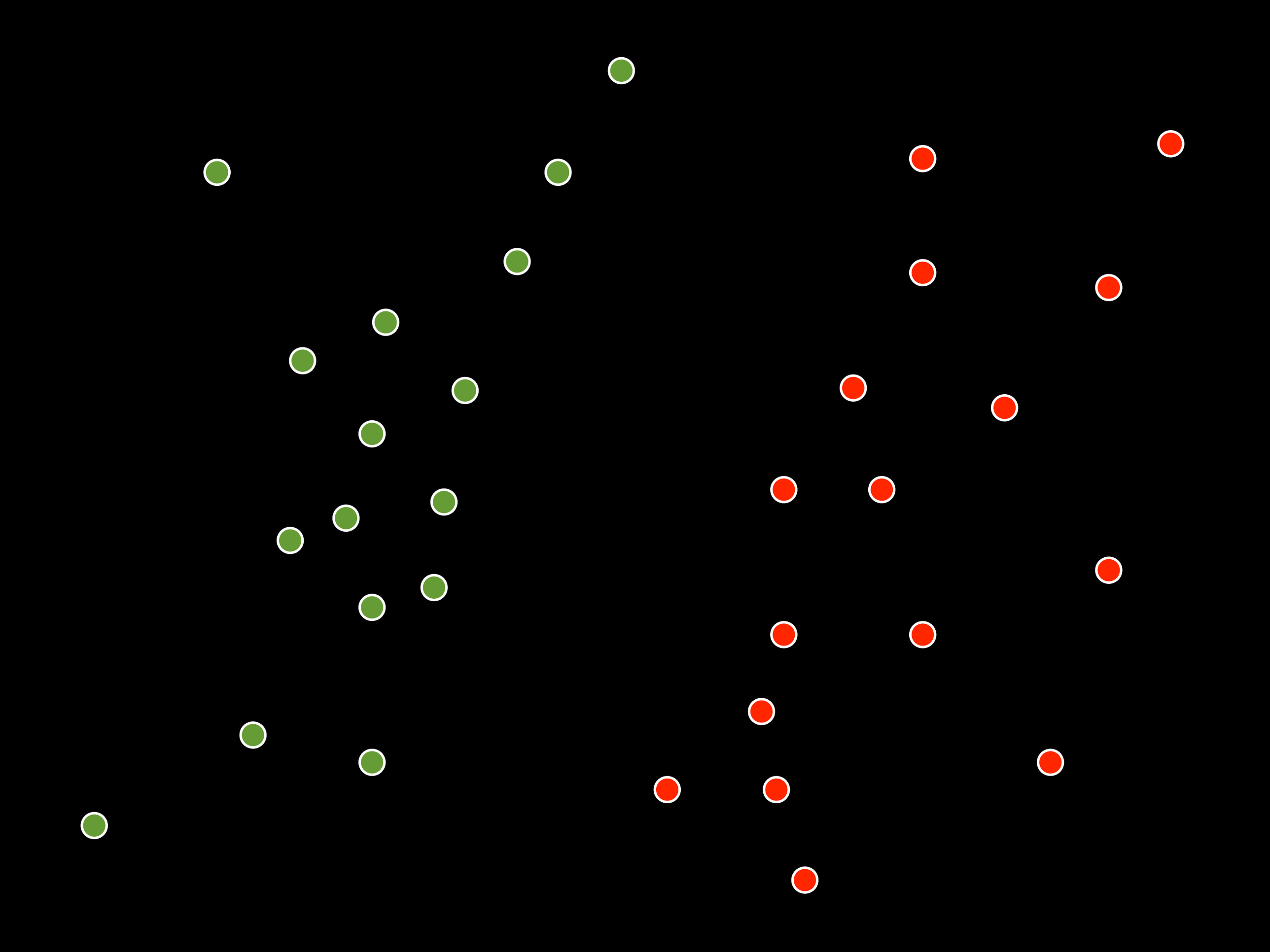
classification

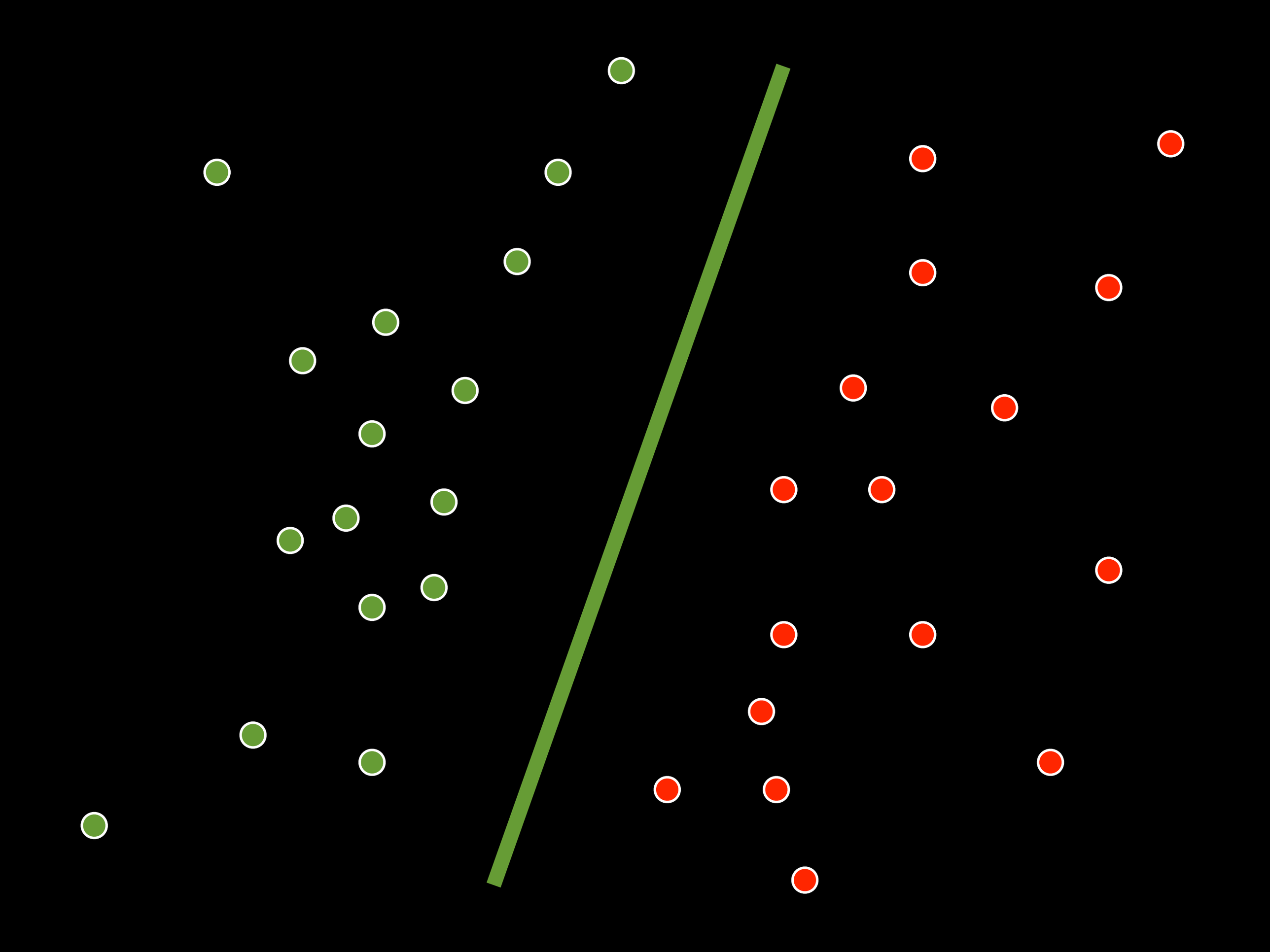
2

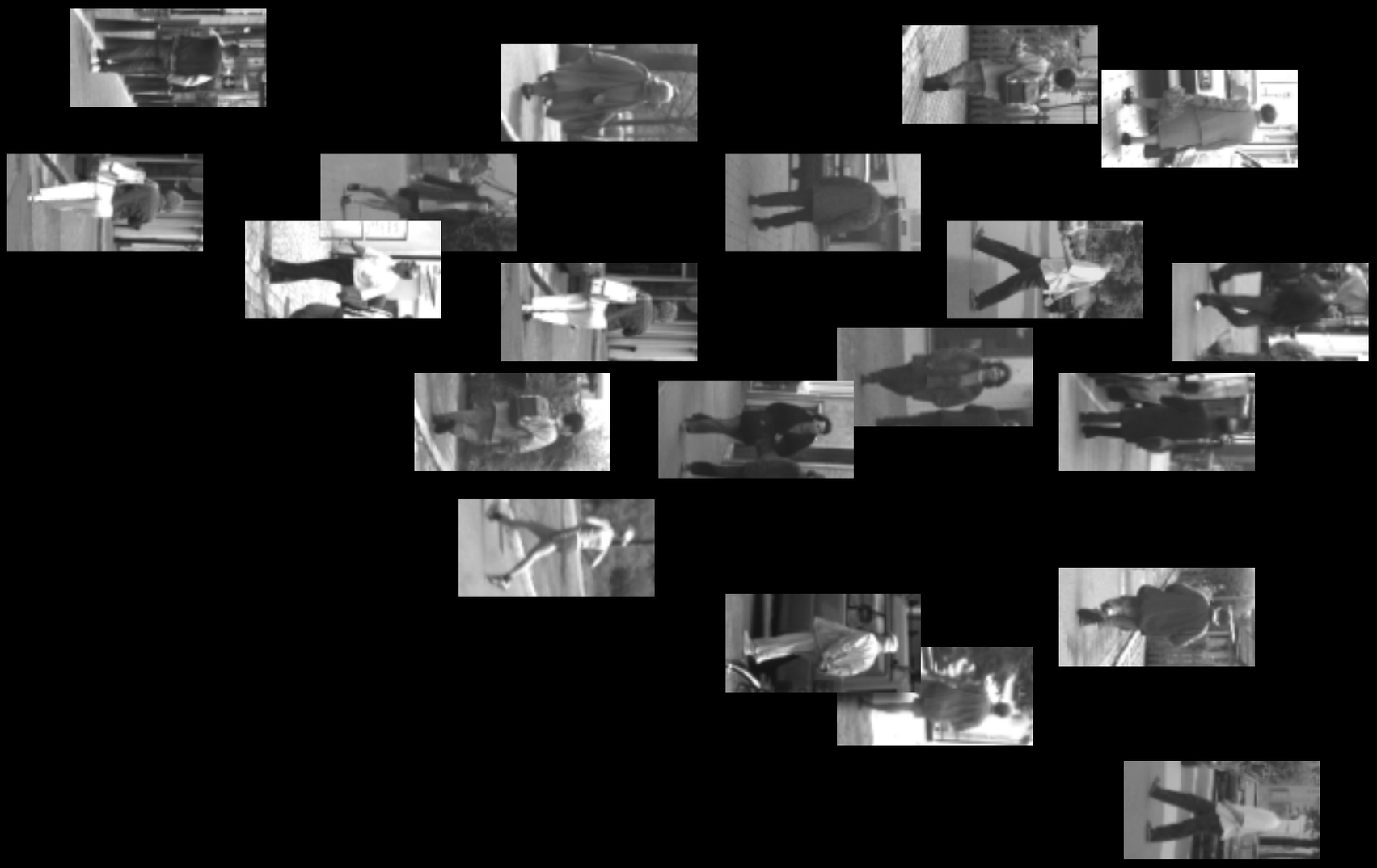
classification

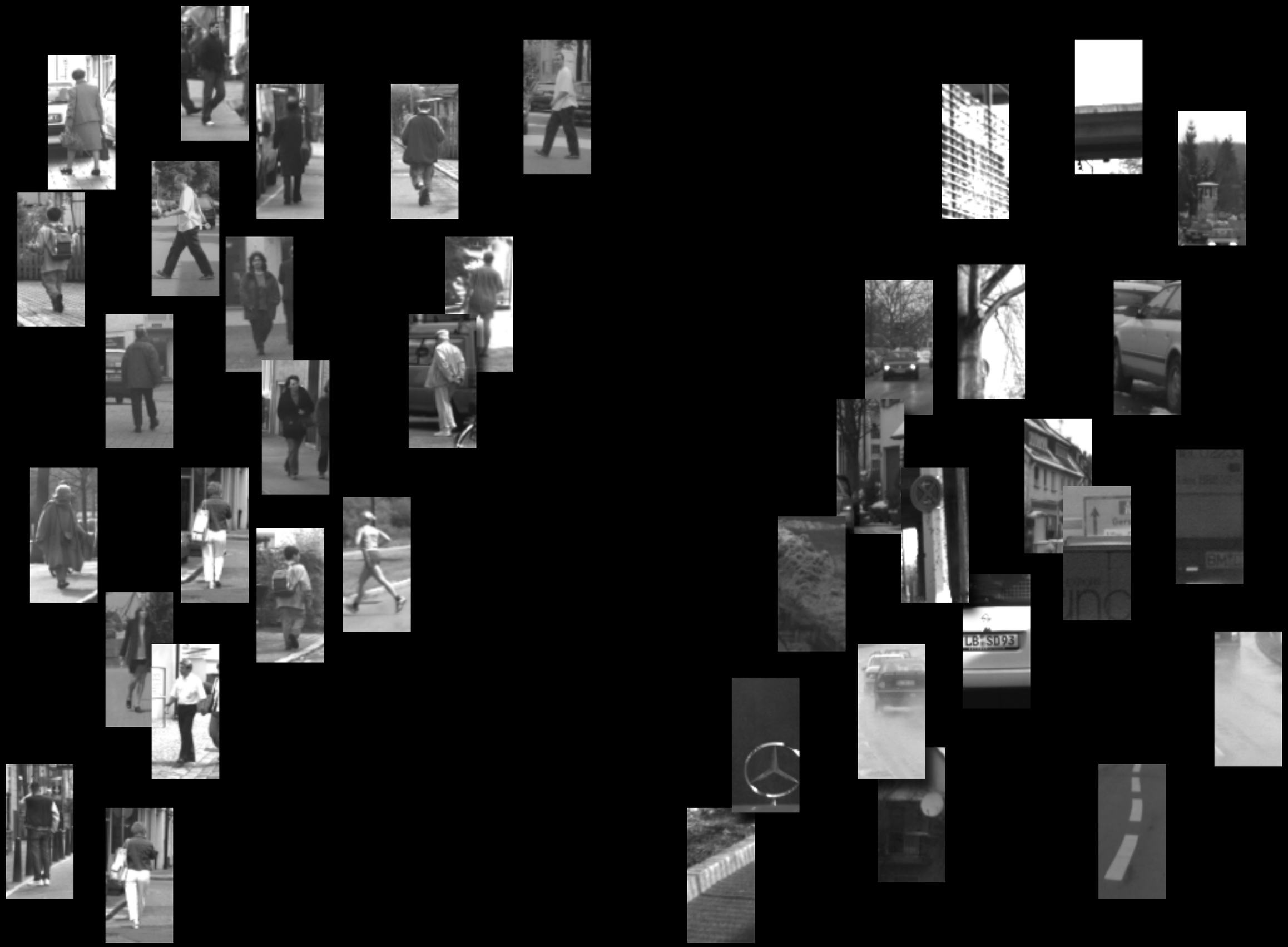
assign input to one of two or more discrete classes

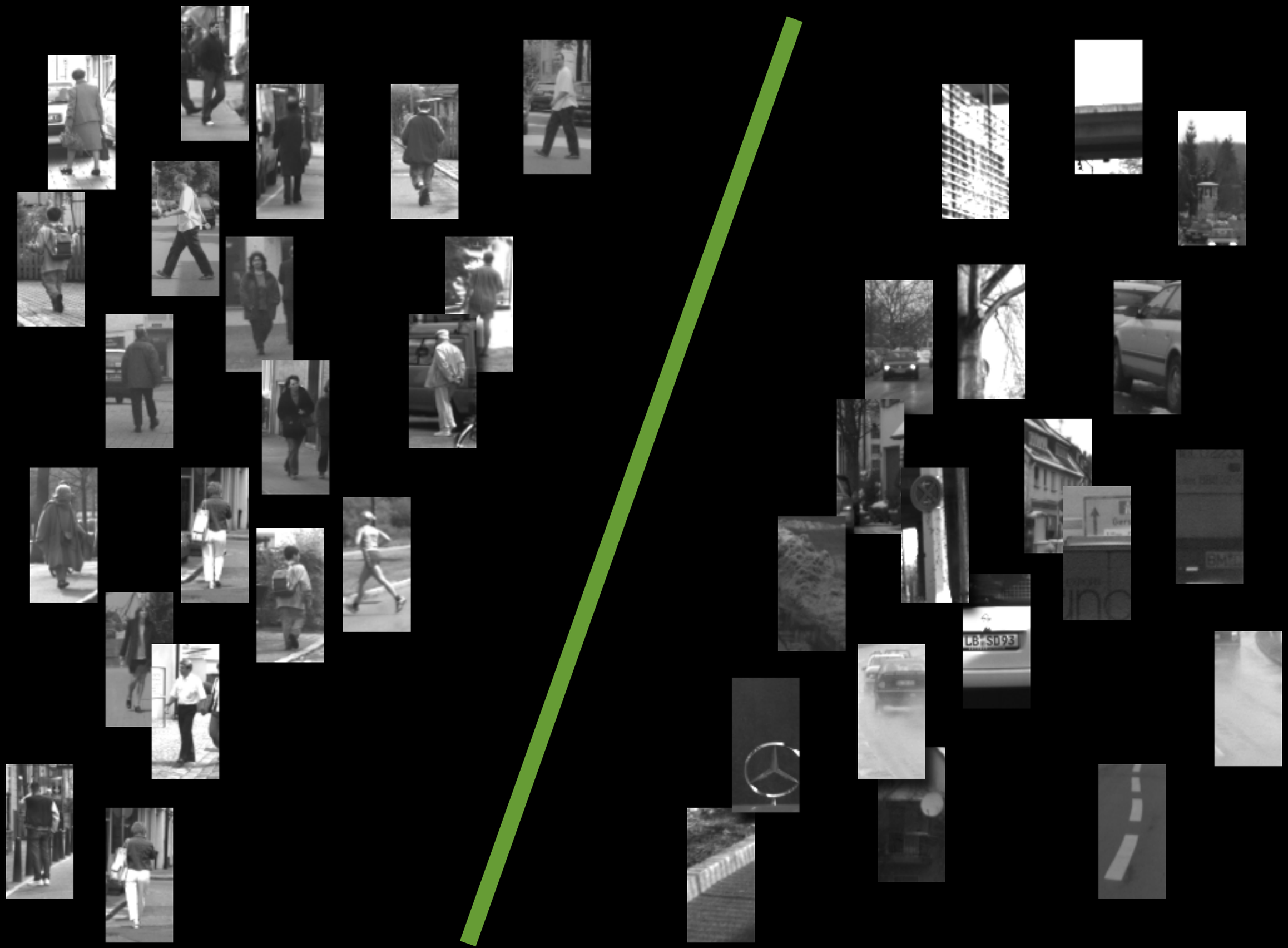


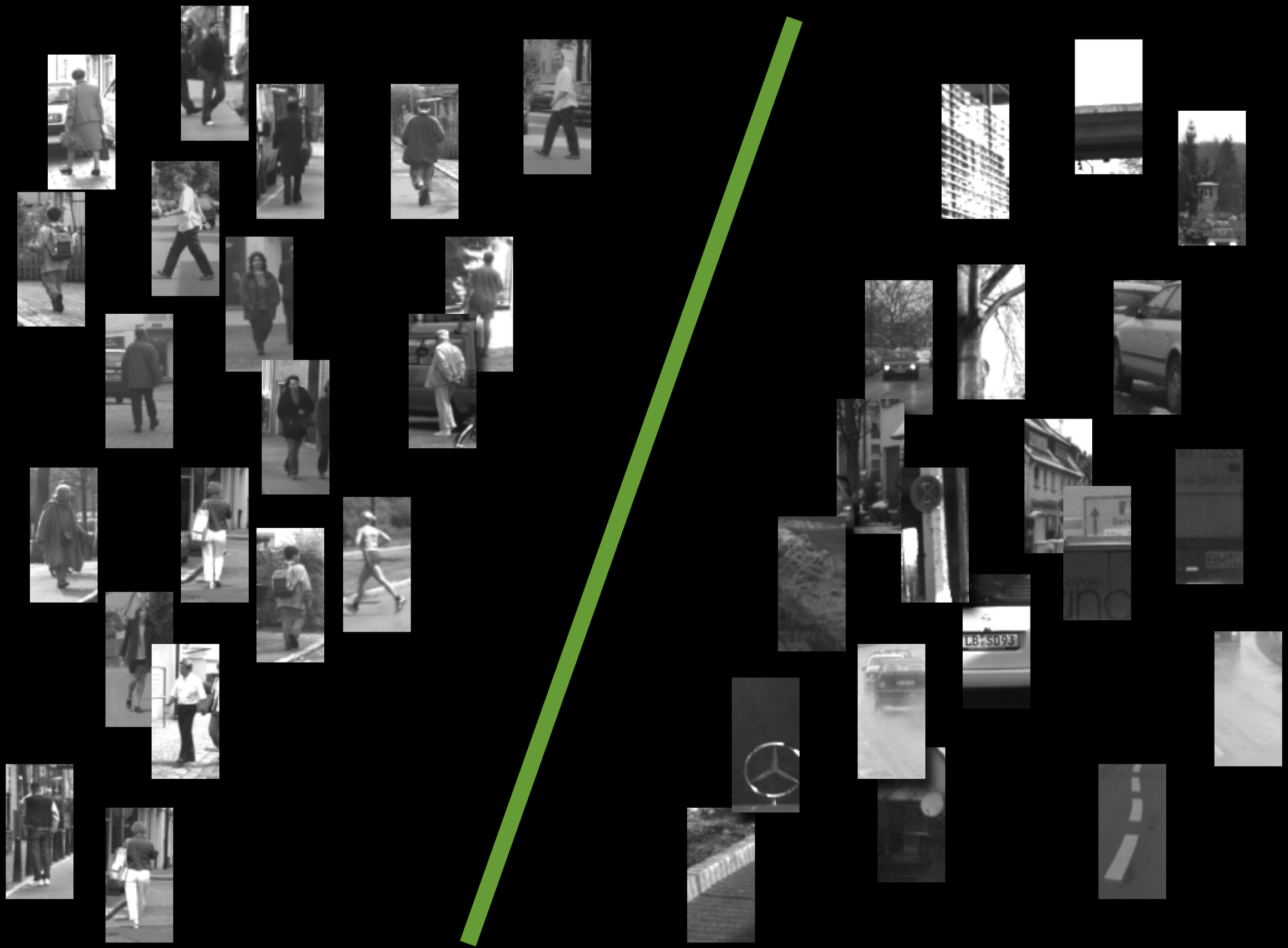


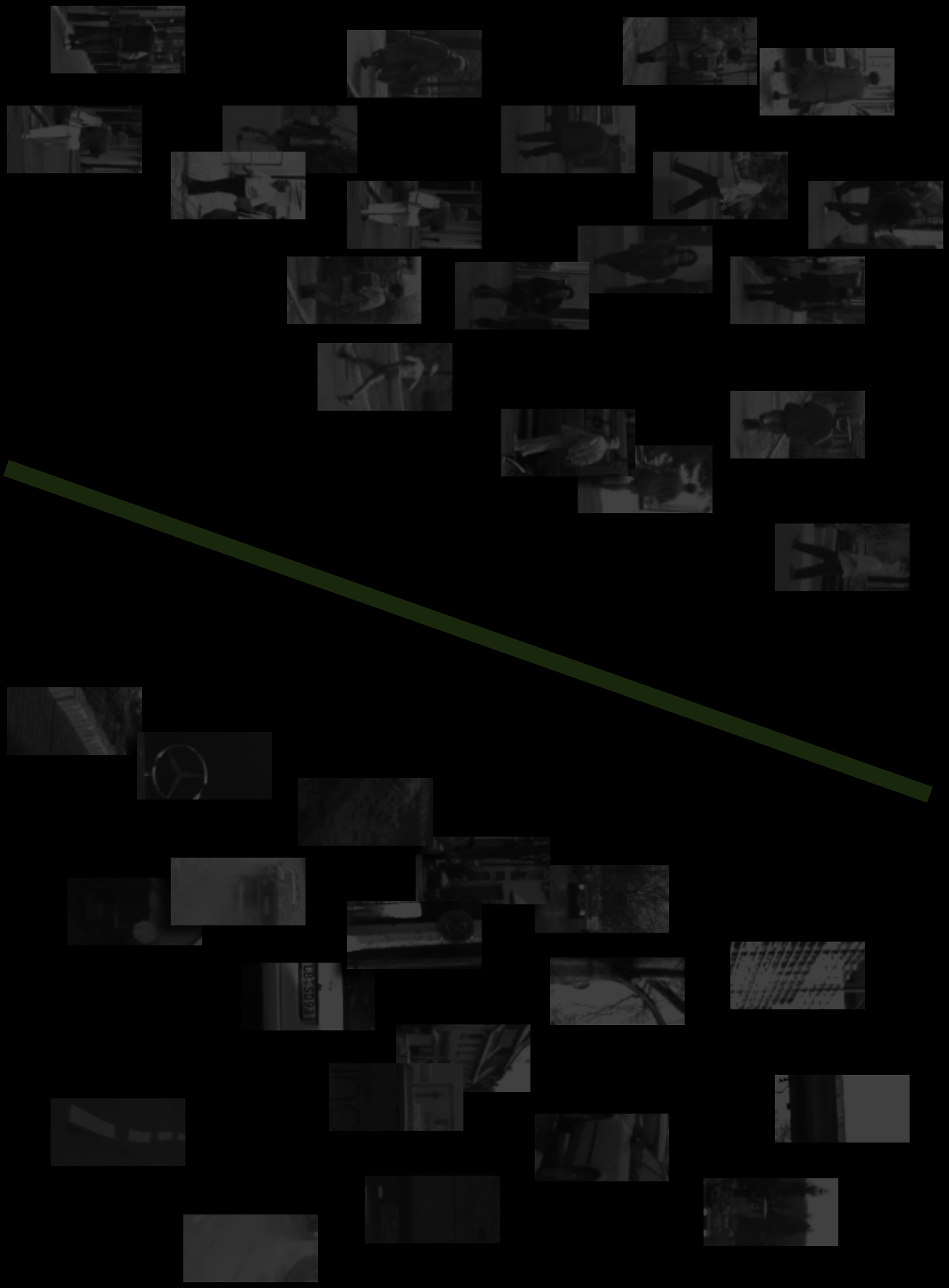


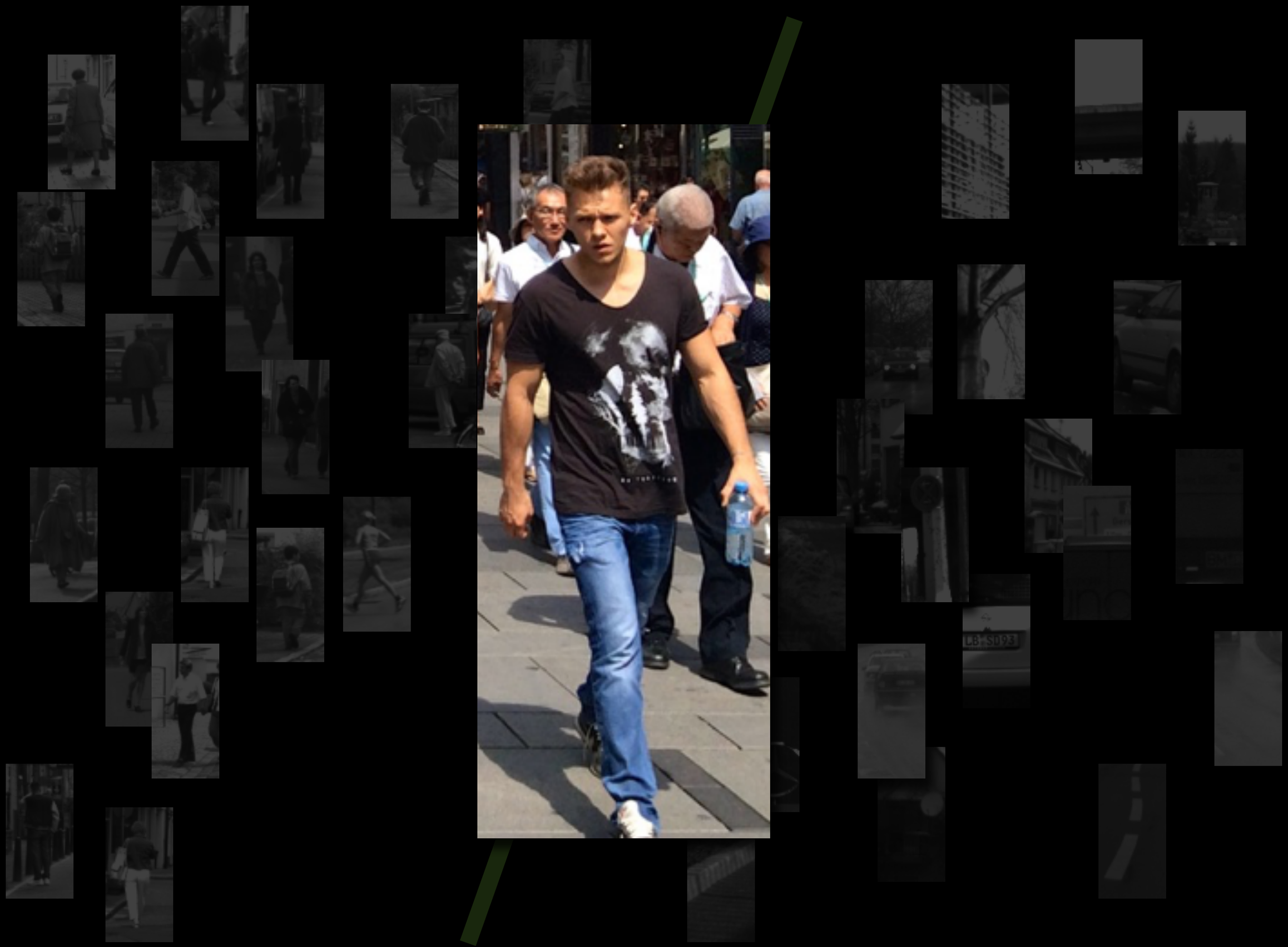


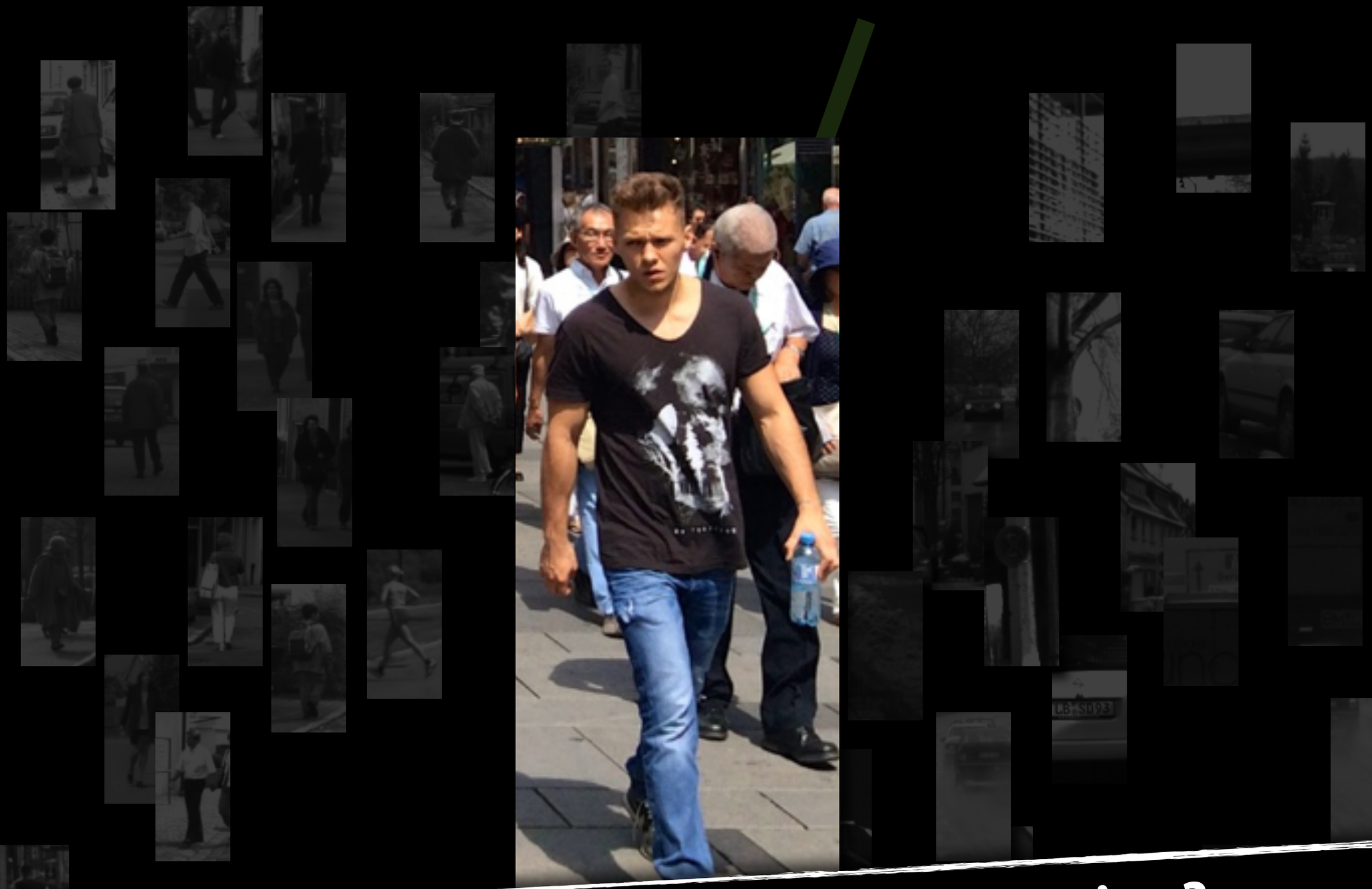

LB SD93


LB-SD93


LB-SD93







Does this image contain a pedestrian?

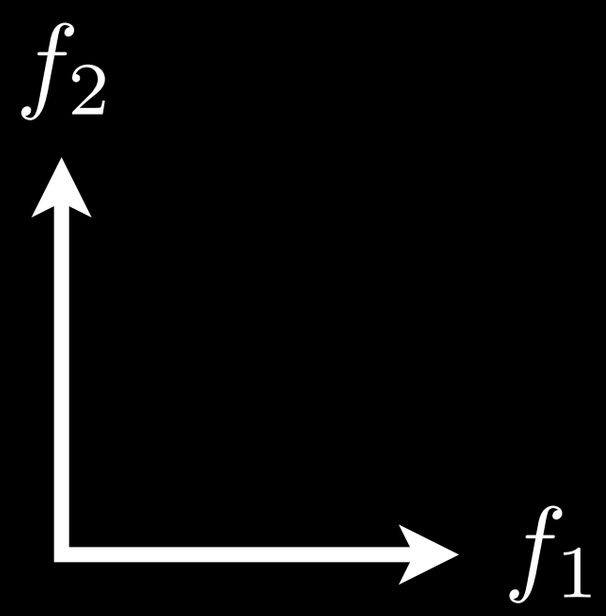
3

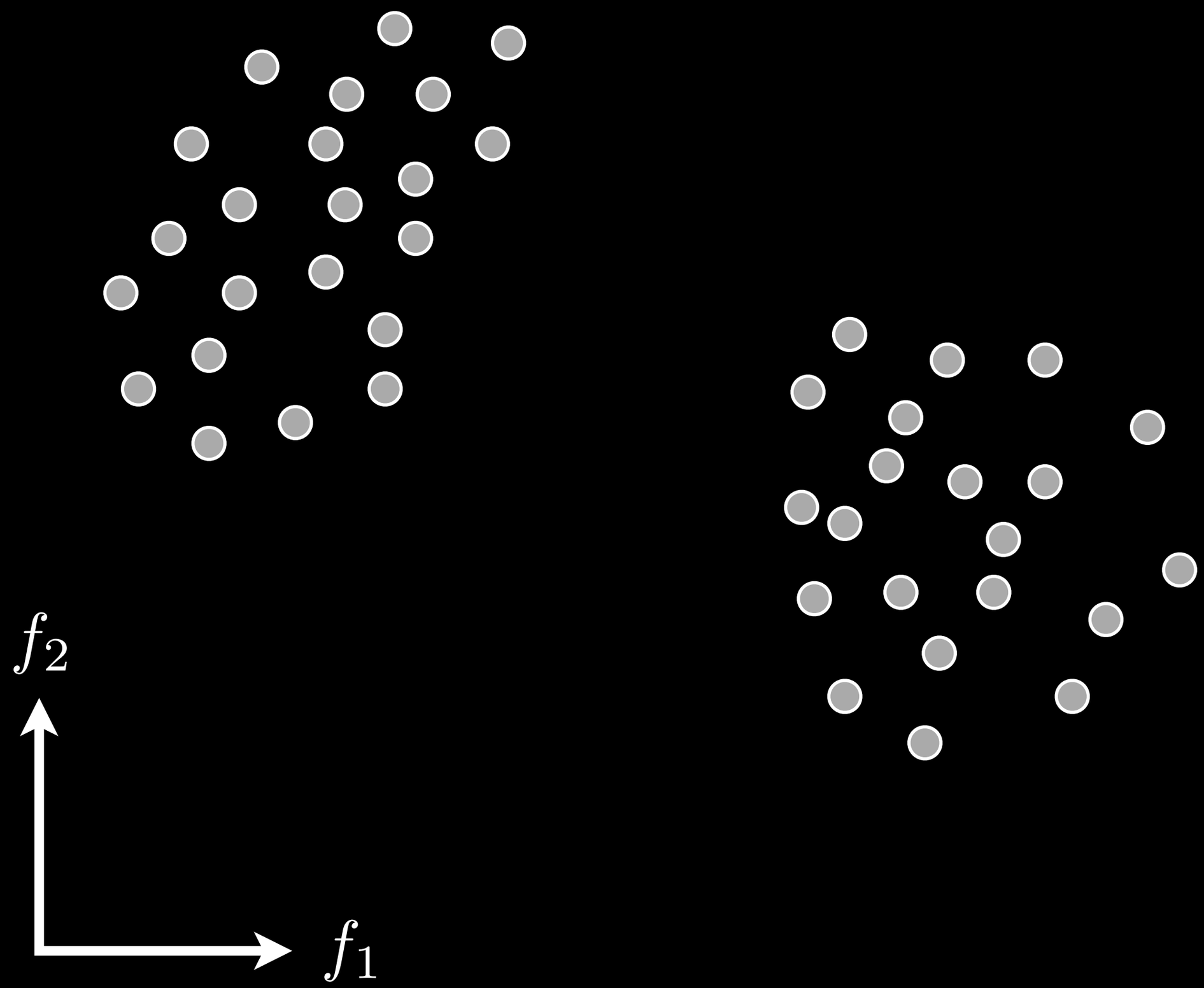
clustering

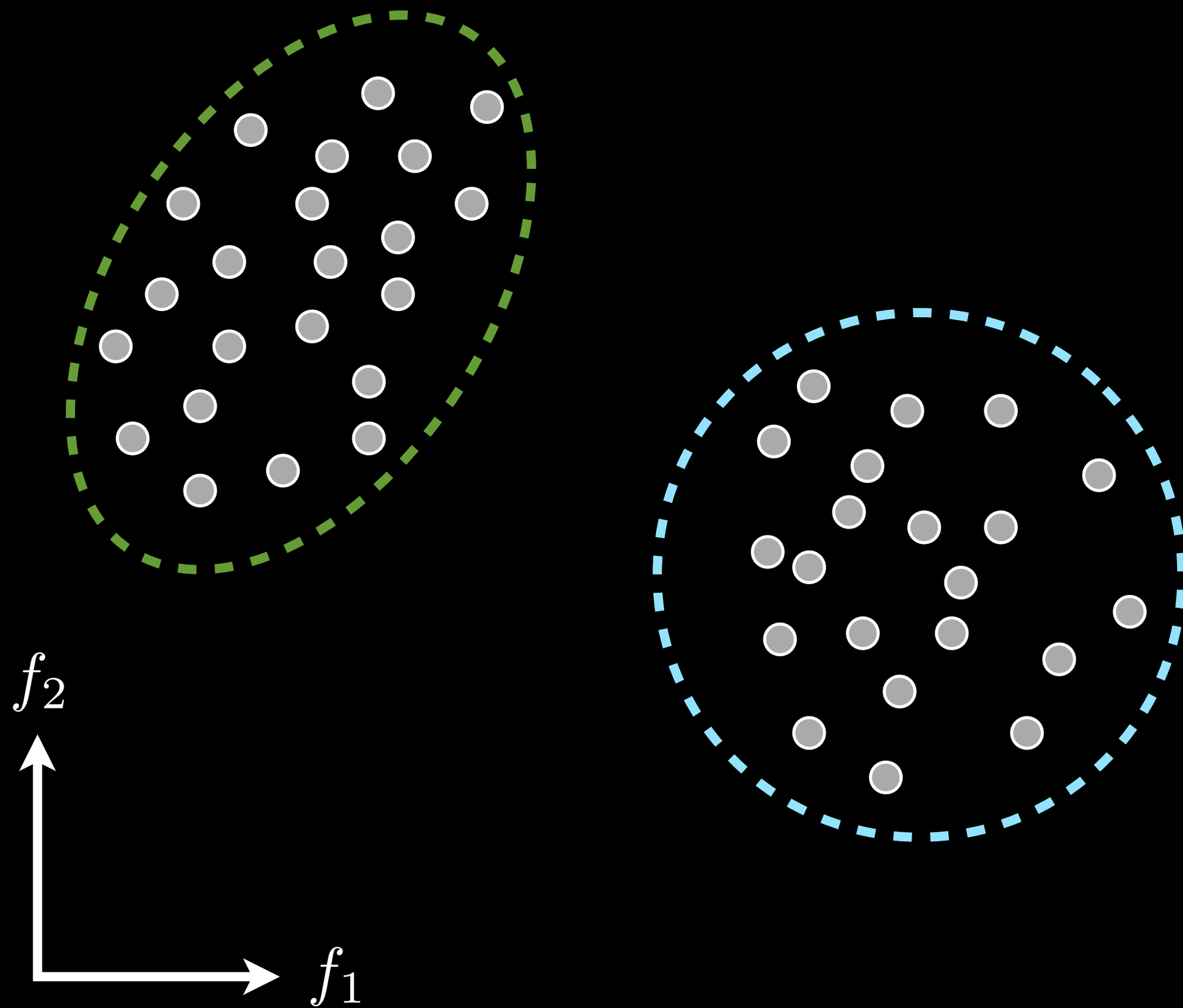
3

clustering

find hidden discrete structure in unlabelled data


f_2
f_1


f_2
f_1


f_2
f_1

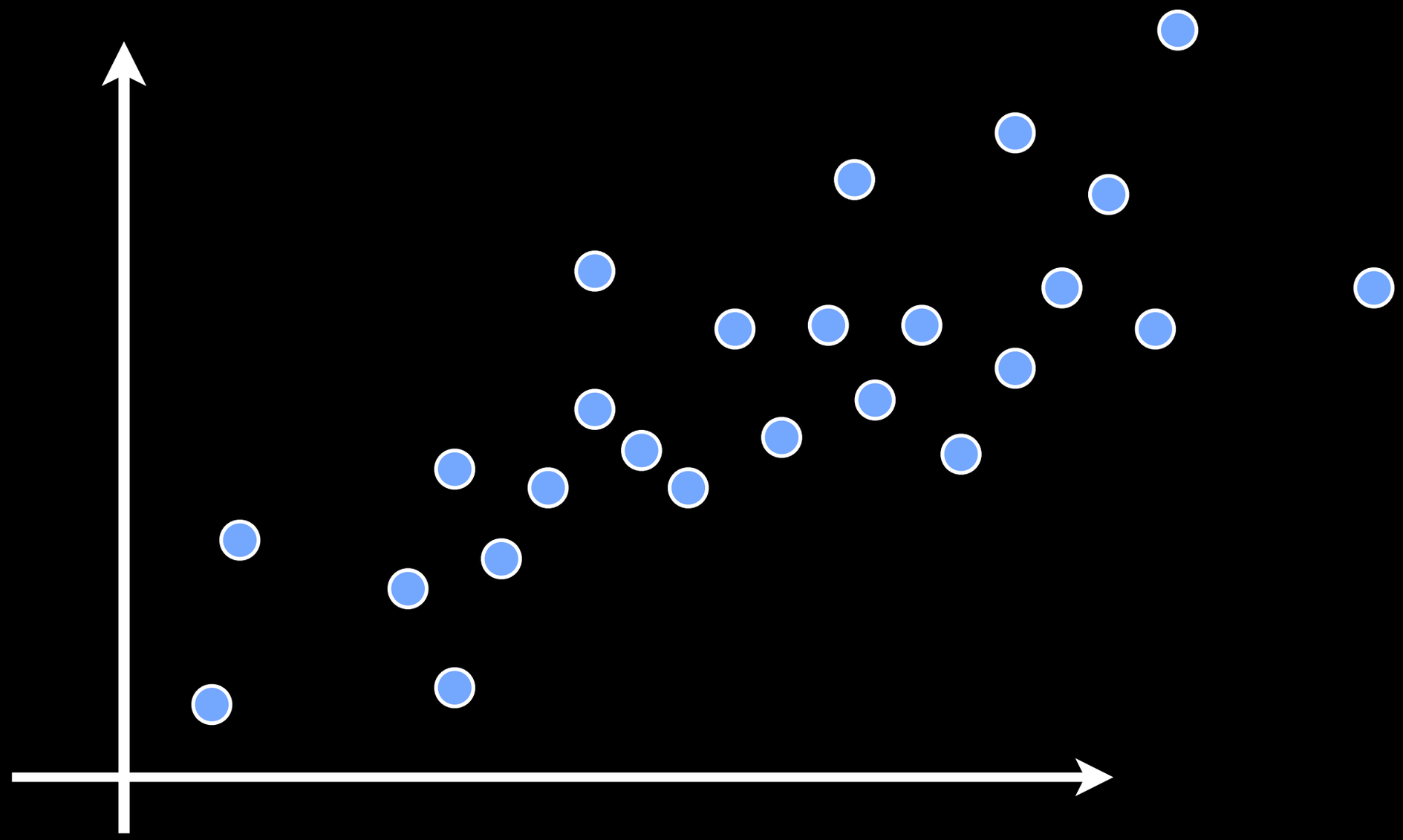
4

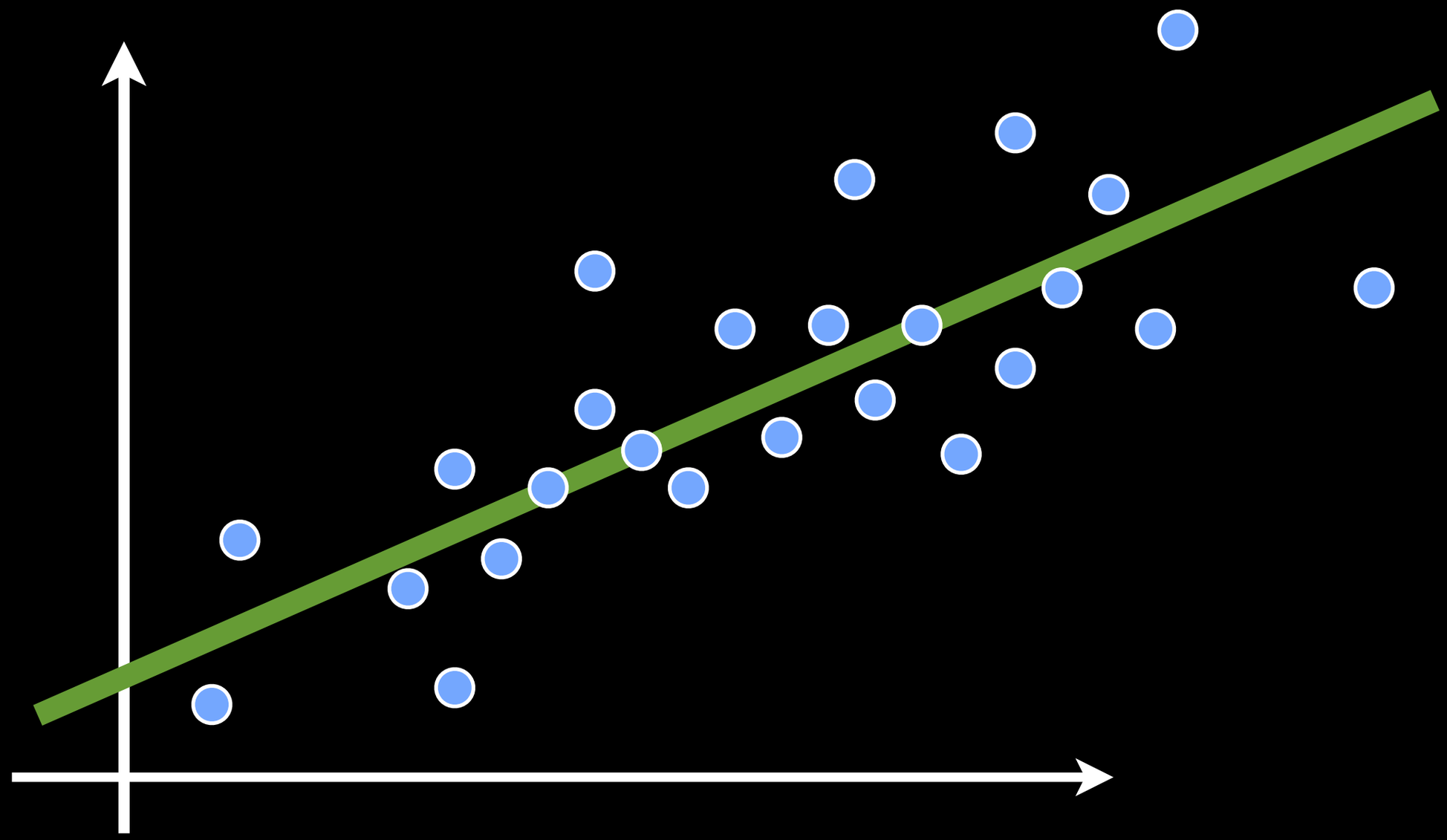
dimensionality reduction

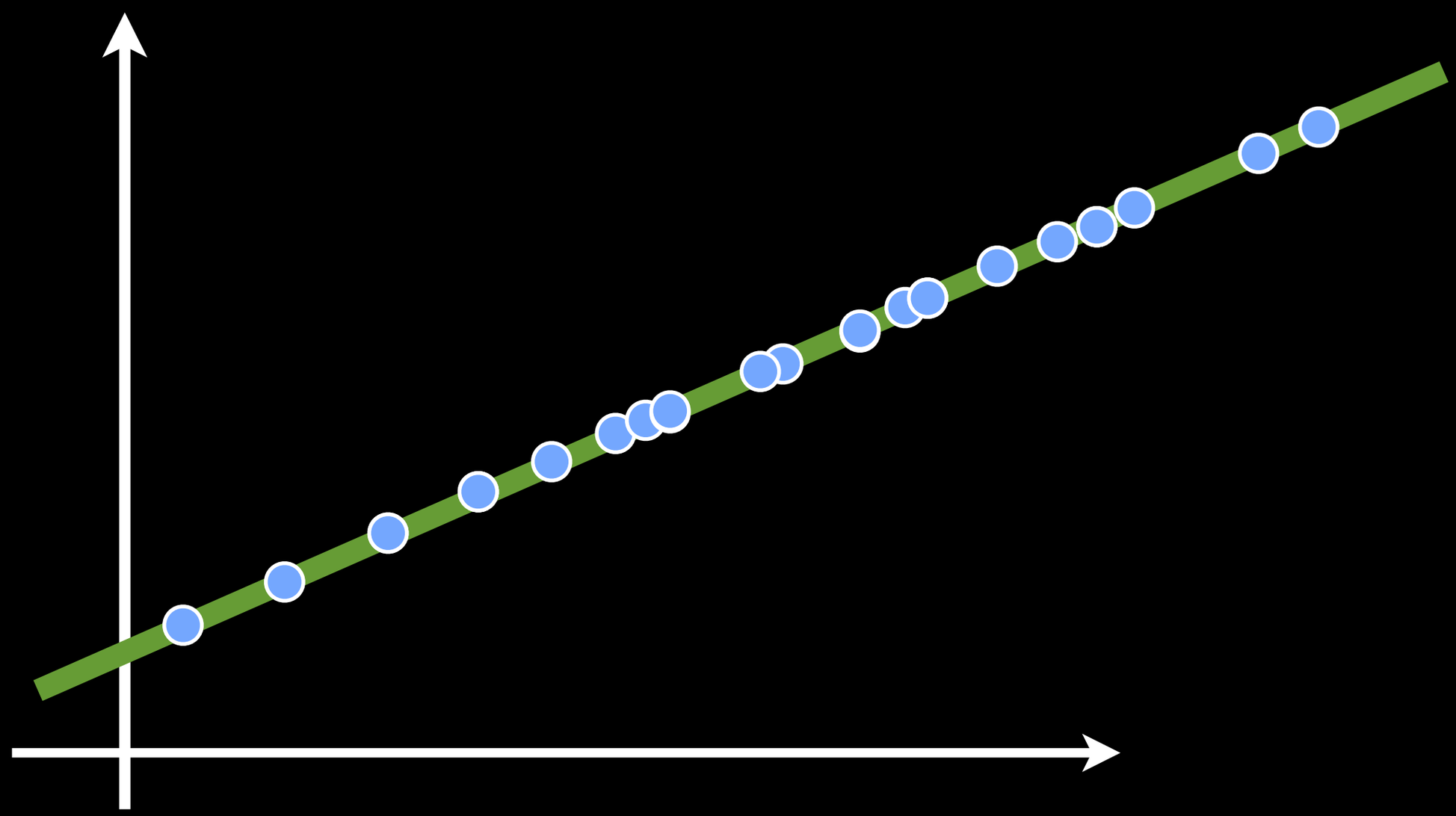
4

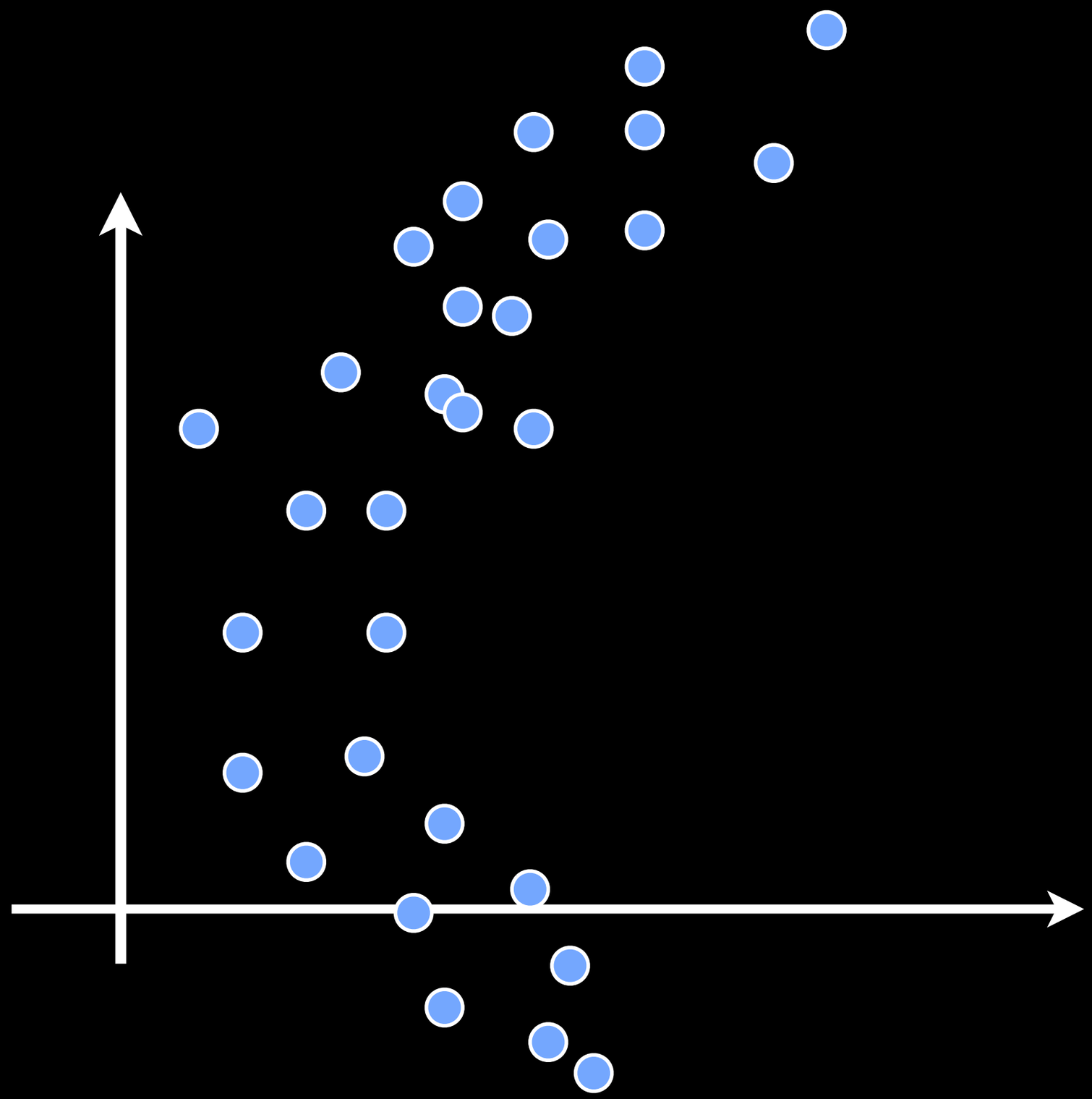
dimensionality reduction

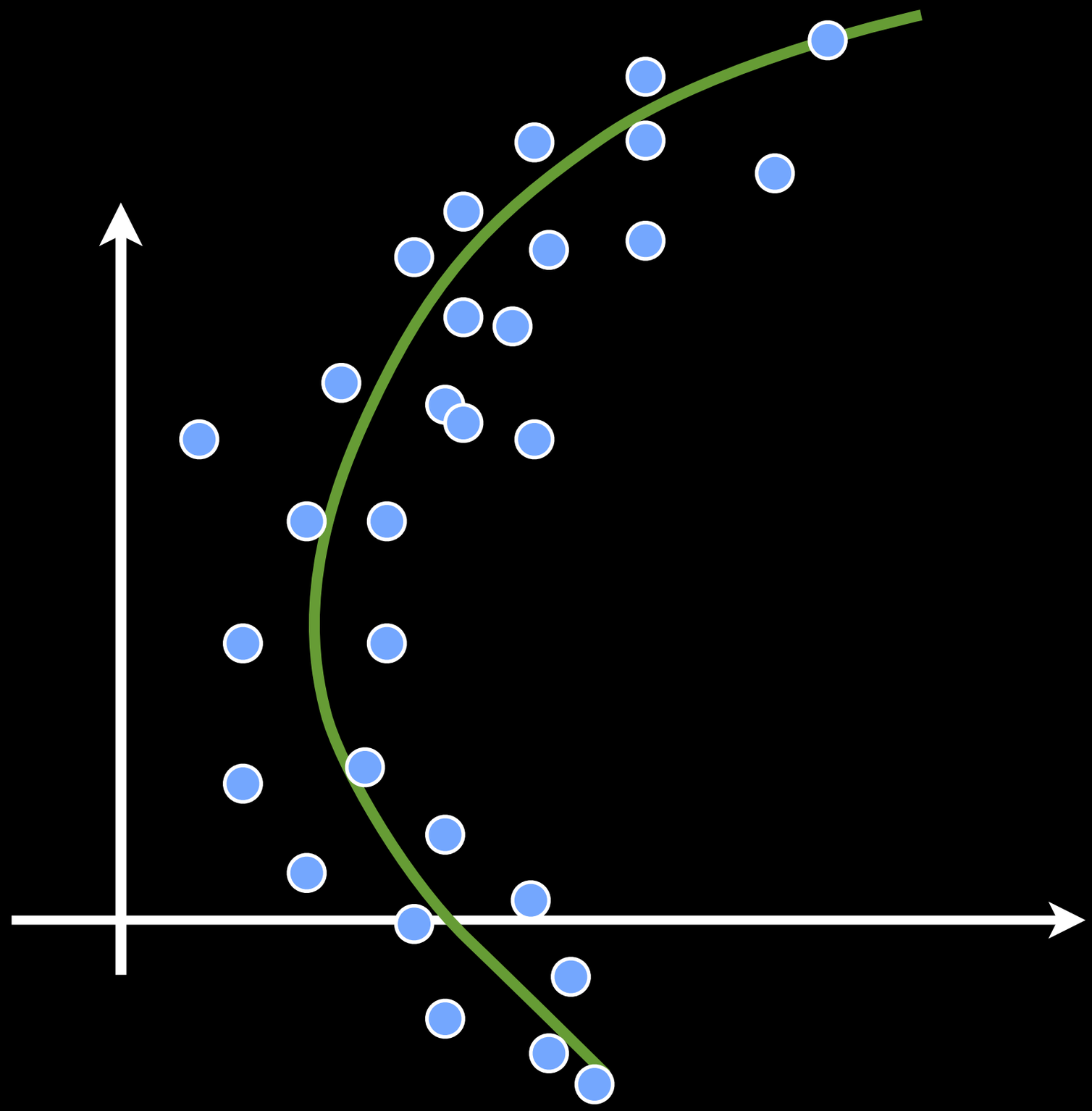
find hidden continuous structure in data

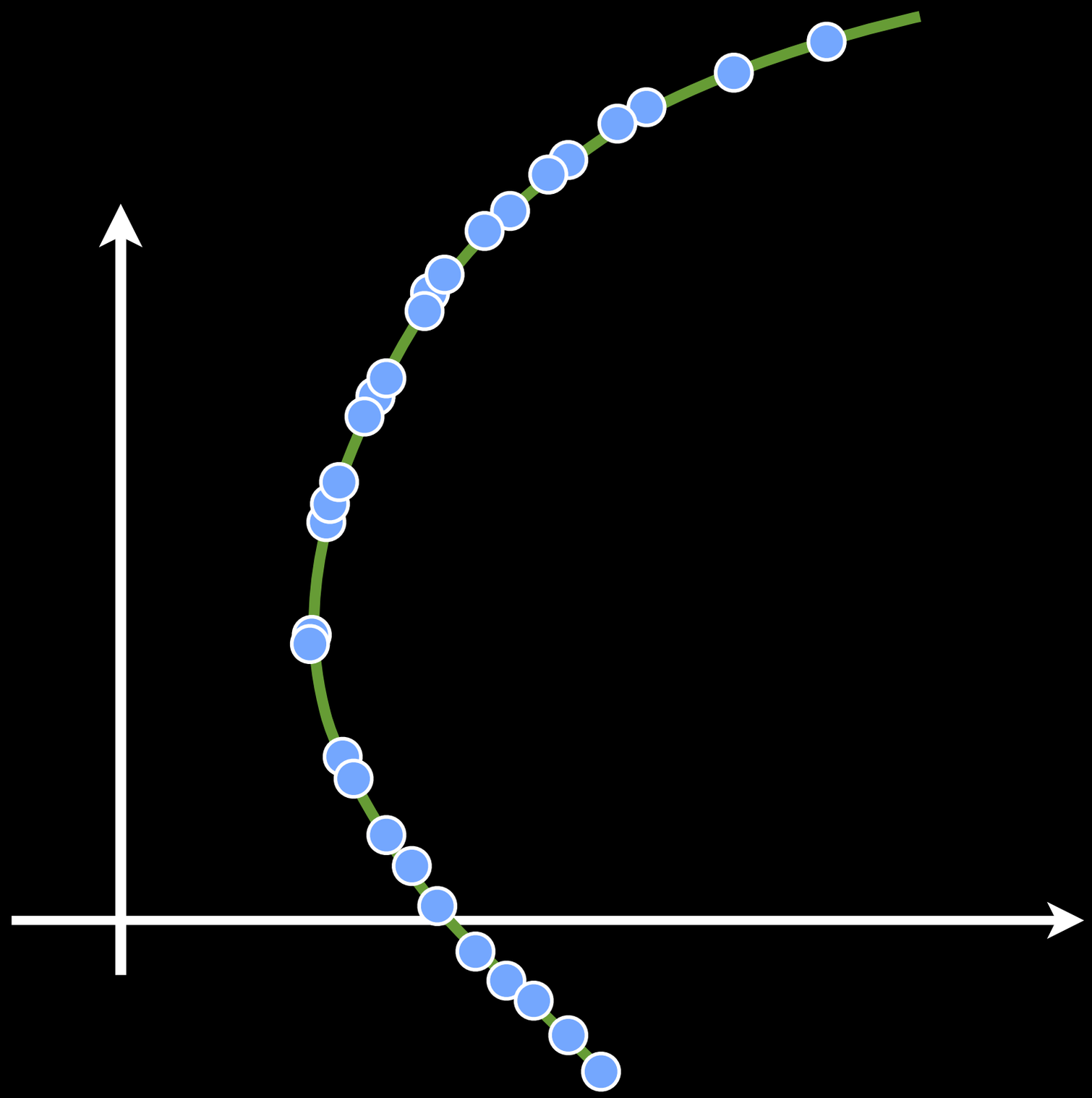












Summary

	supervised	unsupervised
discrete		
continuous		

Summary

	supervised	unsupervised
discrete	classification	
continuous		

Summary

	supervised	unsupervised
discrete	classification	
continuous	regression	

Summary

	supervised	unsupervised
discrete	classification	clustering
continuous	regression	

Summary

	supervised	unsupervised
discrete	classification	clustering
continuous	regression	dimensionality reduction

5

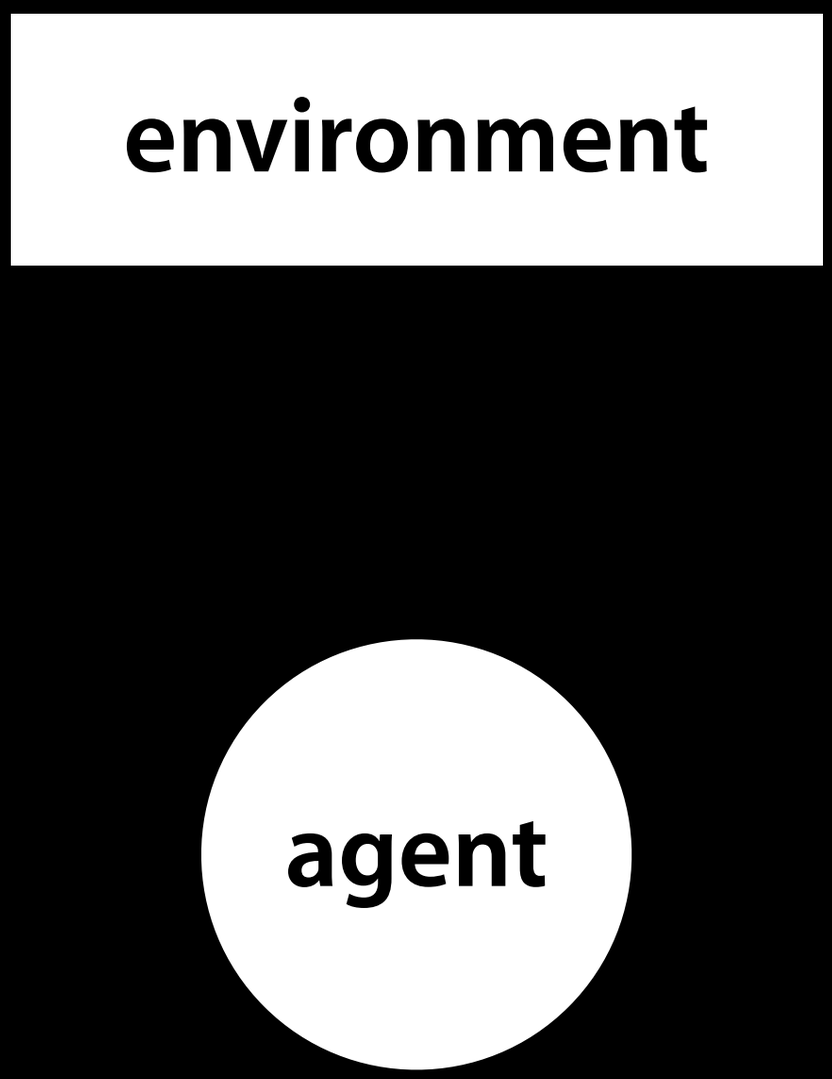
reinforcement learning

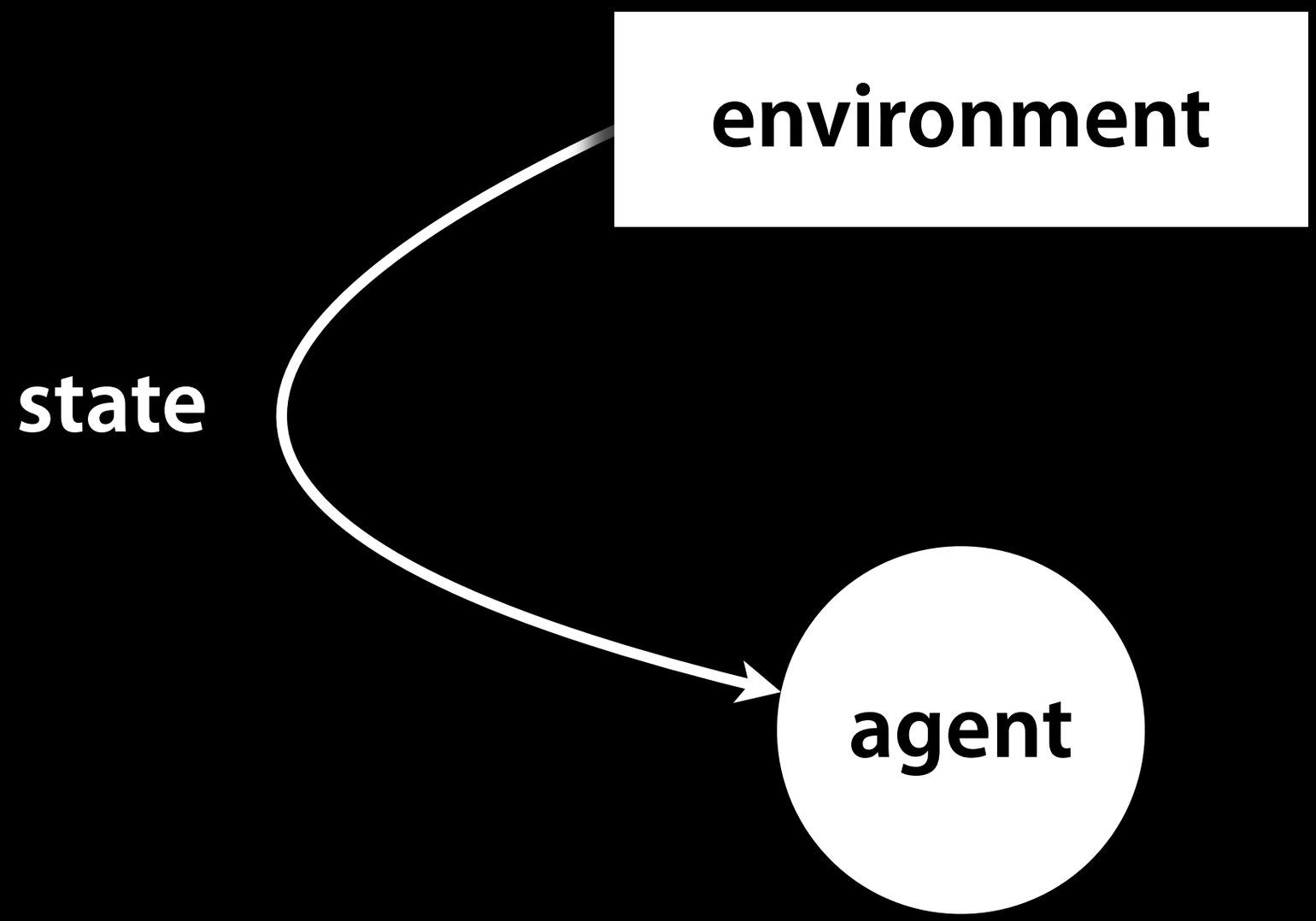
5

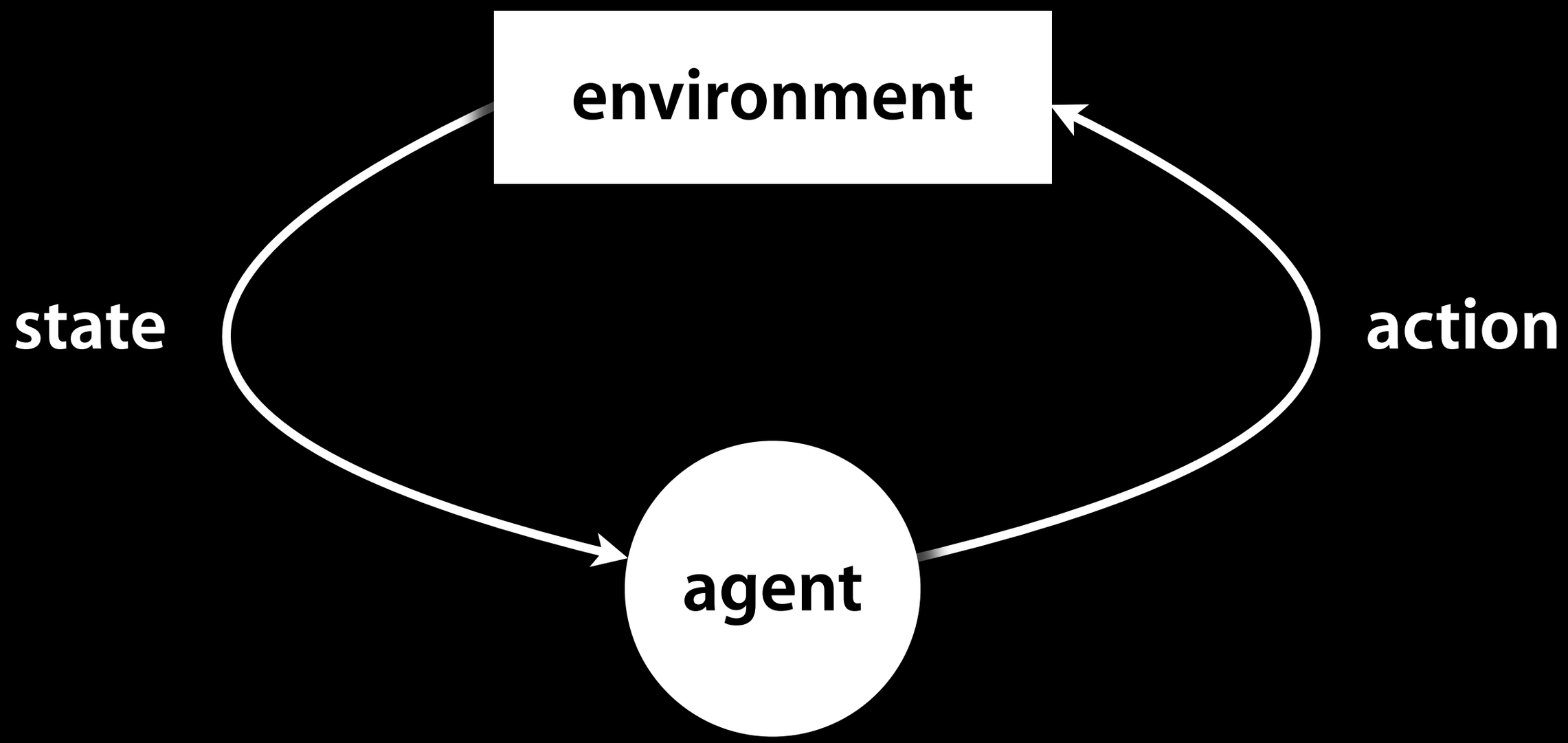
reinforcement learning

learning by interacting with environment

agent


environment
agent


environment
state
agent


environment
state
action
agent

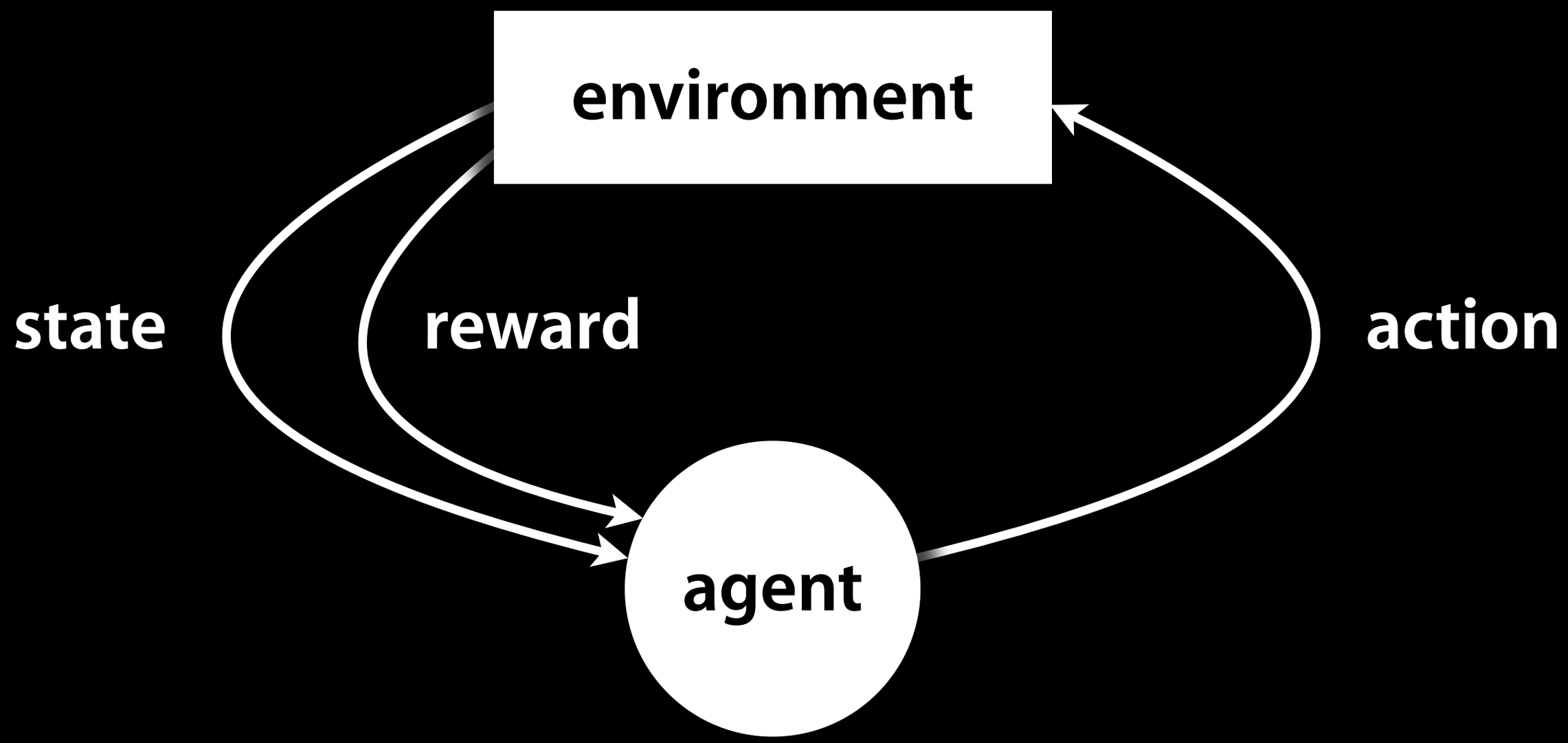
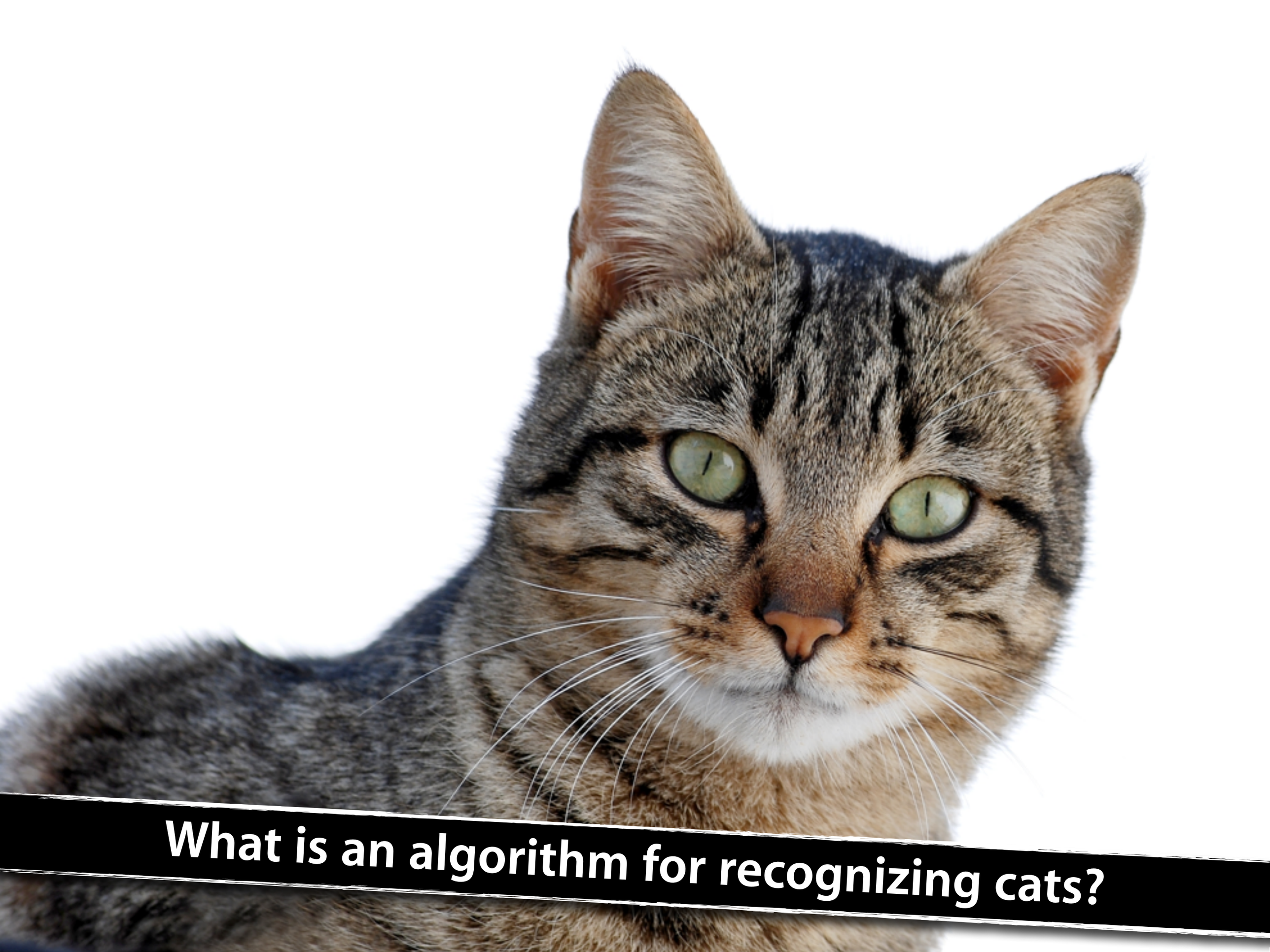

environment
state
reward
action
agent

Image Recognition


What is an algorithm for recognizing cats?

Rule-based classifier?

10,000 - 30,000 object categories

Challenges

viewpoint differences

intraclass variability

deformation

illumination

background clutter

occlusion

image classification

assign an image to a known class based on the objects present in the image

What objects are present in the image?

ImageNet Challenge

ImageNet Challenge
1000 categories, a million images

scene classification

assign an image to a known class based on the scene in the image

What is the scene in the image?

What is the scene in the image?

GOOD FOOD
MADE IN CHINA
TORONTO'S FIRST KOREAN BARBEQUE IS STILL THE HOTTEST GRILL IN TOWN.
KOREAN GRILL HOUSE
KOREAN GRILL HOUSE
TORONTO'S ORIGINAL KOREAN TABLE GRILL RESTAURANT
salon
liberté
ZANZIBAR
ZANZIBAR
NAILS
2nd
HOTELS
ENTERTAINMENT
MURRAY DEMOLITION
Manulife
WEAR THE PANTS
DOCKERS
THE SECOND DAWN OF MAN
FUTURE
Notice something missing?

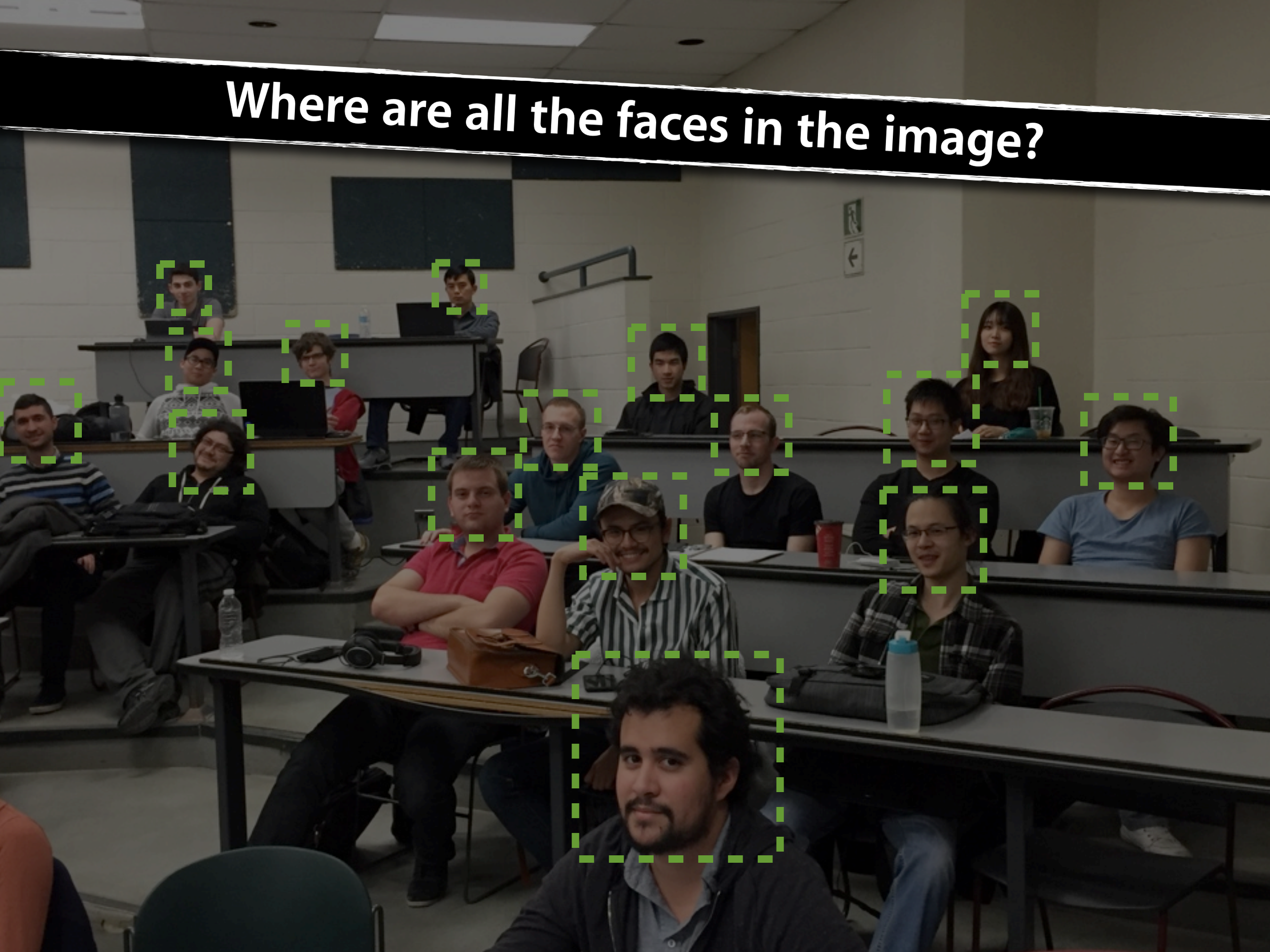
MURRAY
DEMOLITION
FUTURE
ENTERTAINMENT

object detection

locate where all object instances of known classes are present in the image

Where are all the faces in the image?

Where are all the faces in the image?



semantic segmentation

assign each image pixel to a known class

Label each pixel by semantic category

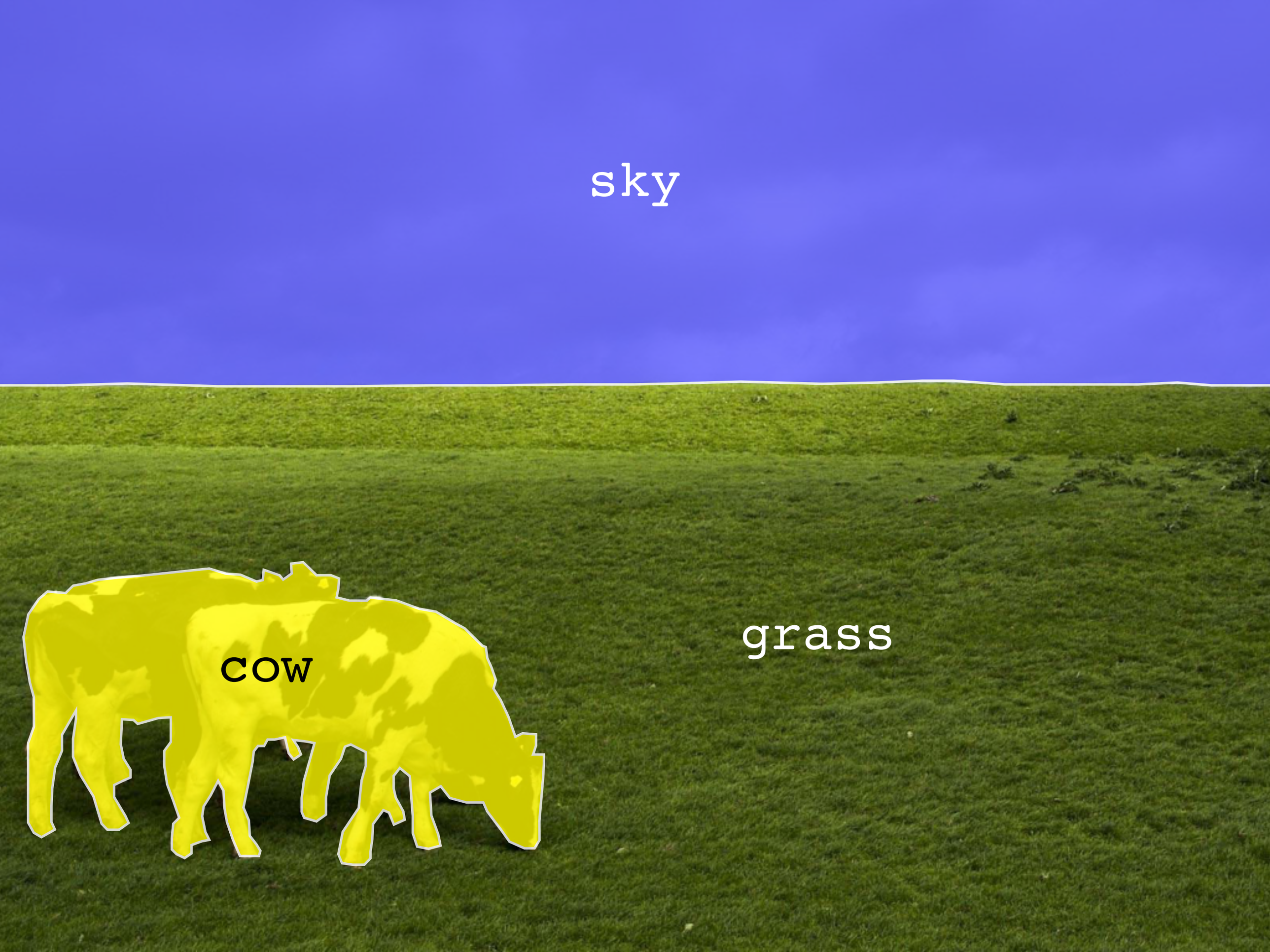

sky
cow
grass

image captioning

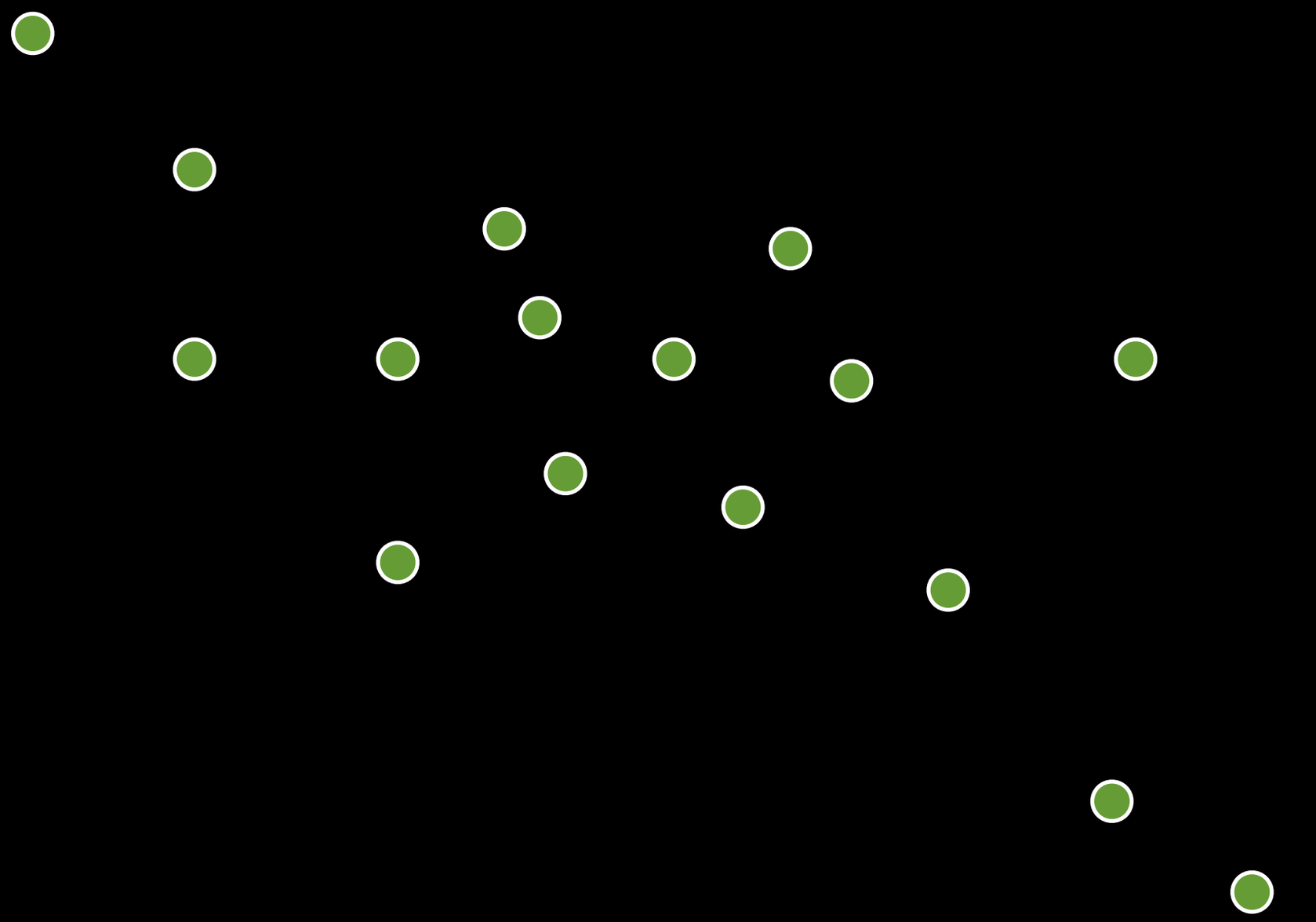
provide a short sentence describing the image

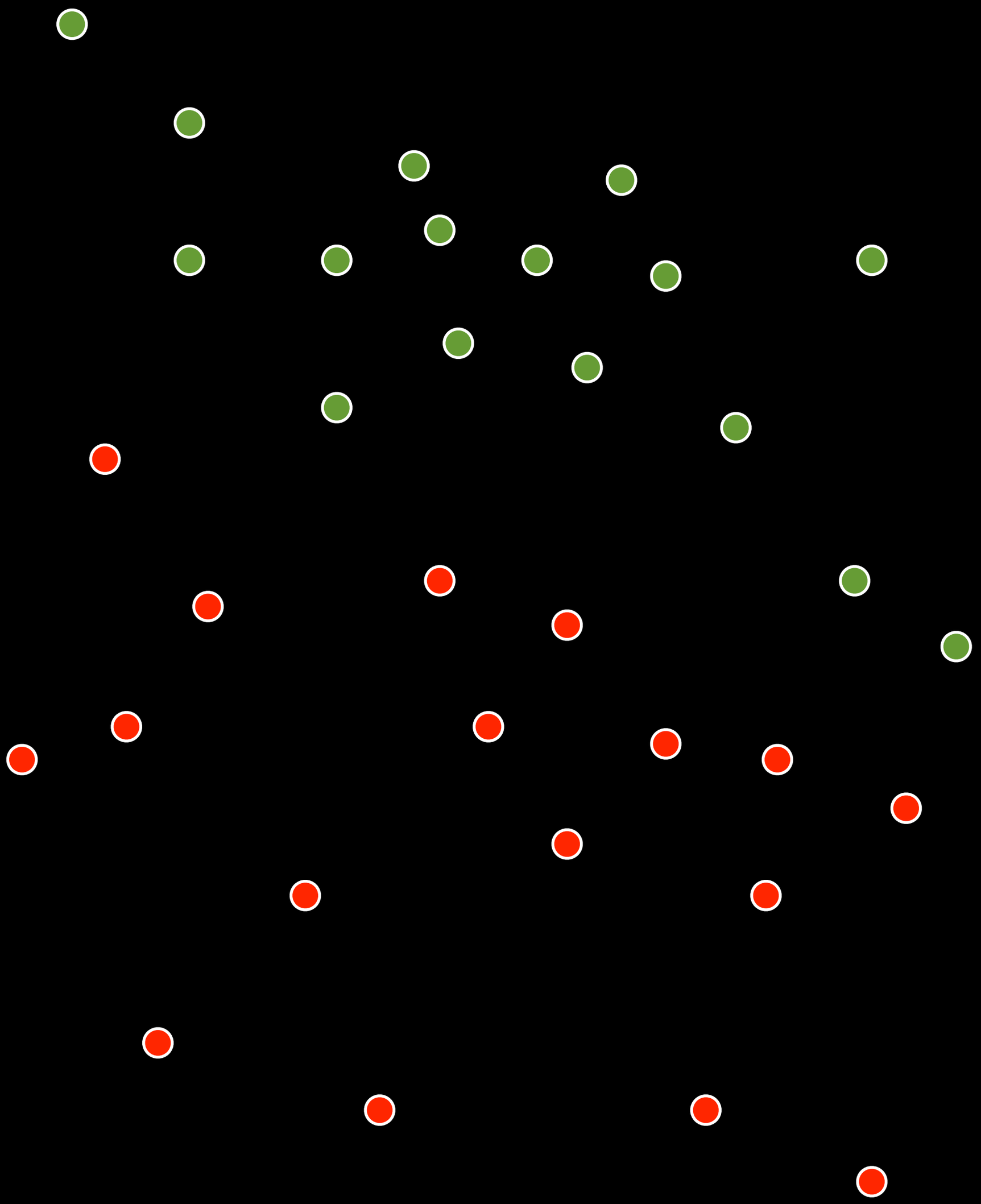
“a group of zebras drinking water”

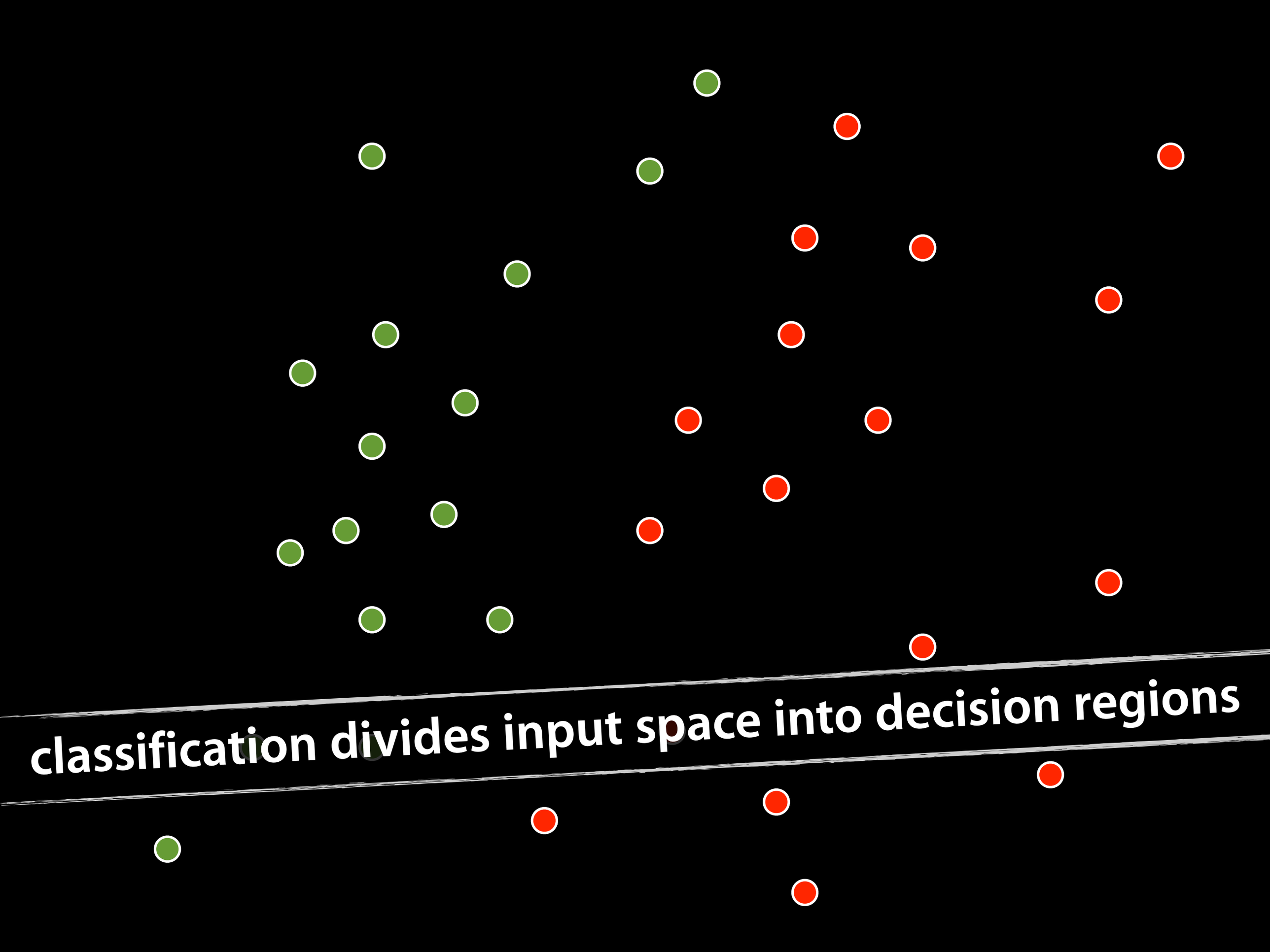
Classification

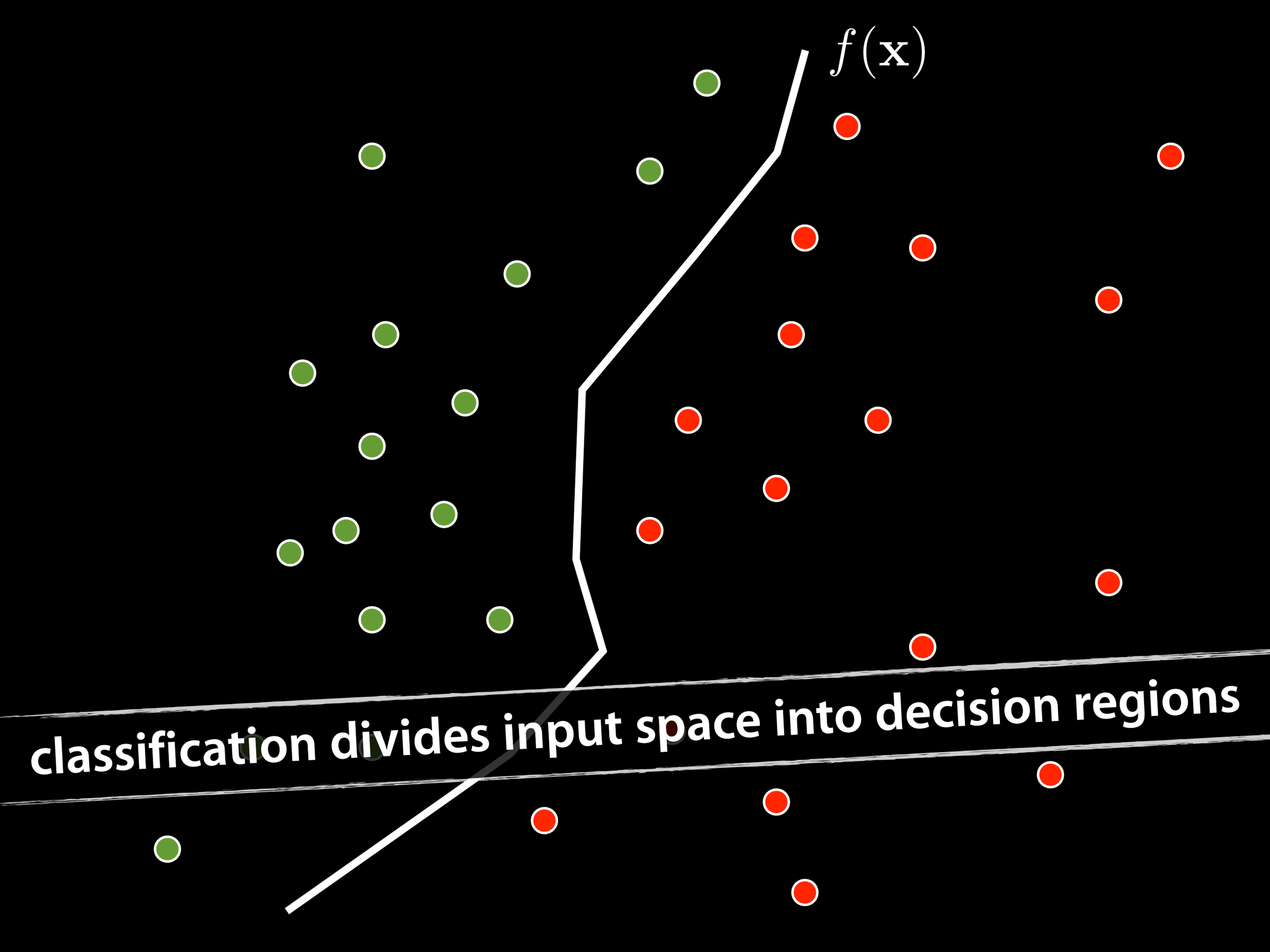
Given a training set of N observations

$$\{(\mathbf{x}_i, y_i)\}_{i=1}^{N} \quad \text{where} \quad \mathbf{x}_i \in \mathbb{R}^D \quad \text{and} \quad y_i \in \{-1, 1\}$$






classification divides input space into decision regions


$f(\mathbf{x})$
classification divides input space into decision regions

Support Vector Machines

Restricted Boltzmann Machine

Naive Bayes

Bayesian Network

Convolutional Neural Network

Randomized Forest

Boosted Decision Tree

AdaBoost

K-Nearest Neighbour

Decision Tree

K-Nearest Neighbour

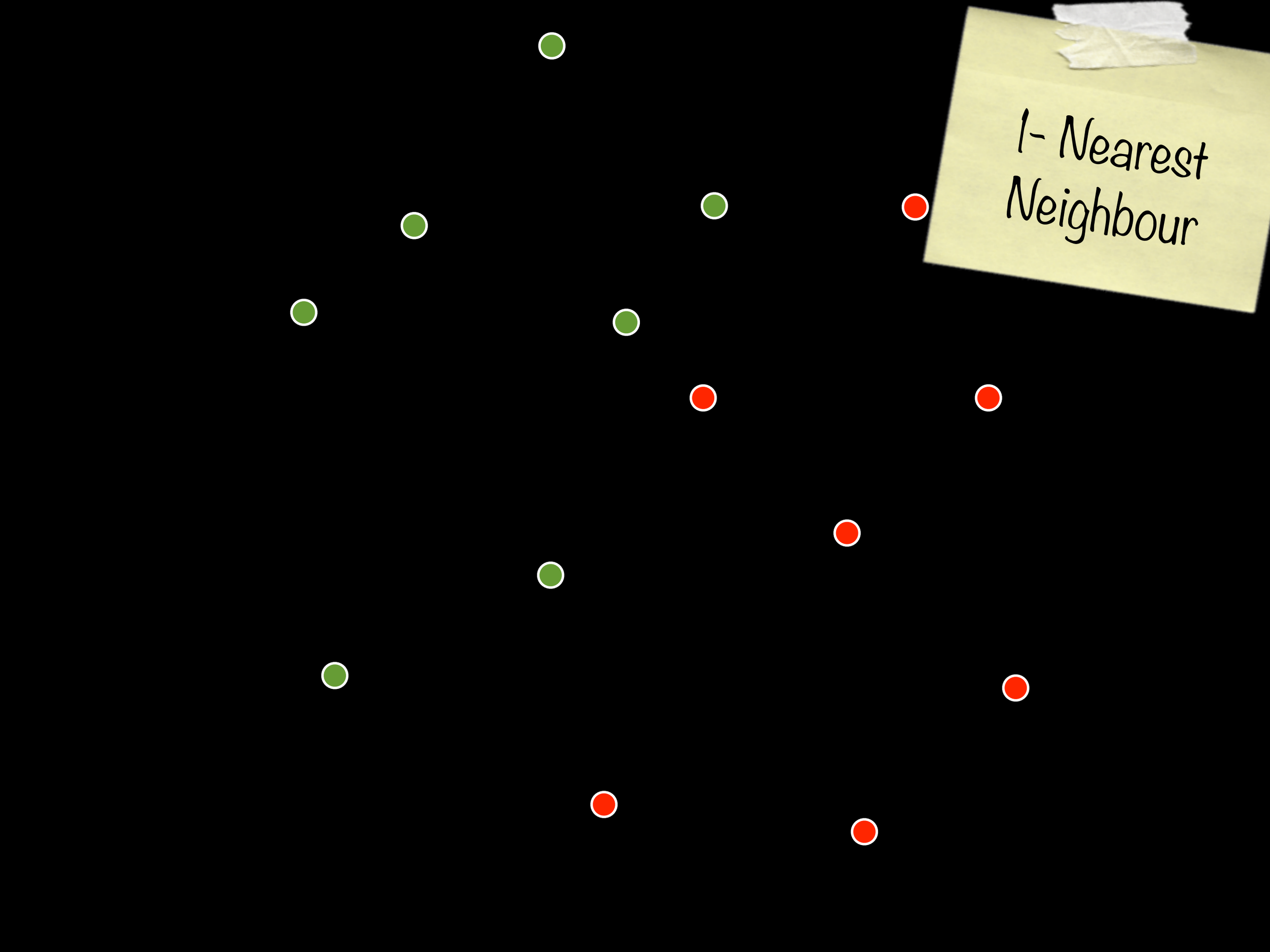
Classification

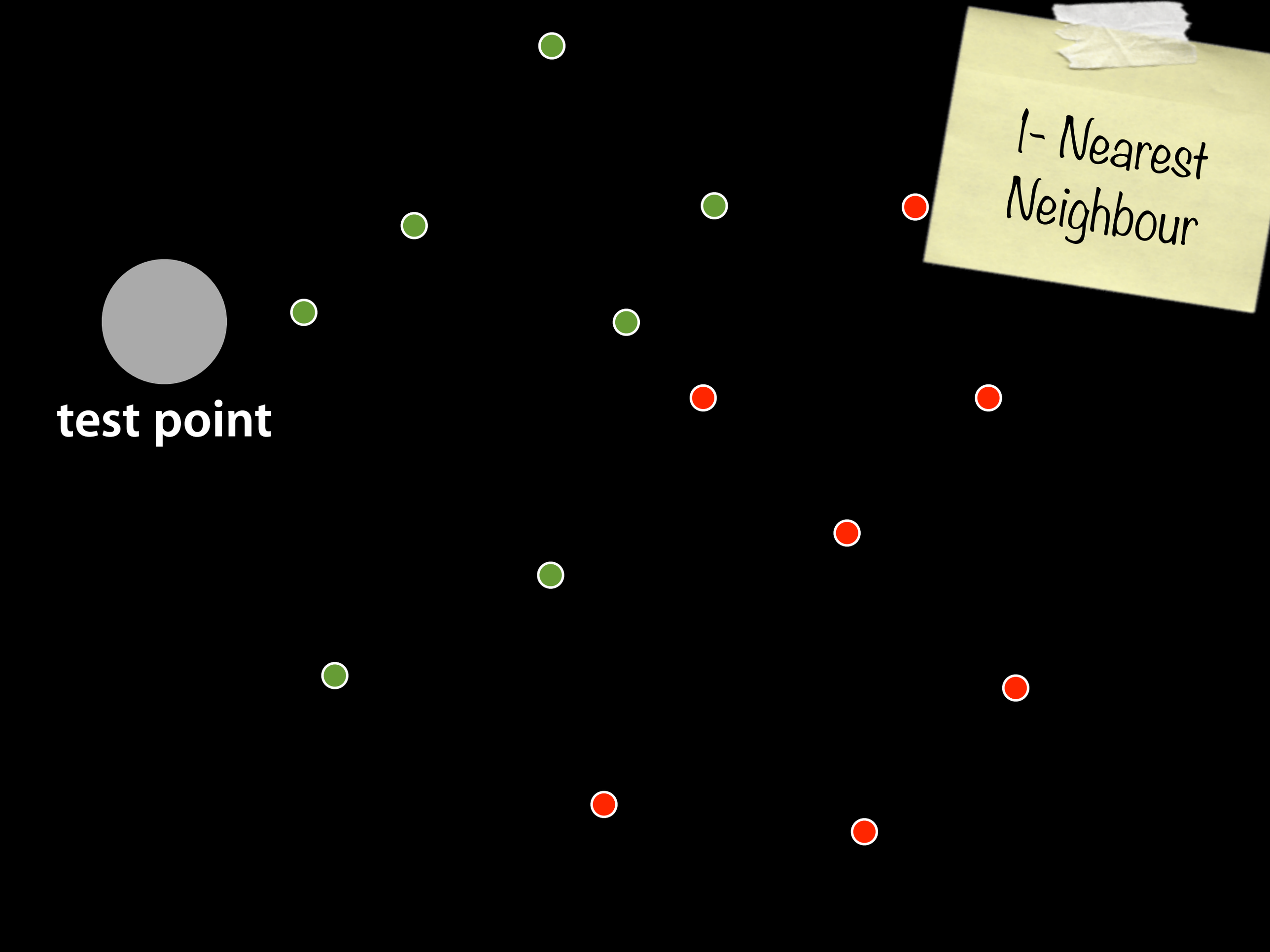
For each test point, find the K nearest samples in the training data

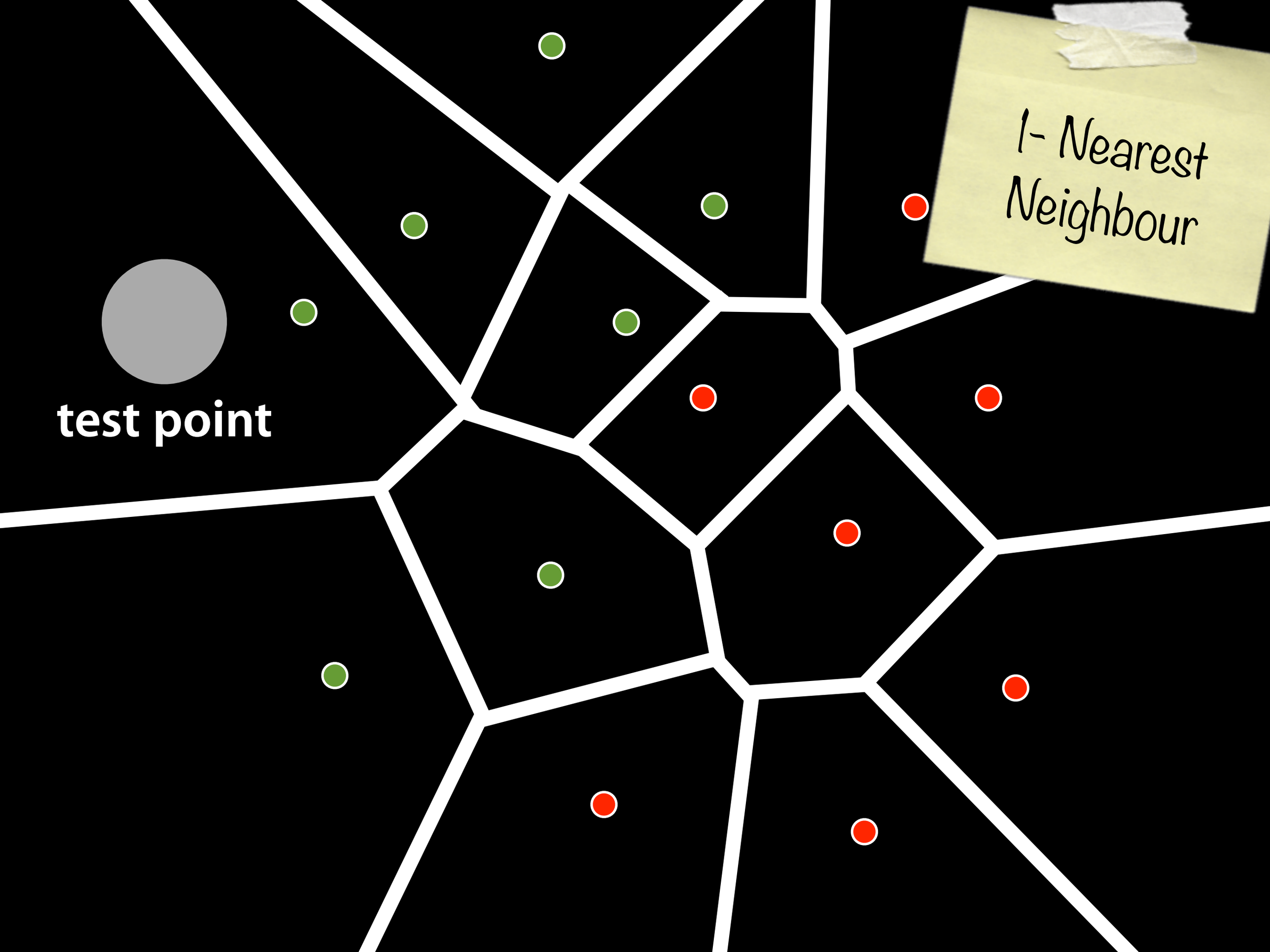
For each test point, find the K nearest samples in the training data

Classify the point according to the majority vote of their class label

1- Nearest Neighbour

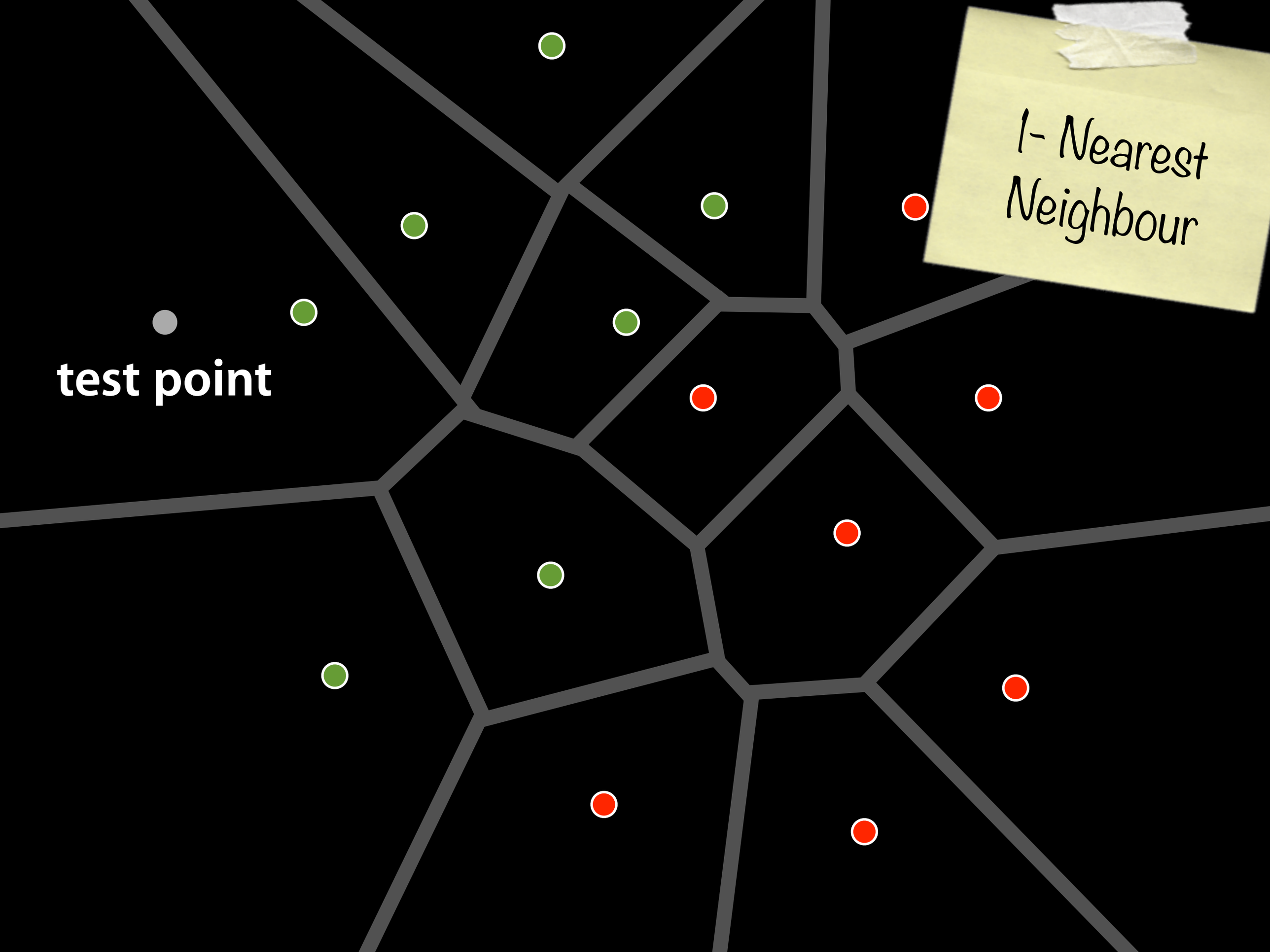

1- Nearest
Neighbour


1- Nearest Neighbour
test point


1- Nearest Neighbour
test point

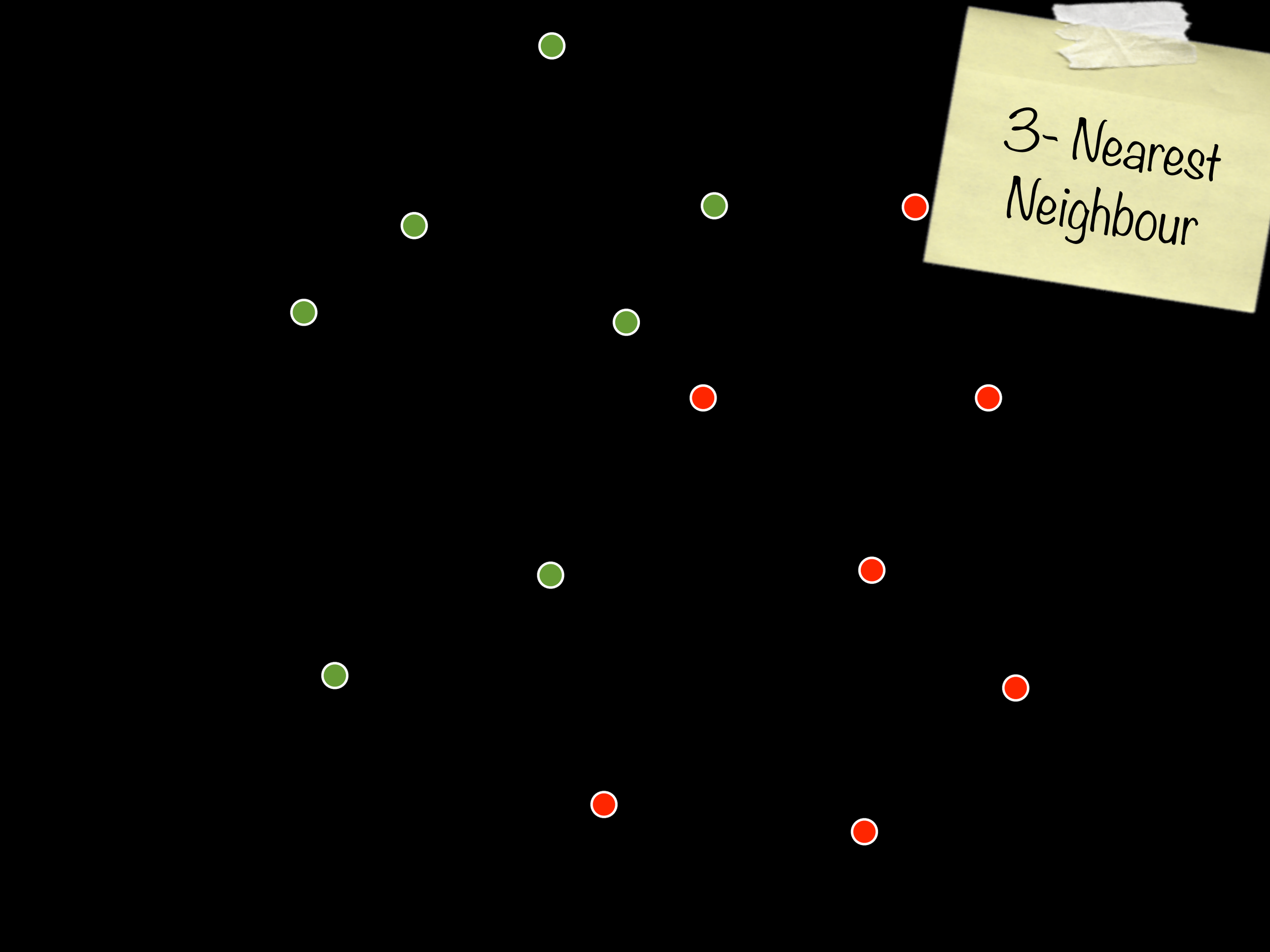
1- Nearest Neighbour
test point

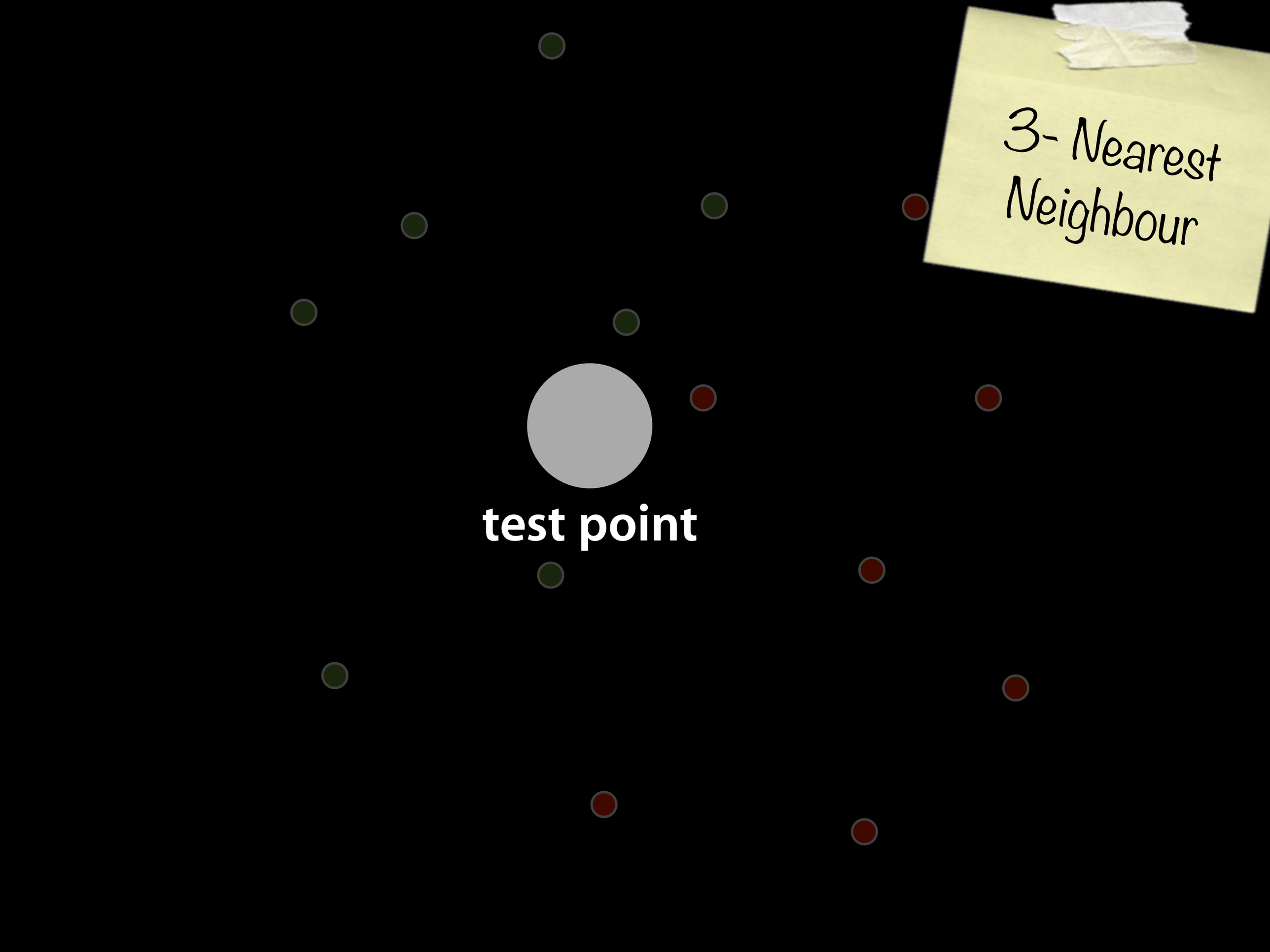
1- Nearest Neighbour
test point

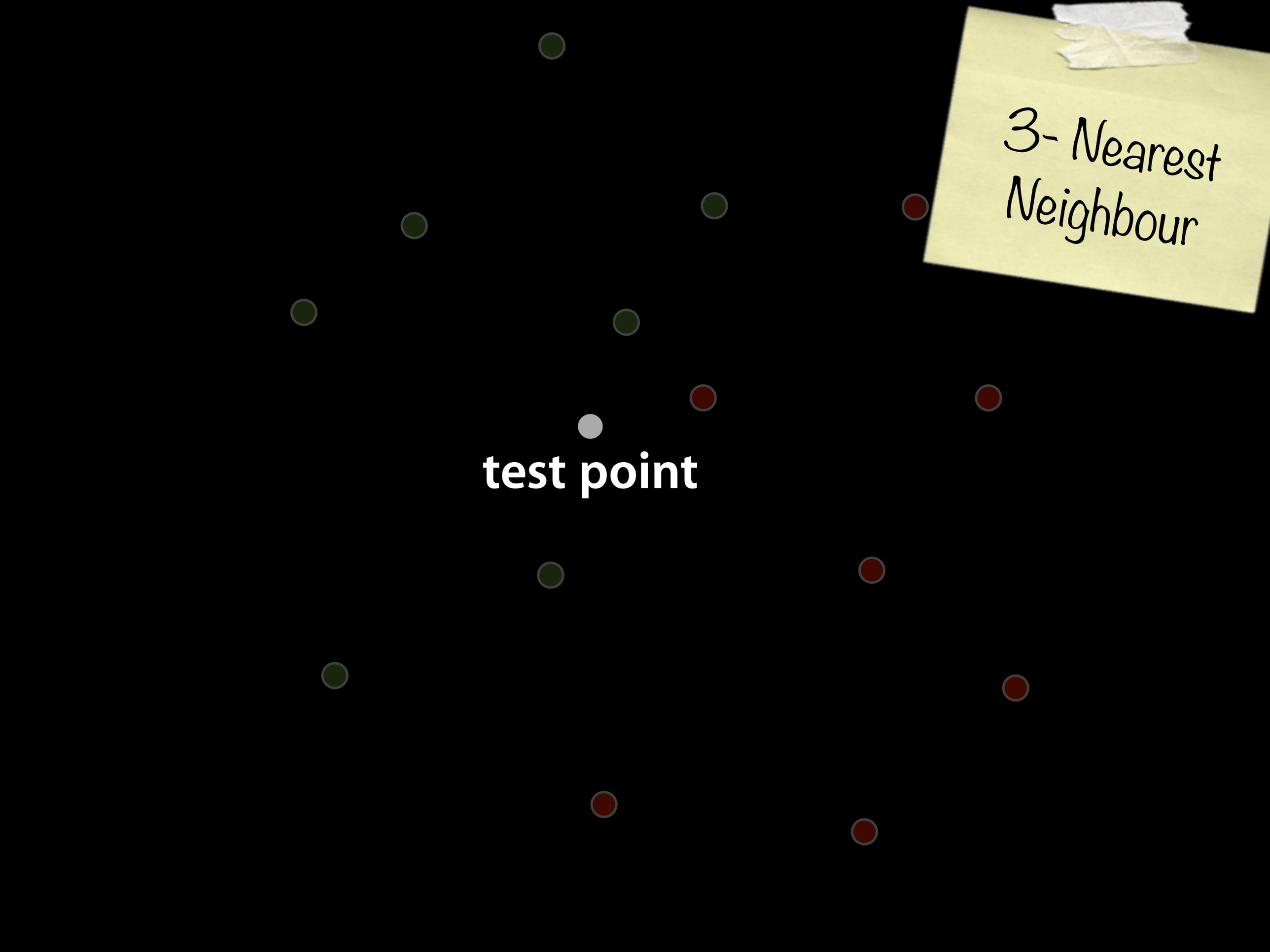

1- Nearest Neighbour
test point

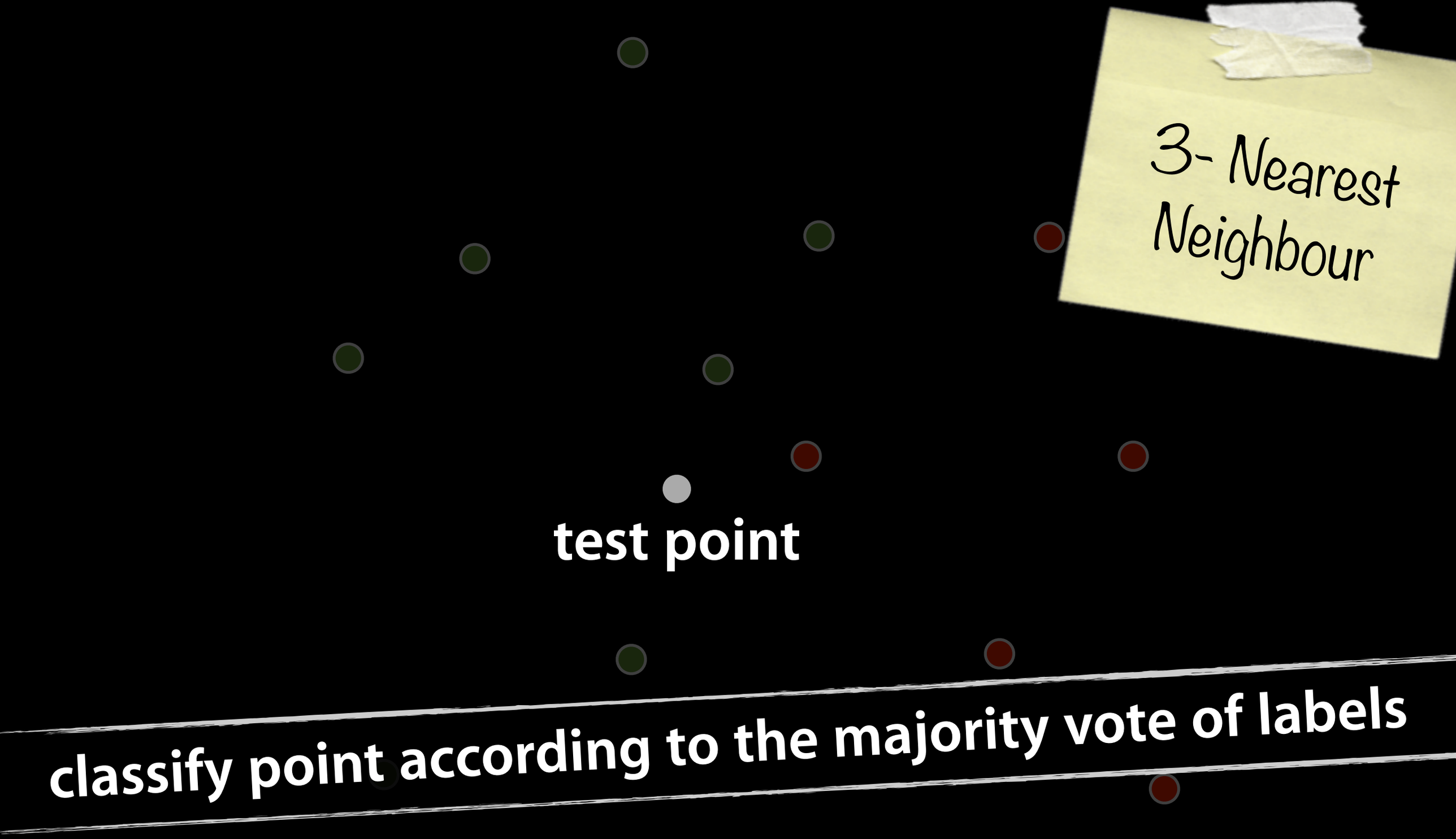
1- Nearest Neighbour
test point

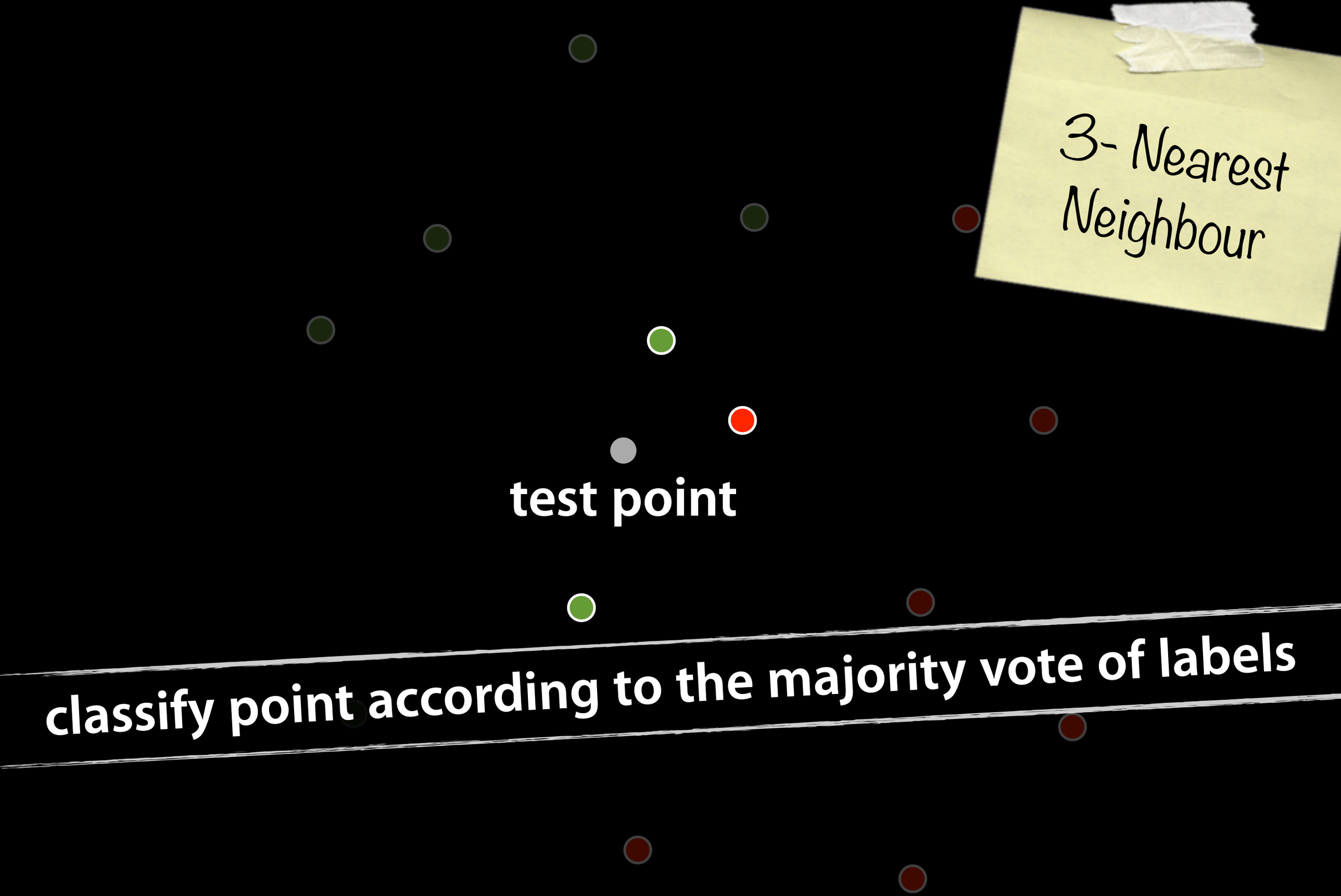
3- Nearest Neighbour


3- Nearest
Neighbour


3- Nearest Neighbour
test point


3- Nearest Neighbour
test point


3- Nearest Neighbour
test point
classify point according to the majority vote of labels

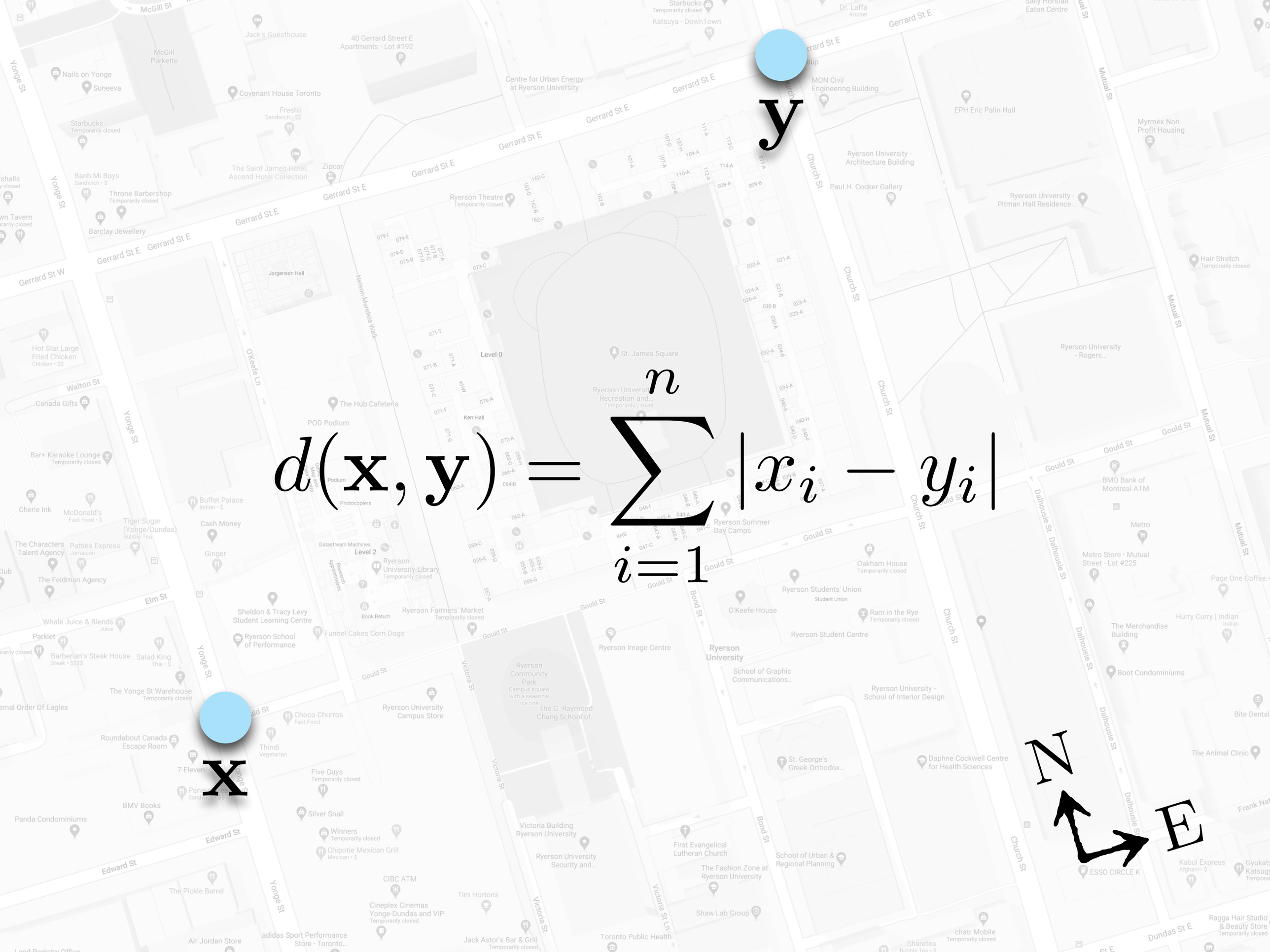

3- Nearest Neighbour
test point
classify point according to the majority vote of labels

Distance Measures

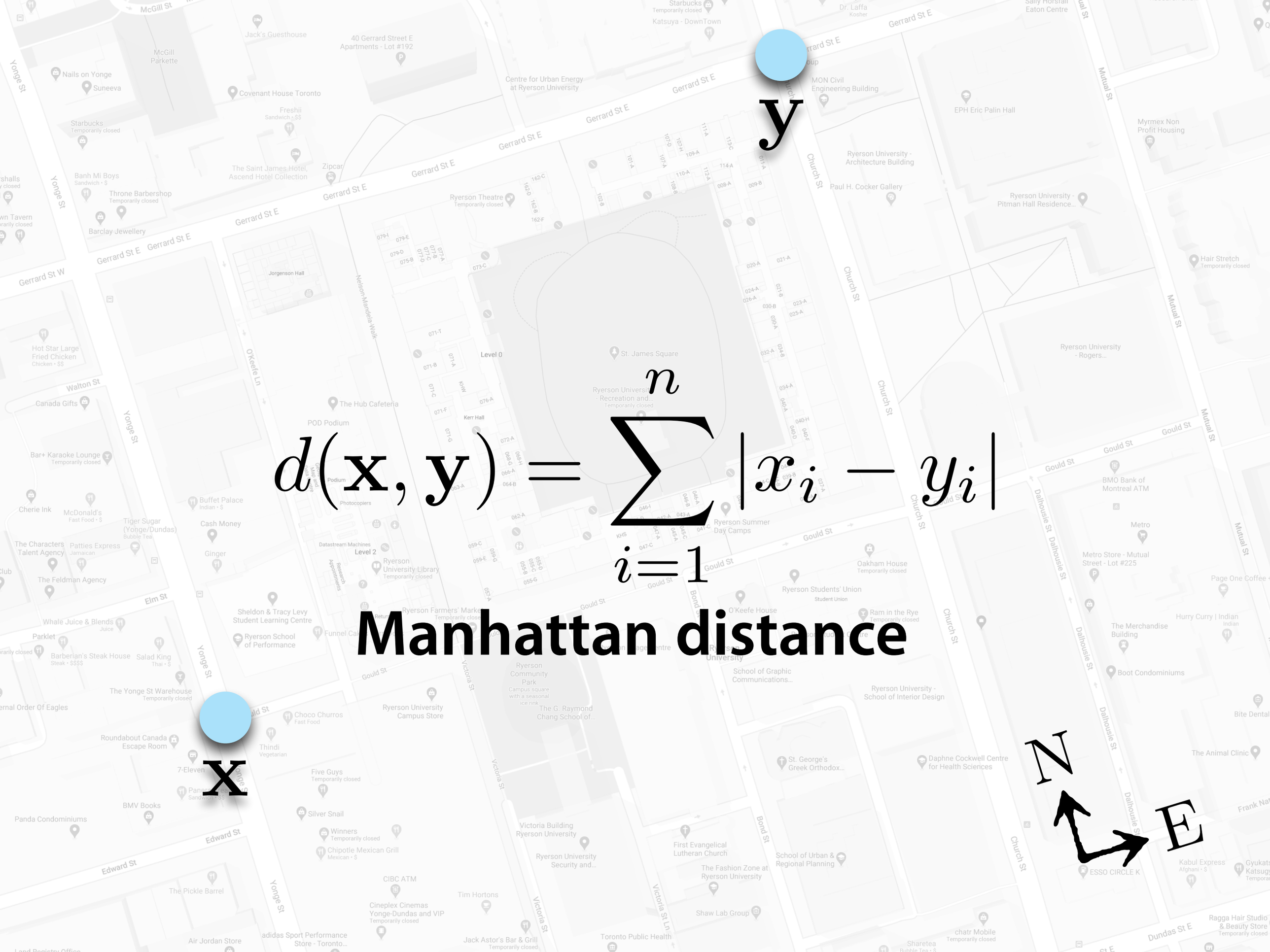
McGill St
Jack's Guesthouse
40 Gerrard Street E Apartments - Lot #192
McGill Parkette
Nails on Yonge
Suneeva
Covenant House Toronto
Freshii
Sandwich • $$
Starbucks
Temporarily closed
Banh Mi Boys
Sandwich • $
Throne Barbershop
Temporarily closed
Barclay Jewellery
The Saint James Hotel, Ascend Hotel Collection
Zipcar
Centre for Urban Energy at Ryerson University
Starbucks
Temporarily closed
Katsuya - DownTown
Dr. Laffa
Kosher
Sally Horsfall Eaton Centre
Brydson Group
MON Civil Engineering Building
EPH Eric Palin Hall
Myrmex Non Profit Housing
Ryerson University - Architecture Building
Paul H. Cocker Gallery
Ryerson University - Pitman Hall Residence...
Hair Stretch
Temporarily closed
Ryerson Theatre
Temporarily closed
Jorgenson Hall
Nelson Mandela Walk
Level 0
St. James Square
Ryerson University - Recreation and...
Temporarily closed
Ryerson University - Rogers...
Hot Star Large Fried Chicken
Chicken • $$
Walton St
Canada Gifts
The Hub Cafeteria
POD Podium
Kerr Hall
Bar+ Karaoke Lounge
Temporarily closed
Cherie Ink
McDonald's
Fast Food • $
Tiger Sugar (Yonge/Dundas)
Bubble Tea
Buffet Palace
Indian • $
Cash Money
Ginger
The Characters Talent Agency
Patties Express
Jamaican
The Feldman Agency
Geospatial, Map and Data Centre
Podium
Photocopiers
Datastream Machines
Level 2
Research Appointments
Ryerson University Library
Temporarily closed
Book Return
Ryerson Summer Day Camps
BMO Bank of Montreal ATM
Metro
Metro Store - Mutual Street - Lot #225
Oakham House
Temporarily closed
Ryerson Students' Union
Student Union
O'Keefe House
Ram in the Rye
Temporarily closed
Ryerson Student Centre
Page One Coffee
Hurry Curry | Indian
Indian
The Merchandise Building
Boot Condominiums
Elm St
Whale Juice & Blends
Juice
Parklet
Barberian's Steak House
Steak • $$$$
Salad King
Thai • $
Sheldon & Tracy Levy Student Learning Centre
Ryerson School of Performance
Funnel Cakes Corn Dogs
Ryerson Farmers' Market
Temporarily closed
Ryerson Image Centre
Ryerson University
School of Graphic Communications...
Ryerson University - School of Interior Design
The Yonge St Warehouse
Temporarily closed
Ryerson Community Park
Campus square with a seasonal ice rink
The G. Raymond Chang School of...
Ryerson University Campus Store
Choco Churros
Fast Food
Thindi
Vegetarian
Roundabout Canada Escape Room
7-Eleven
Five Guys
Temporarily closed
Panera Bread
Sandwich • $$
BMV Books
Panda Condominiums
Silver Snail
Winners
Temporarily closed
Chipotle Mexican Grill
Mexican • $
Edward St
The Pickle Barrel
CIBC ATM
Cineplex Cinemas Yonge-Dundas and VIP
Temporarily closed
Tim Hortons
Victoria Building Ryerson University
Ryerson University Security and...
First Evangelical Lutheran Church
The Fashion Zone at Ryerson University
School of Urban & Regional Planning
St. George's Greek Orthodox...
Daphne Cockwell Centre for Health Sciences
Shaw Lab Group
Toronto Public Health
Jack Astor's Bar & Grill
Temporarily closed
Air Jordan Store
adidas Sport Performance Store - Toronto...
chatr Mobile
Temporarily closed
Sharetea
Bubble Tea • $
ESSO CIRCLE K
Kabul Express
Afghani • $
Ragga Hair Studio & Beauty Store
Temporarily closed
Dundas St E
Frank Nat
Bite Dental
The Animal Clinic
Gerrard St E
Gerrard St W
Yonge St
Church St
Mutual St
Gould St
Dalhousie St
Bond St
Victoria St
Victoria St Ln
O'Keefe Ln
N
E

N
E
X
McGill St
Jack's Guesthouse
40 Gerrard Street E Apartments - Lot #192
McGill Parkette
Nails on Yonge
Suneeva
Covenant House Toronto
Freshii
Sandwich • $$
Starbucks
Temporarily closed
Banh Mi Boys
Sandwich • $
Throne Barbershop
Temporarily closed
Barclay Jewellery
The Saint James Hotel, Ascend Hotel Collection
Zipcar
Centre for Urban Energy at Ryerson University
Starbucks
Temporarily closed
Katsuya - DownTown
Dr. Laffa
Kosher
Sally Horsfall Eaton Centre
Brydson Group
MON Civil Engineering Building
EPH Eric Palin Hall
Myrmex Non Profit Housing
Ryerson University - Architecture Building
Paul H. Cocker Gallery
Ryerson University - Pitman Hall Residence...
Hair Stretch
Temporarily closed
Ryerson Theatre
Temporarily closed
Jorgenson Hall
Nelson Mandela Walk
Level 0
St. James Square
Ryerson University - Recreation and...
Temporarily closed
Ryerson Summer Day Camps
Ryerson University - Rogers...
Hot Star Large Fried Chicken
Chicken • $$
Walton St
Canada Gifts
O'Keefe Ln
The Hub Cafeteria
POD Podium
Kerr Hall
Bar+ Karaoke Lounge
Temporarily closed
Cherie Ink
McDonald's
Fast Food • $
Tiger Sugar (Yonge/Dundas)
Bubble Tea
Buffet Palace
Indian • $
Cash Money
Ginger
The Characters Talent Agency
Patties Express
Jamaican
The Feldman Agency
Podium
Photocopiers
Datastream Machines
Level 2
Ryerson University Library
Temporarily closed
Book Return
Ryerson Farmers' Market
Temporarily closed
Elm St
Sheldon & Tracy Levy Student Learning Centre
Ryerson School of Performance
Funnel Cakes Corn Dogs
Whale Juice & Blends
Juice
Parklet
Barberian's Steak House
Steak • $$$$
Salad King
Thai • $
The Yonge St Warehouse
Temporarily closed
Roundabout Canada Escape Room
Choco Churros
Fast Food
Thindi
Vegetarian
7-Eleven
Ryerson University Campus Store
Ryerson Community Park
Campus square with a seasonal ice rink
The G. Raymond Chang School of...
Ryerson Image Centre
Ryerson University
Oakham House
Temporarily closed
Ryerson Students' Union
Student Union
O'Keefe House
Ram in the Rye
Temporarily closed
Ryerson Student Centre
School of Graphic Communications...
Ryerson University - School of Interior Design
St. George's Greek Orthodox...
Daphne Cockwell Centre for Health Sciences
BMO Bank of Montreal ATM
Metro
Metro Store - Mutual Street - Lot #225
Page One Coffee
The Merchandise Building
Hurry Curry | Indian
Indian
Boot Condominiums
Bite Dental
The Animal Clinic
Frank Nat
Kabul Express
Afghani • $
Gyukats Katsugy
Temporar
ESSO CIRCLE K
Ragga Hair Studio & Beauty Store
Temporarily closed
Dundas St E
Panda Condominiums
BMV Books
Five Guys
Temporarily closed
Silver Snail
Winners
Temporarily closed
Chipotle Mexican Grill
Mexican • $
Edward St
The Pickle Barrel
CIBC ATM
Cineplex Cinemas Yonge-Dundas and VIP
Temporarily closed
Tim Hortons
Air Jordan Store
adidas Sport Performance Store - Toronto...
Jack Astor's Bar & Grill
Temporarily closed
Victoria Building Ryerson University
Ryerson University Security and...
First Evangelical Lutheran Church
The Fashion Zone at Ryerson University
School of Urban & Regional Planning
Shaw Lab Group
Toronto Public Health
chatr Mobile
Temporarily closed
Sharetea
Bubble Tea • $
Yonge St
Gerrard St E
Gerrard St W
Church St
Mutual St
Bond St
Victoria St
Victoria St Ln
Dalhousie St
Gould St

y
X
N
E
McGill St
Jack's Guesthouse
40 Gerrard Street E Apartments - Lot #192
McGill Parkette
Nails on Yonge
Suneeva
Covenant House Toronto
Freshii
Sandwich • $$
Starbucks
Temporarily closed
Centre for Urban Energy at Ryerson University
Starbucks
Temporarily closed
Katsuya - DownTown
Dr. Laffa
Kosher
Sally Horsfall Eaton Centre
MON Civil Engineering Building
EPH Eric Palin Hall
Myrmex Non Profit Housing
Ryerson University - Architecture Building
Paul H. Cocker Gallery
Ryerson University - Pitman Hall Residence...
Hair Stretch
Temporarily closed
Banh Mi Boys
Sandwich • $
Throne Barbershop
Temporarily closed
Barclay Jewellery
The Saint James Hotel, Ascend Hotel Collection
Zipcar
Ryerson Theatre
Temporarily closed
Jorgenson Hall
Nelson Mandela Walk
Level 0
St. James Square
Ryerson University - Recreation and...
Temporarily closed
Ryerson Summer Day Camps
Ryerson University - Rogers...
Hot Star Large Fried Chicken
Chicken • $$
Walton St
Canada Gifts
O'Keefe Ln
The Hub Cafeteria
POD Podium
Kerr Hall
Bar+ Karaoke Lounge
Temporarily closed
Cherie Ink
McDonald's
Fast Food • $
Tiger Sugar (Yonge/Dundas)
Bubble Tea
Buffet Palace
Indian • $
Cash Money
Ginger
Geospatial, Map and Data Centre
Podium
Photocopiers
Datastream Machines
Level 2
Research Appointments
Ryerson University Library
Temporarily closed
Book Return
The Characters Talent Agency
Patties Express
Jamaican
The Feldman Agency
Elm St
Sheldon & Tracy Levy Student Learning Centre
Ryerson School of Performance
Ryerson Farmers' Market
Temporarily closed
Funnel Cakes Corn Dogs
Whale Juice & Blends
Juice
Parklet
Barberian's Steak House
Steak • $$$$
Salad King
Thai • $
The Yonge St Warehouse
Temporarily closed
Ryerson Image Centre
Ryerson University
O'Keefe House
Ryerson Students' Union
Student Union
Oakham House
Temporarily closed
Ram in the Rye
Temporarily closed
Ryerson Student Centre
School of Graphic Communications...
Ryerson University - School of Interior Design
Ryerson Community Park
Campus square with a seasonal ice rink
The G. Raymond Chang School of...
Ryerson University Campus Store
Choco Churros
Fast Food
Thindi
Vegetarian
Roundabout Canada Escape Room
7-Eleven
Five Guys
Temporarily closed
Silver Snail
Winners
Temporarily closed
Chipotle Mexican Grill
Mexican • $
BMV Books
Panda Condominiums
Edward St
The Pickle Barrel
CIBC ATM
Cineplex Cinemas Yonge-Dundas and VIP
Temporarily closed
Tim Hortons
Air Jordan Store
adidas Sport Performance Store - Toronto...
Jack Astor's Bar & Grill
Temporarily closed
Toronto Public Health
Victoria Building Ryerson University
Ryerson University Security and...
First Evangelical Lutheran Church
The Fashion Zone at Ryerson University
Shaw Lab Group
St. George's Greek Orthodox...
Daphne Cockwell Centre for Health Sciences
School of Urban & Regional Planning
chatr Mobile
Temporarily closed
Sharetea
Bubble Tea • $
Gould St
BMO Bank of Montreal ATM
Metro
Metro Store - Mutual Street - Lot #225
Page One Coffee
The Merchandise Building
Hurry Curry | Indian
Indian
Boot Condominiums
Bite Dental
The Animal Clinic
ESSO CIRCLE K
Kabul Express
Afghani • $
Ragga Hair Studio & Beauty Store
Temporarily closed
Dundas St E
Frank Nat
Church St
Yonge St
Mutual St
Dalhousie St
Bond St
Victoria St
Victoria St Ln
Gerrard St E
Gerrard St W

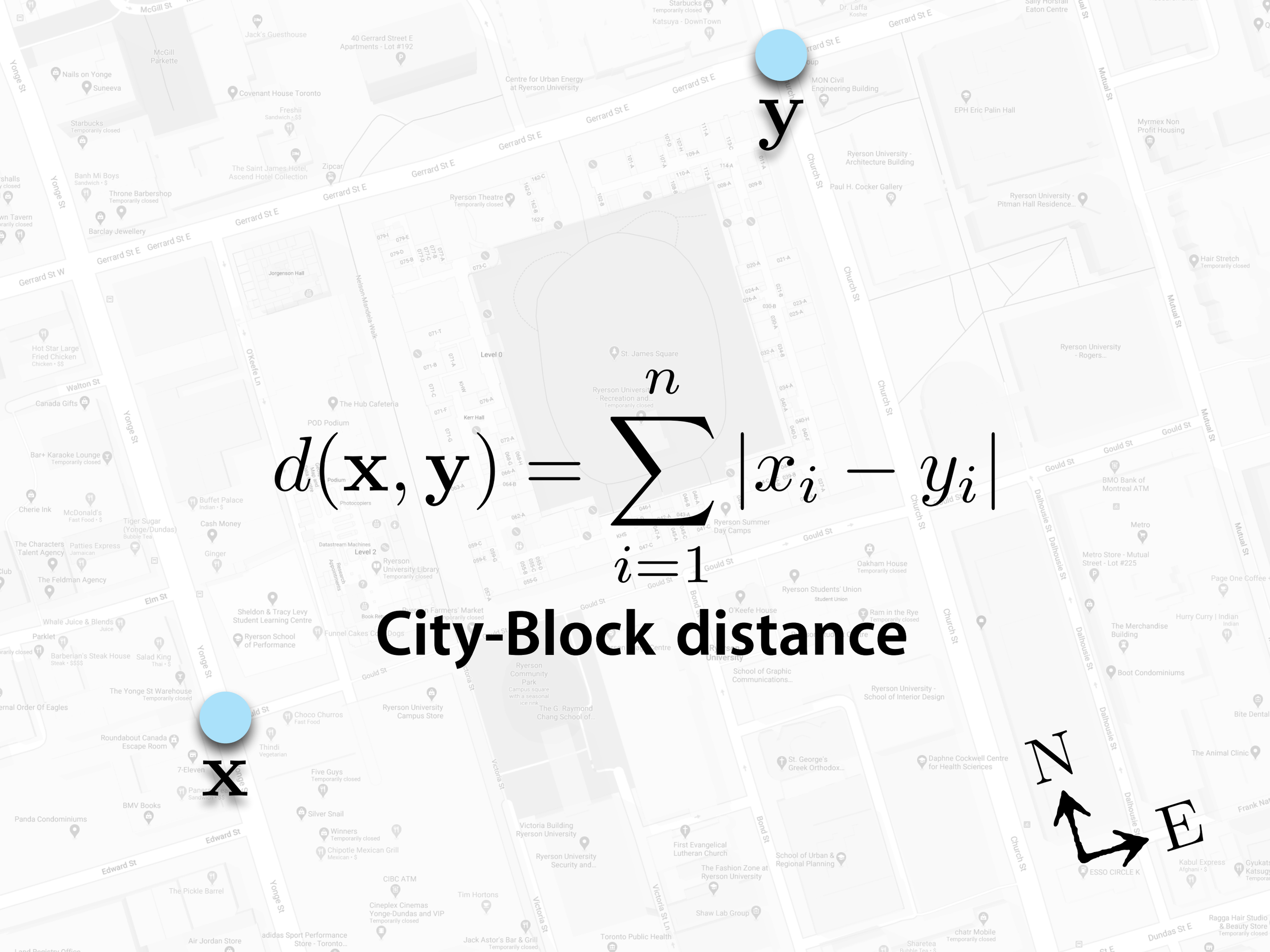


$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$



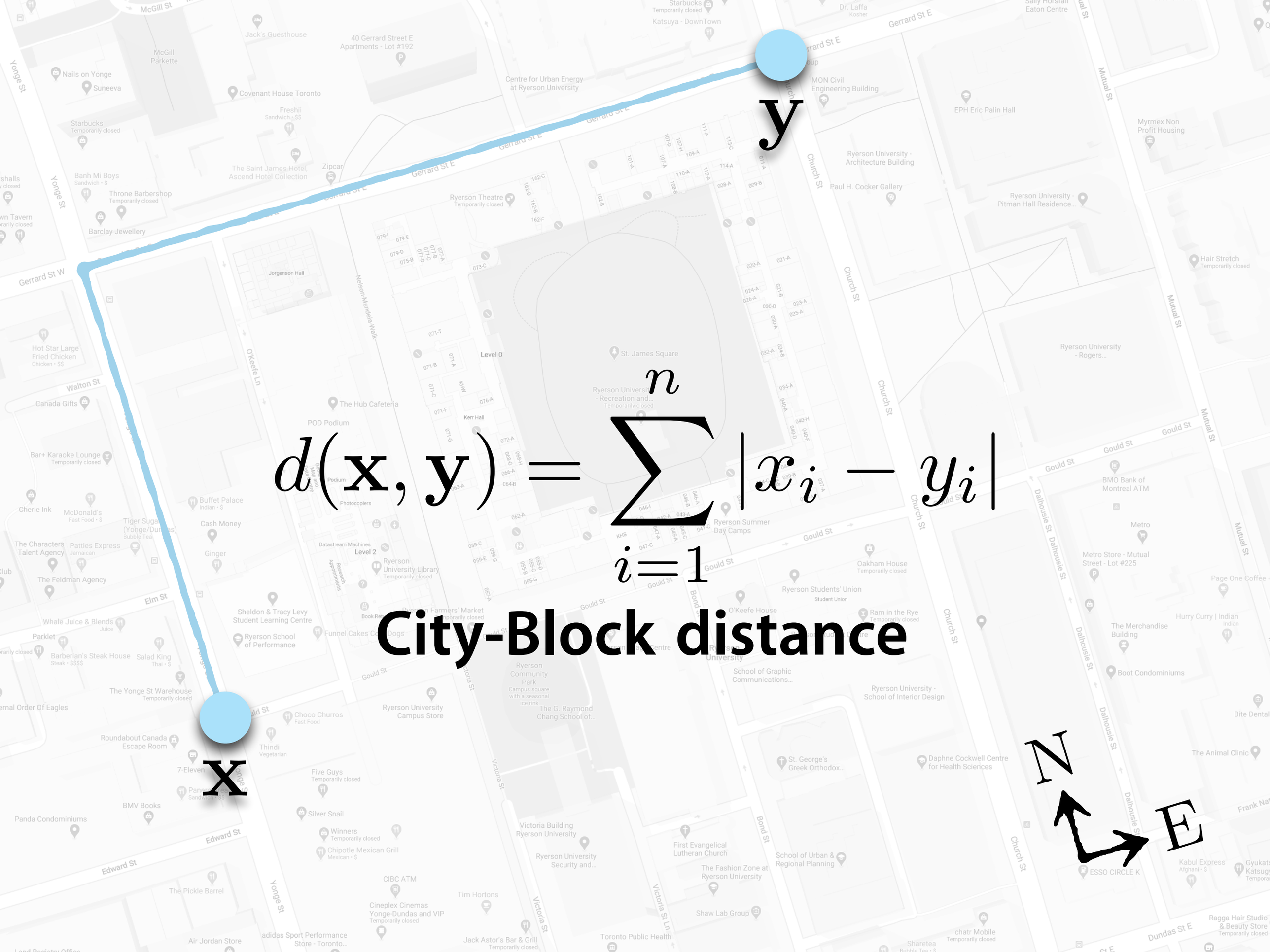
$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

Manhattan distance



$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

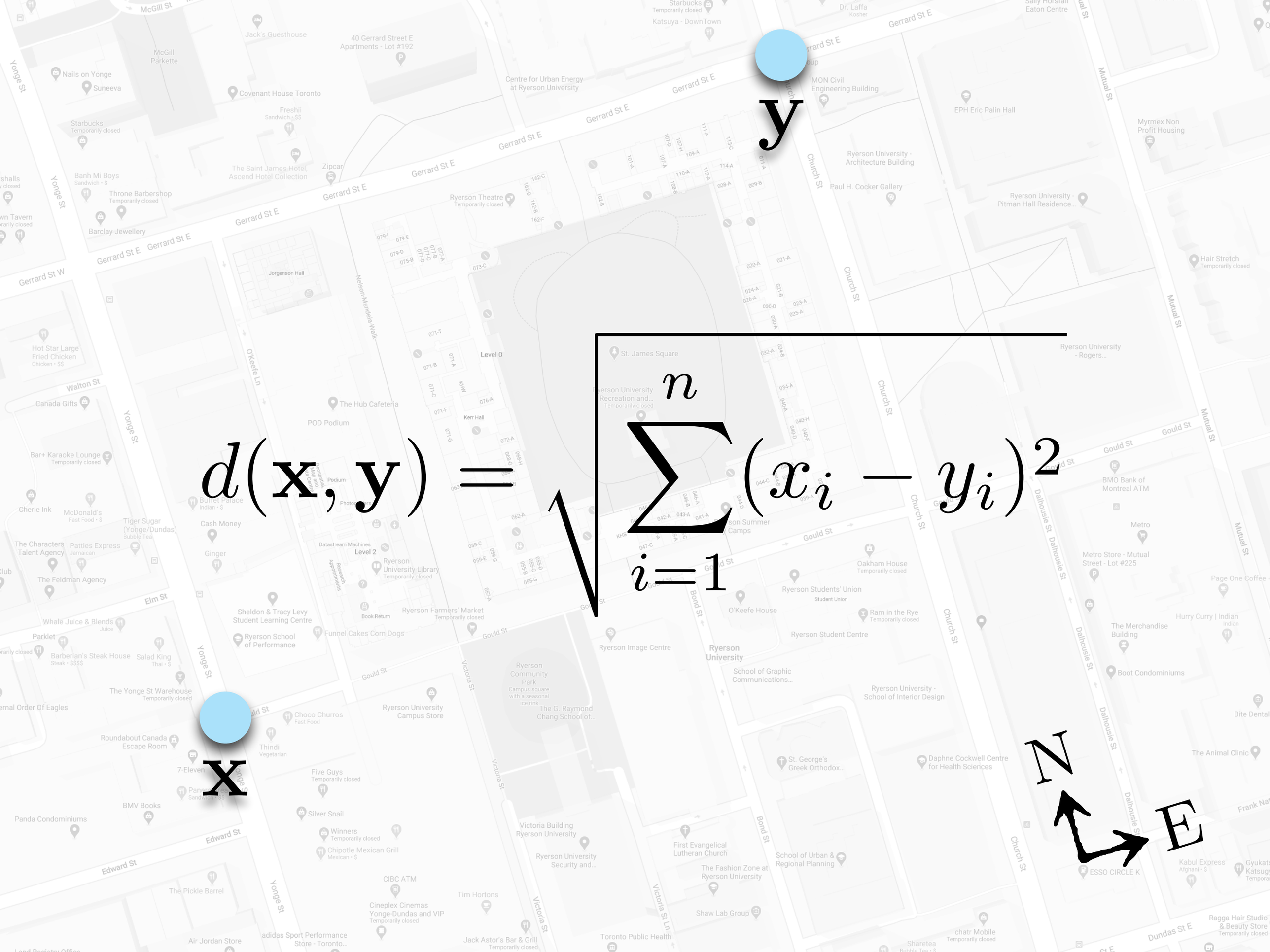
City-Block distance



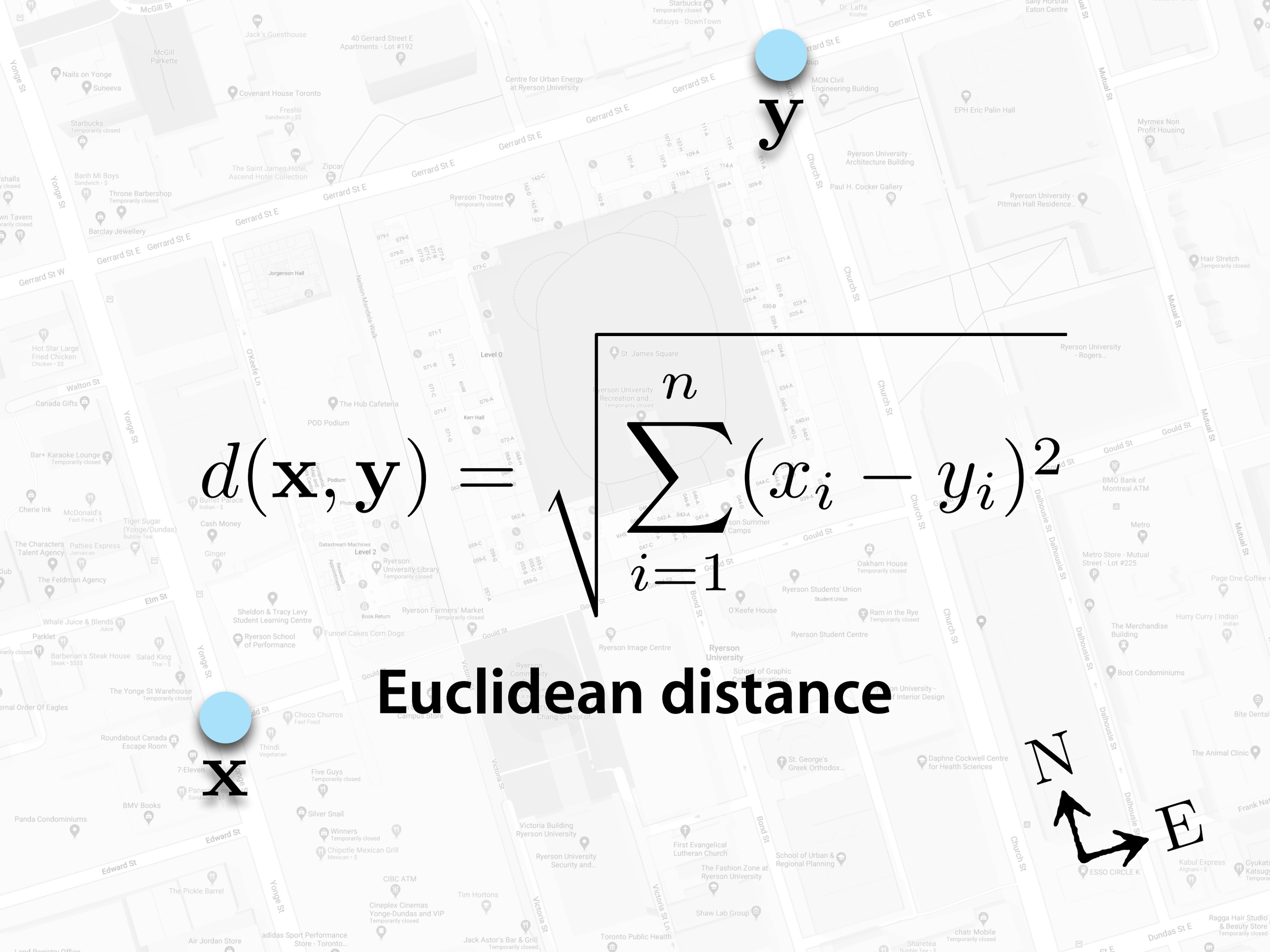
$$d(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{n} |x_i - y_i|$$

City-Block distance

y
X
N
E
McGill St
Gerrard St E
Gerrard St W
Yonge St
Church St
Mutual St
Dalhousie St
Bond St
Victoria St
Victoria St Ln
Gould St
Elm St
Walton St
Edward St
Dundas St E
O'Keefe Ln
Nelson Mandela Walk
Jack's Guesthouse
40 Gerrard Street E Apartments - Lot #192
McGill Parkette
Nails on Yonge
Suneeva
Covenant House Toronto
Freshii
Starbucks
Banh Mi Boys
Throne Barbershop
Barclay Jewellery
The Saint James Hotel, Ascend Hotel Collection
Zipcar
Centre for Urban Energy at Ryerson University
Ryerson Theatre
Jorgenson Hall
Hot Star Large Fried Chicken
Canada Gifts
Bar+ Karaoke Lounge
Cherie Ink
McDonald's
Tiger Sugar (Yonge/Dundas)
The Characters Talent Agency
Patties Express
The Feldman Agency
Buffet Palace
Cash Money
Ginger
The Hub Cafeteria
POD Podium
Photocopiers
Datastream Machines
Level 0
Level 2
Kerr Hall
St. James Square
Ryerson University - Recreation and...
Ryerson University Library
Ryerson Farmers' Market
Book Return
Sheldon & Tracy Levy Student Learning Centre
Ryerson School of Performance
Funnel Cakes Corn Dogs
Whale Juice & Blends
Parklet
Barberian's Steak House
Salad King
The Yonge St Warehouse
Choco Churros
Ryerson University Campus Store
Ryerson Community Park
The G. Raymond Chang School of...
Ryerson Image Centre
Ryerson University
School of Graphic Communications...
Ryerson University - School of Interior Design
Ryerson Summer Day Camps
Oakham House
Ryerson Students' Union
O'Keefe House
Ram in the Rye
Ryerson Student Centre
Dr. Laffa
Katsuya - DownTown
Starbucks
Sally Horsfall Eaton Centre
MON Civil Engineering Building
EPH Eric Palin Hall
Ryerson University - Architecture Building
Paul H. Cocker Gallery
Ryerson University - Pitman Hall Residence...
Myrmex Non Profit Housing
Hair Stretch
Ryerson University - Rogers...
BMO Bank of Montreal ATM
Metro
Metro Store - Mutual Street - Lot #225
Page One Coffee
Hurry Curry | Indian
The Merchandise Building
Boot Condominiums
Bite Dental
The Animal Clinic
Daphne Cockwell Centre for Health Sciences
St. George's Greek Orthodox...
Roundabout Canada Escape Room
7-Eleven
Thindi
Five Guys
BMV Books
Panda Condominiums
Silver Snail
Winners
Chipotle Mexican Grill
The Pickle Barrel
CIBC ATM
Cineplex Cinemas Yonge-Dundas and VIP
Tim Hortons
Jack Astor's Bar & Grill
Air Jordan Store
adidas Sport Performance Store - Toronto...
Victoria Building Ryerson University
Ryerson University Security and...
First Evangelical Lutheran Church
The Fashion Zone at Ryerson University
Shaw Lab Group
Toronto Public Health
School of Urban & Regional Planning
chatr Mobile
Sharetea
ESSO CIRCLE K
Kabul Express
Ragga Hair Studio & Beauty Store

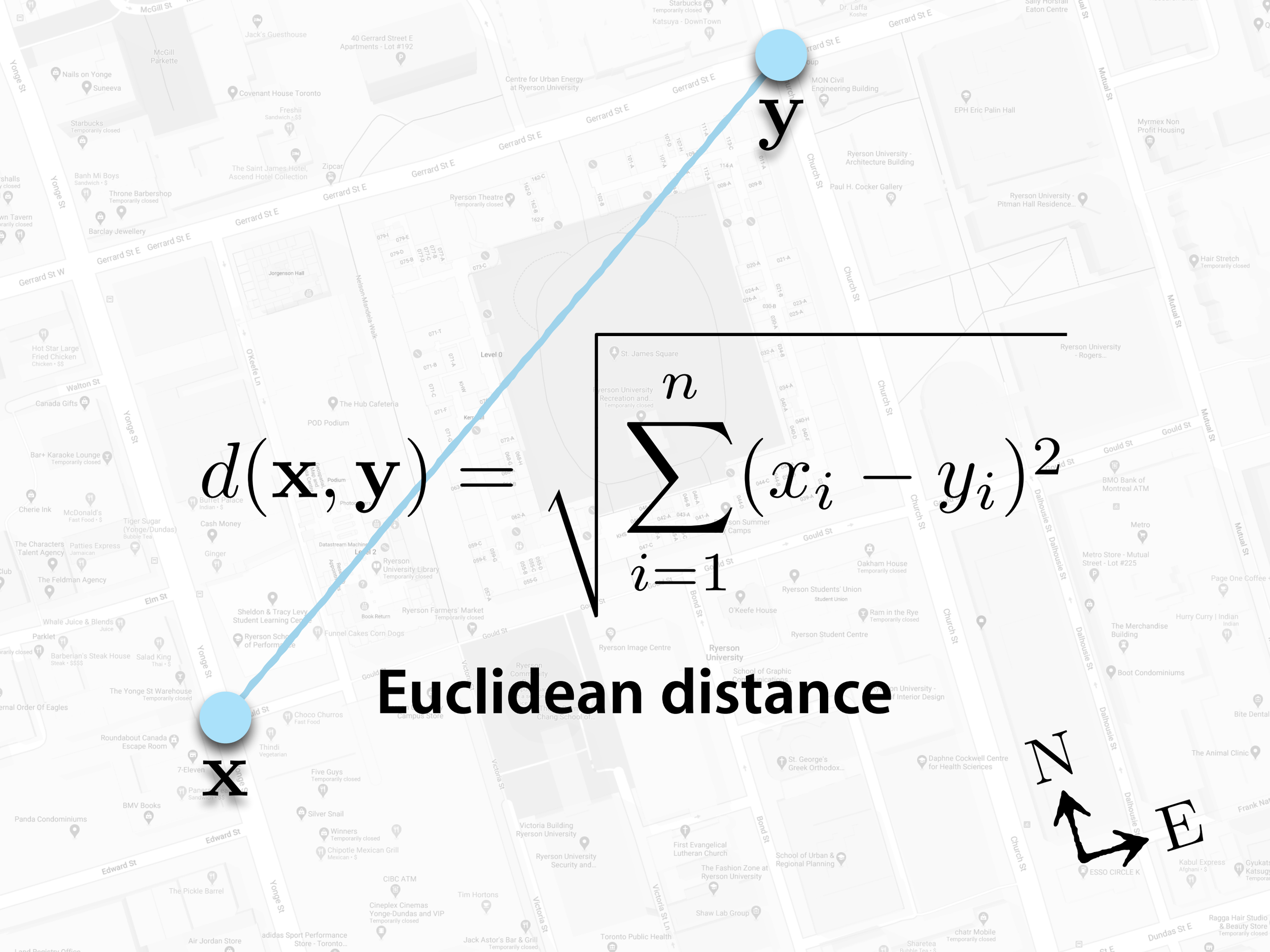


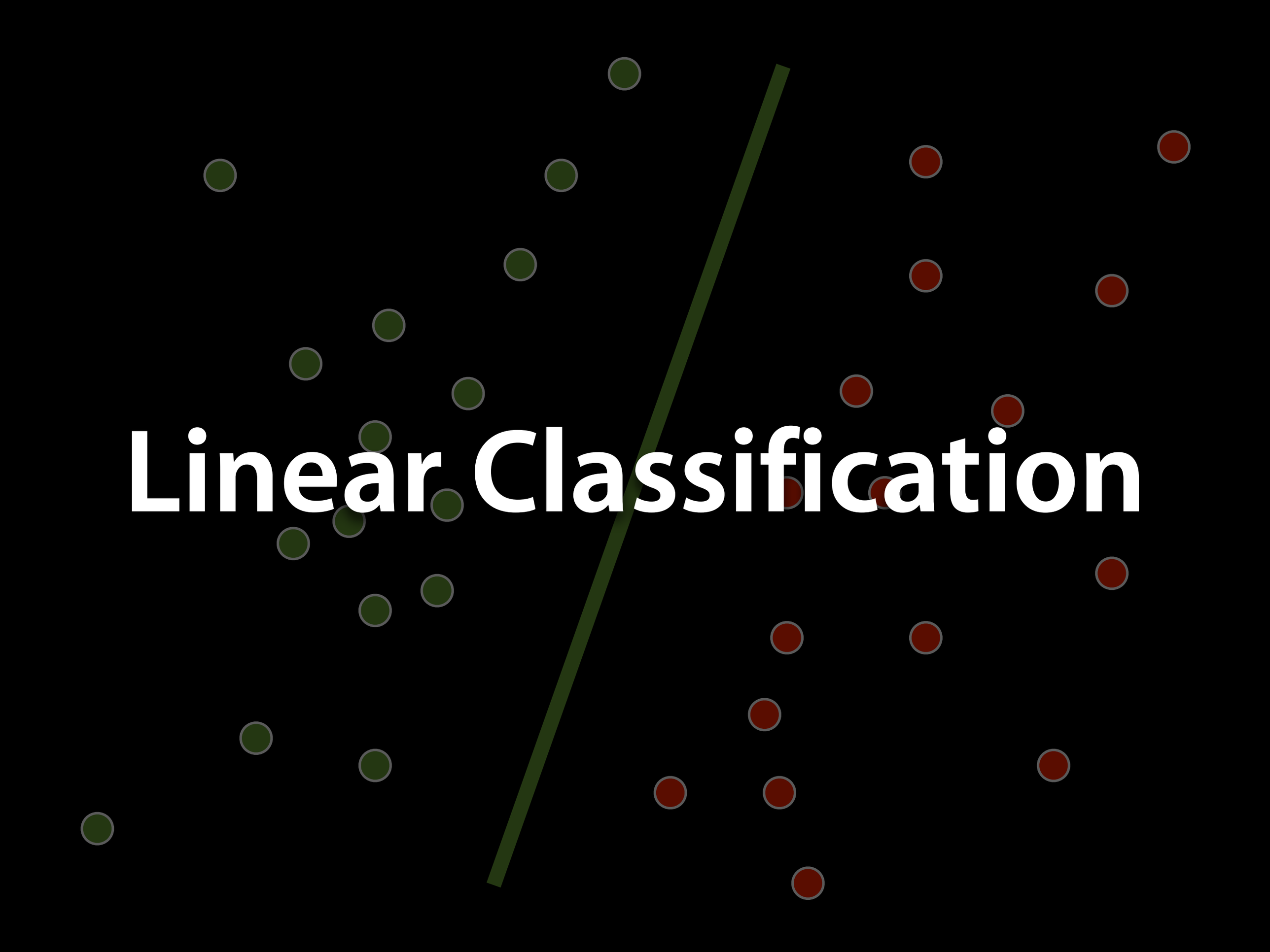
$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

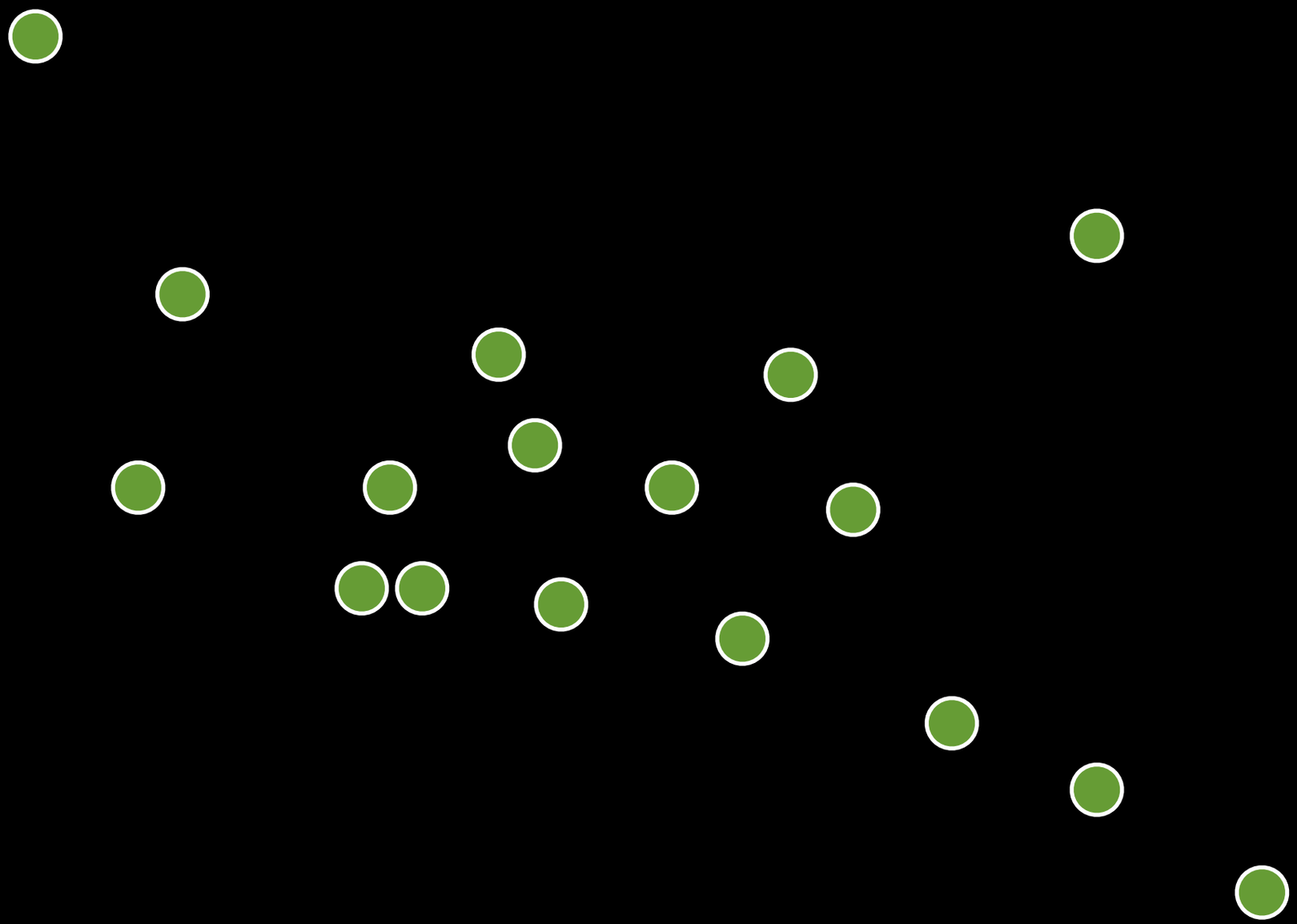


$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}$$

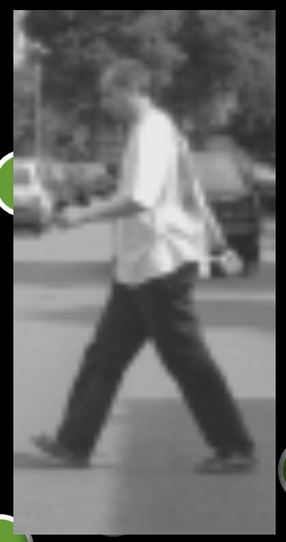
Euclidean distance

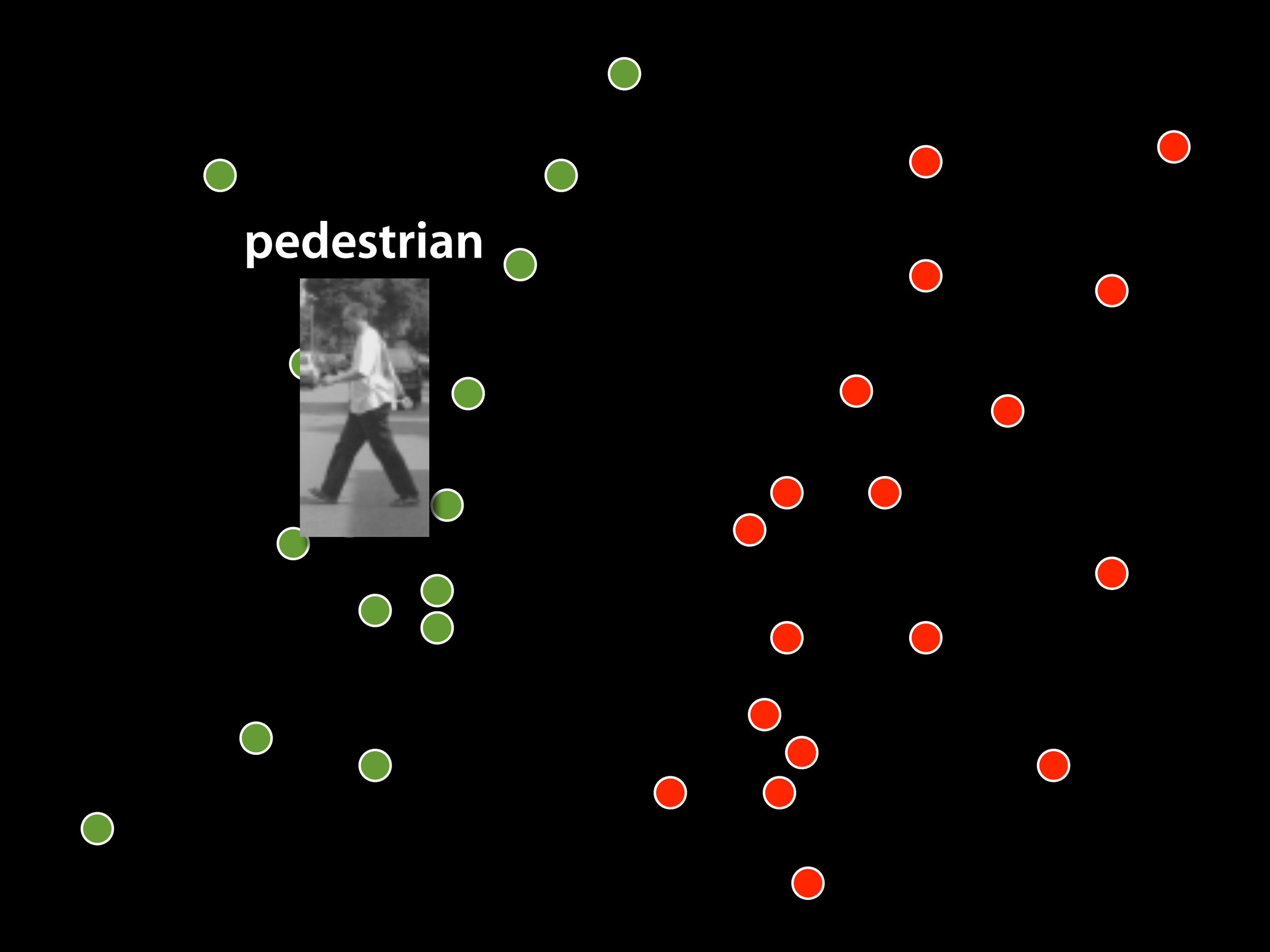

$$d(\mathbf{x},\mathbf{y}) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$
Euclidean distance
$\mathbf{x}$
$\mathbf{y}$
N
E

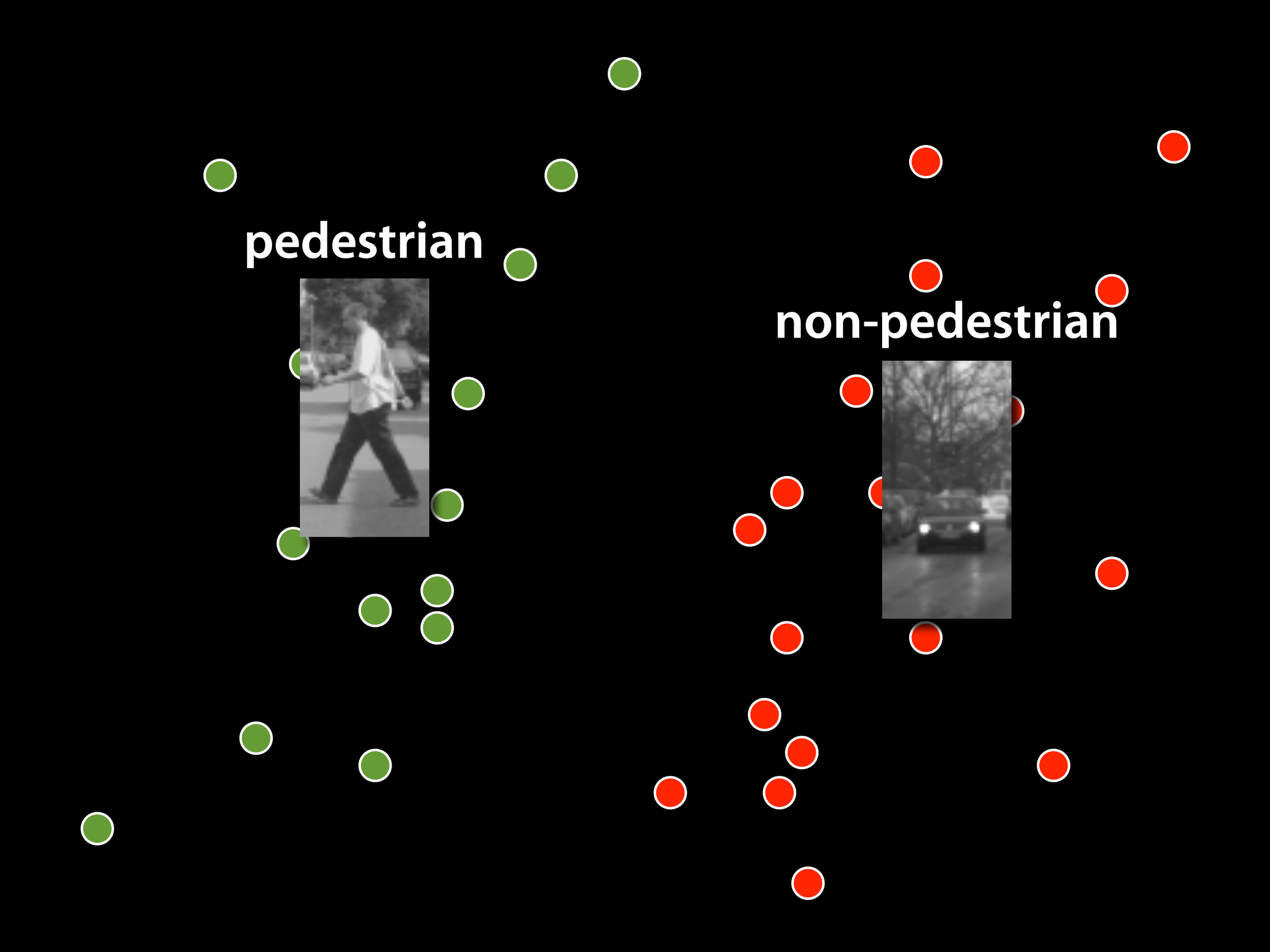

Linear Classification

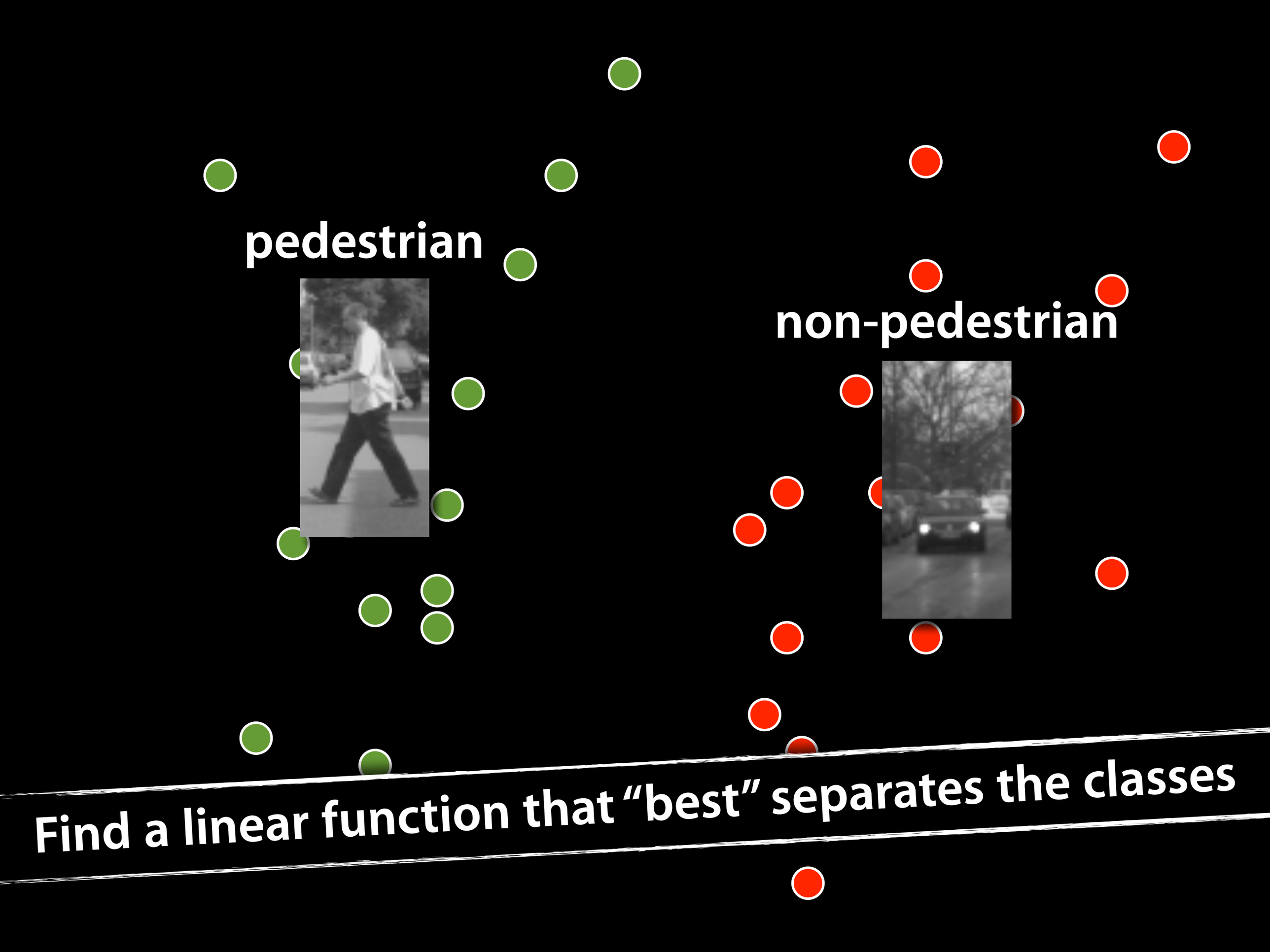


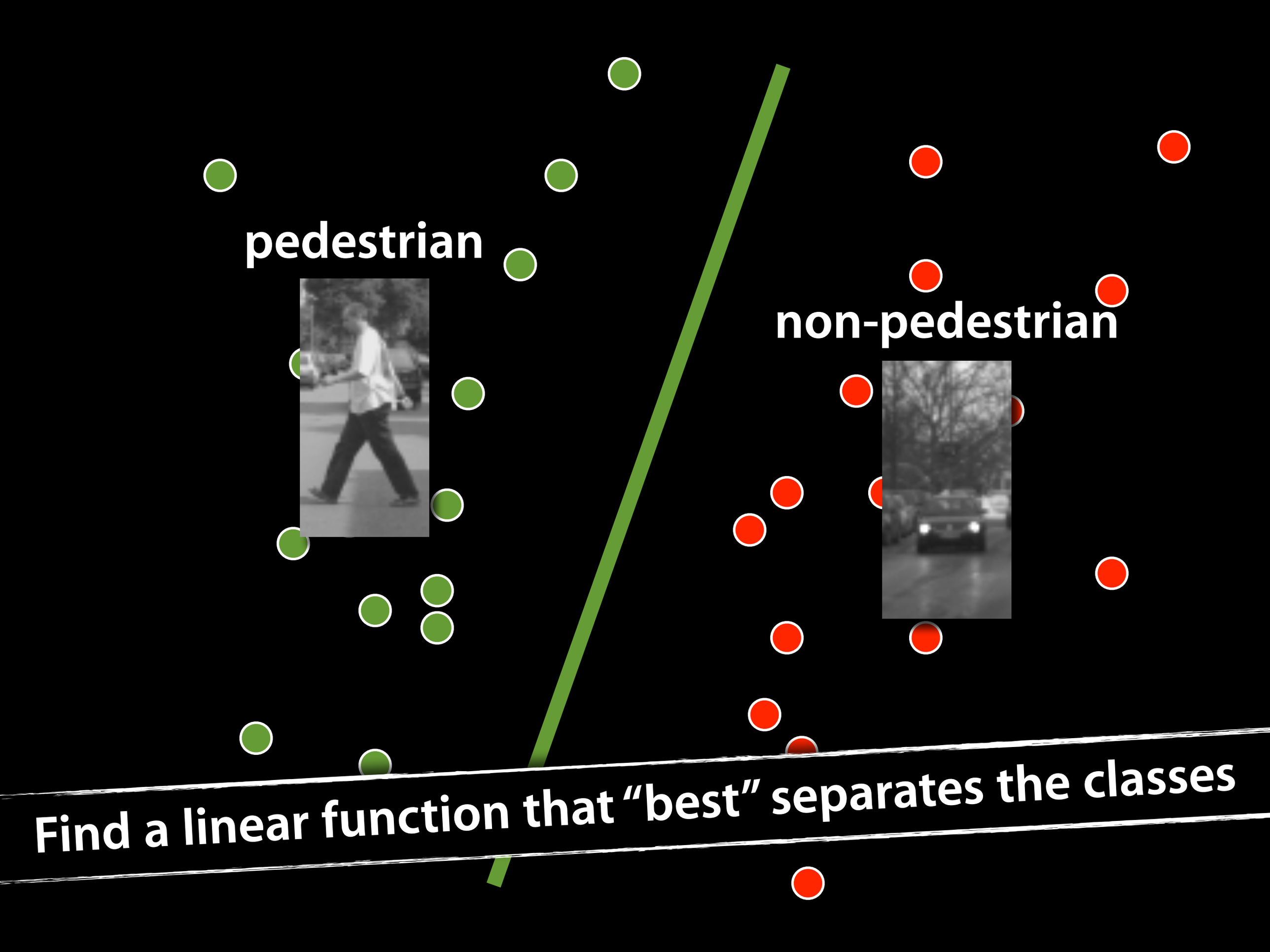
pedestrian




pedestrian

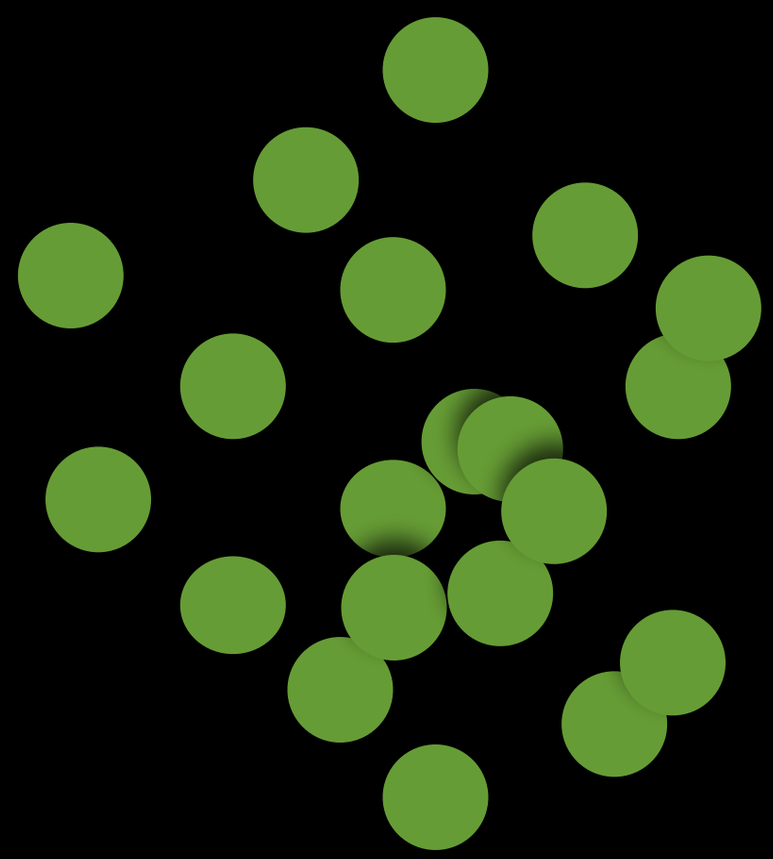

pedestrian
non-pedestrian

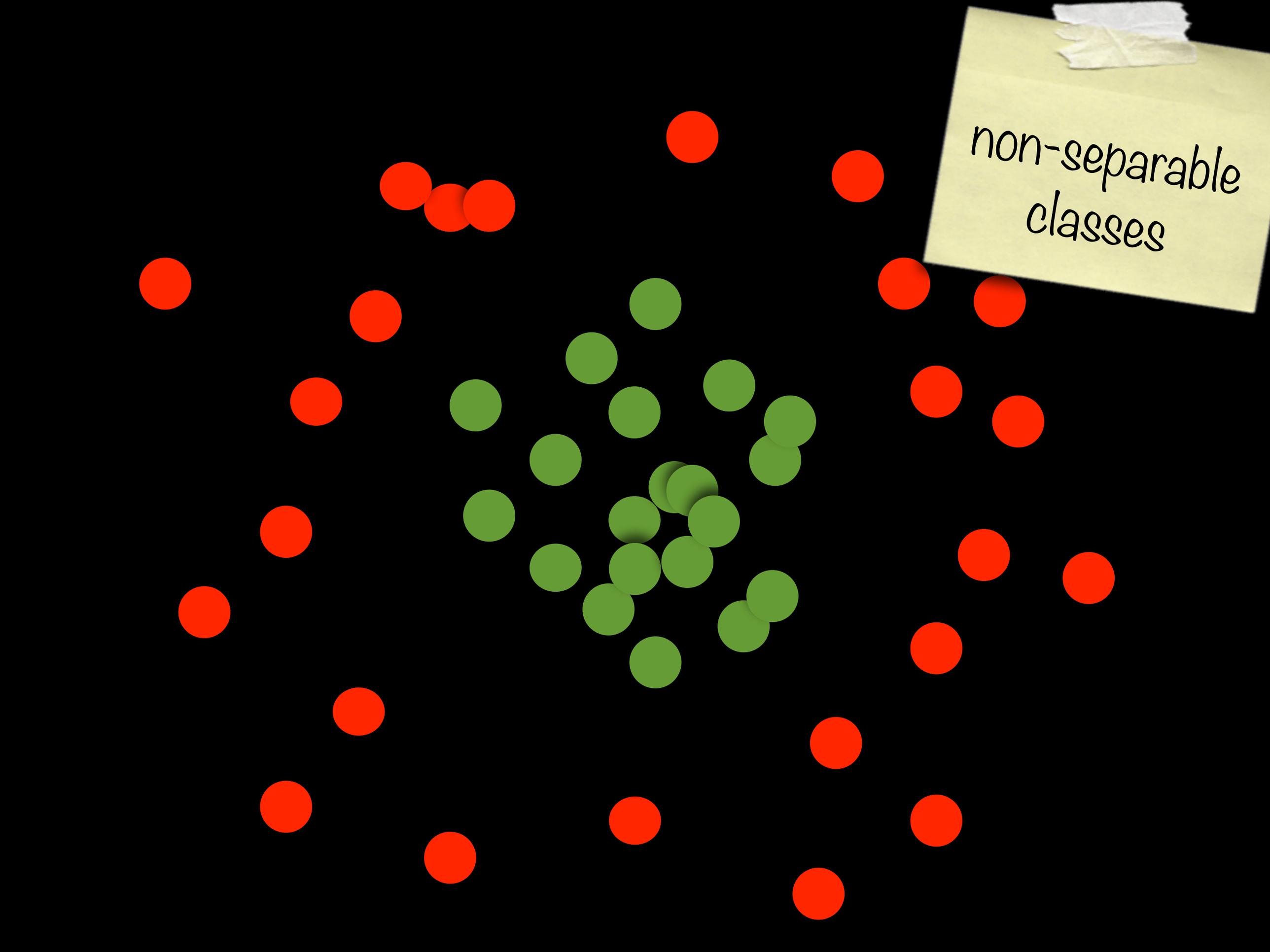

pedestrian
non-pedestrian
Find a linear function that “best” separates the classes

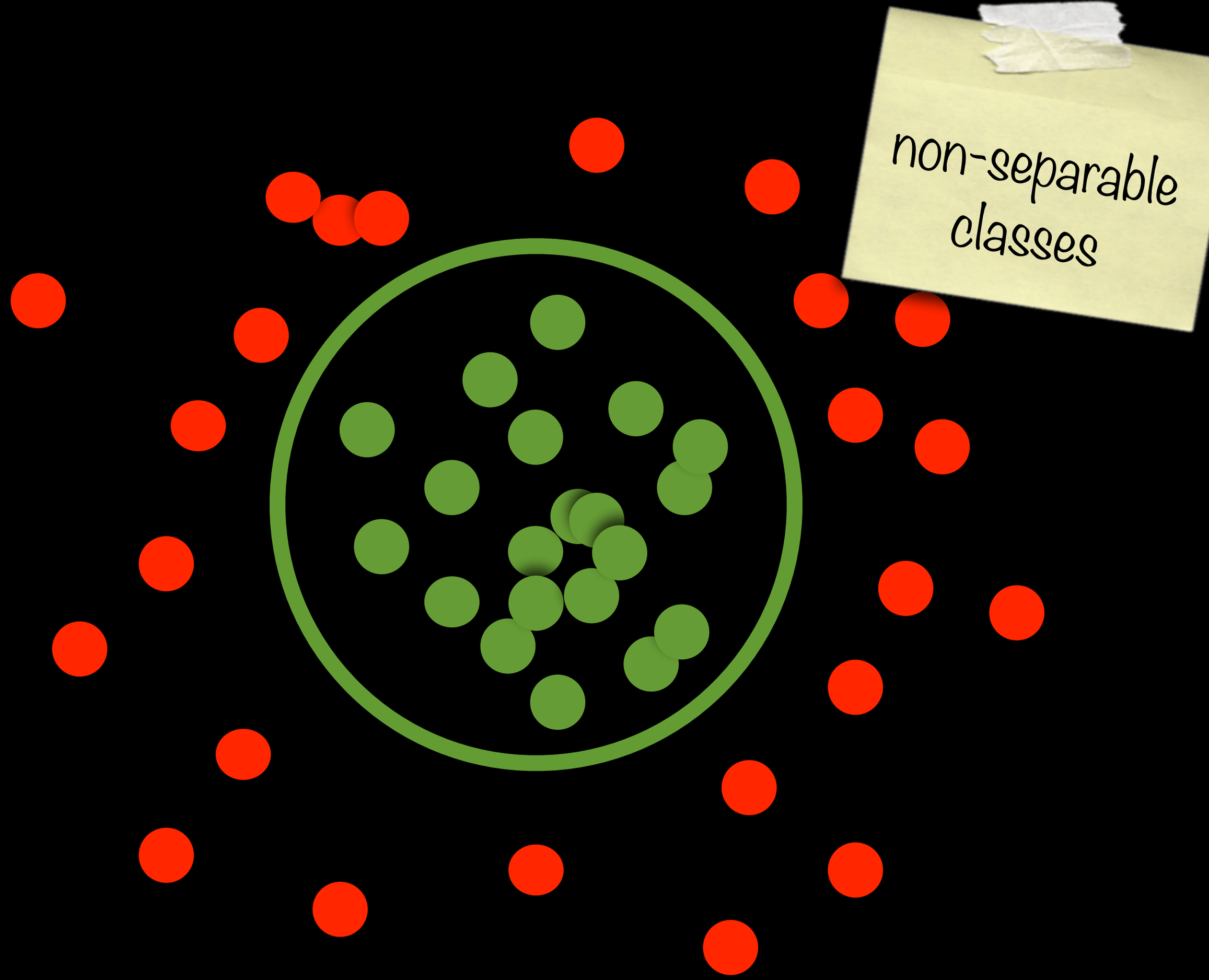

pedestrian
non-pedestrian
Find a linear function that "best" separates the classes

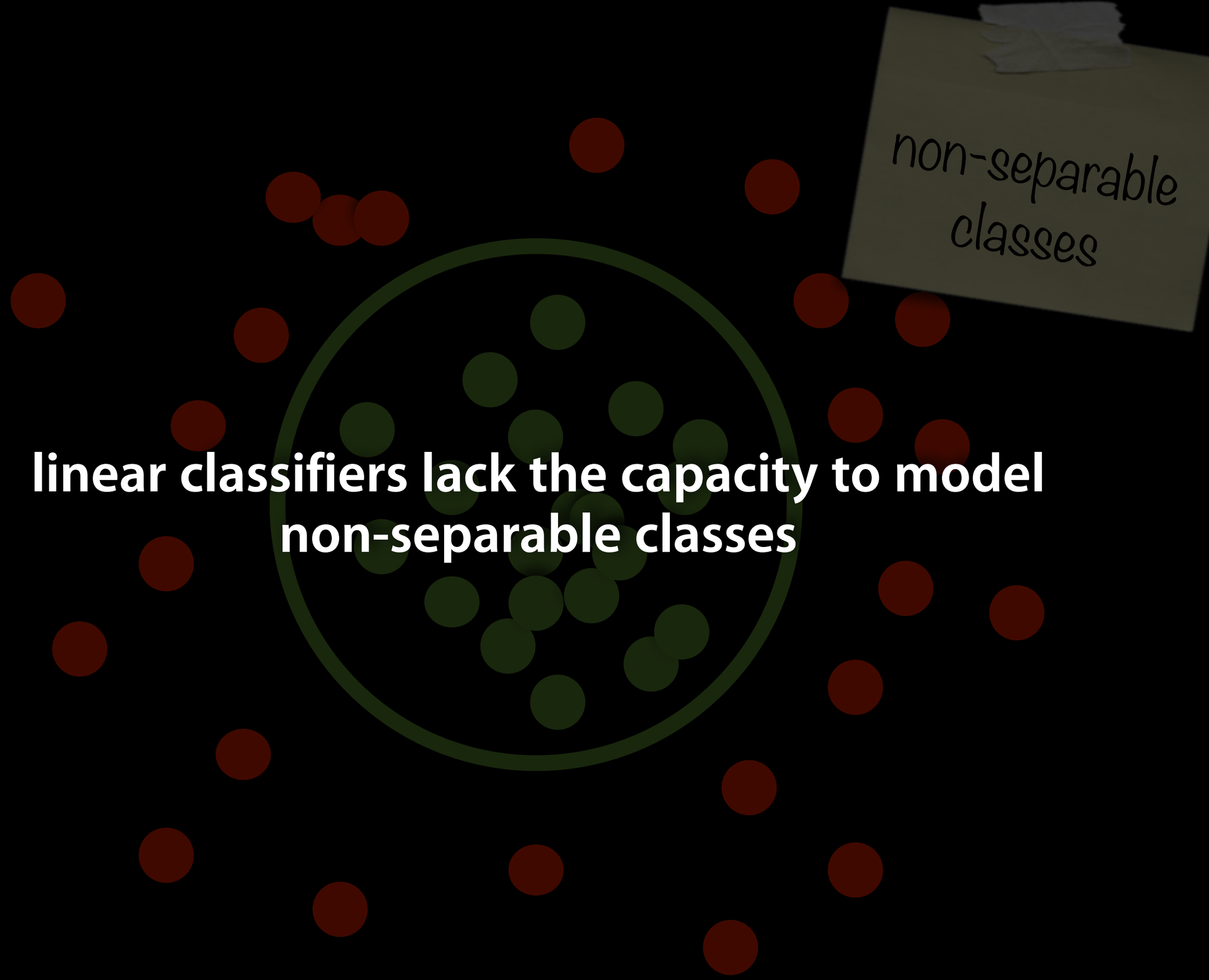
non-separable classes

non-separable classes

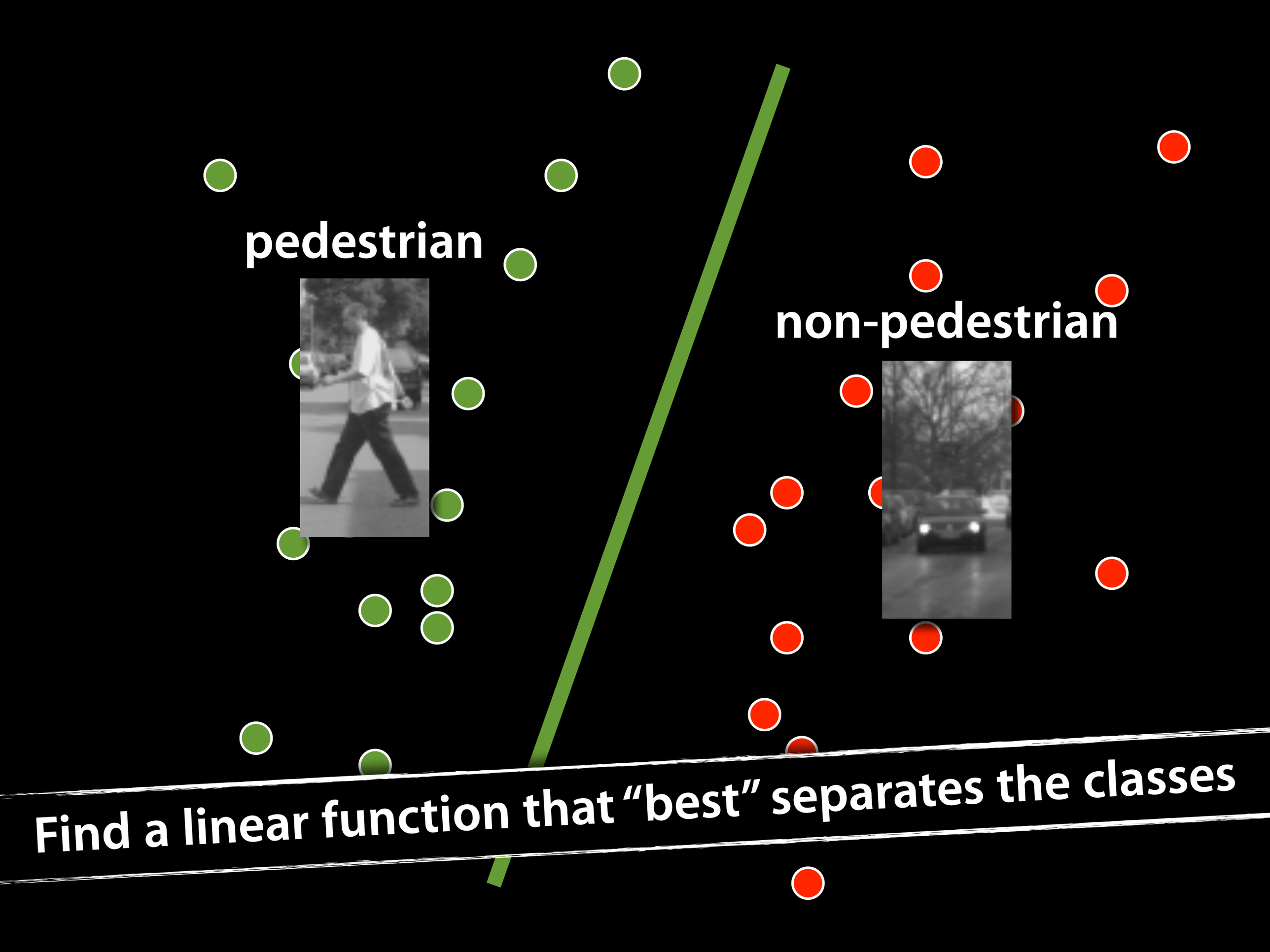


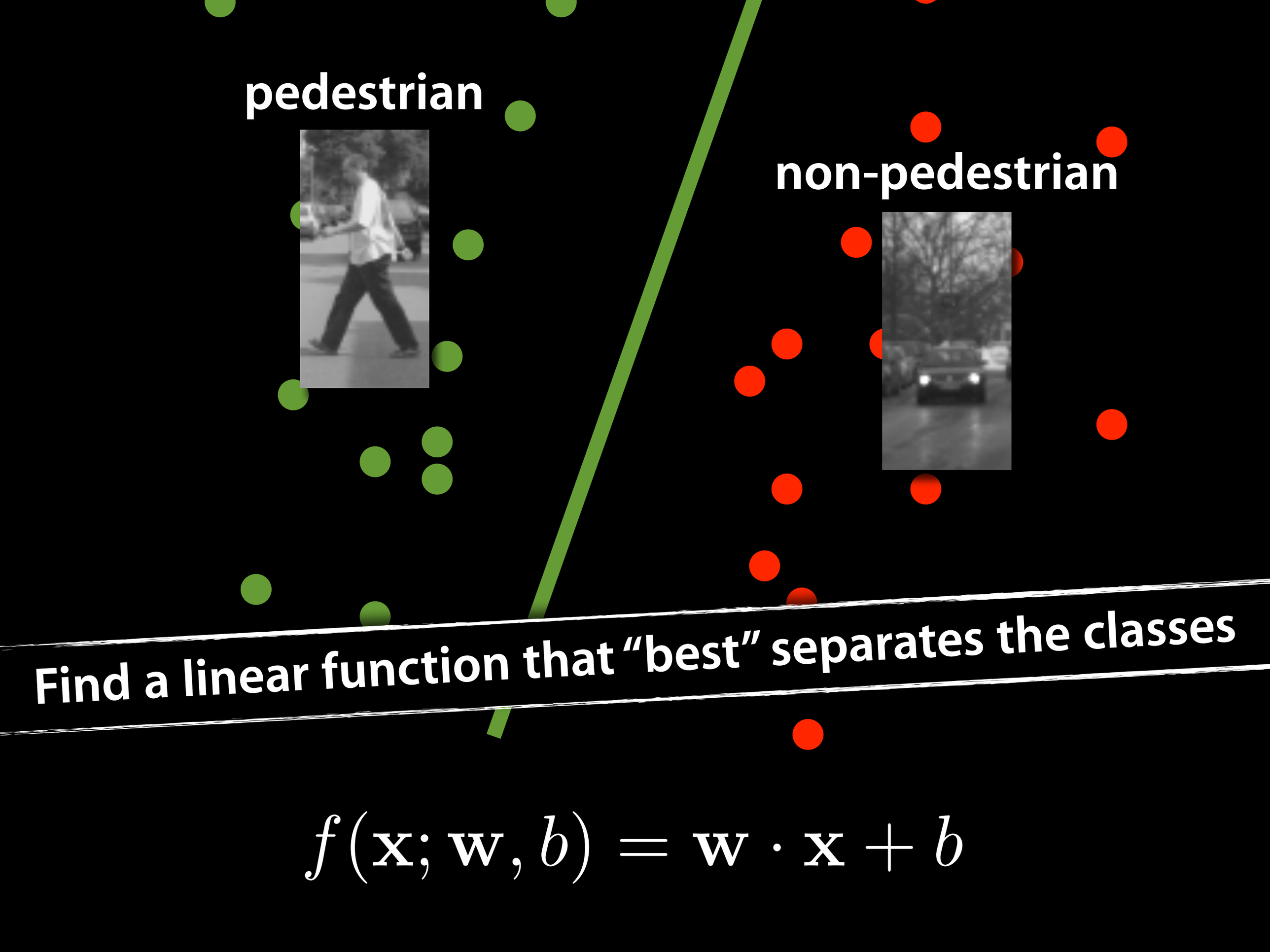

non-separable
classes


non-separable
classes



linear classifiers lack the capacity to model non-separable classes


pedestrian
non-pedestrian
Find a linear function that "best" separates the classes

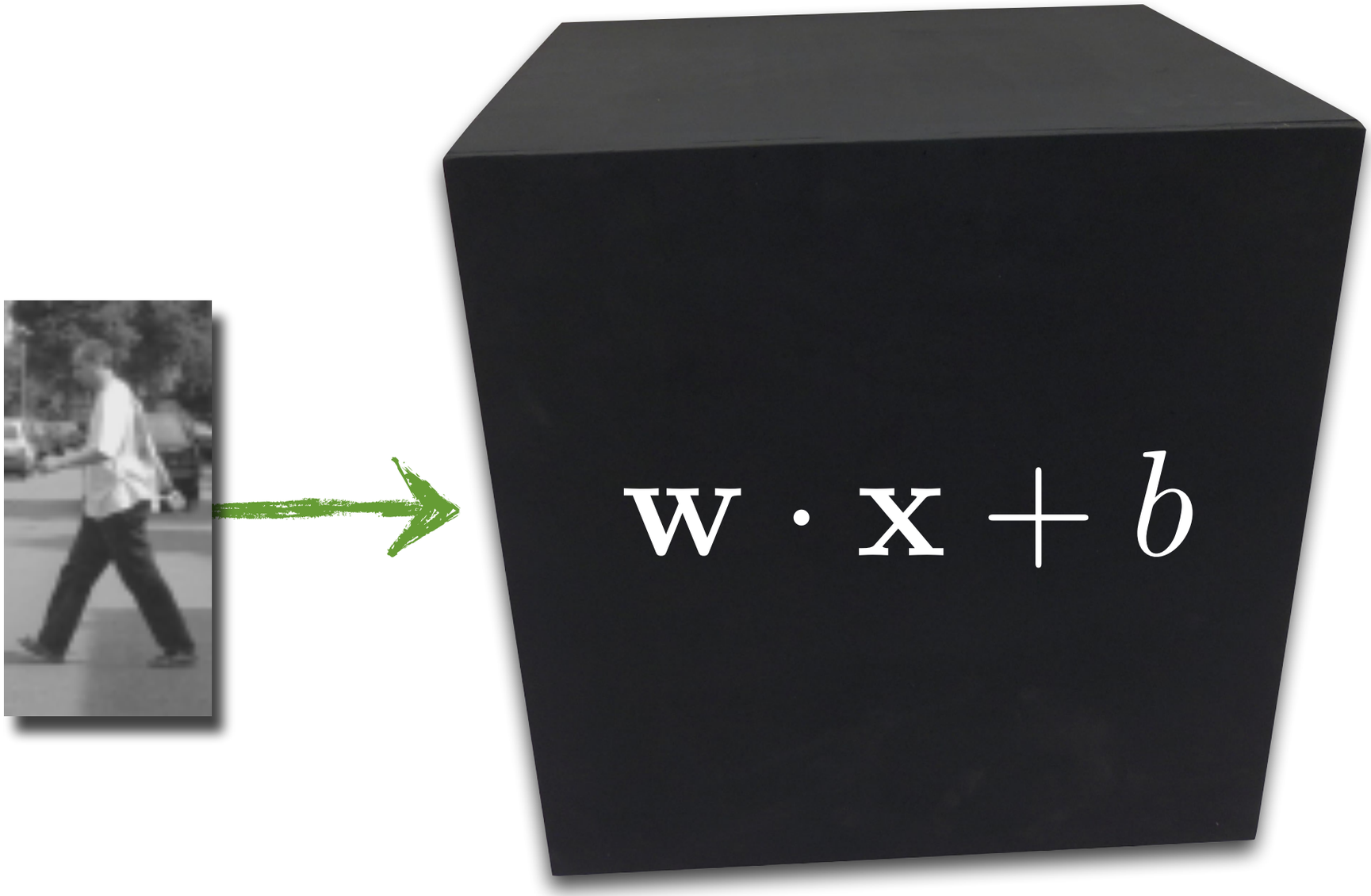


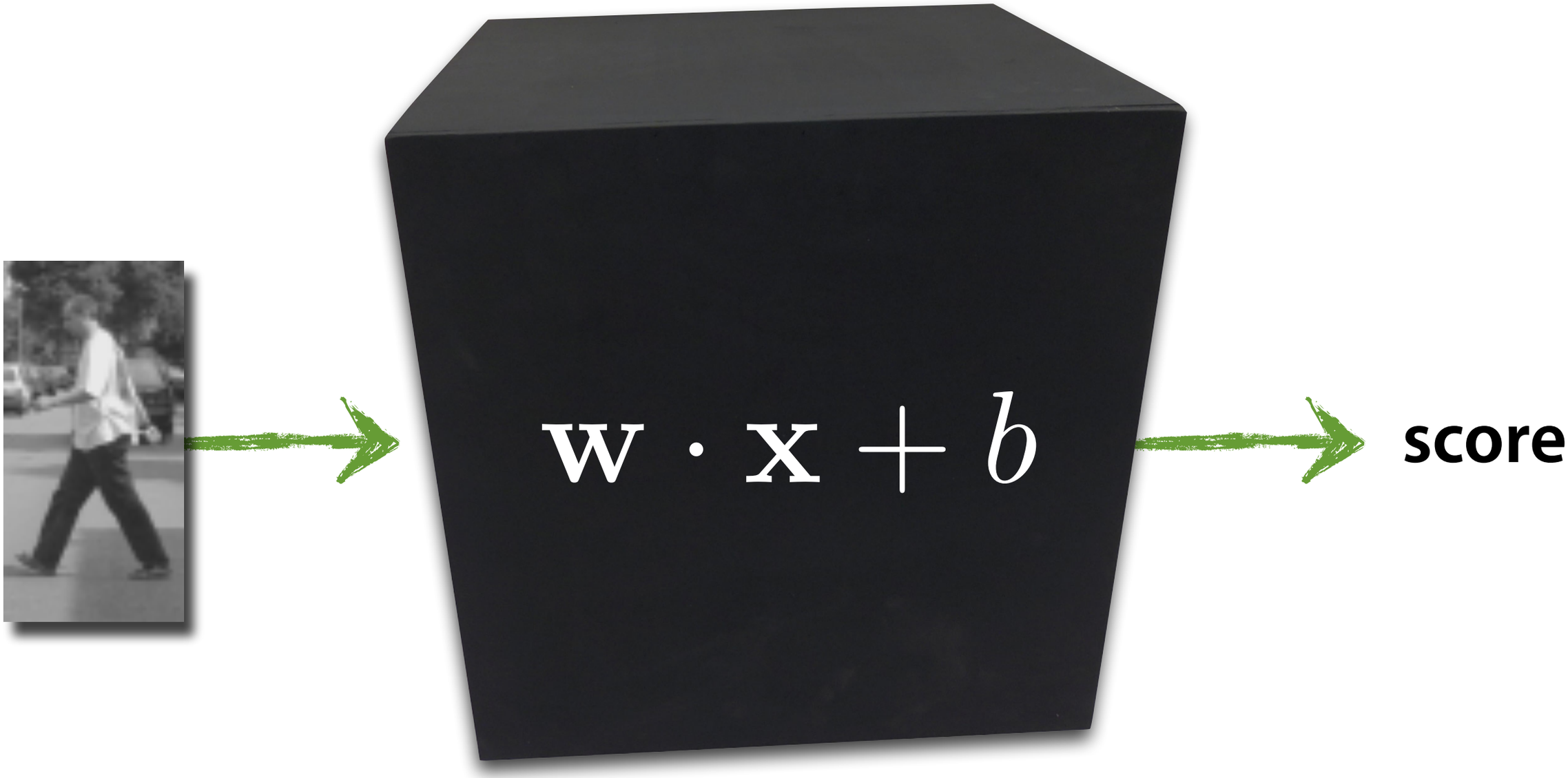
$$f(\mathbf{x}; \mathbf{w}, b) = \mathbf{w} \cdot \mathbf{x} + b$$

At testing, the sign of the score determines the class

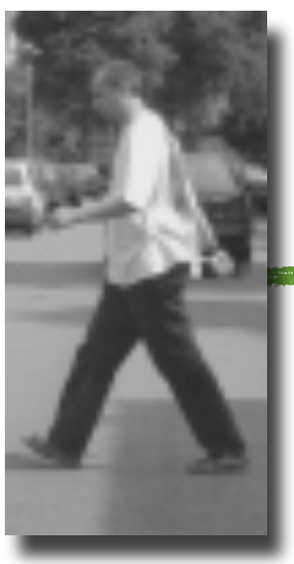
$$f(\mathbf{x}; \mathbf{w}, b) = \mathbf{w} \cdot \mathbf{x} + b$$

$$\mathbf{w} \cdot \mathbf{x} + b$$


$\mathbf{w} \cdot \mathbf{x} + b$


$\mathbf{w} \cdot \mathbf{x} + b$
score

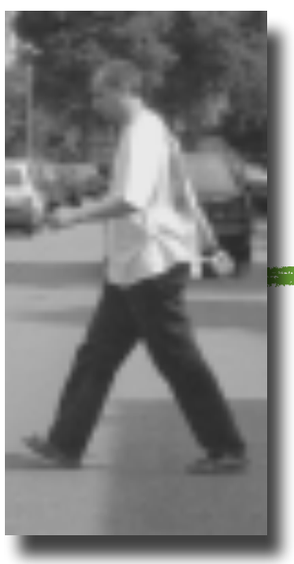
$\mathbf{x}$



$$\underbrace{\begin{pmatrix} 0.2 \\ -0.8 \\ 0.1 \\ 2.0 \end{pmatrix}}_{\mathbf{w}} \cdot \underbrace{\begin{pmatrix} 56 \\ 231 \\ 24 \\ 2 \end{pmatrix}}_{\mathbf{x}} \overset{b}{+ 1.1}$$

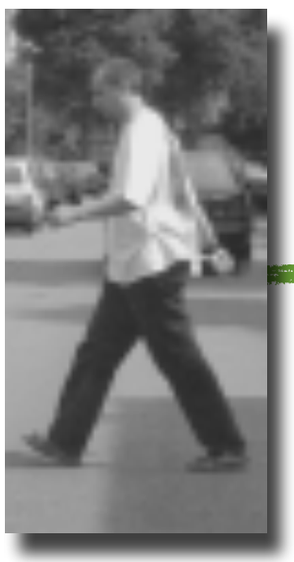
score

$\mathbf{x}$



$$\underbrace{\begin{pmatrix} 0.2 \\ -0.8 \\ 0.1 \\ 2.0 \end{pmatrix}}_{\mathbf{w}} \cdot \underbrace{\begin{pmatrix} 56 \\ 231 \\ 24 \\ 2 \end{pmatrix}}_{\mathbf{x}} \overset{b}{+ 1.1} \longrightarrow -96.8$$

$\mathbf{x}$

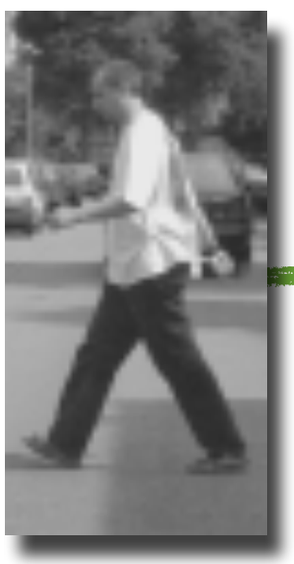


$$\begin{pmatrix} 0.2 \\ -0.8 \\ 0.1 \\ 2.0 \end{pmatrix} \cdot \begin{pmatrix} 56 \\ 231 \\ 24 \\ 2 \end{pmatrix} + 1.1 \longrightarrow -96.8$$

$\mathbf{w}$ $\mathbf{x}$ b

What are the best model parameters?

$$\mathbf{x} \quad \begin{pmatrix} 0.2 \\ -0.8 \\ 0.1 \\ 2.0 \end{pmatrix} \cdot \begin{pmatrix} 56 \\ 231 \\ 24 \\ 2 \end{pmatrix} + 1.1 \quad b \quad -96.8$$



$\mathbf{w}$ $\mathbf{x}$

Loss function quantifies agreement between predicted scores and the ground truth labels

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

scoring function

$$L_i = \max(0, 1 - y_i(\mathbf{w} \cdot \mathbf{x}_i + b))$$

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

training set label, 1 or -1

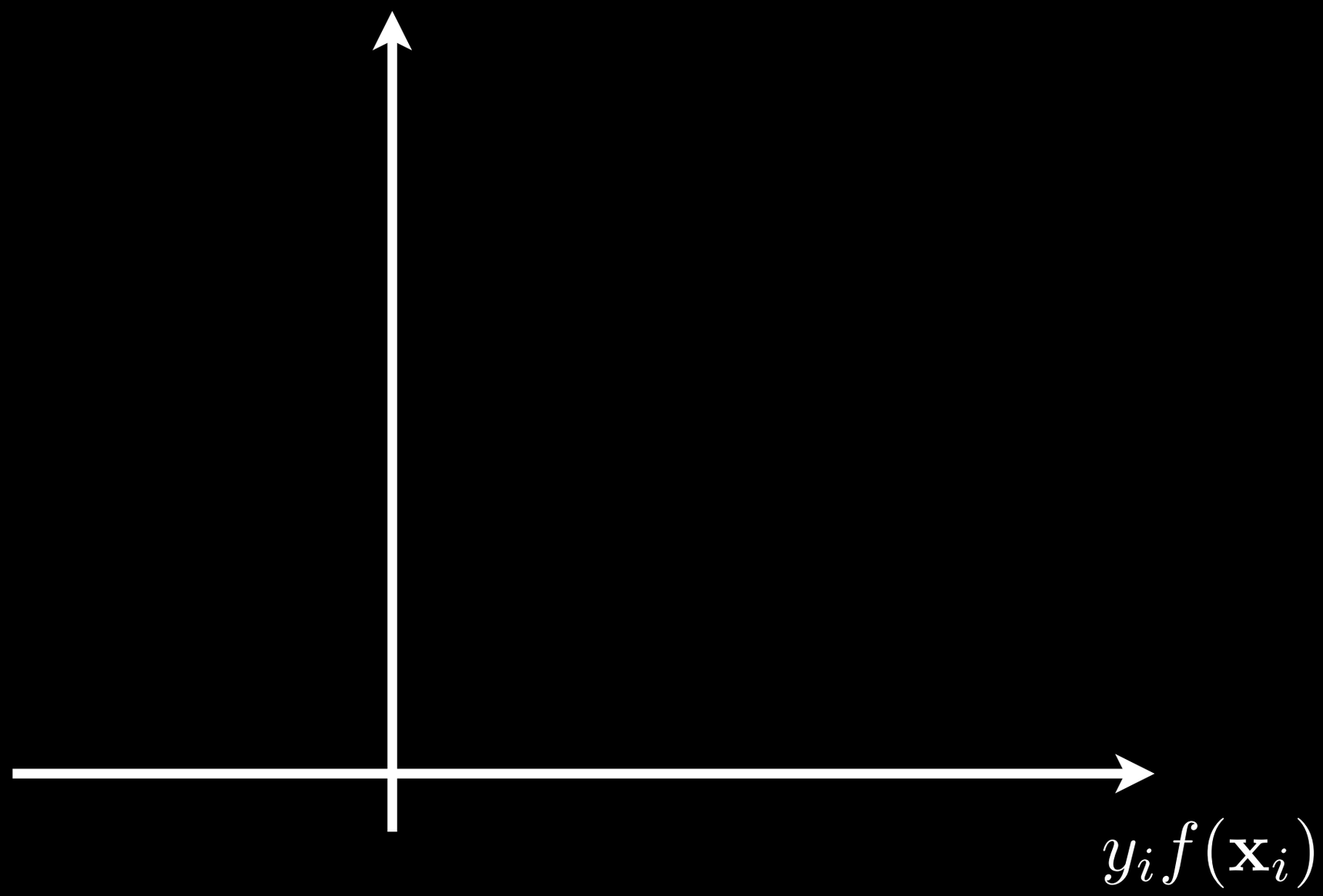
loss margin

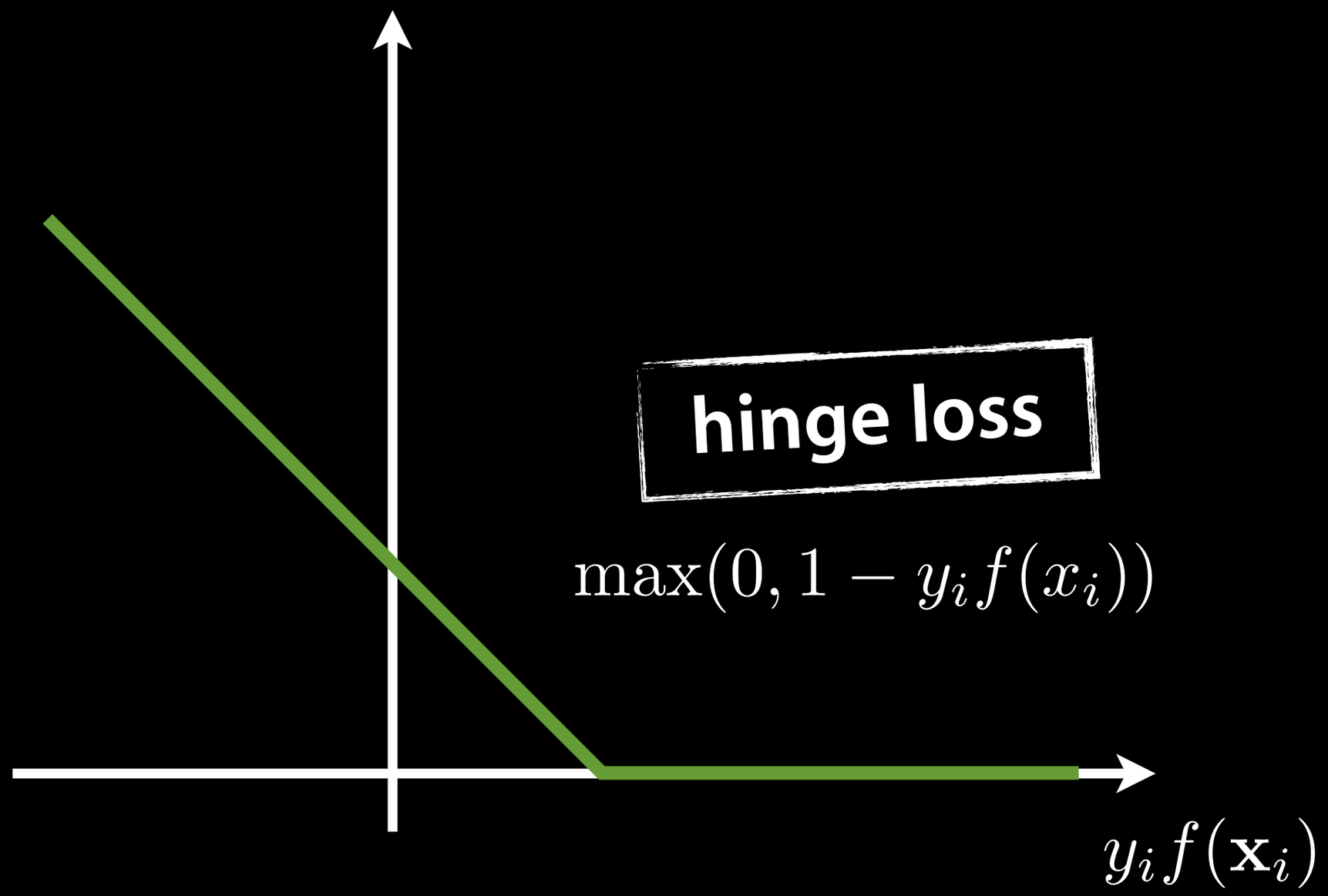
$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

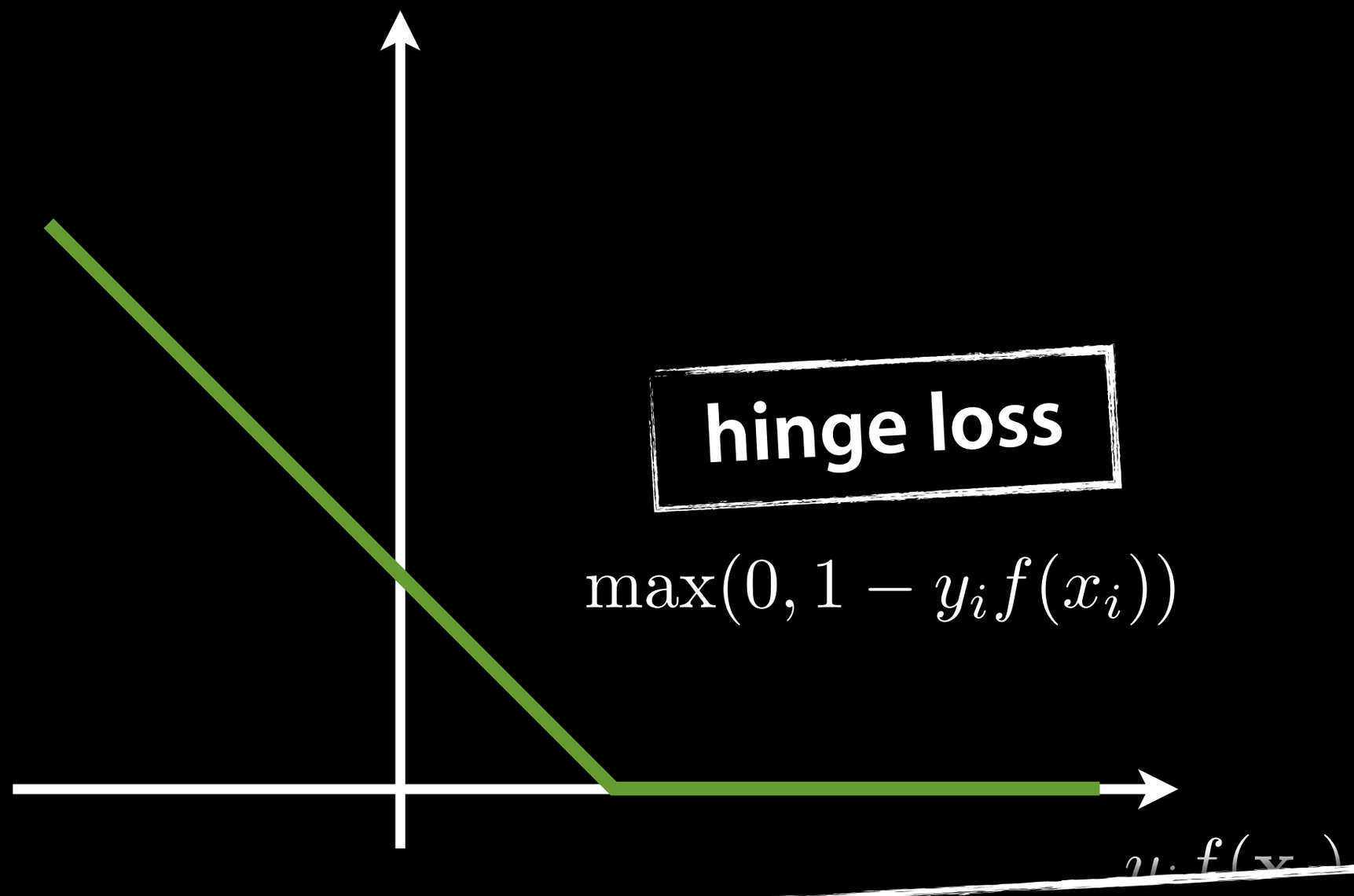
$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

hinge loss function


$y_i f(\mathbf{x}_i)$

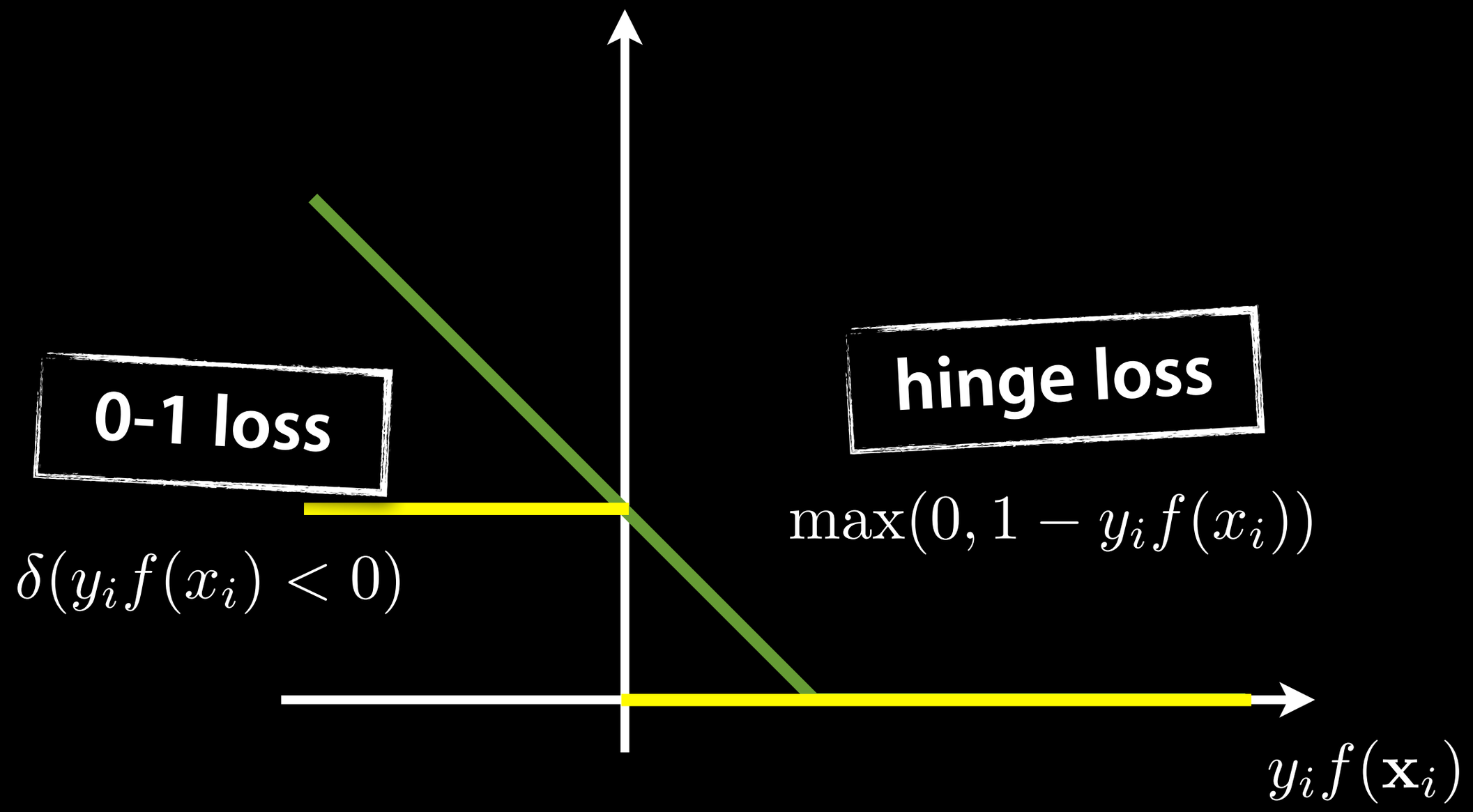

hinge loss
$\max(0, 1 - y_i f(x_i))$
$y_i f(\mathbf{x}_i)$

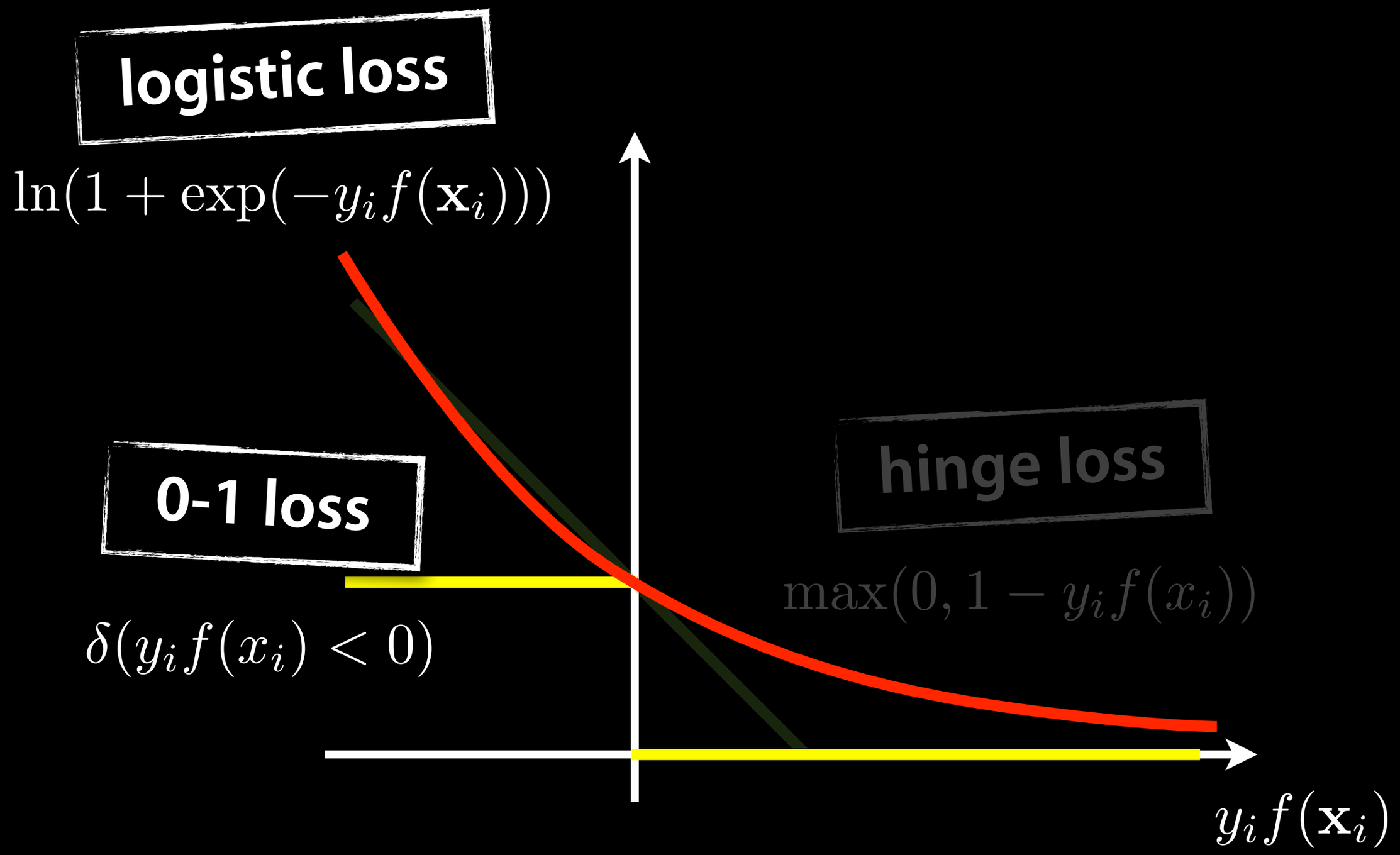


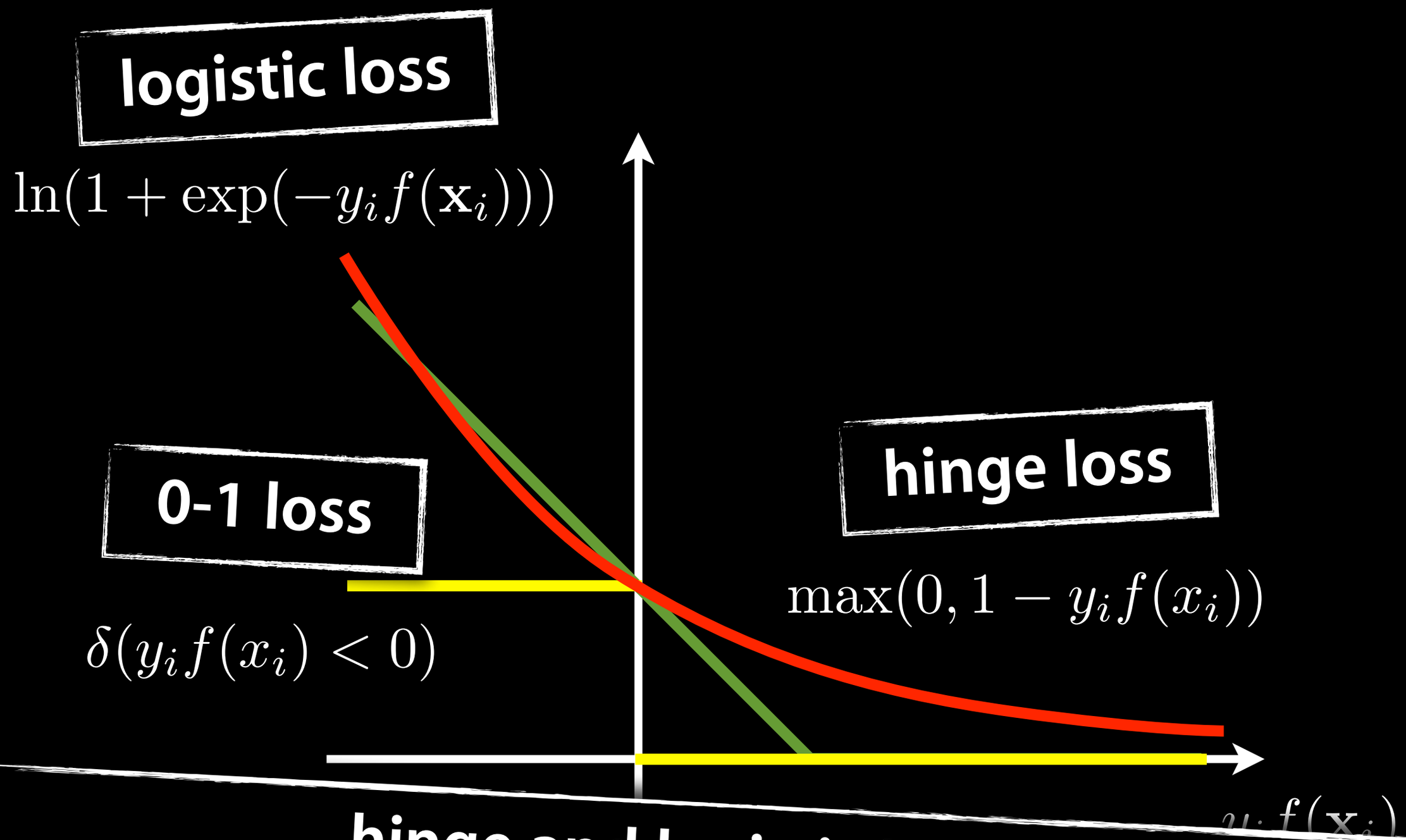
hinge loss

$$\max(0, 1 - y_i f(x_i))$$

linear penalty if point is within margin


0-1 loss
$\delta(y_i f(x_i) < 0)$
hinge loss
$\max(0, 1 - y_i f(x_i))$
$y_i f(\mathbf{x}_i)$


logistic loss
$\ln(1 + \exp(-y_i f(\mathbf{x}_i)))$
0-1 loss
$\delta(y_i f(x_i) < 0)$
hinge loss
$\max(0, 1 - y_i f(x_i))$
$y_i f(\mathbf{x}_i)$

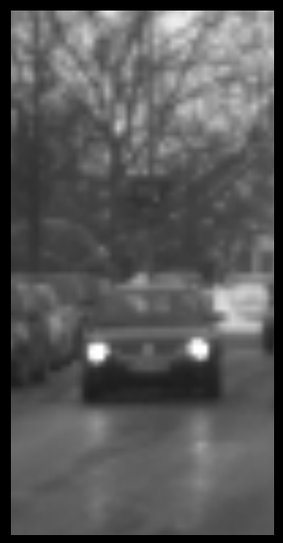


hinge and logistic losses are convex approximations of 0-1 loss

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

hinge loss function

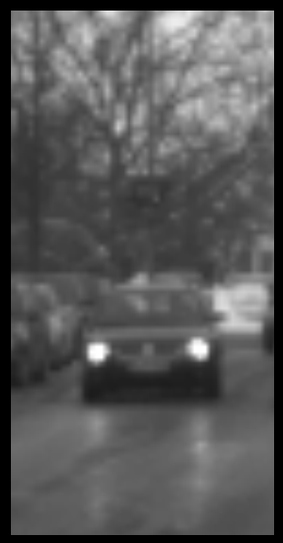
$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

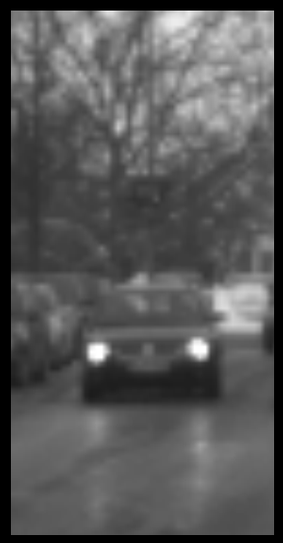
score

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



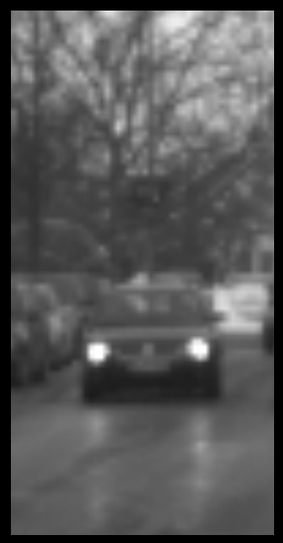
score 0.3

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



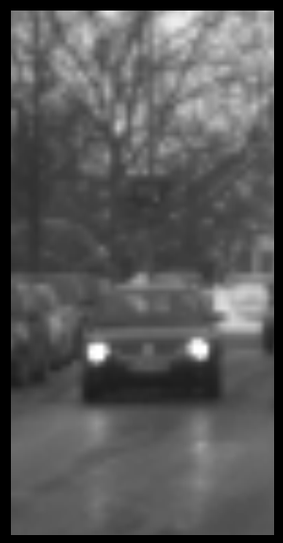
score 0.3 -30

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



score 0.3 -30 -96.8

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



score	0.3	-30	-96.8
loss			

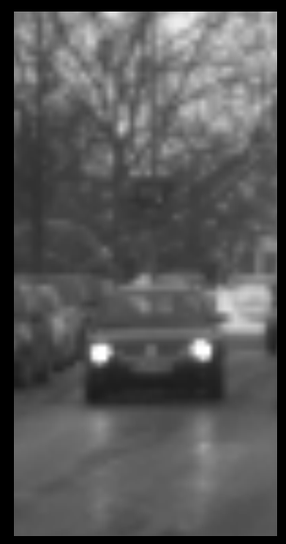
$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

score	0.3	-30	-96.8
loss	0.7		

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

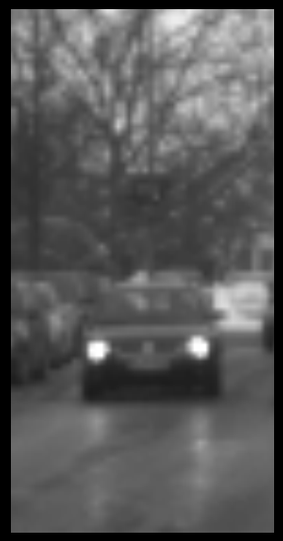
score	0.3	-30	-96.8
loss	0.7	0	

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



score	0.3	-30	-96.8
loss	0.7	0	97.8

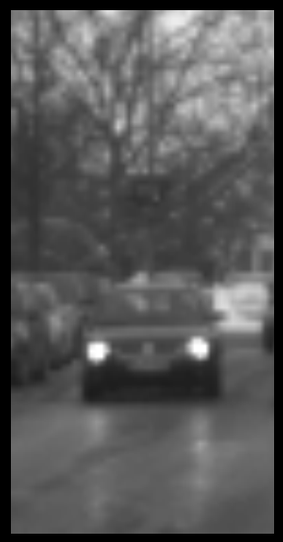
$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



score	0.3	-30	-96.8
loss	0.7	0	97.8

Overall loss is the average loss over the training set

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

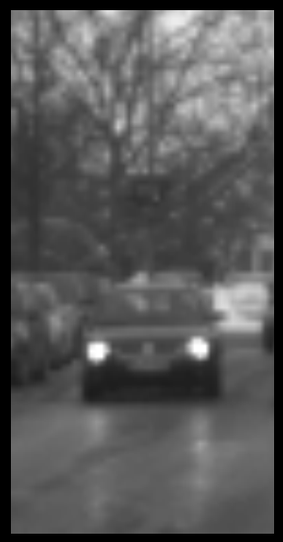


score	0.3	-30	-96.8
loss	0.7	0	97.8

Overall loss is the average loss over the training set

$$L = \frac{1}{N} \sum_{i=1}^{N} L_i$$

$$L_i = \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$



score	0.3	-30	-96.8
loss	0.7	0	97.8

Overall loss is the average loss over the training set

$$L = \frac{1}{N}\sum_{i=1}^{N} L_i = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

$$L = \frac{1}{N} \sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

$$L = \frac{1}{N} \sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b])$$

Loss function admits multiple solutions

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

L2 norm is a common regularizer

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w} \cdot \mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

L2 norm is a common regularizer

$$R(\mathbf{w}) = \|\mathbf{w}\|_2^2$$

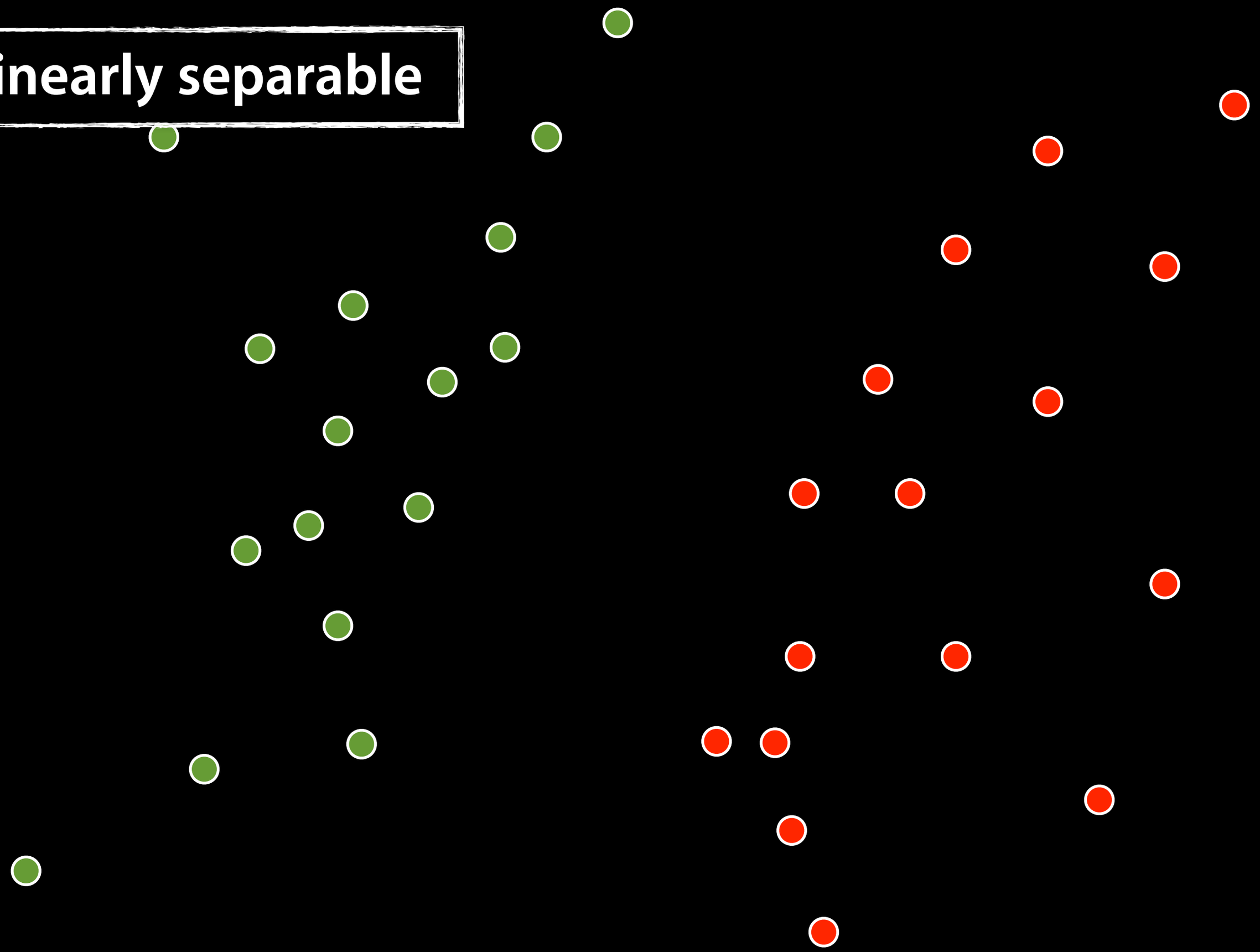
$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

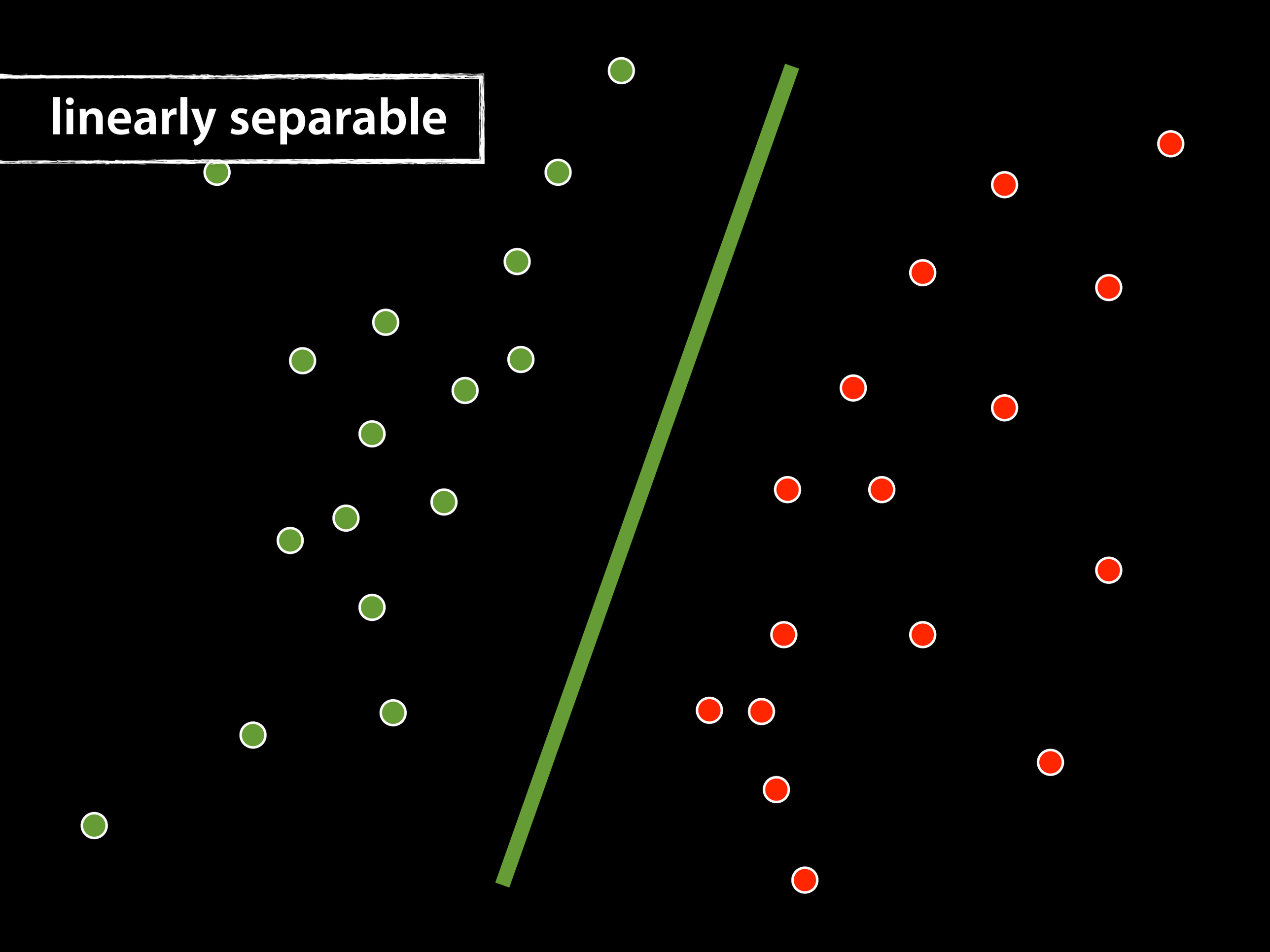
L2 norm is a common regularizer

$$R(\mathbf{w}) = \|\mathbf{w}\|_2^2$$

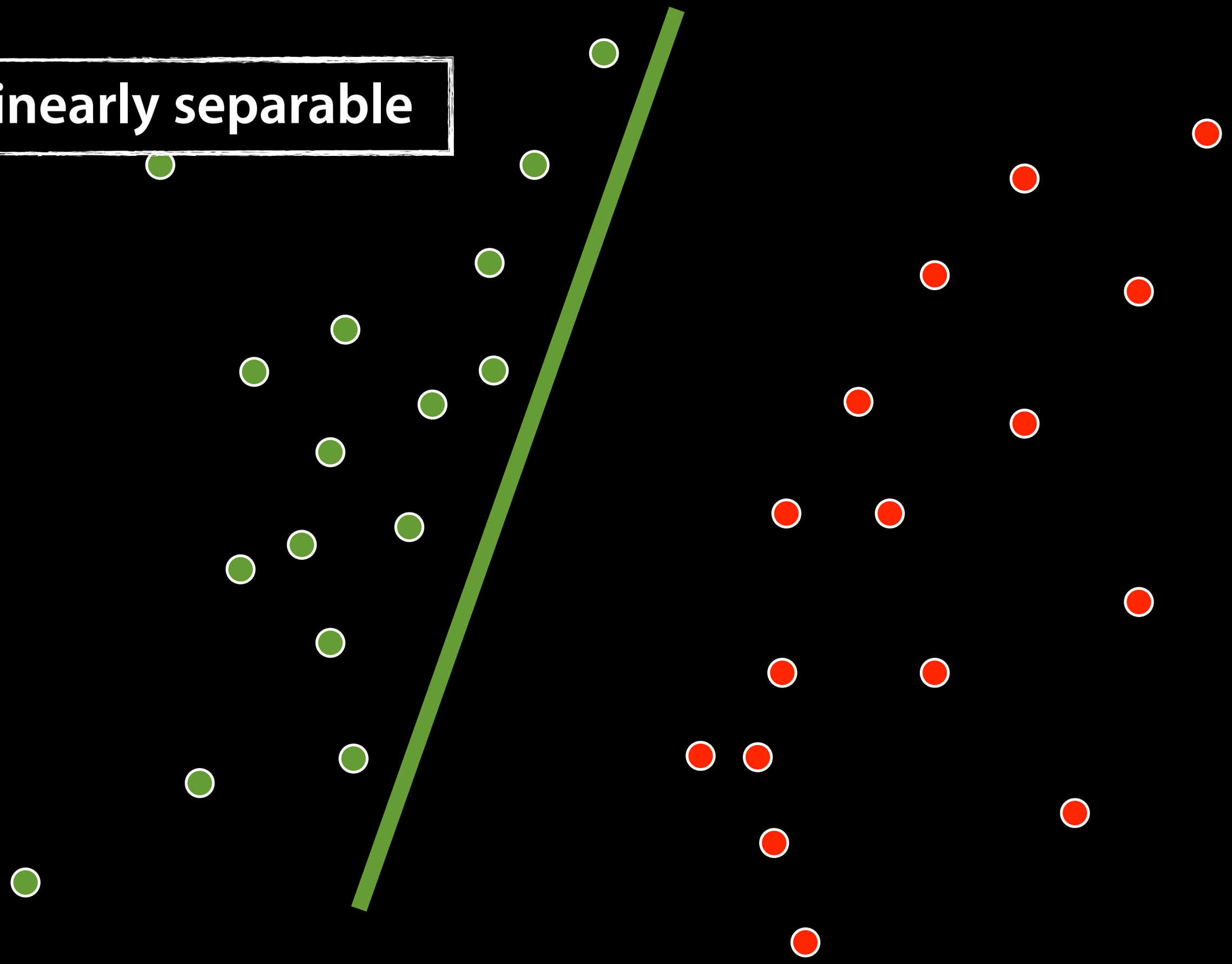
Support Vector Machine (SVM)

linearly separable

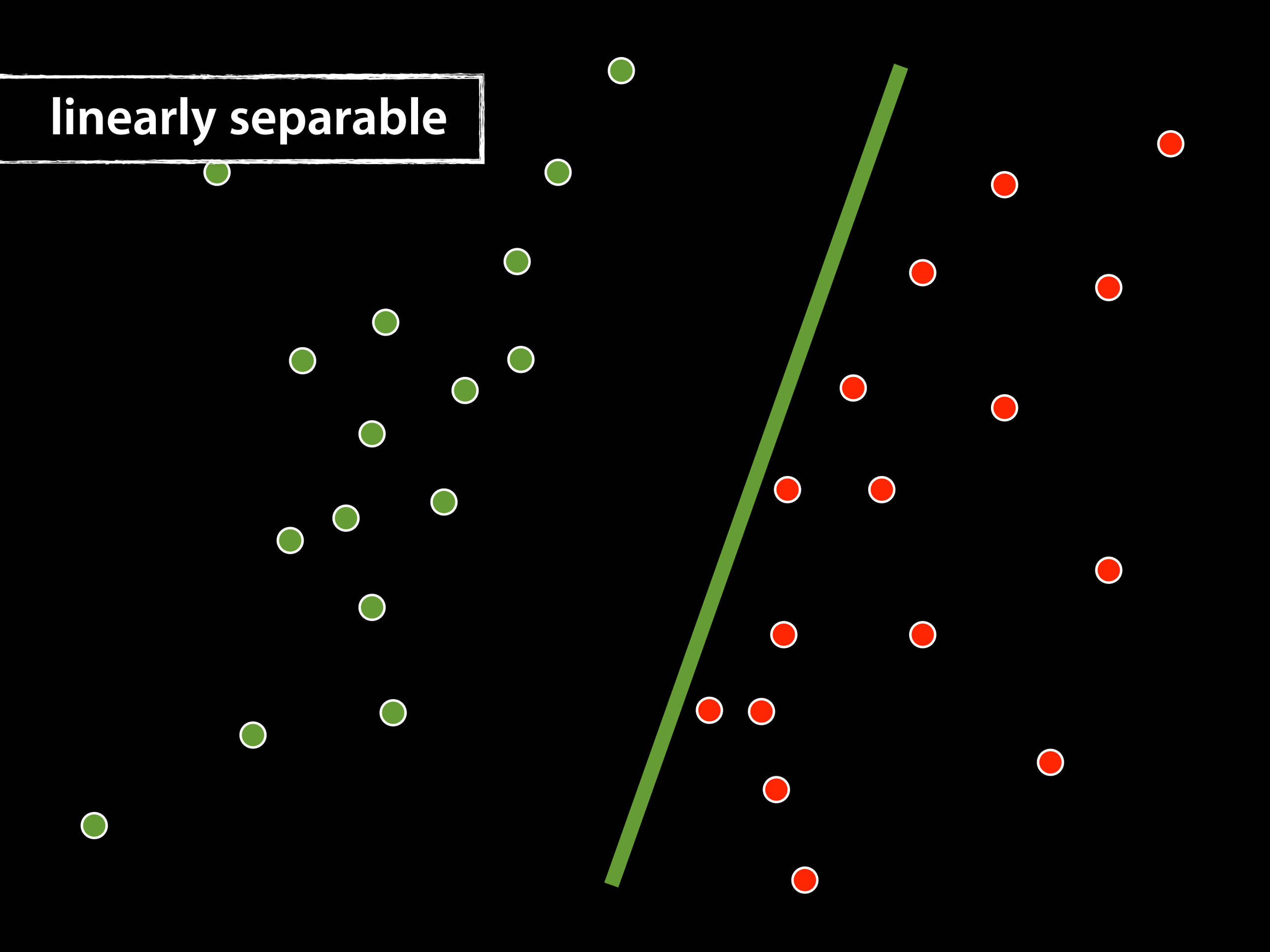


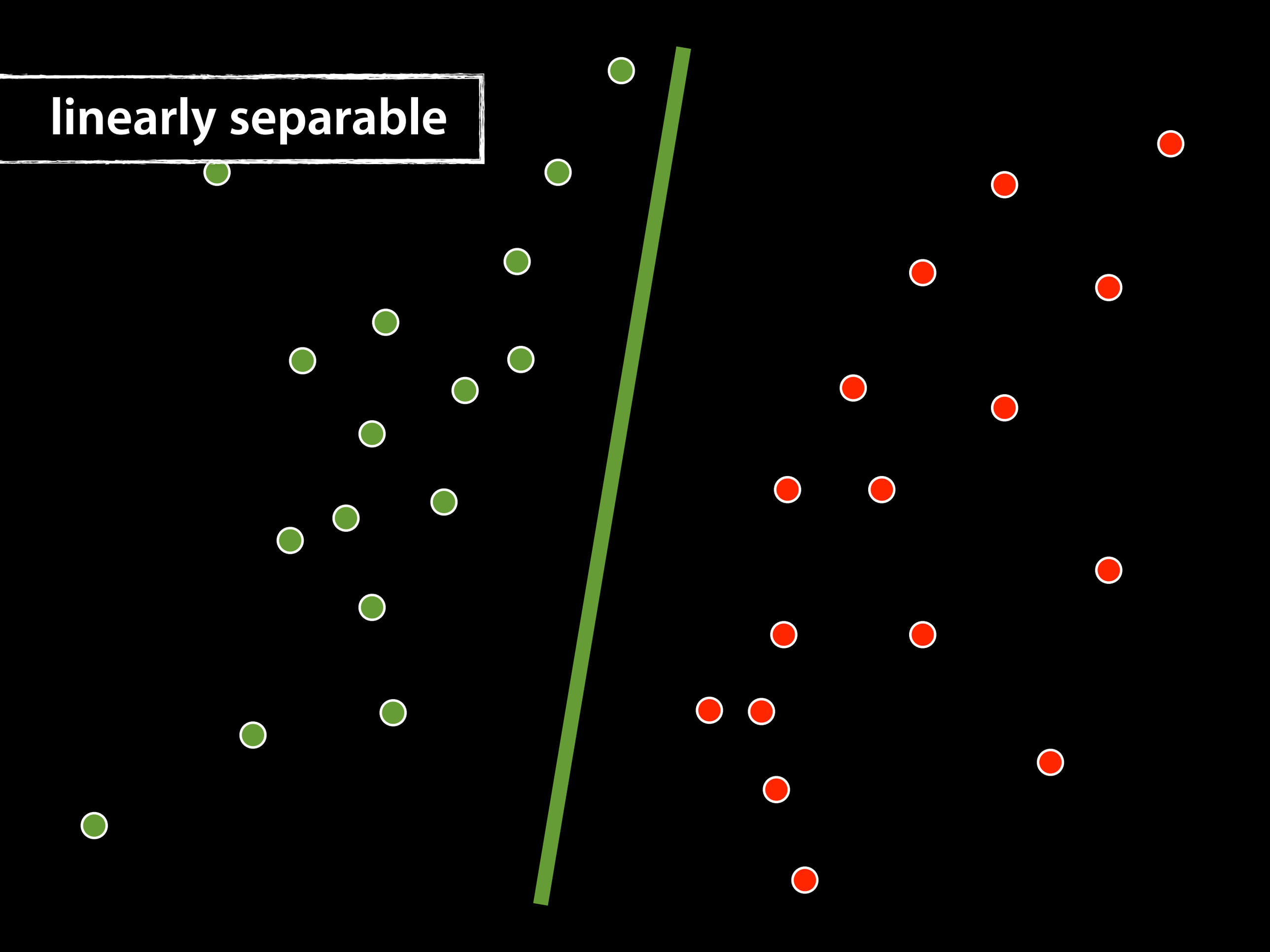

linearly separable

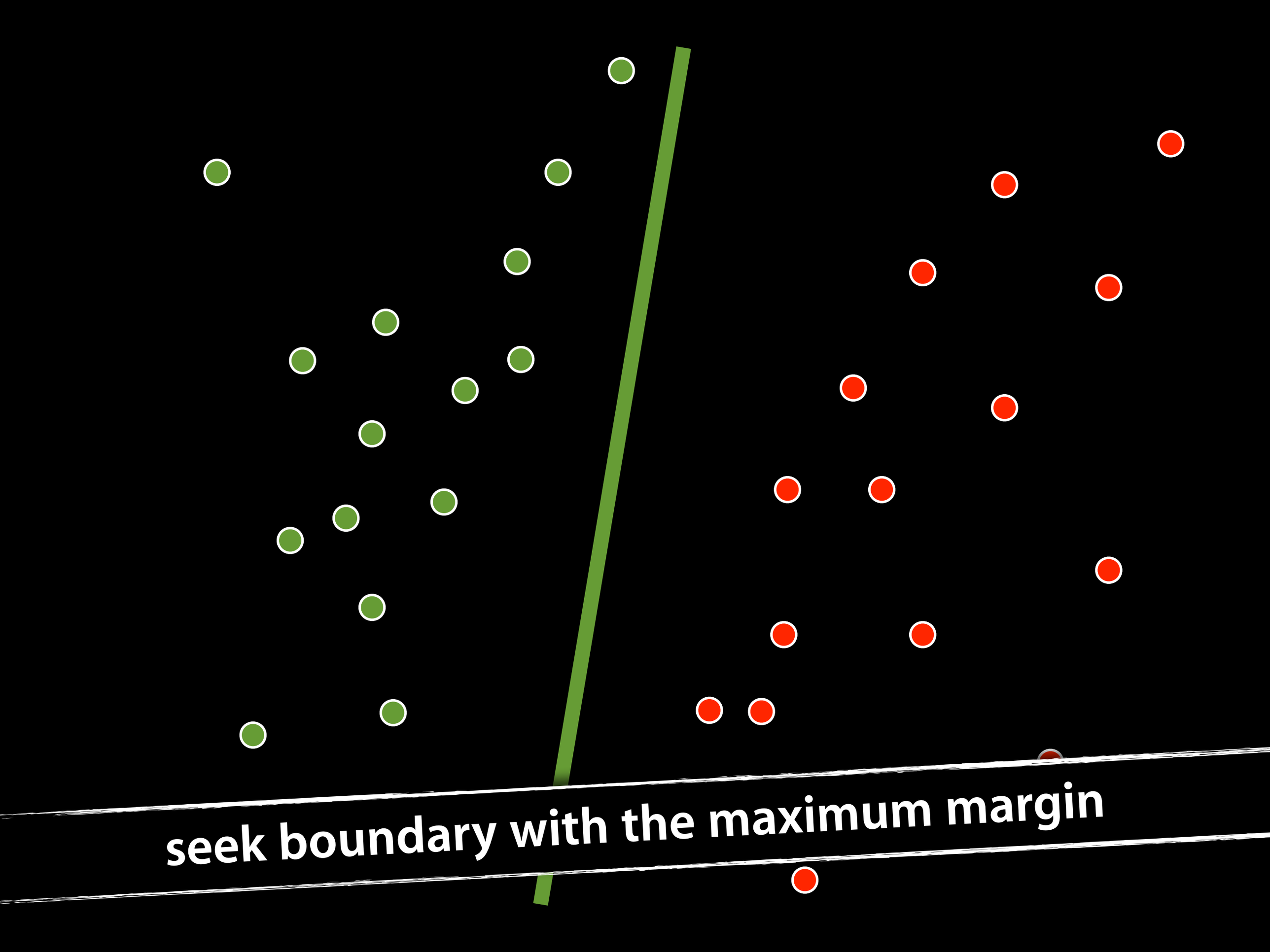
linearly separable

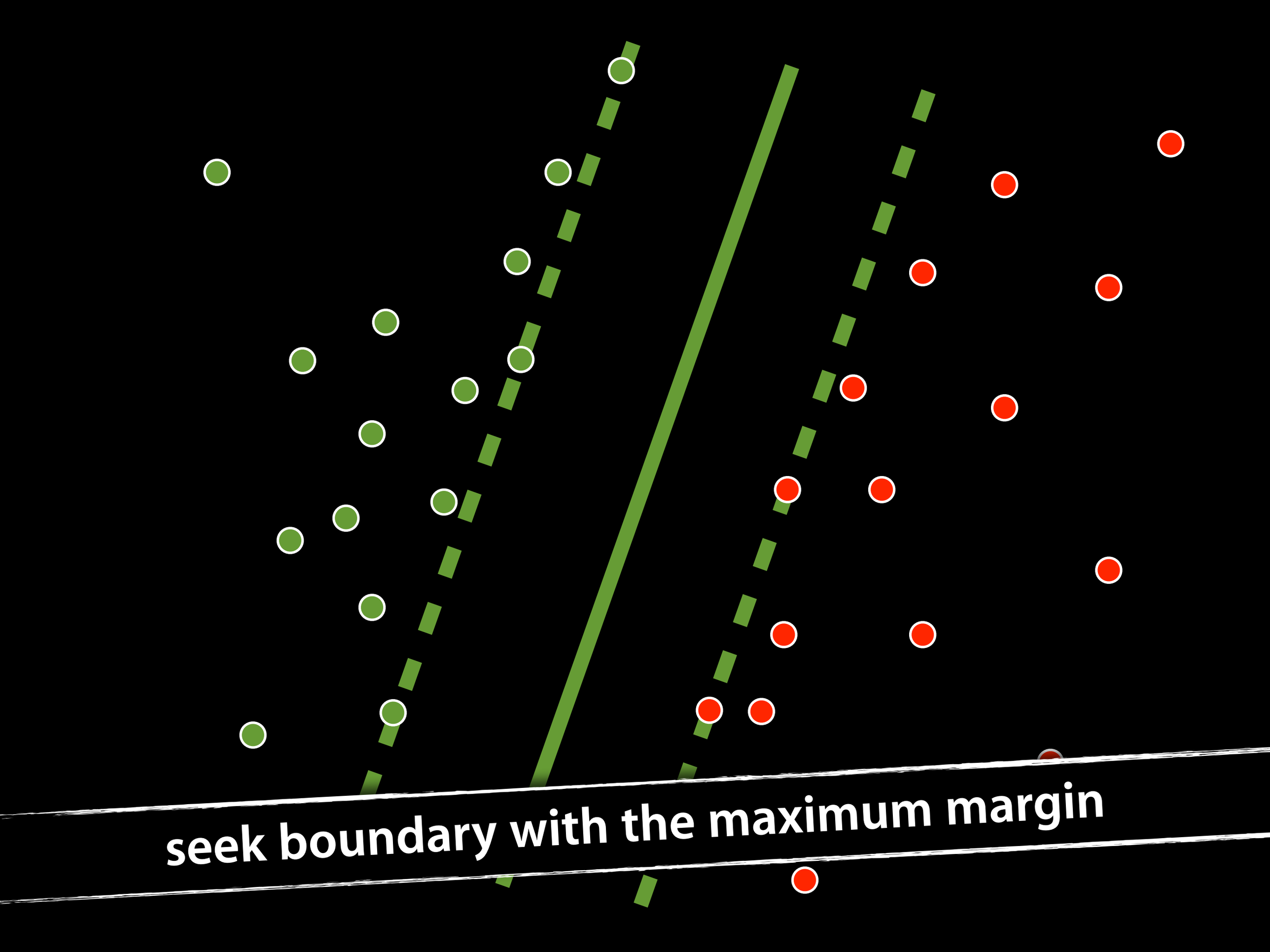


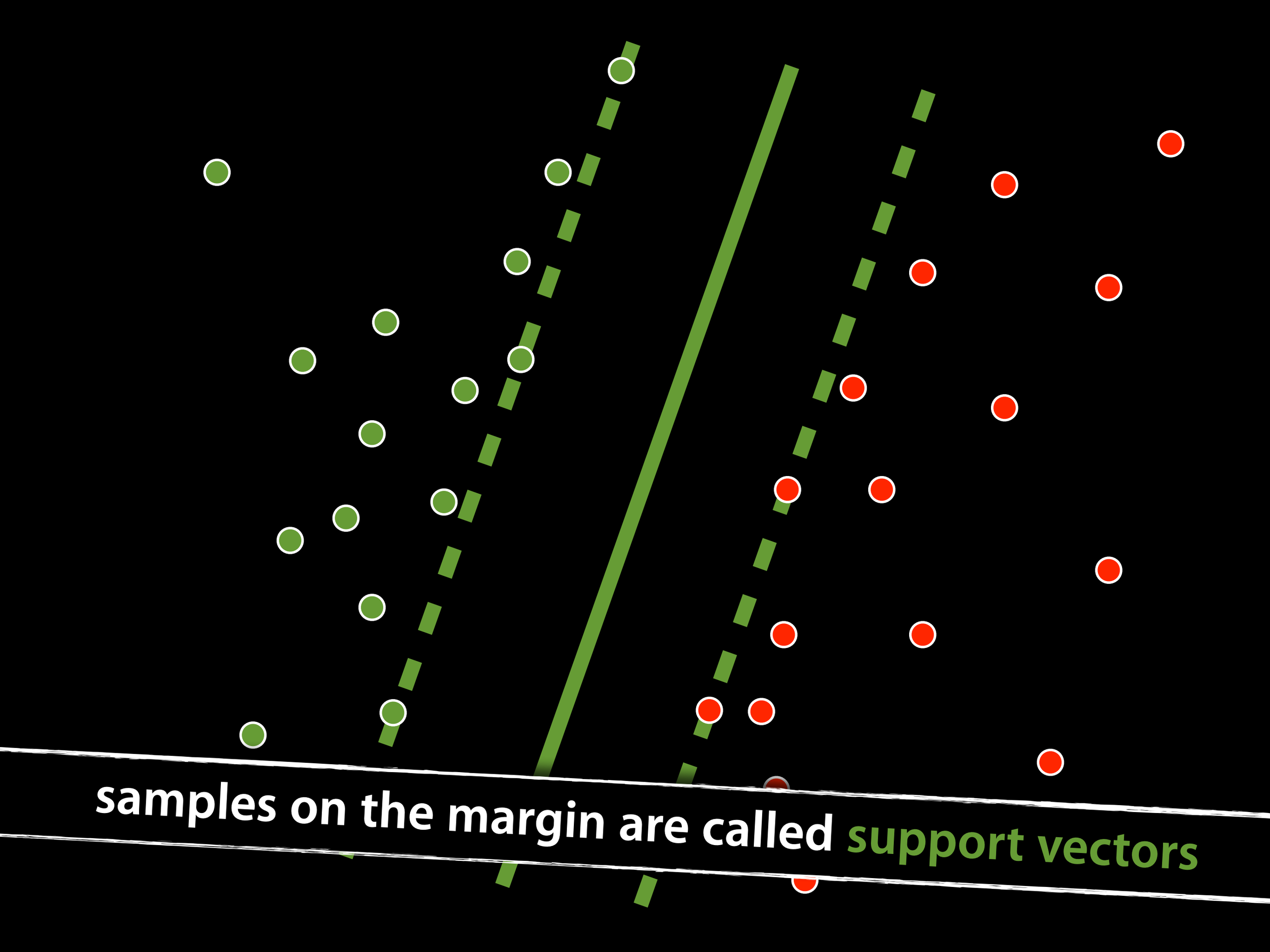
linearly separable

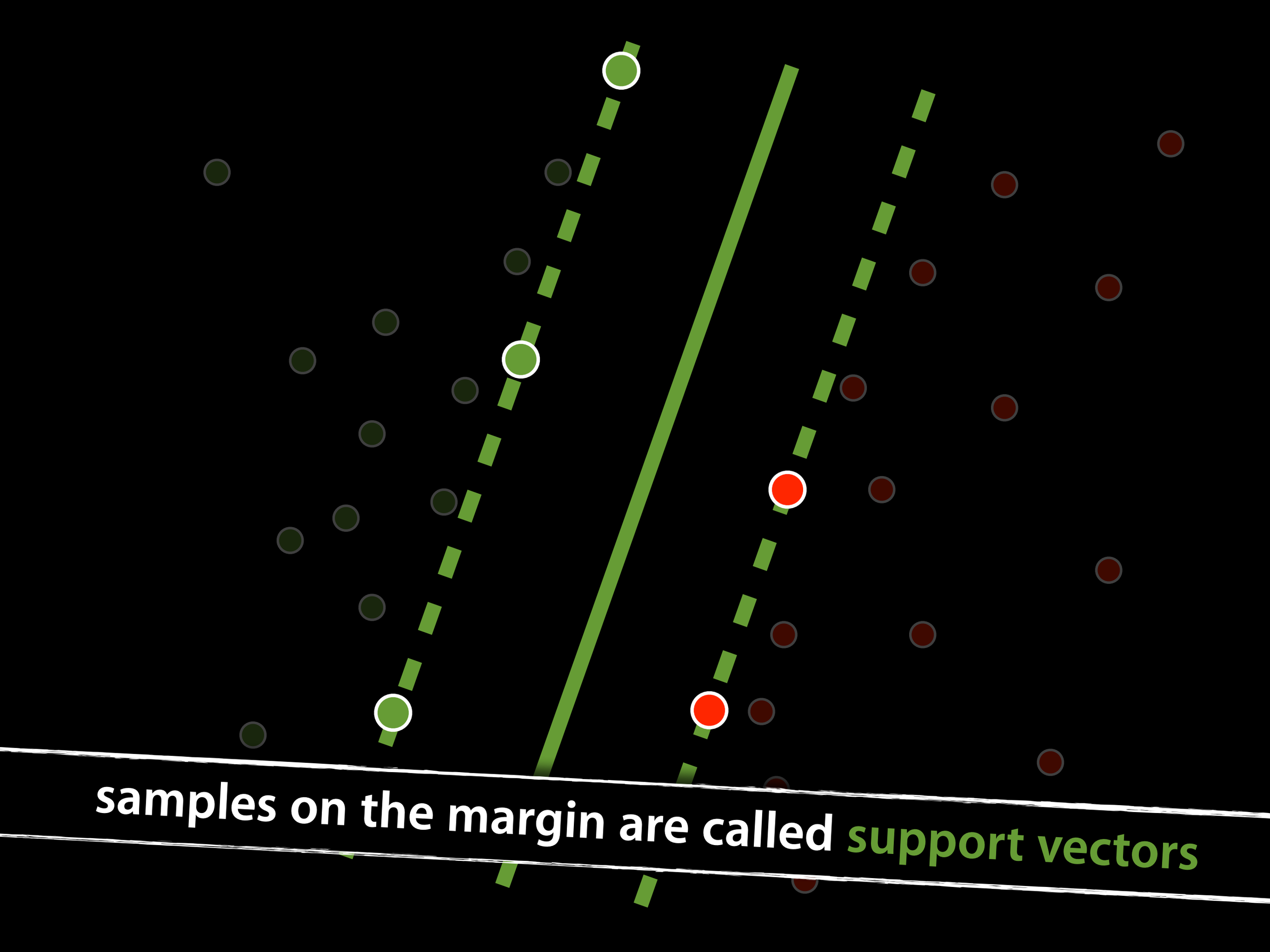



linearly separable


seek boundary with the maximum margin


seek boundary with the maximum margin


samples on the margin are called support vectors


samples on the margin are called support vectors

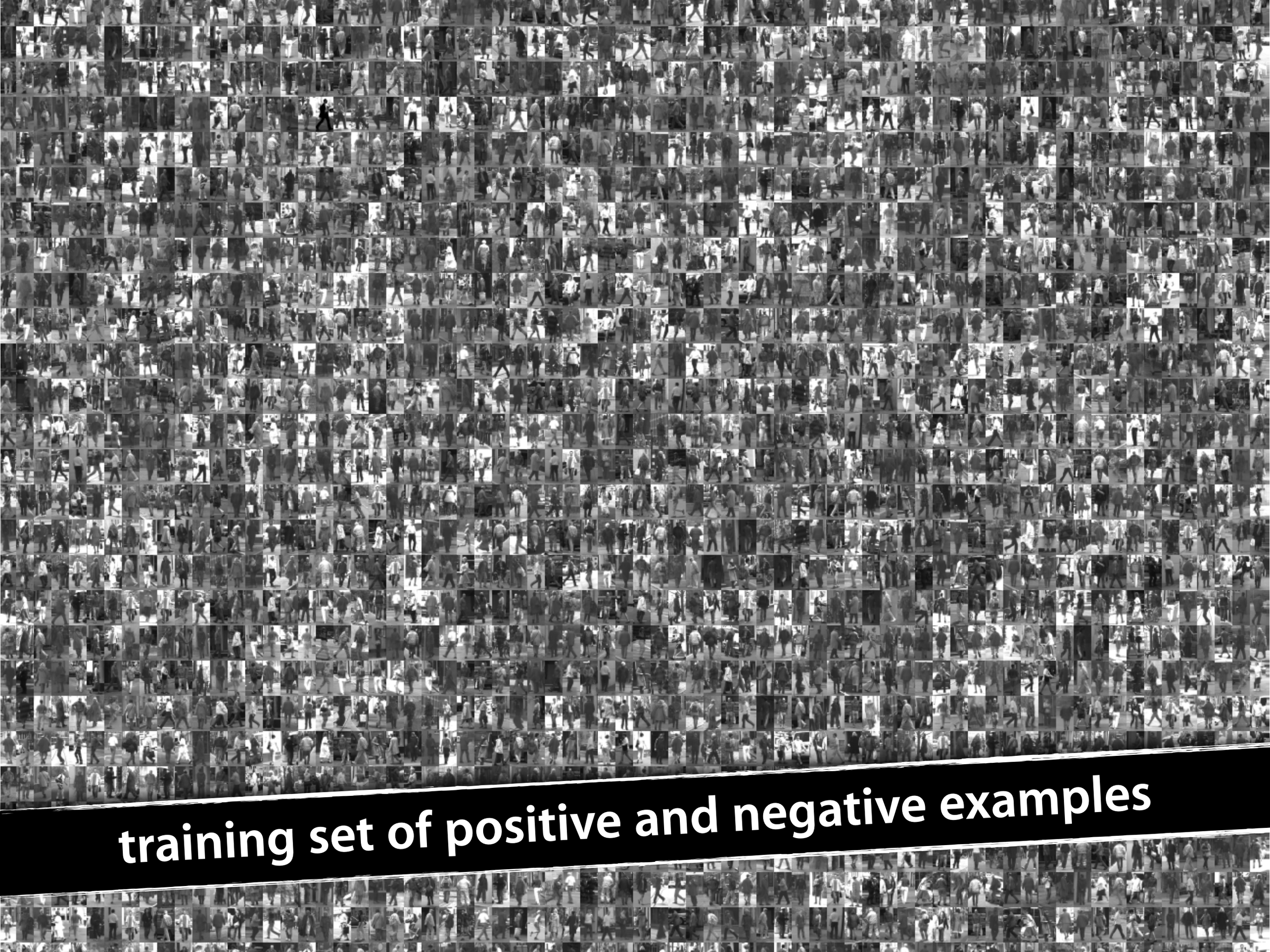
$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

L2 norm is a common regularizer

$$R(\mathbf{w}) = \|\mathbf{w}\|_2^2$$

Support Vector Machine (SVM)

Optimization


training set of positive and negative examples

$$f(\mathbf{x}; \mathbf{w}) = \mathbf{w} \cdot \mathbf{x} + b$$

scoring function

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

loss function

How do we find the optimal model parameters?

random search

random search

Assume there are 1000 binary parameters

Assume there are 1000 binary parameters

$$2^{1000} = 1.0715 \times 10^{301}$$

candidate solutions

random walk

drunkard's walk

What is the direction of greatest descent?

partial derivatives

partial derivatives

$$\frac{\partial f(x,y)}{\partial x} = \lim_{\epsilon \to 0} \frac{f(x+\epsilon, y) - f(x,y)}{\epsilon}$$

partial derivatives

$$\frac{\partial f(x,y)}{\partial x} = \lim_{\epsilon \to 0} \frac{f(x+\epsilon, y) - f(x,y)}{\epsilon}$$

$$\frac{\partial f(x,y)}{\partial y} = \lim_{\epsilon \to 0} \frac{f(x, y+\epsilon) - f(x,y)}{\epsilon}$$

gradient

gradient

$$\nabla f(x_1, x_2, \ldots, x_N)$$

gradient

$$\nabla f(x_1, x_2, \ldots, x_N) = \begin{pmatrix} \dfrac{\partial f}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} \\ \vdots \\ \dfrac{\partial f}{\partial x_N} \end{pmatrix}$$

gradient

$$\nabla f(x_1, x_2, \ldots, x_N) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_N} \end{pmatrix}$$

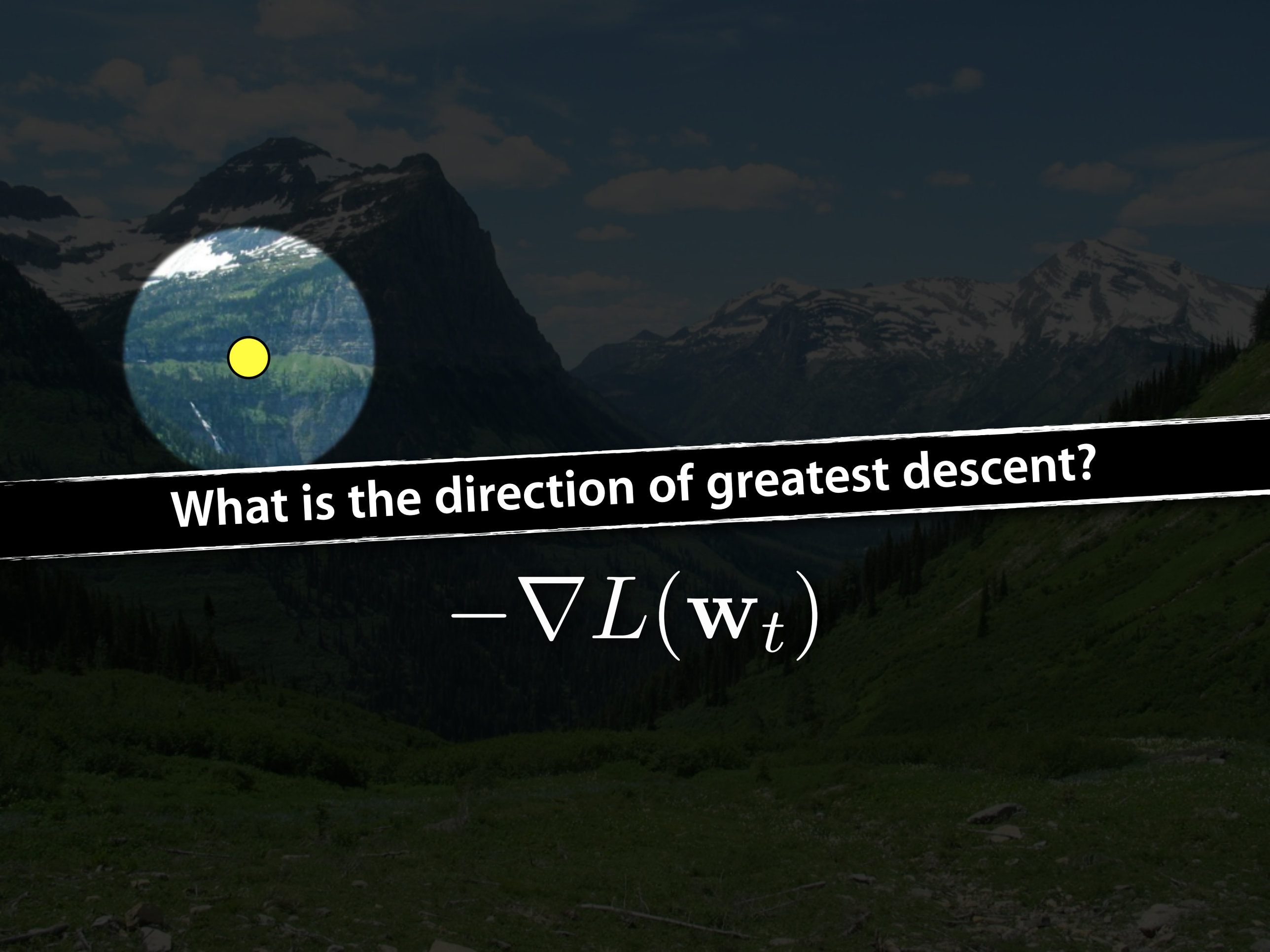
points in the direction of greatest rate of increase

gradient

$$\nabla f(x_1, x_2, \ldots, x_N) = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_N} \end{pmatrix}$$

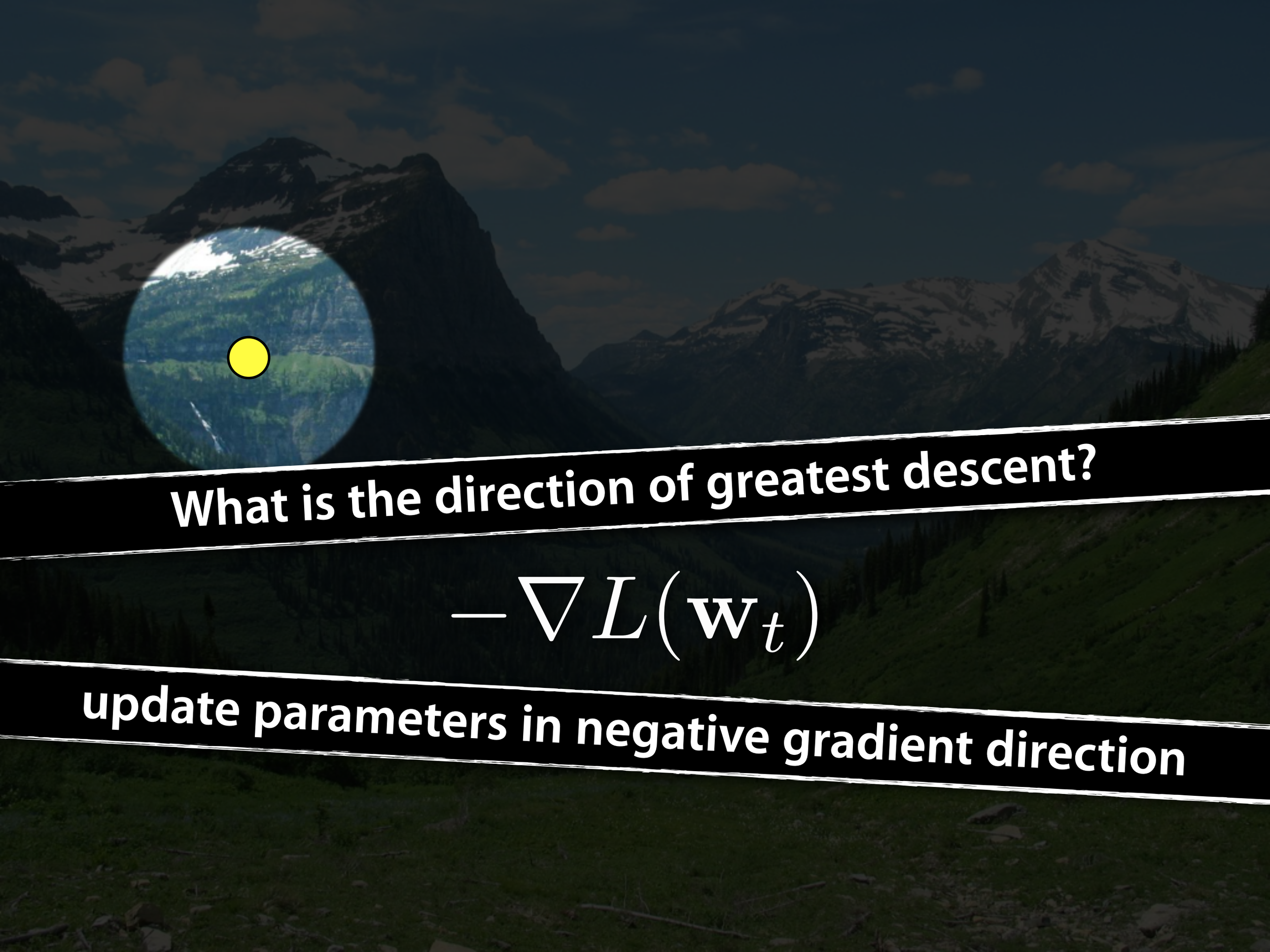
magnitude is the rate of increase

What is the direction of greatest descent?



What is the direction of greatest descent?

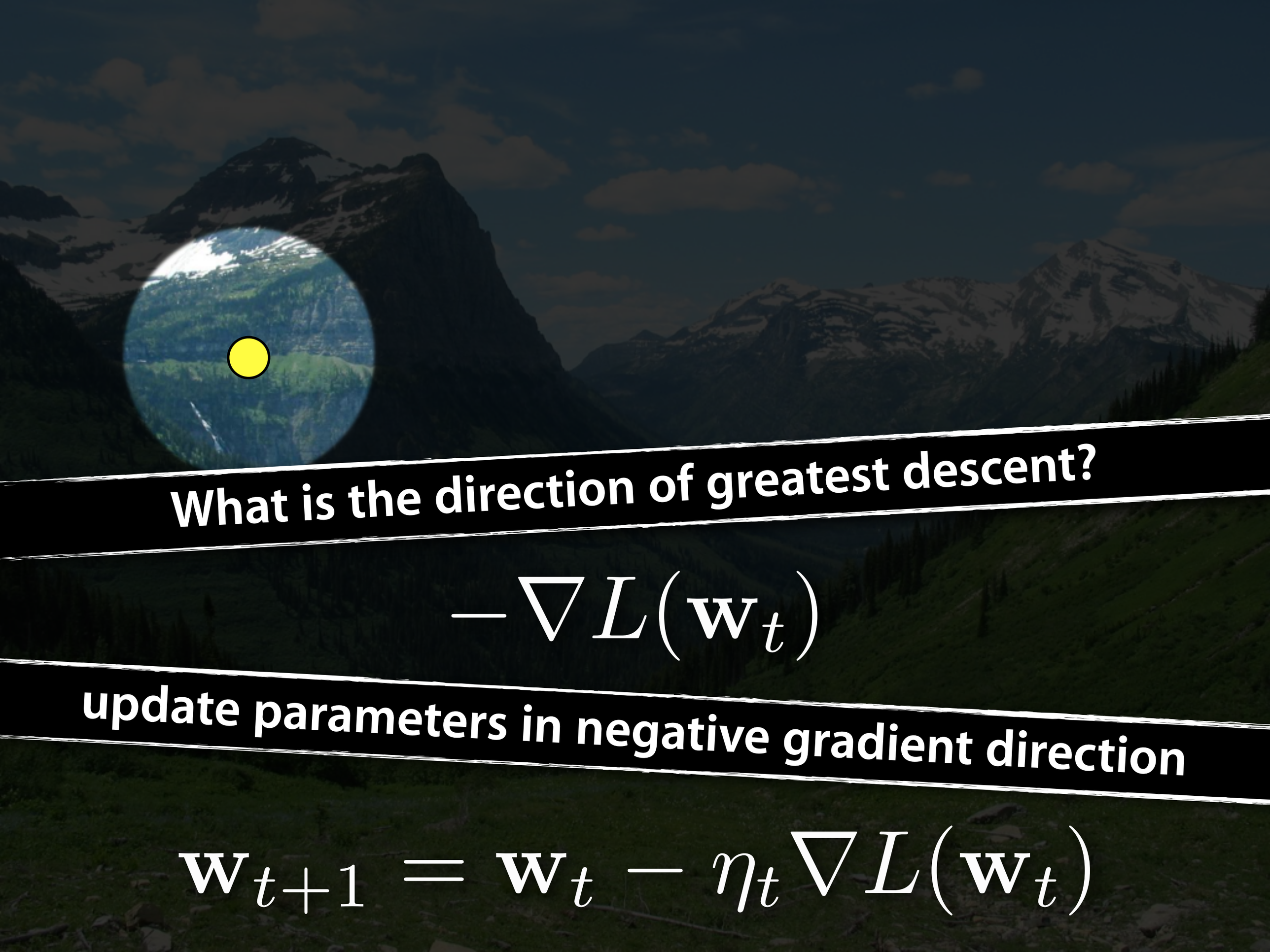
$$-\nabla L(\mathbf{w}_t)$$



What is the direction of greatest descent?

$$-\nabla L(\mathbf{w}_t)$$

update parameters in negative gradient direction

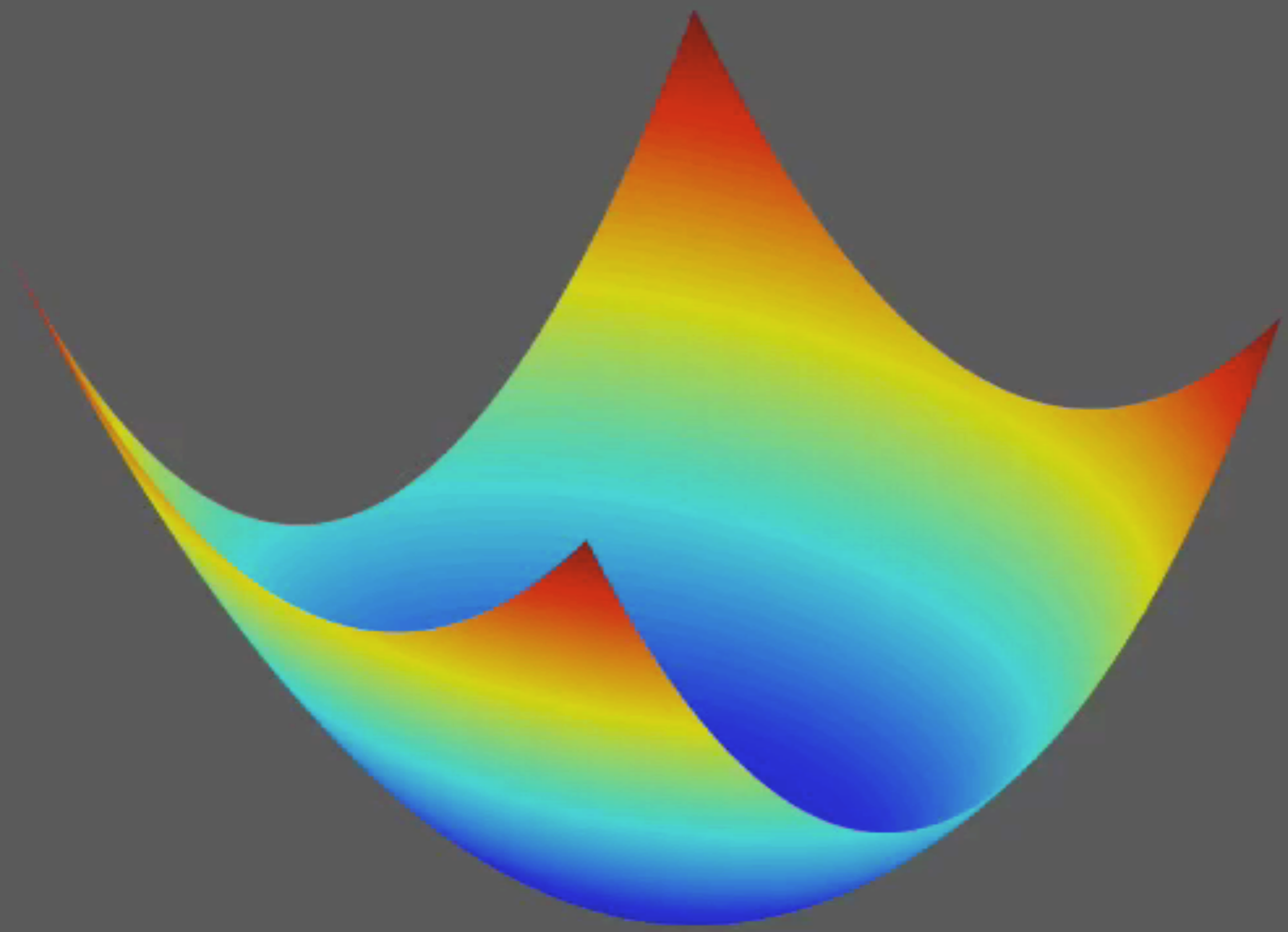


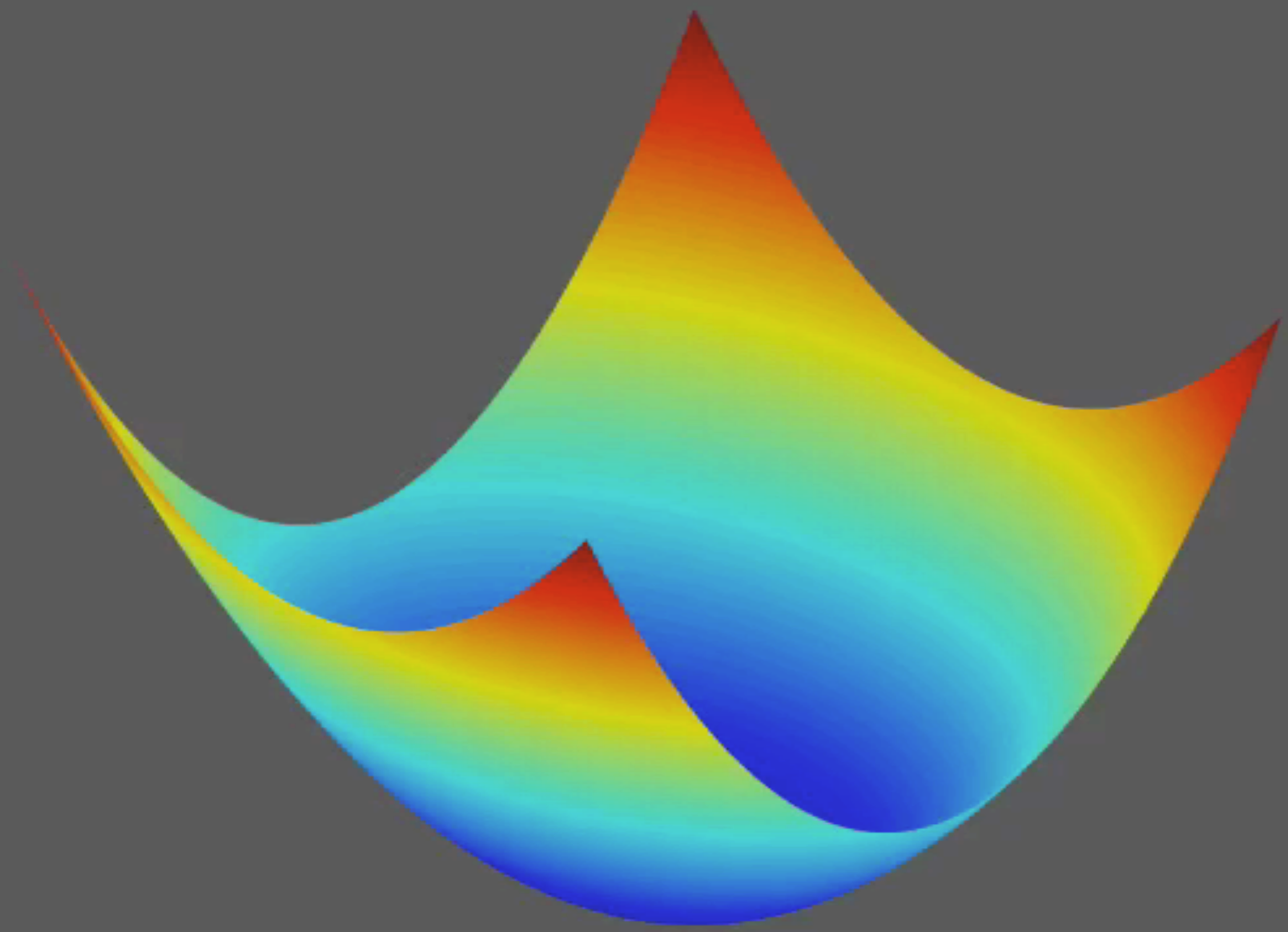
What is the direction of greatest descent?

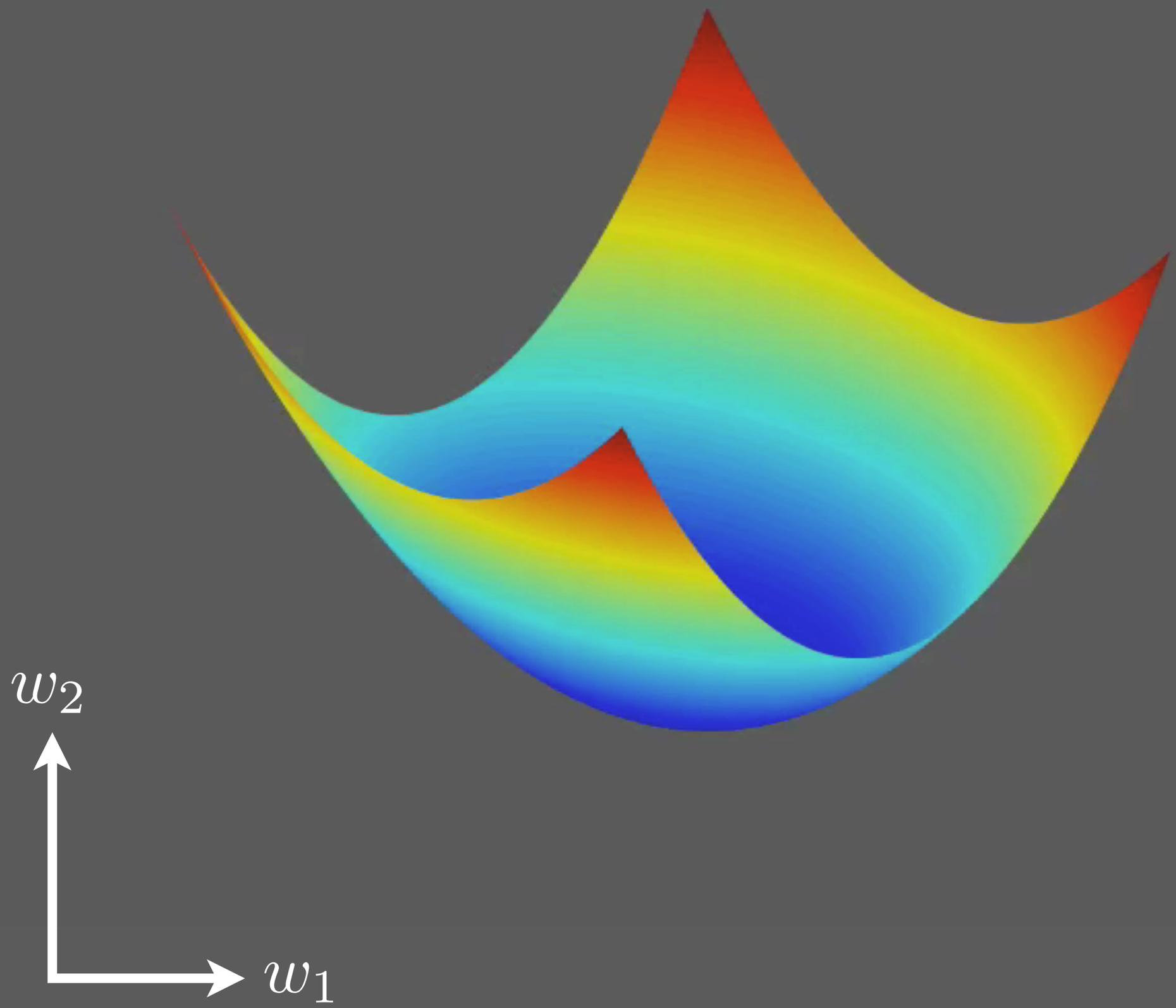
$$-\nabla L(\mathbf{w}_t)$$

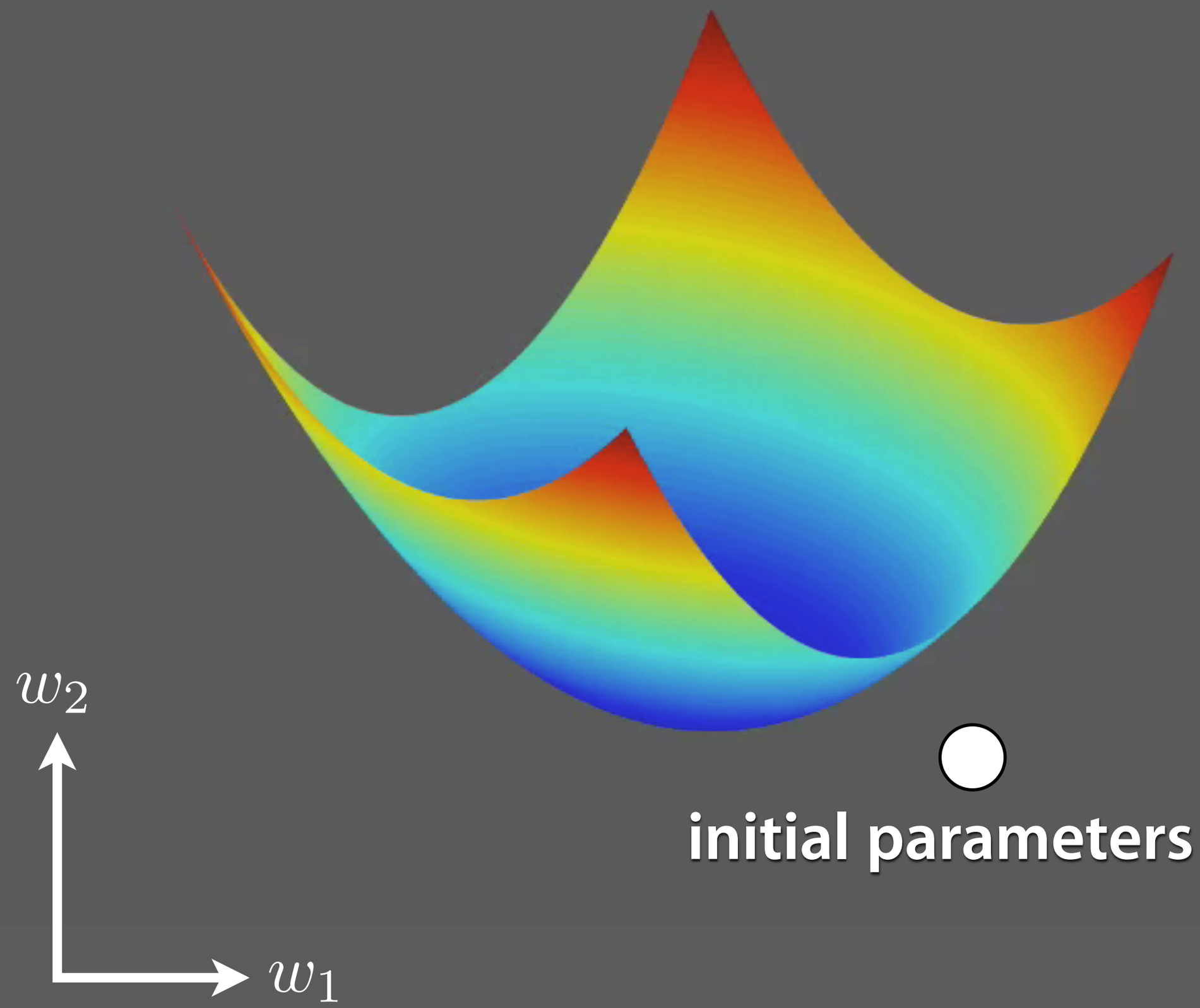
update parameters in negative gradient direction

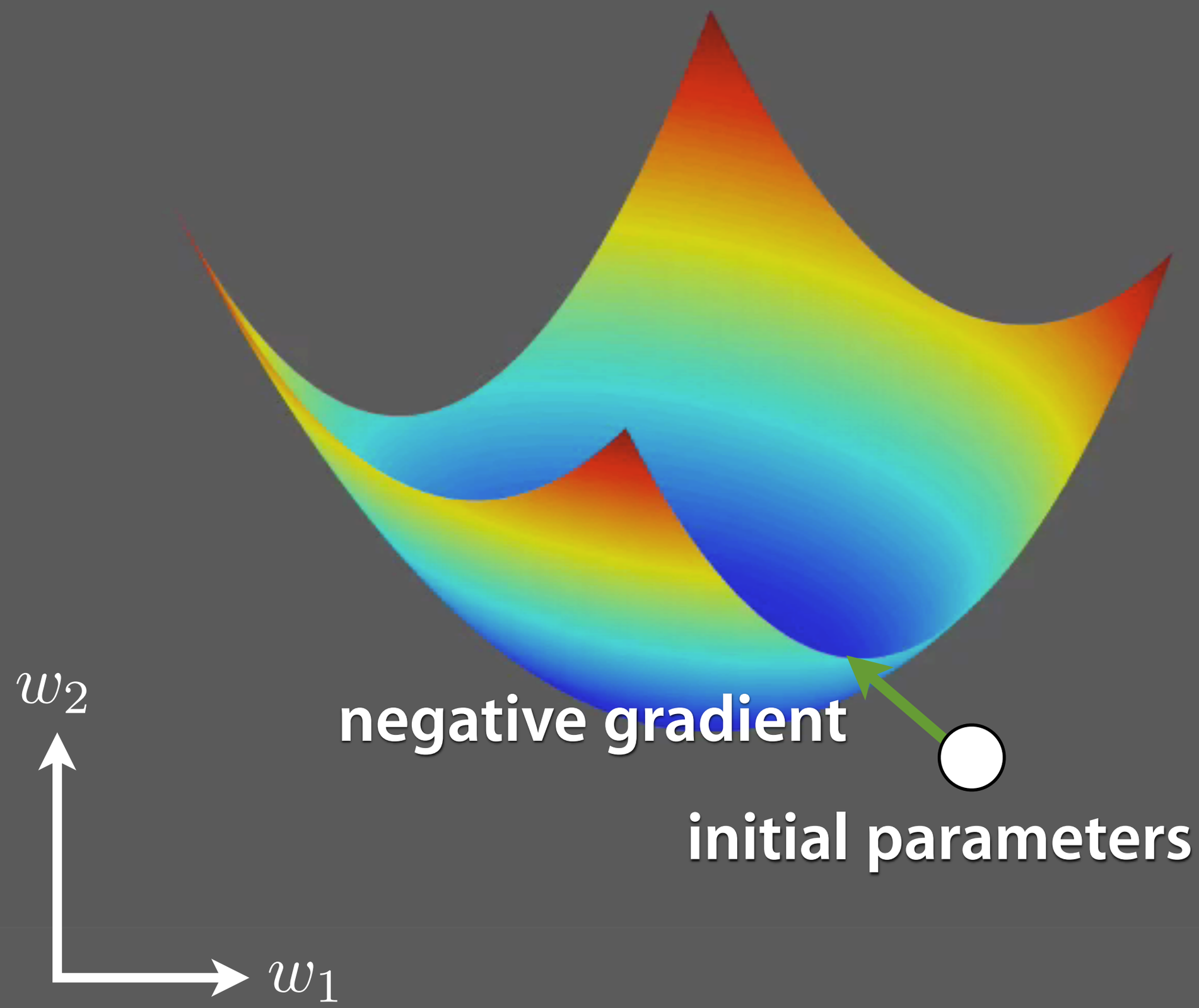
$$\mathbf{w}_{t+1} = \mathbf{w}_t - \eta_t \nabla L(\mathbf{w}_t)$$

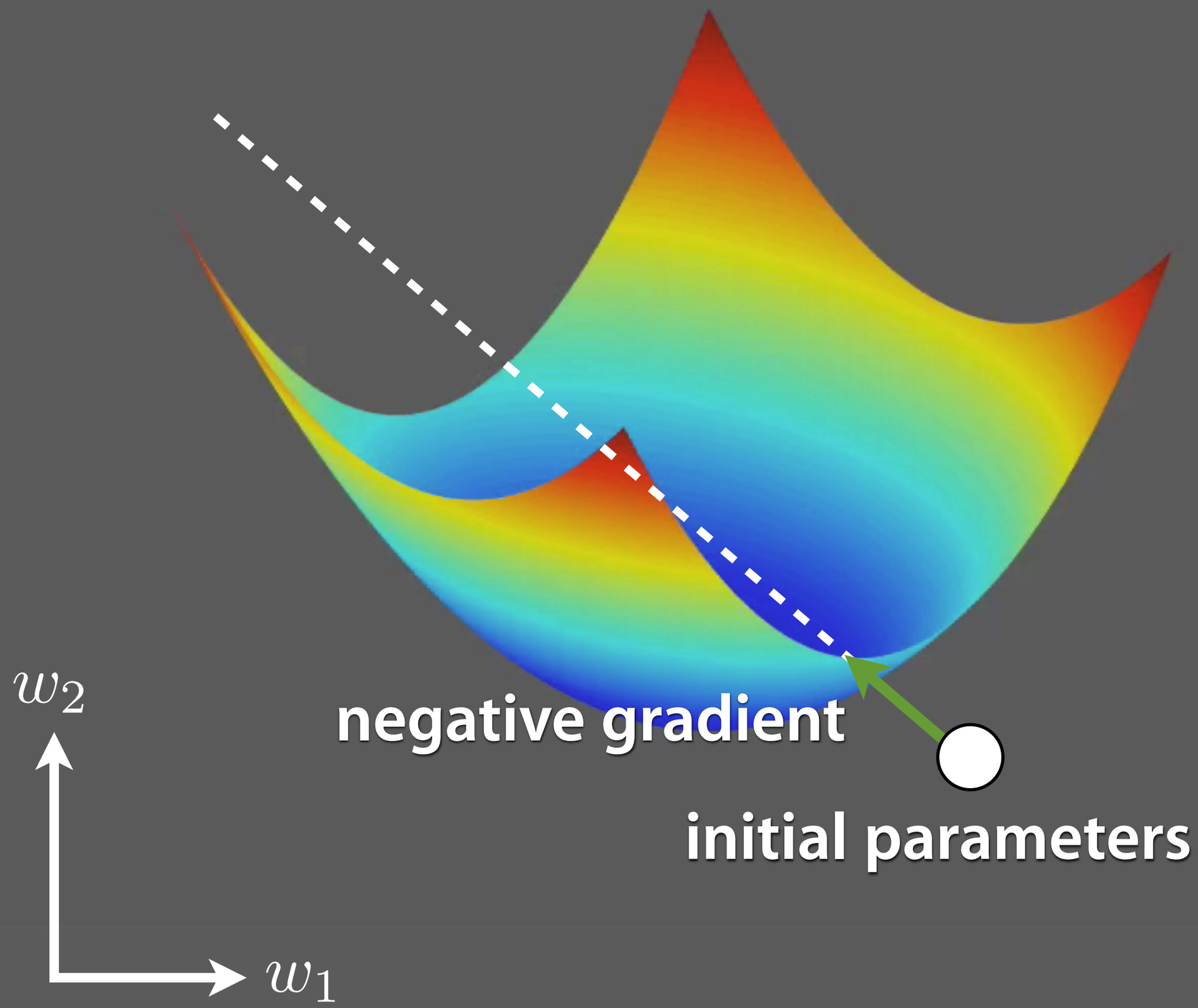


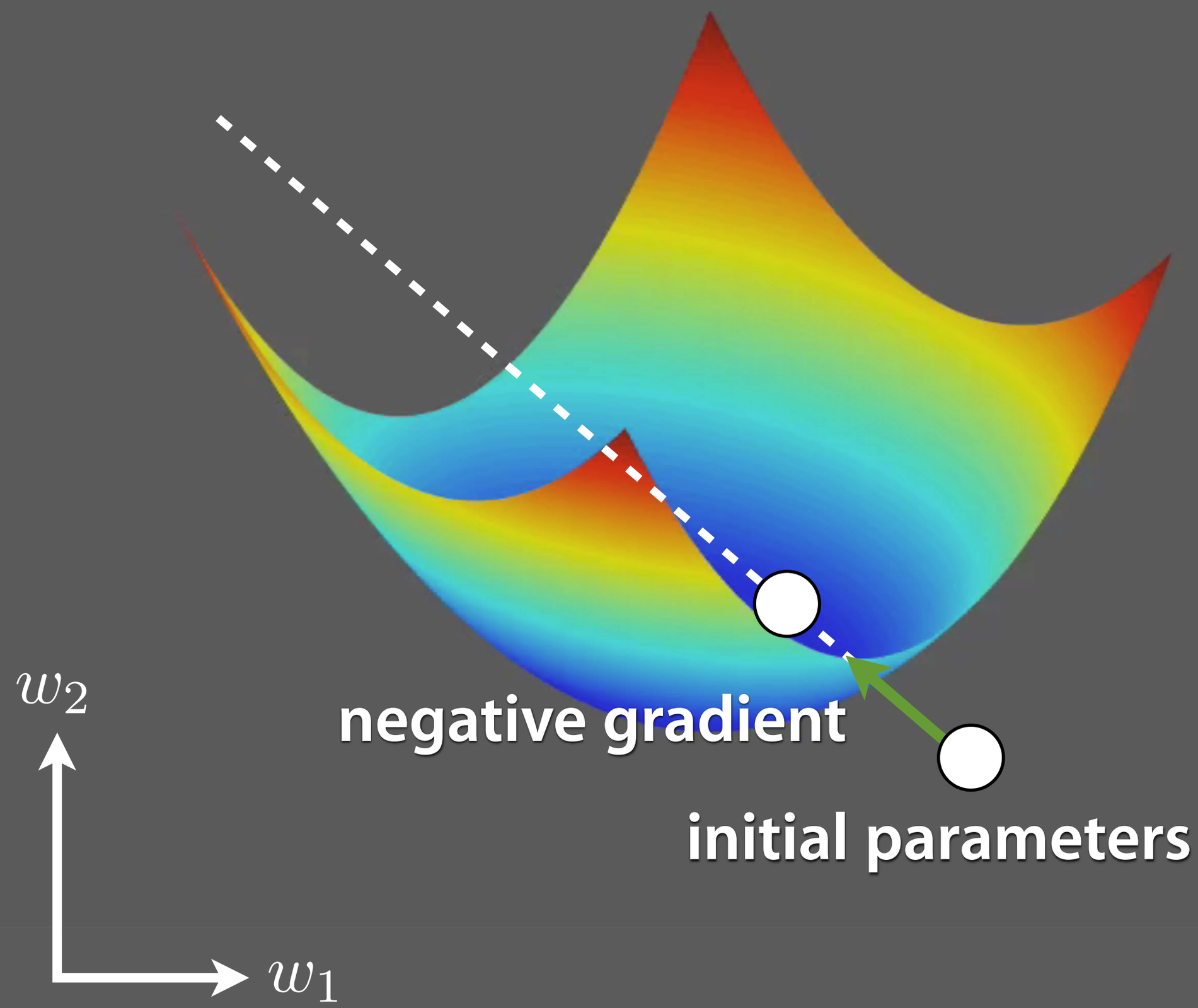


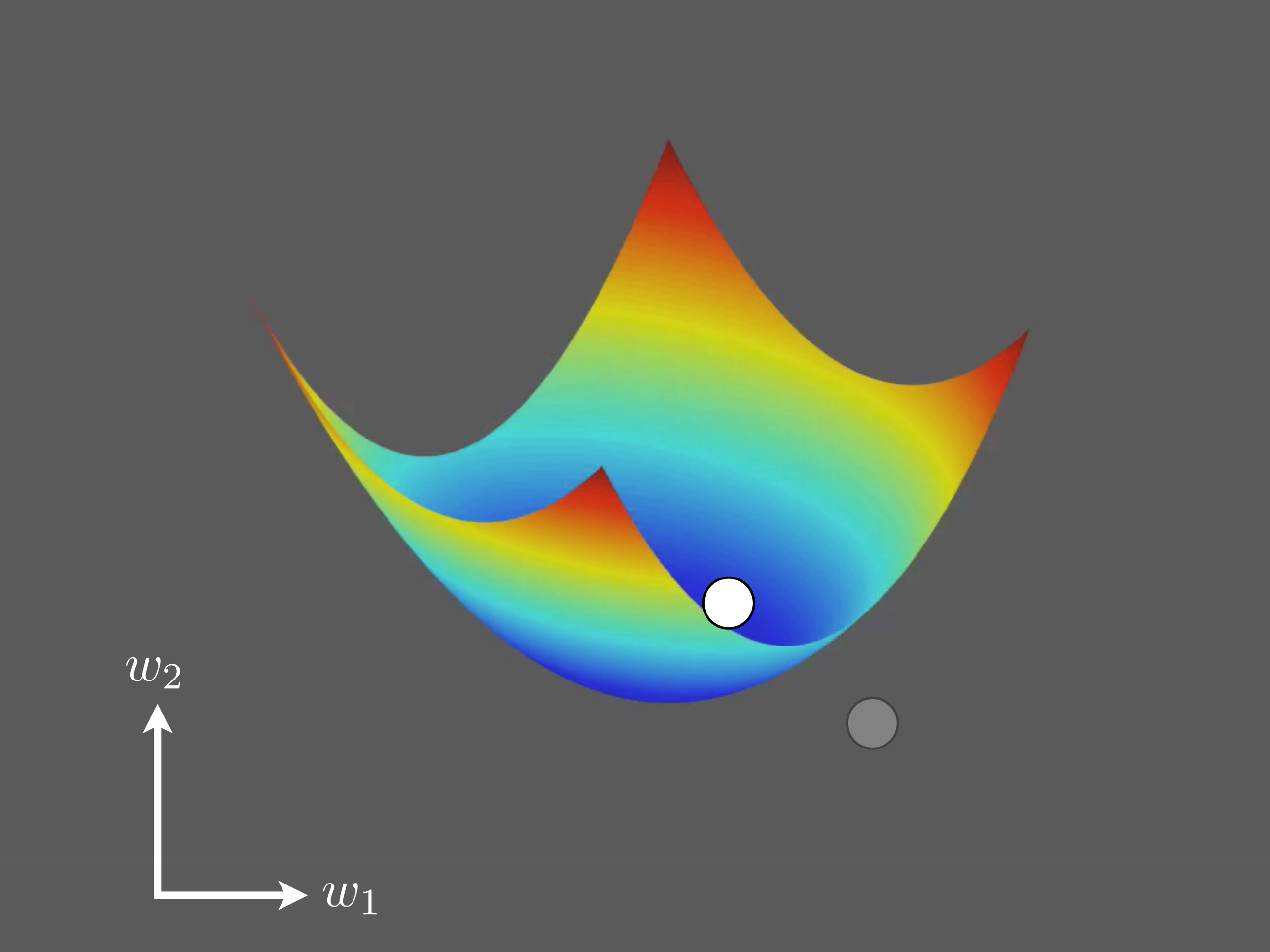

w_2
w_1

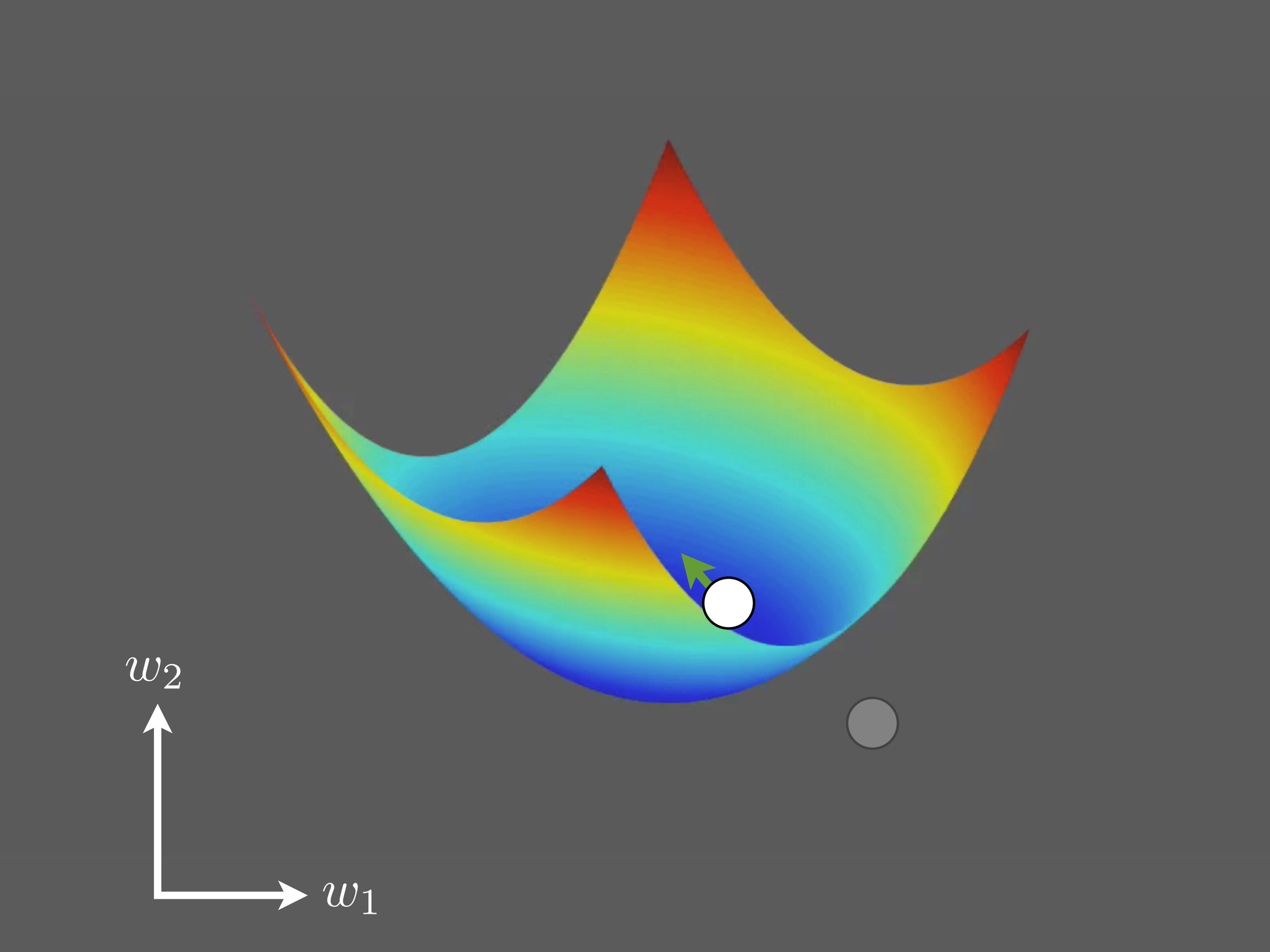

w_2
w_1
initial parameters

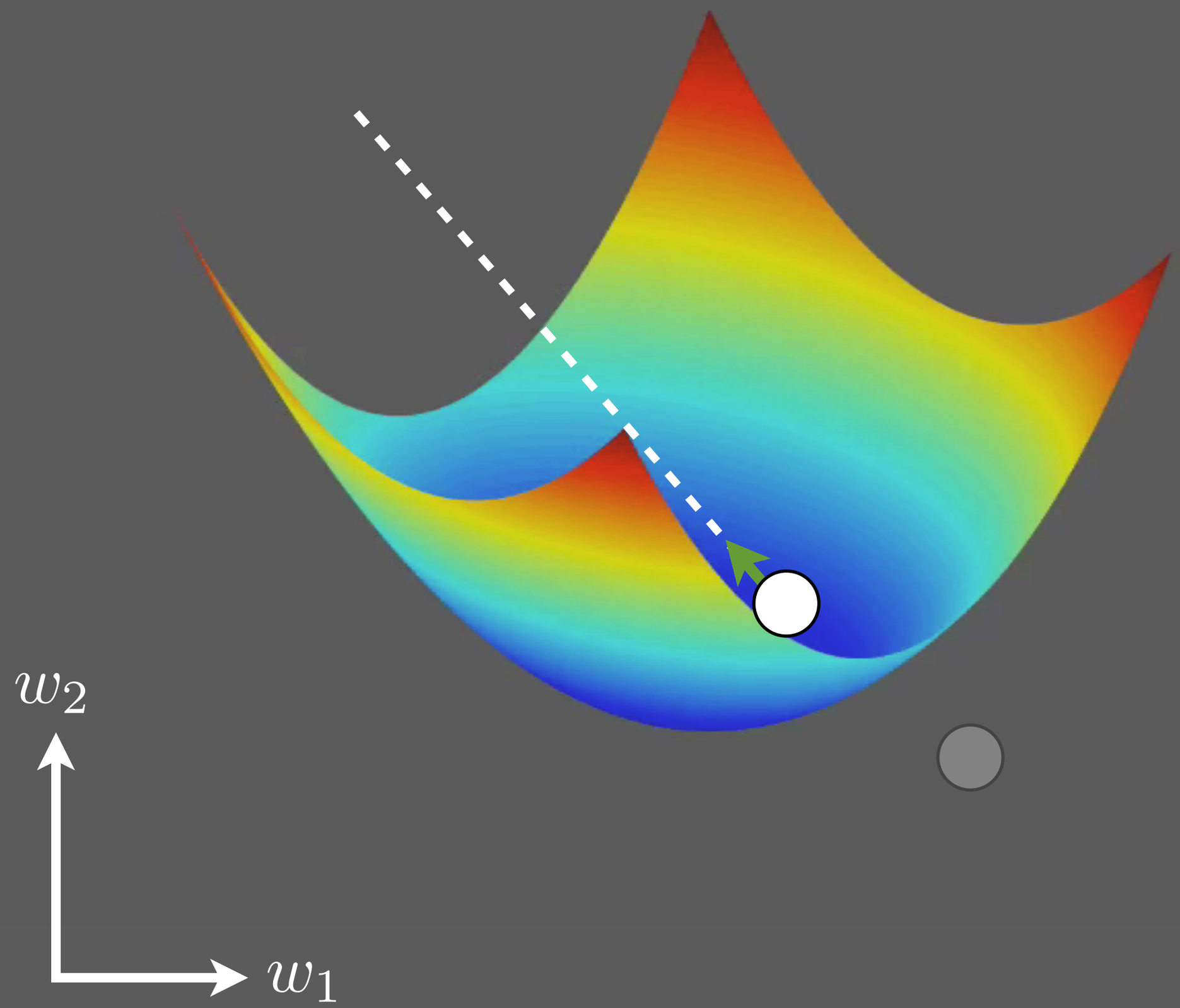

negative gradient
initial parameters
w_2
w_1

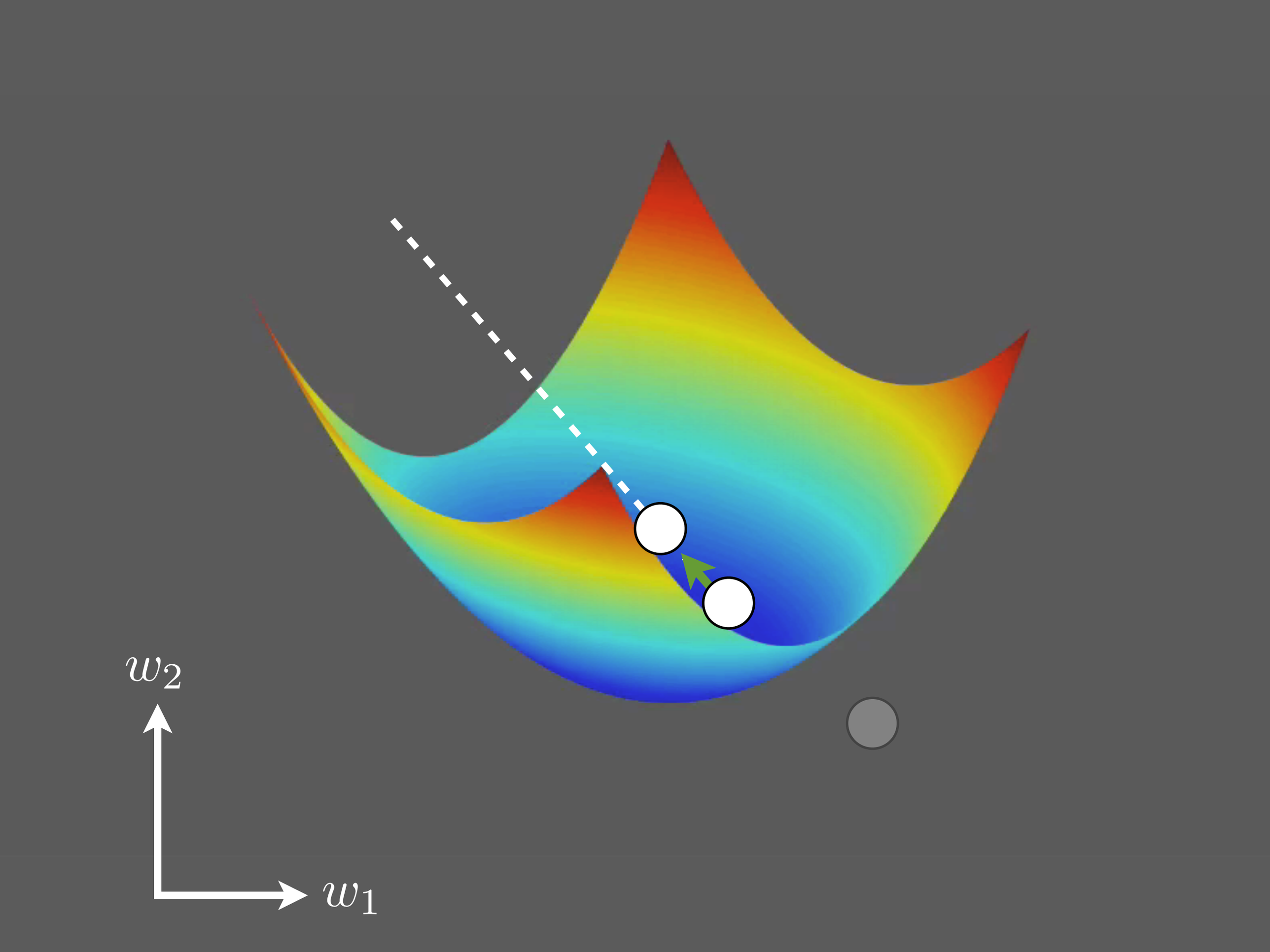

negative gradient
initial parameters
w_2
w_1

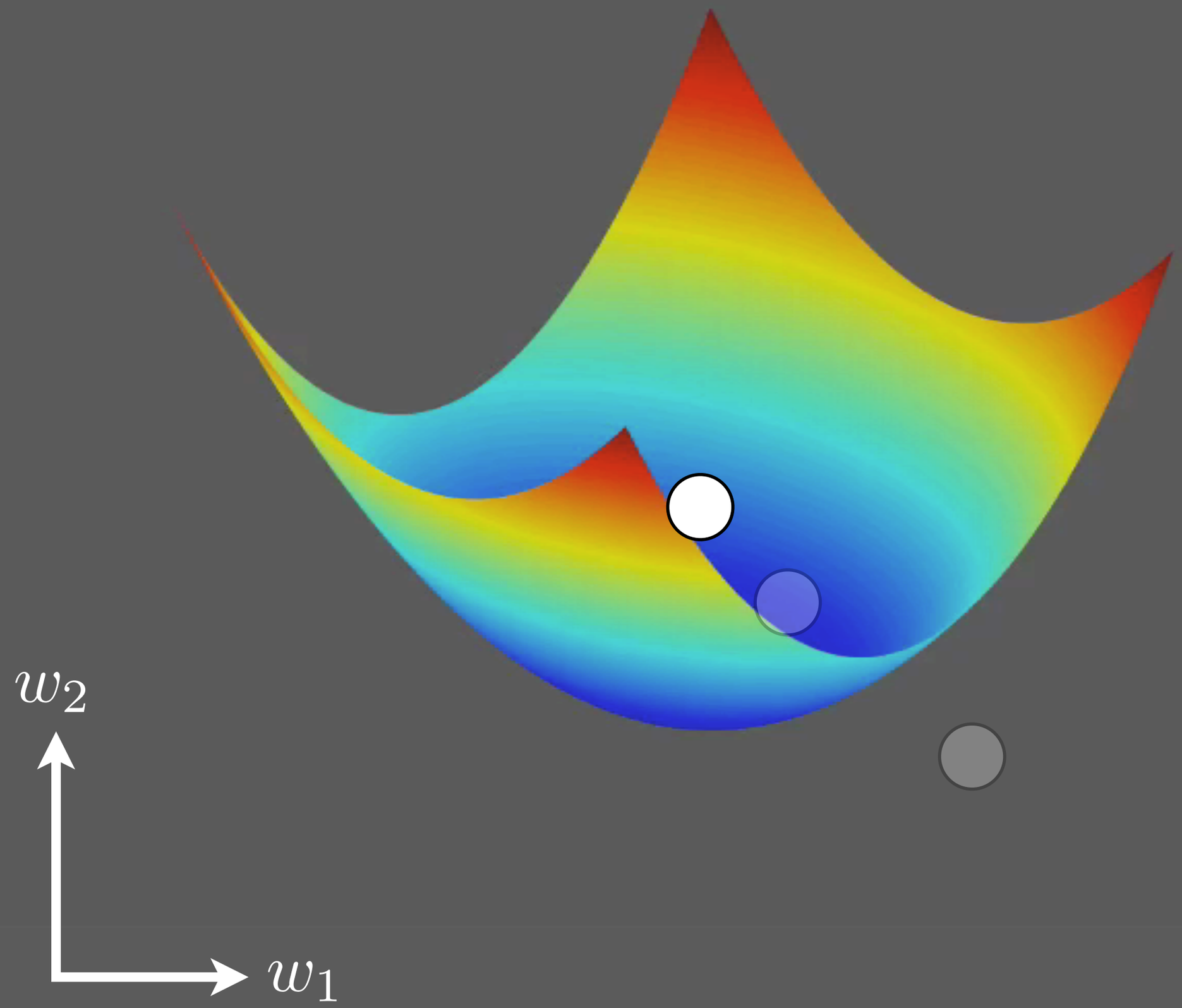

negative gradient
initial parameters
w_2
w_1

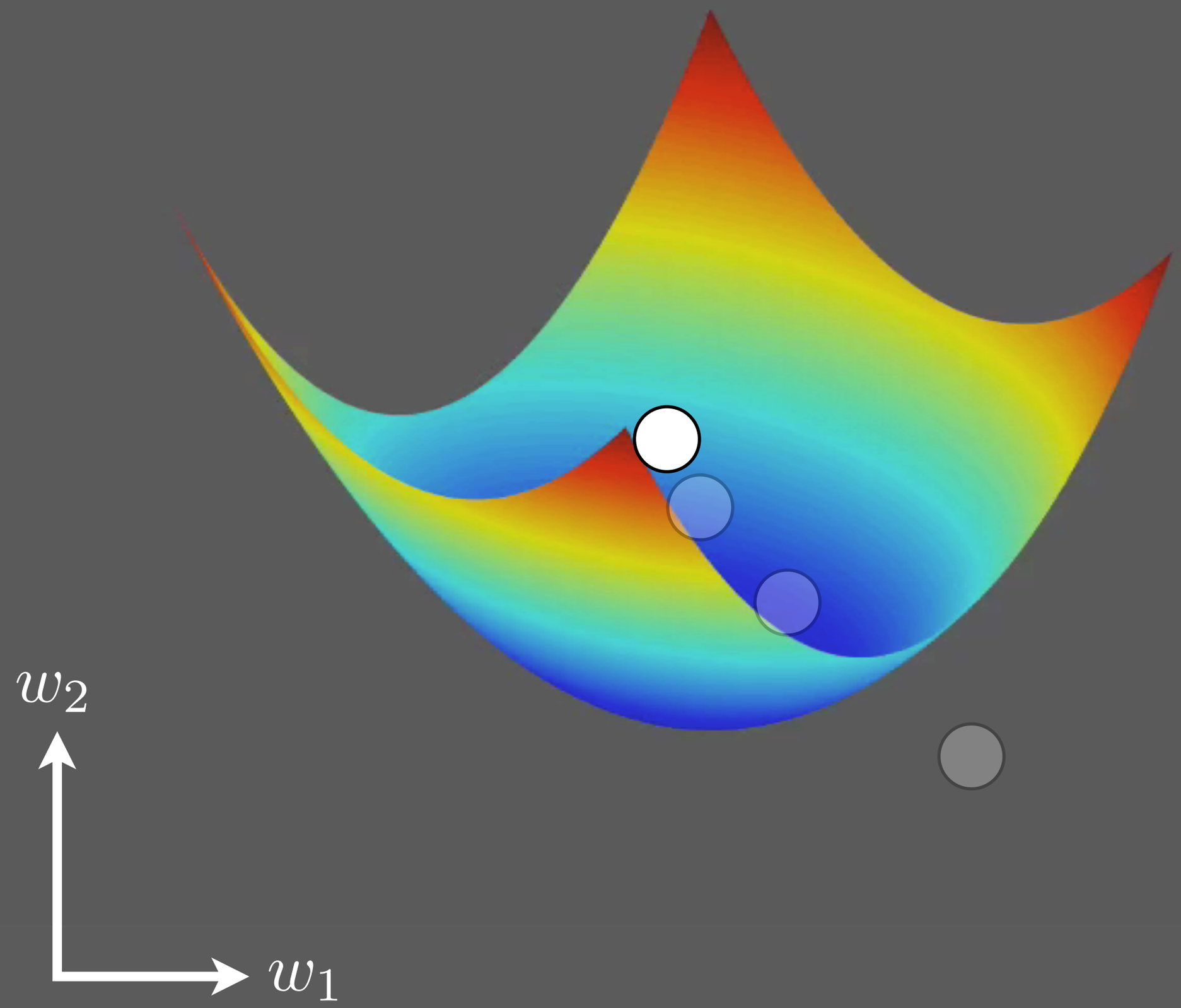

w_2
w_1

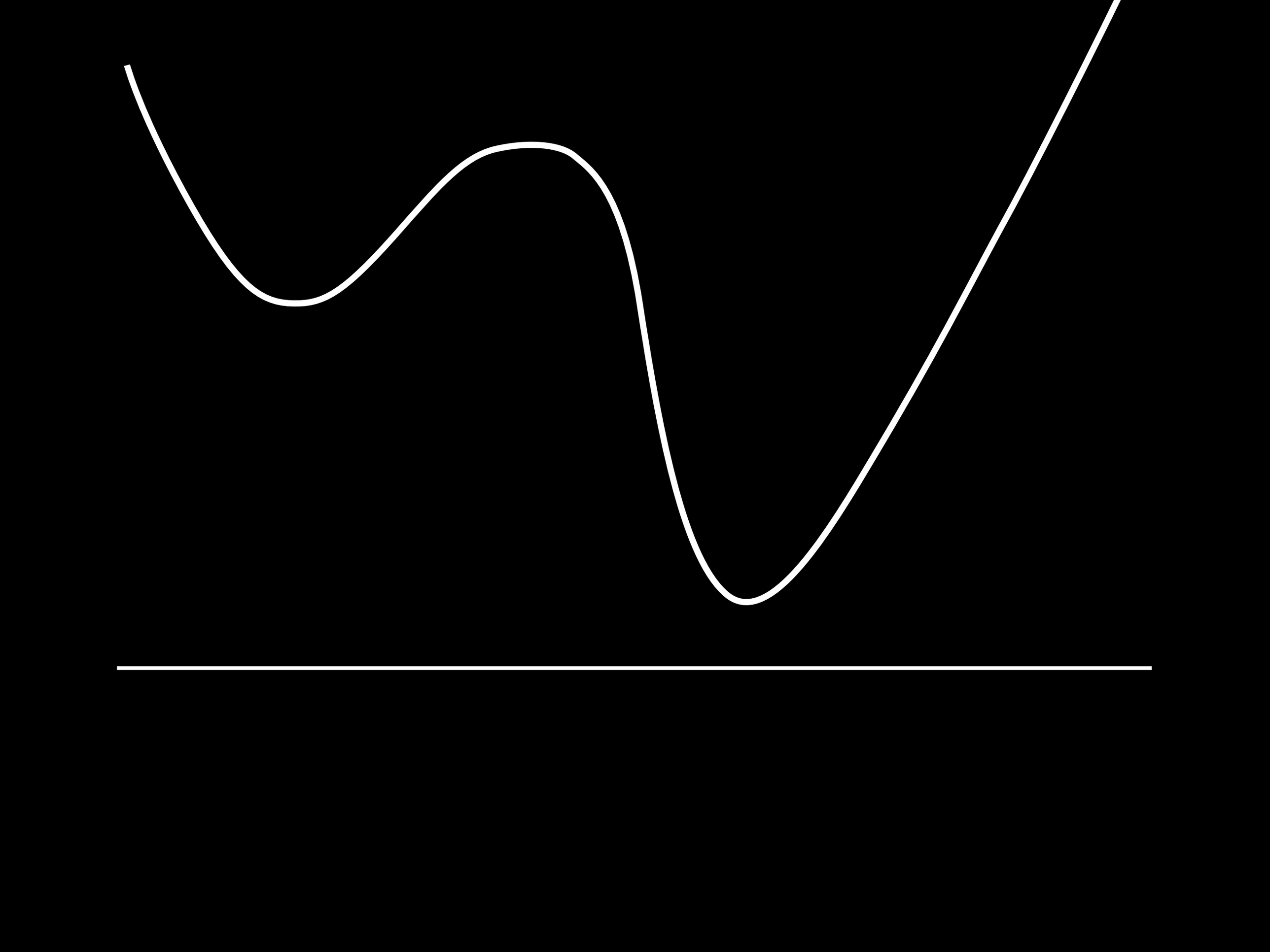

w_2
w_1

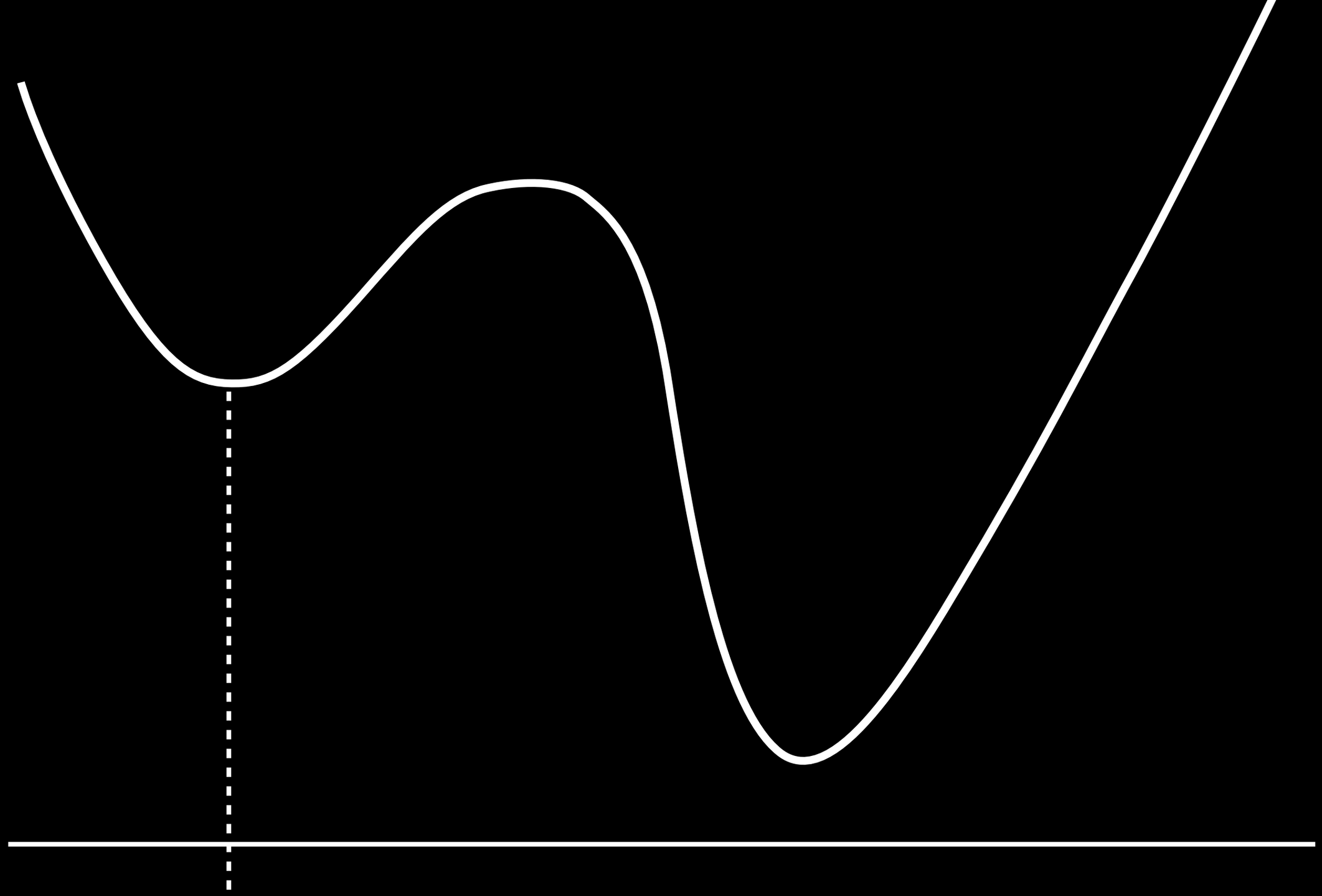

w_2
w_1


w_2
w_1

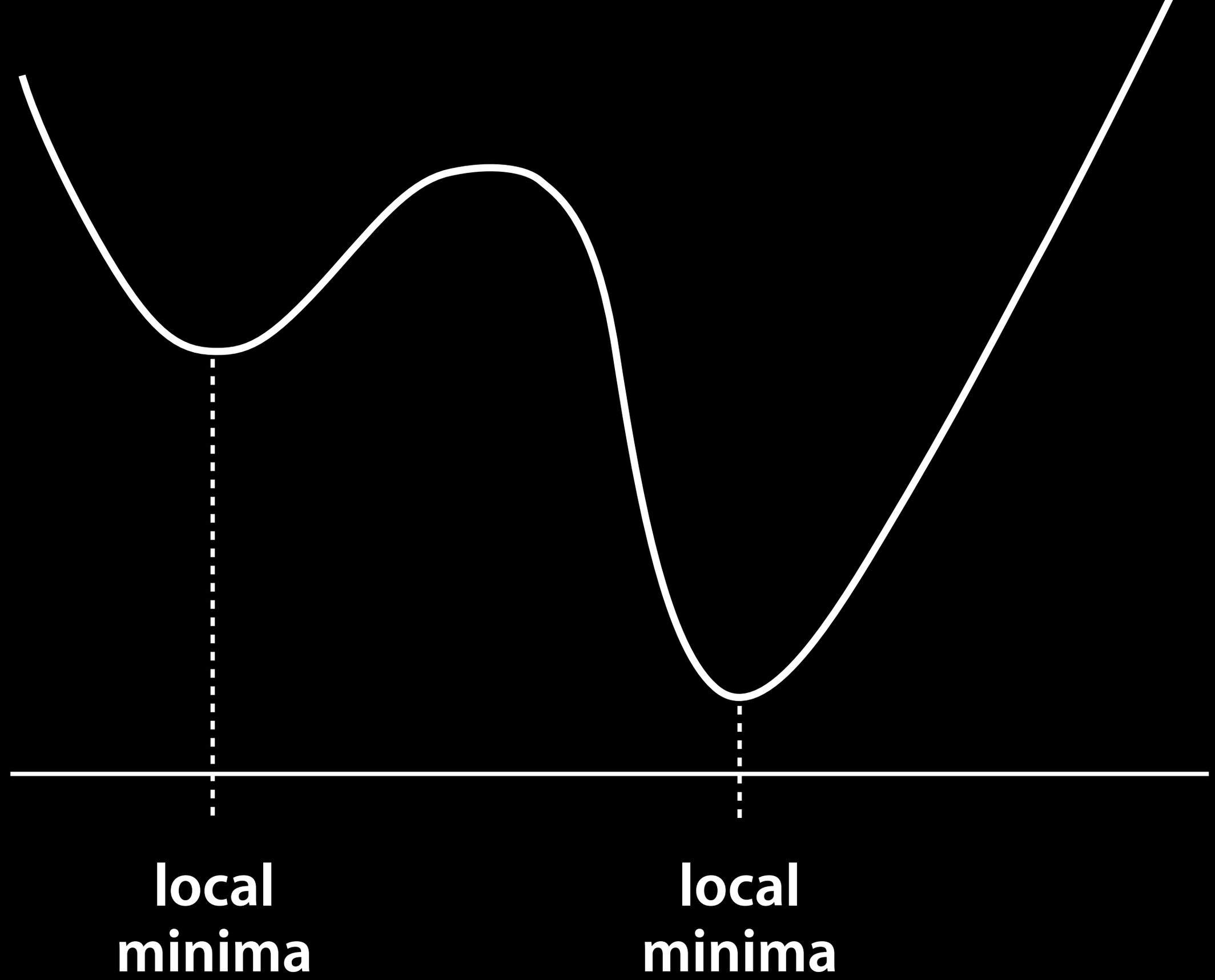

w_2
w_1

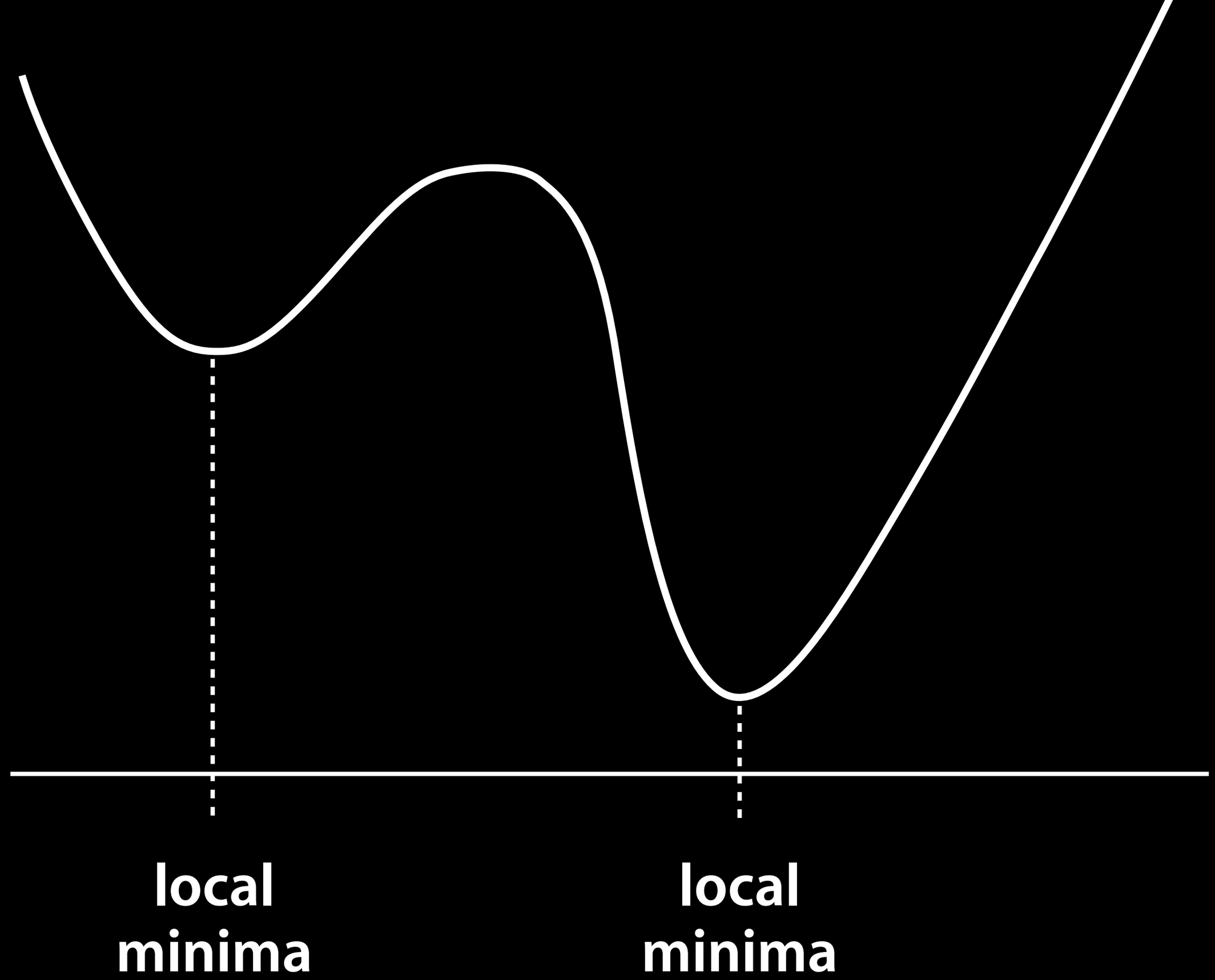

w_2
w_1

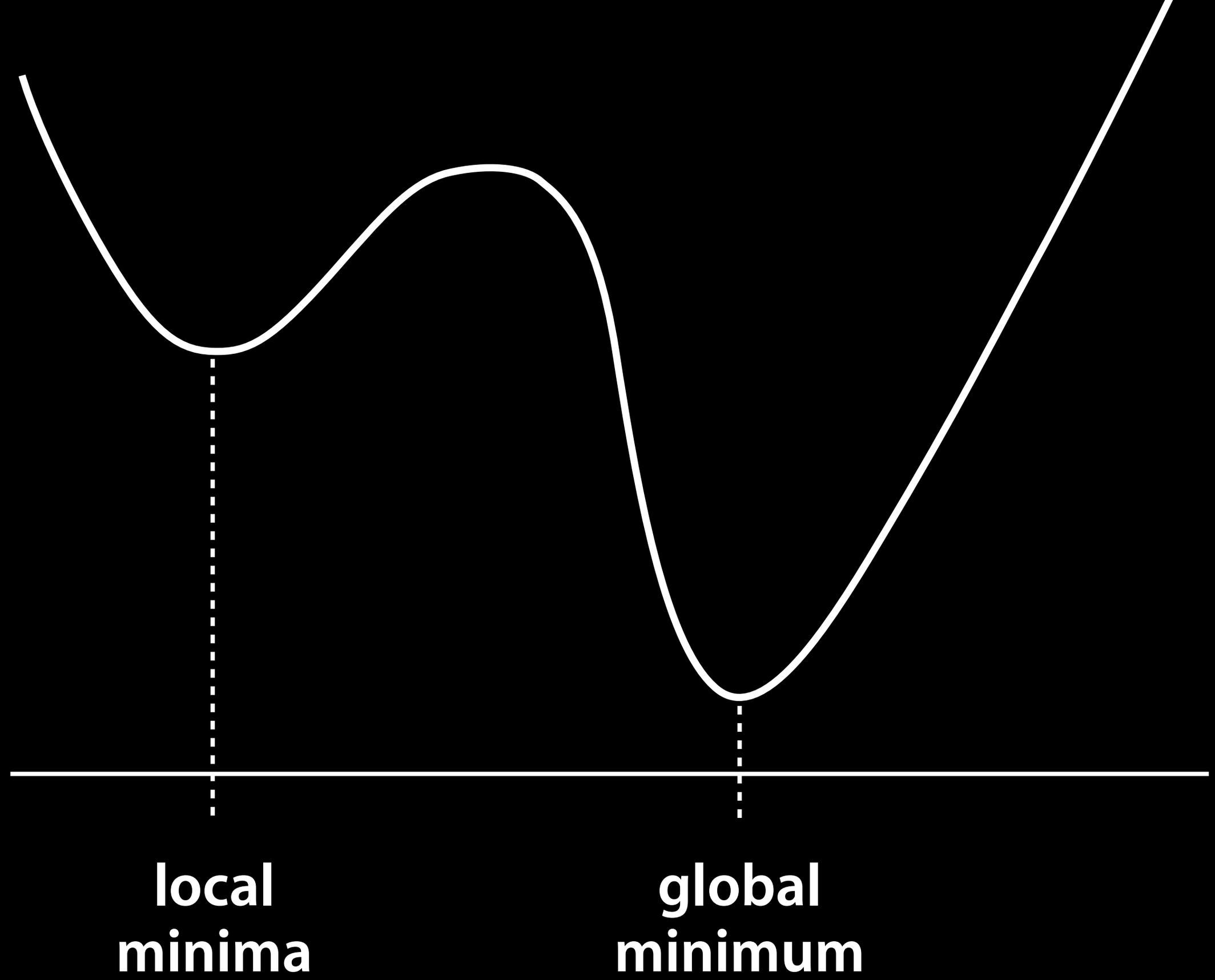


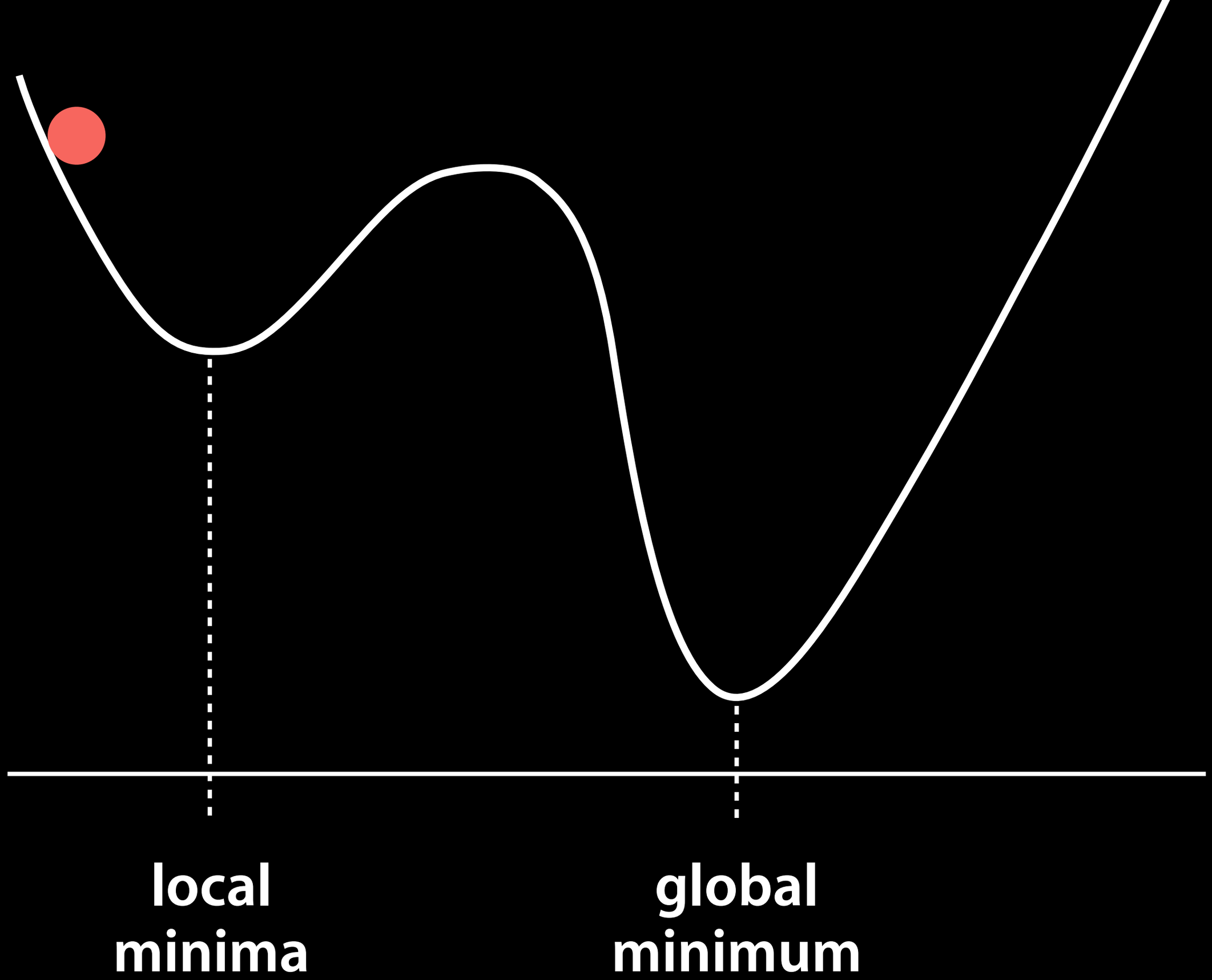


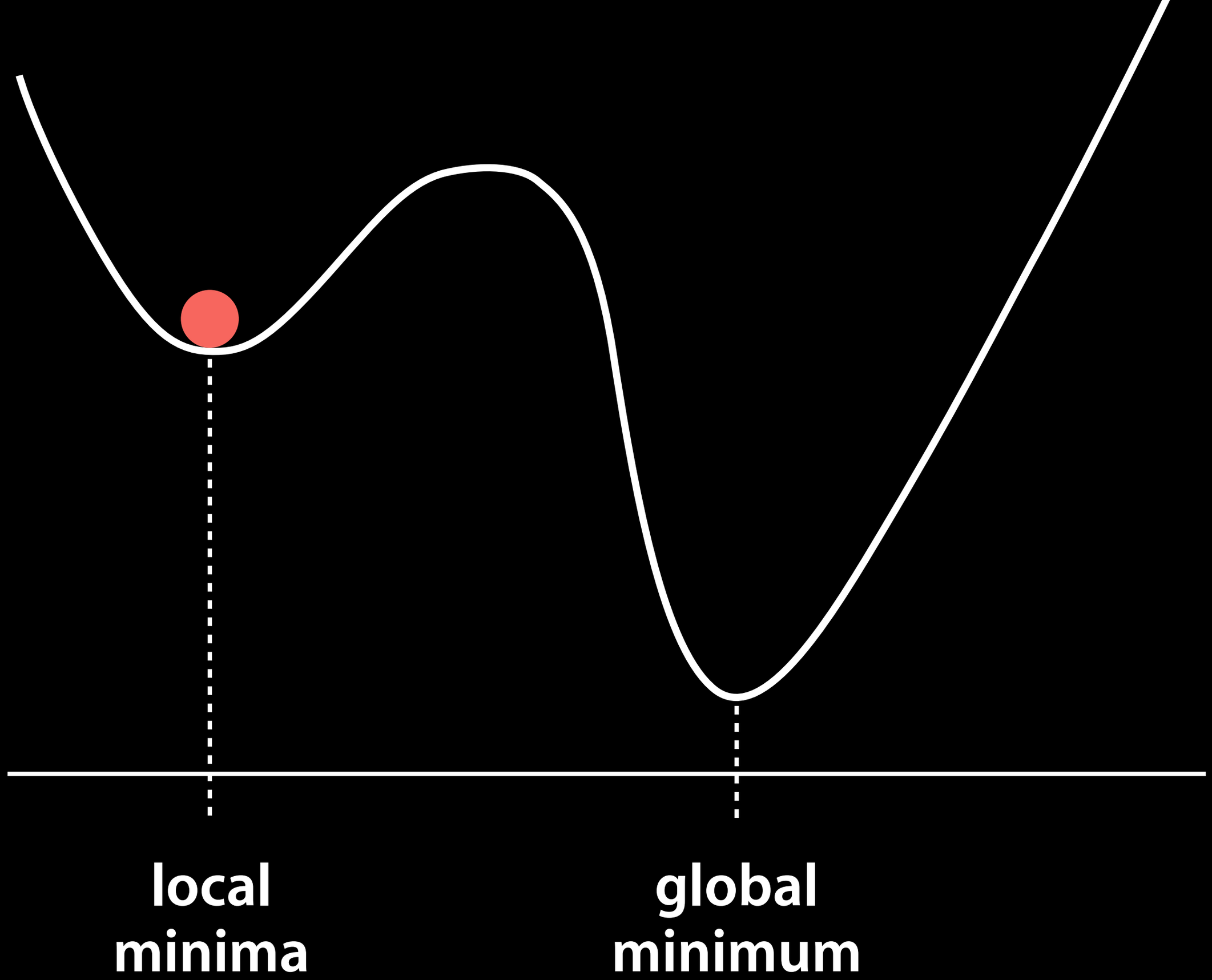
local
minima

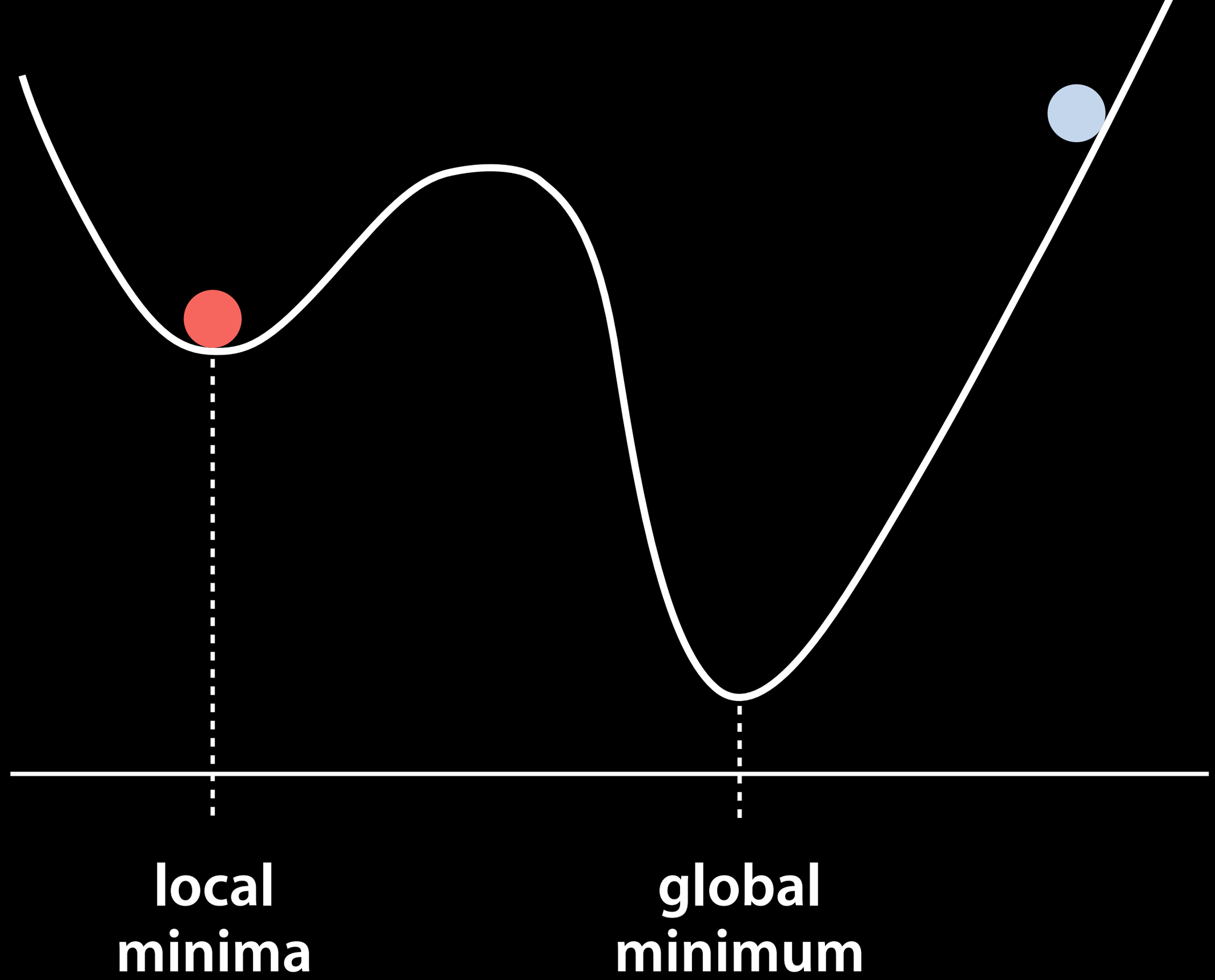

local
minima
local
minima

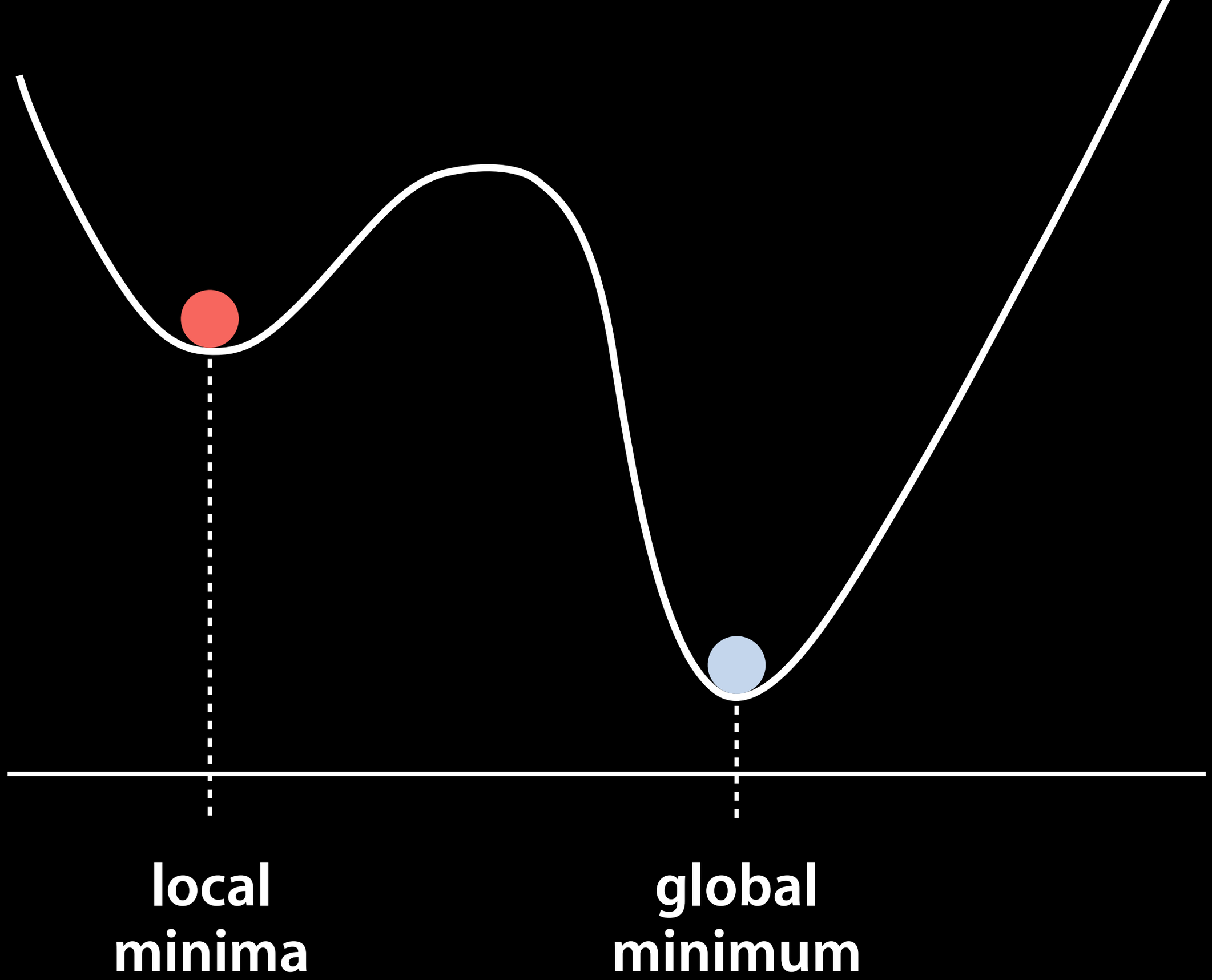

local
minima
local
minima


local
minima
global
minimum


local
minima
global
minimum


local
minima
global
minimum


local
minima
global
minimum


local
minima
global
minimum

Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any

$\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$**:**

Definition:

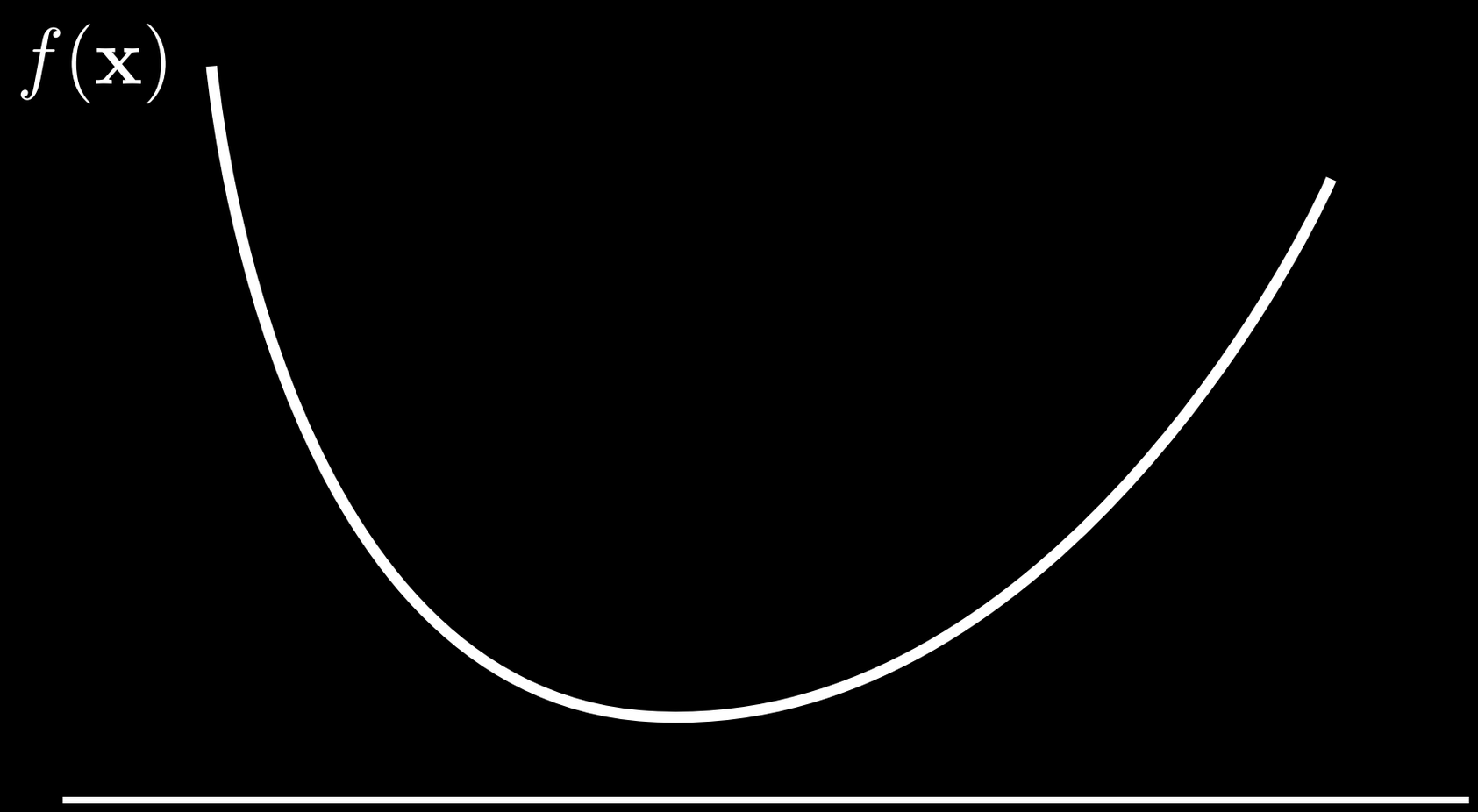
A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$**:**

$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$

Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

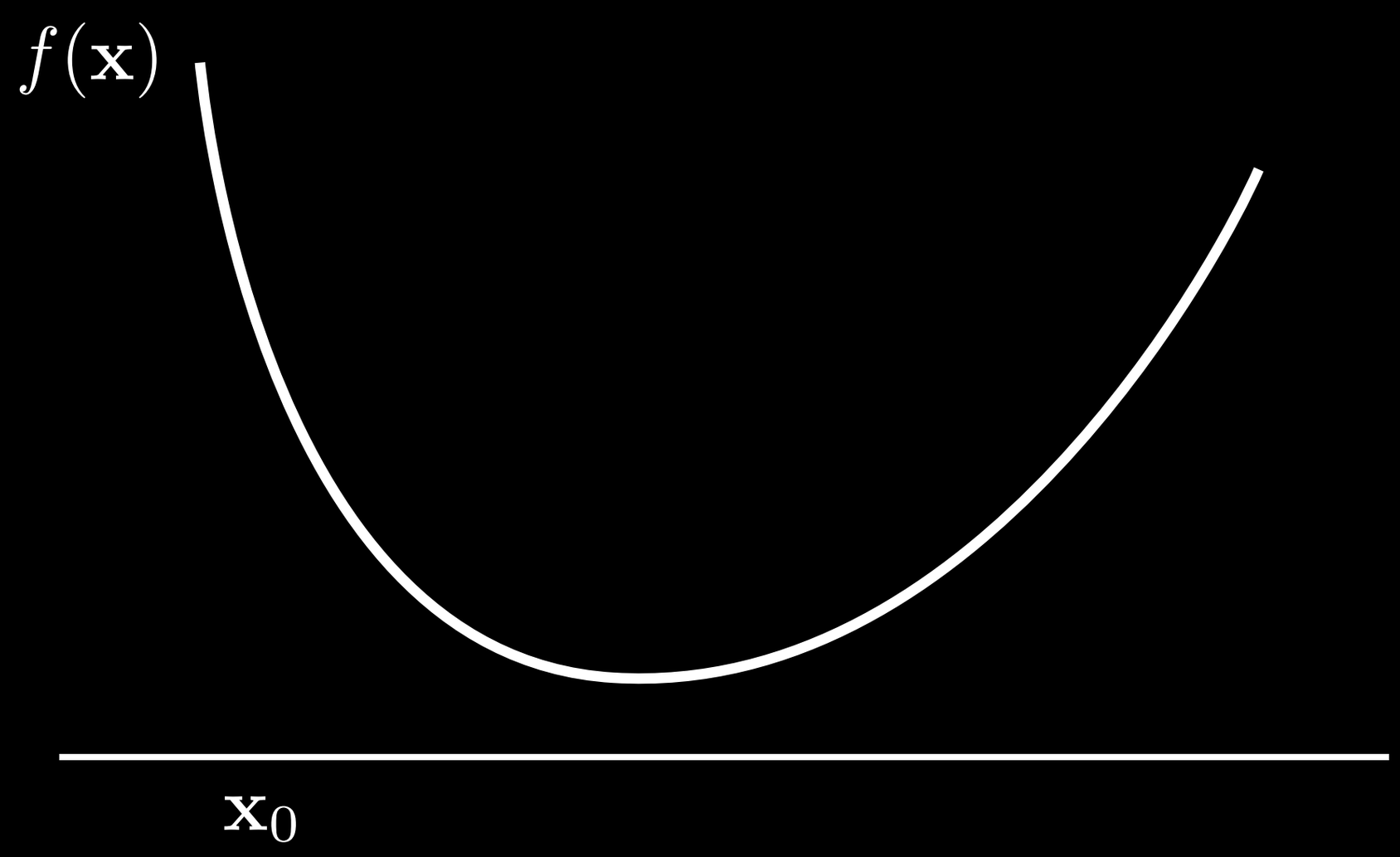
$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$



Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

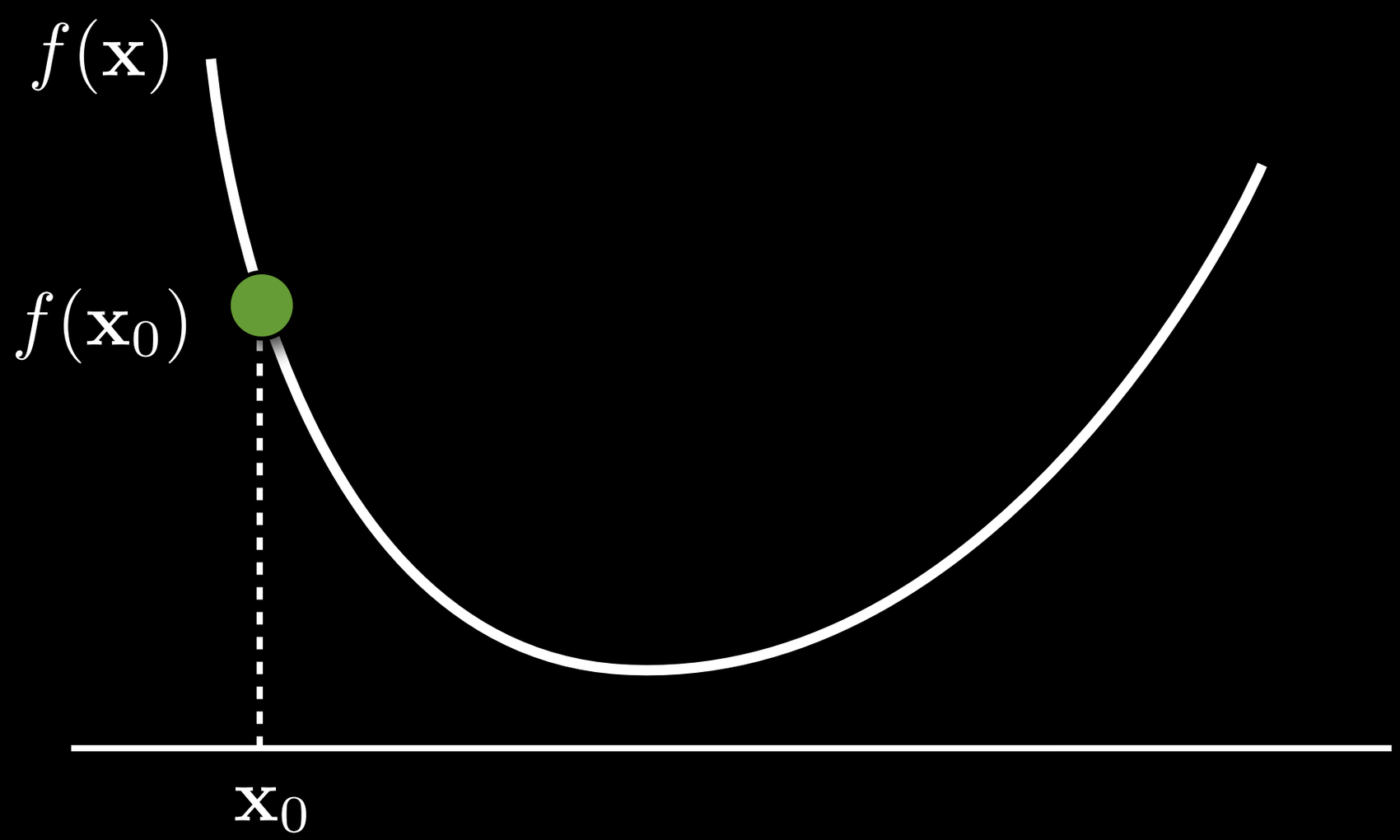
$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$



Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

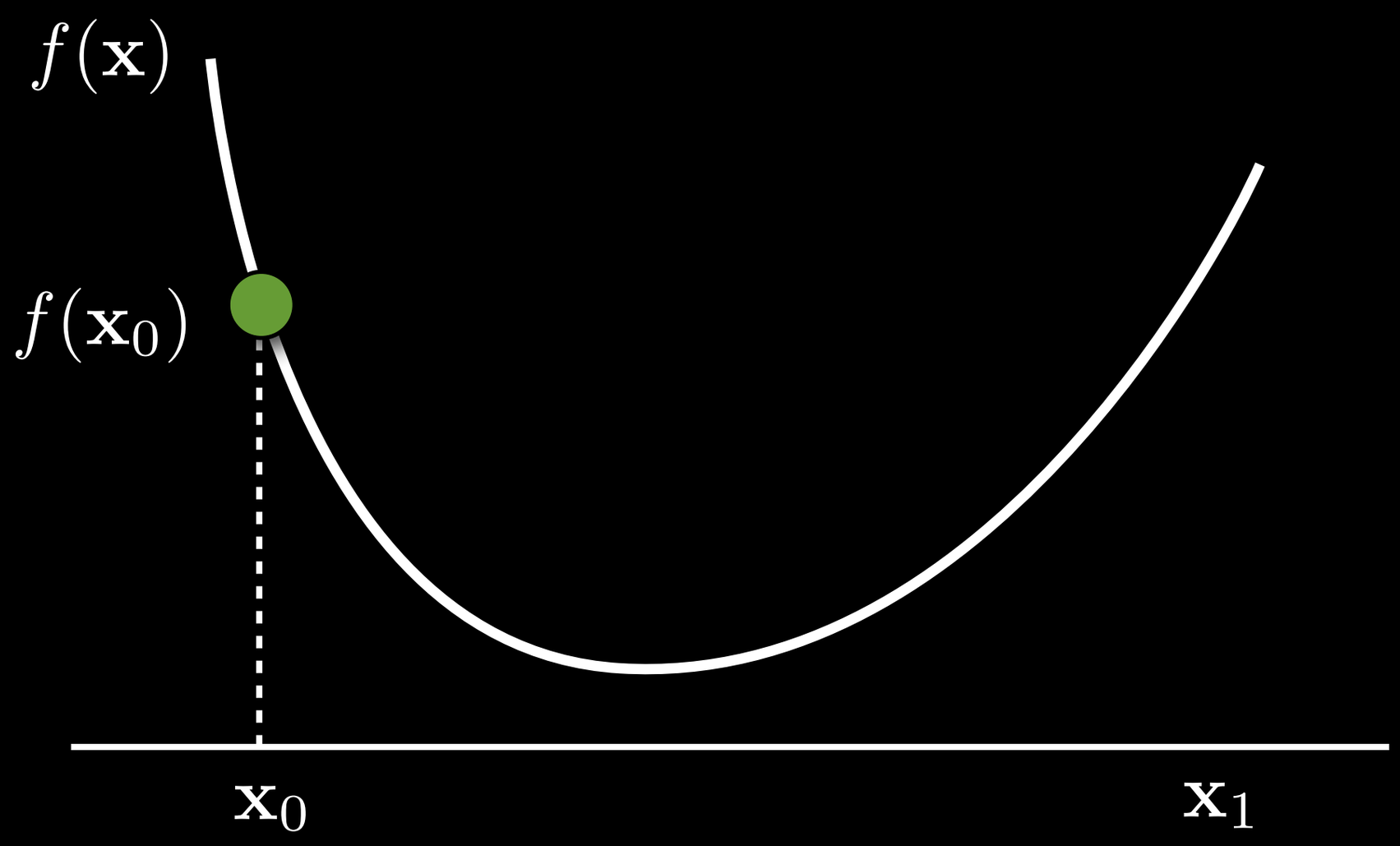
$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$



Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

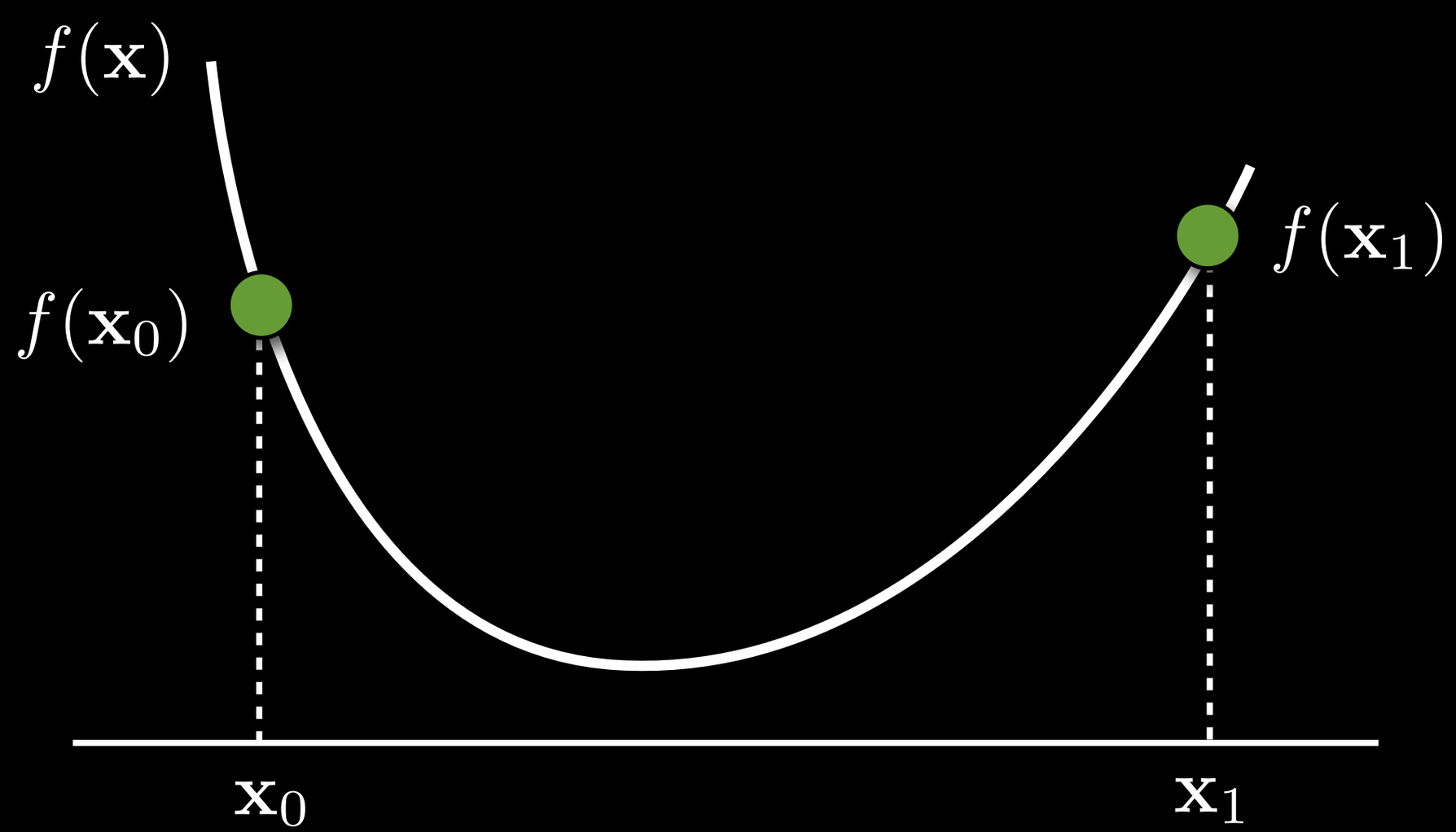
$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$



Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

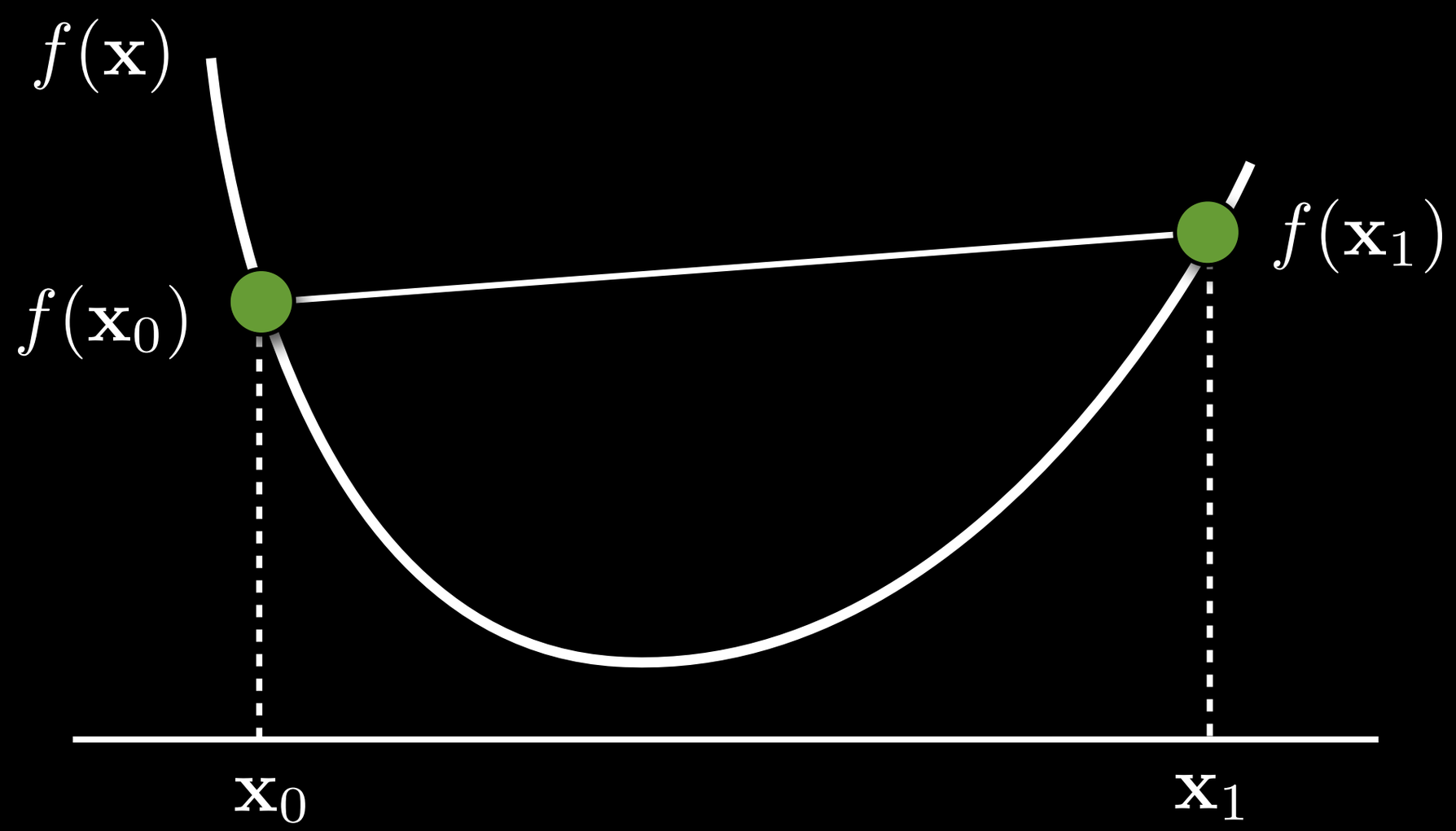
$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$



Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

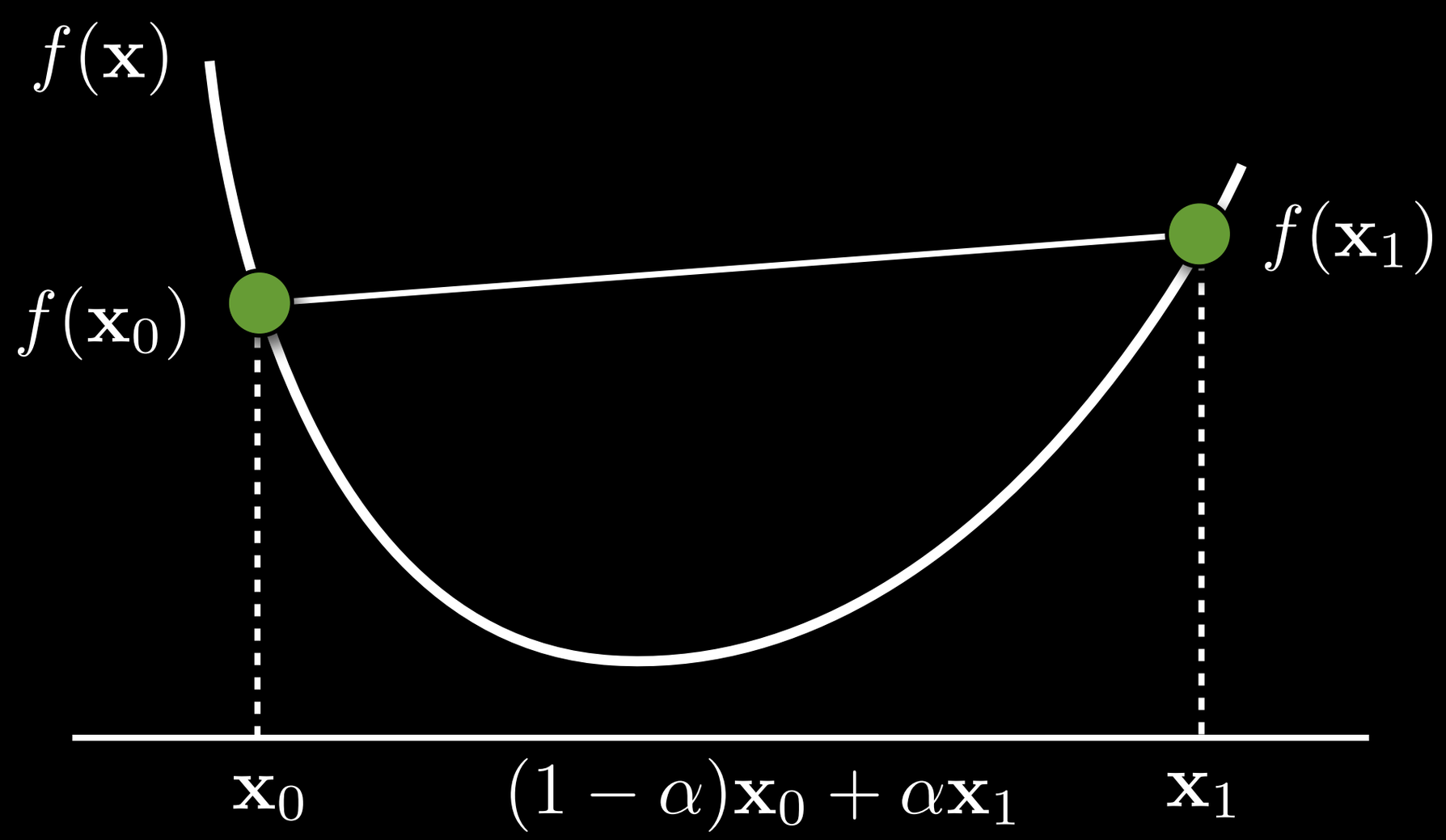
$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$



Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

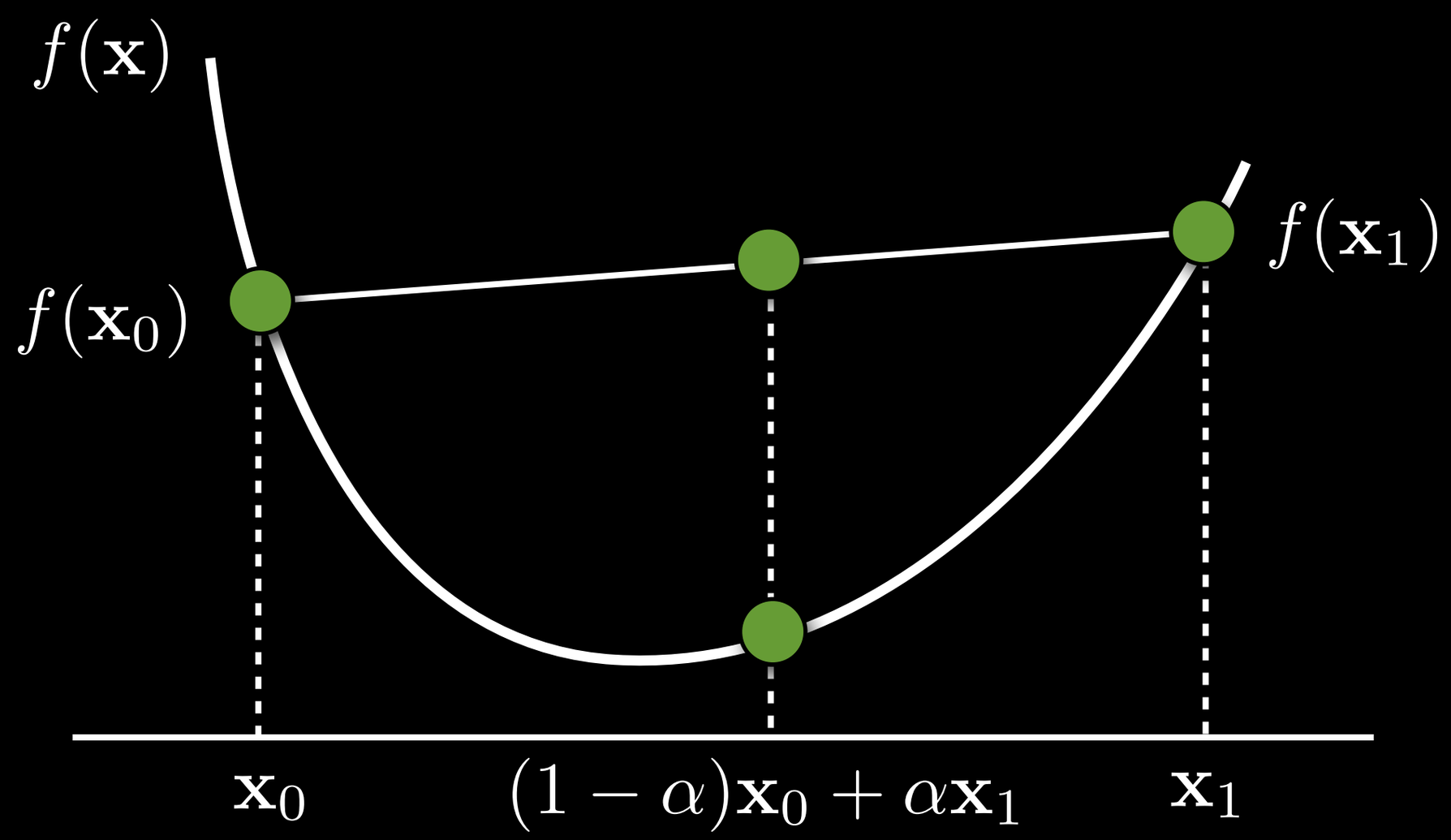
$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$

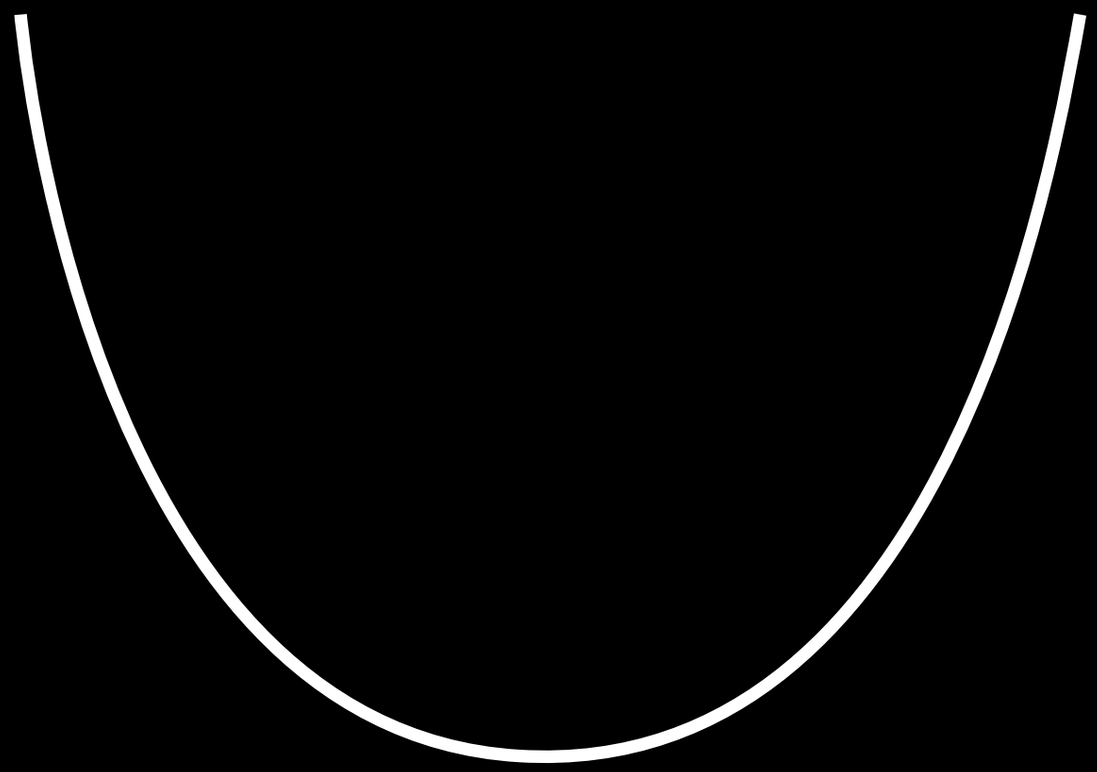


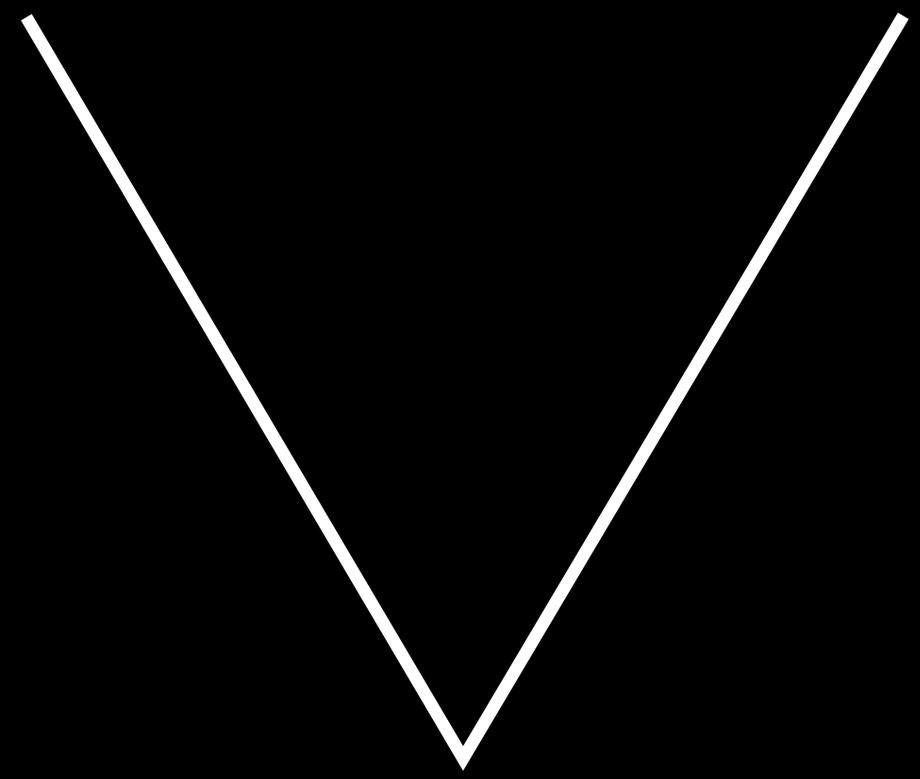
Definition:

A function, $f : D \to \mathbb{R}$, is said to be convex if for any $\mathbf{x}_0$ **and** $\mathbf{x}_1$ **in** D **and** $0 \leq \alpha \leq 1$ **:**

$$f((1-\alpha)\mathbf{x}_0 + \alpha\mathbf{x}_1) \leq (1-\alpha)f(\mathbf{x}_0) + \alpha f(\mathbf{x}_1).$$



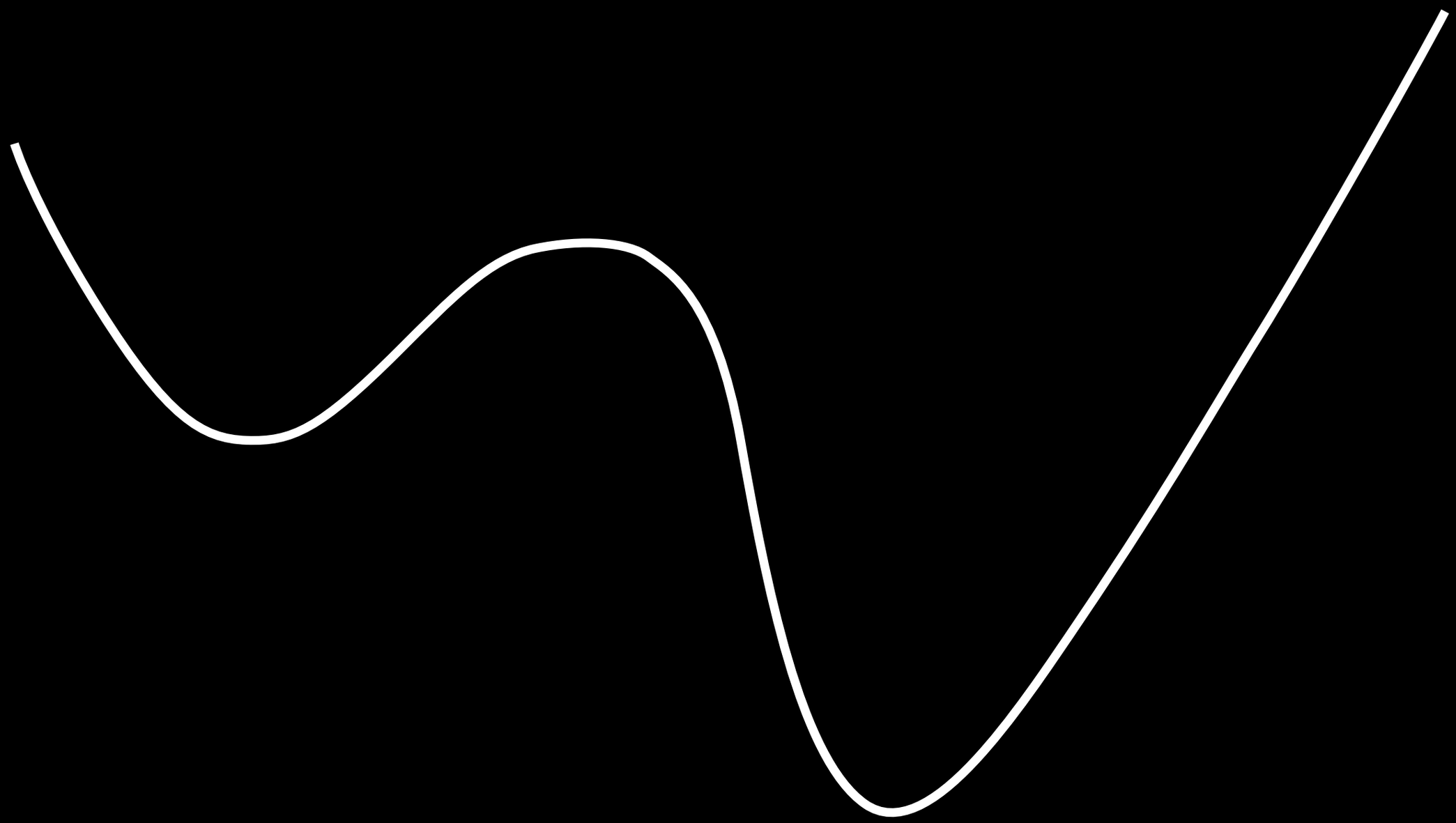




convex functions

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

SVM cost function is convex



non-convex function

Training

Objective of machine learning is generalization

Data

Data

NEVER train and test on the same data

35. A
36. B
37. C
38. A
39. B
40. E
41. B
42. B
43. A
44. B
45. C
46. A
RESCORE
MARK ✓
TOTAL ONLY/BO
KEY ITEM COUNT

Data

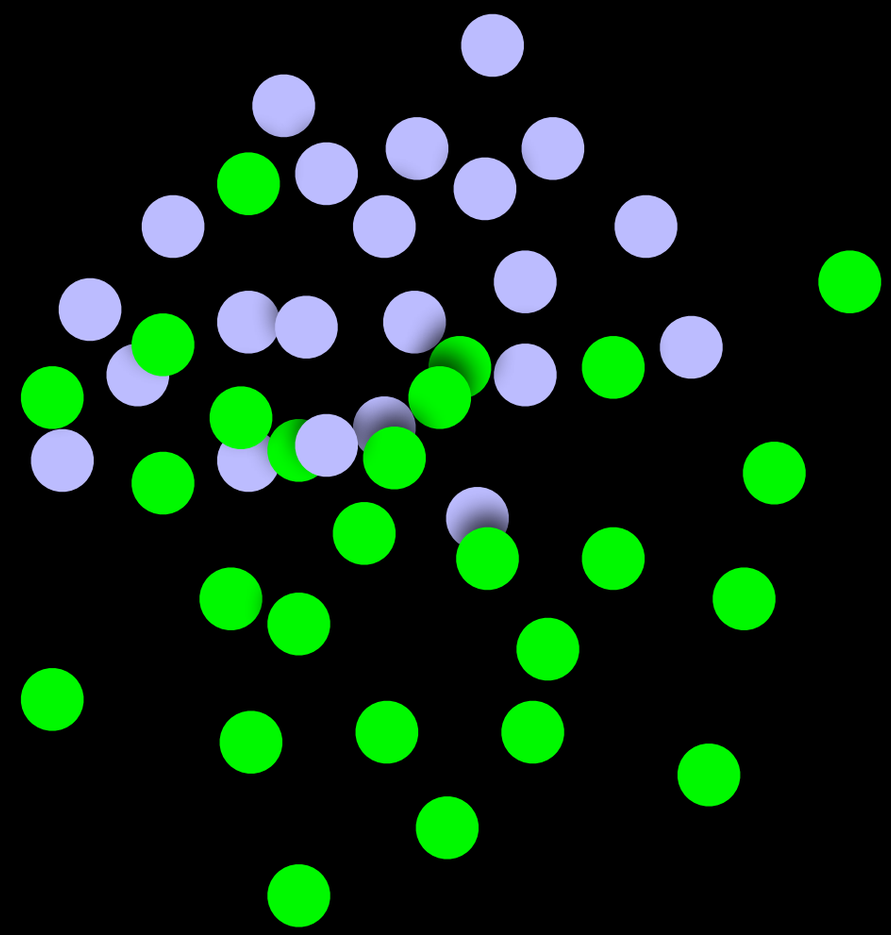
NEVER train and test on the same data

Training Data	Test

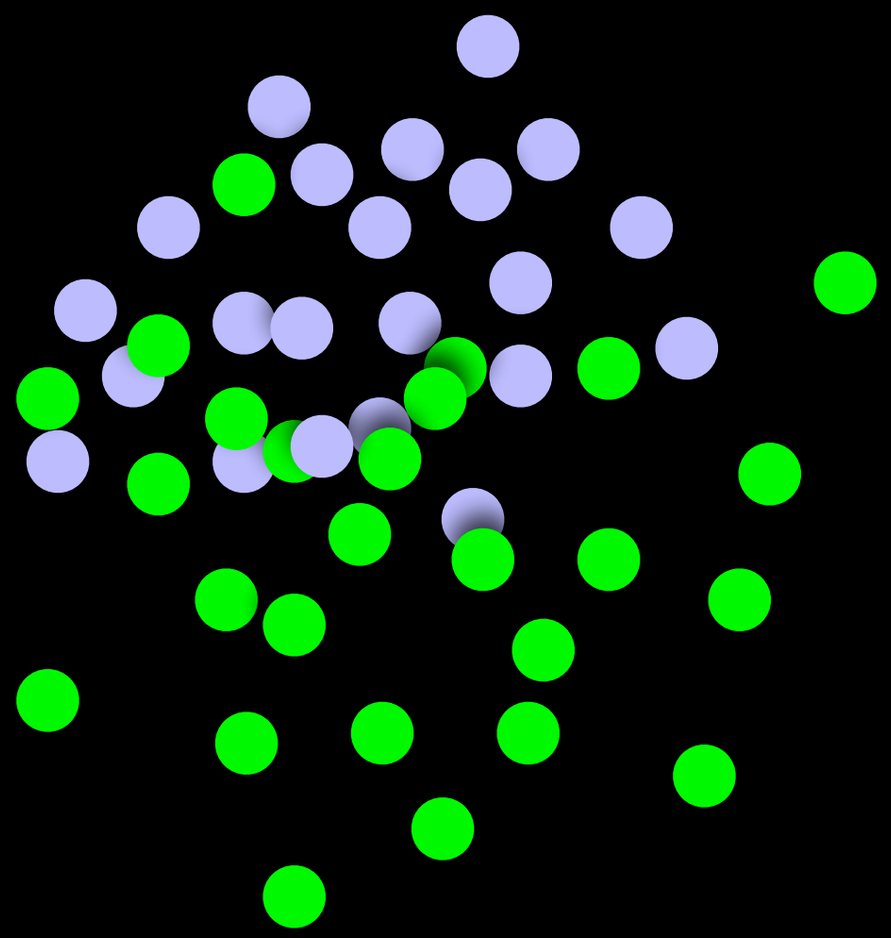
Train on training data and test on withheld test set

Training Data	Test

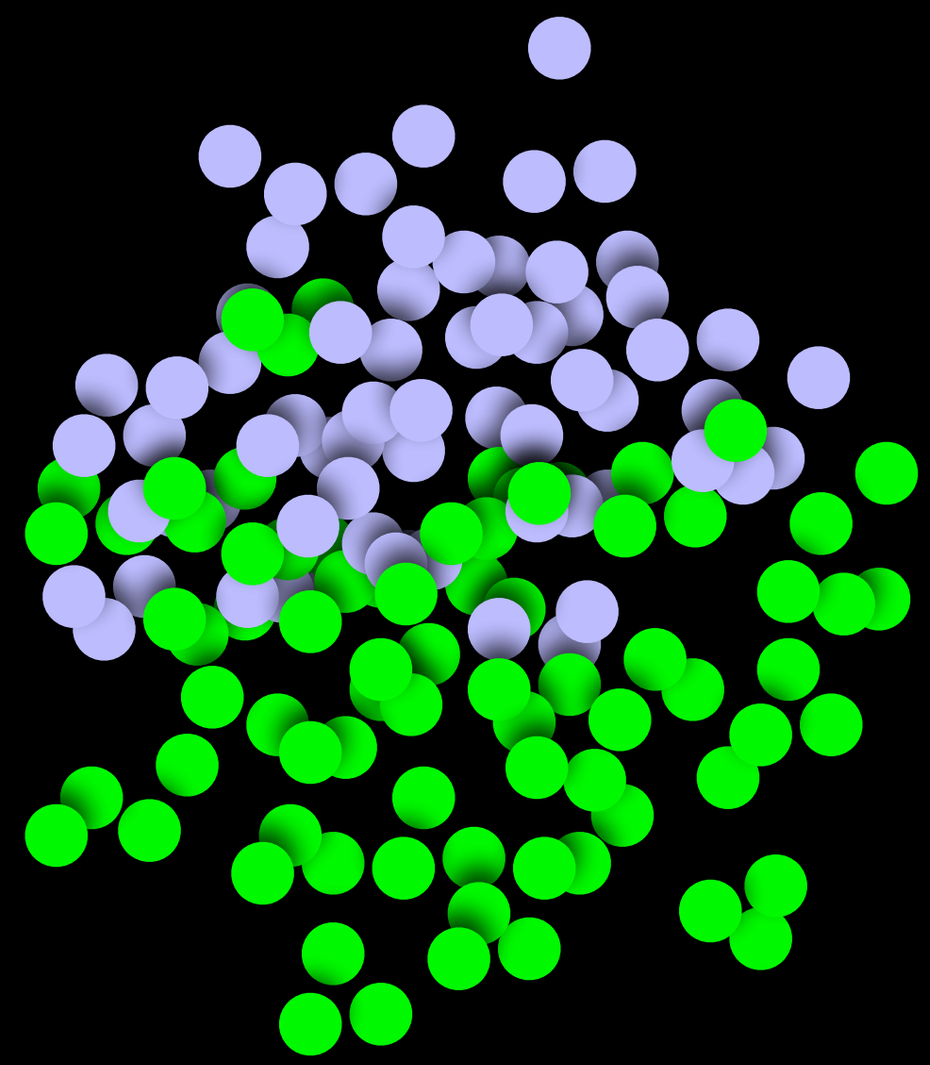
Assume training and testing data drawn independently from the same distribution



training data



training data

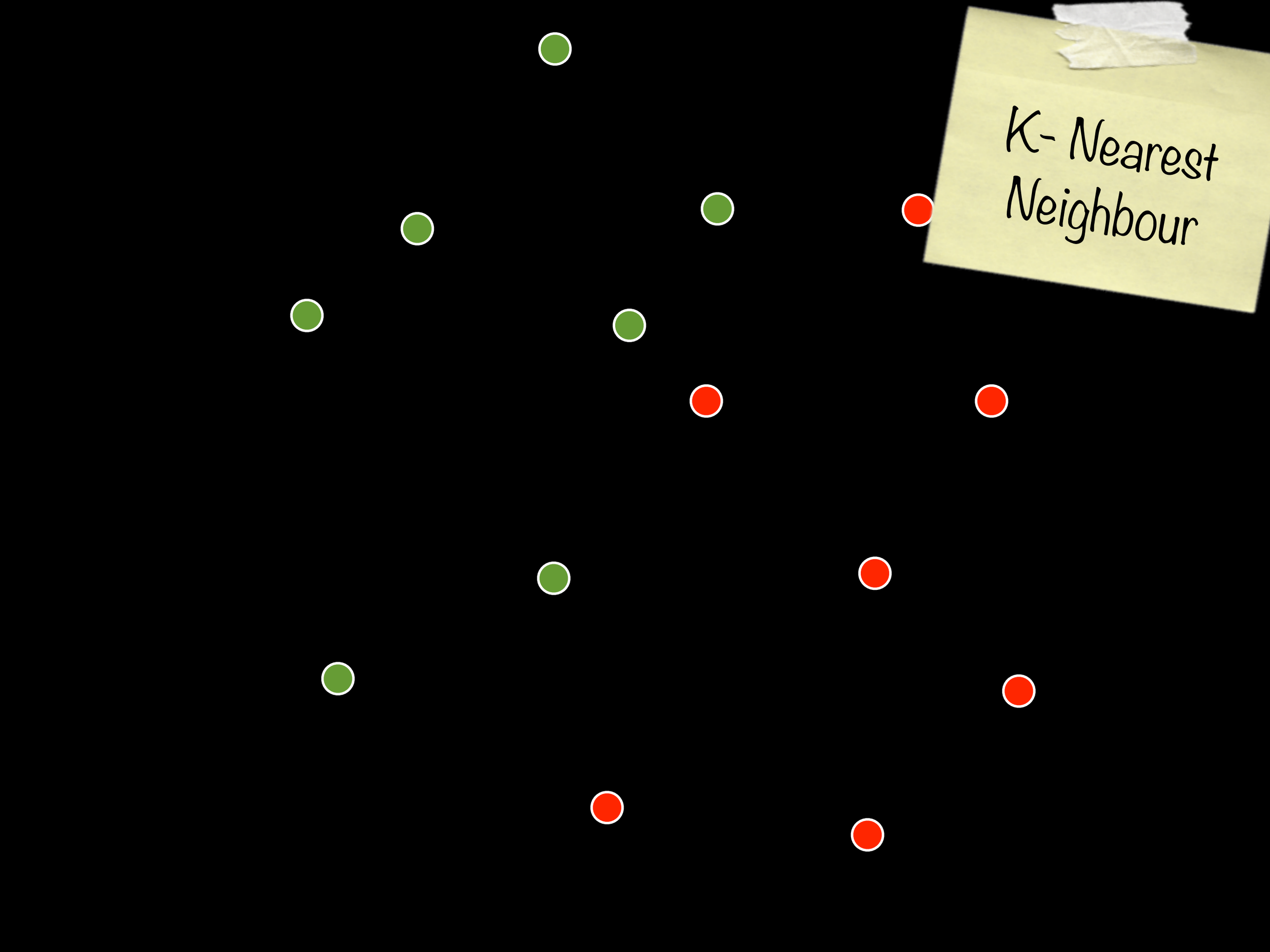


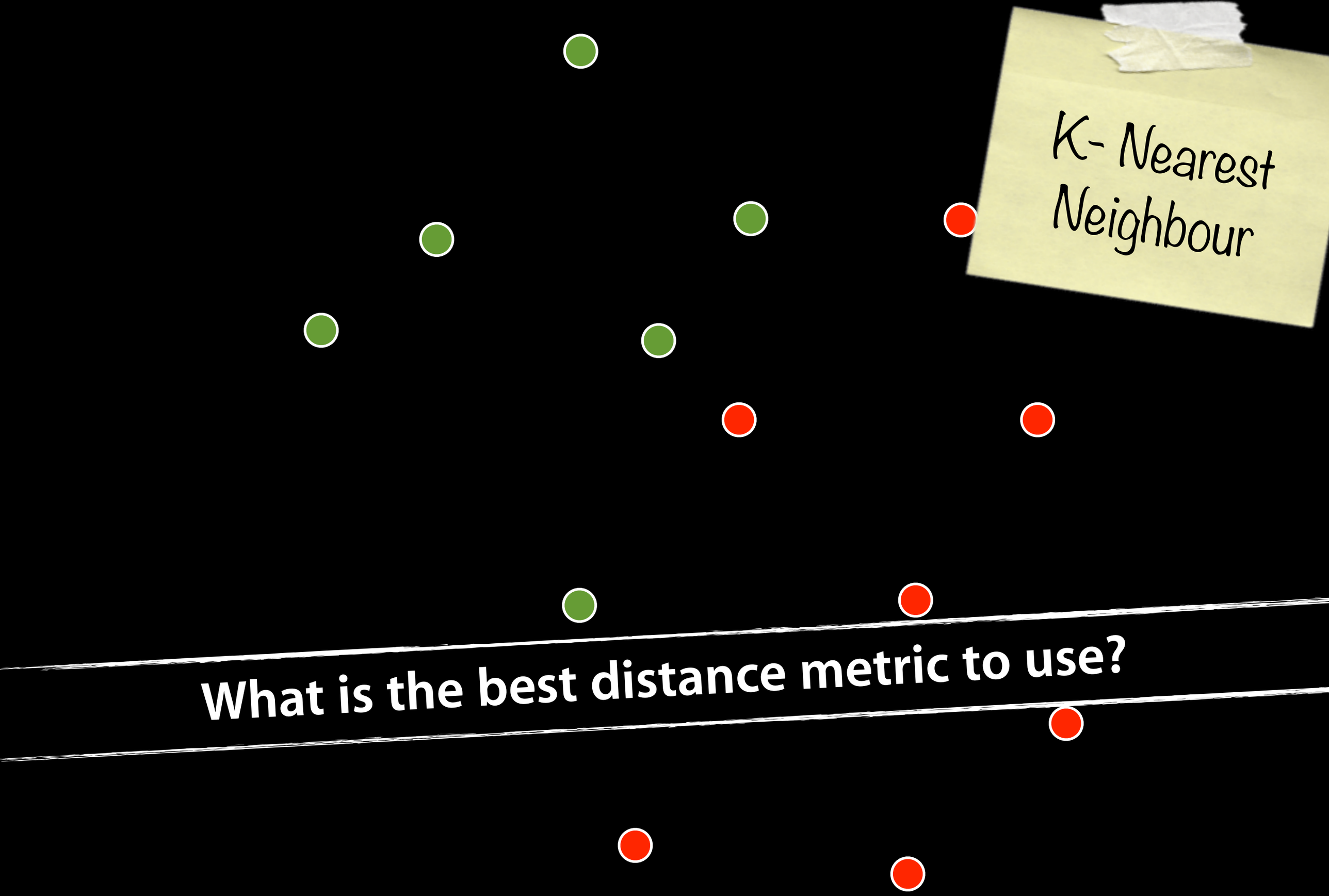
testing data

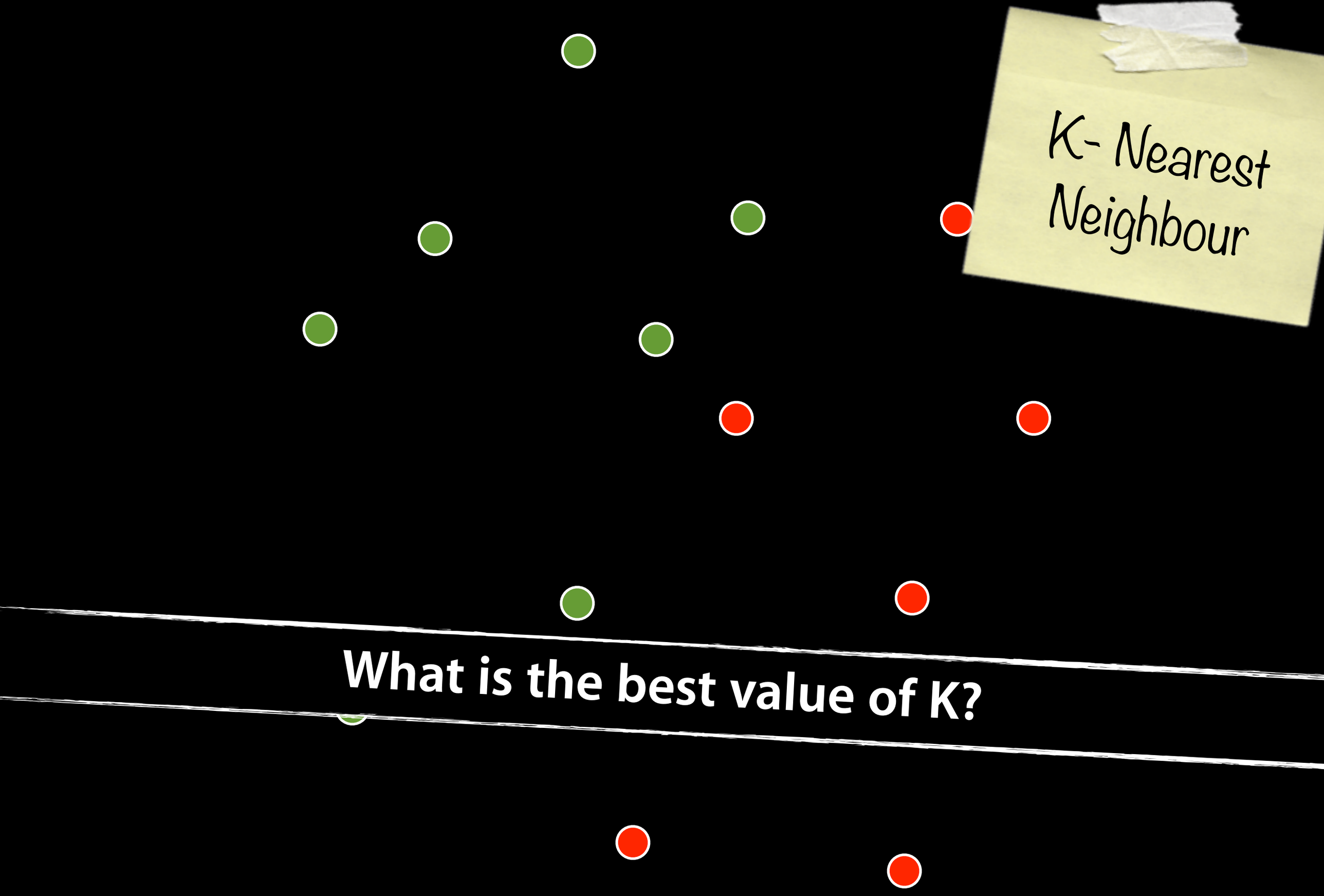
Training Data	Test

Assume training and testing data drawn independently from the same distribution

How do we set the hyperparameters?


K- Nearest
Neighbour


K- Nearest Neighbour
What is the best distance metric to use?


K- Nearest Neighbour
What is the best value of K?

Linear Classifiers

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

Linear Classifiers

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

What is the best regularizer?

Linear Classifiers

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

What is the best regularizer?

Linear Classifiers

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

What is the best regularization parameter?

Linear Classifiers

$$L = \frac{1}{N}\sum_{i=1}^{N} \max(0, 1 - y_i[\mathbf{w}\cdot\mathbf{x}_i + b]) + \lambda R(\mathbf{w})$$

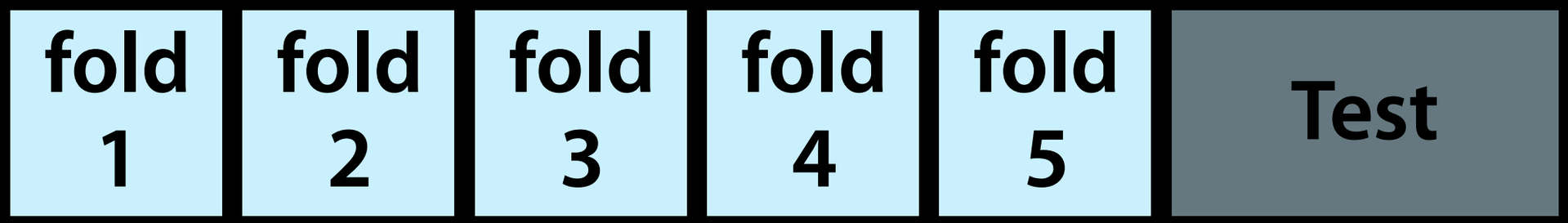
What is the best regularization parameter?

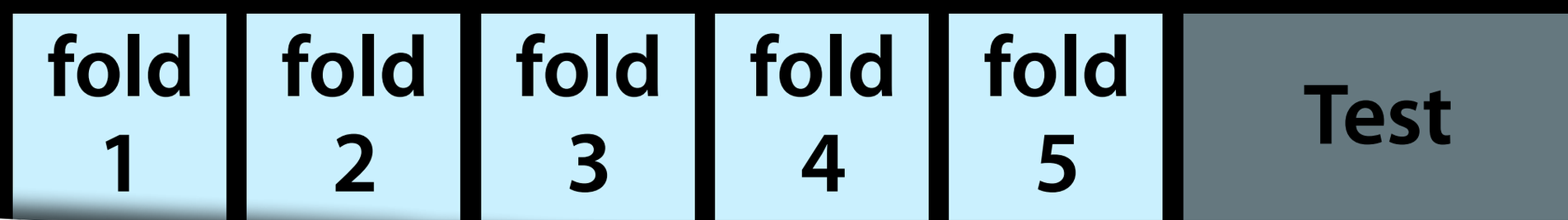
How do we set the hyperparameters?

Training Data
Test

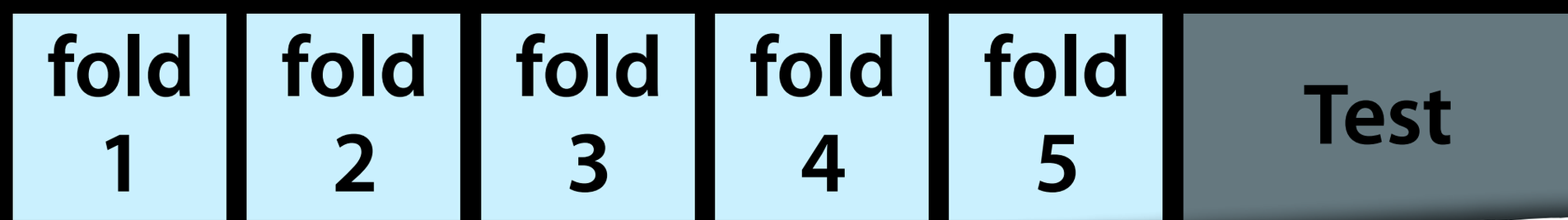
Training Data
Test
Validation

Training Data	Test


fold 1
fold 2
fold 3
fold 4
fold 5
Test

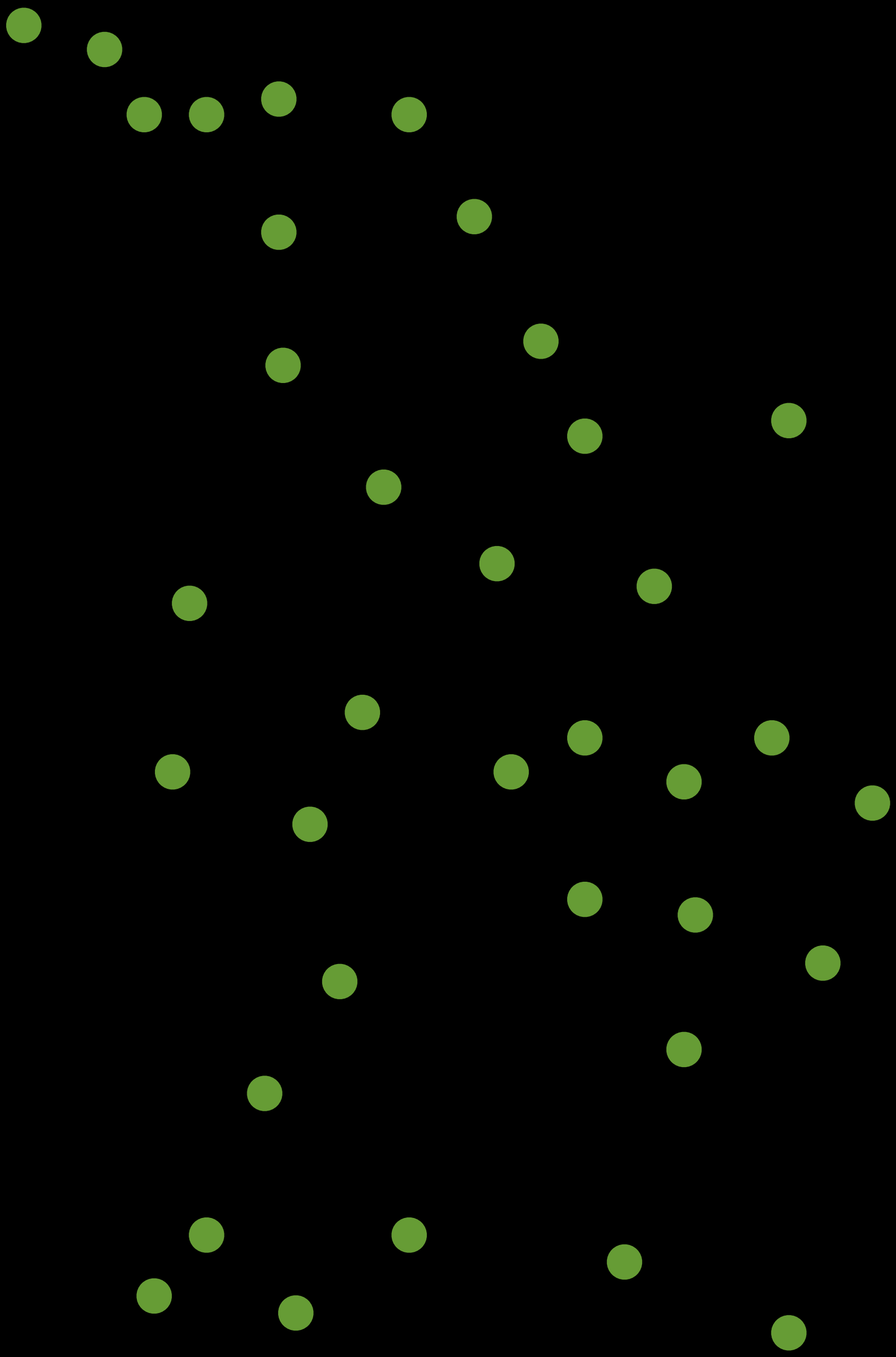


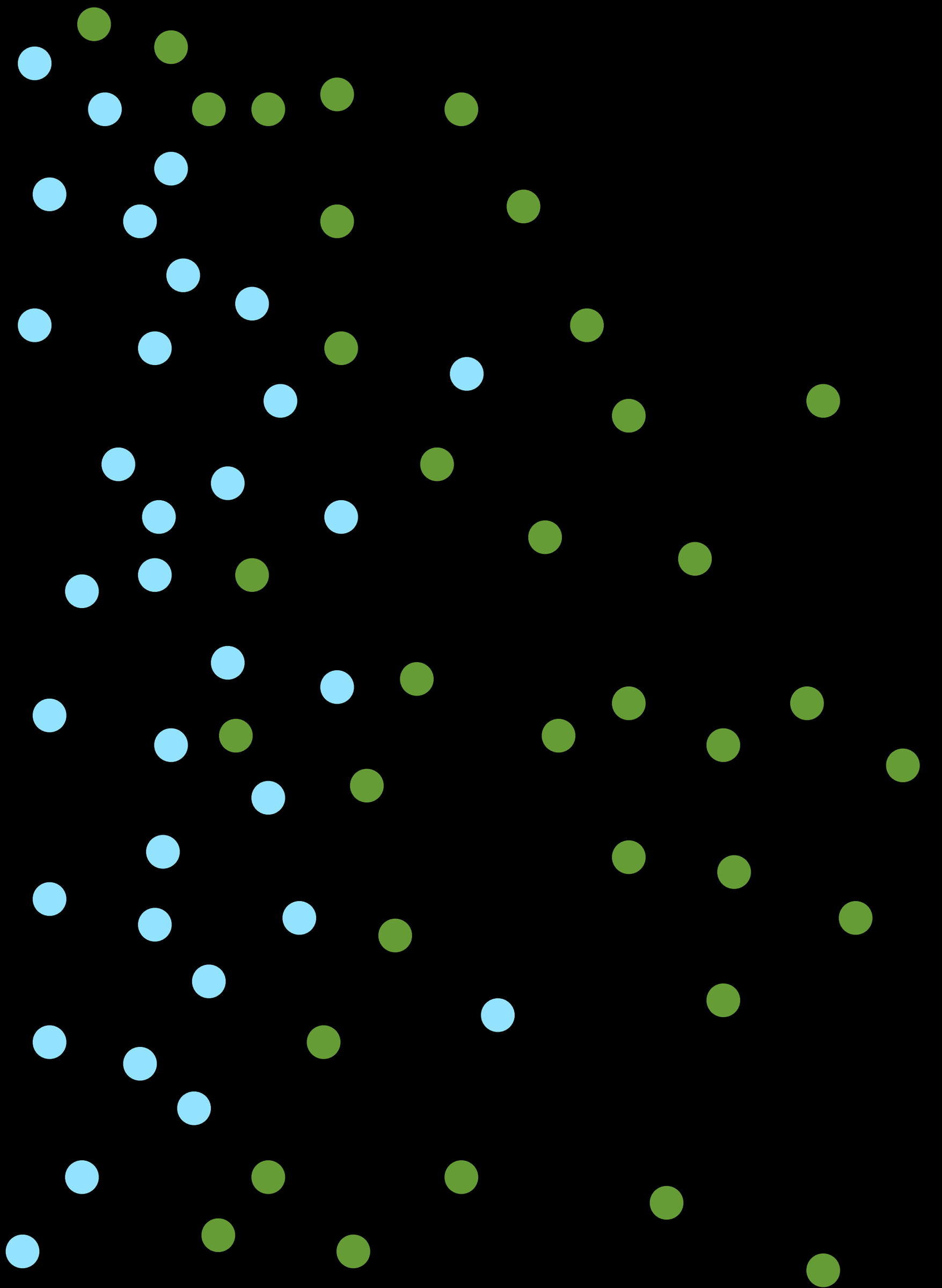
Use a fold as validation data and train on other folds

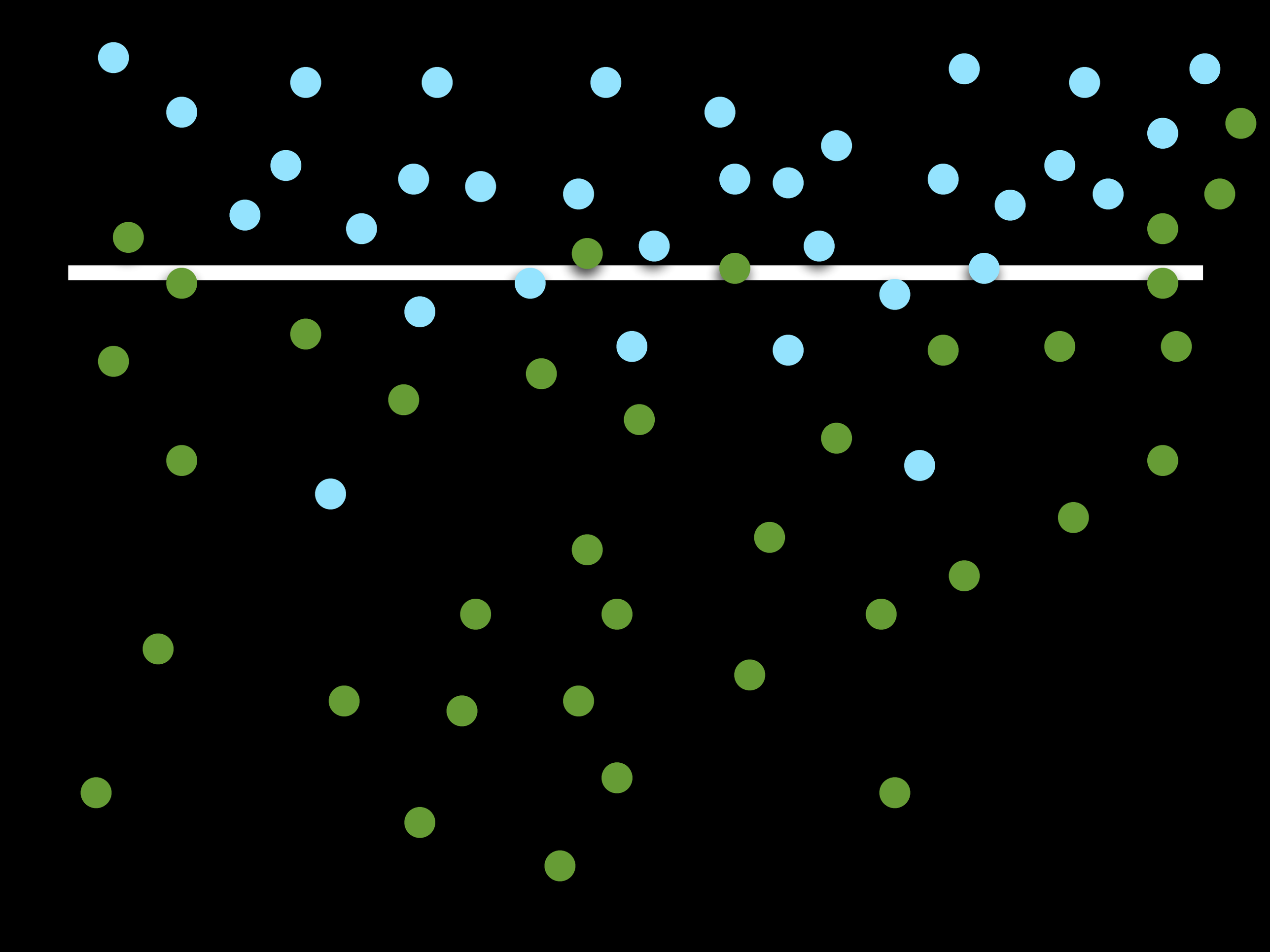


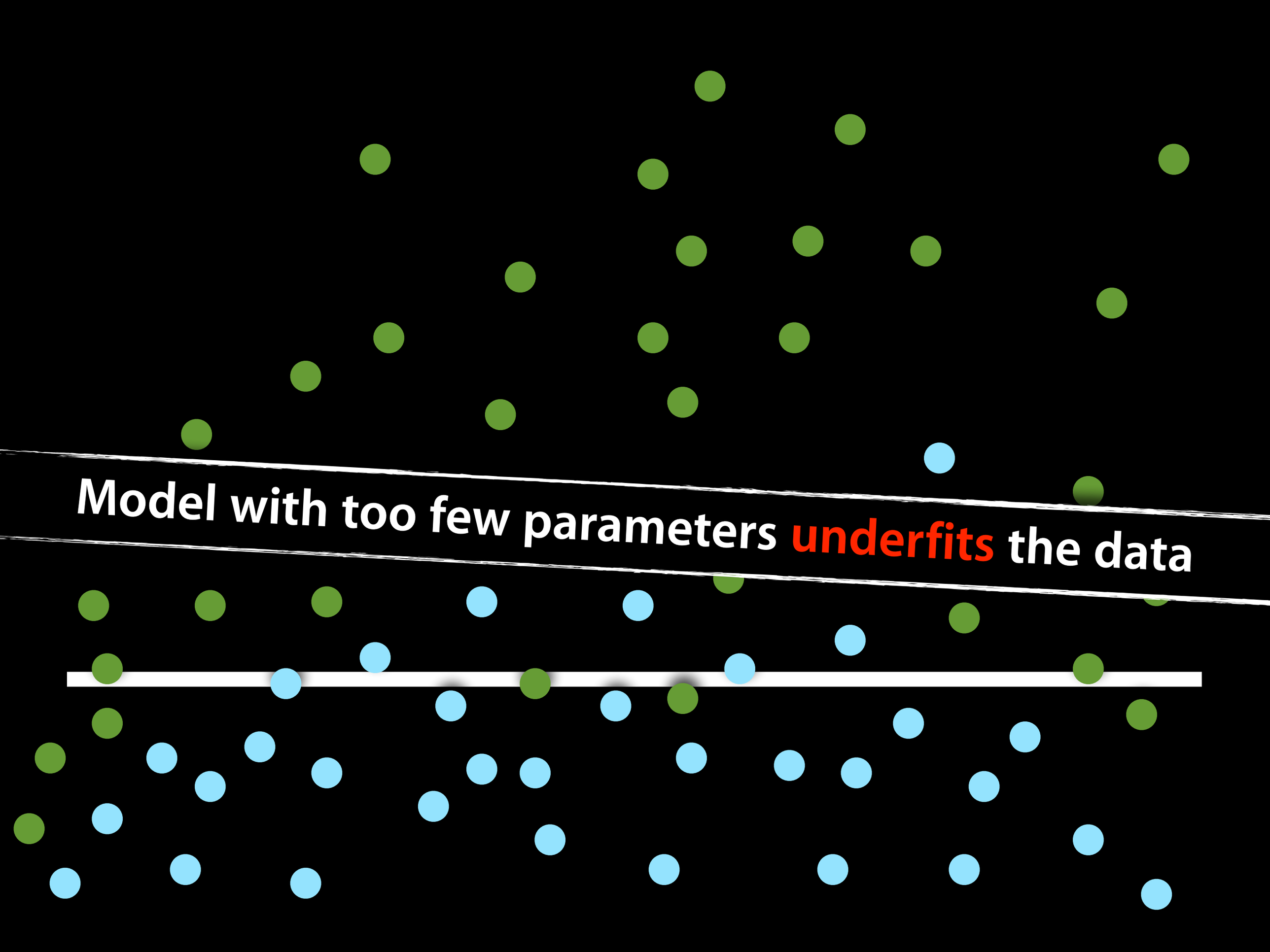
Cross-validation

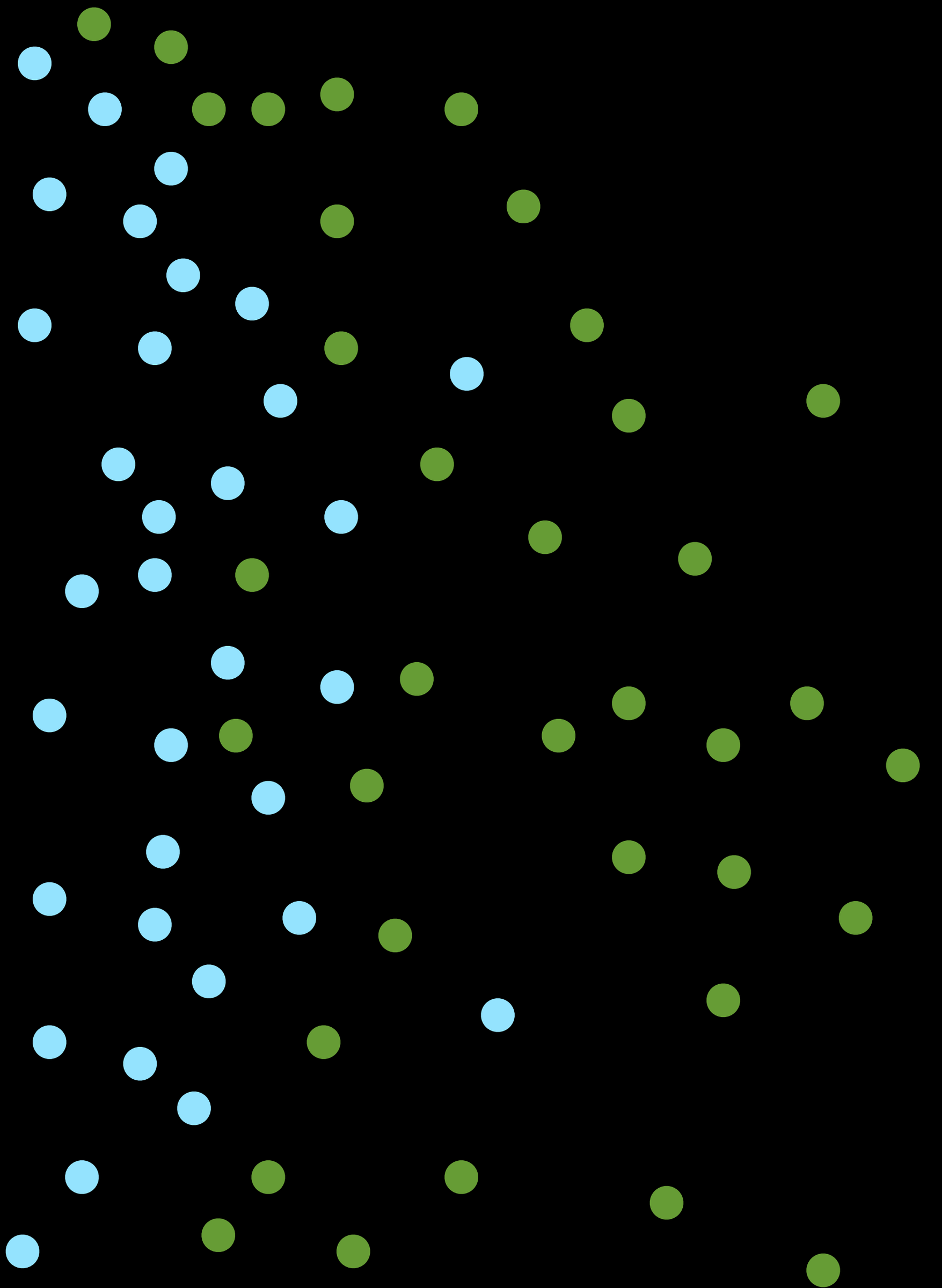
Select each fold in turn as validation data and average results

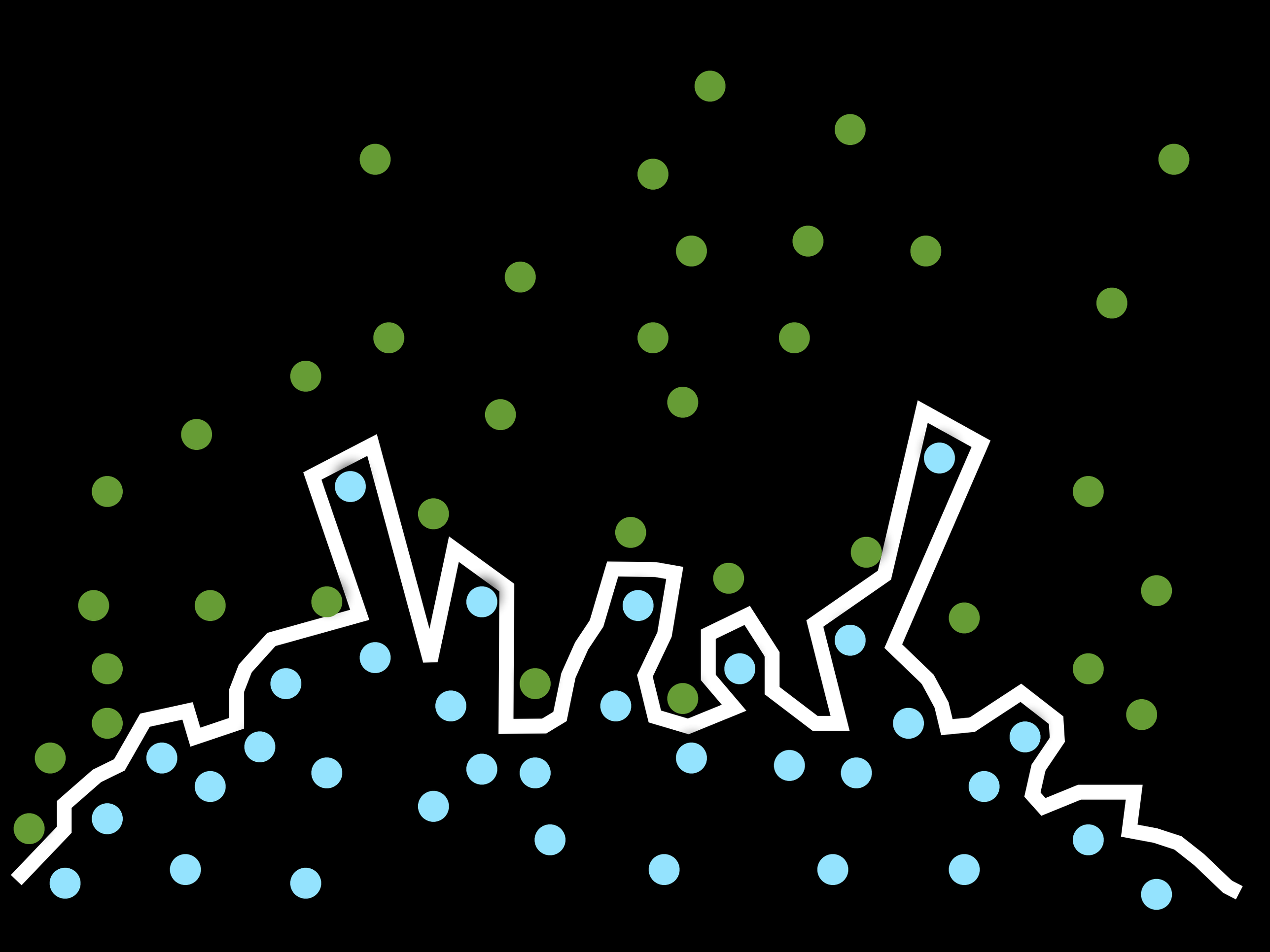




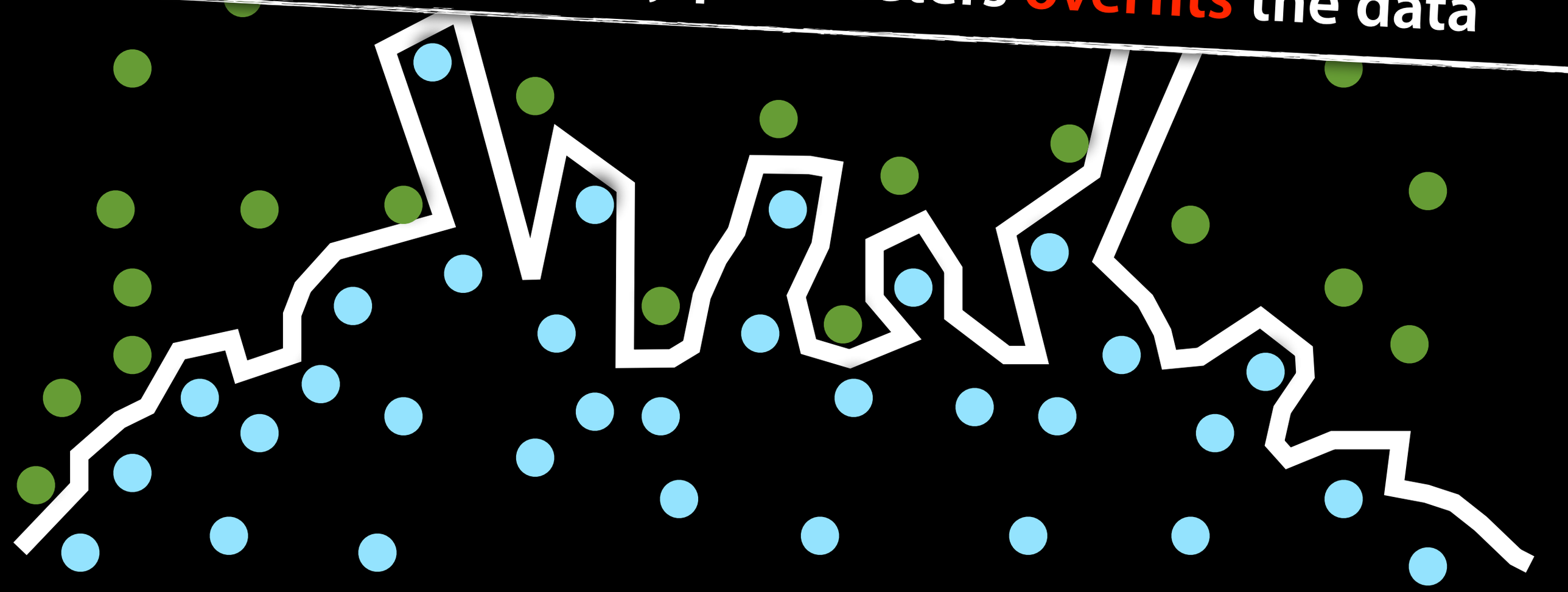


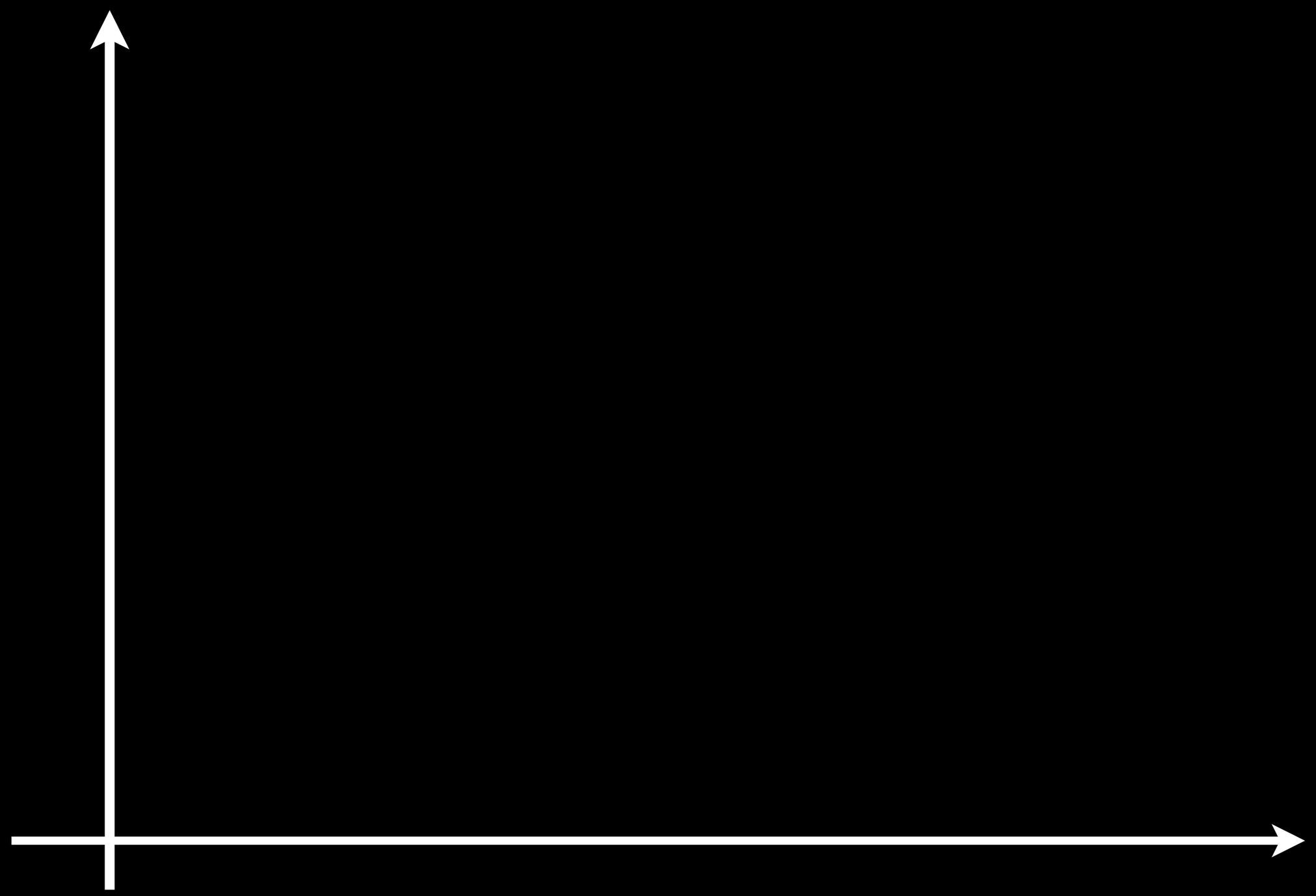

Model with too few parameters underfits the data

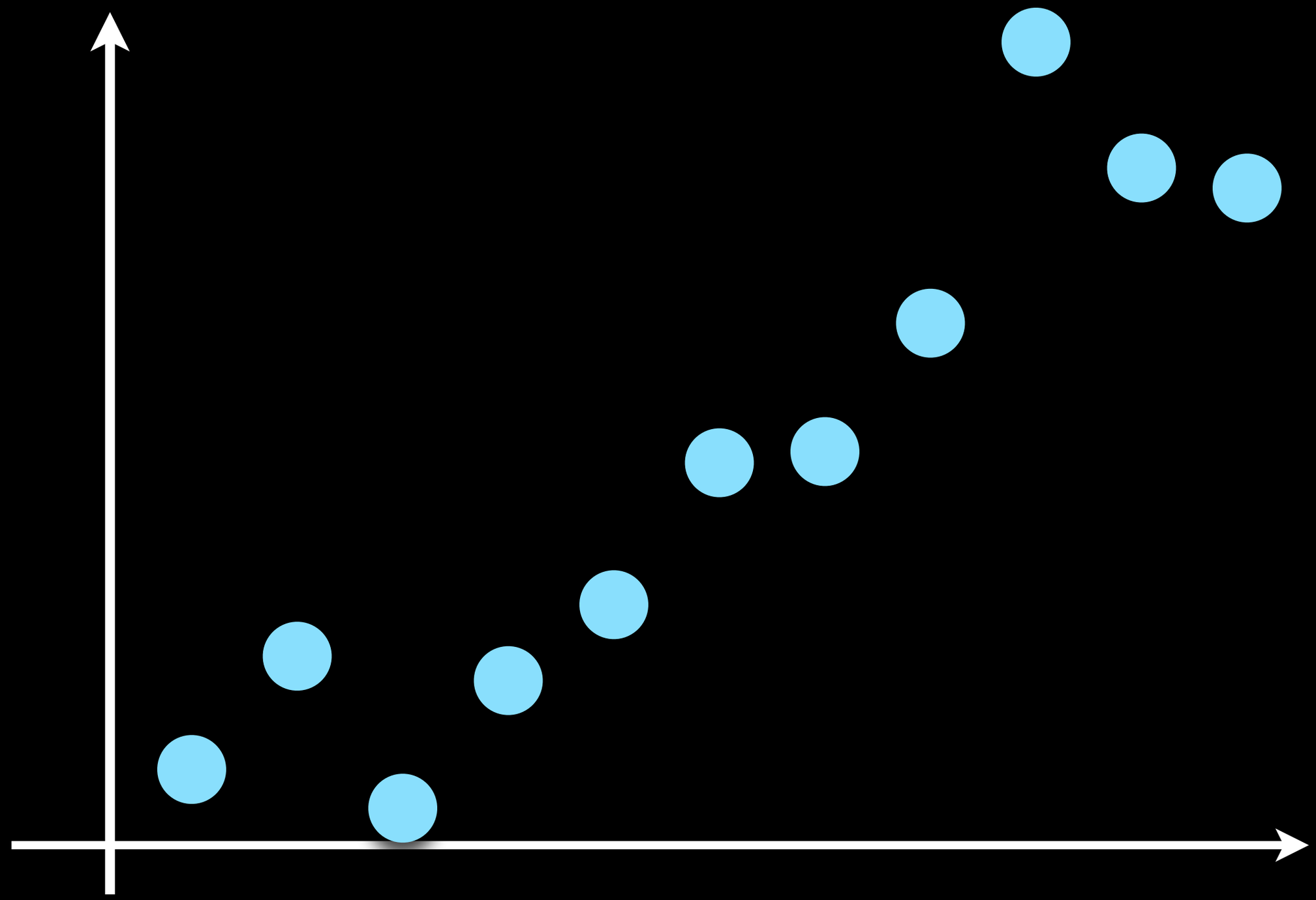


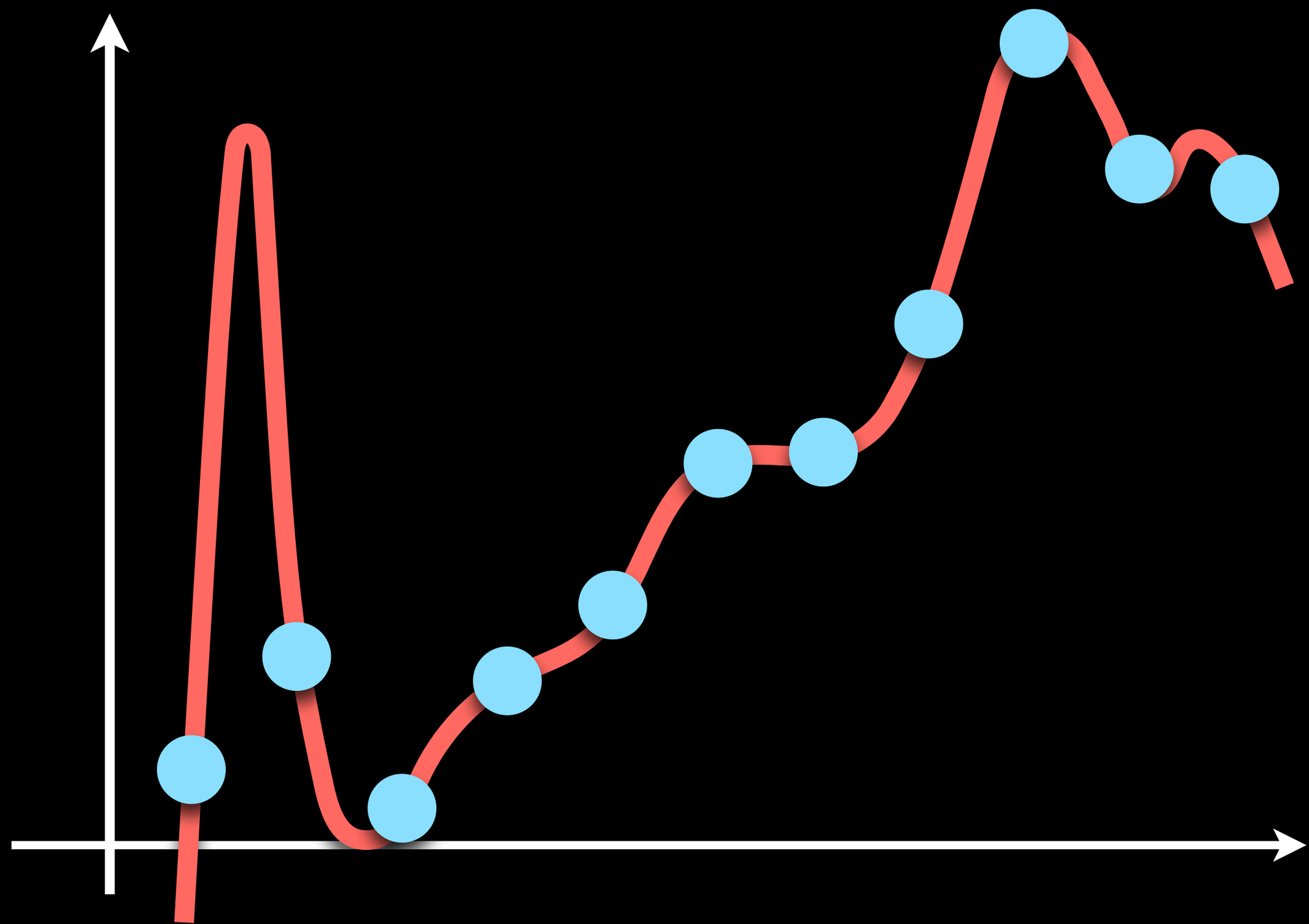

work

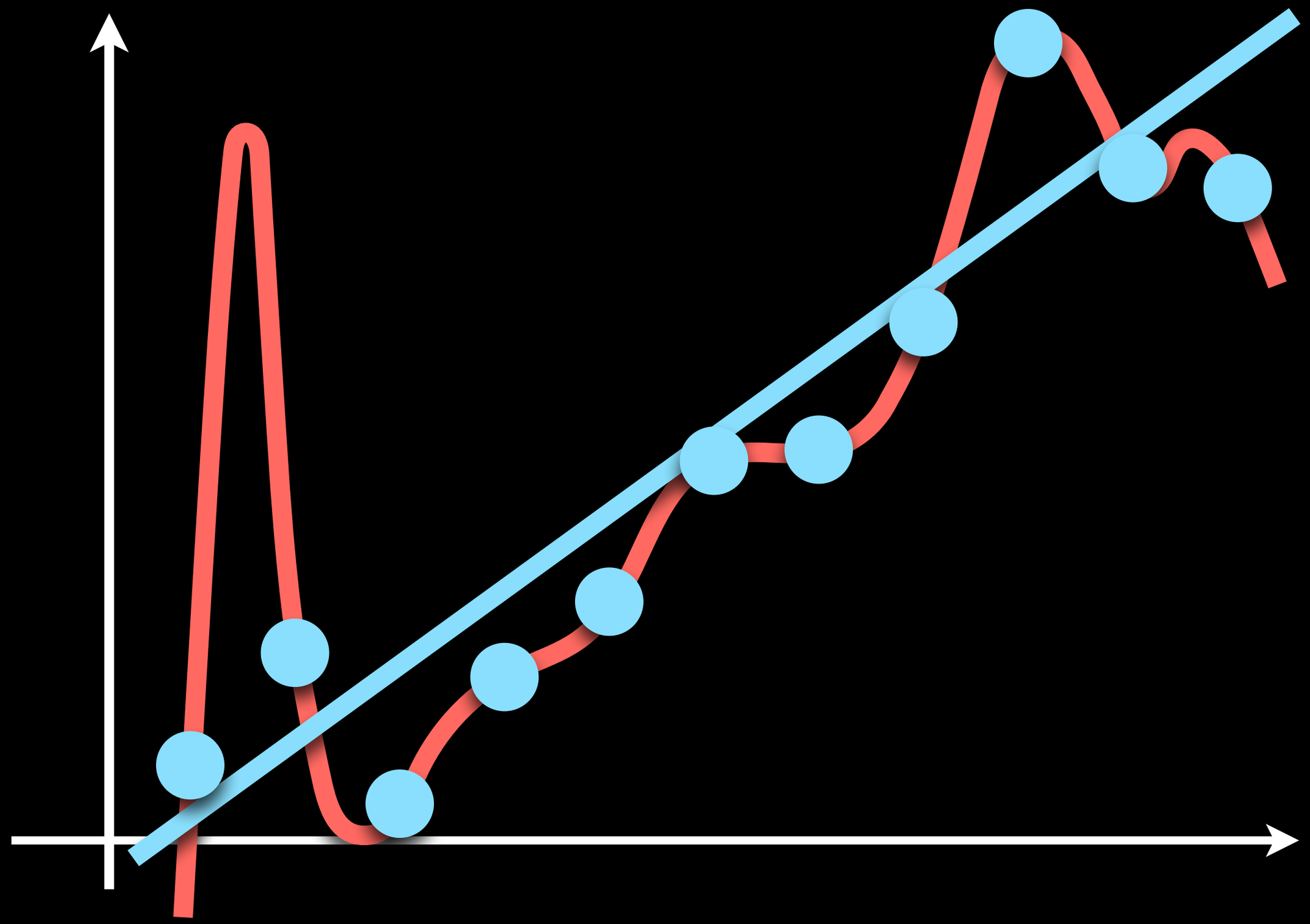
Model with too many parameters **overfits** the data





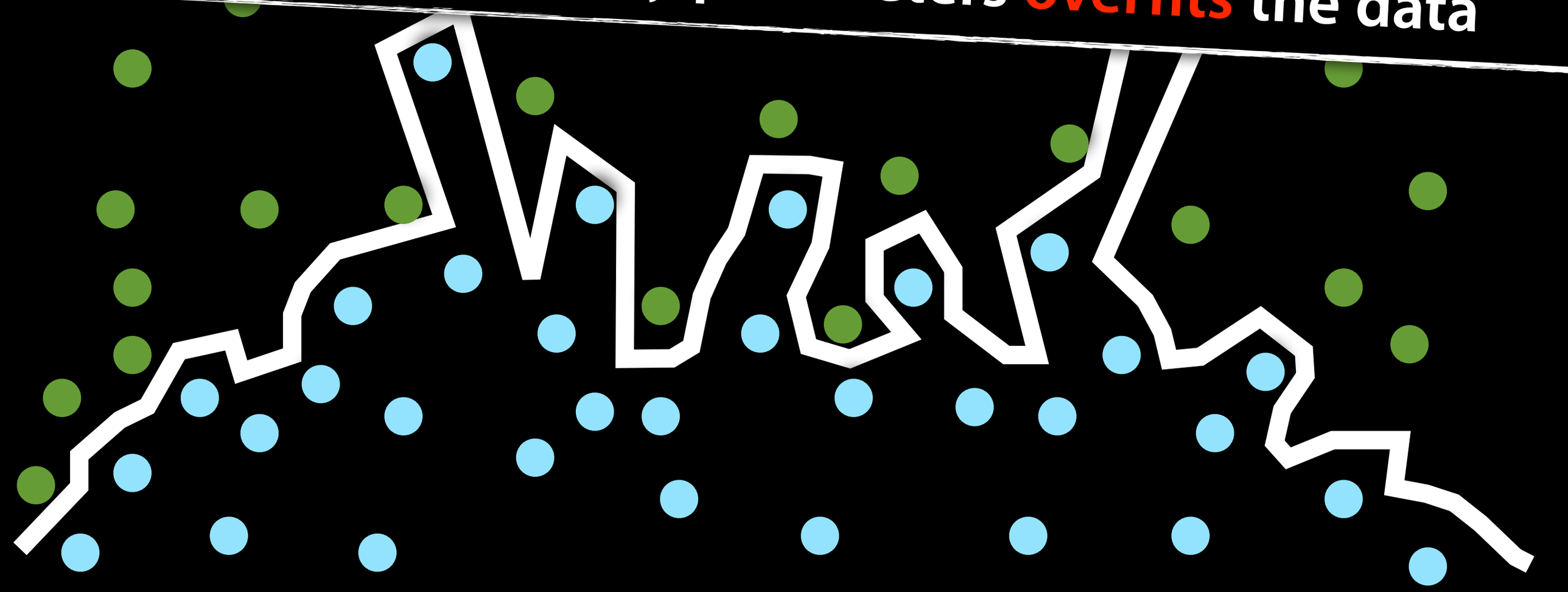


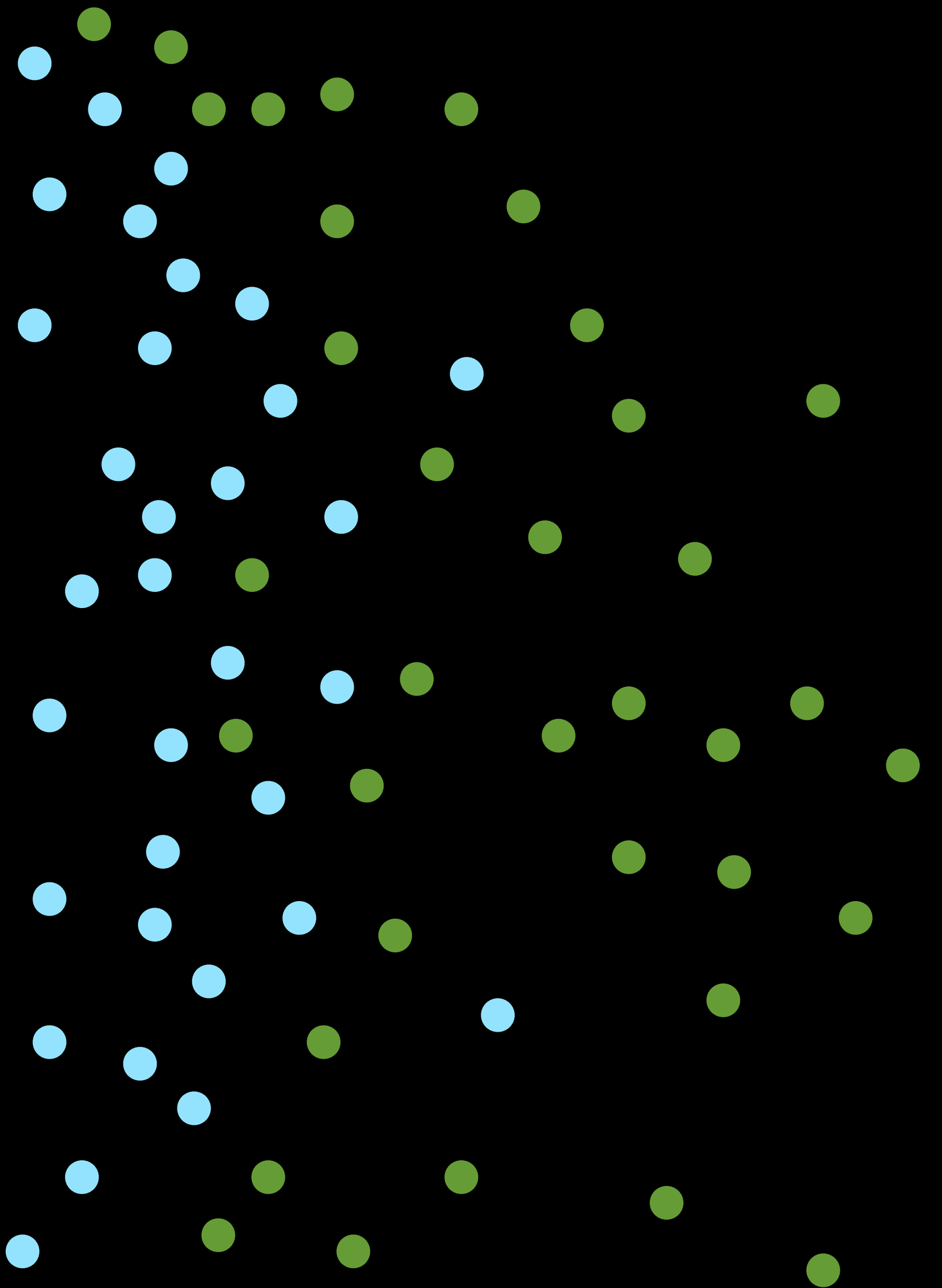


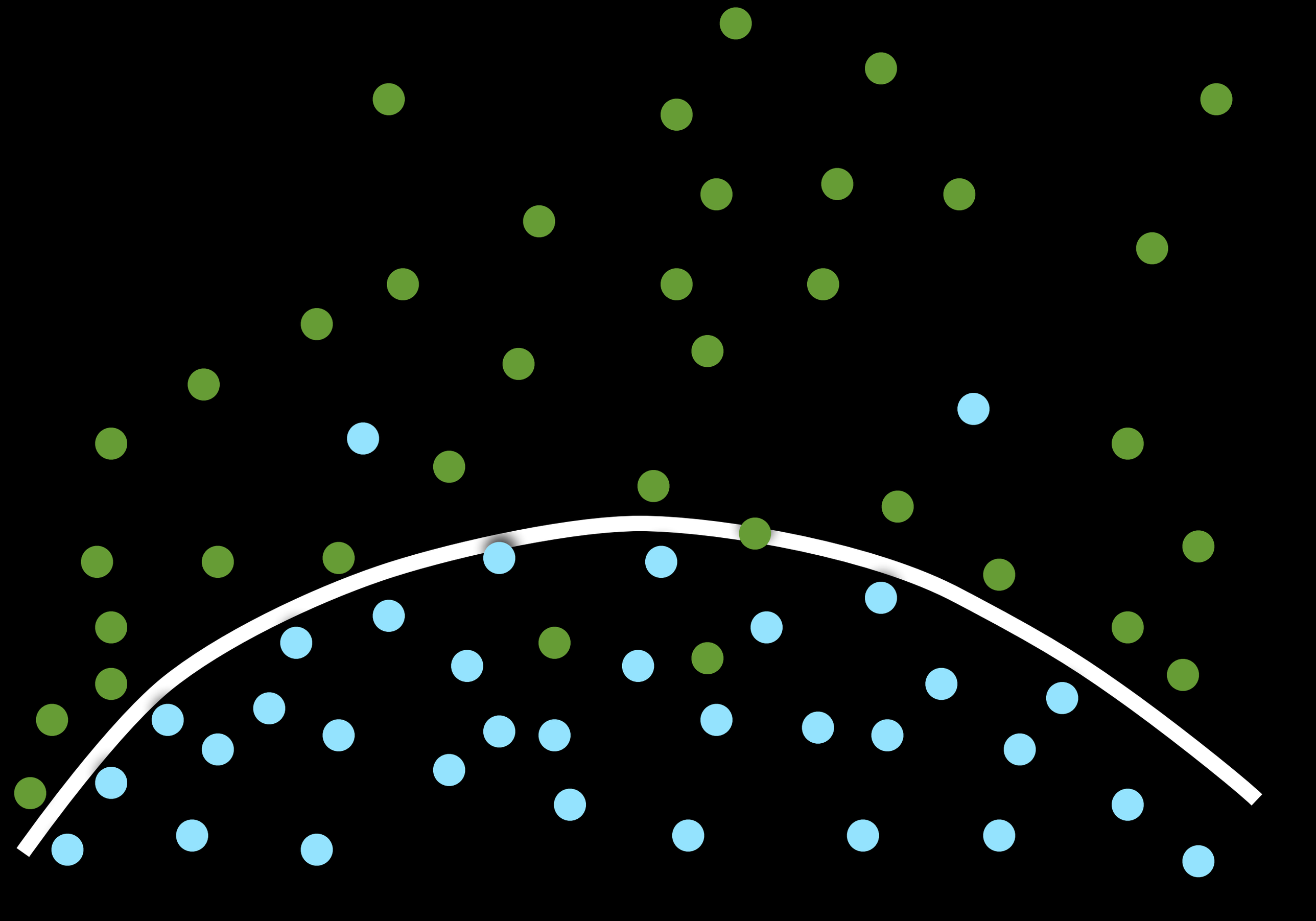


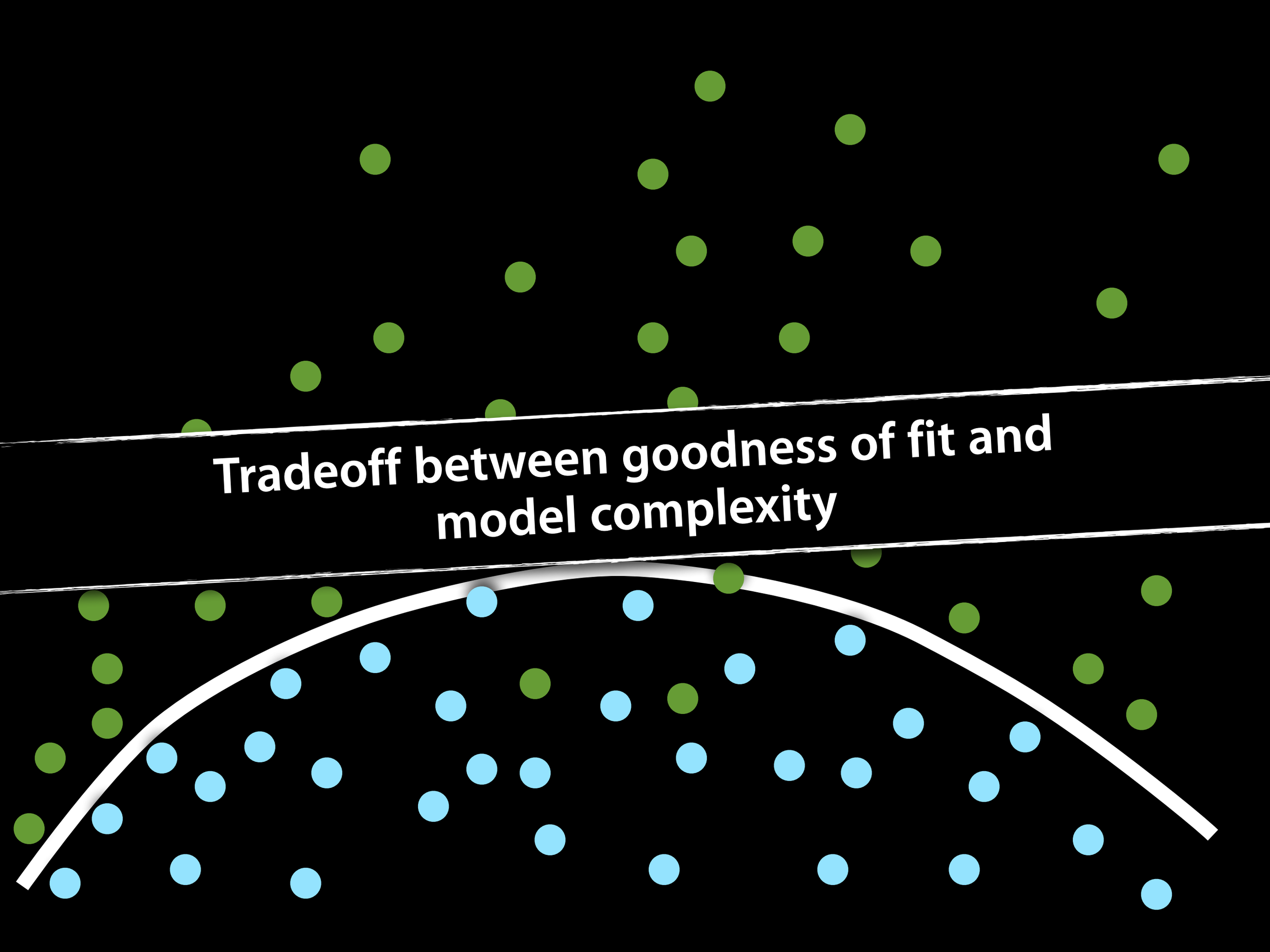
OVERFITTING

Model with too many parameters **overfits** the data








Tradeoff between goodness of fit and
model complexity

“All things being equal, the simplest solution tends to be the best one.”

— William of Ockham