Machine Learning - Part 1

Apr. 8, 2025

Recap question:

April 8, 2025

Is the strategy $0.4a_1 + 0.6a_2$ for player A and b_1 for player B a Nash equilibrium or would either player want to change their strategy?

Player B

Player A	b_1	b_2
a_1	(2,1)	(1,0)
a_2	(1,0)	(-3,1)

Recap question:

The reward for player A is

$$0.4 \times 2 + 0.6 \times 1 = 1.4$$

so player A would want to change their strategy to a_1.
It is therefore not a Nash equilibrium.

Player B

Player A	b_1	b_2
a_1	(2,1)	(1,0)
a_2	(1,0)	(-3,1)

How to find Mixed-strategy Nash equilibrium ?

When there are two players:

- **Step 1:** Player 1 needs to identify the mixed strategy that will make player 2's strategies have equal payoff.

- **Step 2:** Player 2 needs to identify the mixed strategy that will make player 1's strategies have equal payoff.

We find the mixed-strategy Nash equilibrium

		player 2	
		b_1	b_2
Player 1	a_1	(4, 7)	(-1, 8)
	a_2	(2, 4)	(0, 3)

Step 1: We find the strategy for player 1. Assume that player 1 plays strategy a_1 with probability p and a_2 with probability $1-p$. We compute the reward of player 2 if she uses strategy b_1 and the reward of player 2 if she uses strategy b_2 .

We find the mixed-strategy Nash equilibrium

		player 2	
		b_1	b_2
Player 1	a_1	(4, 7)	(-1, 8)
	a_2	(2, 4)	(0, 3)

If player 2 plays b_1, reward is $7p + 4(1-p) = 3p + 4$

If player 2 plays b_2, reward is $8p + 3(1-p) = 5p + 3$

We find the mixed-strategy Nash equilibrium

		player 2	
		b_1	b_2
Player 1	a_1	(4, 7)	(-1, 8)
	a_2	(2, 4)	(0, 3)

We now equate the two rewards and solve for p .

We find the mixed-strategy Nash equilibrium

		player 2 b_1	player 2 b_2
Player 1	a_1	(4, 7)	(-1, 8)
Player 1	a_2	(2, 4)	(0, 3)

We now equate the two rewards and solve for p.

If the two rewards are equal:

$$3p + 4 = 5p + 3 \Rightarrow p = 0.5$$

This means that the strategy for player 1 in the Nash equilibrium is $pa_1 + (1-p)a_2 = \boxed{0.5a_1 + 0.5a_2.}$

We find the mixed-strategy Nash equilibrium

		player 2 b_1	player 2 b_2
Player 1	a_1	(4, 7)	(-1, 8)
Player 1	a_2	(2, 4)	(0, 3)

Step 2: We find the strategy for player 2. Assume that player 2 plays strategy b_1 with probability q and b_2 with probability $1-q$. We compute the reward of player 1 if he uses strategy a_1 and the reward of player 1 if he uses strategy a_2 .

We find the mixed-strategy Nash equilibrium

		player 2	
		b_1	b_2
Player 1	a_1	(4, 7)	(-1, 8)
	a_2	(2, 4)	(0, 3)

If player 1 plays a_1, reward is $4q - (1-q) = 5q - 1$

If player 1 plays a_2, reward is $2q + 0(1-q) = 2q$

We find the mixed-strategy Nash equilibrium

		player 2	
		b_1	b_2
Player 1	a_1	(4, 7)	(-1, 8)
	a_2	(2, 4)	(0, 3)

We now equate the two rewards and solve for q .

If the two rewards are equal:

$$5q - 1 = 2q \Rightarrow q = 1/3$$

This means that the strategy for player 2 in the Nash equilibrium is $qb_1 + (1-q)b_2 = \boxed{\frac{1}{3}b_1 + \frac{2}{3}b_2.}$

player 2

Player 1	b_1	b_2	$\frac{b_1}{3} + \frac{2b_2}{3}$
a_1	(4, 7)	(-1, 8)	(2/3, 23/3)
a_2	(2, 4)	(0, 3)	(2/3, 10/3)
$\frac{a_1}{2} + \frac{a_2}{2}$	(3, 5.5)	(-0.5, 5.5)	(2/3, 5.5)

Nash equilibrium

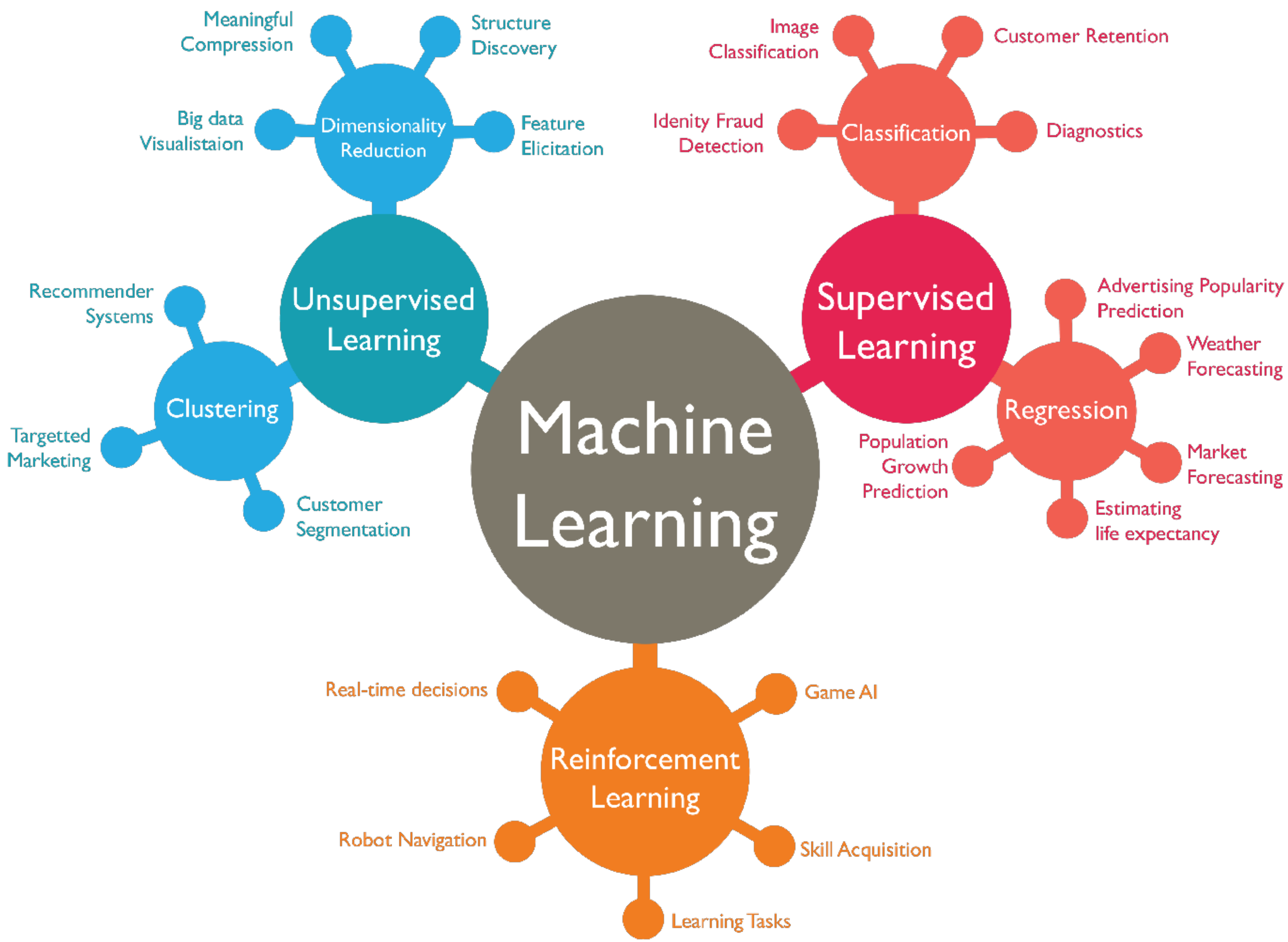
Machine Learning - Part 1

April 8, 2025

By the end of this lecture, you will be able to:

1. Explain binary classification
2. Explain regression
3. Define supervised and unsupervised learning

Machine learning (**ML**) is the study of computer algorithms that can improve automatically through experience and by the use of data



We will discuss:

1. Classification

Coins vs cats

We will discuss:

1. Classification
2. Facial recognition

We will discuss:

1. Classification
2. Facial recognition
3. Natural language processing

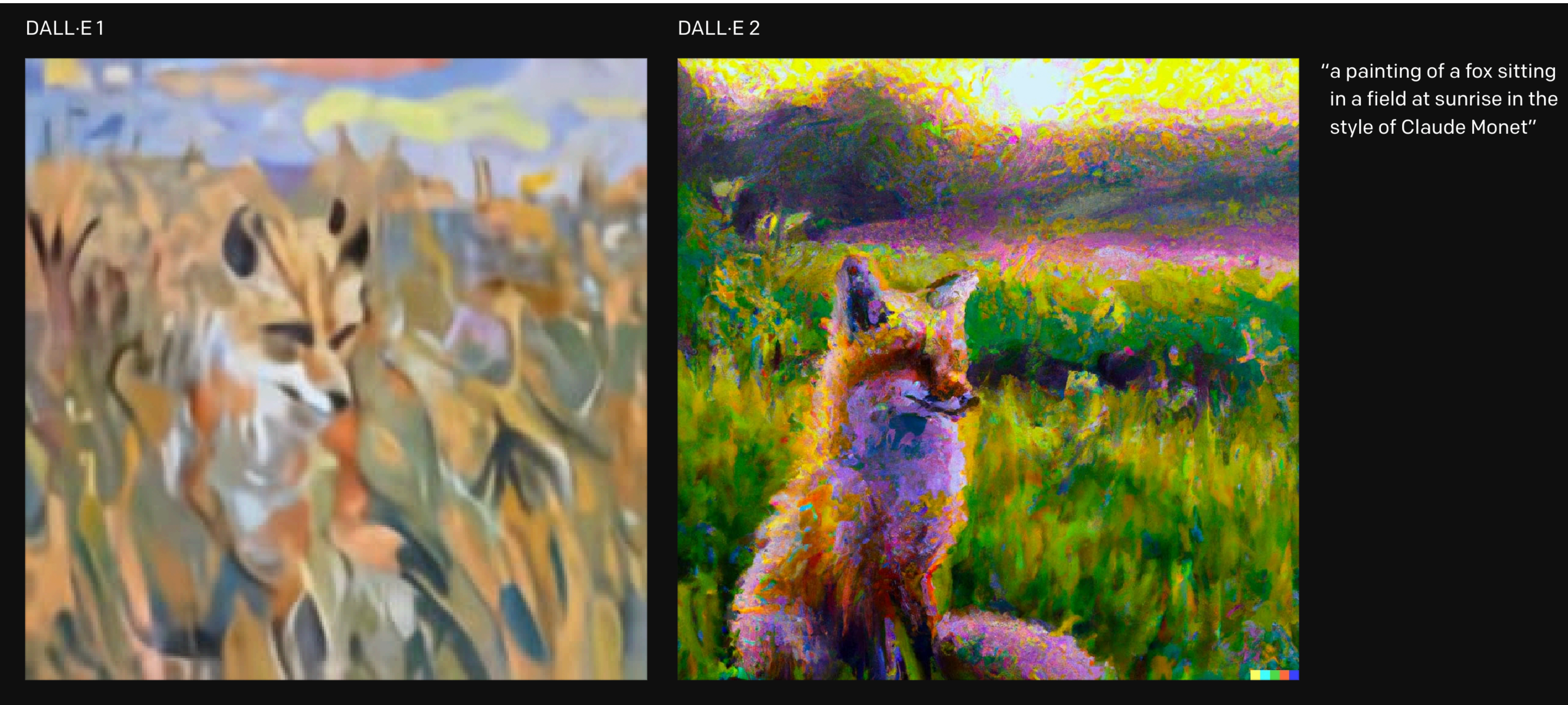
HM Explain why math is important

Math is a fundamental part of our lives and has a pervasive influence on the world around us. It helps us understand patterns, quantities, and relationships. It is the language of science, technology, economics, and many other fields. Here are a few reasons why math is important:

1. Math helps us reason and think logically: Doing math helps develop our problem-solving and critical thinking skills. It teaches us to break problems down into smaller parts, look for patterns, and come up with logical solutions.
2. Math is used in everyday life: From shopping and cooking to sports and finance, we use math in our daily activities. It helps us make informed decisions and understand the world around us.
3. Math is used in a variety of fields: Math is used in many different fields, including science, engineering, medicine, and finance. It is a critical tool for understanding and solving problems in these fields.
4. Math helps us make sense of the world: From predicting the weather to understanding the behavior of a virus, math helps us understand and make predictions about the world around us.

We will discuss:

1. Classification
2. Facial recognition
3. Natural language processing
4. Generative Adversarial Networks



Classification

How does your phone distinguish between two faces (or two objects)?

Coins vs cats

Classification

The coin looks gray and the cat not as gray, so we could look at the average pixel color and try to classify based on this one number.

Classification

The coin looks gray and the cat not as gray, so we could look at the average pixel color and try to classify based on this one number.

Assume that we have 5 pictures that we know are either cats or coins and we compute their average pixel colors

(0.6, coin)   (0.1, coin)   (3.8, cat)   (2.4, cat)   (0.8, cat)

Classification

The coin looks gray and the cat not as gray, so we could look at the average pixel color and try to classify based on this one number.

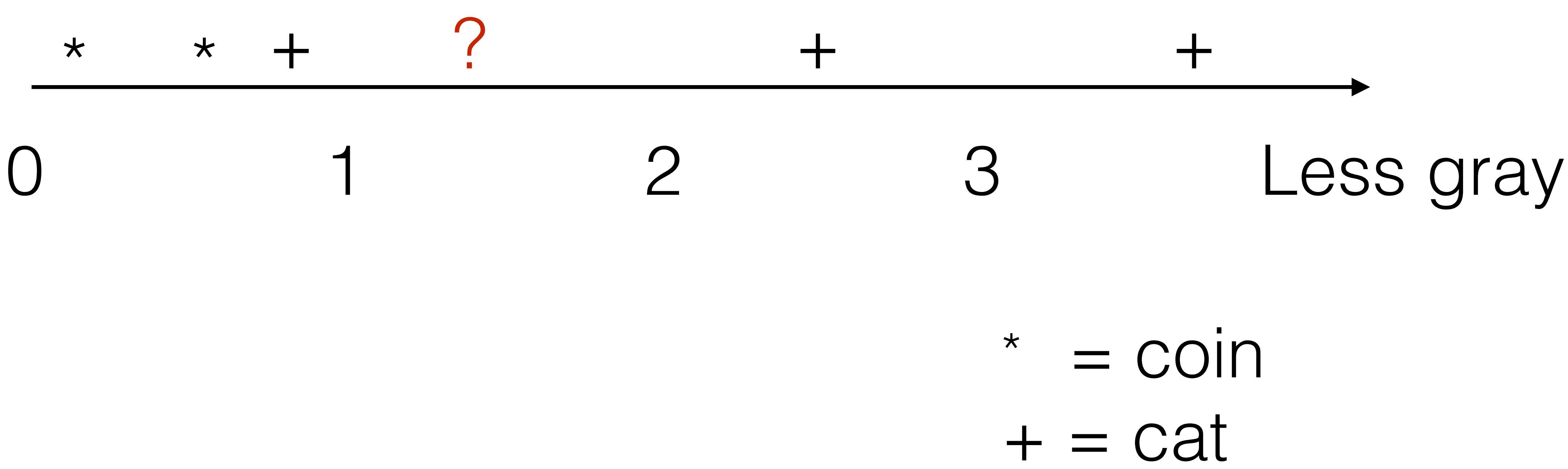
Assume that we have 5 pictures that we know are either cats or coins and we compute their average pixel colors

(0.6, coin) (0.1, coin) (3.8, cat) (2.4, cat) (0.8, cat)

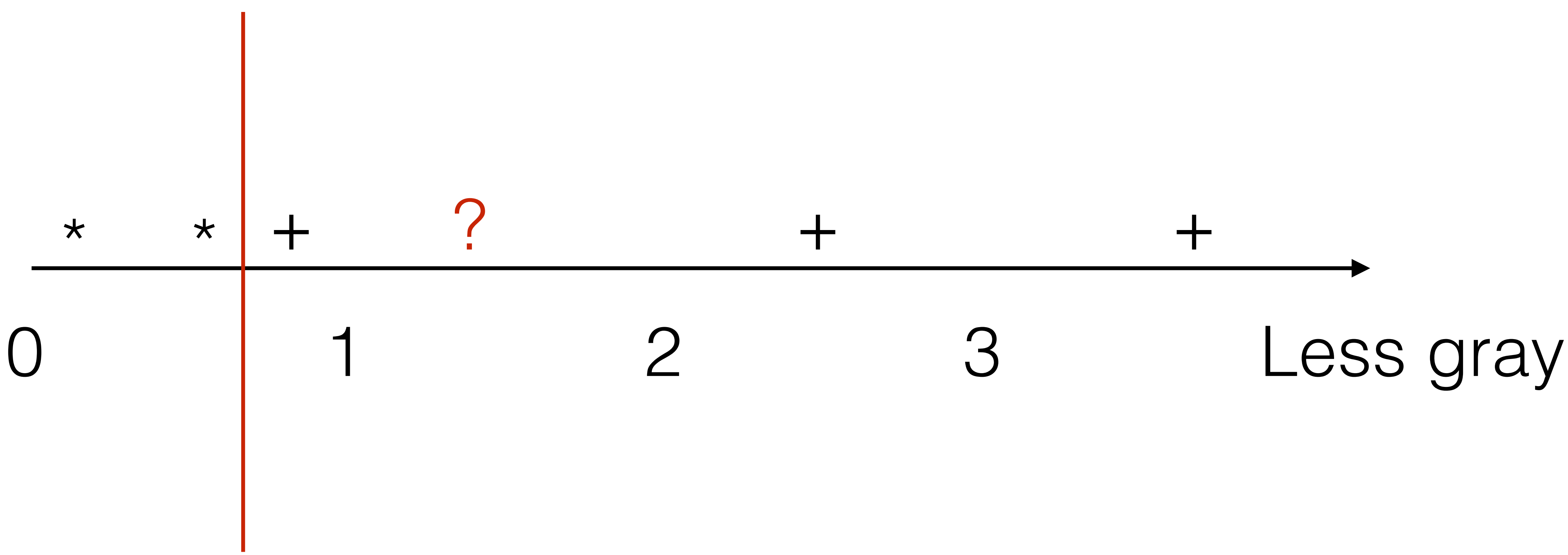
We then take a picture of a new object and the phone computes its average pixel color to 1.3. Should we call this a cat or a coin?

(0.6, coin) (0.1, coin) (3.8, cat) (2.4, cat) (0.8, cat)

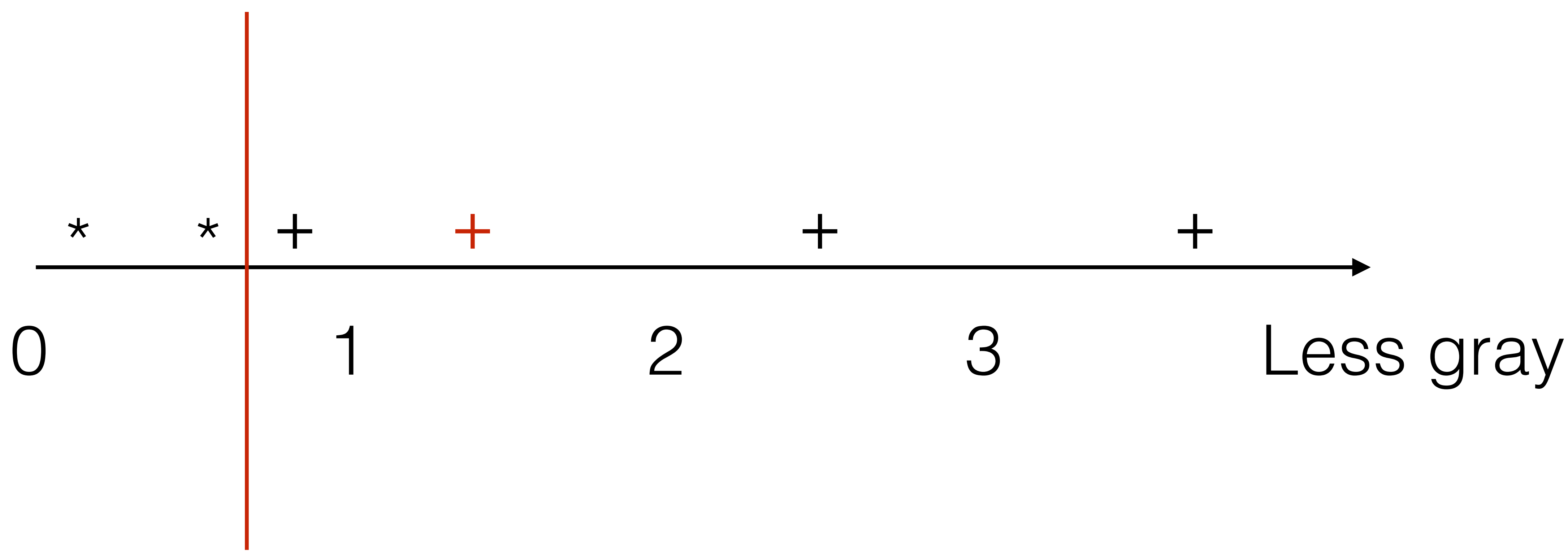
(1.3, ?)



Decision boundary: curve separating the cats from the coins. Everything to the left is classified as a coin, everything to the right is classified as a cat.

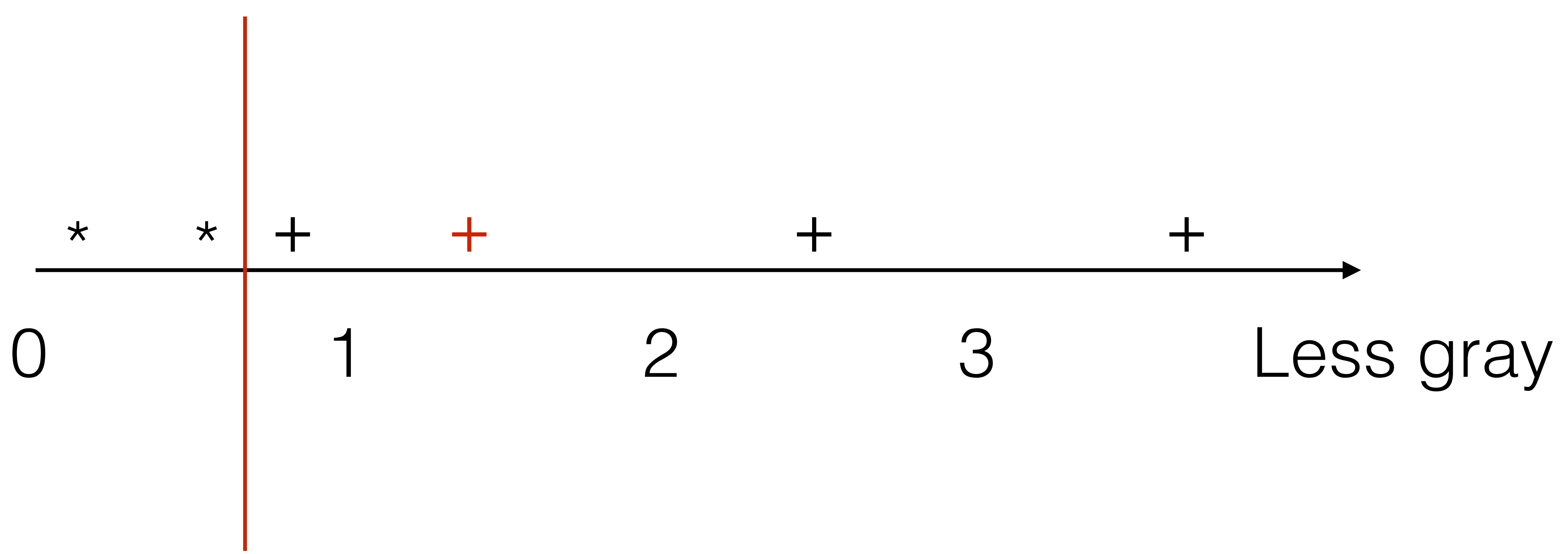


Decision boundary: curve separating the cats from the coins. Everything to the left is classified as a coin, everything to the right is classified as a cat.



This is an example of **binary classification:** classifying into two categories.

1-D case: from a single value, predict the binary label (* or +)



How can this fail?

How can this fail?

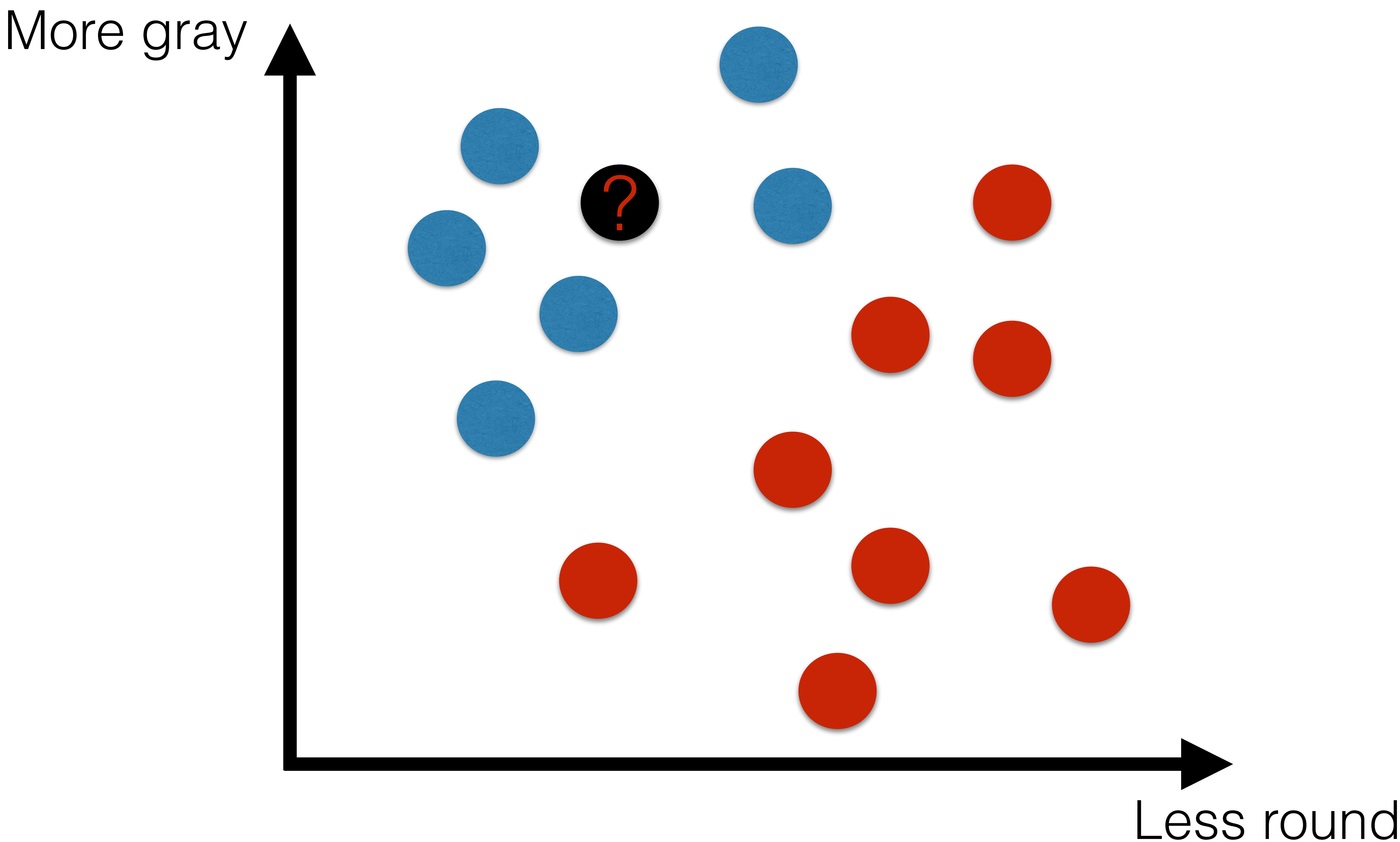
There are gray cats and they would be misclassified as coins.

How can this fail?

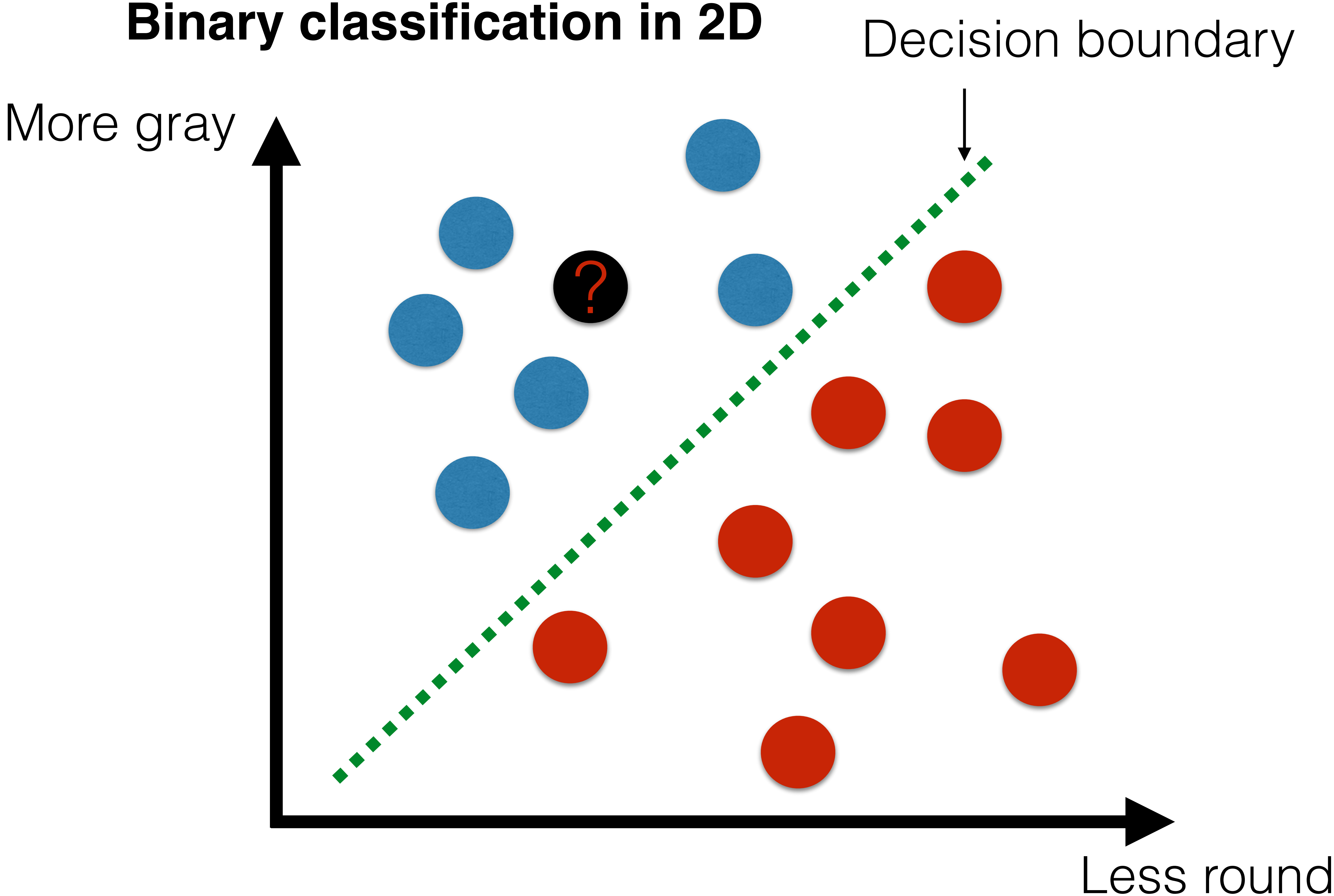
There are gray cats and they would be misclassified as coins.

To improve the classification algorithm, let’s assume the phone also measures how round the object is. The grayer and rounder, the more likely it is to be a coin.

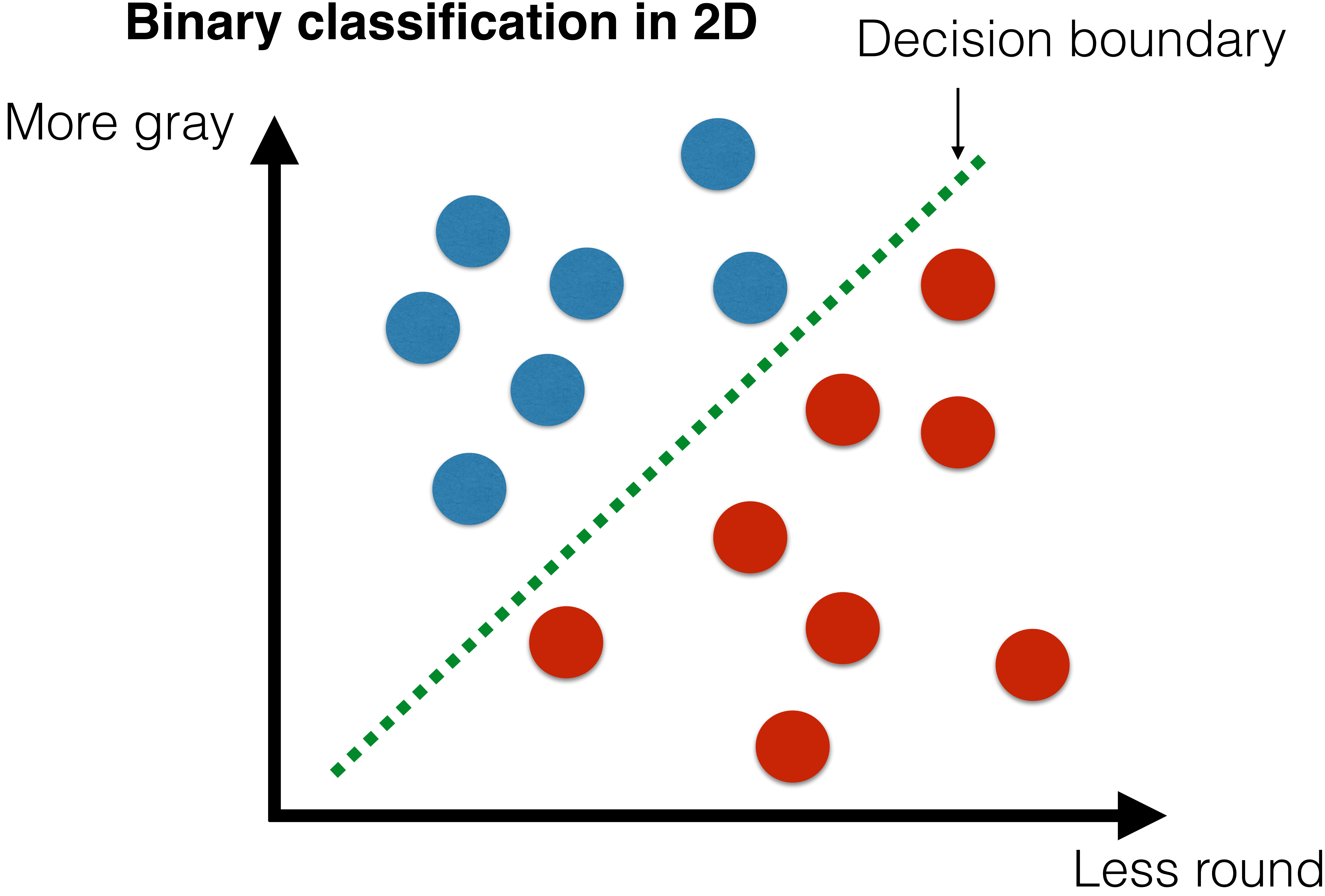
Binary classification in 2D



Binary classification in 2D



Binary classification in 2D



How can this fail?

How can this fail?

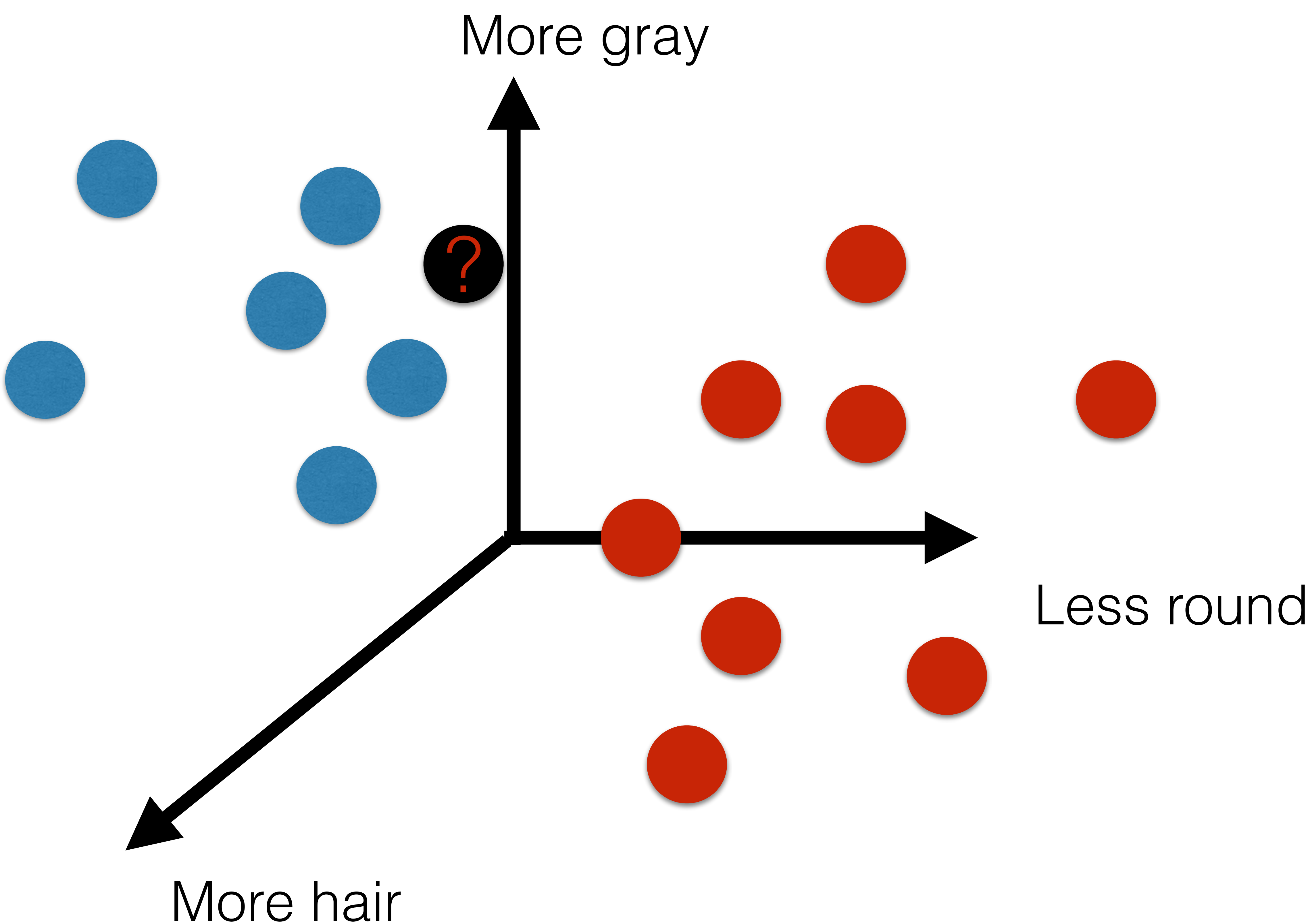
A gray cat curled up into a ball would be misclassified as a coin.

How can this fail?

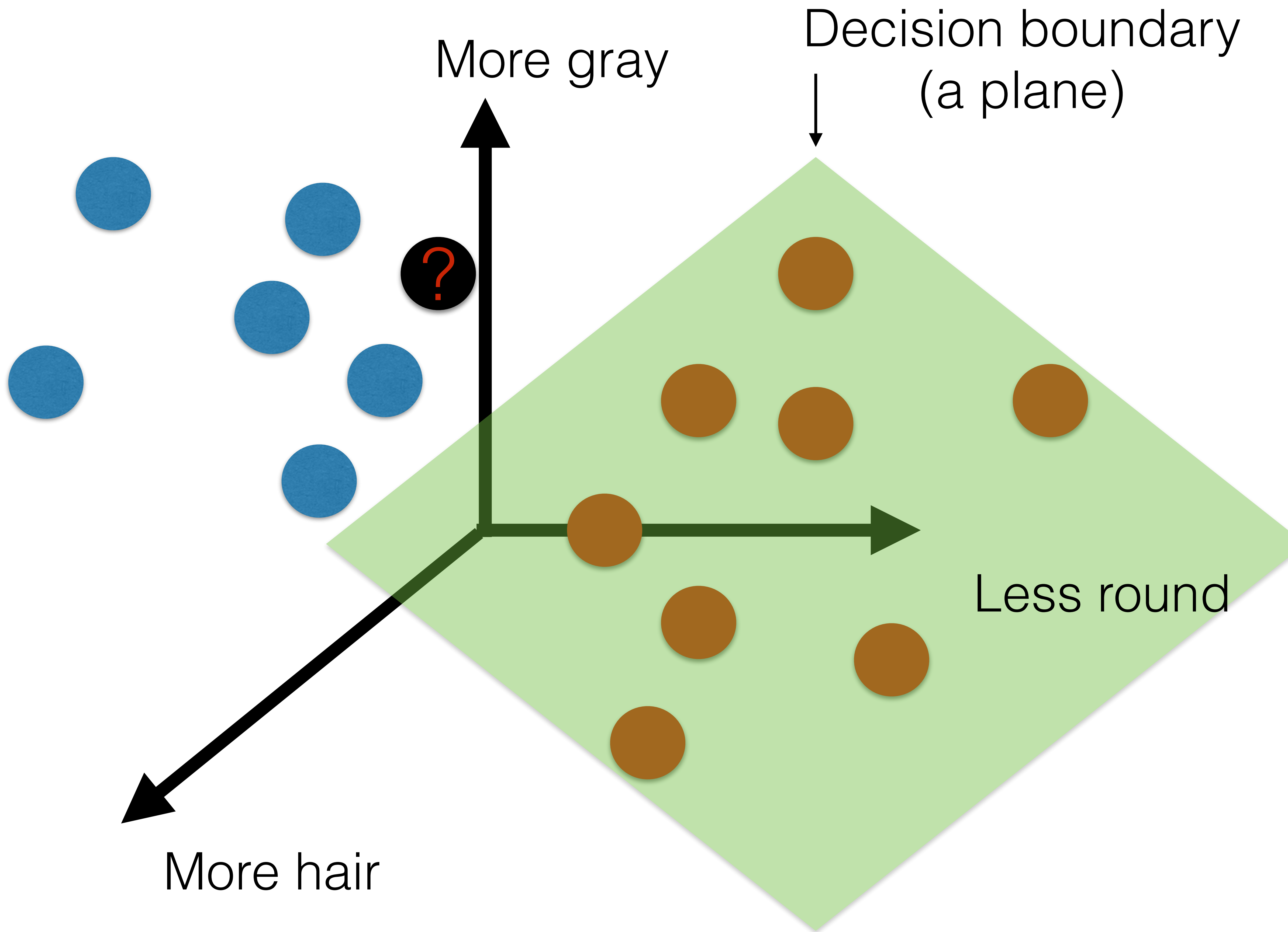
A gray cat curled up into a ball would be misclassified as a coin.

To improve the classification algorithm, let’s assume the phone also (somehow) measures how hairy the object is. The grayer and rounder and less hairy, the more likely it is to be a coin.

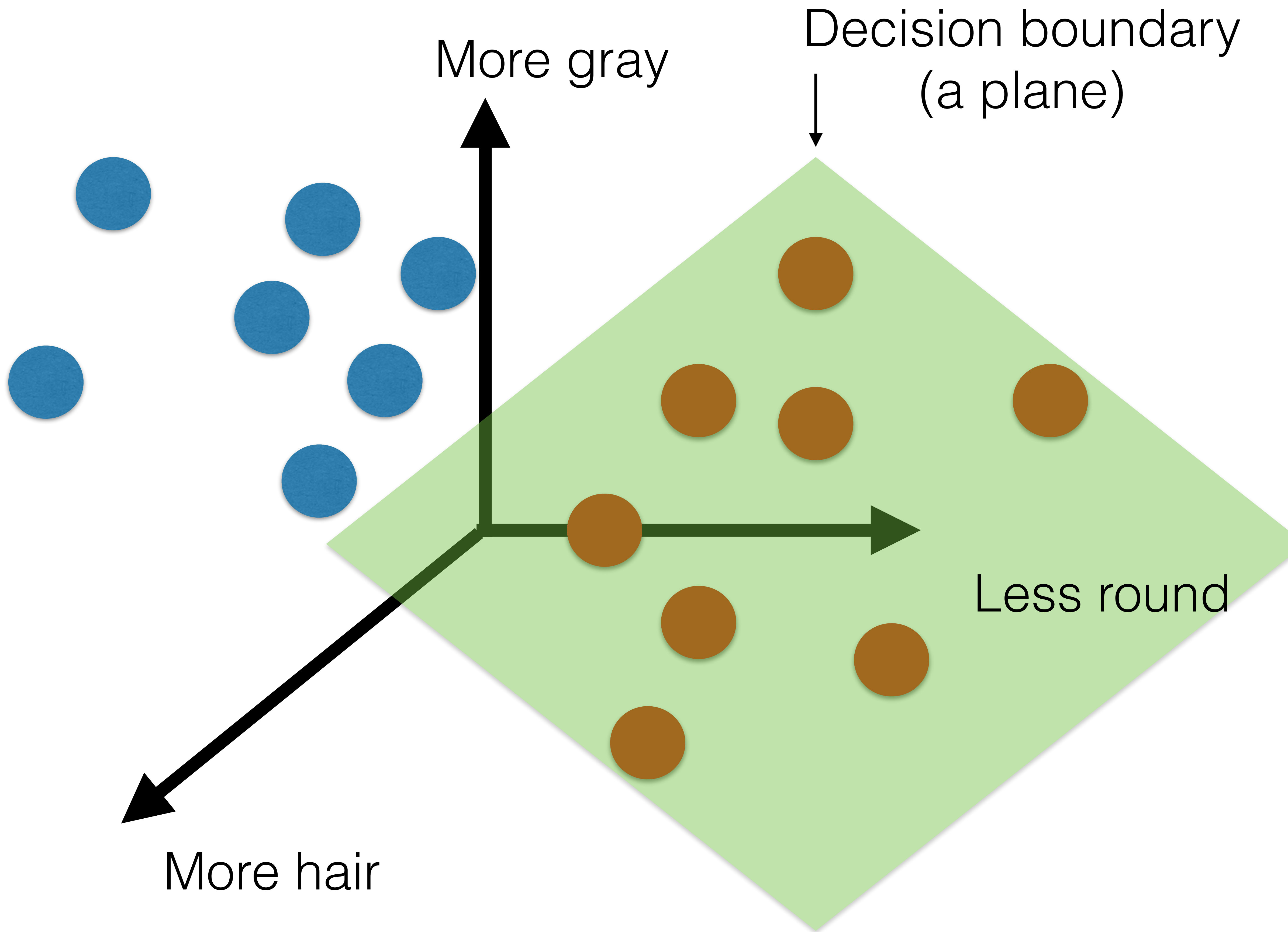
Binary classification in 3D



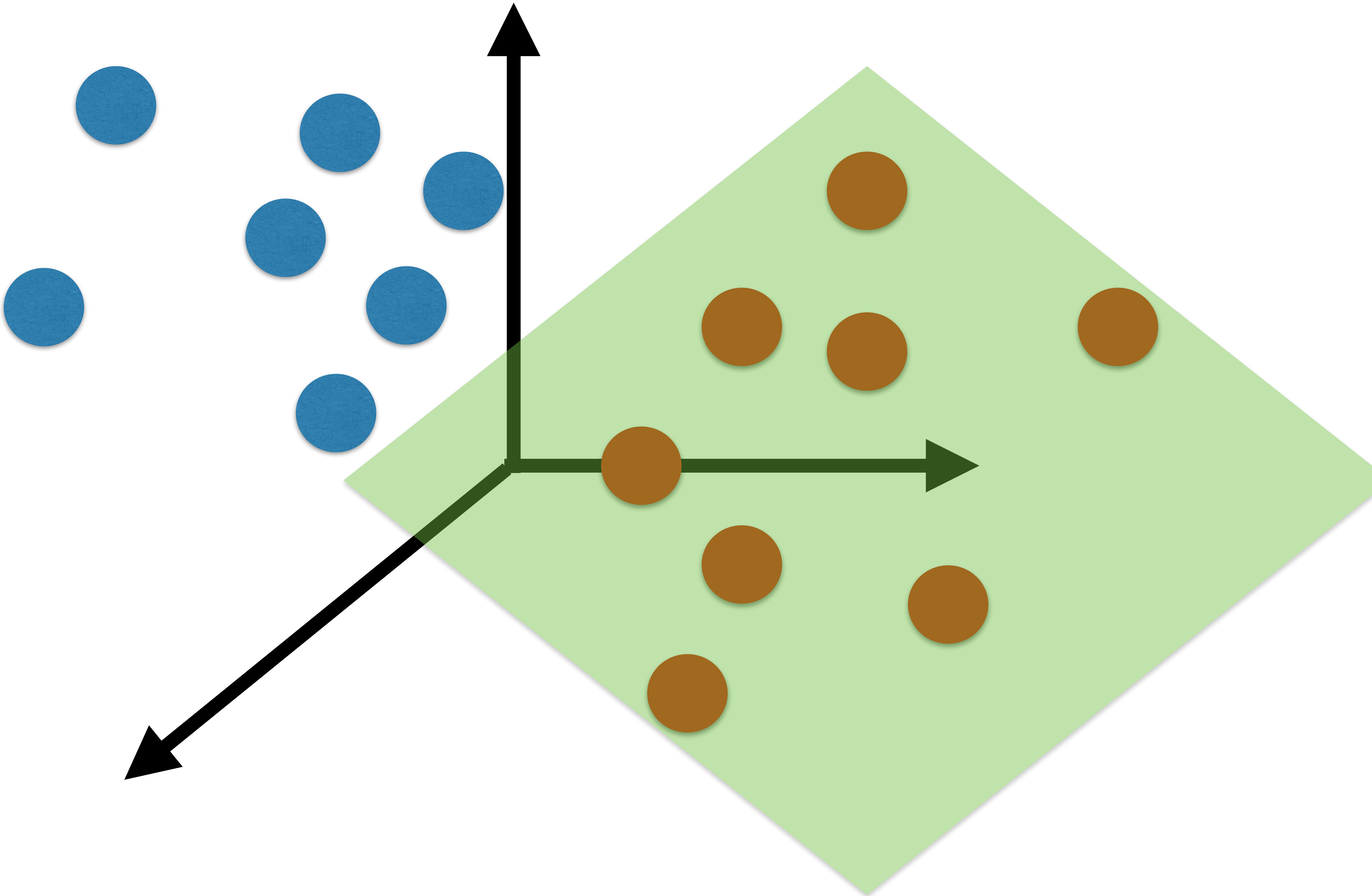
Binary classification in 3D



Binary classification in 3D



Learn from data to find the decision boundary



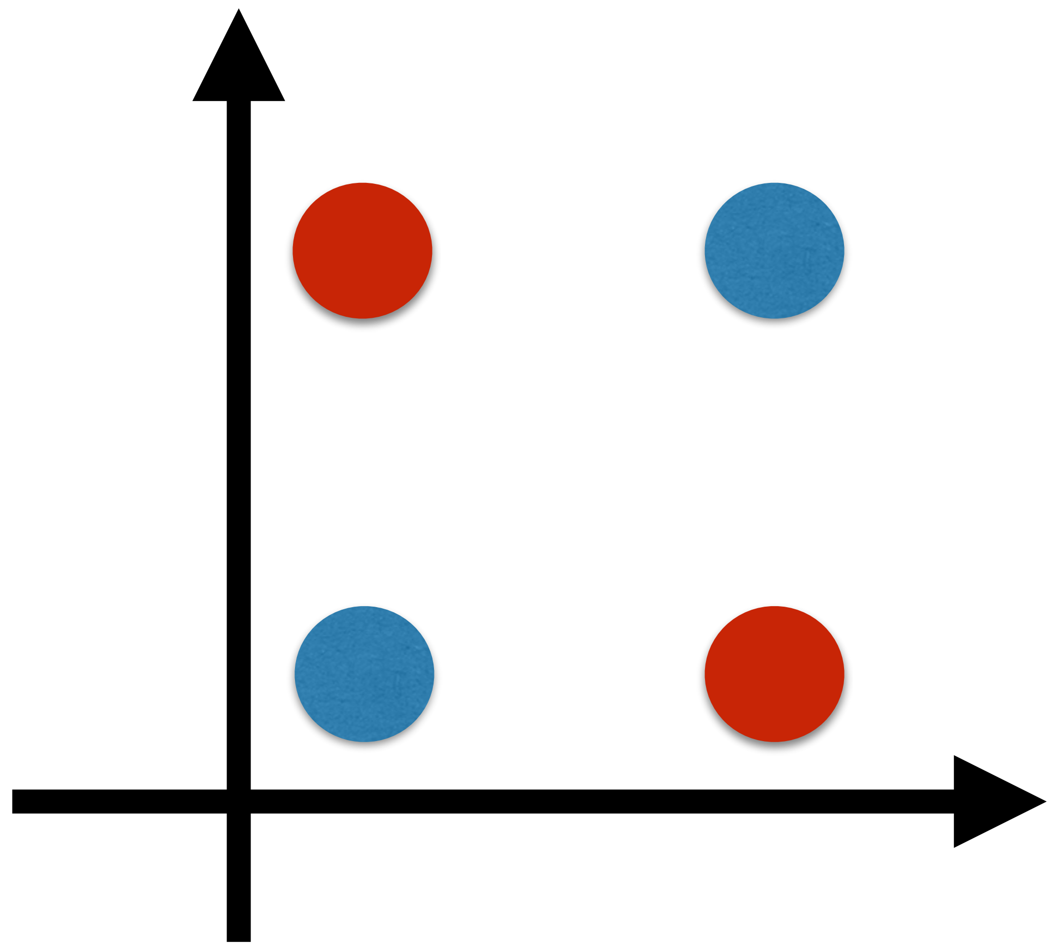
Learn from data to find the decision boundary

We could continue in the same way to improve the classification algorithm by:

1. Adding more and more distinguishing features of cats vs coins
2. Coming up with decision boundaries

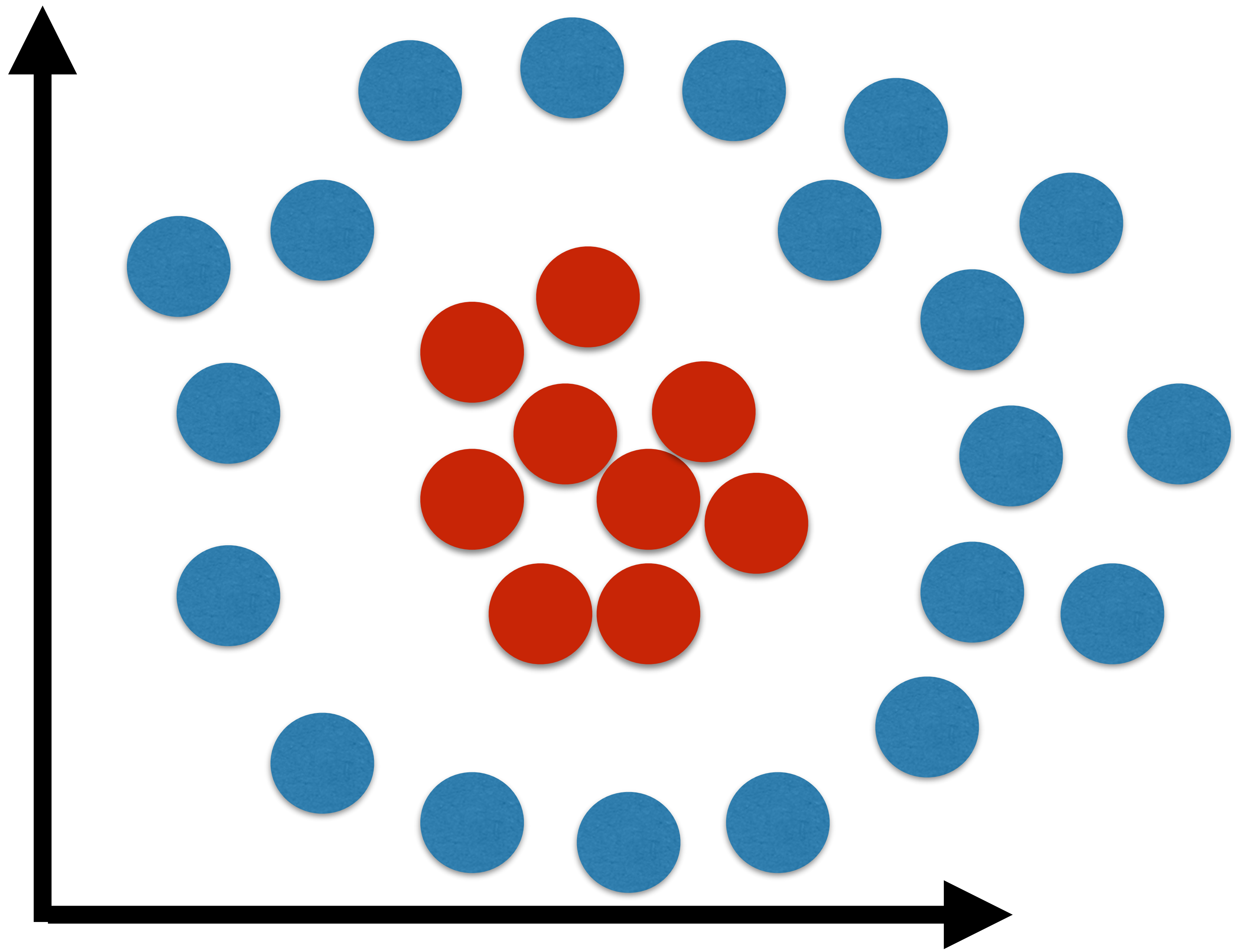
Observation 1: the choice of decision boundary is important!

Can you find any straight line to separate red from blue?



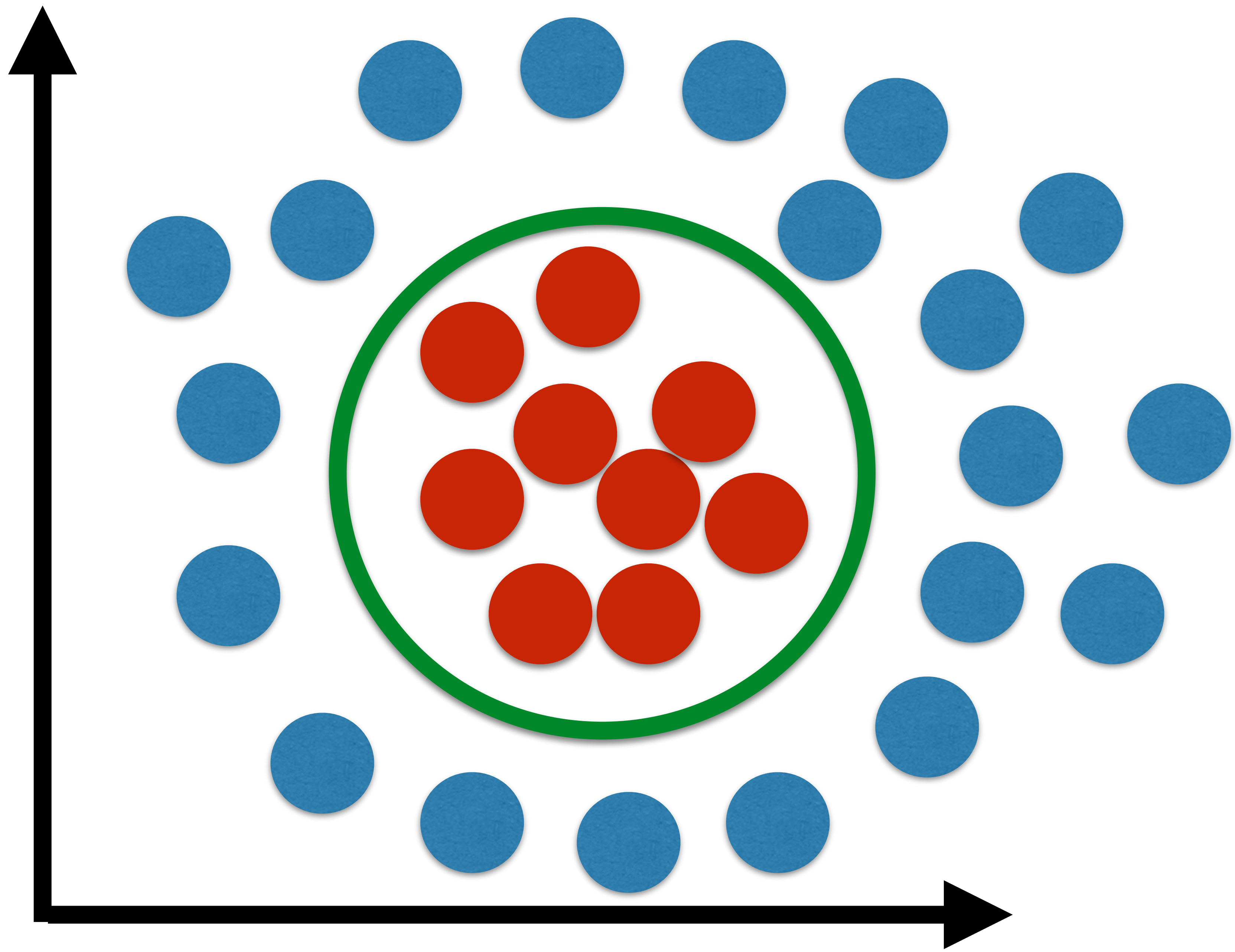
Observation 1: the choice of decision boundary is important!

Can you find a decision boundary to separate red from blue?



Observation 1: the choice of decision boundary is important!

Can you find a decision boundary to separate red from blue?



Observation 2: data is important!

(0.6, coin) (0.1, coin) (3.8, cat) (2.4, cat) (0.8, cat)

Assume that we look at some more pictures and see the following average pixel colors:

(1.5, coin) (4, coin) (0.4, cat) (1.0, cat)

Observation 2: data is important!

(0.6, coin)   (0.1, coin)   (3.8, cat)   (2.4, cat)   (0.8, cat)

Assume that we look at some more pictures and see the following average pixel colors:

(1.5, coin)   (4, coin)   (0.4, cat)   (1.0, cat)

How do we classify (1.3, ?) now?

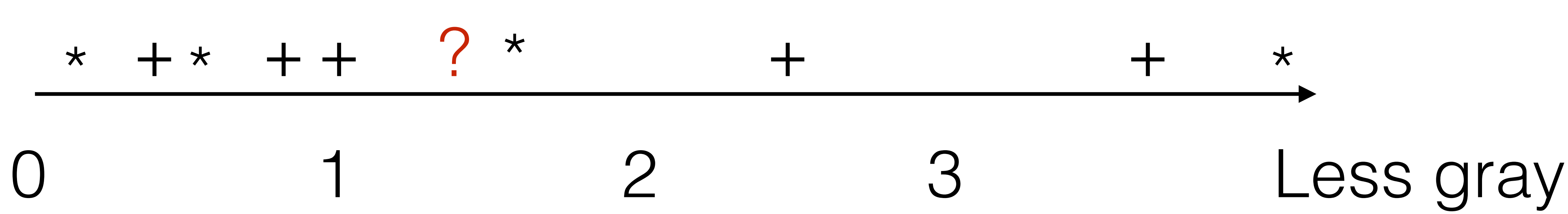
Observation 2: data is important!

(0.6, coin)   (0.1, coin)   (3.8, cat)   (2.4, cat)   (0.8, cat)

Assume that we look at some more pictures and see the following average pixel colors:

(1.5, coin)   (4, coin)   (0.4, cat)   (1.0, cat)

How do we classify (1.3, ?) now?



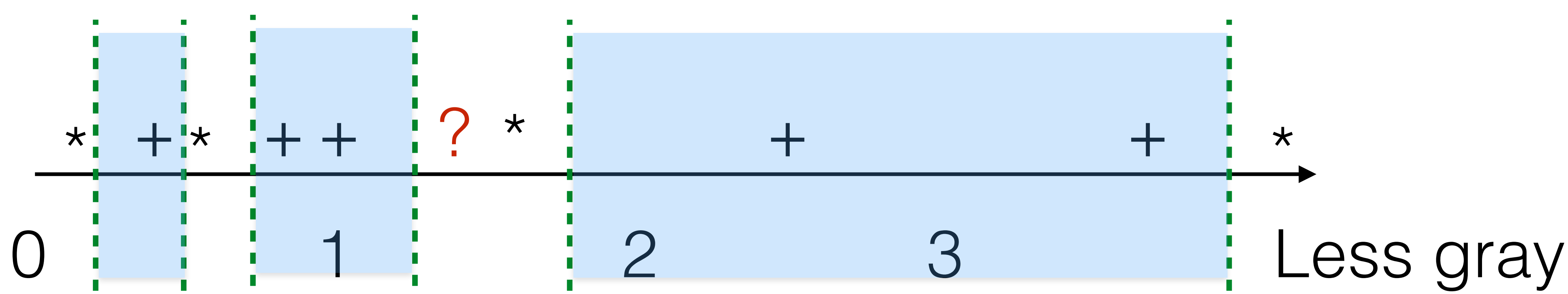
Observation 2: data is important!

(0.6, coin)   (0.1, coin)   (3.8, cat)   (2.4, cat)   (0.8, cat)

Assume that we look at some more pictures and see the following average pixel colors:

(1.5, coin)   (4, coin)   (0.4, cat)   (1.0, cat)

How do we classify (1.3, ?) now?



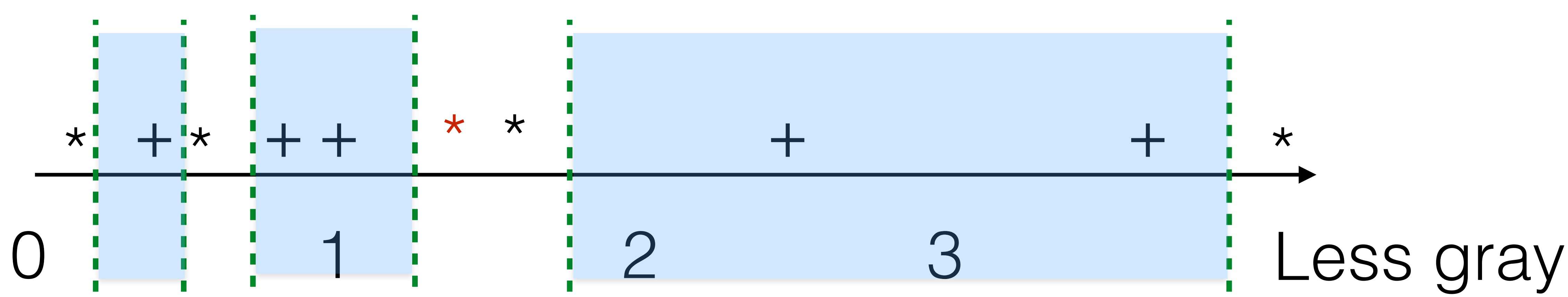
Observation 2: data is important!

(0.6, coin)   (0.1, coin)   (3.8, cat)   (2.4, cat)   (0.8, cat)

Assume that we look at some more pictures and see the following average pixel colors:

(1.5, coin)   (4, coin)   (0.4, cat)   (1.0, cat)

How do we classify (1.3, ?) now?



Applications of Classification

1. Facial recognition
2. Fraud detection
3. Spam email filter
4. Sentiment analysis

Applications of Classification

1. Facial recognition
2. Fraud detection
3. Spam email filter
4. Sentiment analysis

Example: Is “You should see their decadent dessert menu” a positive or negative statement?

Regression

We want to investigate the effect of the number of hours math majors study (during the entire term) on their grade

Regression

We want to investigate the effect of the number of hours math majors study (during the entire term) on their grade

We sample five math majors and get the following data:
Student 1: 0 hours, 0 points on the final (out of 100)
Student 2: 2 hours, 1 points
Student 3: 20 hours, 11 points
Student 4: 40 hours, 19 points
Student 5: 100 hours, 53 points

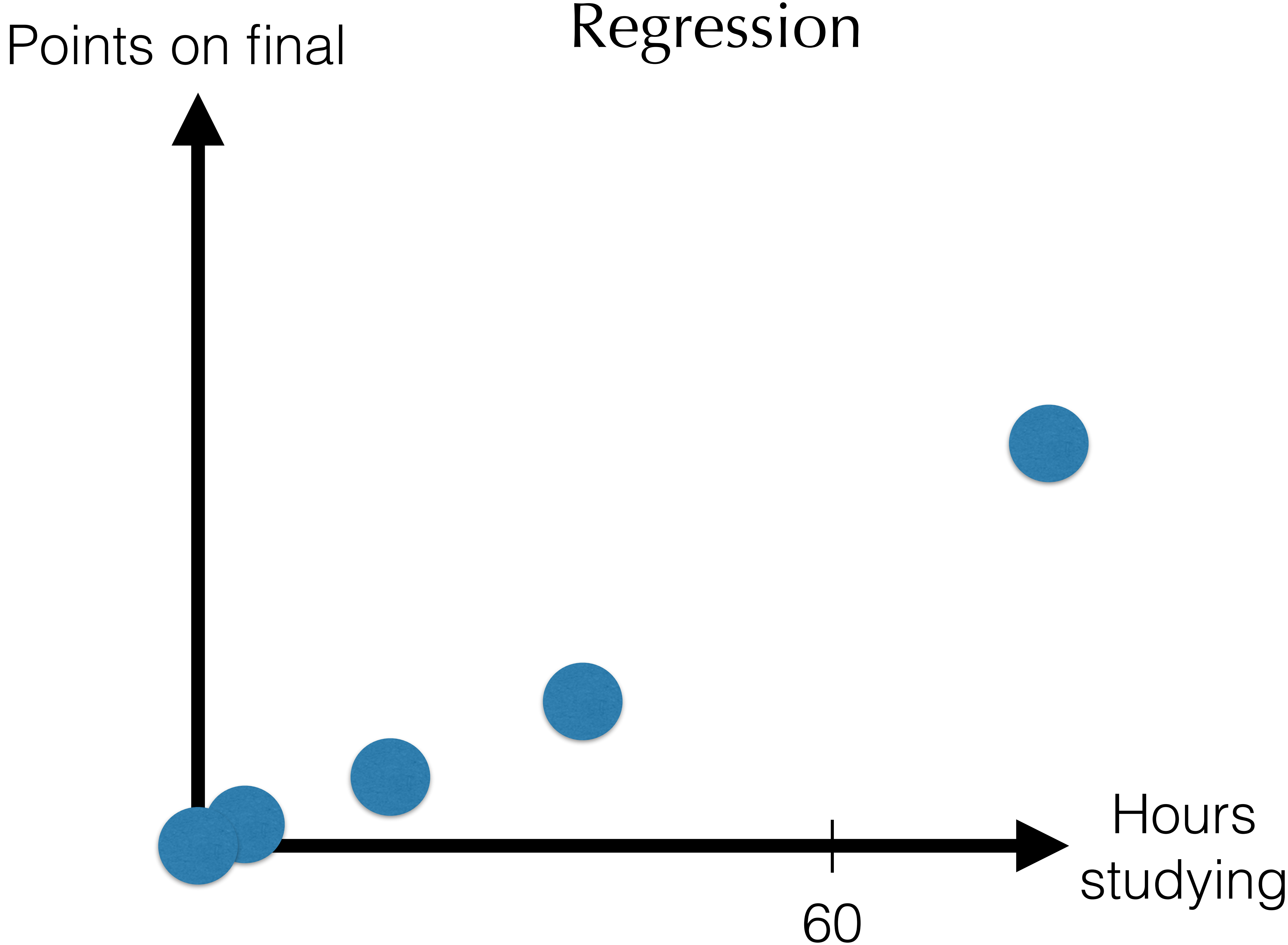
Regression

We want to investigate the effect of the number of hours math majors study (during the entire term) on their grade

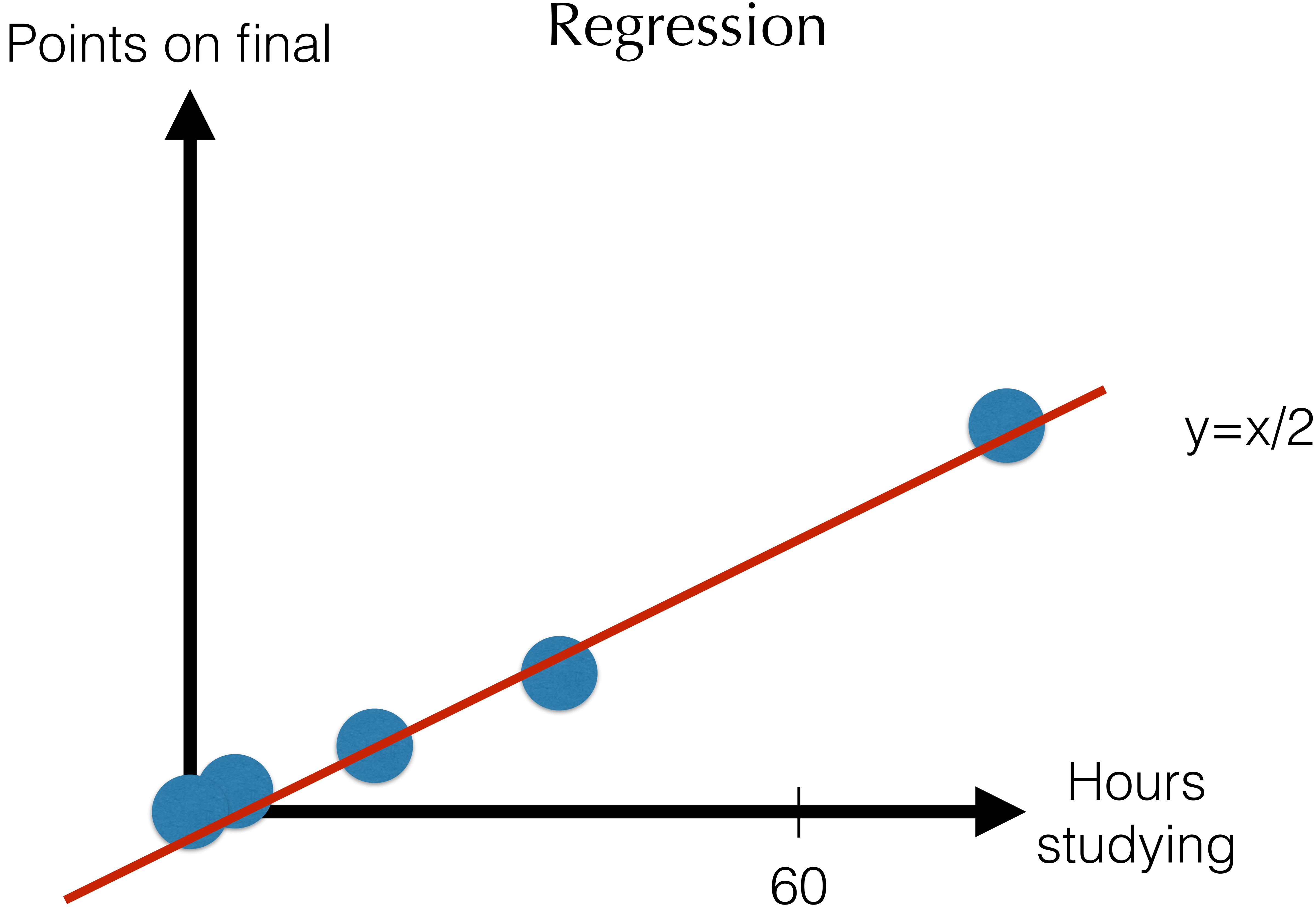
We sample five math majors and get the following data:
Student 1: 0 hours, 0 points on the final (out of 100)
Student 2: 2 hours, 1 points
Student 3: 20 hours, 11 points
Student 4: 40 hours, 19 points
Student 5: 100 hours, 53 points

From this data, how many points would we expect for a student studying 60 hours?

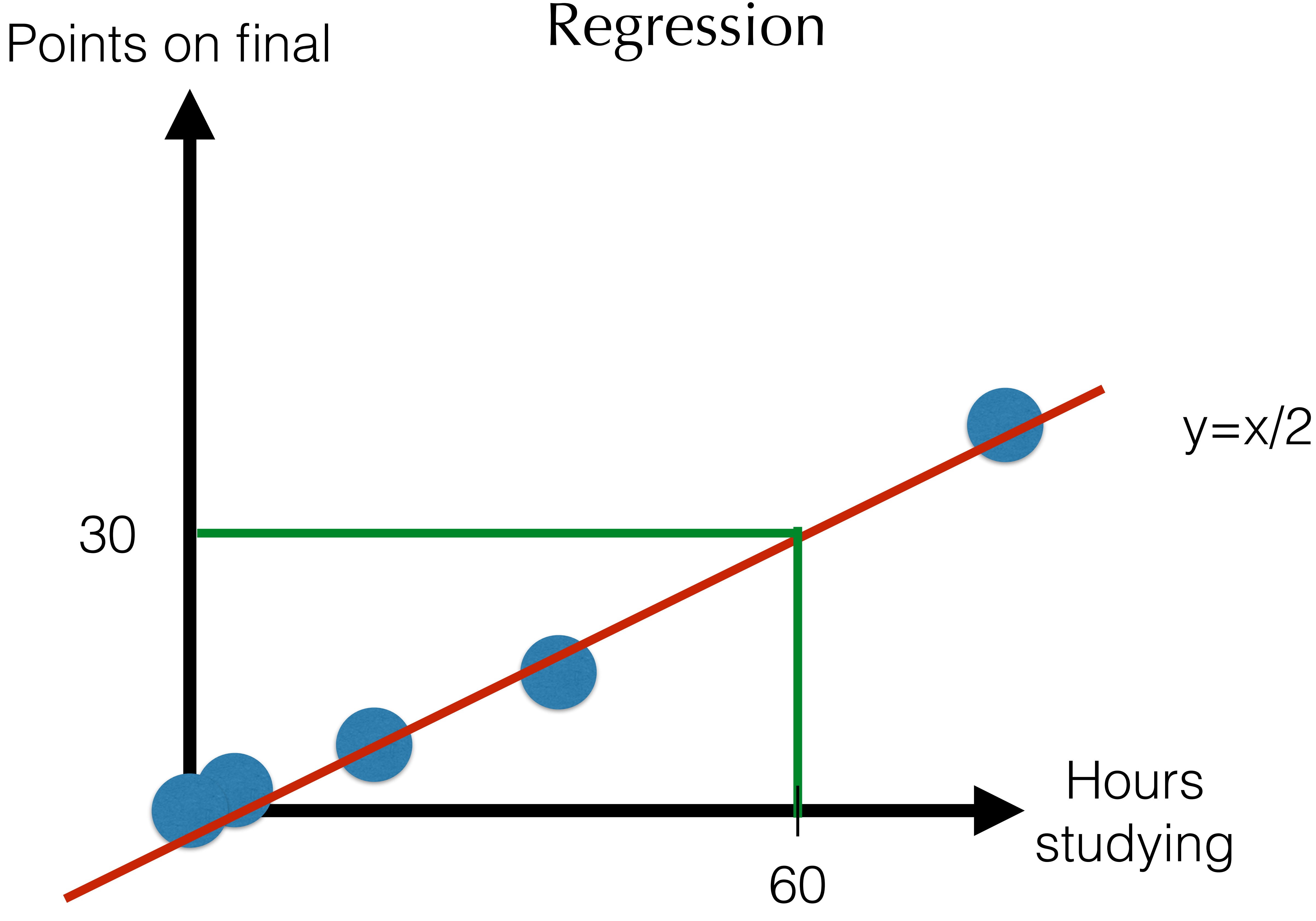
Regression



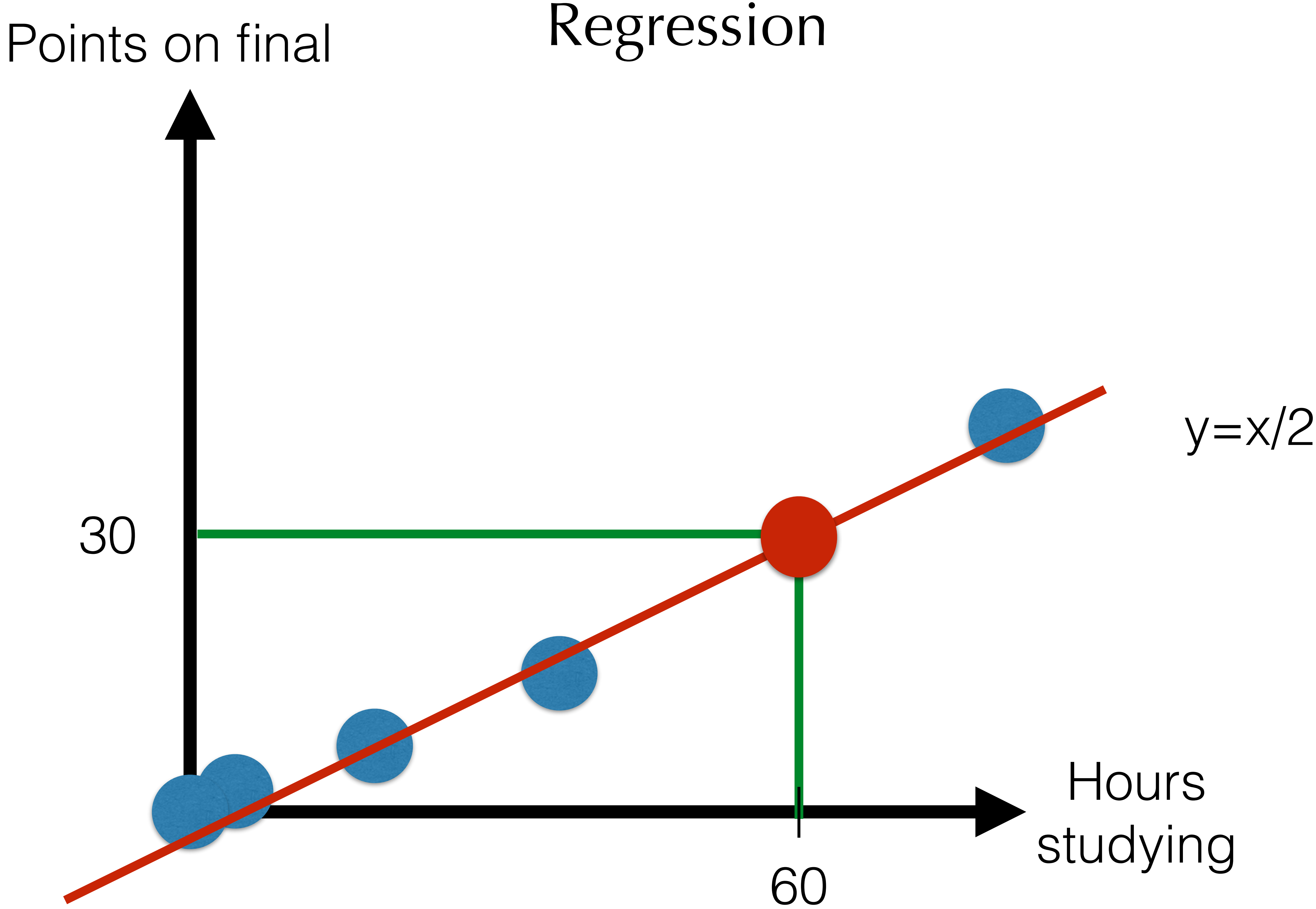
Regression



Regression



Regression



Regression

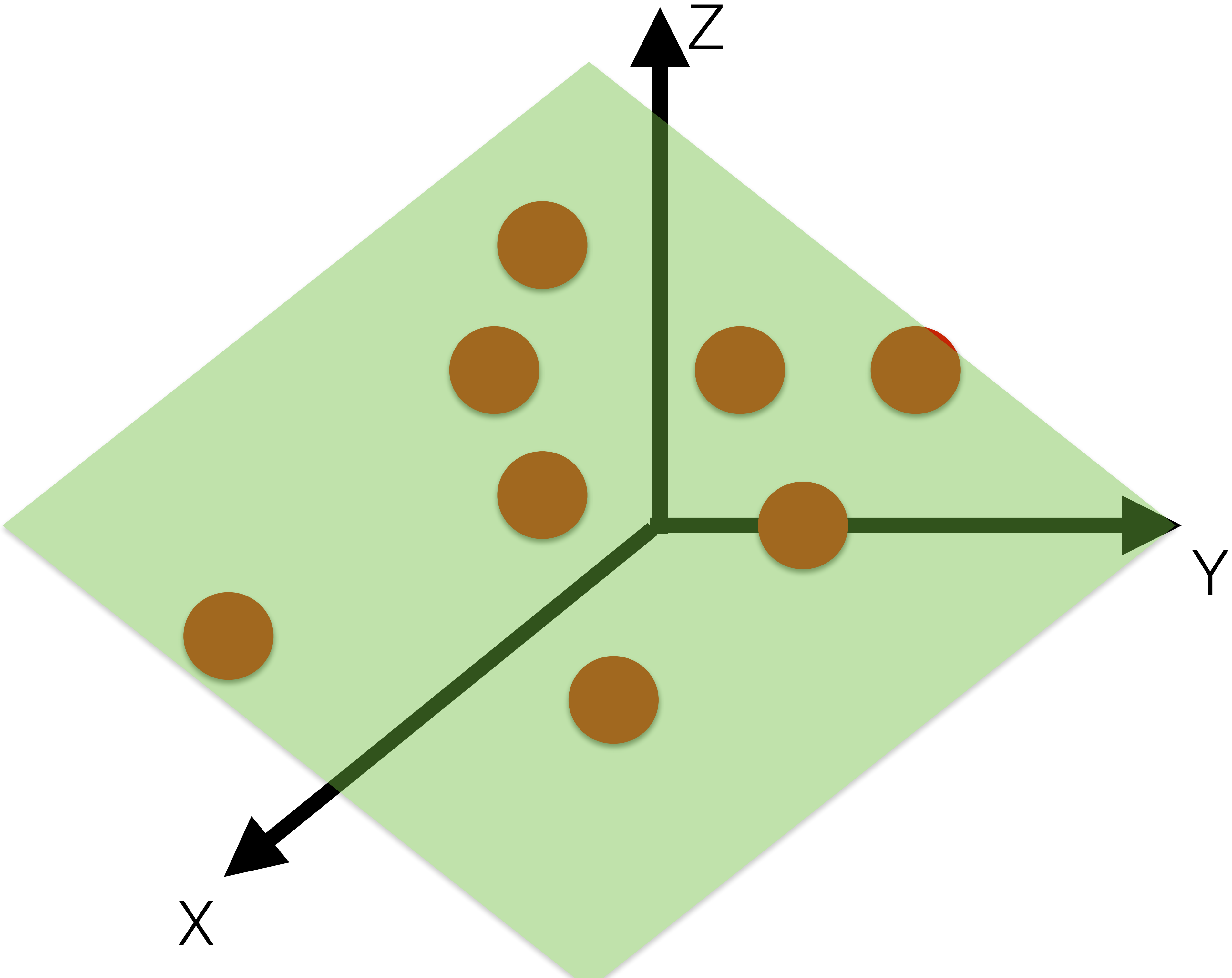
The **response variables** are what we are interested in and are unknown

The **independent variables** (factors/ covariates) are the known variables

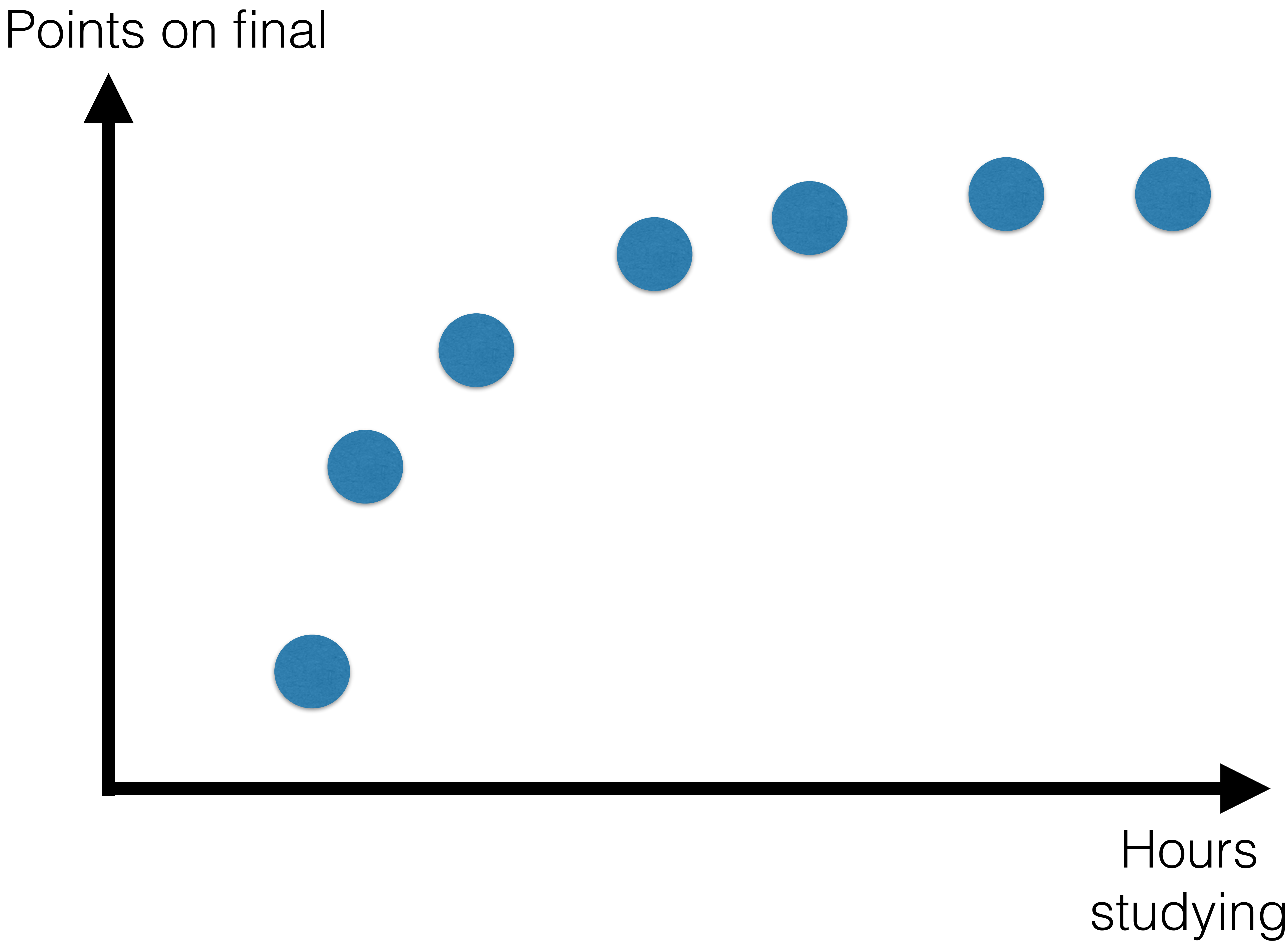
Regression: estimating response variable(s) from the independent variable(s).

Linear regression: regression using linear functions (= straight lines)

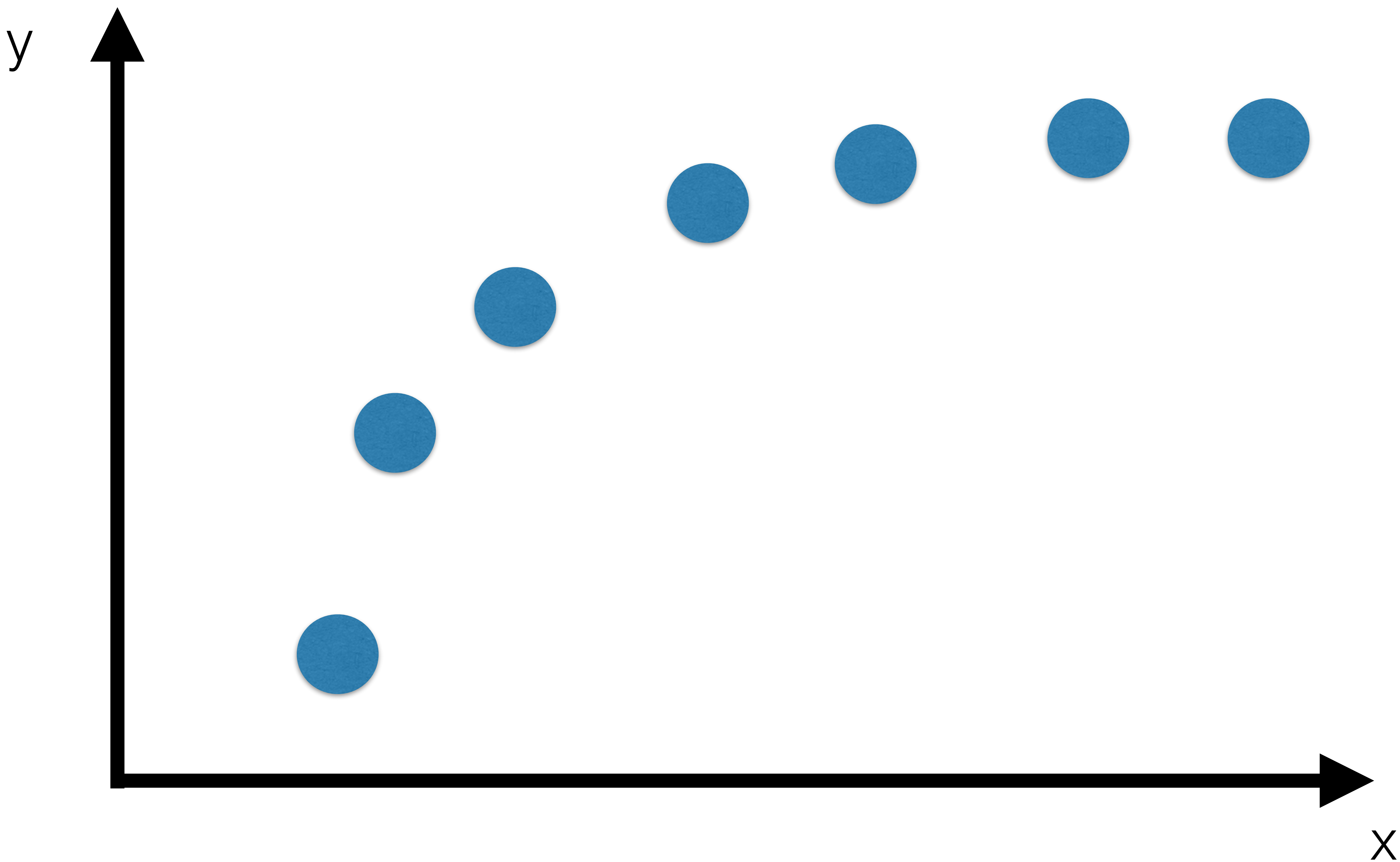
What about 2-D? Estimate Z from X and Y



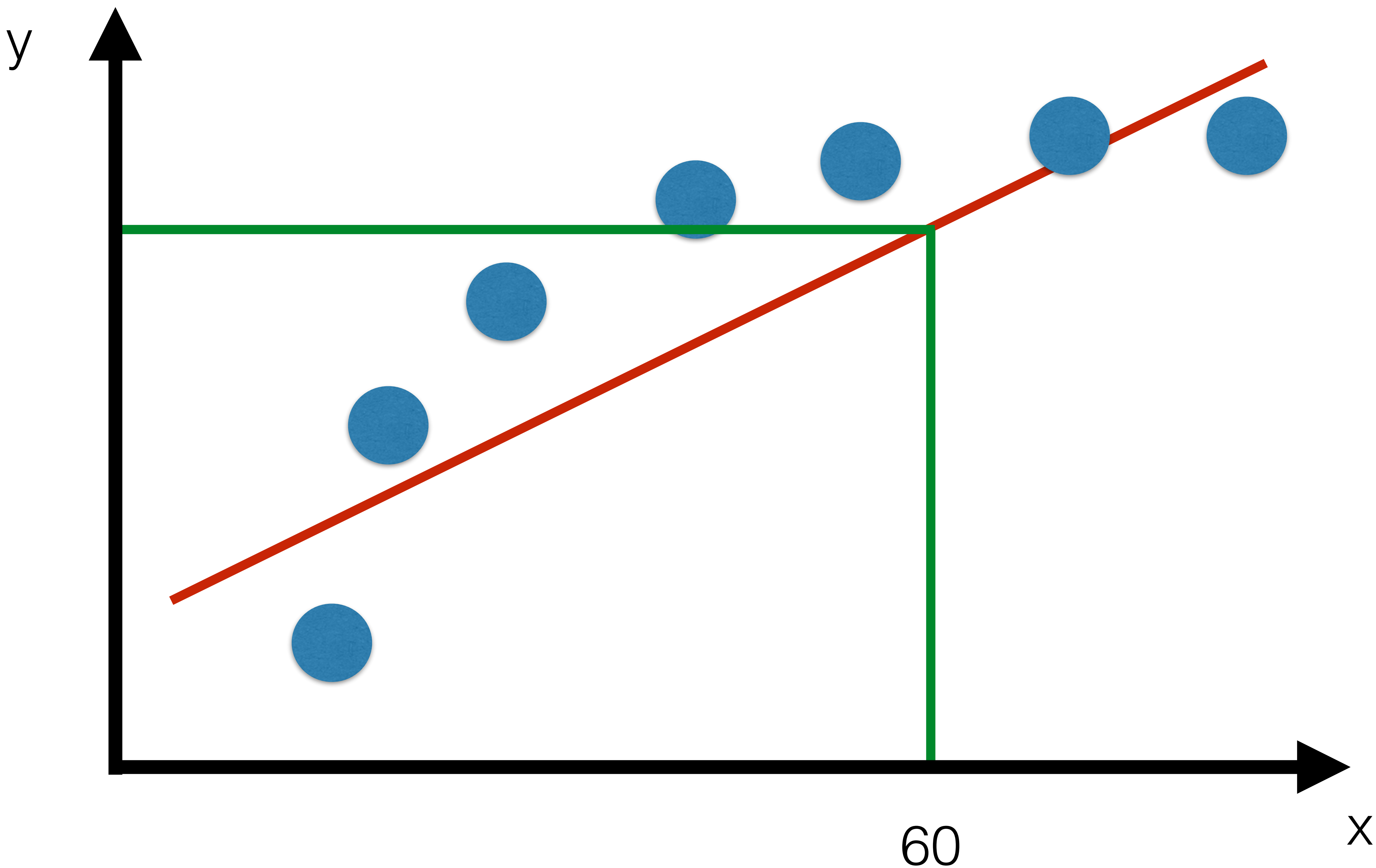
Non-linear Regression:



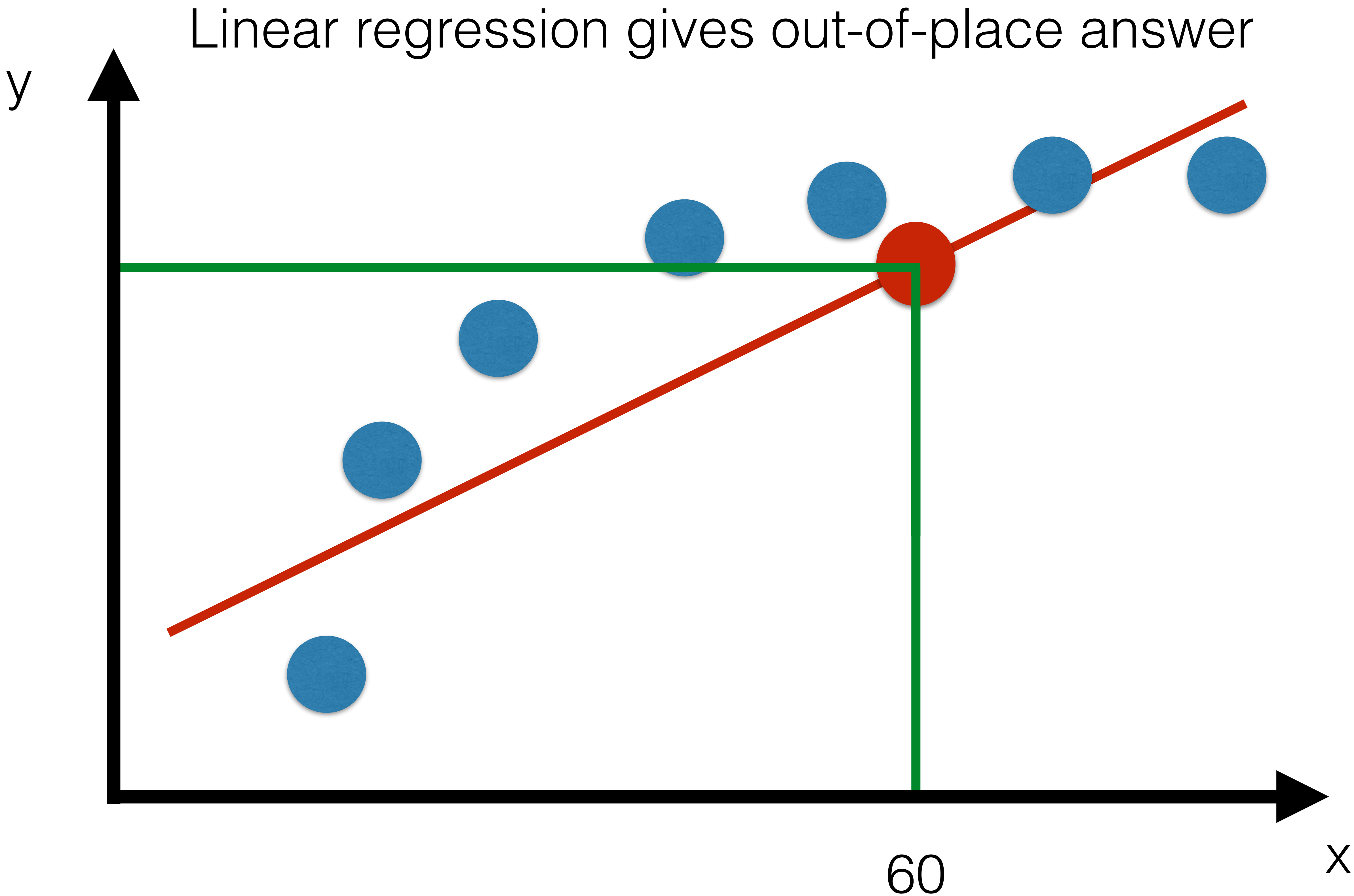
Non-linear Regression:



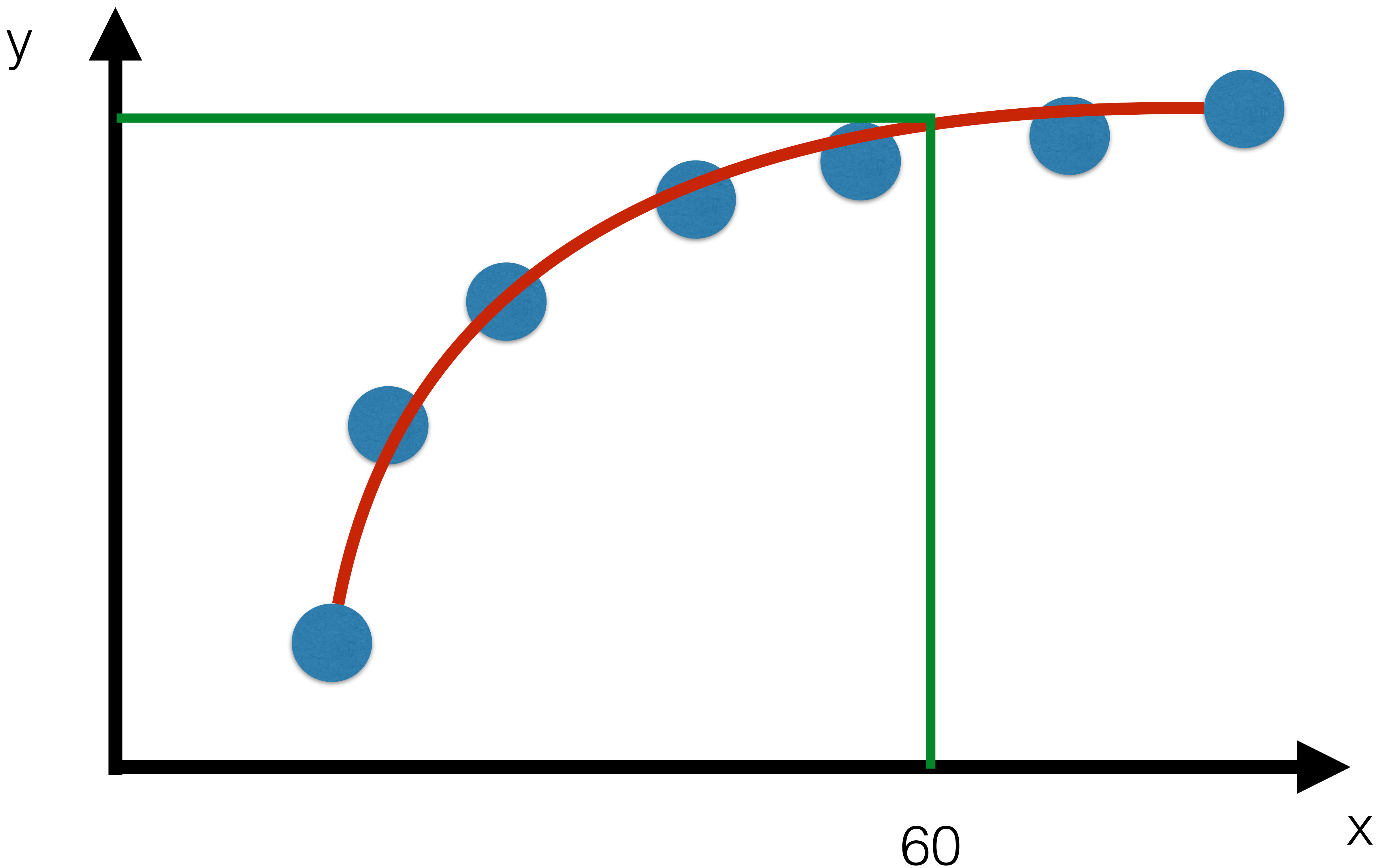
Non-linear Regression:



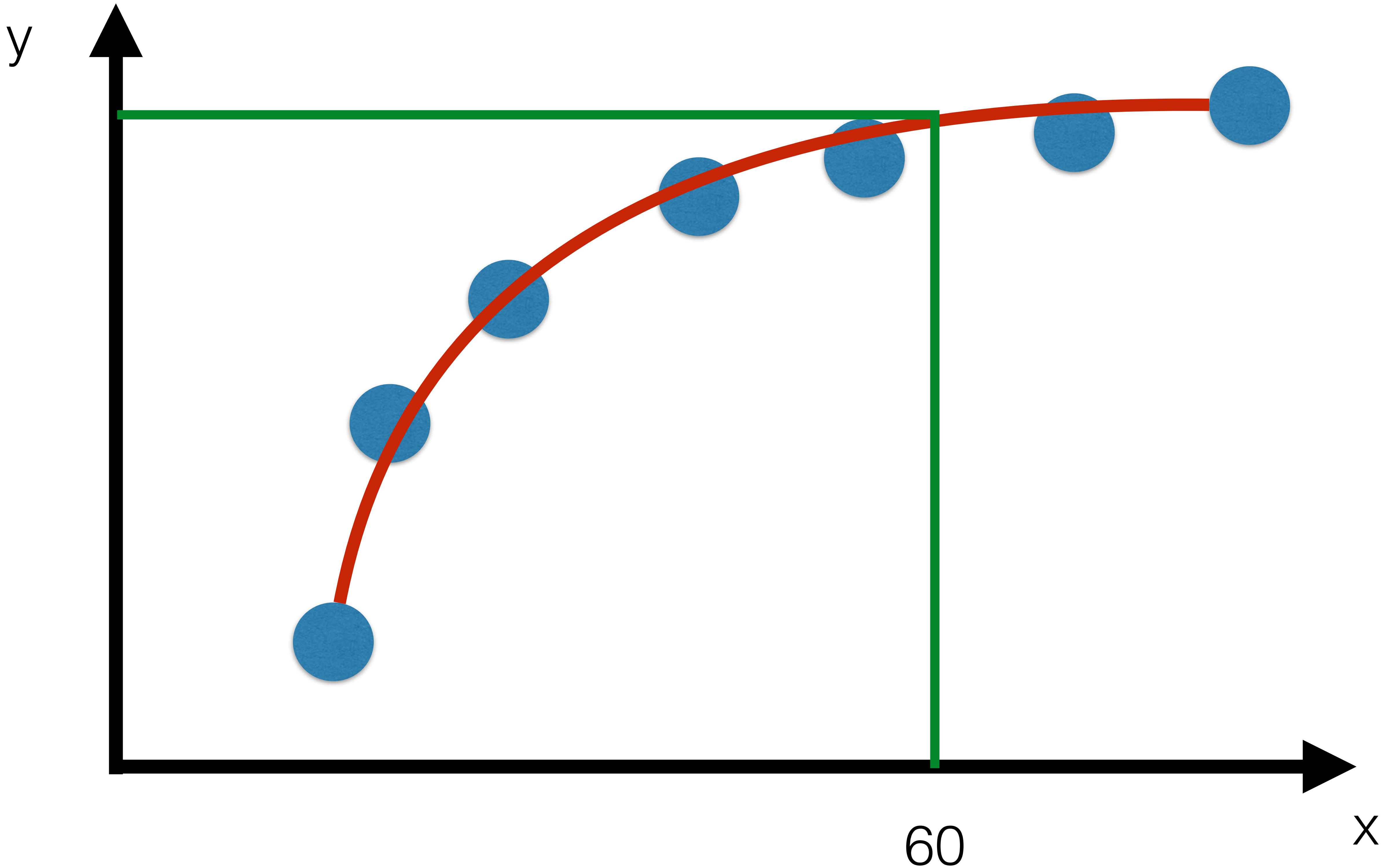
Non-linear Regression:



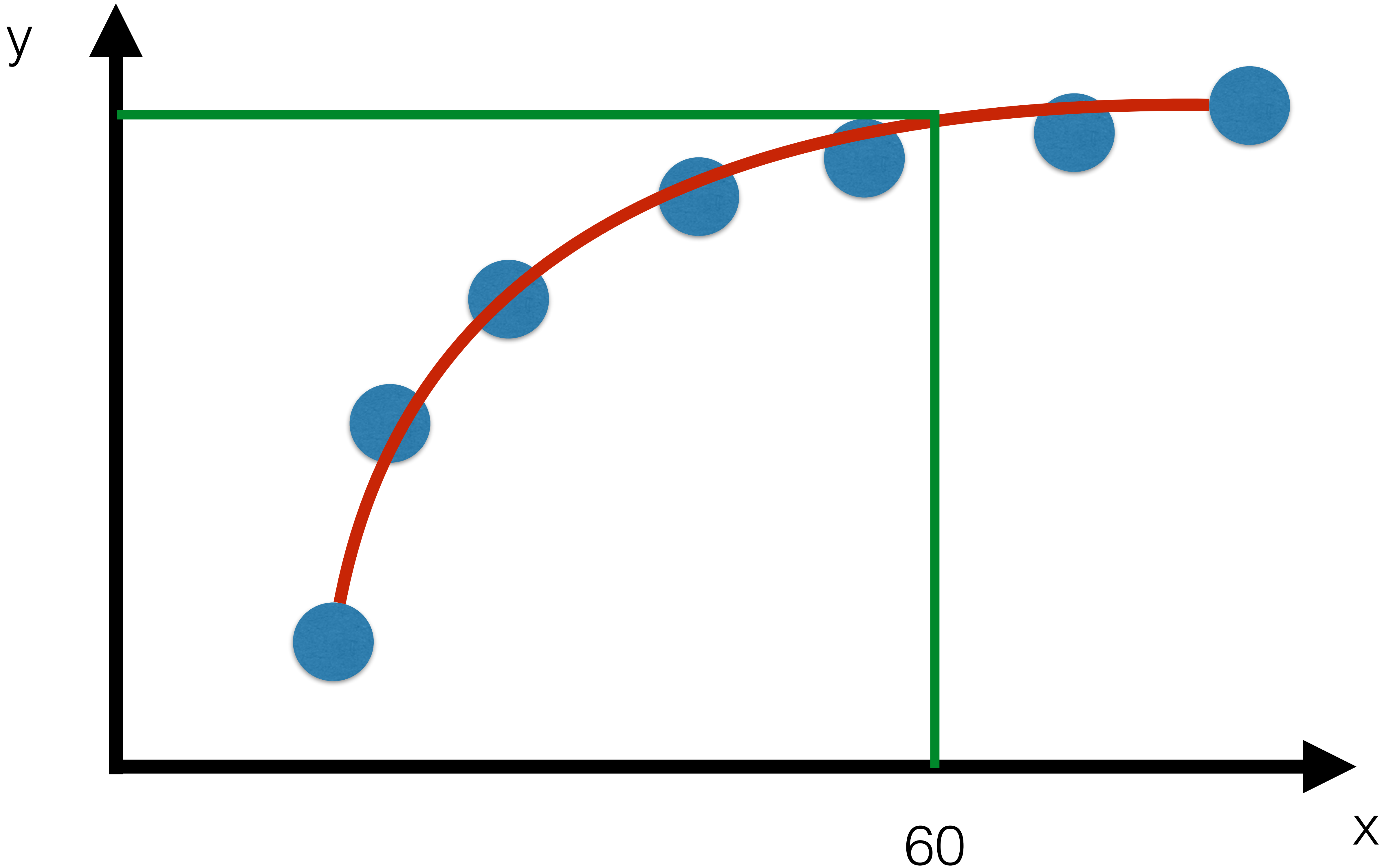
Non-linear Regression:



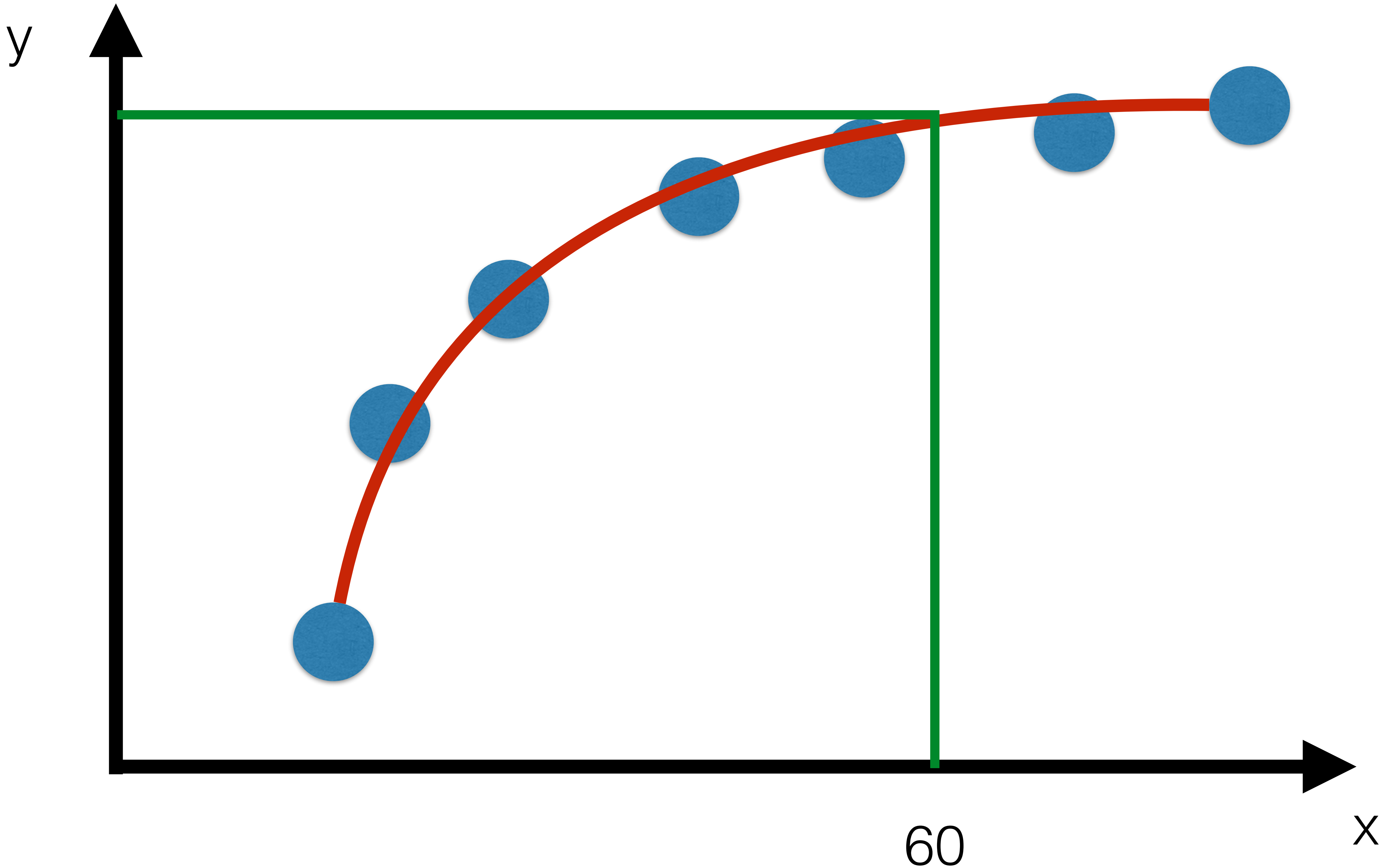
The procedure of finding a relationship between x and y is called **model-fitting**



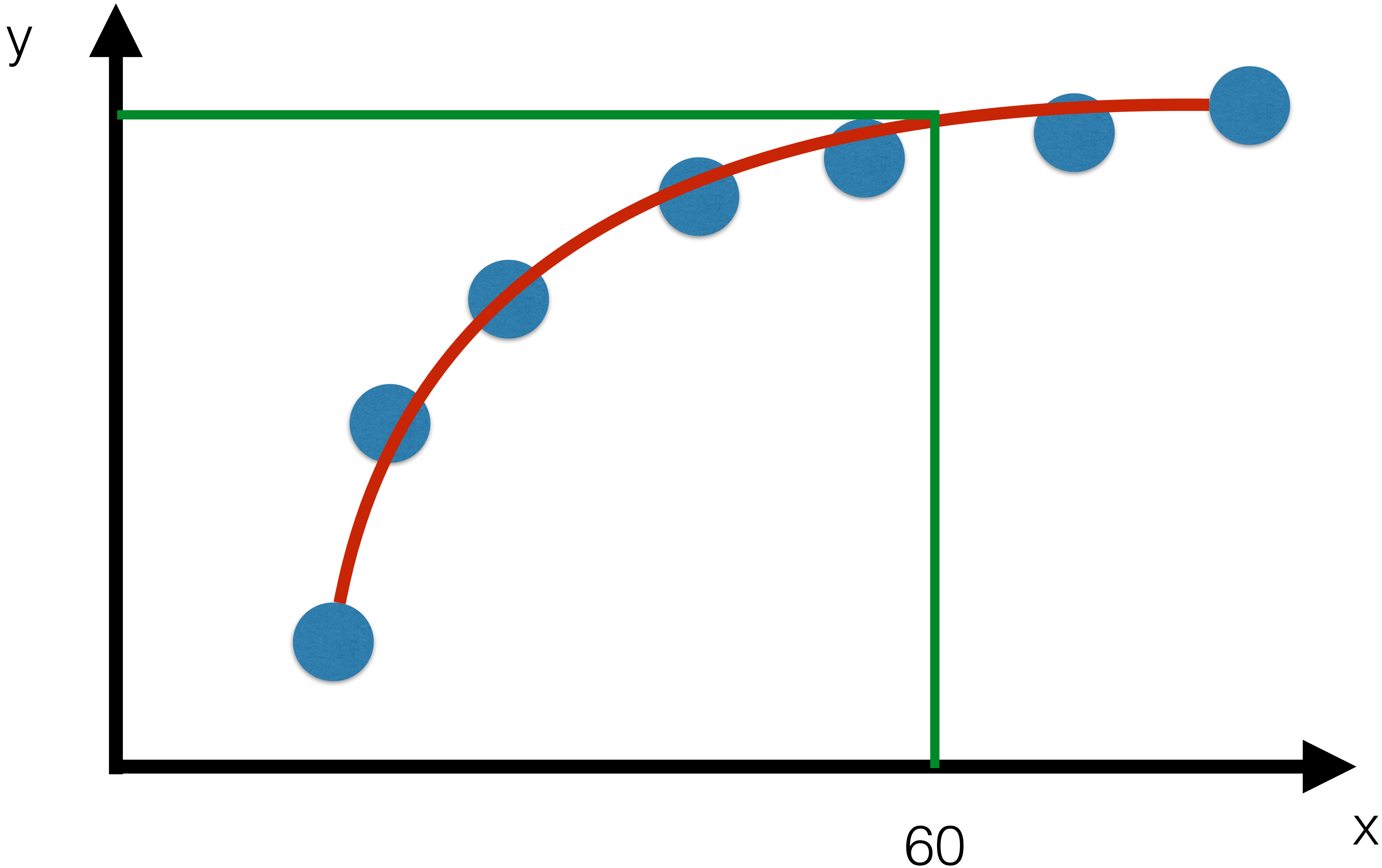
The data used for model-fitting (training) is called **training data**



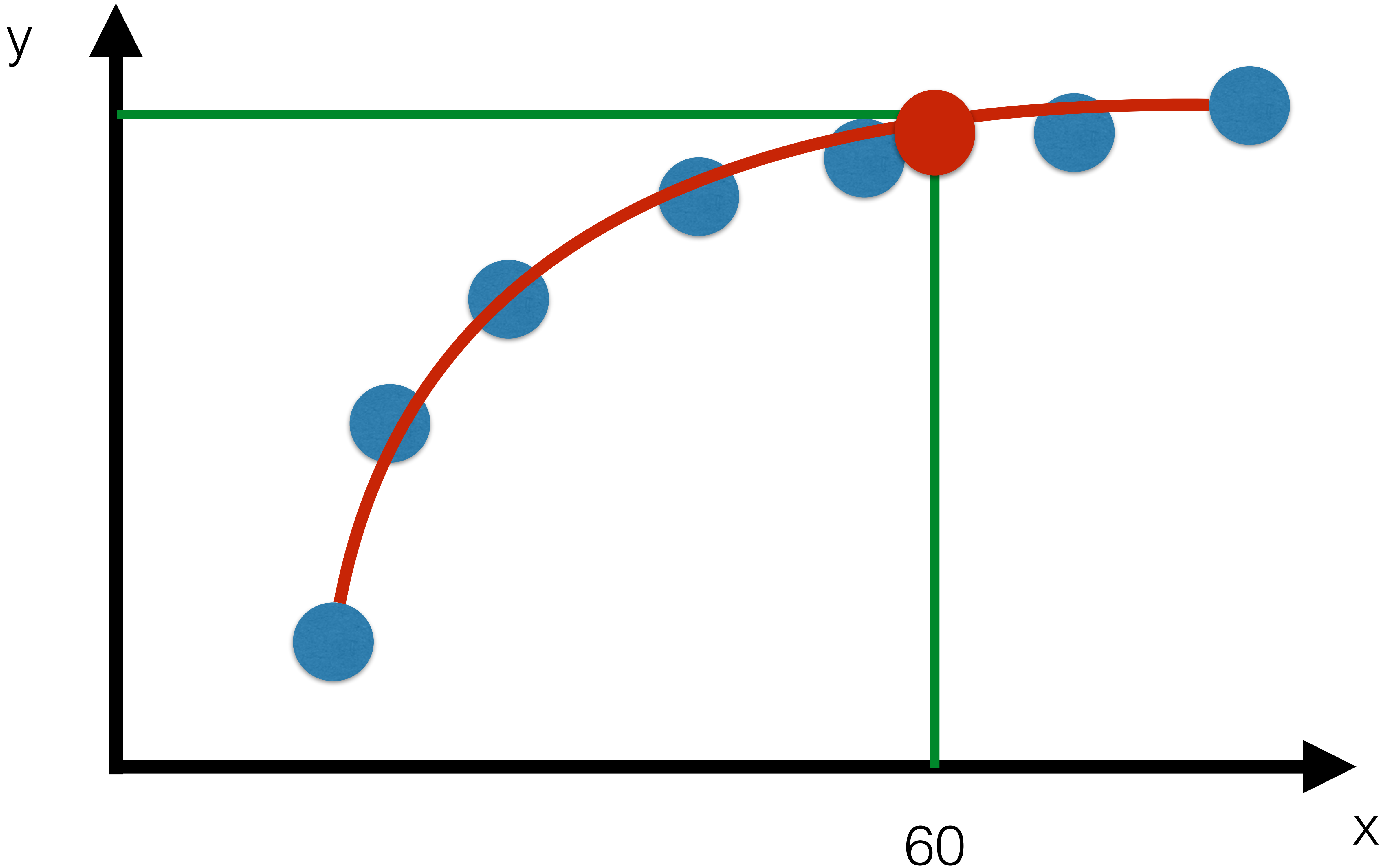
Training data makes our model more accurate, so the computer learns



Then… new data (x, no y) come in, and we use our trained model to predict their values of y



This is called testing. The new data is called **testing data**



Examples of Regression

1. Zestimate
2. Recommendation systems that suggest what movies or television shows to watch next based on user preferences

Two categories of ML algorithms

1. **Supervised learning**: algorithms are trained based on example inputs that are labeled with their desired outputs by humans.

Coins vs cats

Two categories of ML algorithms

1. **Supervised learning**: algorithms are trained based on example inputs that are labeled with their desired outputs by humans.

2. **Unsupervised learning**: the input data is not labeled and algorithms are expected to find structure within the input data by itself.

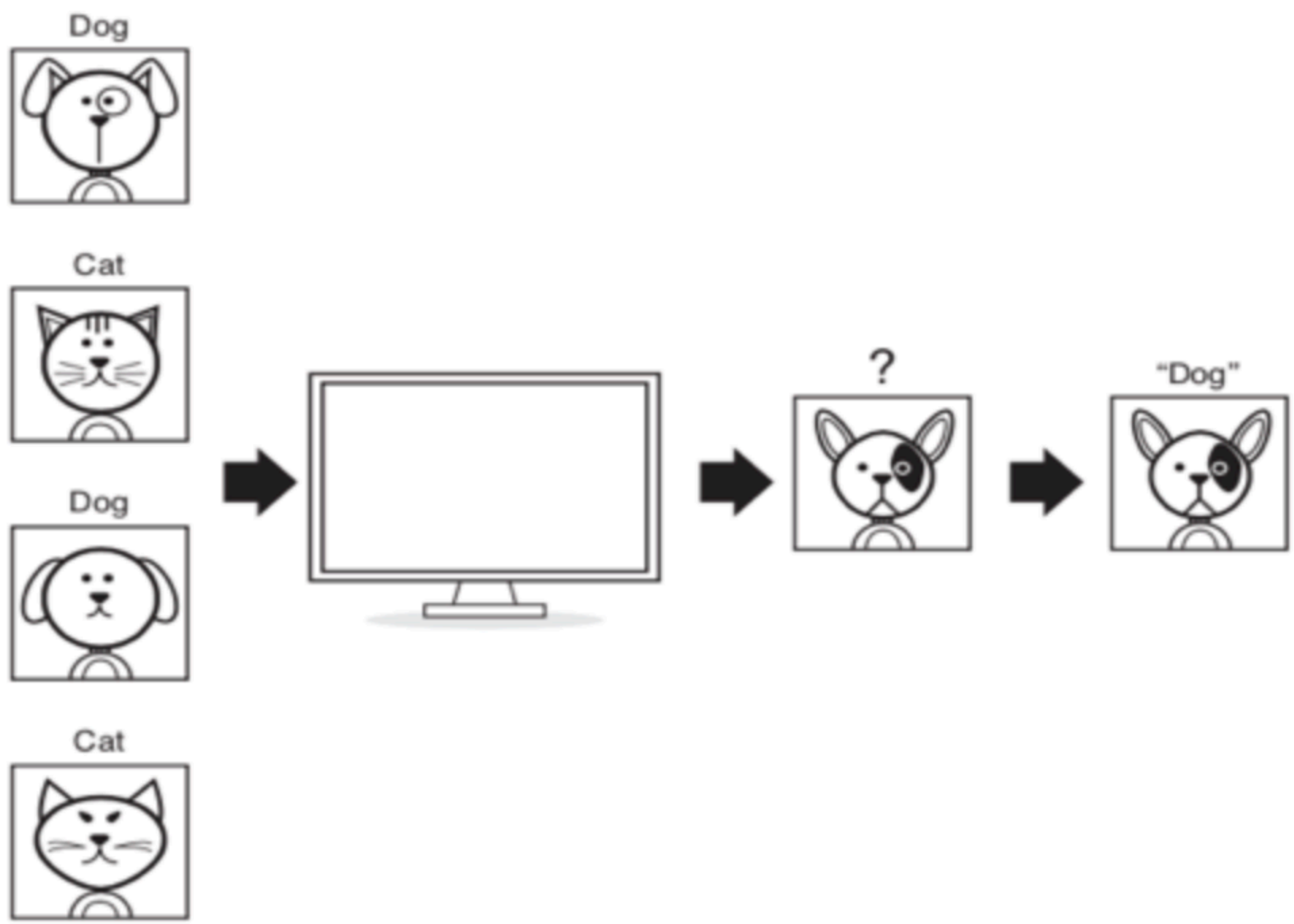
Supervised learning

Example 1: Recovery of the missing arithmetic symbols.

Math Quiz #1 - Teacher's Answer Key

1)	2	4	5	=	3	5)	6	2	2	=	10
2)	5	2	8	=	2	6)	3	1	1	=	2
3)	2	2	1	=	3	7)	5	3	4	=	11
4)	4	2	2	=	6	8)	1	8	1	=	7

Example 2: Image classification

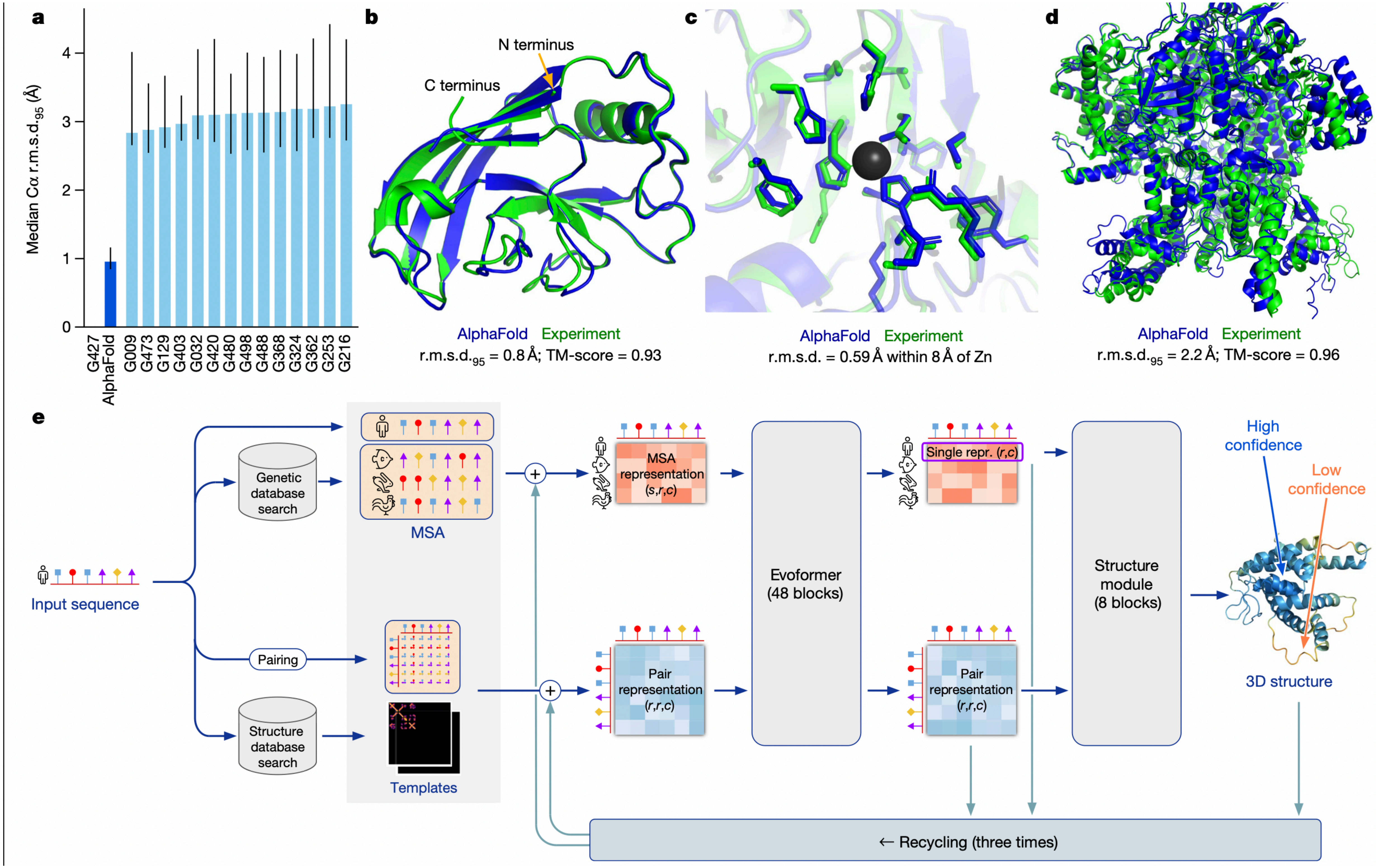


Example 3: Real estate pricing

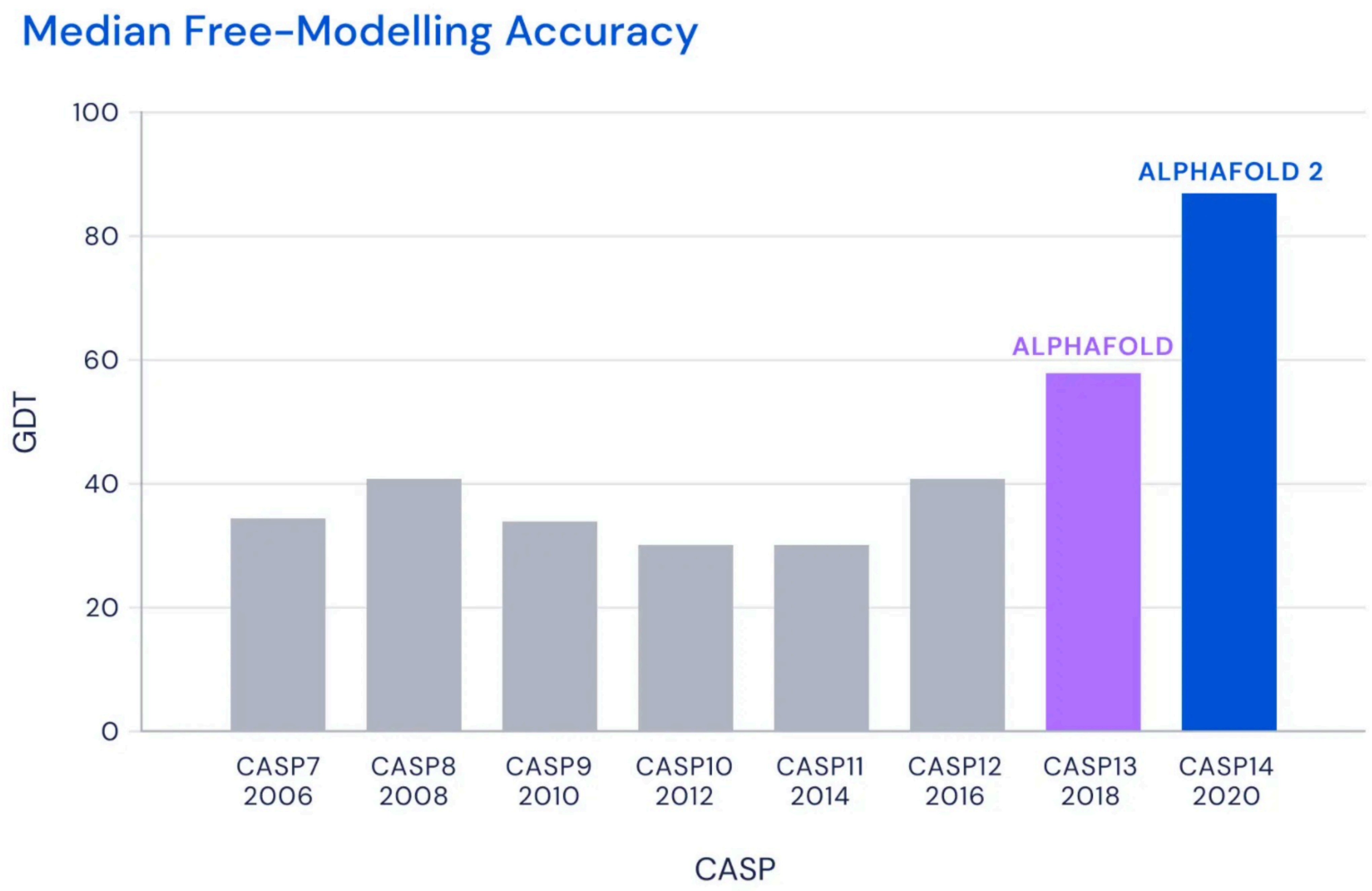
Bedrooms	Sq. feet	Neighborhood	Sale price
3	2000	Normaltown	$250,000
2	800	Hipsterton	$300,000
2	850	Normaltown	$150,000
1	550	Normaltown	$78,000
4	2000	Skid Row	$150,000

3	2000	Hipsterton	???

Example 4: AlphaFold 2



Example 4: AlphaFold 2



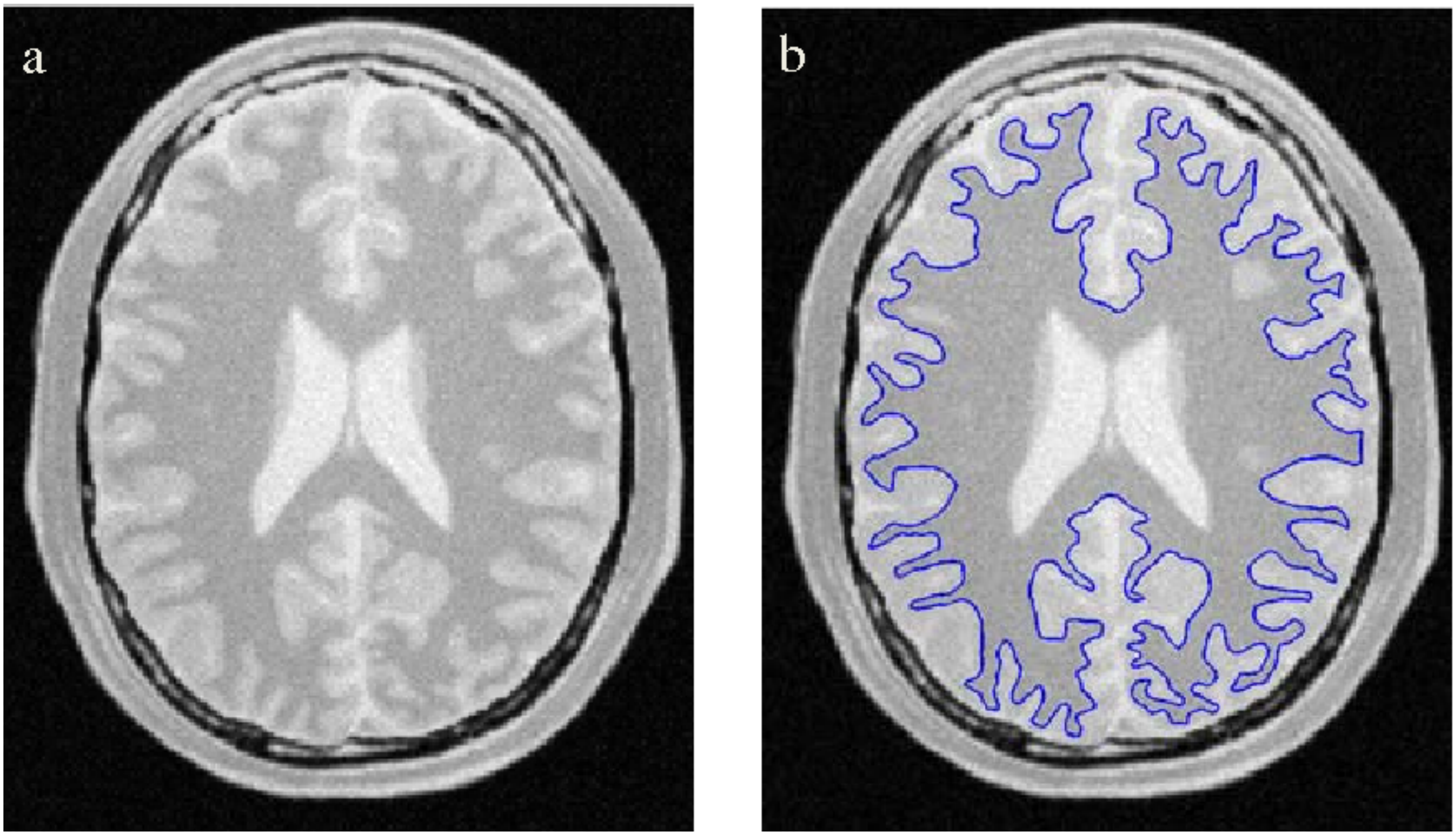
https://www.deepmind.com/blog/alphafold-a-solution-to-a-50-year-old-grand-challenge-in-biology

Example 5: Autonomous Driving

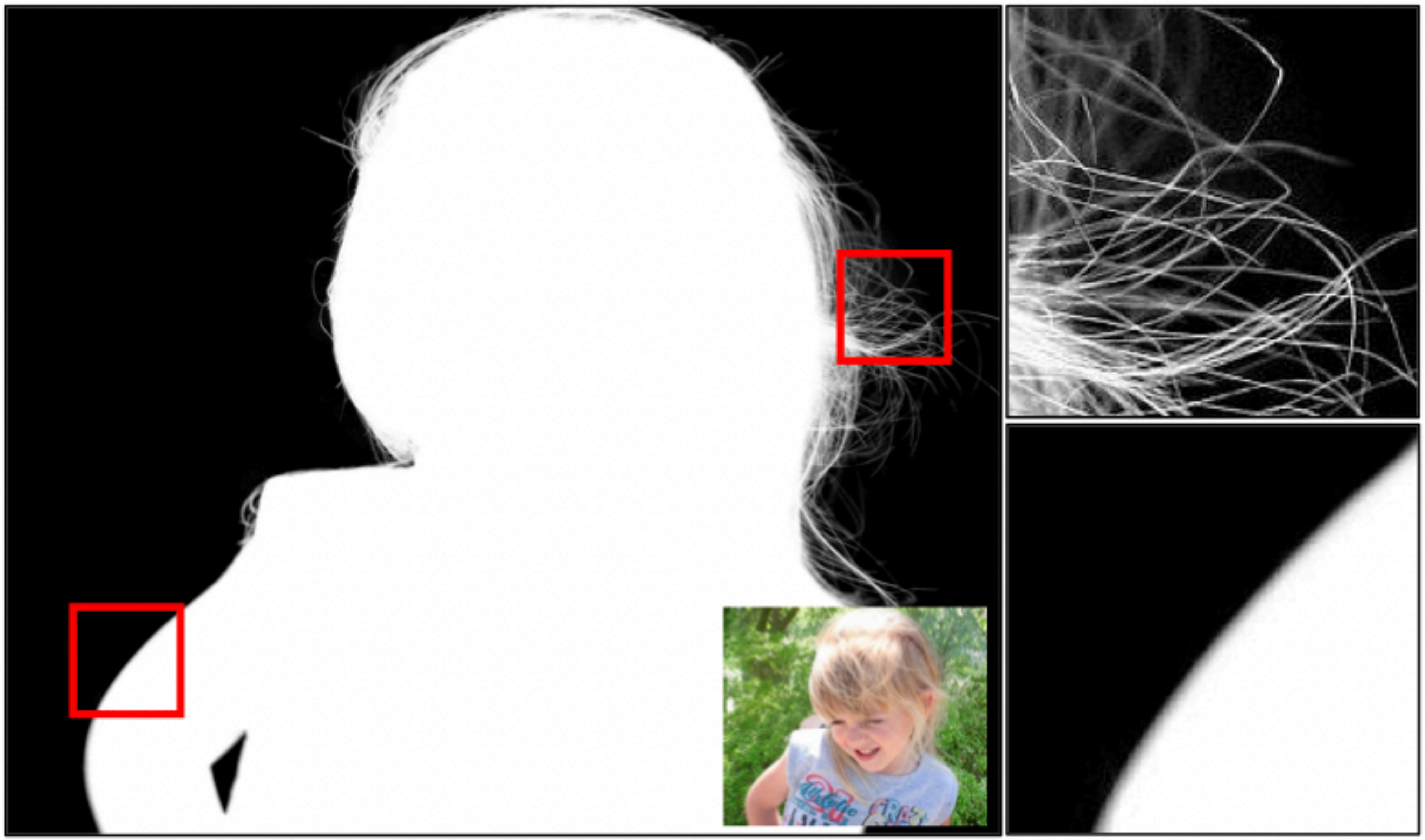
Unsupervised learning

Example 1: Clustering

Example 2: Medical image segmentation



Example 3: Image Matting



Extra: Reinforcement Learning

Example : AlphaZero / AlphaGo