

DECISION TREES AND RANDOM FORESTS

Lesson 15

Prepared by Vlada Rozova

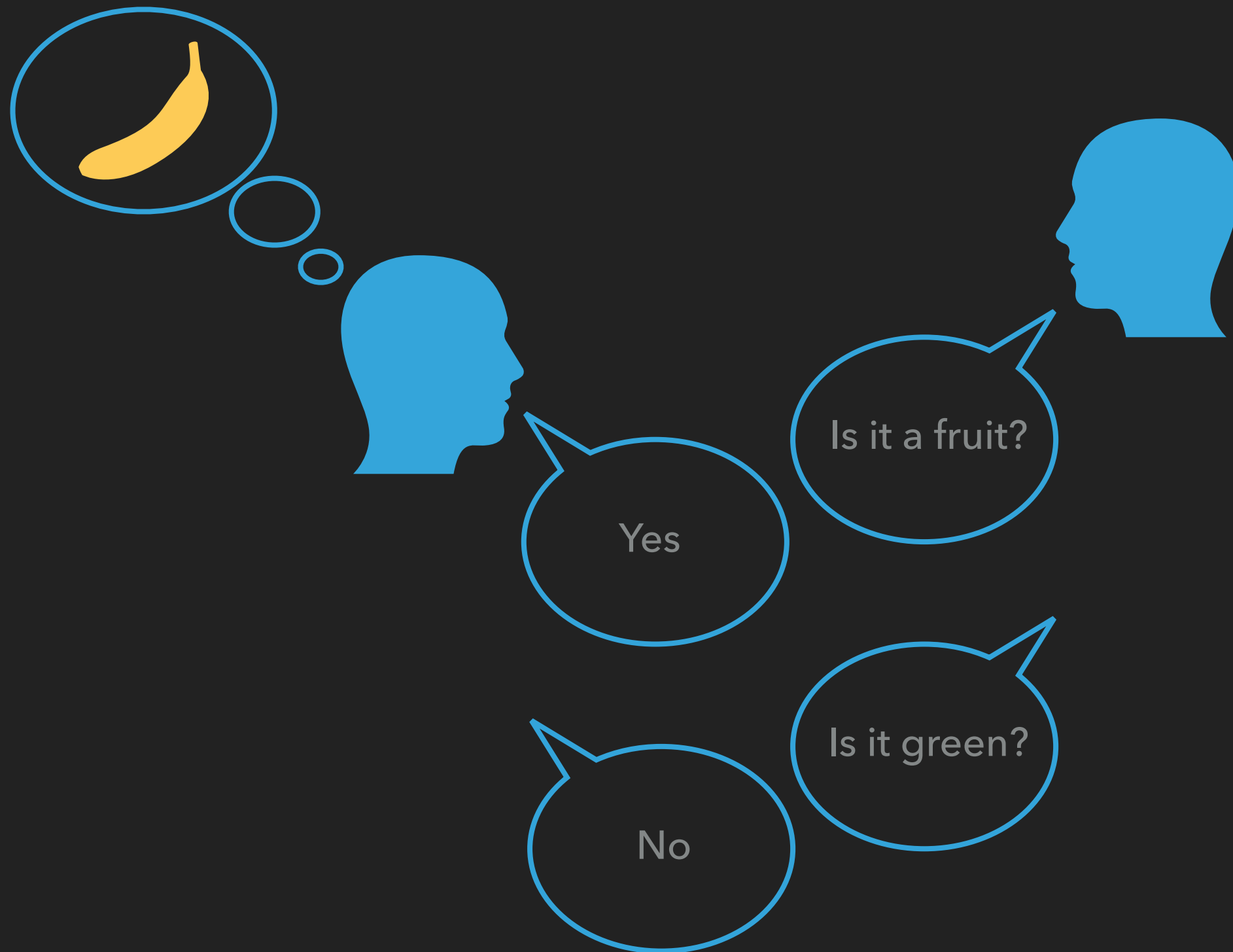
OBJECTIVES

- ▶ Define and build a decision tree
- ▶ Visually represent a decision tree
- ▶ Determine when using decision trees is appropriate

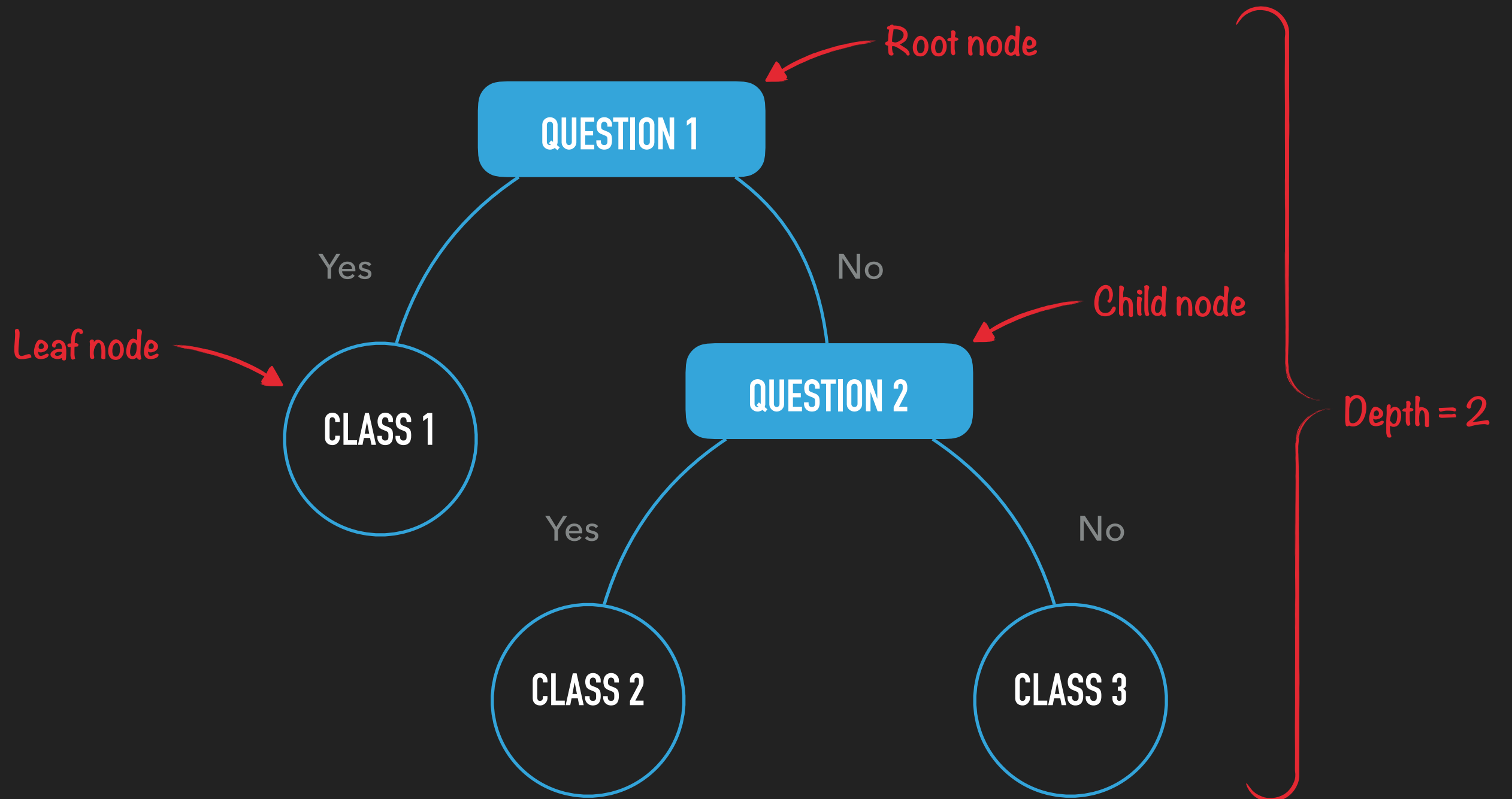
OUTLINE OF THE LESSON

- ▶ Intuition behind decision trees
- ▶ What are decision trees?
- ▶ How to build a decision tree?
- ▶ Build your first decision tree!
- ▶ When and how to use decision trees?

INTUITION BEHIND DECISION TREES



WHAT ARE DECISION TREES?



HOW TO BUILD A DECISION TREE?

- ▶ Toy problem: classify fruits from the fruit dataset
- ▶ Use a Decision Tree Classifier with Depth = 2

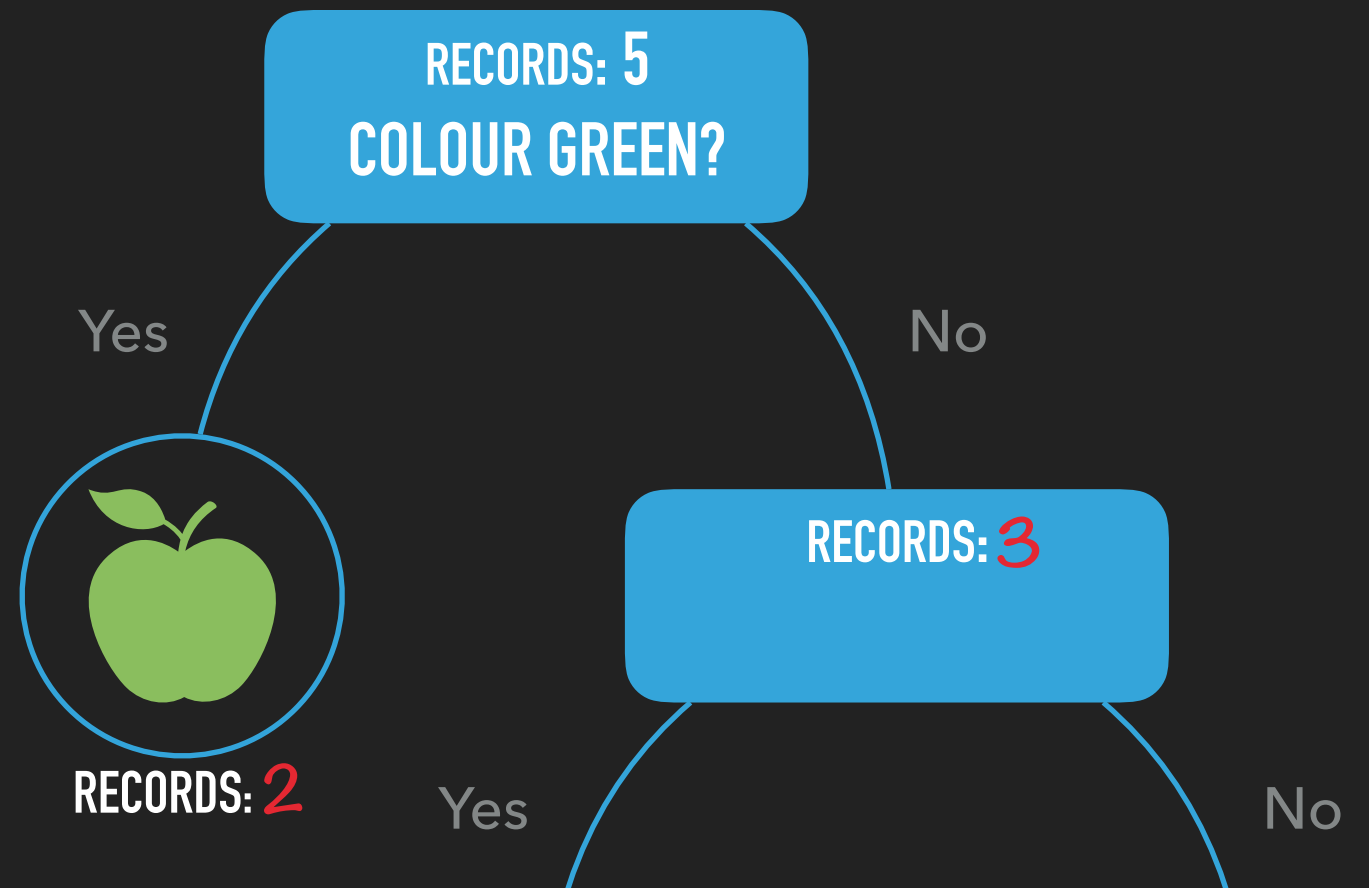
Fruit	Colour	Length
Apple	Green	5.5
Apple	Green	7
Lemon	Yellow	6
Lemon	Yellow	4.5
Banana	Yellow	16



HOW TO BUILD A DECISION TREE?

- ▶ Toy problem: classify fruits from the fruit dataset
- ▶ Use a Decision Tree Classifier with Depth = 2

Fruit	Colour	Length
Apple	Green	5.5
Apple	Green	7
Lemon	Yellow	6
Lemon	Yellow	4.5
Banana	Yellow	16



HOW TO BUILD A DECISION TREE?

- ▶ Toy problem: classify fruits from the fruit dataset
- ▶ Use a Decision Tree Classifier with Depth = 2

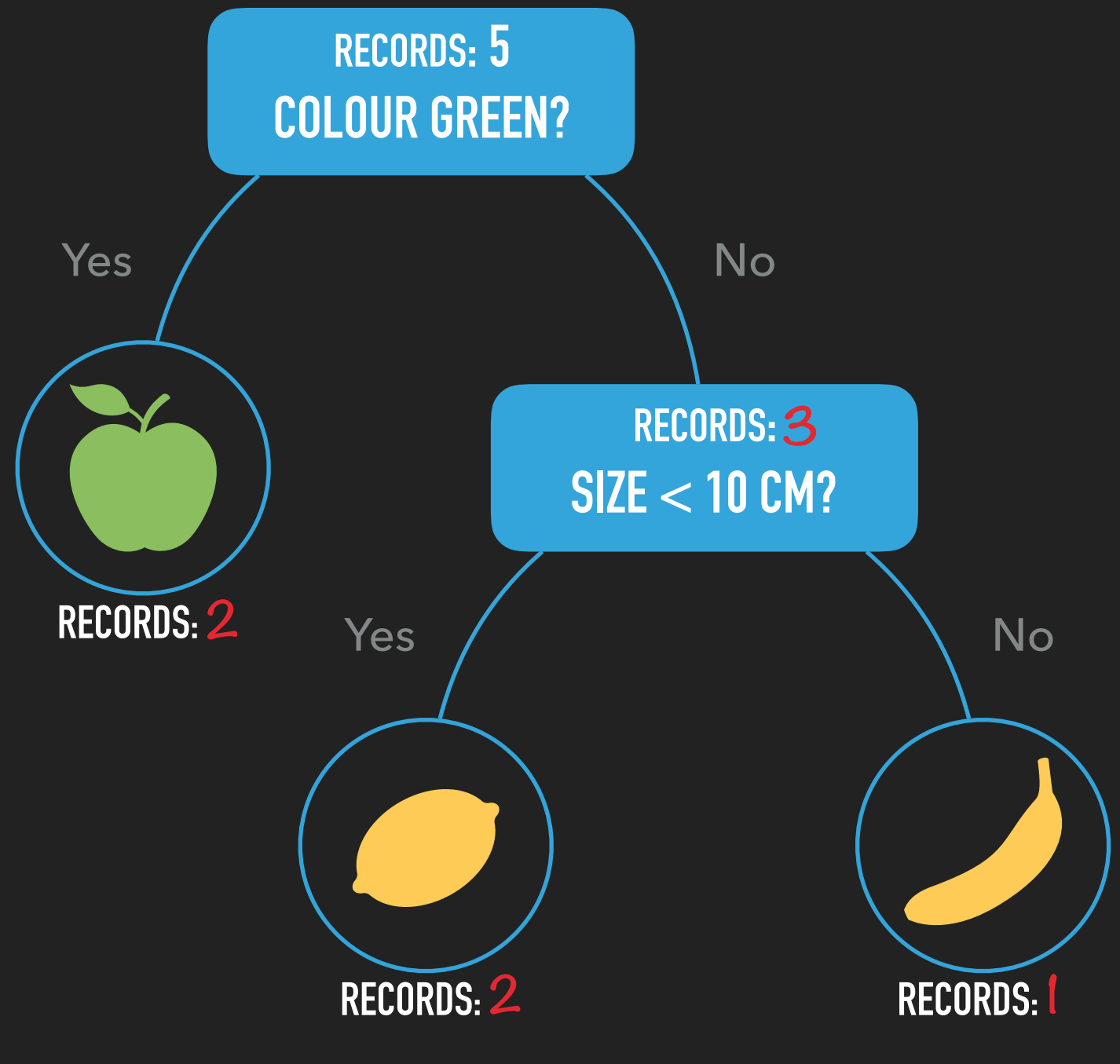
Fruit	Colour	Size
Apple	Green	5.5
Apple	Green	7
Lemon	Yellow	6
Lemon	Yellow	4.5
Banana	Yellow	16



HOW TO BUILD A DECISION TREE?

- ▶ Toy problem: classify fruits from the fruit dataset
- ▶ Use a Decision Tree Classifier with Depth = 2

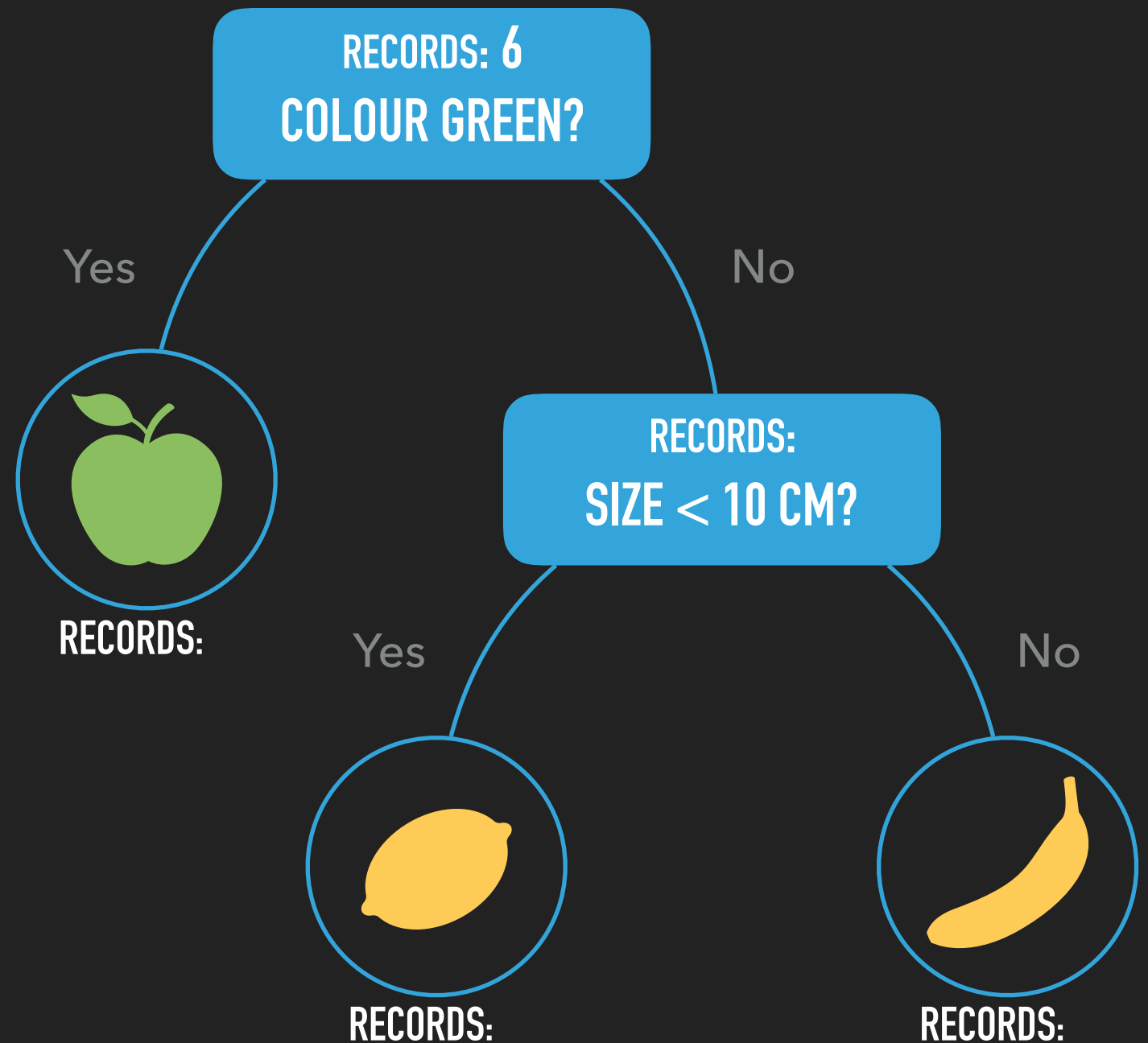
Fruit	Colour	Size
Apple	Green	5.5
Apple	Green	7
Lemon	Yellow	6
Lemon	Yellow	4.5
Banana	Yellow	16



HOW TO BUILD A DECISION TREE?

- ▶ Toy problem: classify fruits from the fruit dataset
- ▶ Use a Decision Tree Classifier with Depth = 2

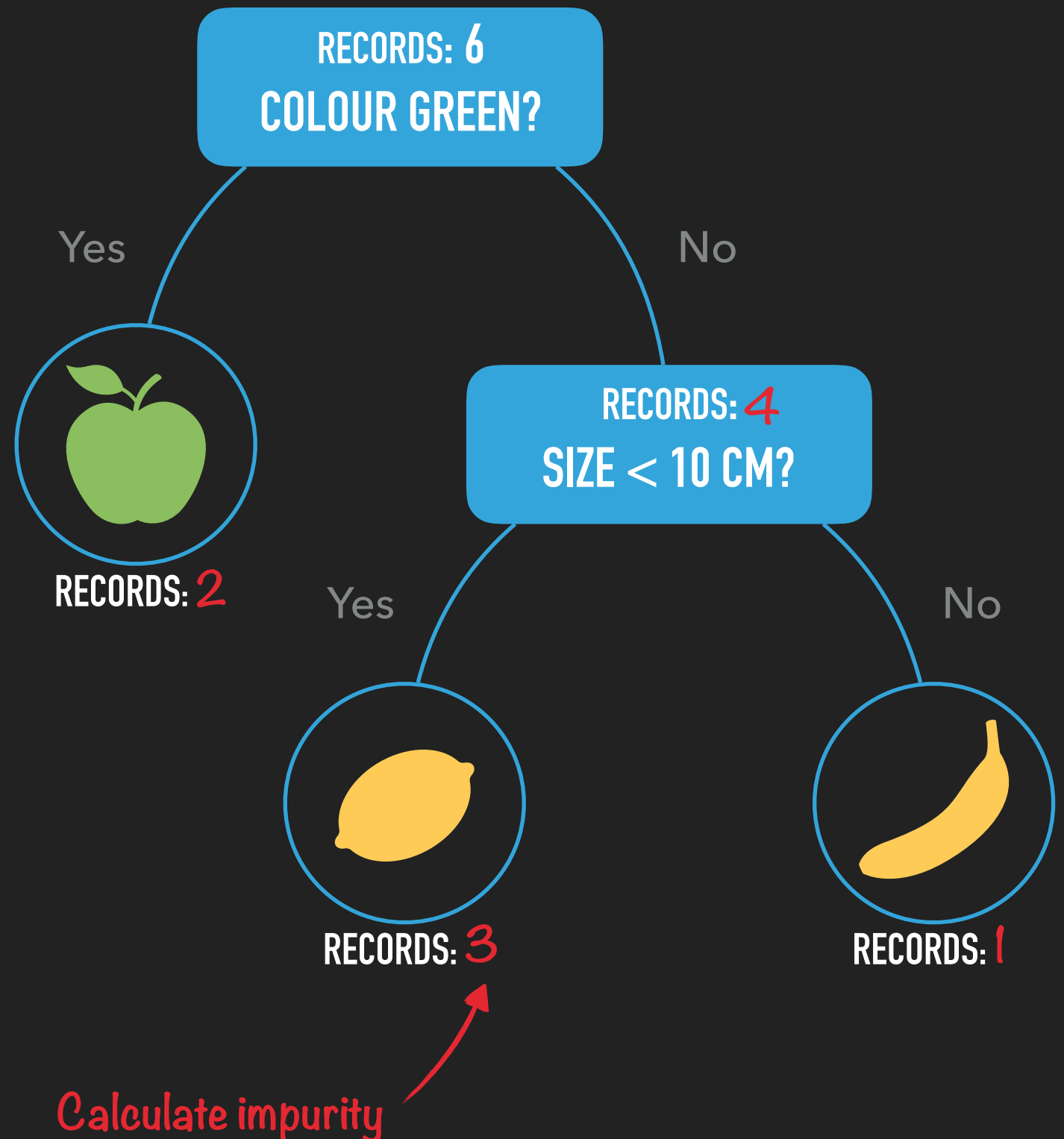
Fruit	Colour	Size
Apple	Green	5.5
Apple	Green	7
Lemon	Yellow	6
Lemon	Yellow	4.5
Banana	Yellow	16
Banana	Yellow	8



HOW TO BUILD A DECISION TREE?

- ▶ Toy problem: classify fruits from the fruit dataset
- ▶ Use a Decision Tree Classifier with Depth = 2

Fruit	Colour	Size
Apple	Green	5.5
Apple	Green	7
Lemon	Yellow	6
Lemon	Yellow	4.5
Banana	Yellow	16
Banana	Yellow	8

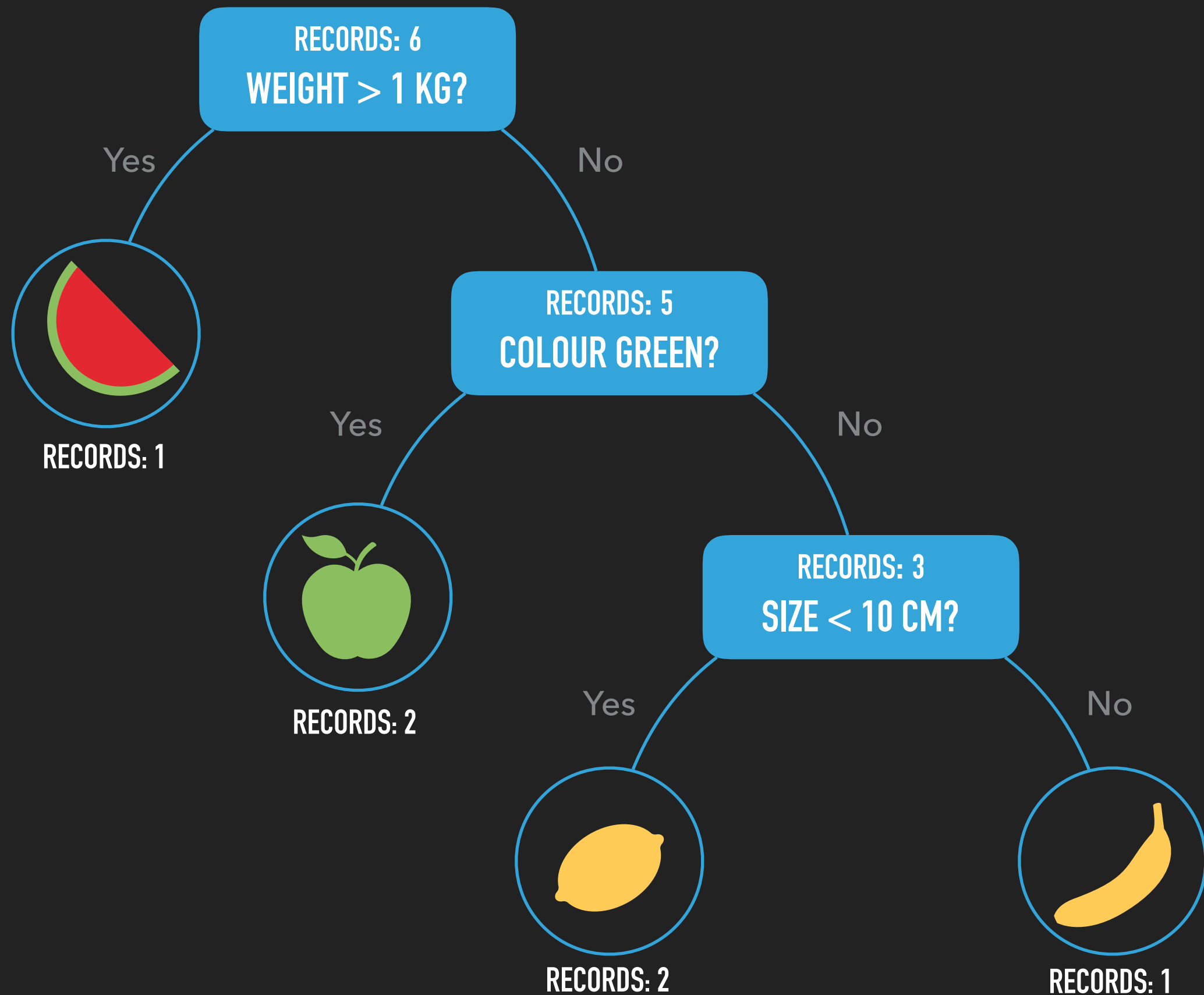


BUILD YOUR FIRST DECISION TREE!

- ▶ Task: classify fruits from the fruit dataset into 4 classes: apples, bananas, lemons, watermelons
- ▶ Use a Decision Tree Classifier with Depth = 3

Fruit	Colour	Length	Weight
Apple	Green	5.5	0.17
Apple	Green	7	0.2
Lemon	Yellow	6	0.08
Lemon	Yellow	4.5	0.11
Banana	Yellow	16	0.18
Watermelon	Green	24	8.5

BUILD YOUR FIRST DECISION TREE!



WHEN AND HOW TO USE DECISION TREES

- ▶ Decision trees can be used both for classification and regression
- ▶ Visually and explicitly represent decisions and provide classification rules that can be applied manually
- ▶ Require very little data preparation (no centring, scaling, etc.)
- ▶ Sensitive to small variations in the data: a slight change can result in a drastically different tree