Decision Tree

Regression

Regression tree vs classification tree

Low pressure	High Temperature	High humidity	Wind Speed	Rain
No	No	No	10.0	No
Yes	Yes	Yes	30.0	Yes
Yes	Yes	No	20.0	No
Yes	No	Yes	50.0	No
No	No	Yes	70.0	Yes

classification tree

Discrete data

Low pressure	High Temperature	High humidity	Wind Speed	Rain amount
No	No	No	10.0	0.1
Yes	Yes	Yes	30.0	0.3
Yes	Yes	No	20.0	1.0
Yes	No	Yes	50.0	2.1
No	No	Yes	70.0	0.7

Regression tree

Continuous data

Regression tree vs classification tree

Regression tree is a type of decision tree

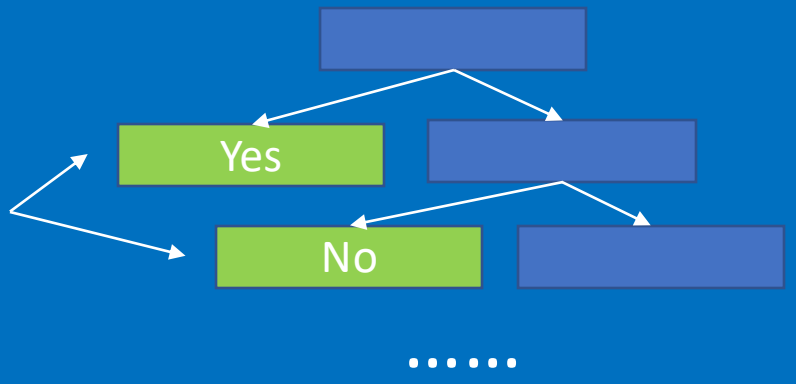
In regression tree, each leaf represents a numeric value

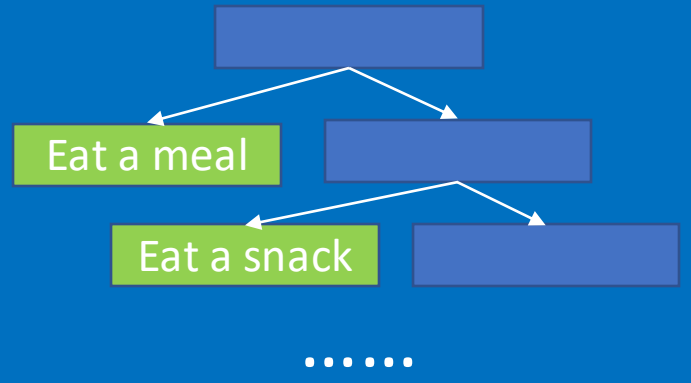
Regression tree vs classification tree

Regression tree is a type of decision tree

In regression tree, each leaf represents a numeric value

In Classification tree, each leaf represents a Yes/No, or True/False, or some other discrete category





Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let’s assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

Dosage	Drug effect.
10	58
20	60
35	57
5	44
…	…

Regression tree workflow

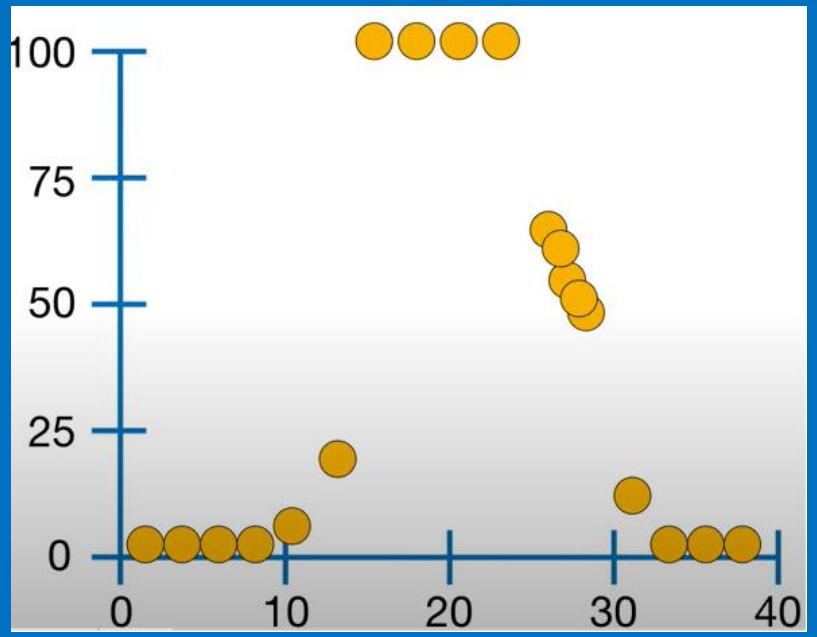
Similar to classification tree, regression tree is built from top down

Let’s assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

Dosage	Drug effect.
10	58
20	60
35	57
5	44
…	…

We can plot it out as

Drug effect.

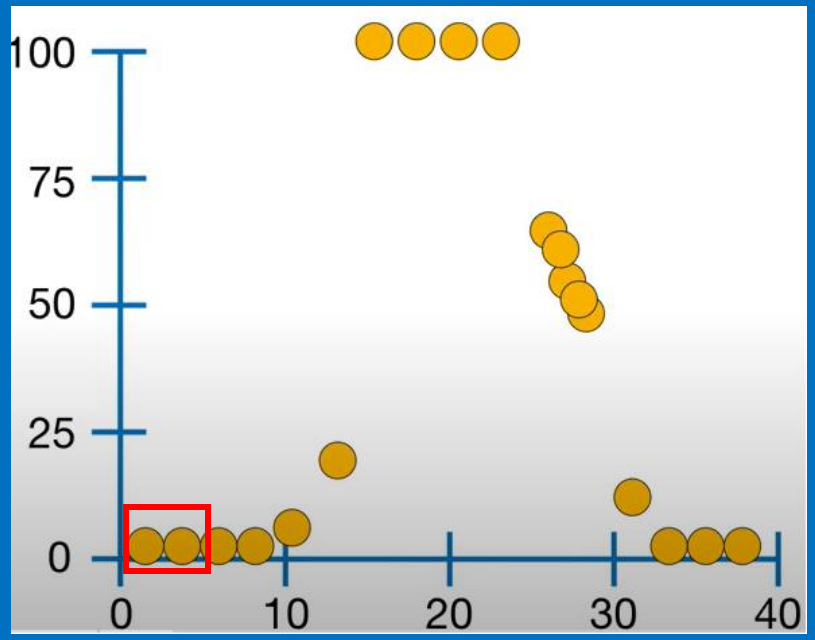


Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let’s assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

Drug effect.



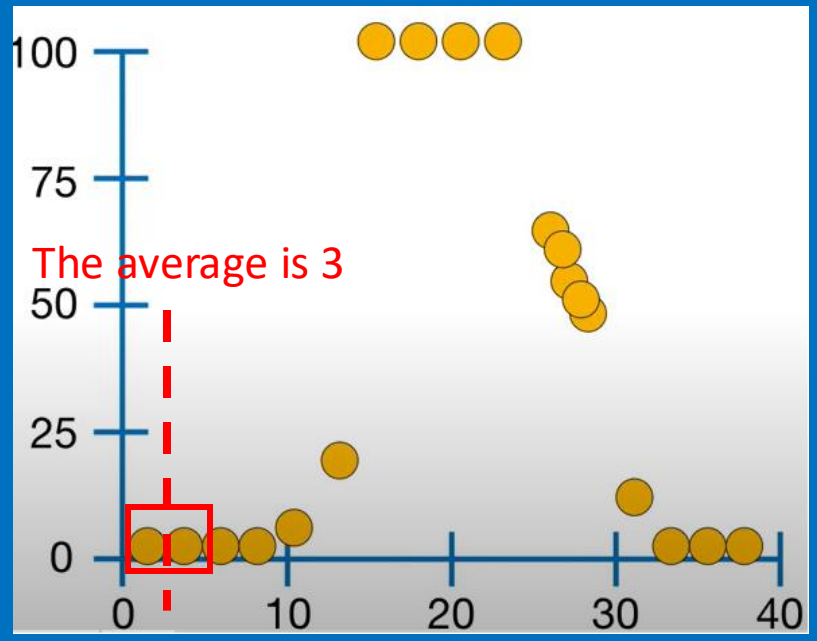
Dosage

Step 1: let’s look at the first two data points (smallest two values)

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let’s assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



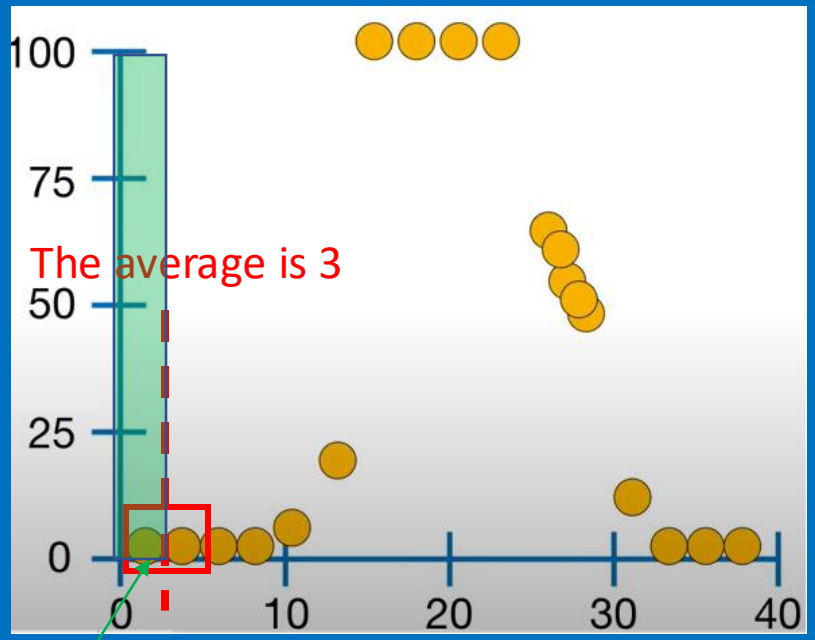
Step 1: let’s look at the first two data points (smallest two values), The average is 3.

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

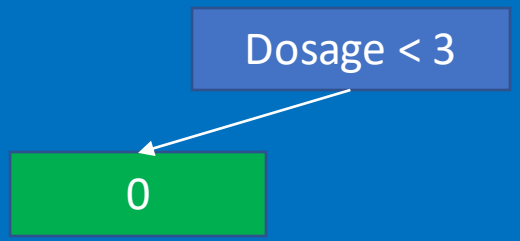
Drug effect.



The average drug effect. is 0

Step 1: let's look at the first two data points (smallest two values), The average is 3.

Now we can build a simple tree based on if the dosage is smaller than 3

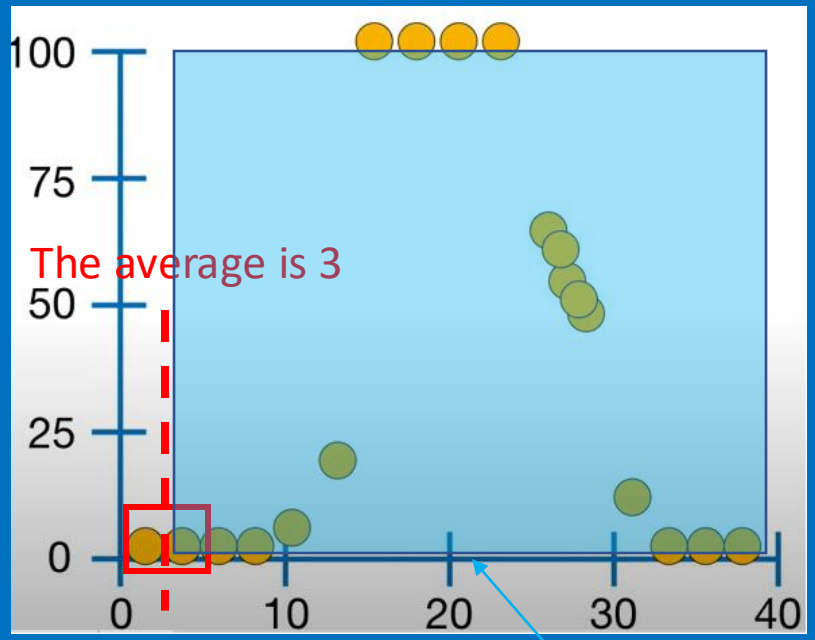


Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

Drug effect.

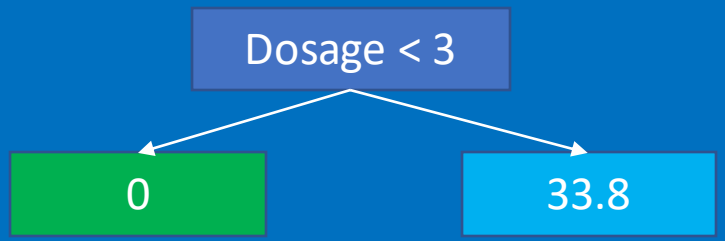


Dosage

The average drug effect. Is 38.8

Step 1: let's look at the first two data points (smallest two values), The average is 3.

Now we can build a simple tree based on if the dosage is smaller than 3

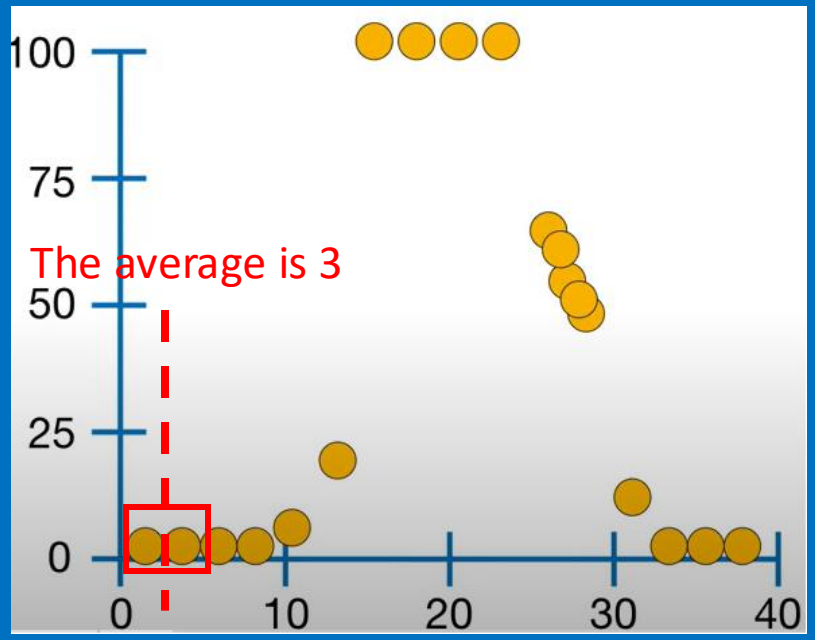


Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

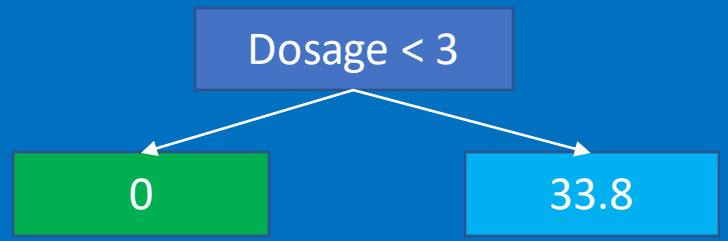
Drug effect.



Dosage

Step 1: let's look at the first two data points (smallest two values), The average is 3.

Now we can build a simple tree based on if the dosage is smaller than 3



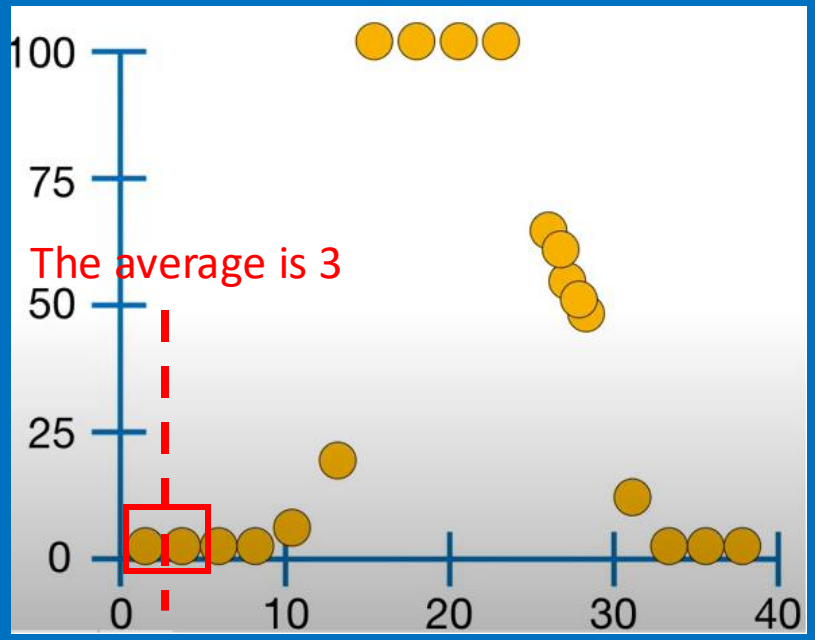
Step 2: let's estimate how accurate this tree is

Regression tree workflow

Similar to classification tree, regression tree is built from top down

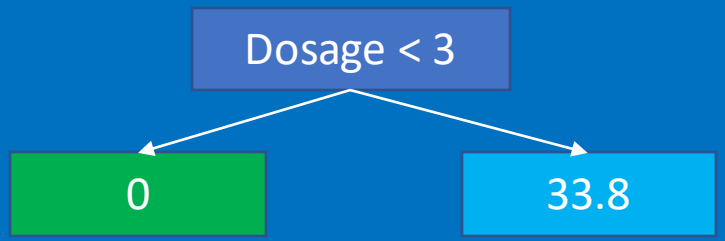
Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

Drug effect.



Step 1: let's look at the first two data points (smallest two values), The average is 3.

Now we can build a simple tree based on if the dosage is smaller than 3



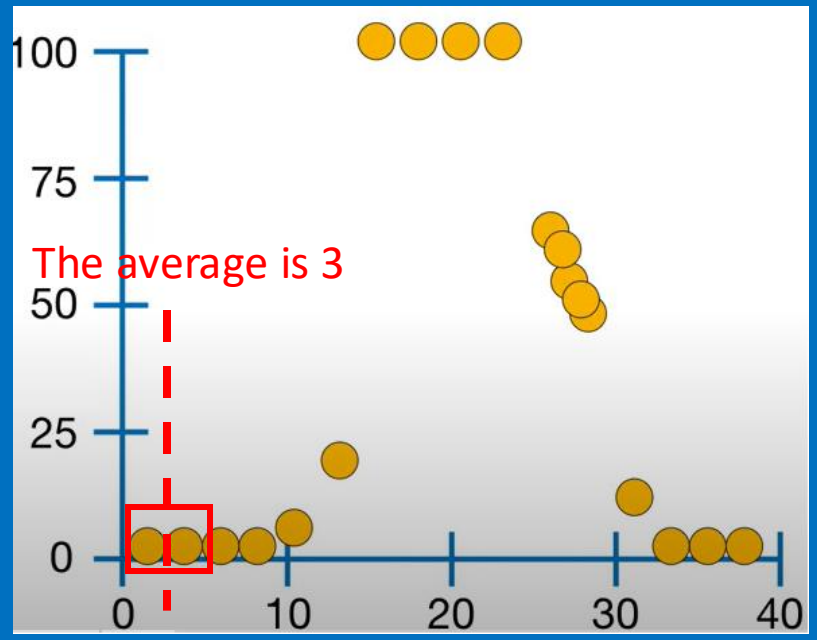
Step 2: let's estimate how accurate this tree is

We go through all the points in the graph in the graph with the tree above, the error for this tree is 27468.5

Regression tree workflow

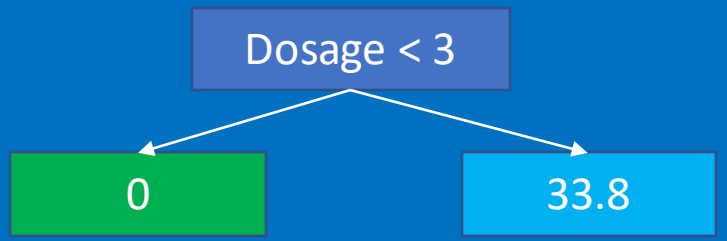
Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



Step 1: let's look at the first two data points (smallest two values), The average is 3.

Now we can build a simple tree based on if the dosage is smaller than 3



Step 2: let's est

We go thr
graph with t e tree above, the error for this tree is 27468.5

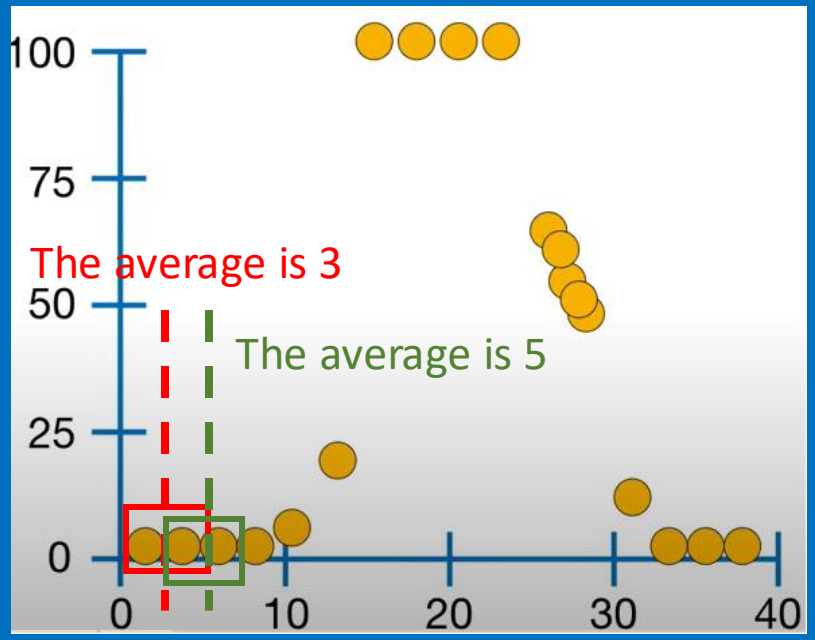
This value is calculated by adding all the error together, e.g., $(Pred - Actual)^2_{point1} + (Pred - Actual)^2_{point2} + \cdots$

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let’s assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

Drug effect.



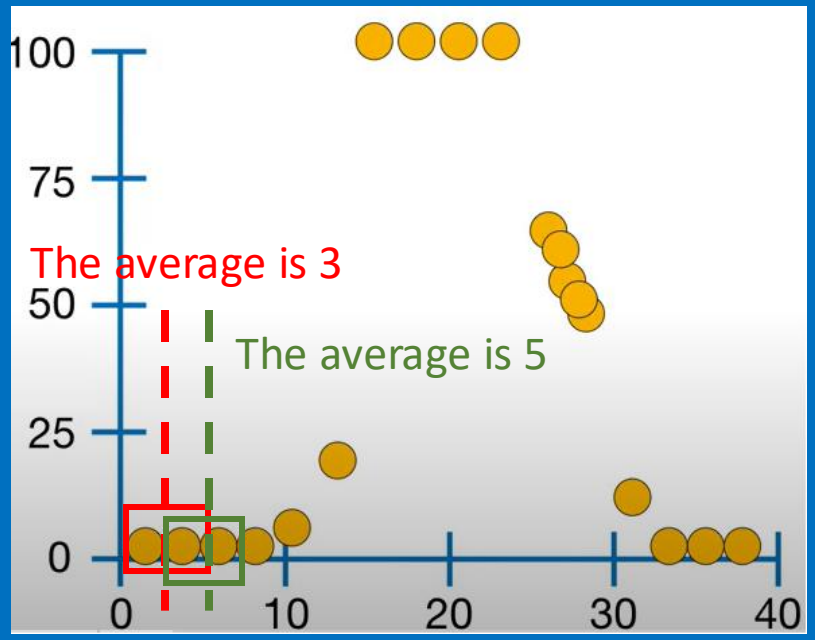
Step 3: then we can move to the next 2 points, and use the average there as the new root threshold, which is 5

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

Drug effect.



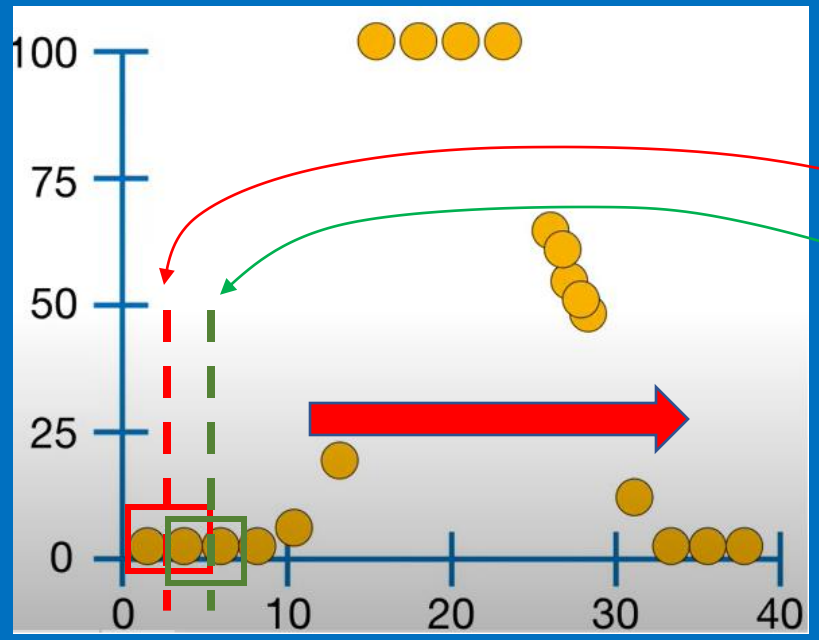
Step 3: then we can move to the next 2 points, and use the average there as the new root threshold, which is 5

Then we can get a new error, which is 23763.2

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



By repeating the above steps, we can get a series of root node threshold, and the corresponding errors

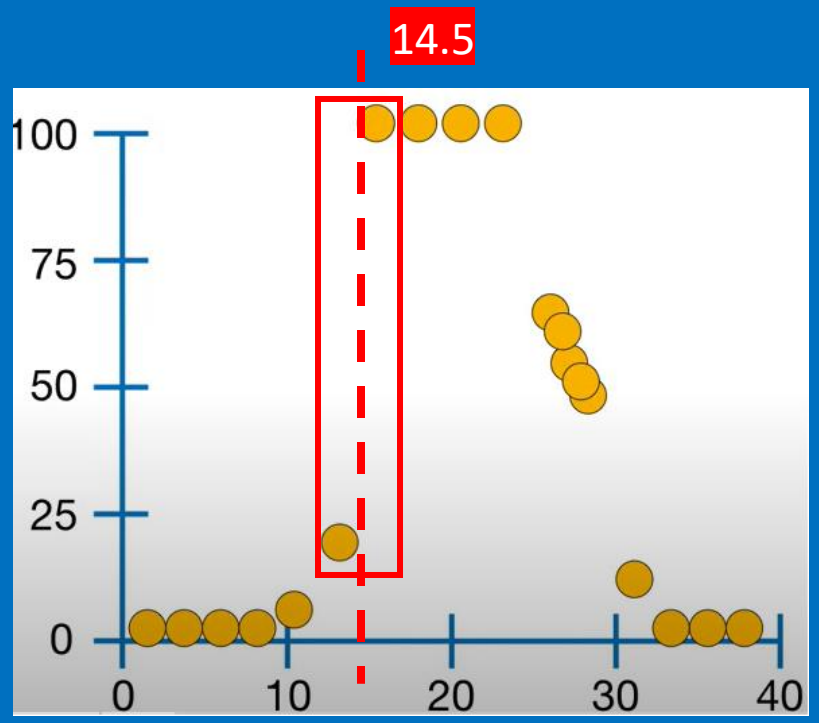
For example

Threshold 3	27468.5
Threshold 5	23763.2
Threshold 8	……
……	

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let’s assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



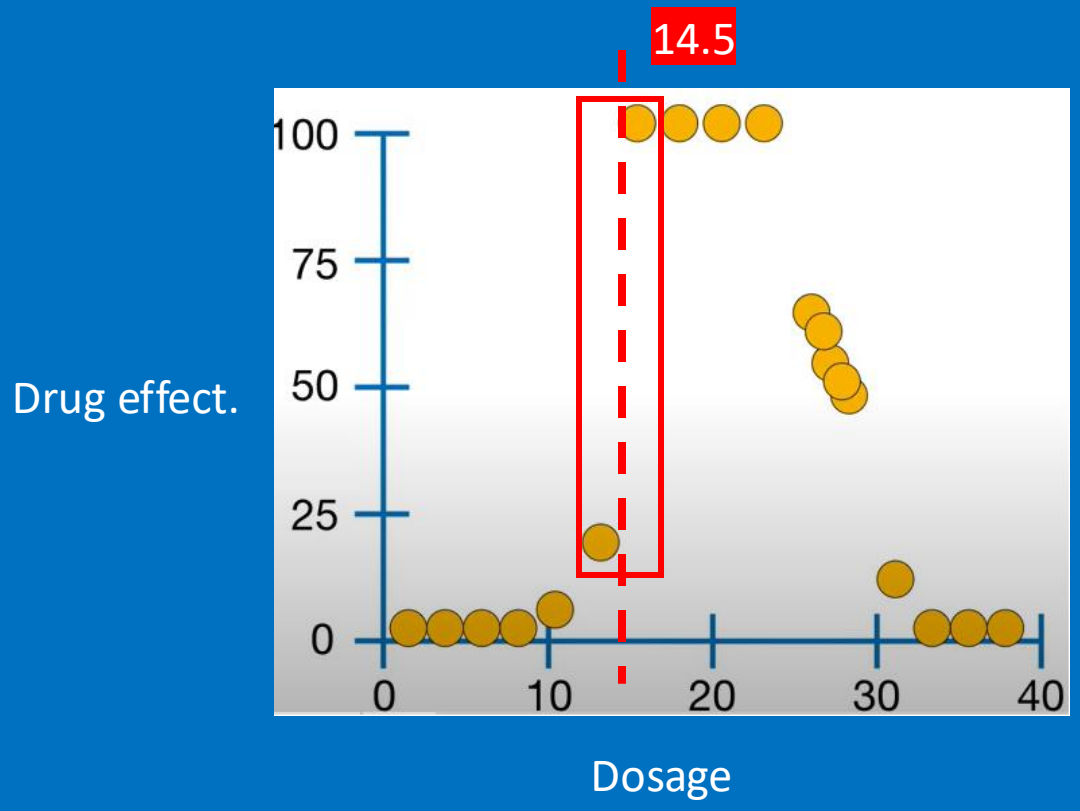
By repeating the above steps, we can get a series of root node threshold, and the corresponding errors

The dosage threshold of 14.5 has the smallest error, and it will be used to build our actual tree

Regression tree workflow

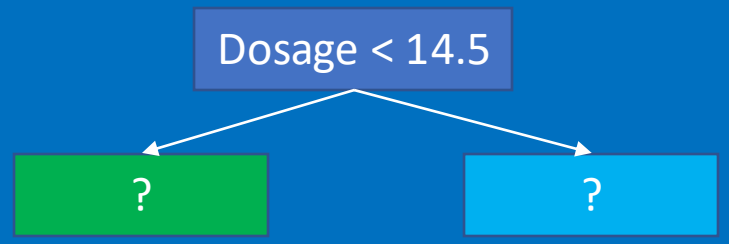
Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



By repeating the above steps, we can get a series of root node threshold, and the corresponding errors

The dosage threshold of 14.5 has the smallest error, and it will be used to build our actual tree

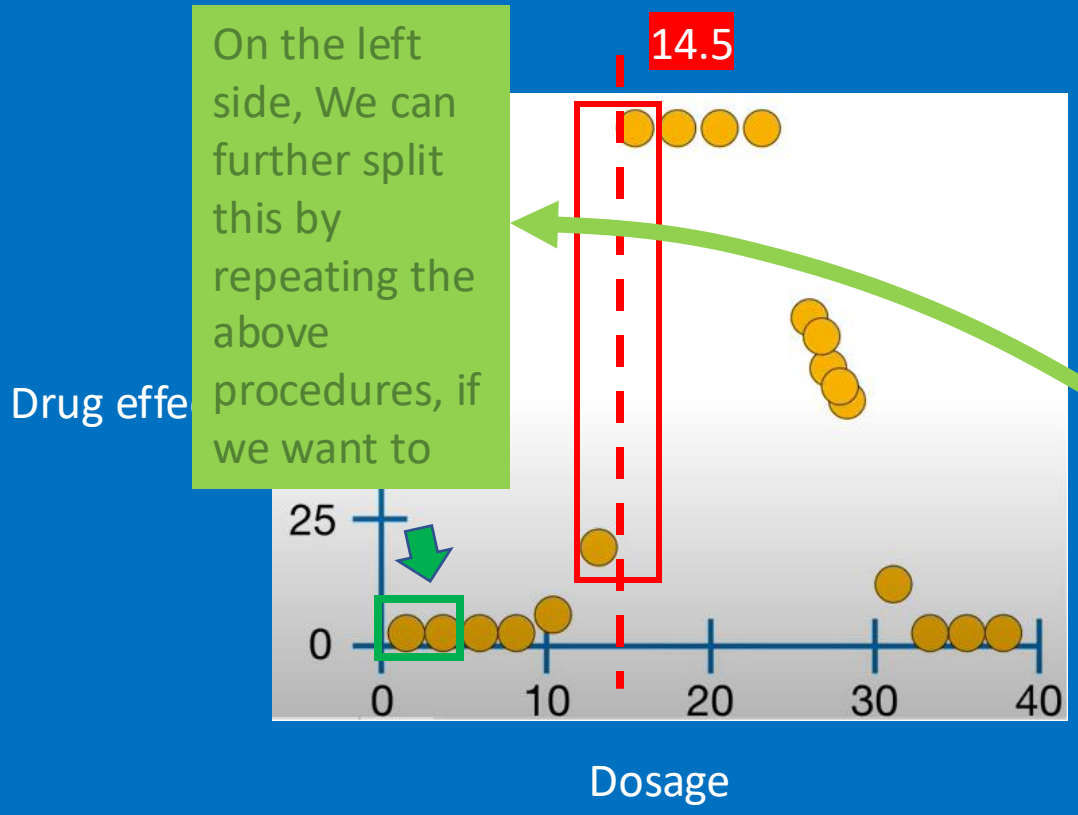


In summary, we split the data into two groups by finding the threshold that give us the smallest sum of squared residuals (errors)

Regression tree workflow

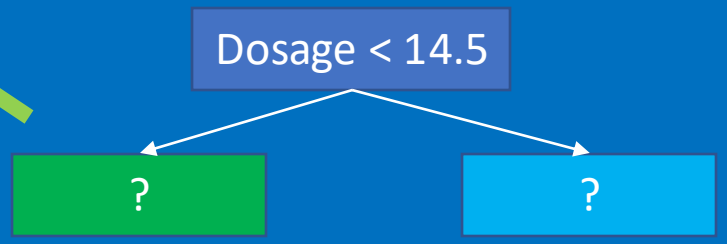
Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



By repeating the above steps, we can get a series of root node threshold, and the corresponding errors

The dosage threshold of 14.5 has the smallest error, and it will be used to build our actual tree

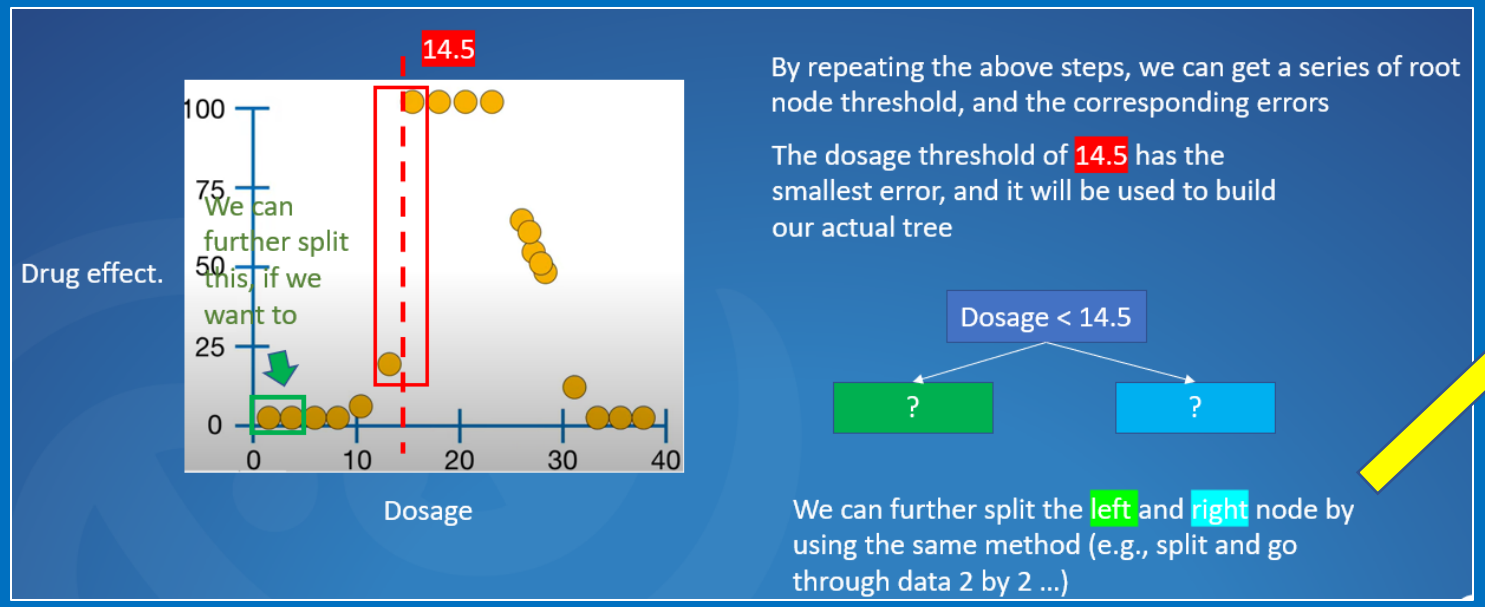


We can further split the left and right node by using the same method (e.g., split and go through data 2 by 2 …)

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio

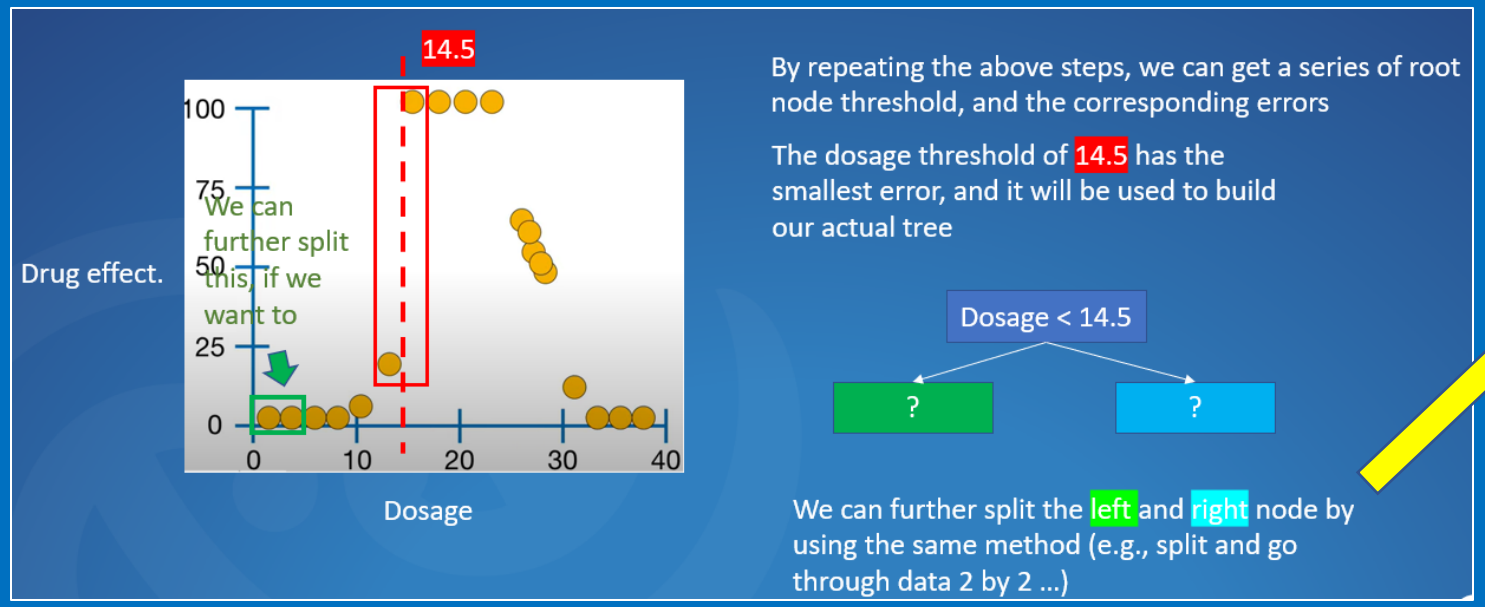


However, we may not want to do so (in order to avoid over-fitting). Usually we can constrain how many data points still being left and decide if we still want to further split the tree …

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



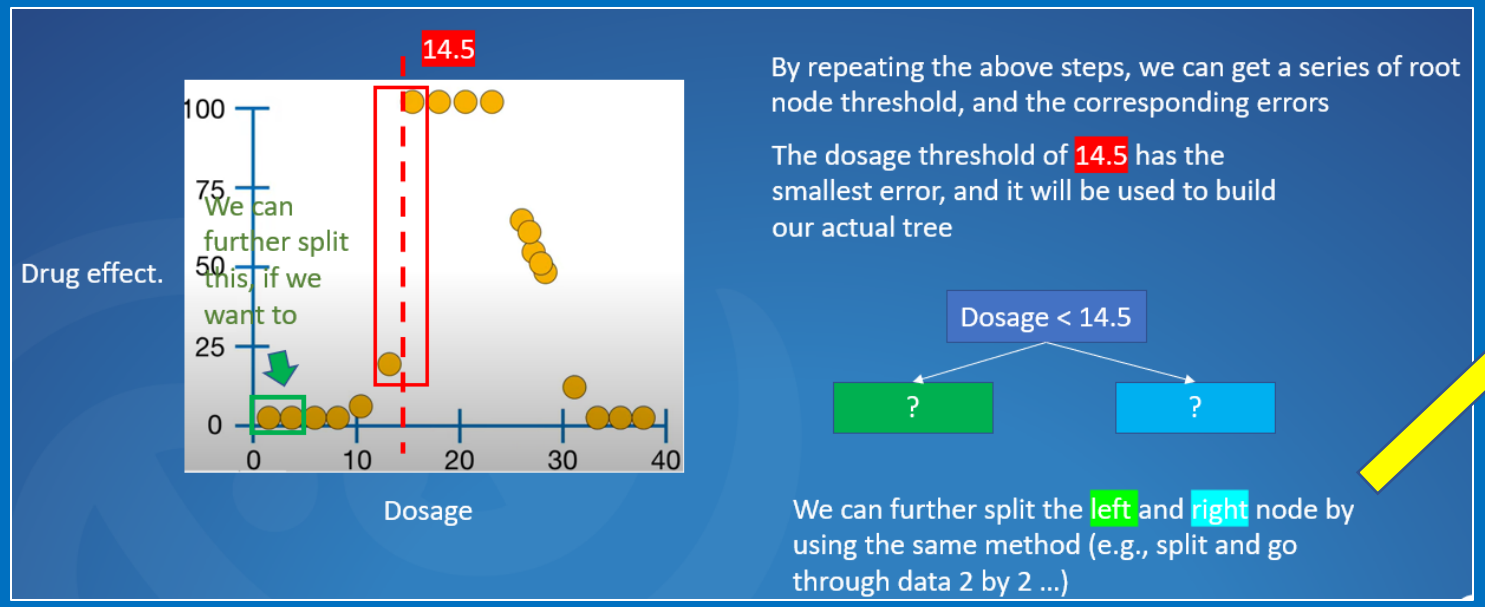
However, we may not want to do so (in order to avoid over-fitting). Usually we can constrain how many datapoints still being left and decide if we still want to further split the tree …

For example, on the left side of "dosage < 14.5", we still have 6 points

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



However, we may not want to do so (in order to avoid over-fitting). Usually we can constrain how many datapoints still being left and decide if we still want to further split the tree …

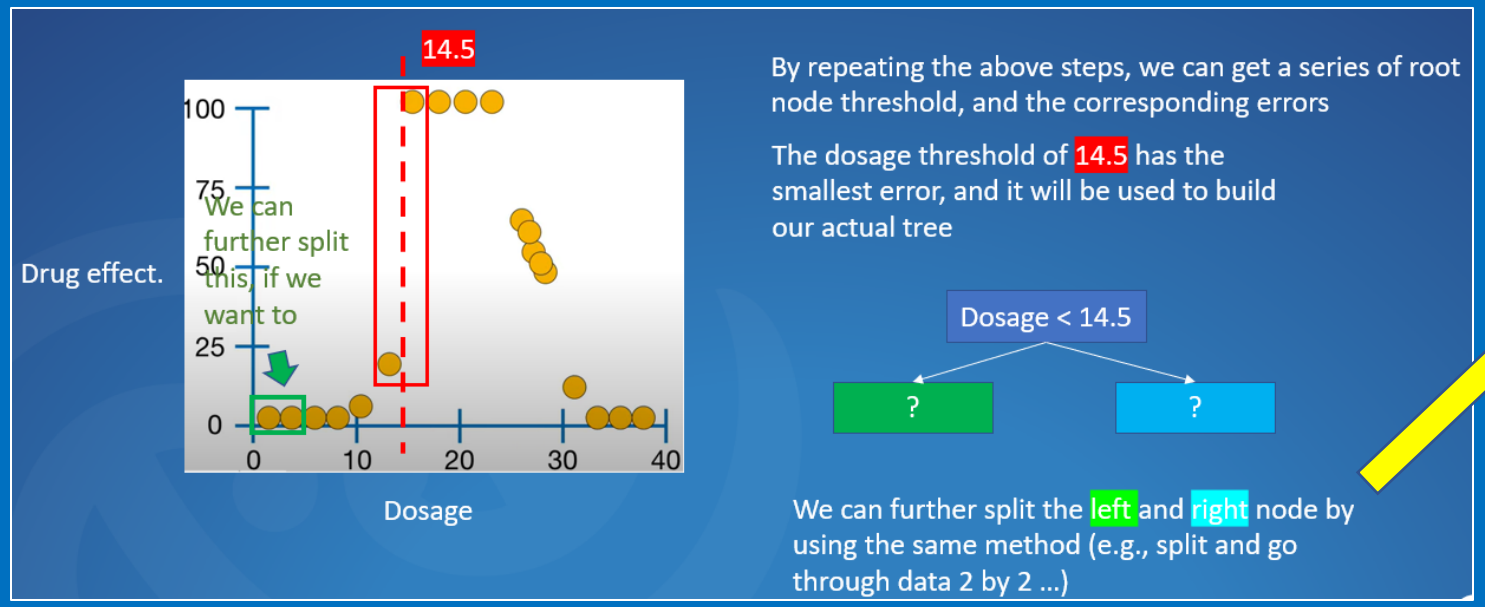
For example, on the left side of "dosage < 14.5", we still have 6 points

- If the condition says that we stop splitting when the "minimum data points == 6", then we stop splitting

Regression tree workflow

Similar to classification tree, regression tree is built from top down

Let's assume that we have only one predictors, which is drug dosage, and the target is drug effective ratio



However, we may not want to do so (in order to avoid over-fitting). Usually we can constrain how many datapoints still being left and decide if we still want to further split the tree …

For example, on the left side of "dosage < 14.5", we still have 6 points

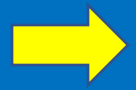
- If the condition says that we stop splitting when the "minimum data points == 6", then we stop splitting
- If the condition says that we stop splitting when the "minimum data points == 4", then we can keep splitting further (since we still have more than 4 data points …) …

Regression tree workflow

Similar to classification tree, regression tree is built from top down

For multiple predictors, it is the similar process, we basically go through the predictor one by one, and calculate their best threshold, and compare the errors, respectively. Then decide which one to be used as the root node …

Dosage	Drug effect.
10	58
20	60
35	57
5	44
…	…



Dosage		Drug effect.
10	…	58
20	…	60
35	…	57
5	…	44
…		…