

$$Y \approx f(X)$$



Hiperparam	valor
------------	-------

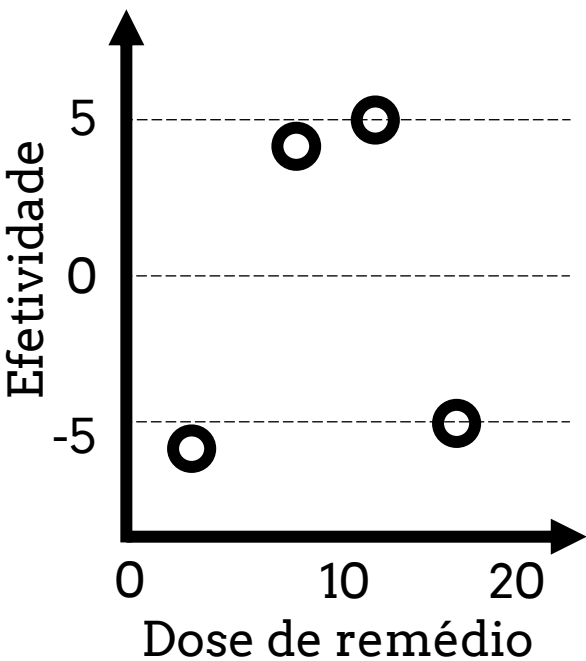
$\lambda$

$\gamma$

$\epsilon$

Tree Depth

Trees



$f(\mathbf{x}) =$

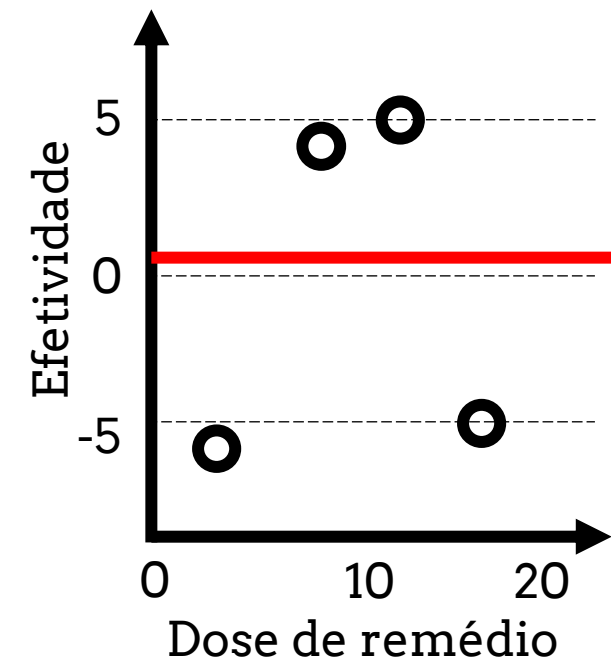


Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	

Tree Depth

Trees 2

$$f(\mathbf{x}) = 0.5$$



Hiperparam	valor
------------	-------

$\lambda$

$\gamma$

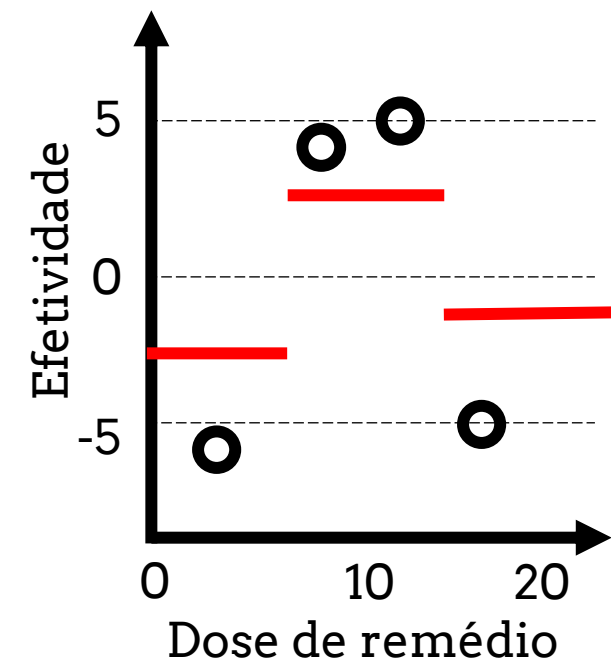
$\varepsilon$

Tree Depth

Trees

2

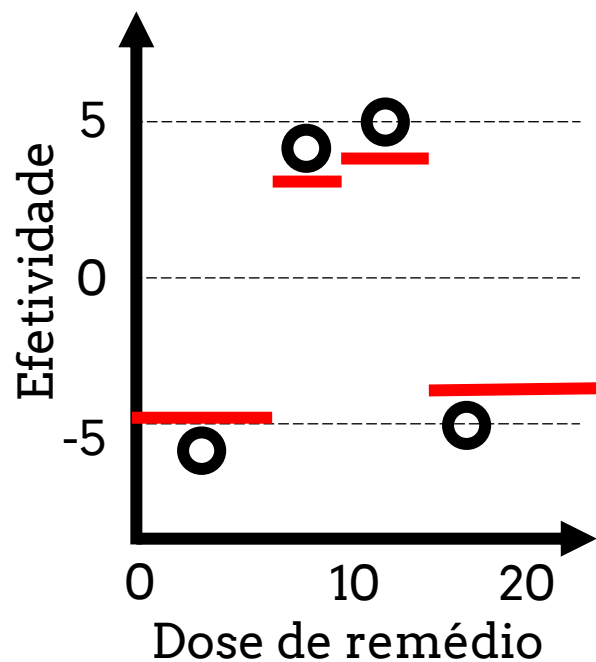
$$f(x) = 0.5 + \varepsilon \times \text{tree}$$



Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	
Tree Depth	
Trees	2



$$f(x) = 0.5 + \varepsilon \times \text{tree} + \varepsilon \times \text{tree}$$



Hiperparam	valor
------------	-------

$\lambda$

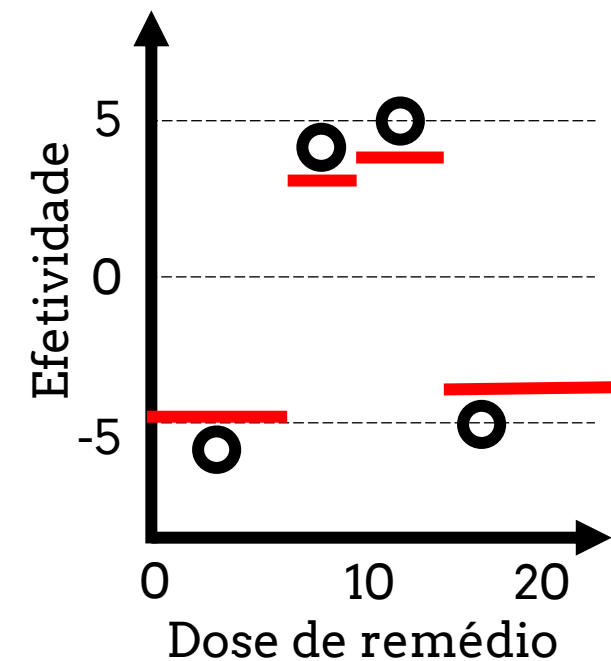
$\gamma$

$\varepsilon$

Tree Depth

Trees

2

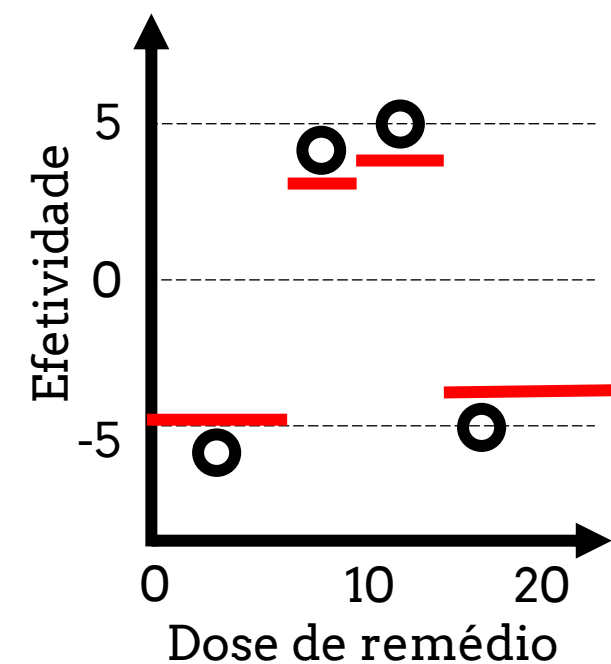


$$f(x) = 0.5 + \varepsilon \times \text{tree} + \varepsilon \times \text{tree}$$

"Learning Rate"



Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2



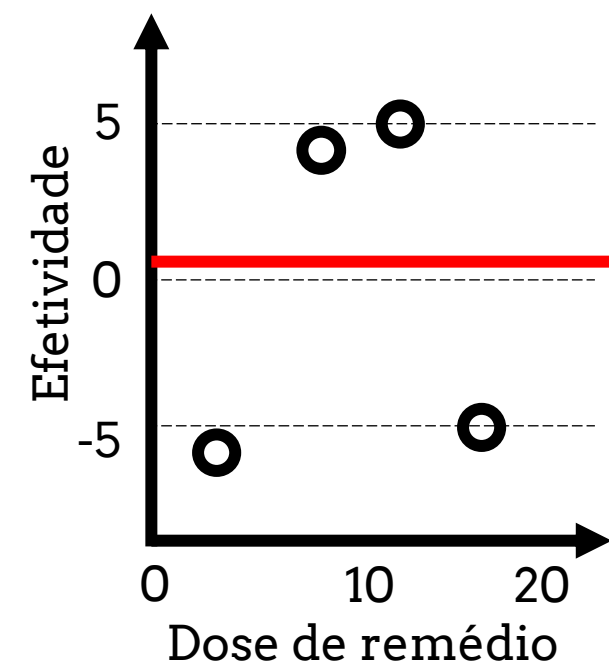
$$f(x) = 0.5 + 0.3 \times \text{[Tree Diagram]} + 0.3 \times \text{[Tree Diagram]}$$

"Learning Rate"



Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

$$f(\mathbf{x}) = 0.5$$

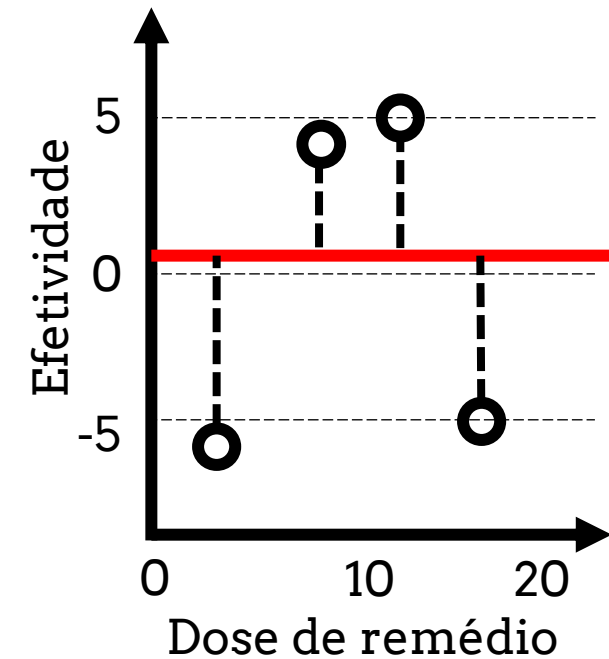


Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

$$f(\mathbf{x}) = 0.5$$



$$resíduo_i = y_i - f(x_i)$$



Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

$$f(\mathbf{x}) = 0.5$$



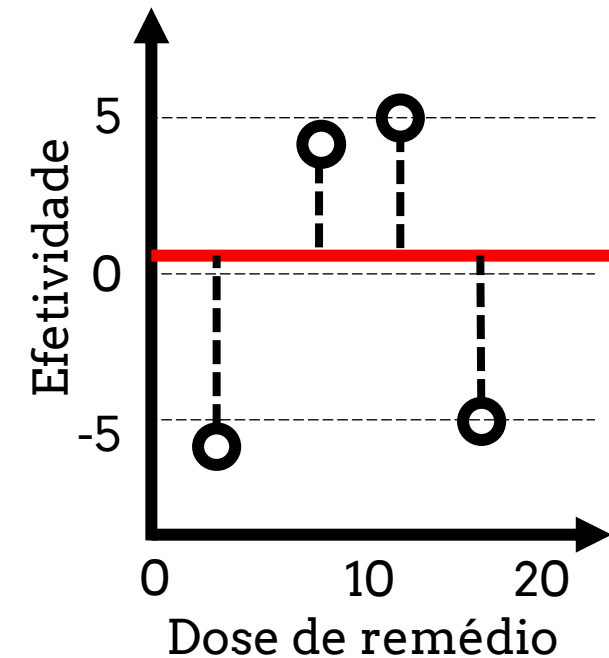
$$resíduo_i = y_i - f(x_i)$$

$$resíduo_1 = -6 - 0.5 = -6.5$$

$$resíduo_2 = 4 - 0.5 = 3.5$$

$$resíduo_3 = 5 - 0.5 = 4.5$$

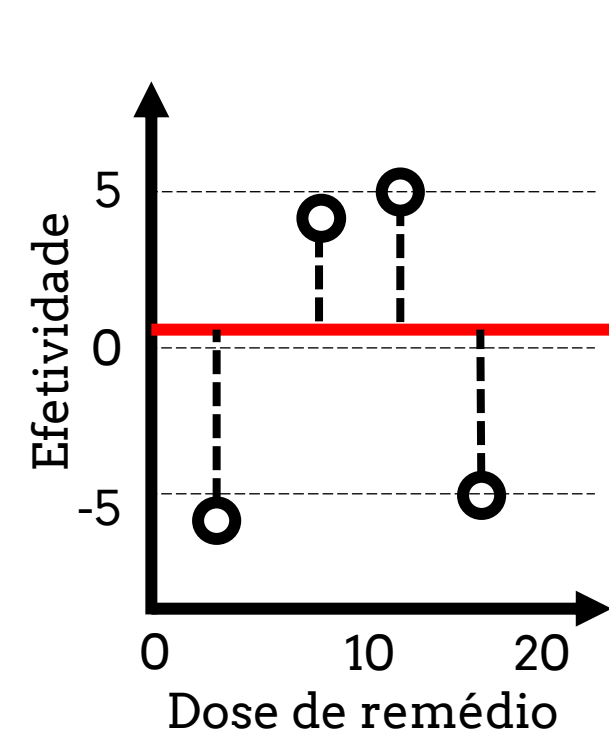
$$resíduo_4 = -5 - 0.5 = -5.5$$



Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

$$f(\mathbf{x}) = 0.5$$

-6.5, 3.5, 4.5, -5.5

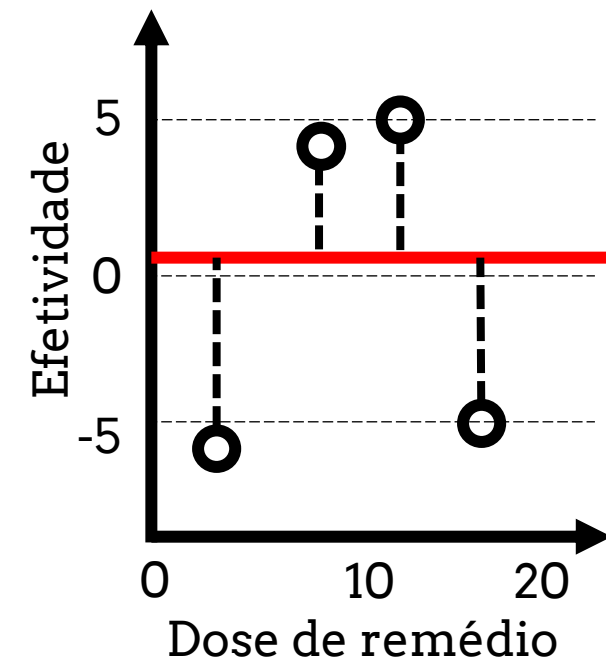


$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
$\lambda$	
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

$$f(\mathbf{x}) = 0.5$$

-6.5, 3.5, 4.5, -5.5



$$Similaridade = \frac{(\sum res\acute{i}duos)^2}{\#res\acute{i}duos + \lambda}$$

“Regularization Parameter”  
Ou “loss reduction”

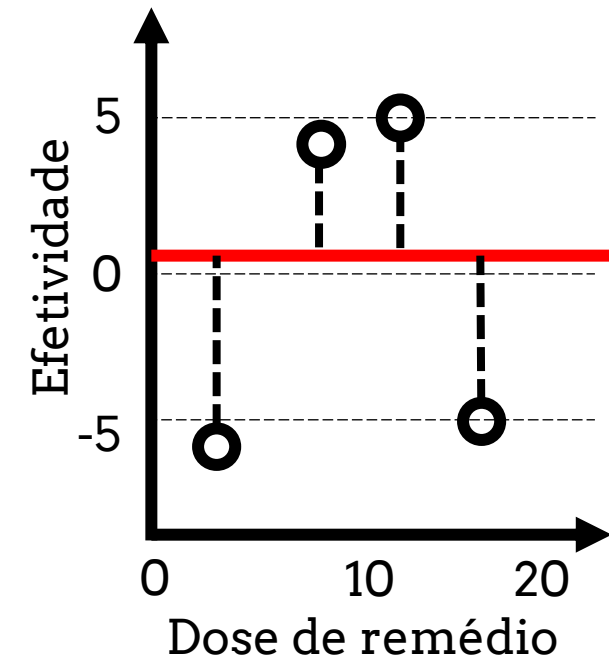
Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

$$f(\mathbf{x}) = 0.5$$

-6.5, 3.5, 4.5, -5.5



Similaridade = \_\_\_\_\_



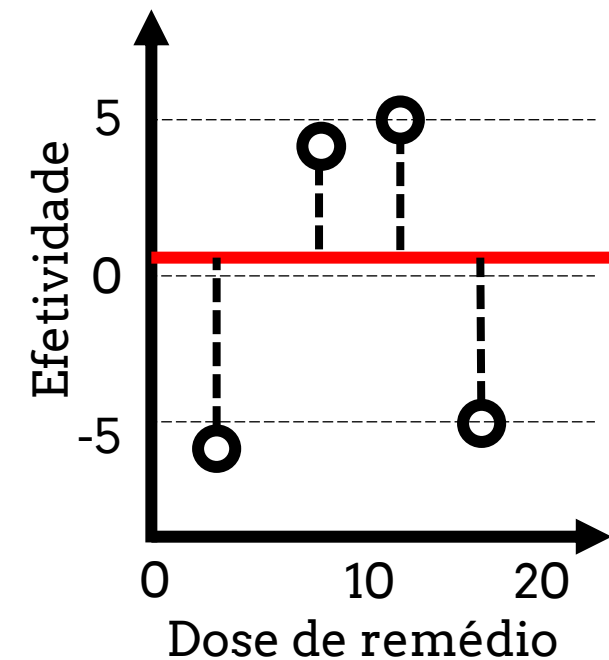
$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

$$f(\mathbf{x}) = 0.5$$

$$\text{Similaridade} = 4$$

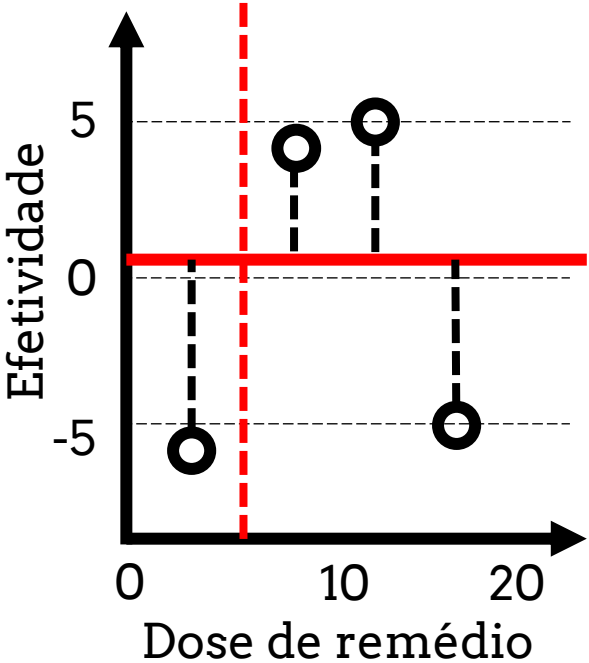
$$-6.5, 3.5, 4.5, -5.5$$



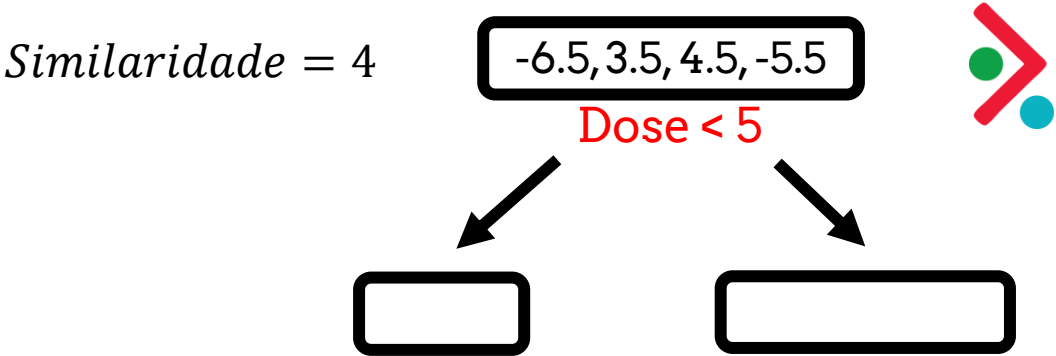
# Hora da primeira árvore

$$\text{Similaridade} = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2



$f(\mathbf{x}) = 0.5$



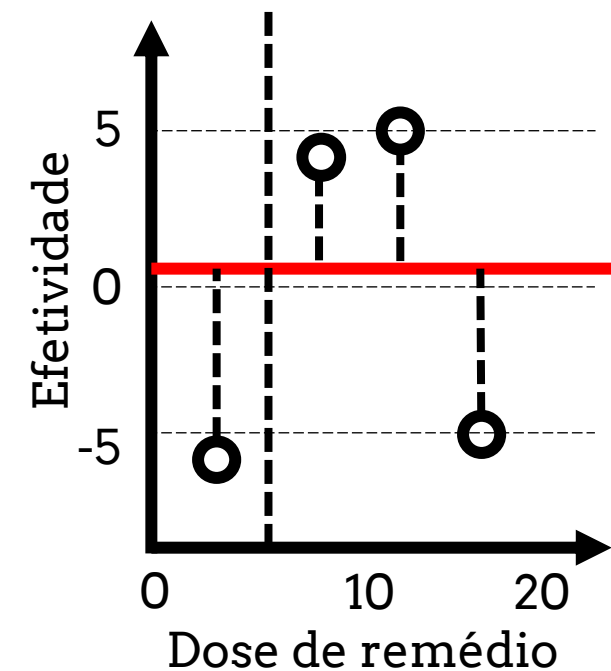
$Similaridade_{esq} = \rule{1.5cm}{0.4pt} =$

$Similaridade_{dir} = \rule{1.5cm}{0.4pt} =$

$$Similaridade = \frac{(\sum res\acute{i}duos)^2}{\#res\acute{i}duos + \lambda}$$

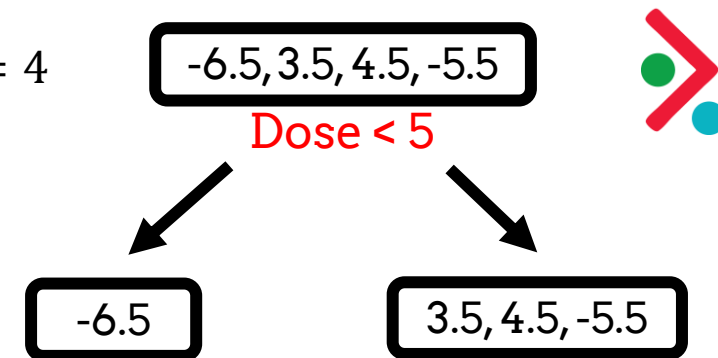


Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2



$$f(\mathbf{x}) = 0.5$$

$$Similaridade = 4$$

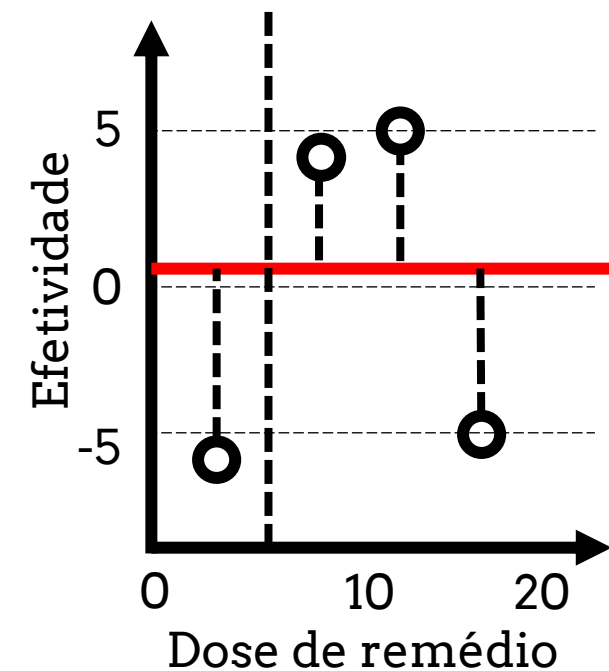


$$Similaridade_{esq} = \frac{(-6.5)^2}{1 + 0} = 42.25$$

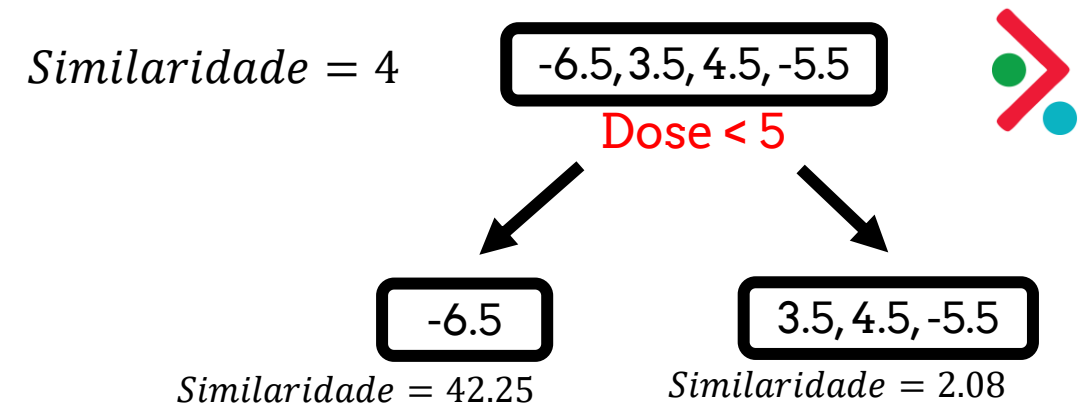
$$Similaridade_{dir} = \frac{(3.5 + 4.5 - 5.5)^2}{3 + 0} = 2.08$$

$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2



$$f(\mathbf{x}) = 0.5$$



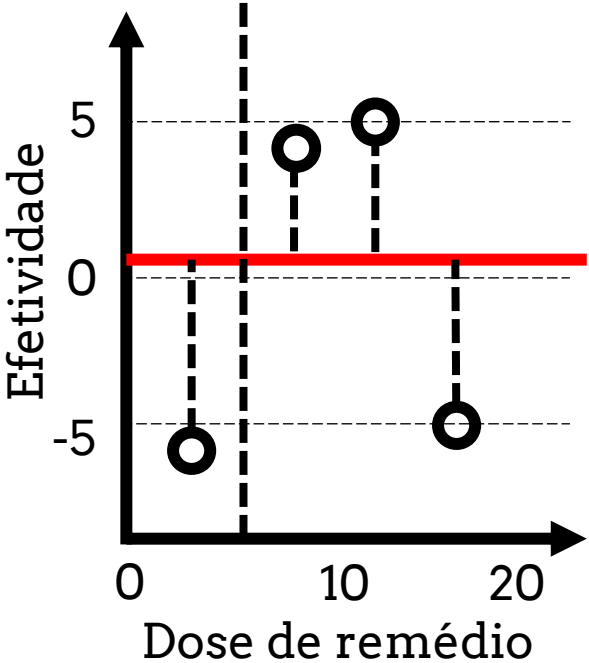
$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3

Tree Depth

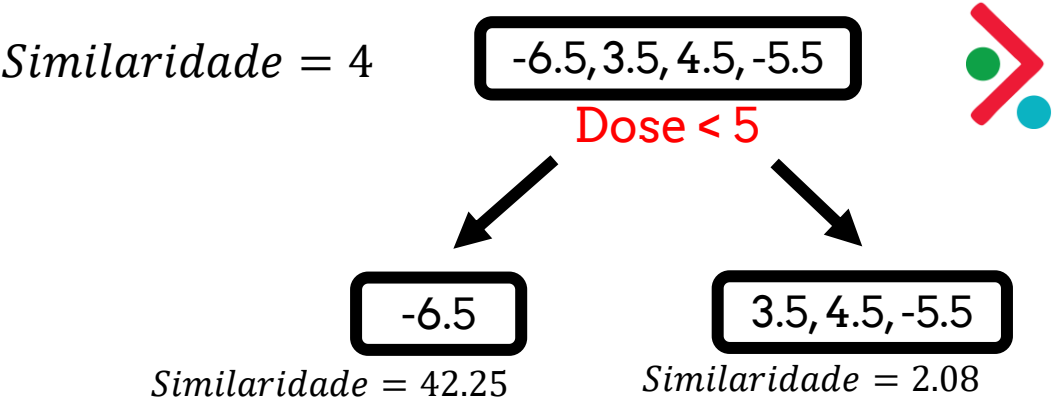
Trees 2



$f(\mathbf{x}) = 0.5$

$Gain =$

Pergunta	Gain
Dose < 5	



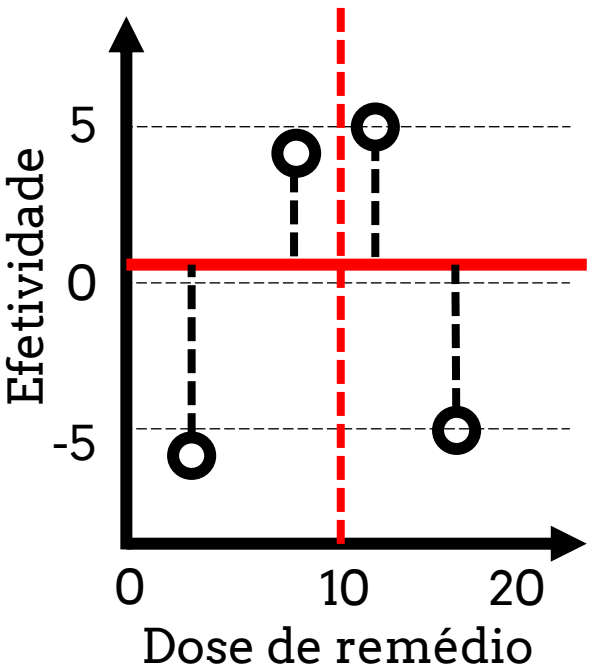
$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$

$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3

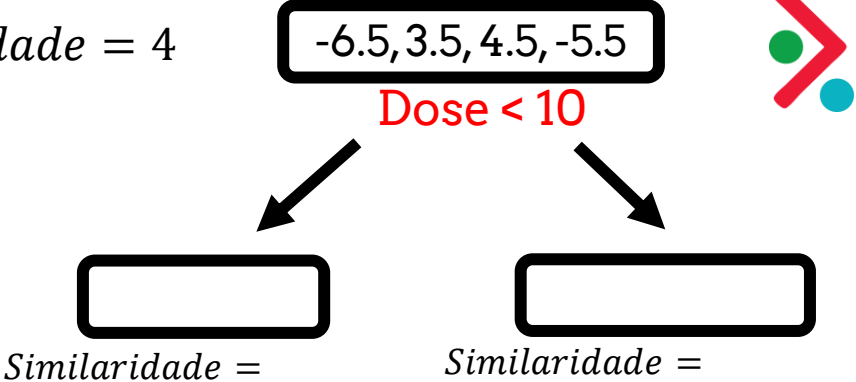
Tree Depth

Trees 2



$f(x) = 0.5$

Similaridade = 4



$Similaridade_{esq} = \rule{1cm}{0.4pt} =$

$Similaridade_{dir} = \rule{1cm}{0.4pt} =$

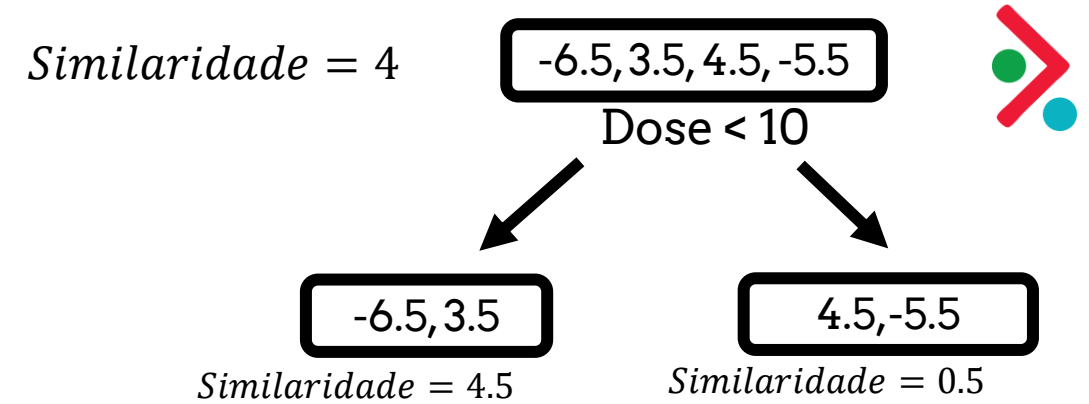
Pergunta	Gain
Dose < 5	40.33
Dose < 10	

$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$

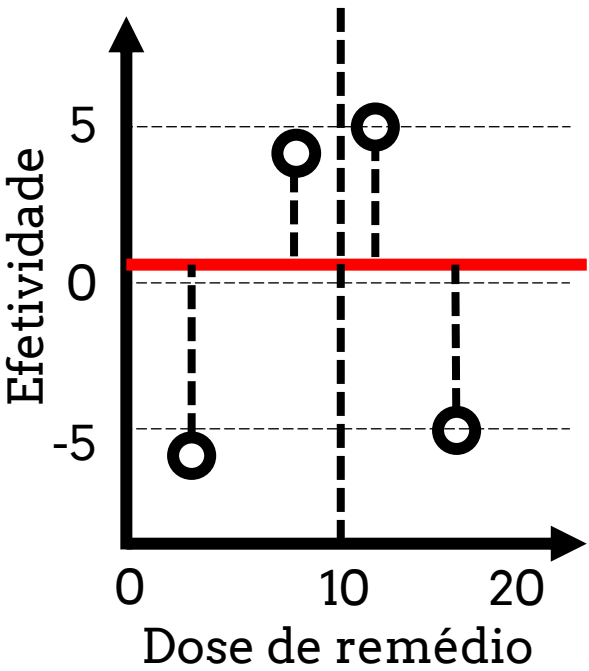
$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3

$$f(\mathbf{x}) = 0.5$$



Tree Depth  
Trees 2



$$Similaridade_{esq} = \frac{(-6.5 + 3.5)^2}{2 + 0} = 4.5$$

$$Similaridade_{dir} = \frac{(4.5 - 5.5)^2}{2 + 0} = 0.5$$

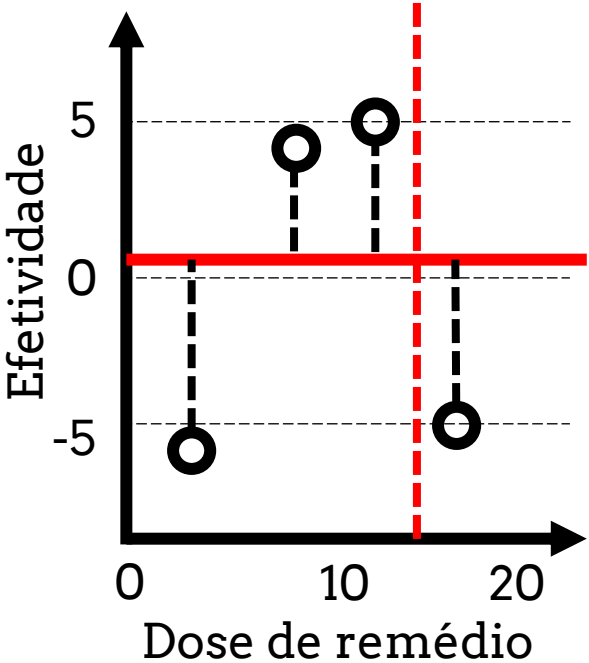
Pergunta	Gain
Dose < 5	40.33
Dose < 10	

Gain =

$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

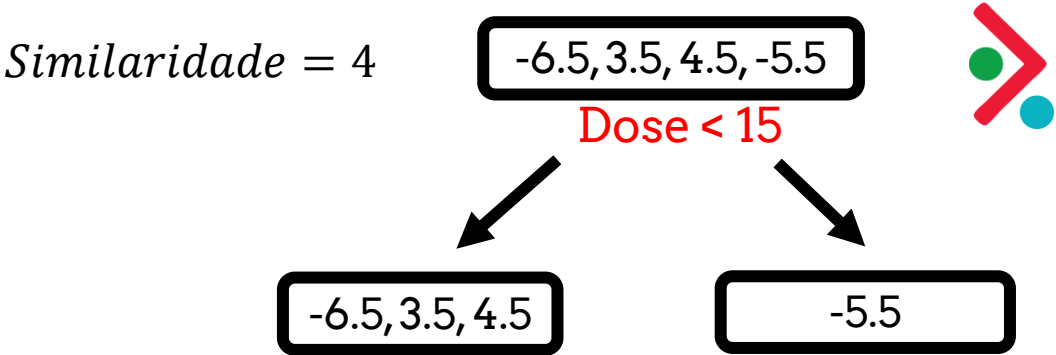
$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2



$f(\mathbf{x}) = 0.5$

Pergunta	Gain
Dose < 5	40.33
Dose < 10	1
Dose < 15	

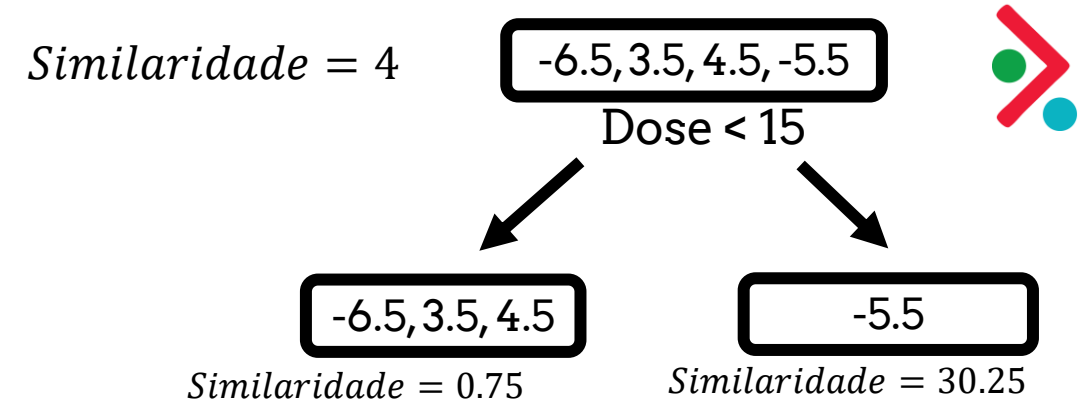


$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

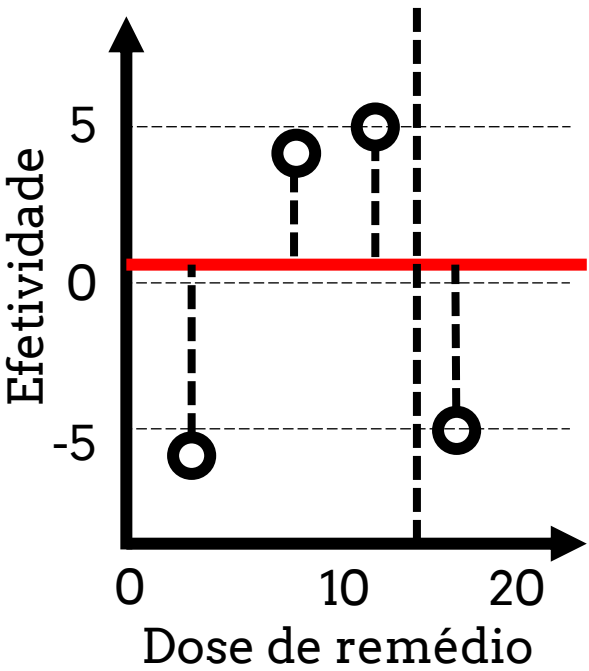
$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3

$$f(\mathbf{x}) = 0.5$$



Tree Depth  
Trees 2



$$Similaridade_{esq} = \frac{(-6.5 + 3.5 + 4.5)^2}{3 + 0} = 0.75$$

$$Similaridade_{dir} = \frac{(-5.5)^2}{1 + 0} = 30.25$$

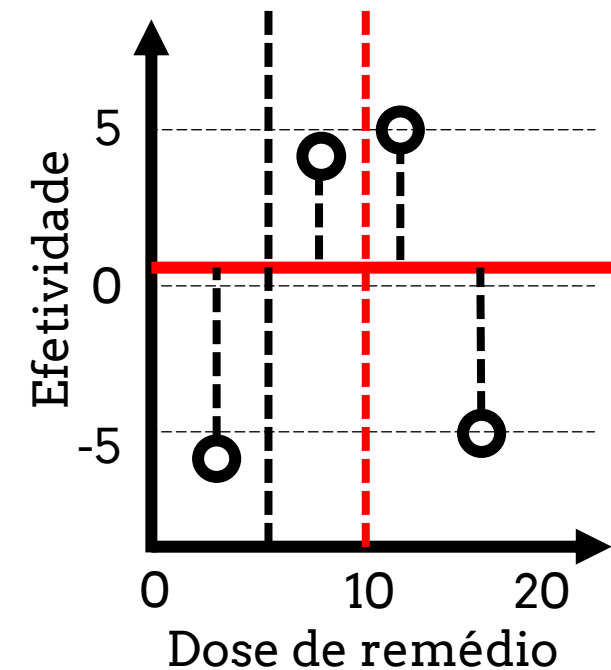
Pergunta	Gain
Dose < 5	40.33
Dose < 10	1
Dose < 15	27

$$Gain = 30.25 + 0.75 - 4 = 27$$

$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

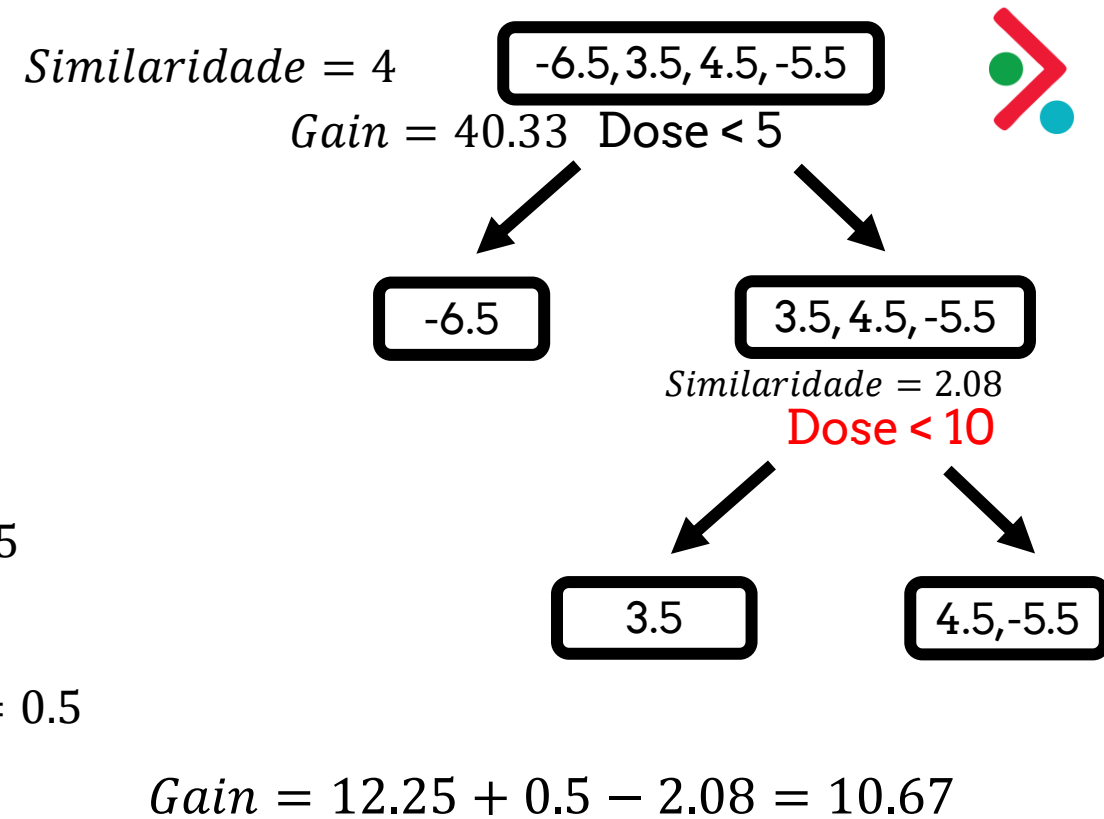


$$f(\mathbf{x}) = 0.5$$

$$Similaridade_{esq} = \frac{(3.5)^2}{1 + 0} = 12.25$$

$$Similaridade_{dir} = \frac{(4.5 - 5.5)^2}{2 + 0} = 0.5$$

Pergunta	Gain
Dose < 10	10.67
Dose < 15	

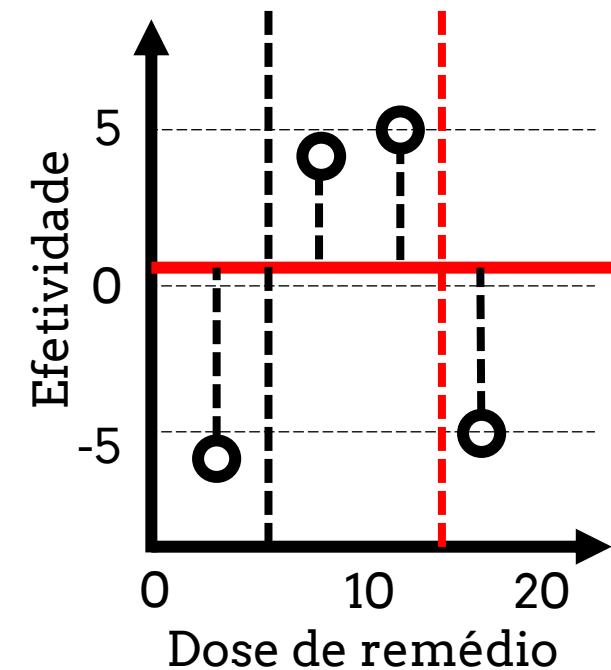


$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$



Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	
Trees	2

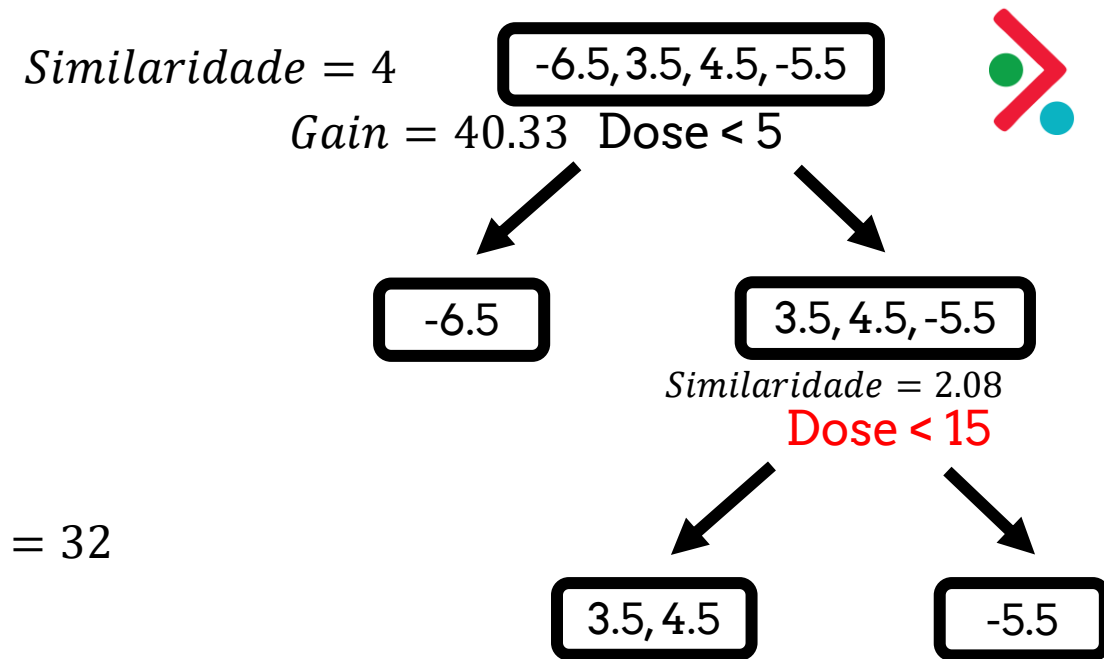


$$f(\mathbf{x}) = 0.5$$

$$Similaridade_{esq} = \frac{(+3.5 + 4.5)^2}{2 + 0} = 32$$

$$Similaridade_{dir} = \frac{(-5.5)^2}{1 + 0} = 30.25$$

Pergunta	Gain
Dose < 10	10.67
Dose < 15	60.17

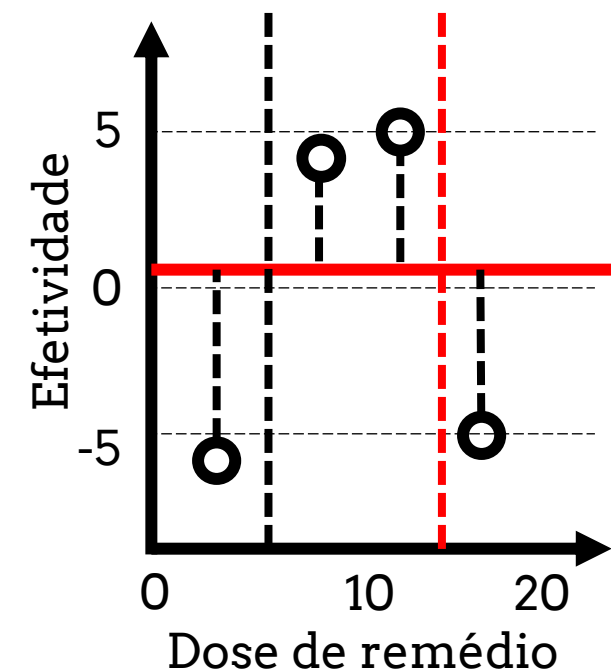


$$Gain = 30.25 + 32 - 2.08 = 60.17$$

$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	2
Trees	2

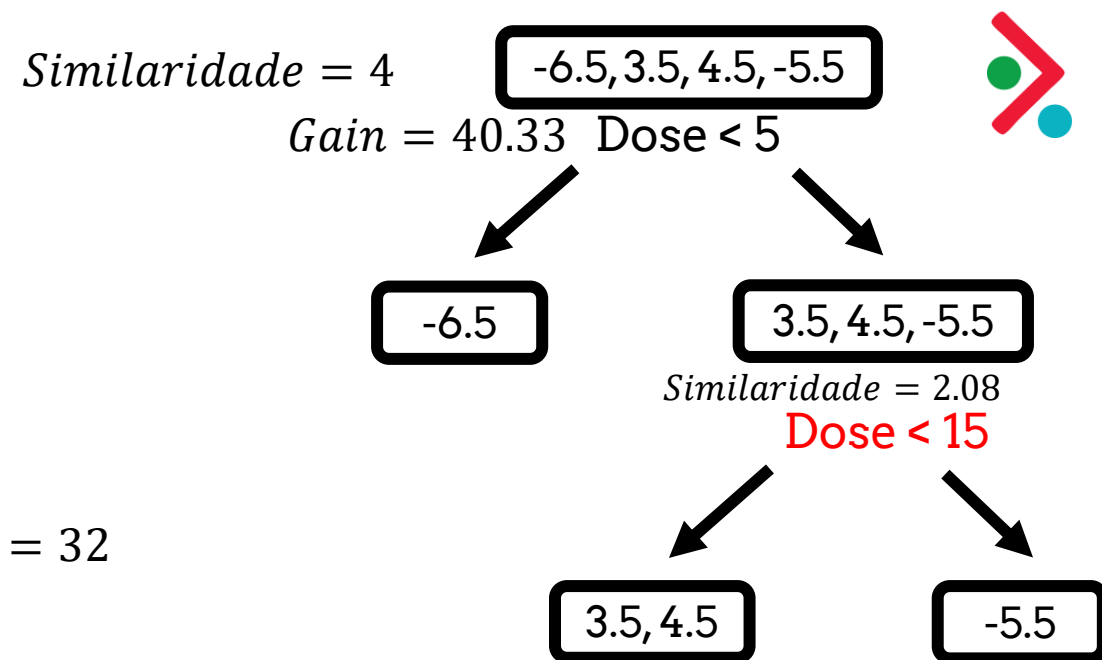


$$f(\mathbf{x}) = 0.5$$

$$Similaridade_{esq} = \frac{(+3.5 + 4.5)^2}{2 + 0} = 32$$

$$Similaridade_{dir} = \frac{(-5.5)^2}{1 + 0} = 30.25$$

Pergunta	Gain
Dose < 10	10.67
Dose < 15	60.17



$$Gain = 30.25 + 32 - 2.08 = 60.17$$

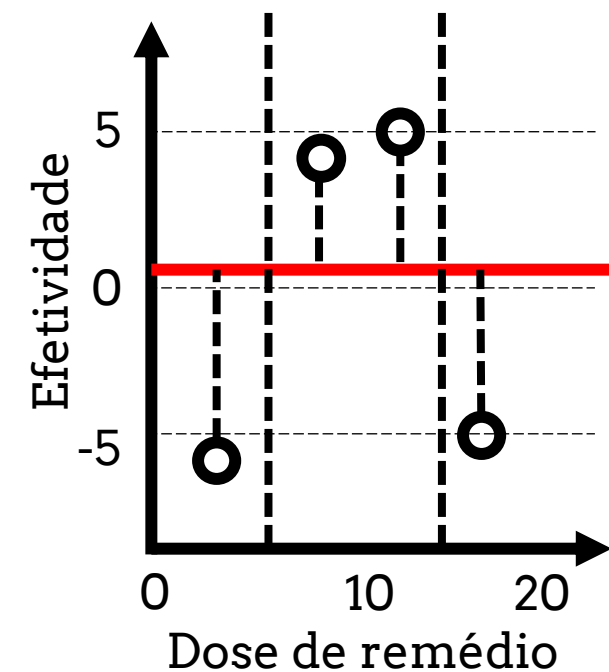
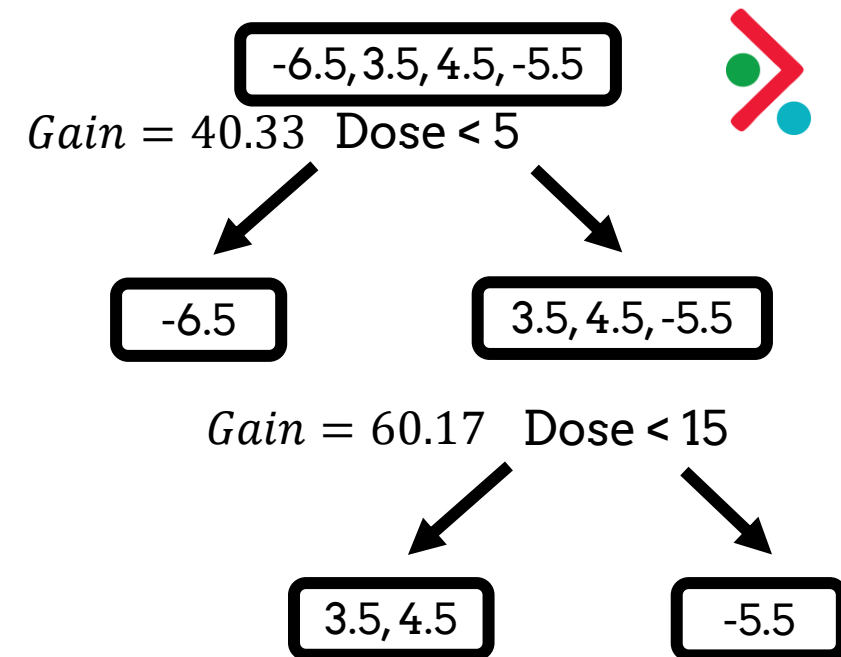
$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(\mathbf{x}) = 0.5$$

Hora da poda



Hiperparam	valor
$\lambda$	0
$\gamma$	
$\varepsilon$	0.3
Tree Depth	2
Trees	2

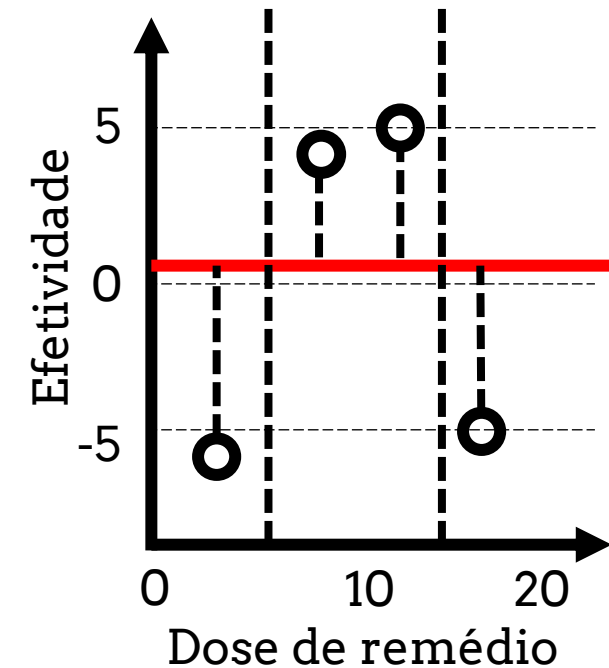
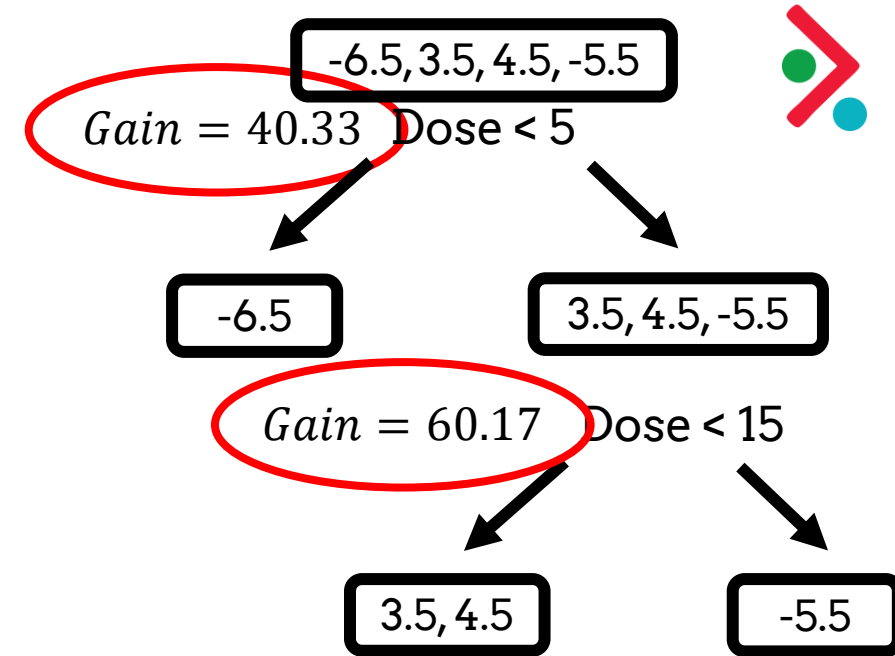
$$f(x) = 0.5$$

## Hora da poda

XGBoost usa o Gain para fazer a poda das árvores.

$\gamma$

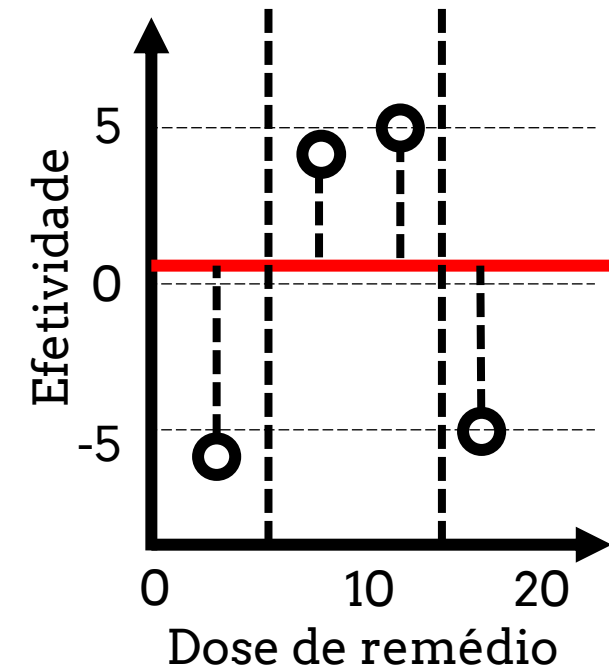
"gamma": nota de corte para o Gain.  
Se  $\text{gain} - \gamma$  for positivo, então não poda!



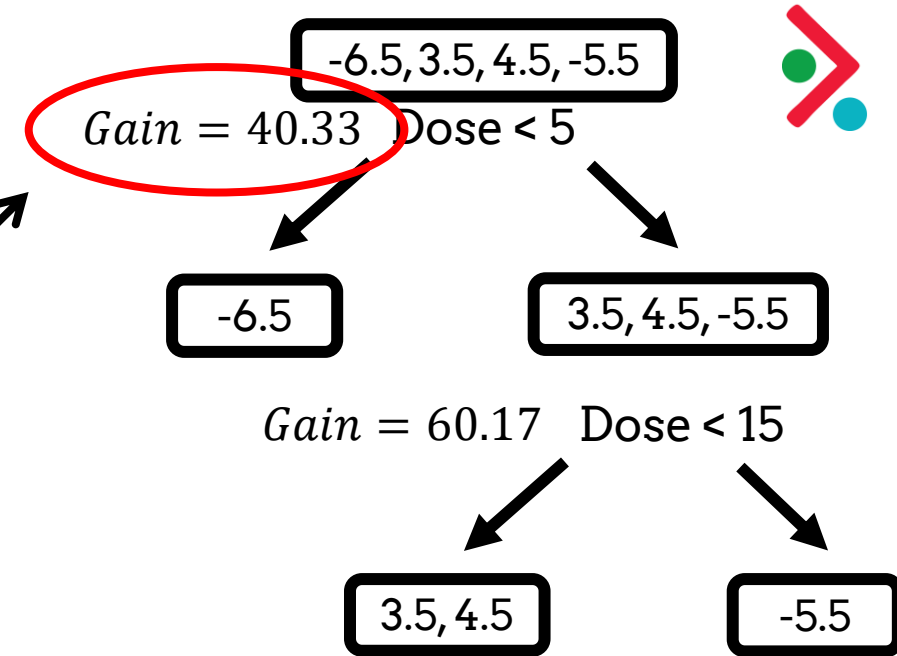


Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(\mathbf{x}) = 0.5$$



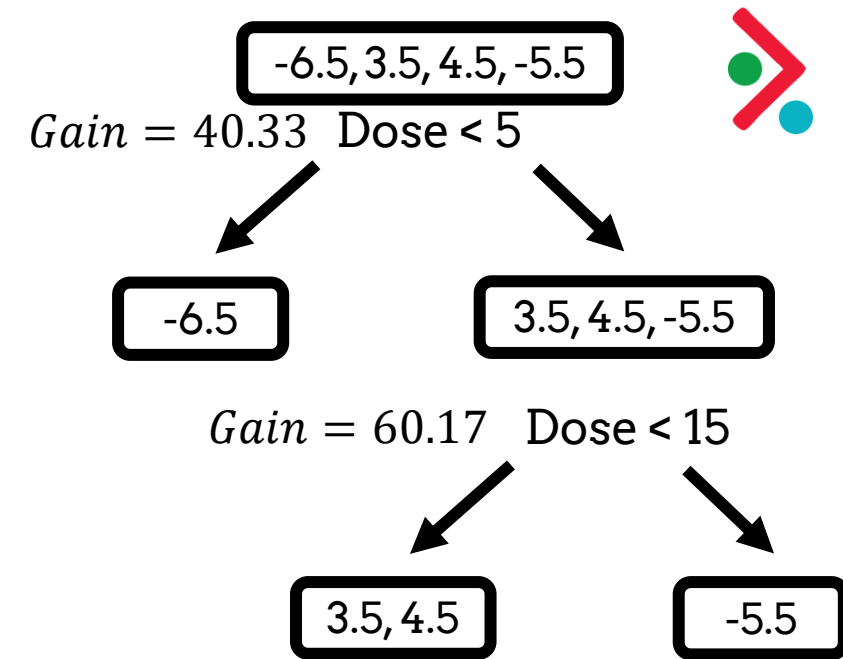
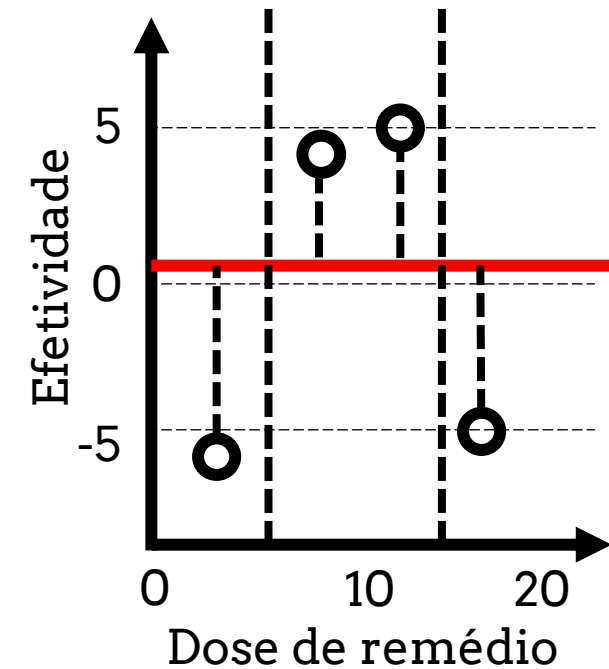
OBSERVAÇÃO: o gain do primeiro nó é menor que 50, indicando para podar. Porém o ramo filho não foi podado, por isso não podemos o pai também.



Se **gain** -  $\gamma$  for positivo, então não poda!

Hiperparam	valor
$\lambda$	0
$\gamma$	70
$\varepsilon$	0.3
Tree Depth	2
Trees	2

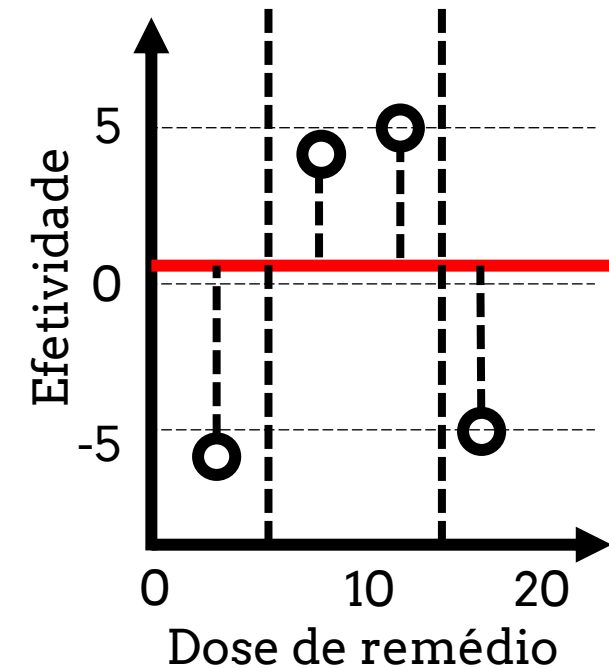
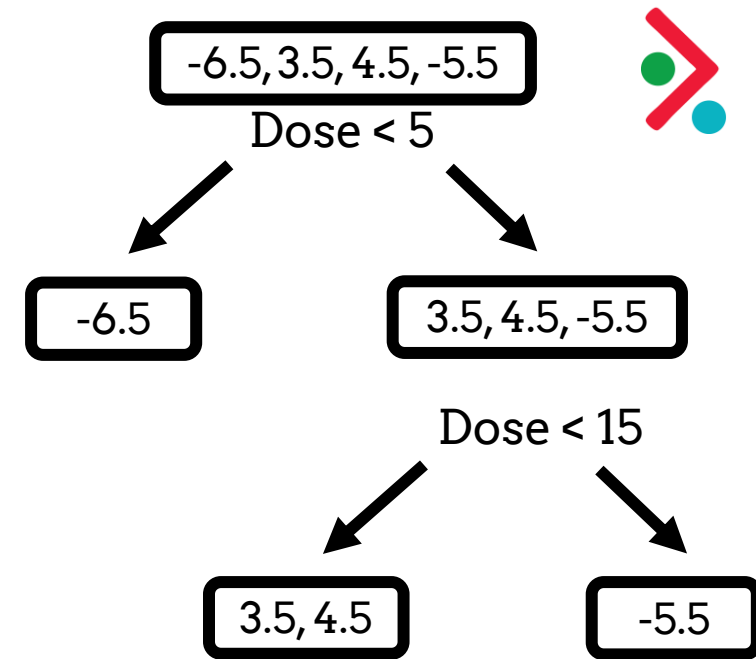
$$f(\mathbf{x}) = 0.5$$



Se **gain** -  $\gamma$  for positivo,  
então não poda!

Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

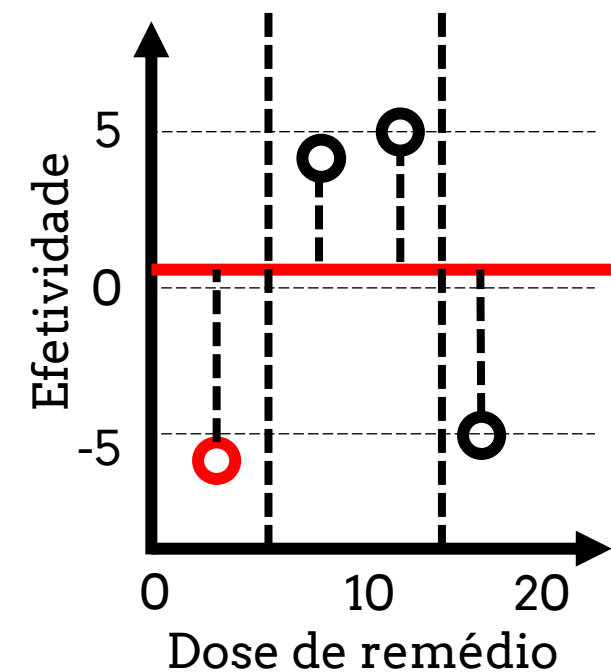
$$f(\mathbf{x}) = 0.5$$



Hora das predições  
Ou "escoragem"

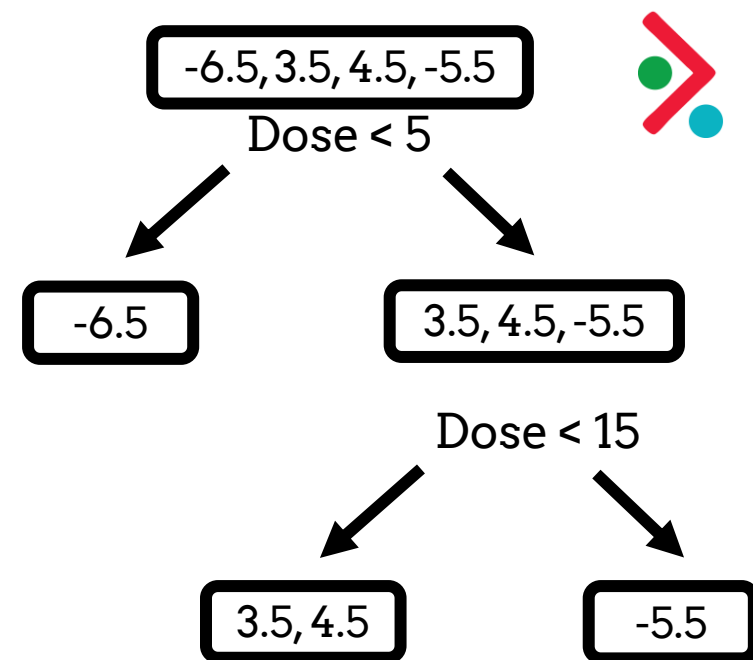


Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

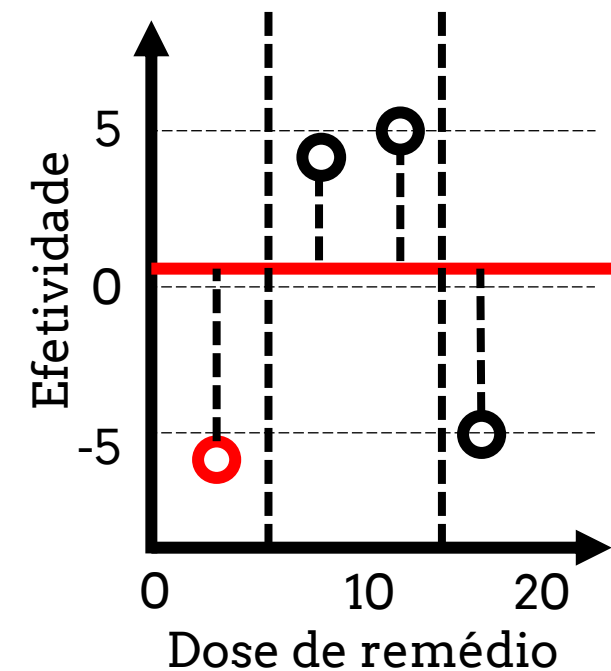


$$f(\mathbf{x}) = 0.5 + 0.3 \times \text{[icon]}$$

$f(\mathbf{x}_1)$



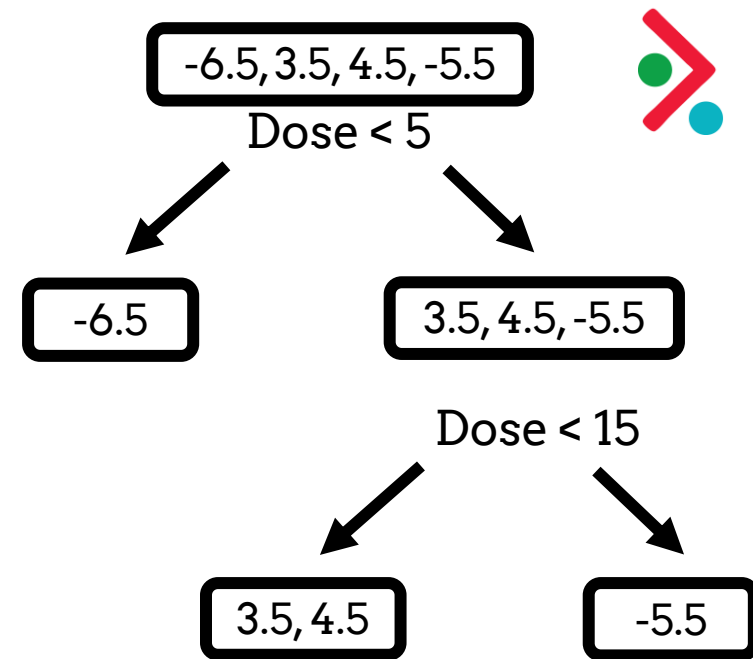
Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



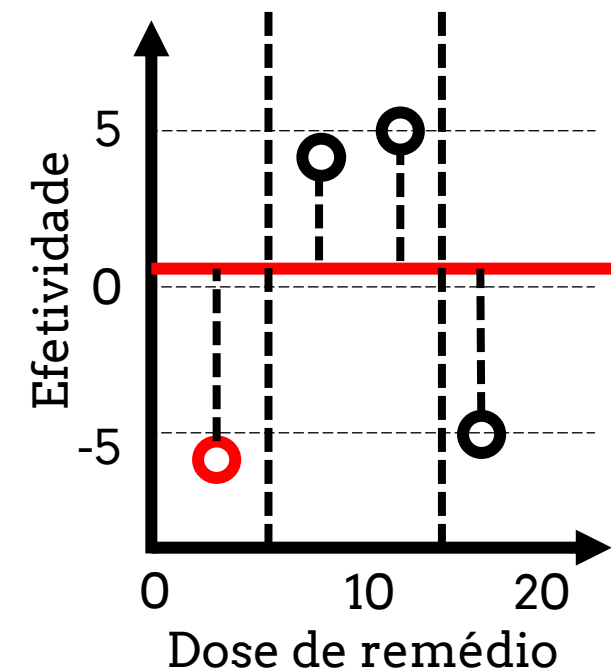
$$f(\mathbf{x}) = 0.5 + 0.3 \times \text{[tree icon]}$$

$$f(2) = 0.5 + 0.3 \times$$

$$predição = \frac{\sum resíduos}{\#resíduos + \lambda}$$



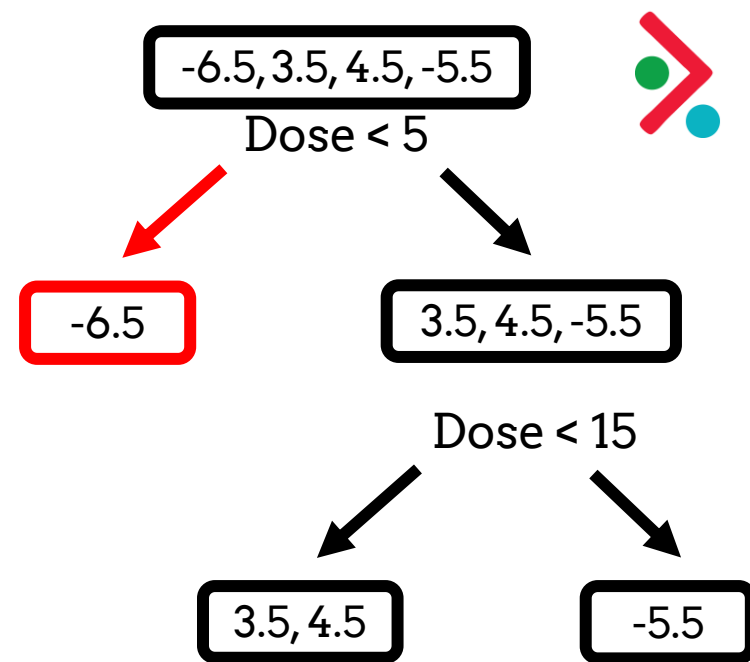
Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(x) = 0.5 + 0.3 \times \begin{array}{c} \square \\ \square \quad \square \\ \square \end{array}$$

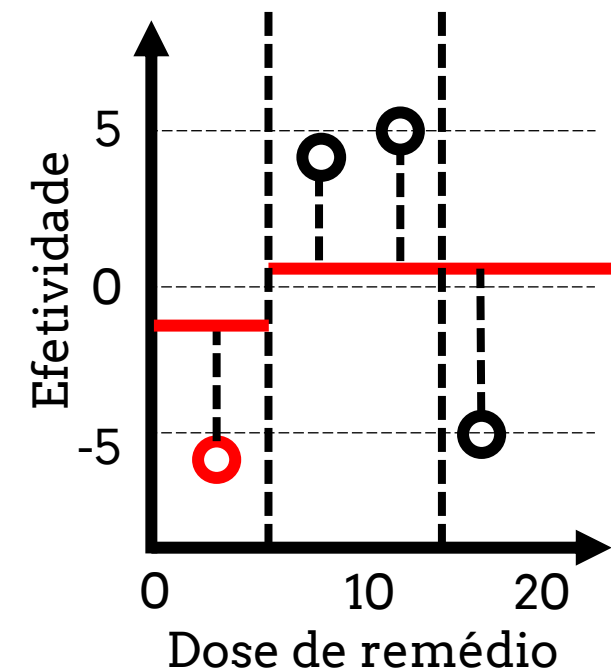
$$f(2) = 0.5 + 0.3 \times -6.5$$

$$predição = \frac{-6.5}{1 + 0} = -6.5$$



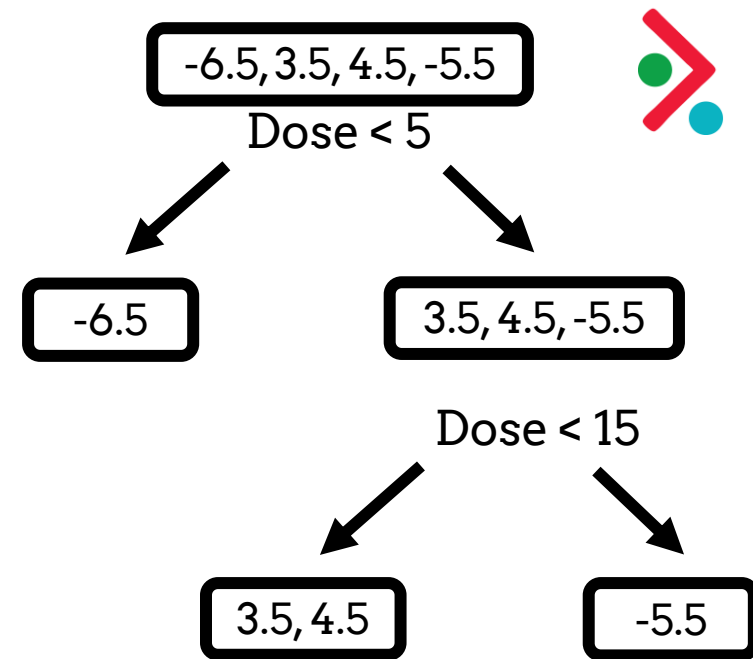
$$predição = \frac{\sum resíduos}{\#resíduos + \lambda}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(x) = 0.5 + 0.3 \times \begin{matrix} \square \\ \square \quad \square \\ \square \end{matrix}$$

$$f(2) = 0.5 + 0.3 \times -6.5 = -1.56$$



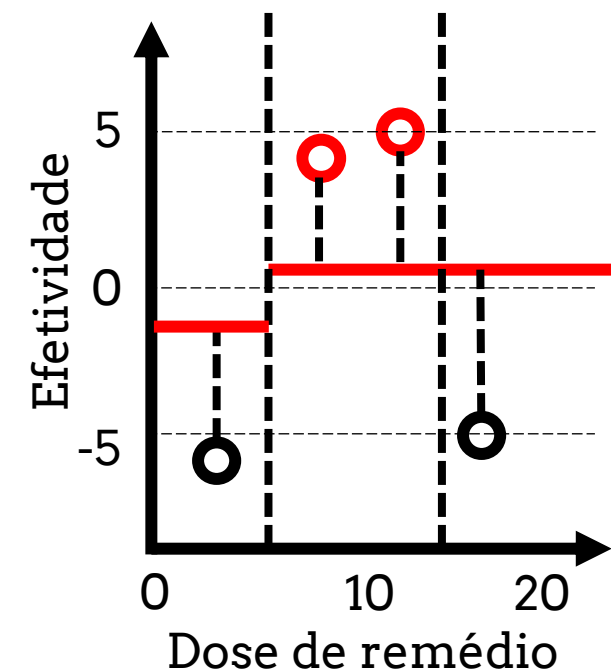
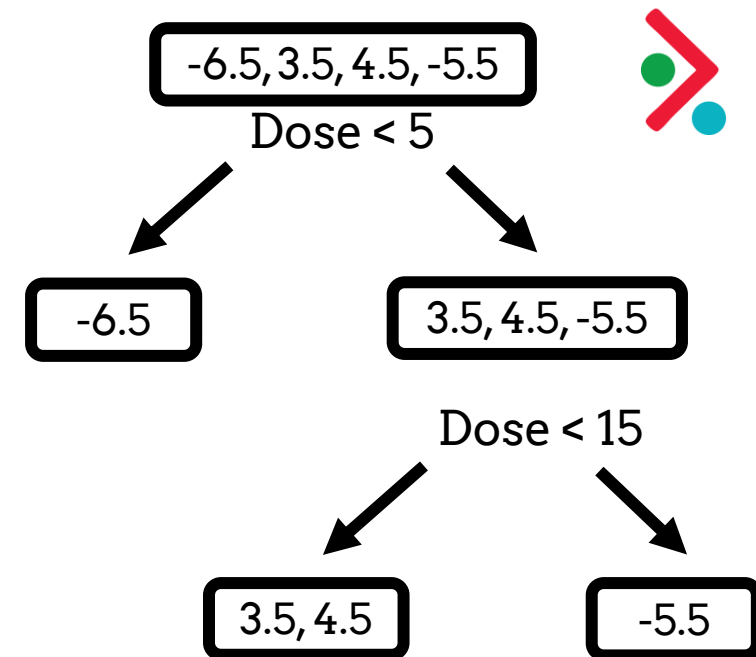
Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(x) = 0.5 + 0.3 \times \begin{array}{c} \square \\ \square \quad \square \\ \square \end{array}$$

$$f(2) = 0.5 + 0.3 \times -6.5 = -1.56$$

$$f(x_2) = 0.5 + 0.3 \times$$

$$f(x_3) = 0.5 + 0.3 \times$$



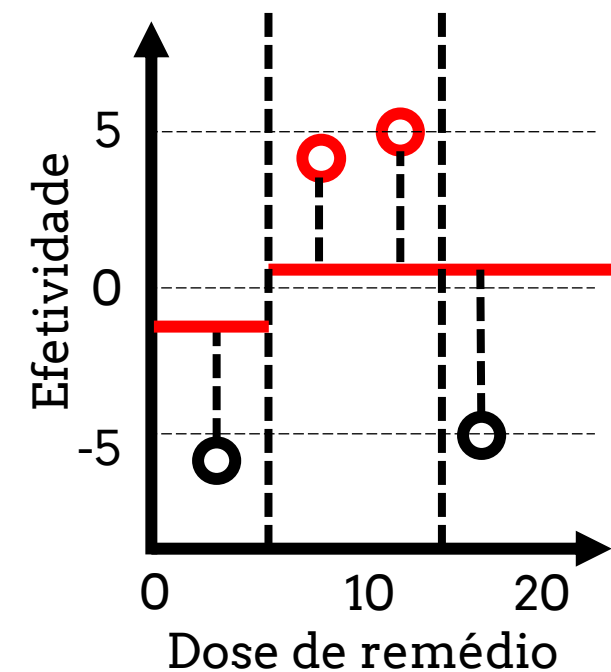
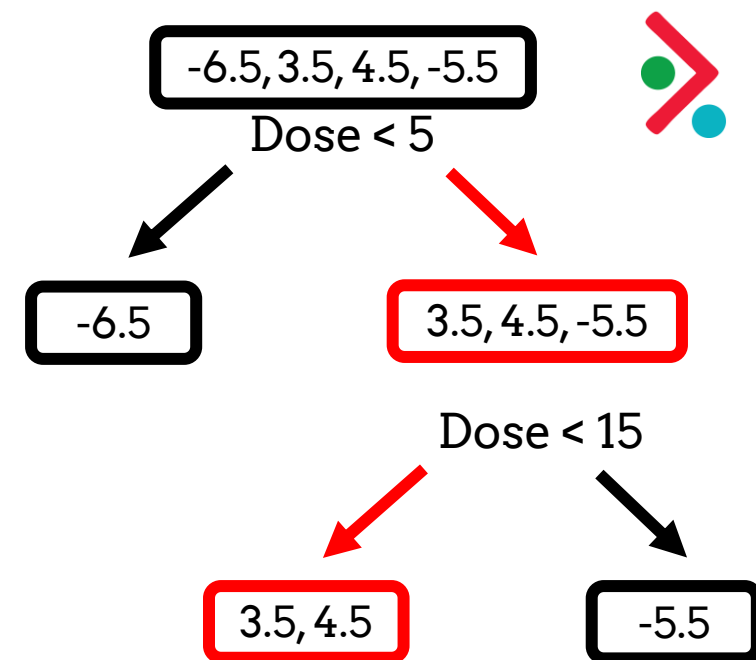
Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(x) = 0.5 + 0.3 \times \begin{array}{c} \square \\ \square \quad \square \\ \square \end{array}$$

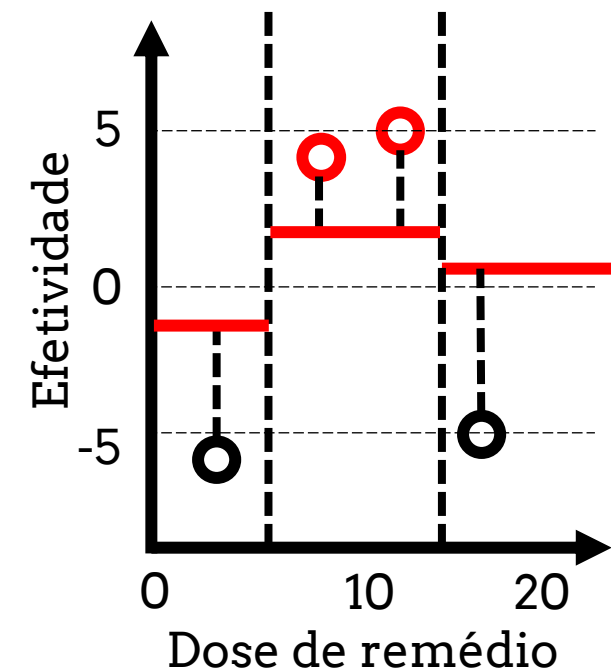
$$f(2) = 0.5 + 0.3 \times -6.5 = -1.56$$

$$f(8) = 0.5 + 0.3 \times$$

$$f(12) = 0.5 + 0.3 \times$$



Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



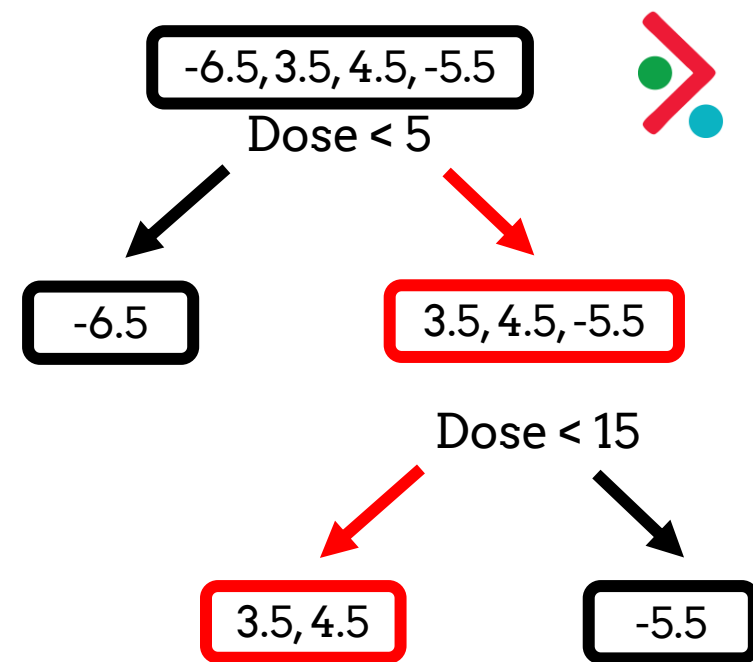
$$f(x) = 0.5 + 0.3 \times \begin{array}{c} \square \\ \square \quad \square \\ \square \end{array}$$

$$f(2) = 0.5 + 0.3 \times -6.5 = -1.56$$

$$f(8) = 0.5 + 0.3 \times 4 = 1.7$$

$$f(12) = 0.5 + 0.3 \times 4 = 1.7$$

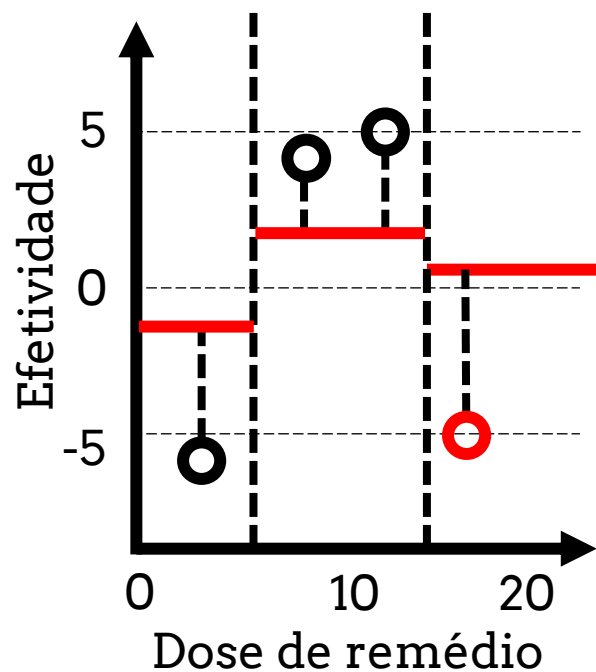
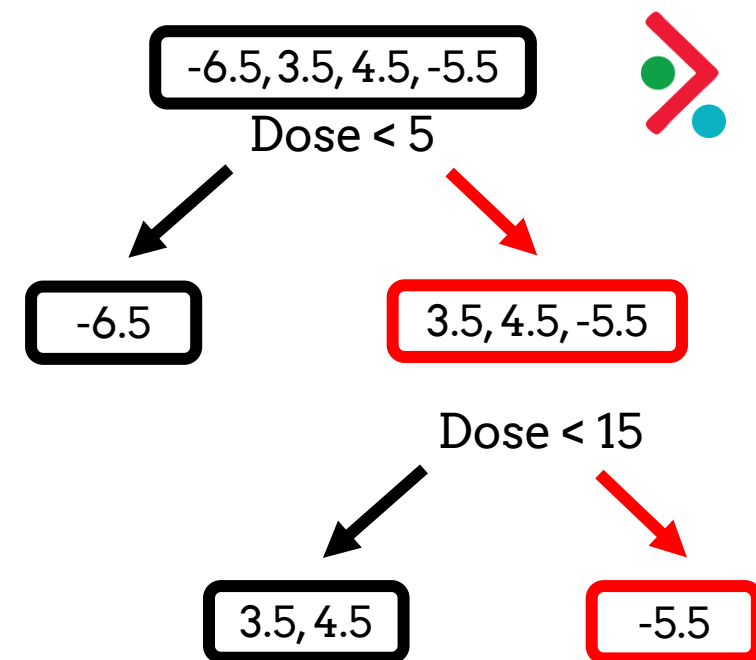
$$predição = \frac{3.5 + 4.5}{2 + 0} = 4$$



Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

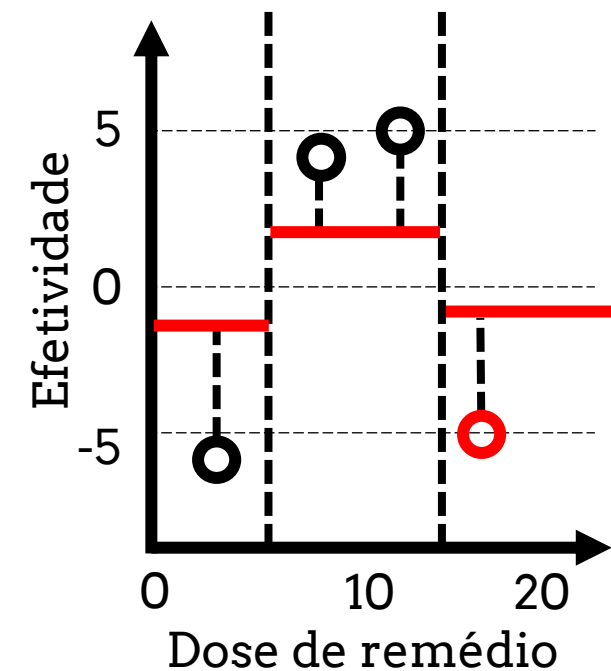
$$f(x) = 0.5 + 0.3 \times \begin{array}{c} \square \\ \square \quad \square \\ \square \end{array}$$

$$\begin{aligned} f(2) &= 0.5 + 0.3 \times -6.5 = -1.56 \\ f(8) &= 0.5 + 0.3 \times 4 = 1.7 \\ f(12) &= 0.5 + 0.3 \times 4 = 1.7 \\ f(16) &= 0.5 + 0.3 \times \end{aligned}$$





Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(x) = 0.5 + 0.3 \times \begin{matrix} \square \\ \square \quad \square \\ \square \end{matrix}$$

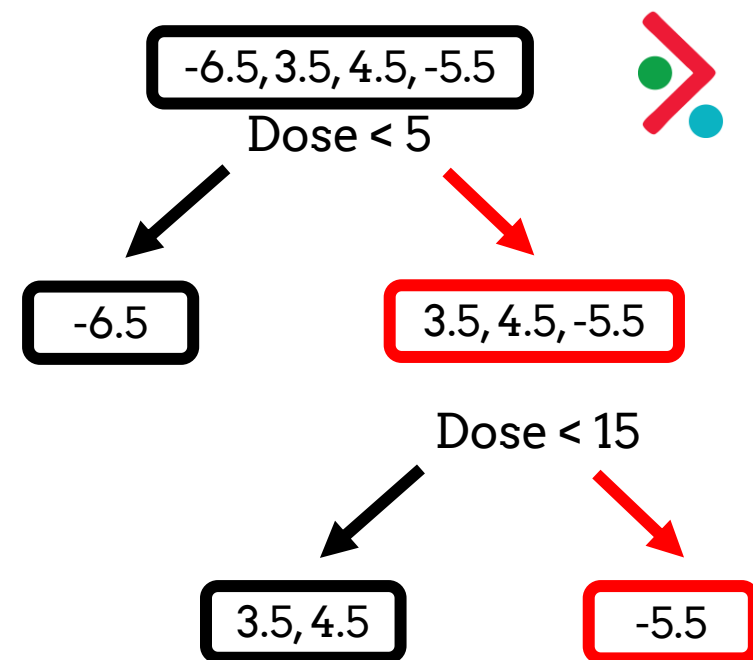
$$f(2) = 0.5 + 0.3 \times -6.5 = -1.56$$

$$f(8) = 0.5 + 0.3 \times 4 = 1.7$$

$$f(12) = 0.5 + 0.3 \times 4 = 1.7$$

$$f(16) = 0.5 + 0.3 \times -5.5 = -1.15$$

$$predição = \frac{-5.5}{1 + 0} = -5.5$$



Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(x) = 0.5 + 0.3 \times \text{[Diagram of a small neural network with 3 input nodes, 2 hidden nodes, and 1 output node]}$$



$$f(2) = 0.5 + 0.3 \times -6.5 = -1.56$$

$$f(8) = 0.5 + 0.3 \times 4 = 1.7$$

$$f(12) = 0.5 + 0.3 \times 4 = 1.7$$

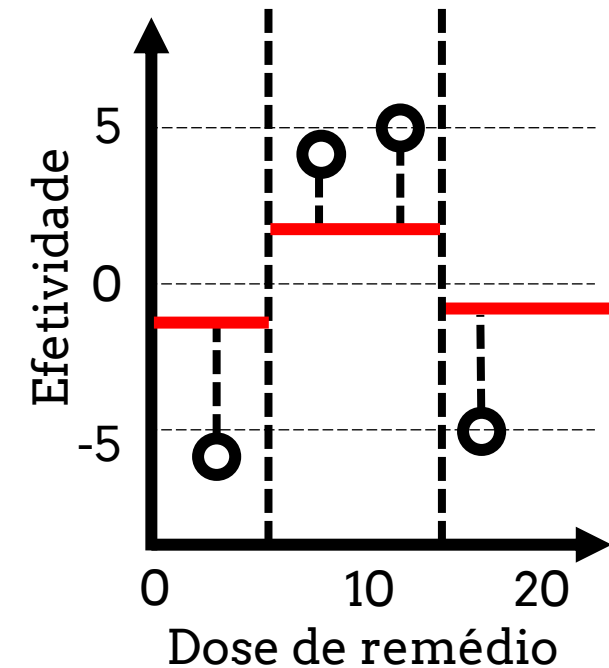
$$f(16) = 0.5 + 0.3 \times -5.5 = -1.15$$

$$\text{resíduo1} = -6 - (-1.56) = -4.44$$

$$\text{resíduo2} = 4 - 1.7 = 2.3$$

$$\text{resíduo3} = 5 - 1.7 = 3.3$$

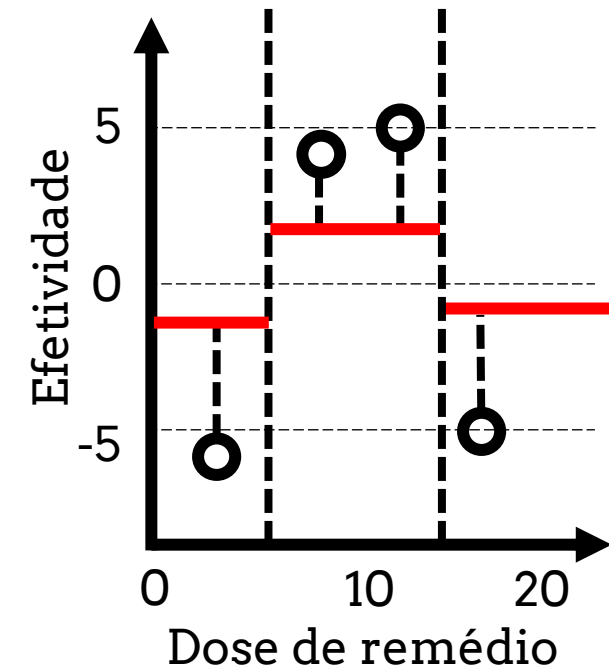
$$\text{resíduo4} = -5 - (-1.15) = -3.85$$



Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(\mathbf{x}) = 0.5 + 0.3 \times \text{[Diagram of a small decision tree with 5 nodes]}$$

-4.44, 2.3, 3.3, -3.85

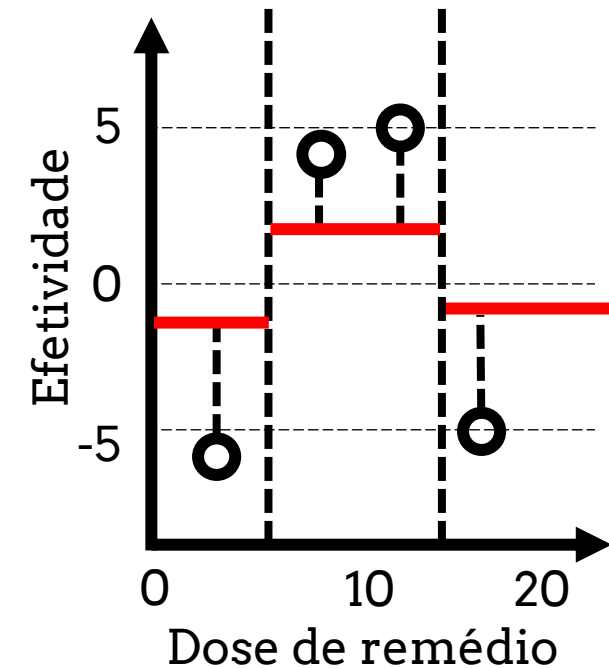


Hora da segunda árvore

Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2

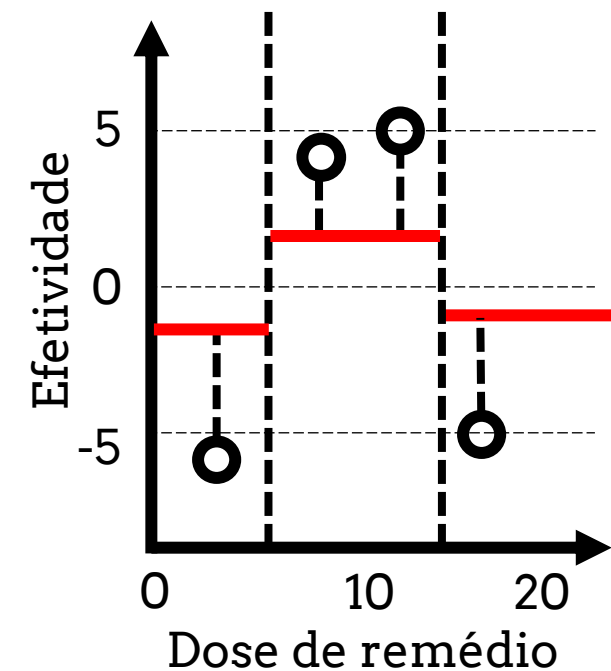
$$f(\mathbf{x}) = 0.5 + 0.3 \times \text{[diagram of a small neural network with 4 input nodes, 2 hidden nodes, and 1 output node]}$$

-4.44, 2.3, 3.3, -3.85

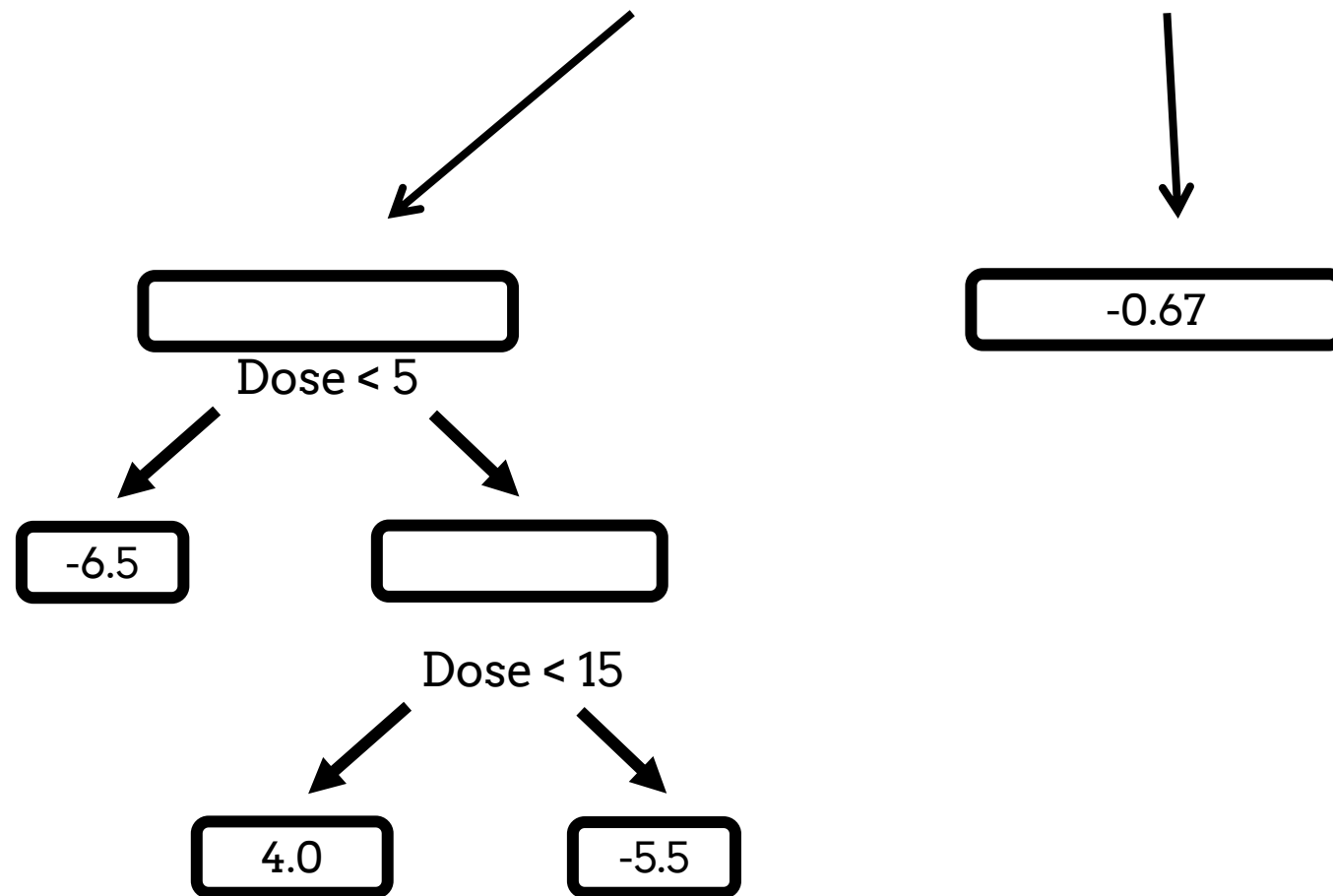


(sim salamin...)

Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(x) = 0.5 + 0.3 \times \text{[Tree 1]} + 0.3 \times \text{[Tree 2]}$$

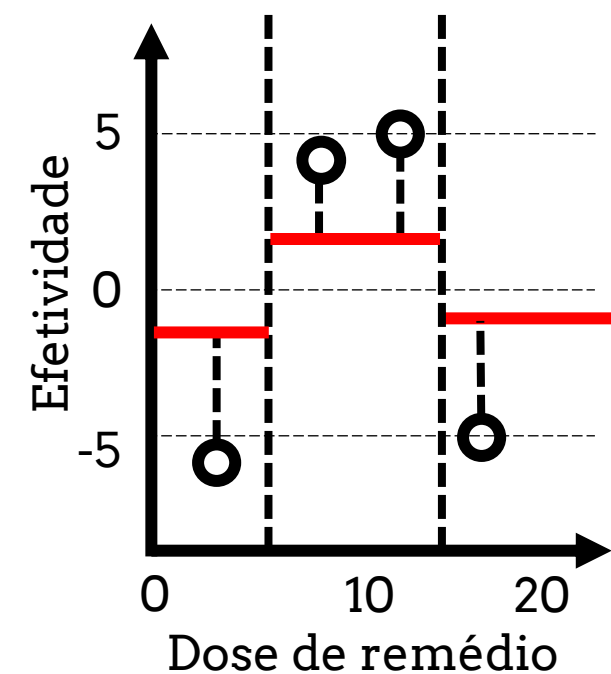


Modelo Final!





Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



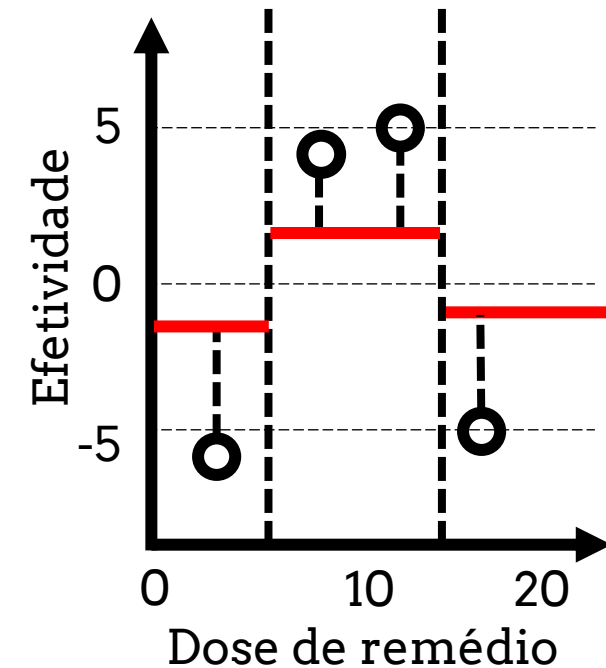
$$f1(x) = 0.5$$

$$f2(x) = 0.5 + 0.3 \times \text{Tree 1}$$

$$f3(x) = 0.5 + 0.3 \times \text{Tree 1} + 0.3 \times \text{Tree 2}$$



Hiperparam	valor
$\lambda$	0
$\gamma$	50
$\varepsilon$	0.3
Tree Depth	2
Trees	2



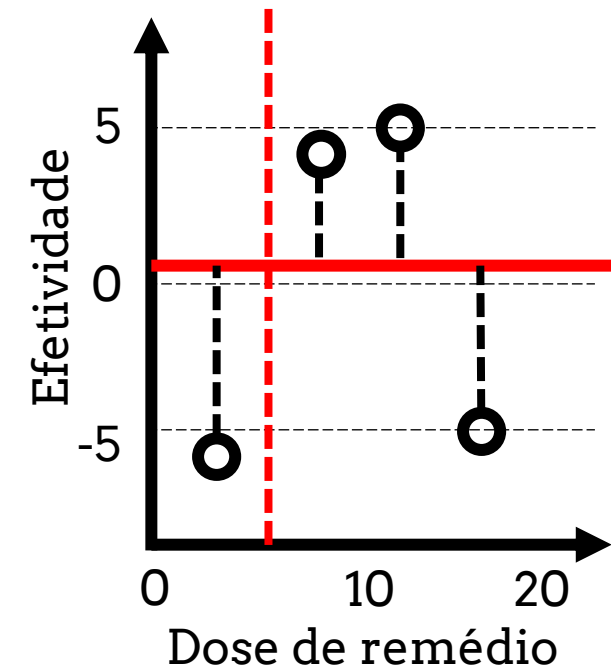
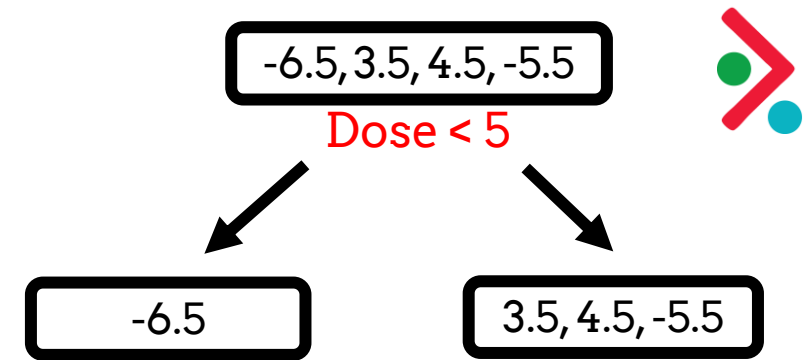
$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times \text{[Tree Diagram]}$$

$$f3(x) = f2(x) + 0.3 \times \text{[Tree Diagram]}$$

Hiperparam	valor
$\lambda$	1
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(\mathbf{x}) = 0.5$$



Pergunta	$\lambda=0$	$\lambda=1$
Dose < 5	40.33	
Dose < 10	1	
Dose < 15	27	

$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$



Hiperparam	valor
$\lambda$	1
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$f(\mathbf{x}) = 0.5$$

$$Similaridade_{pai} = \frac{(-6.5 + 3.5 + 4.5 - 5.5)^2}{4 + 1} = 3.2$$

$$Similaridade_{esq} = \frac{(-6.5)^2}{3 + 1} = 21.125$$

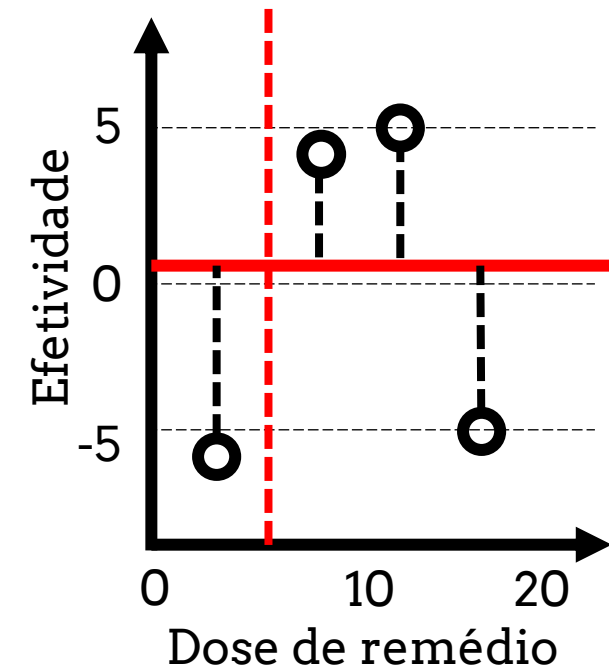
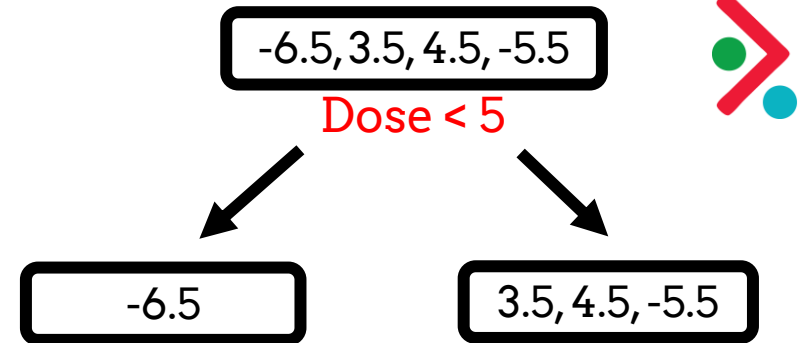
$$Similaridade_{dir} = \frac{(3.5 + 4.5 - 5.5)^2}{1 + 1} = 3.125$$

$$Gain = 3.125 + 21.125 - 3.2 = 21.05$$

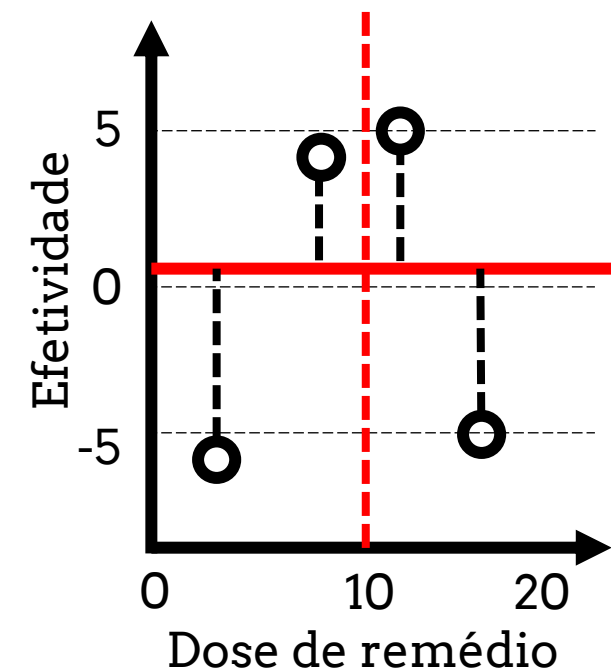
Pergunta	$\lambda=0$	$\lambda=1$
Dose < 5	40.33	21.05
Dose < 10	1	
Dose < 15	27	

$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

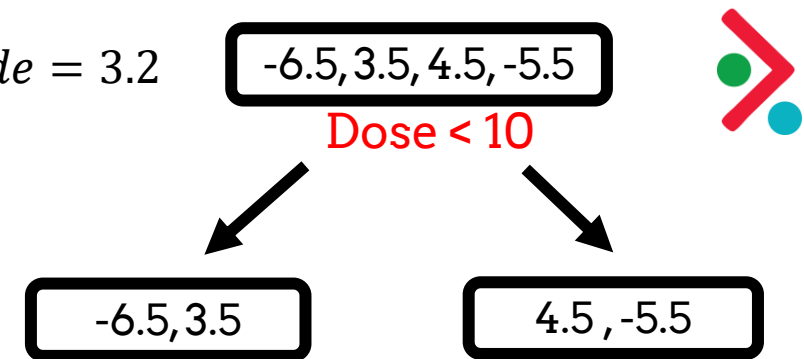


Hiperparam	valor
$\lambda$	1
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(\mathbf{x}) = 0.5$$

$$Similaridade = 3.2$$



$$Similaridade_{esq} = \frac{(-6.5 + 3.5)^2}{2 + 1} = 3$$

$$Similaridade_{dir} = \frac{(4.5 - 5.5)^2}{2 + 1} = 0.33$$

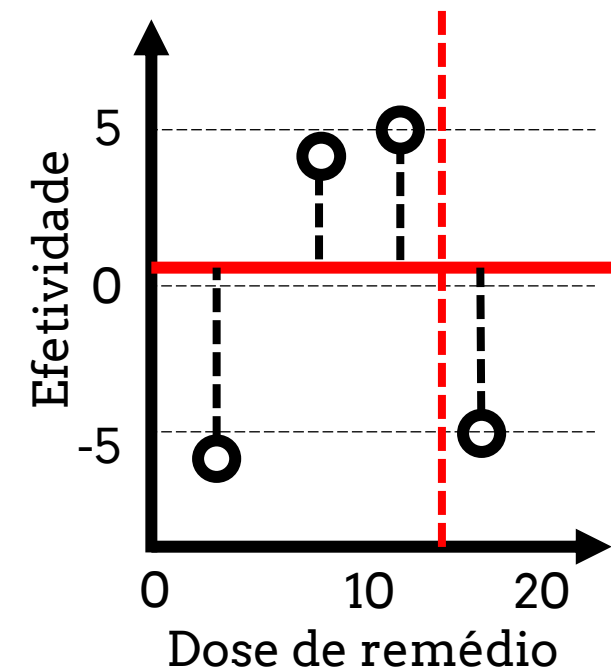
$$Gain = 3 + 0.33 - 3.2$$

Pergunta	$\lambda=0$	$\lambda=1$
Dose < 5	40.33	21.05
Dose < 10	1	0.13
Dose < 15	27	

$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

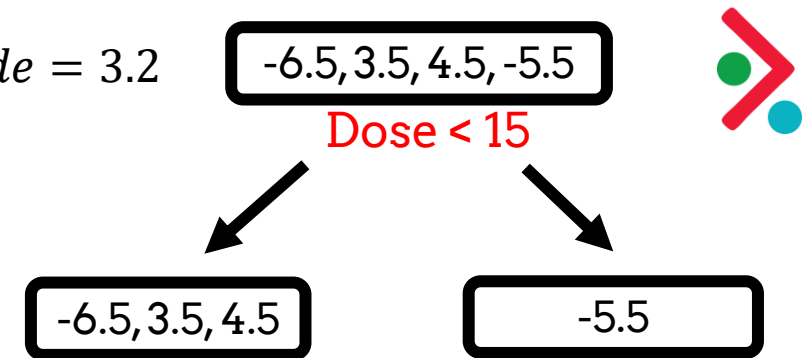
$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	1
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(\mathbf{x}) = 0.5$$

$$Similaridade = 3.2$$



$$Similaridade_{esq} = \frac{(-6.5 + 3.5 + 4.5)^2}{3 + 1} = 0.56$$

$$Similaridade_{dir} = \frac{(-5.5)^2}{1 + 1} = 15.12$$

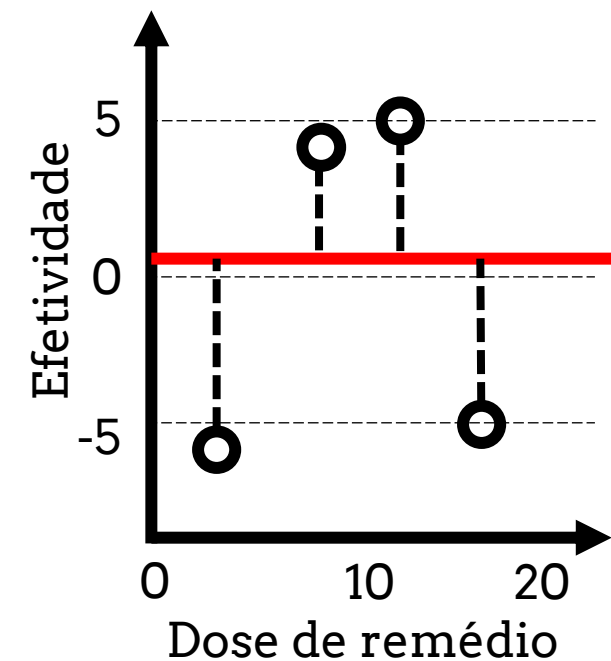
$$Gain = 15.12 + 0.56 - 3.2 = 12.48$$

Pergunta	$\lambda=0$	$\lambda=1$
Dose < 5	40.33	21.05
Dose < 10	1	0.13
Dose < 15	27	12.48

$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

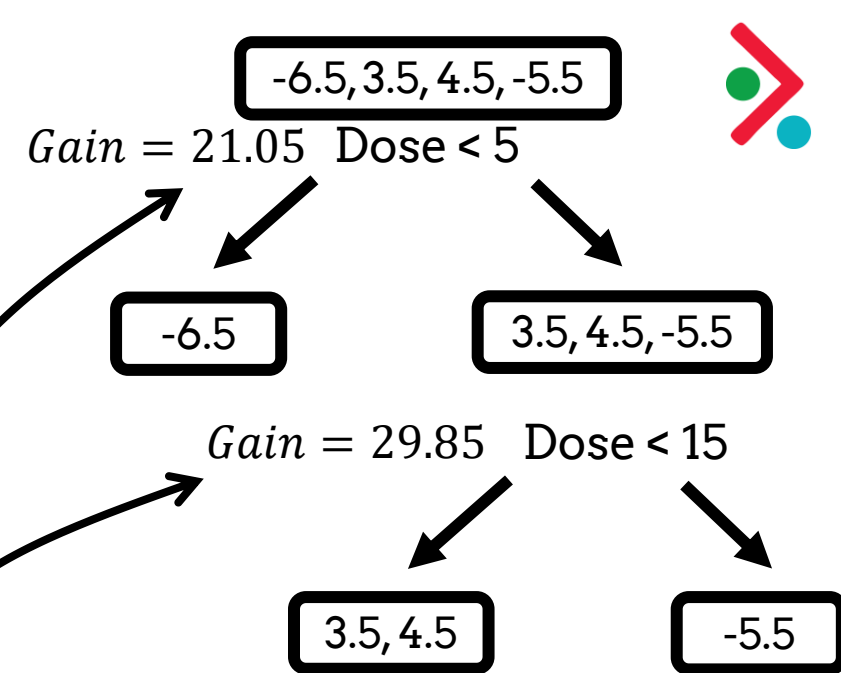
$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

Hiperparam	valor
$\lambda$	1
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(x) = 0.5 + 0.3 \times \text{[Tree Structure]}$$

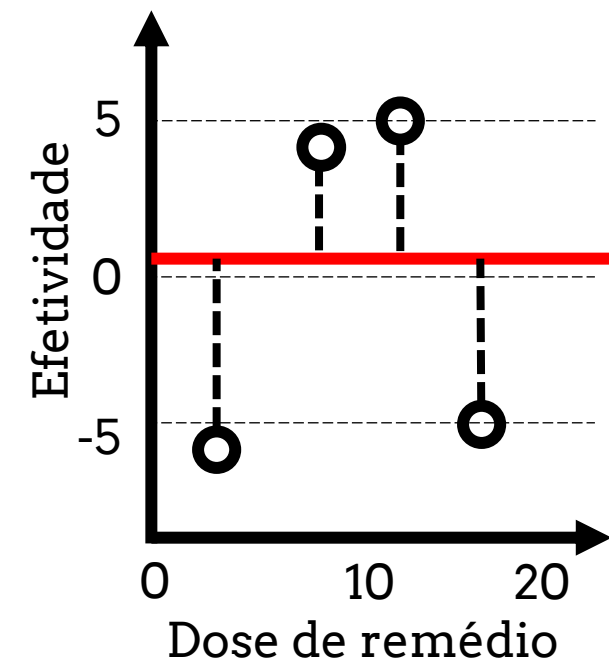
Gains menores,  
mais fáceis de podar!



$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

$$Similaridade = \frac{(\sum \text{resíduos})^2}{\# \text{resíduos} + \lambda}$$

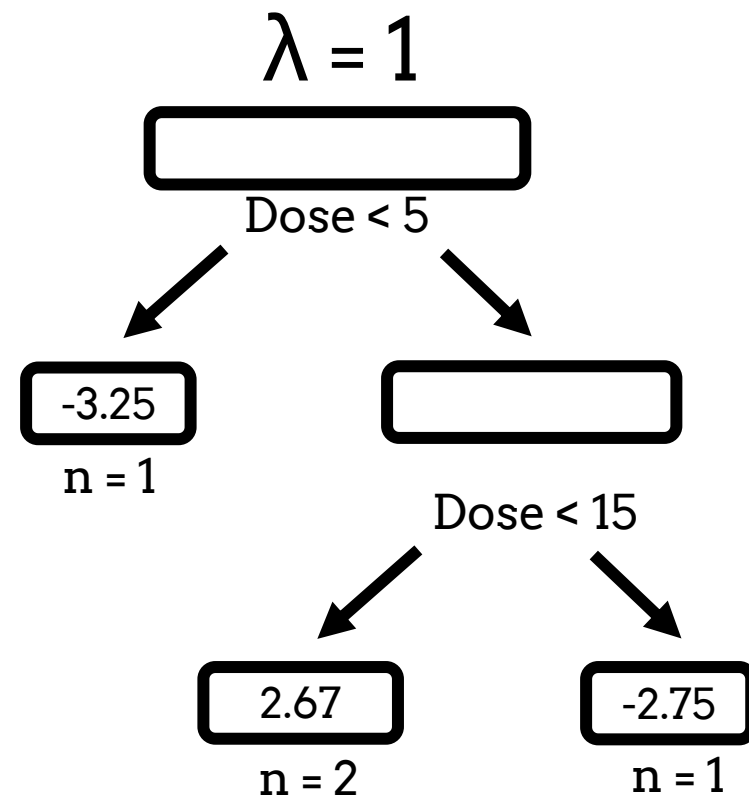
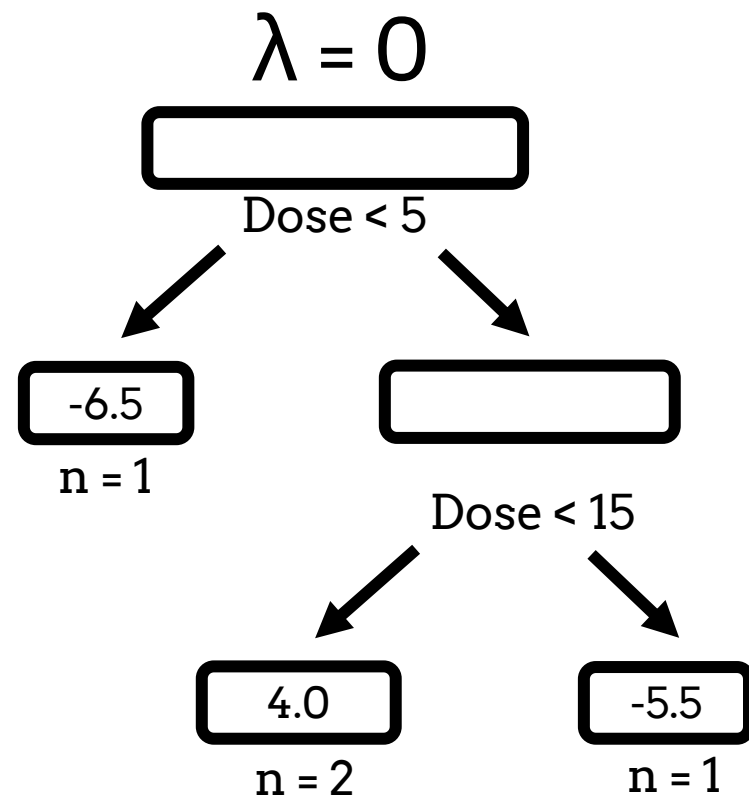
Hiperparam	valor
$\lambda$	1
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$f(x) = 0.5 + 0.3 \times \text{[tree icon]}$$



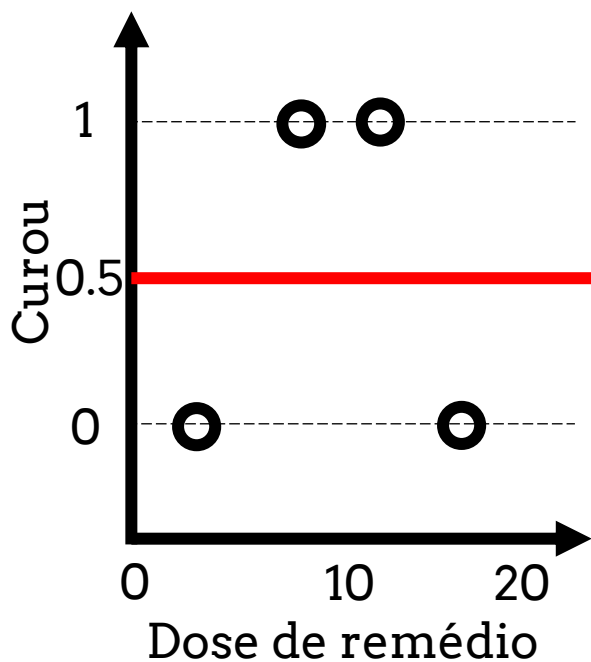
Além disso, os scores também diminuiriam...



$$predição = \frac{\sum resíduos}{\#resíduos + \lambda}$$



Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



Regressão

$$\frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

$$\frac{\sum resíduos}{\#resíduos + \lambda}$$

$$f(x) = 0.5 + \begin{matrix} \square & \square \\ \square & \square \end{matrix}$$

Classificação

$$\frac{(\sum resíduos)^2}{\sum p(1-p) + \lambda}$$

$$\frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

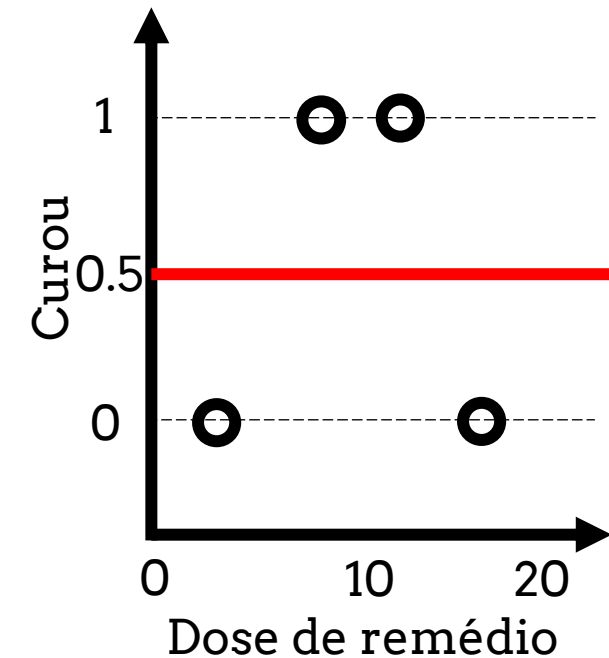
$$\log\left(\frac{f(x)}{1-f(x)}\right) = 0.0 + \begin{matrix} \square & \square \\ \square & \square \end{matrix}$$



Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

No caso de classificação, vamos trocar  $f()$  por  $p()$  para relacionar com o fato de que estamos calculando probabilidades.

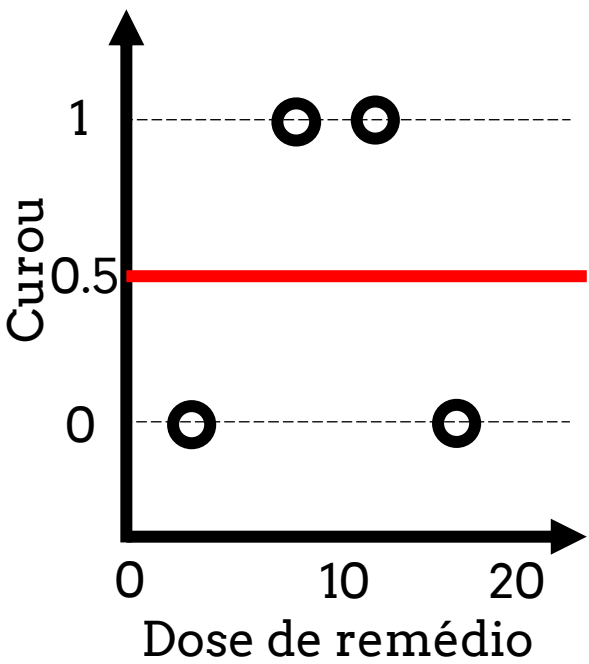




Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

No caso de classificação, vamos trocar  $f()$  por  $p()$  para relacionar com o fato de que estamos calculando probabilidades.



E uma rápida revisão sobre a função logística:

$$\log\left(\frac{p(x)}{1-p(x)}\right) = x$$



Logaritmo da chance,  
ou log-odds,  
ou logit

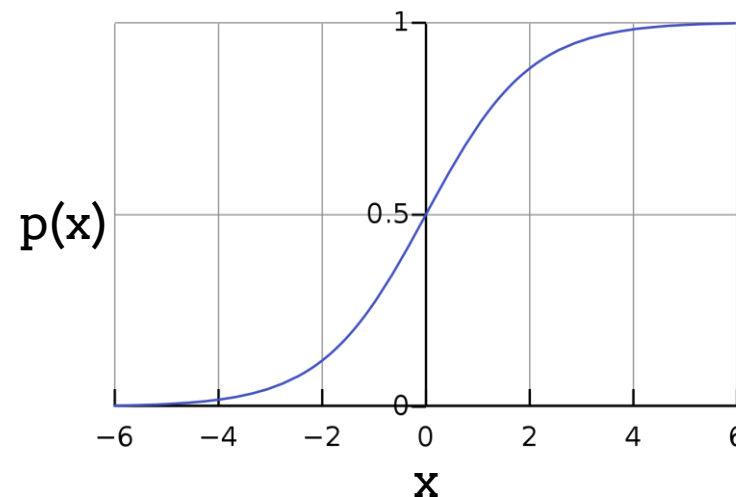




Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

No caso de classificação, vamos trocar  $f()$  por  $p()$  para relacionar com o fato de que estamos calculando probabilidades.



E uma rápida revisão sobre a função logística:

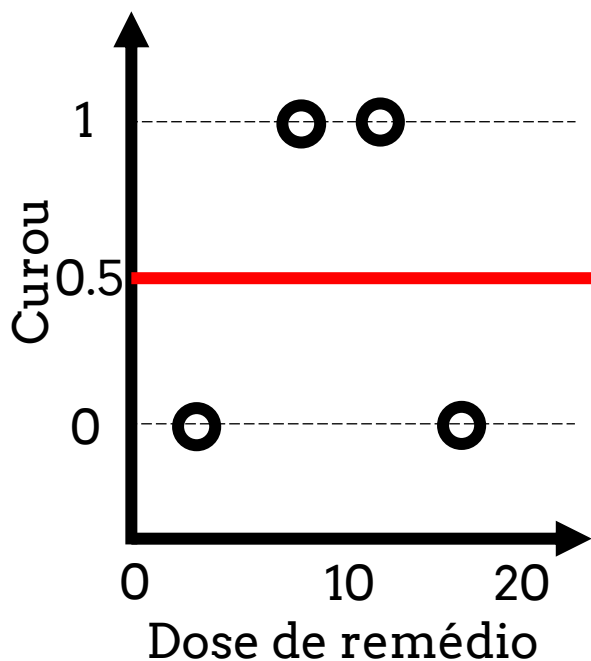
$$\log\left(\frac{p(x)}{1-p(x)}\right) = x$$

inversa

$$p(x) = \frac{1}{1 + e^{-x}}$$

Logaritmo da chance,  
ou log-odds,  
ou logit

Função logística,  
ou sigmoide

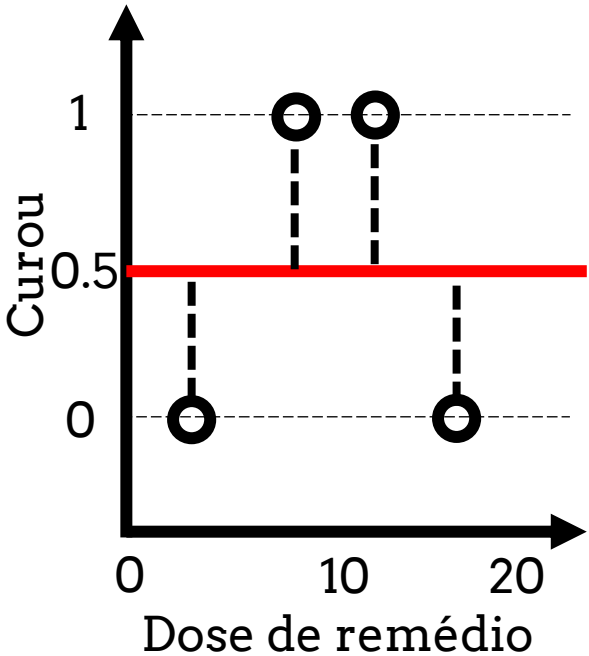


Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2

$$\log\left(\frac{p(x)}{1 - p(x)}\right) = 0.0$$



$$\text{res\u00edduo} = y - p(x)$$

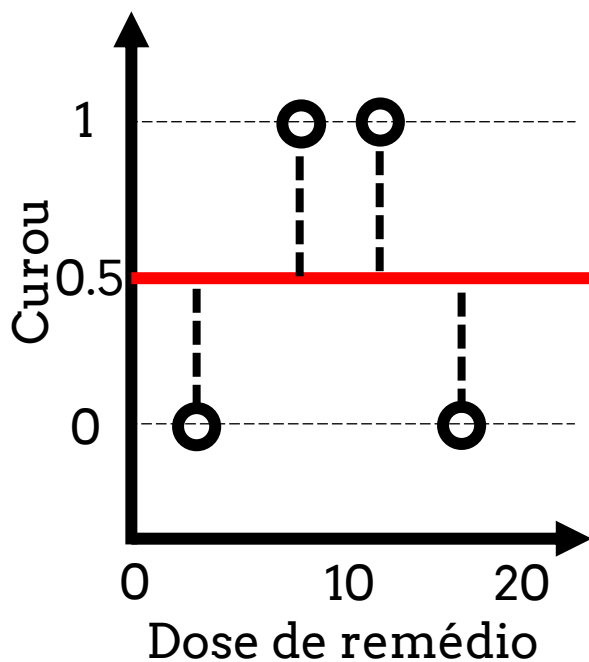


$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2

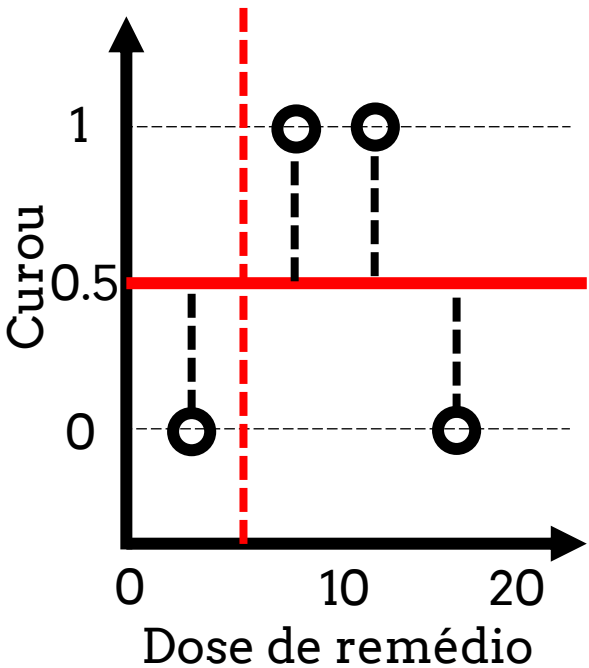
$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

-0.5,0.5,0.5,-0.5

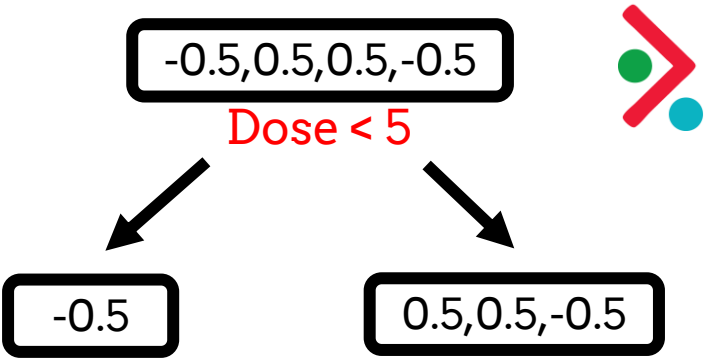


$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$\log\left(\frac{p(x)}{1 - p(x)}\right) = 0.0$$



$$Similaridade_{pai} = \rule{10cm}{0.4pt} =$$

$$Similaridade_{esq} = \rule{10cm}{0.4pt} =$$

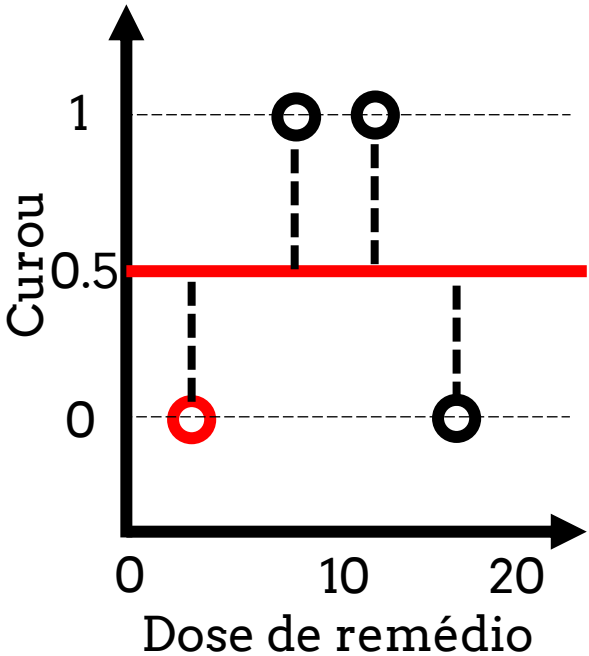
$$Similaridade_{dir} = \rule{10cm}{0.4pt} =$$

$$Gain =$$

$$Similaridade = \frac{(\sum res\acute{i}duos)^2}{\sum p(1 - p) + \lambda}$$

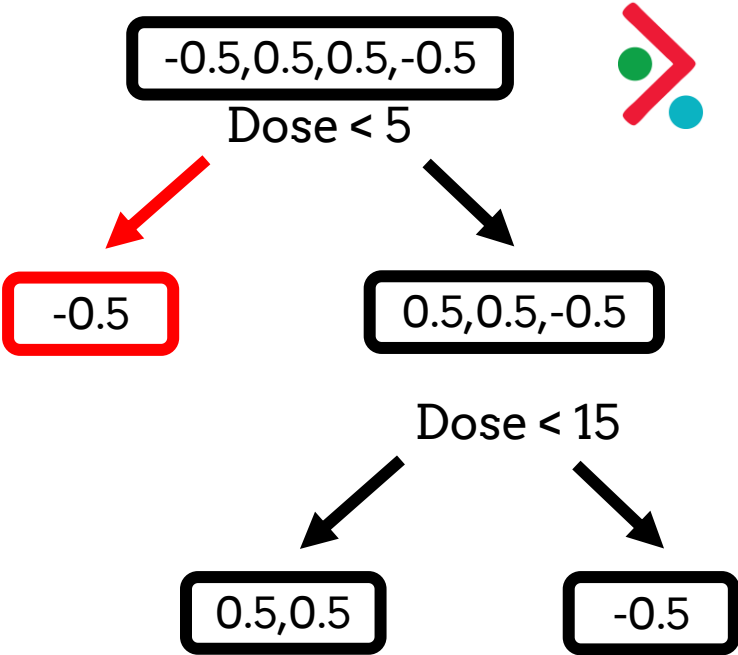
$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$\log\left(\frac{p(x)}{1 - p(x)}\right) = 0.0$$

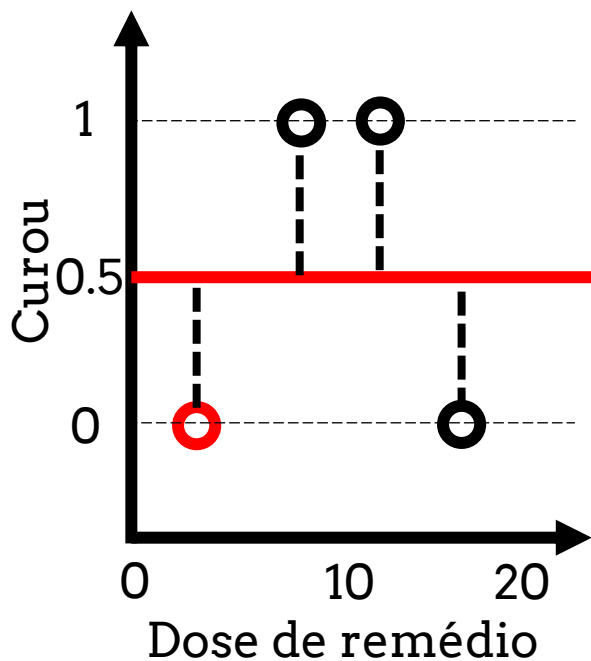
$$\log\left(\frac{p(2)}{1 - p(2)}\right) = 0.0$$



$$predi\c{c}\tilde{a}o = \frac{\sum res\acute{i}duos}{\sum p(1 - p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



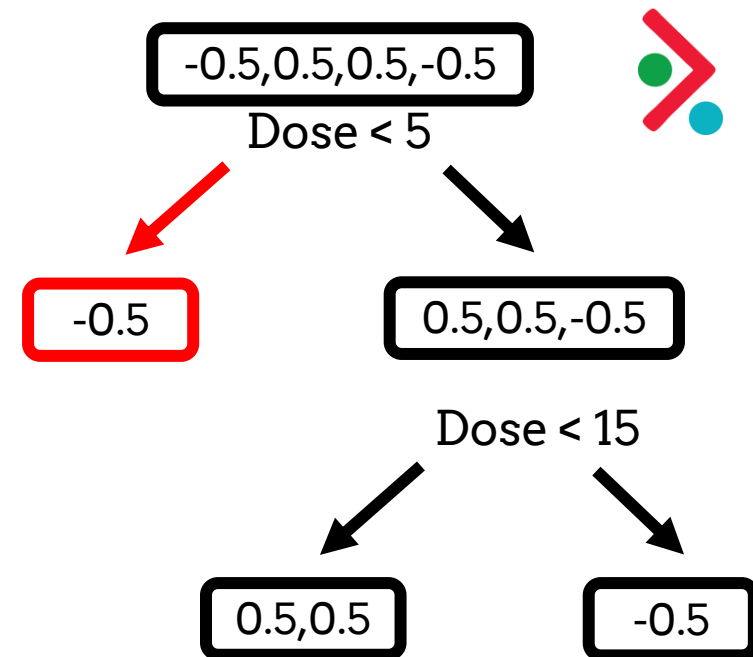
$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

$$\log\left(\frac{p(2)}{1-p(2)}\right) = 0.0 + 0.3 \times (-2)$$

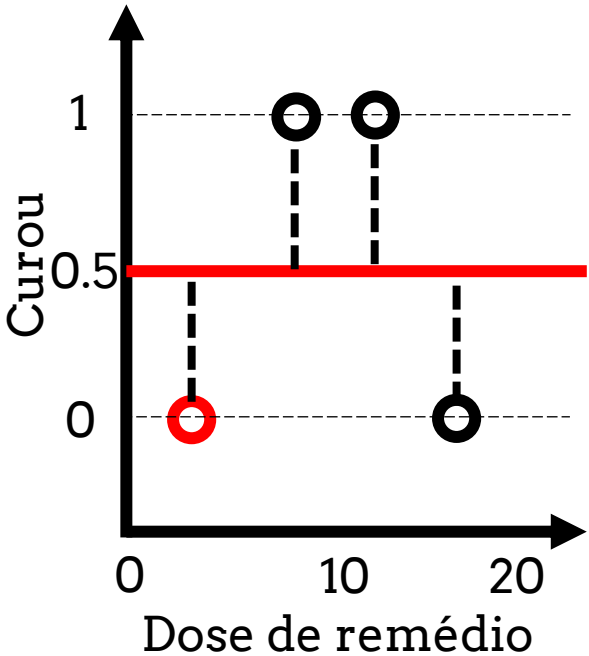
$$\frac{-0.5}{0.5(1-0.5) + 0} = -2$$

$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$

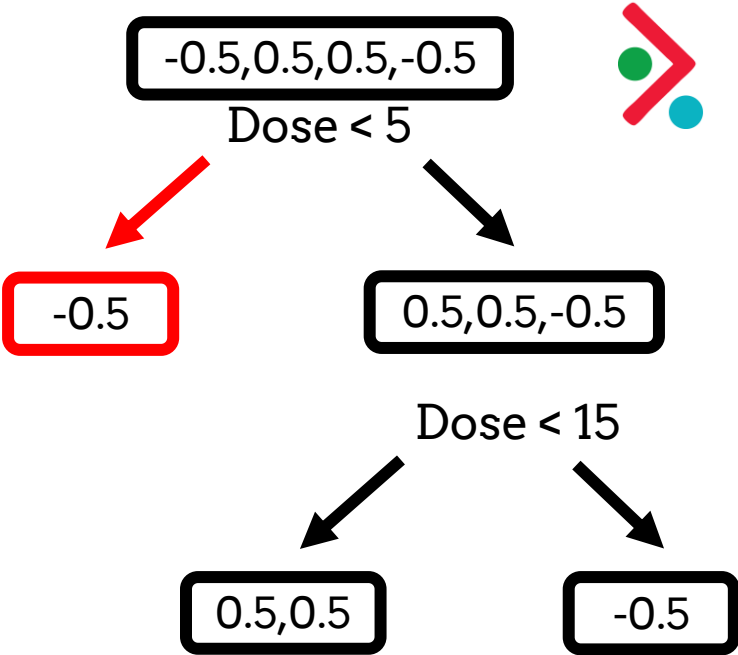


Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$\log\left(\frac{p(x)}{1 - p(x)}\right) = 0.0$$

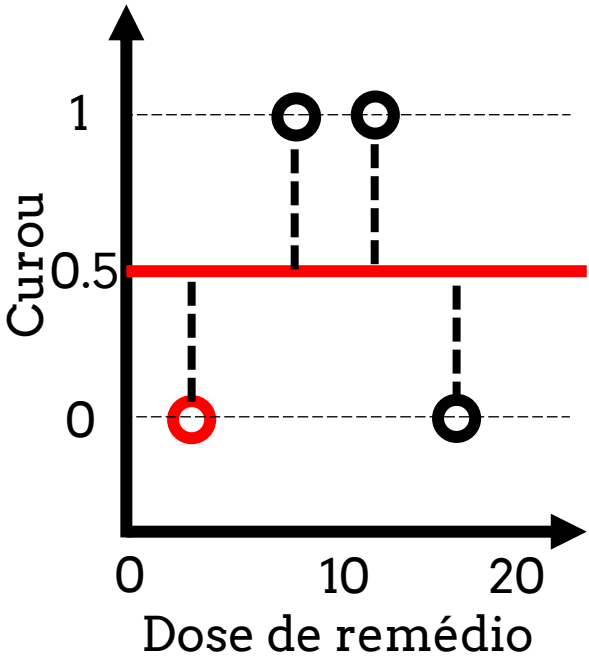
$$\log\left(\frac{p(2)}{1 - p(2)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$



$$predi\c{c}\tilde{a}o = \frac{\sum res\acute{i}duos}{\sum p(1 - p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



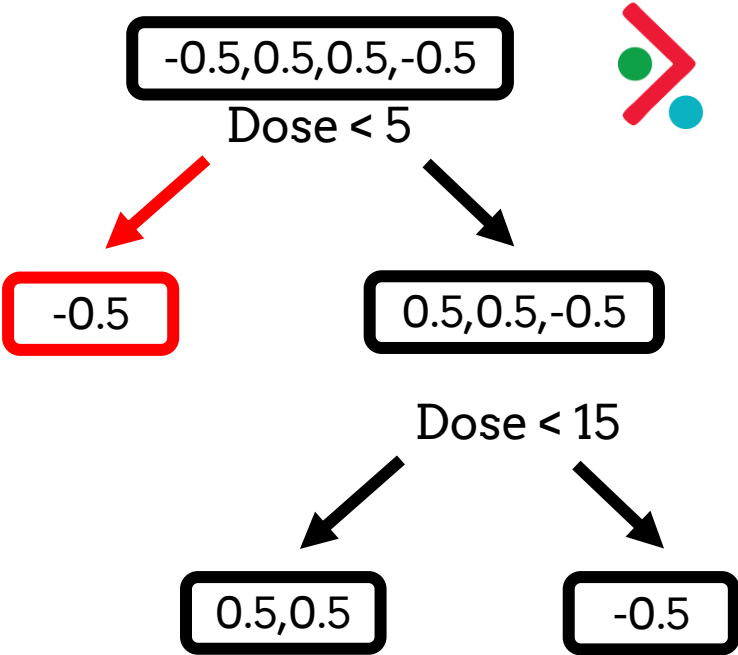
$$\log\left(\frac{p(x)}{1 - p(x)}\right) = 0.0$$

$$\log\left(\frac{p(2)}{1 - p(2)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$

$$p(2) = \frac{1}{1 + e^{(-0.6)}} = 0.354$$

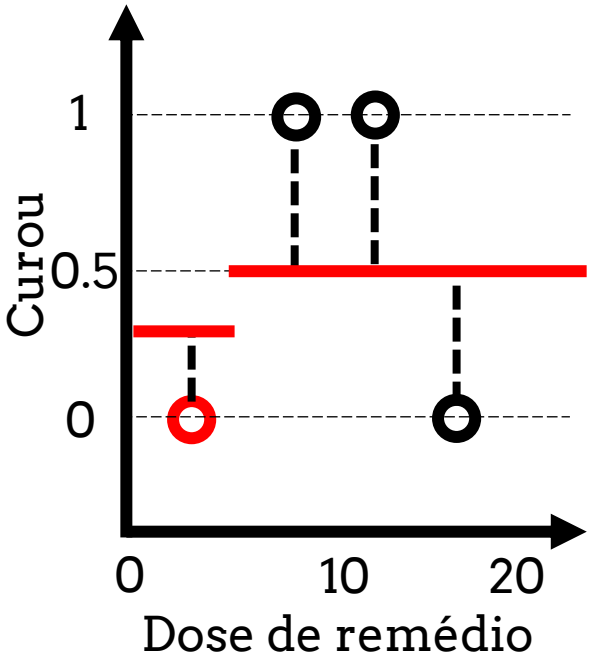
$$predi\c{c}ao = \frac{\sum res\acute{i}duos}{\sum p(1 - p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$





Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



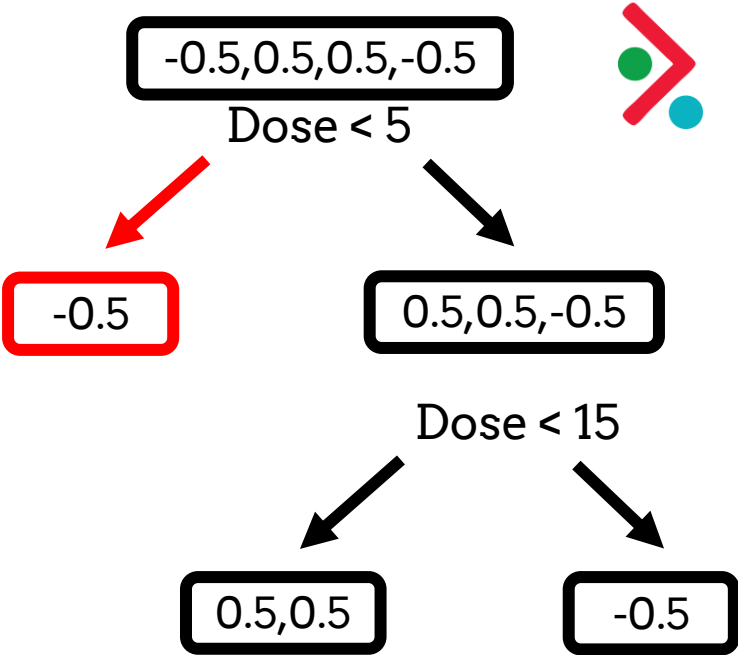
$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

$$\log\left(\frac{p(2)}{1-p(2)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$

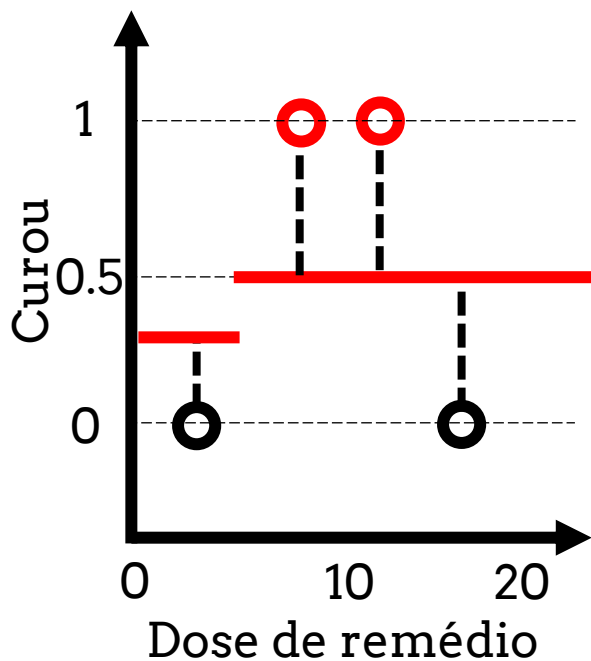
$$p(2) = \frac{1}{1 + e^{(-0.6)}} = 0.354$$

$$predi\c{c}a\tilde{o} = \frac{\sum res\acute{i}duos}{\sum p(1-p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$



Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

$$\log\left(\frac{p(2)}{1-p(2)}\right) = 0.0 + 0.3 \times (-2)$$

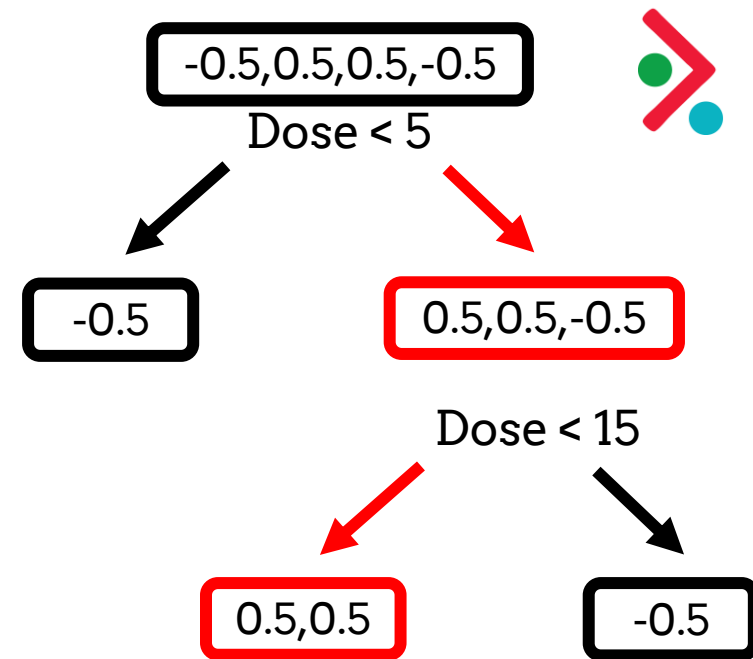
$$\log\left(\frac{p(8)}{1-p(8)}\right) = 0.0 + 0.3 \times 2$$

$$\log\left(\frac{p(12)}{1-p(12)}\right) = 0.0 + 0.3 \times 2$$

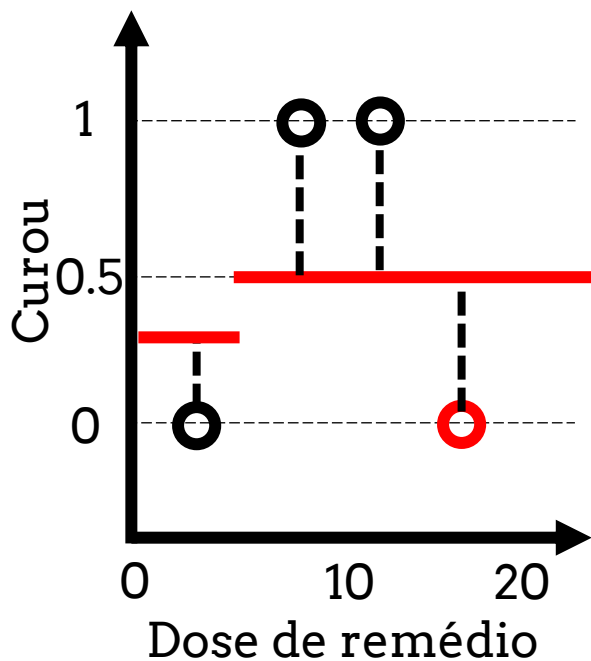
$$\frac{0.5 + 0.5}{0.5(1 - 0.5) + 0.5(1 - 0.5) + 0} = 2$$

$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$



Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

$$\log\left(\frac{p(2)}{1-p(2)}\right) = 0.0 + 0.3 \times (-2)$$

$$\log\left(\frac{p(8)}{1-p(8)}\right) = 0.0 + 0.3 \times 2$$

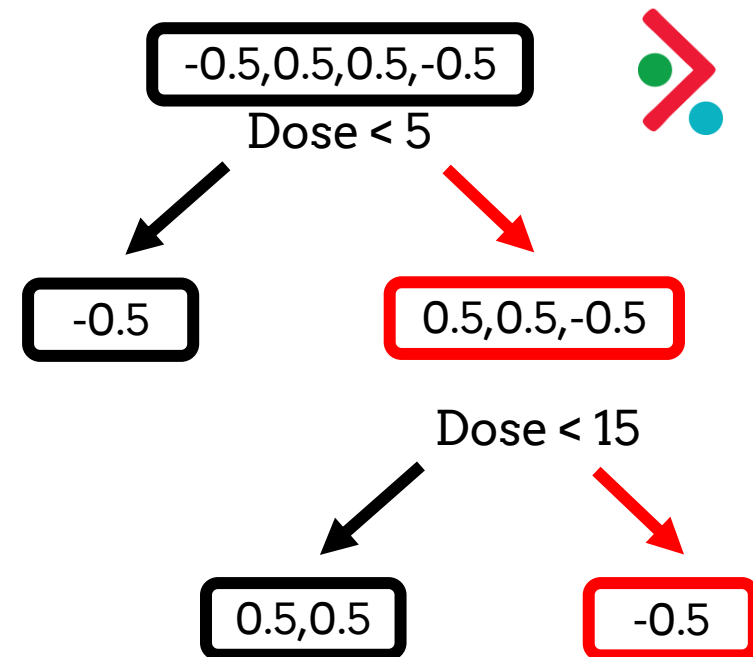
$$\log\left(\frac{p(12)}{1-p(12)}\right) = 0.0 + 0.3 \times 2$$

$$\log\left(\frac{p(16)}{1-p(16)}\right) = 0.0 + 0.3 \times (-2)$$

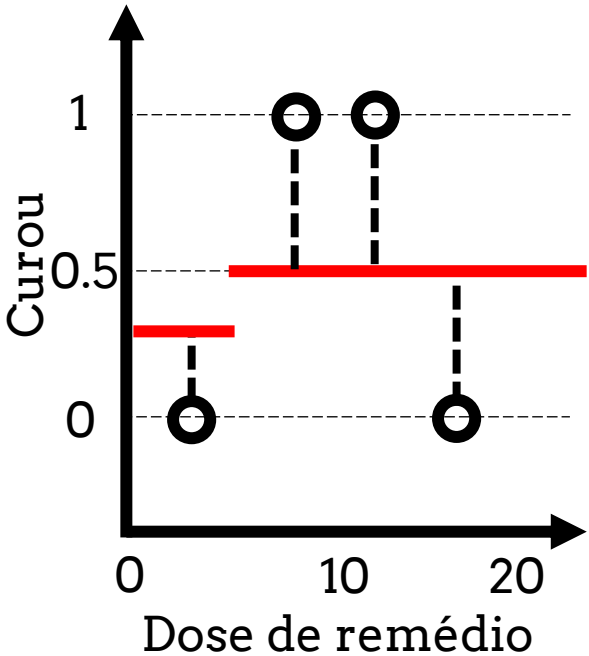
$$\frac{-0.5}{0.5(1-0.5) + 0} = -2$$

$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$



Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



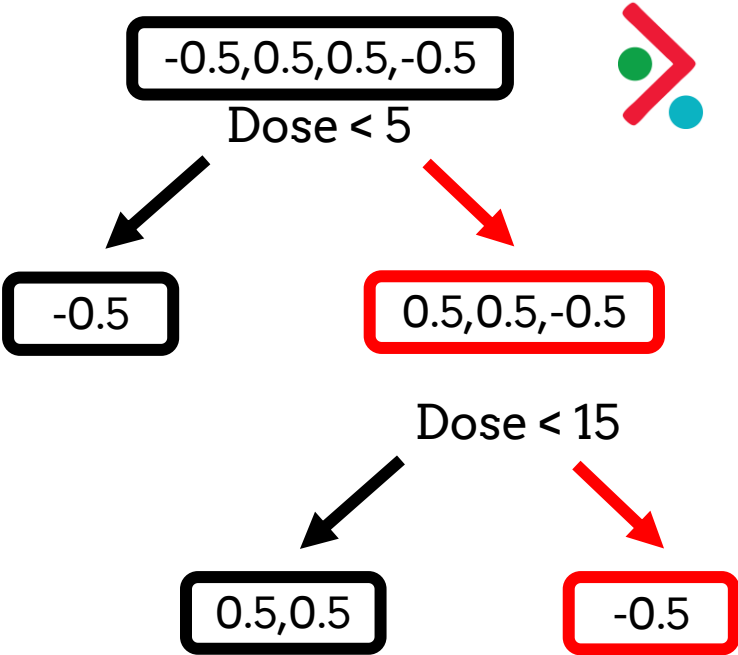
$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0$$

$$\log\left(\frac{p(2)}{1-p(2)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$

$$\log\left(\frac{p(8)}{1-p(8)}\right) = 0.0 + 0.3 \times 2 = 0.6$$

$$\log\left(\frac{p(12)}{1-p(12)}\right) = 0.0 + 0.3 \times 2 = 0.6$$

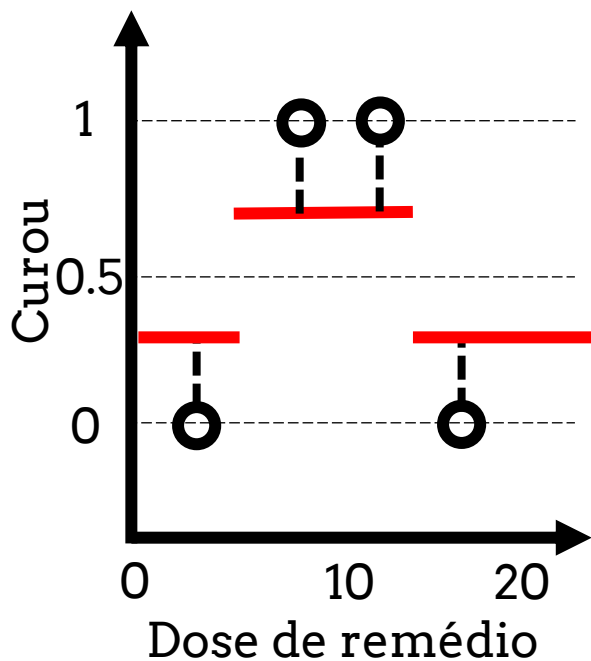
$$\log\left(\frac{p(16)}{1-p(16)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$



$$predi\c{c}a\tilde{o} = \frac{\sum res\acute{i}duos}{\sum p(1-p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
$\lambda$	0
$\gamma$	20
$\varepsilon$	0.3
Tree Depth	2
Trees	2



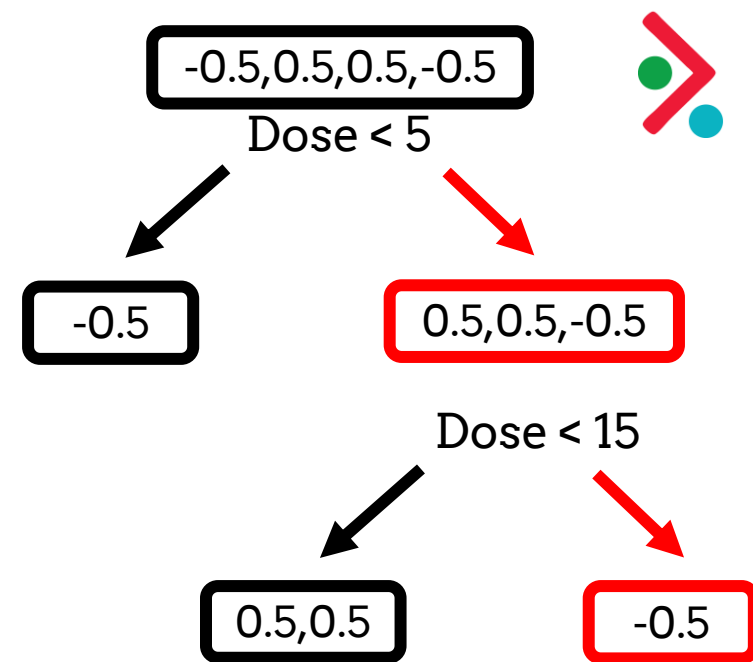
$$\log\left(\frac{p(x)}{1-p(x)}\right) = 0.0 + 0.3 \times \text{[Diagram of a small tree structure]}$$

$$p(2) = \frac{1}{1 + e^{(-0.6)}} = 0.35$$

$$p(8) = \frac{1}{1 + e^{(0.6)}} = 0.65$$

$$p(12) = \frac{1}{1 + e^{(0.6)}} = 0.65$$

$$p(16) = \frac{1}{1 + e^{(-0.6)}} = 0.35$$



$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

$$p(x) = \frac{1}{1 + e^{-x}}$$

