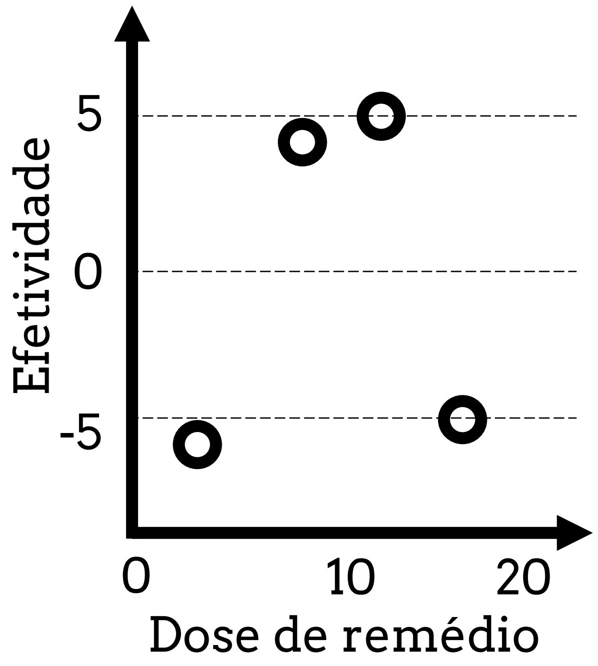
Dose de remédio	Efetividade
2	-6
8	4
12	5
16	-5

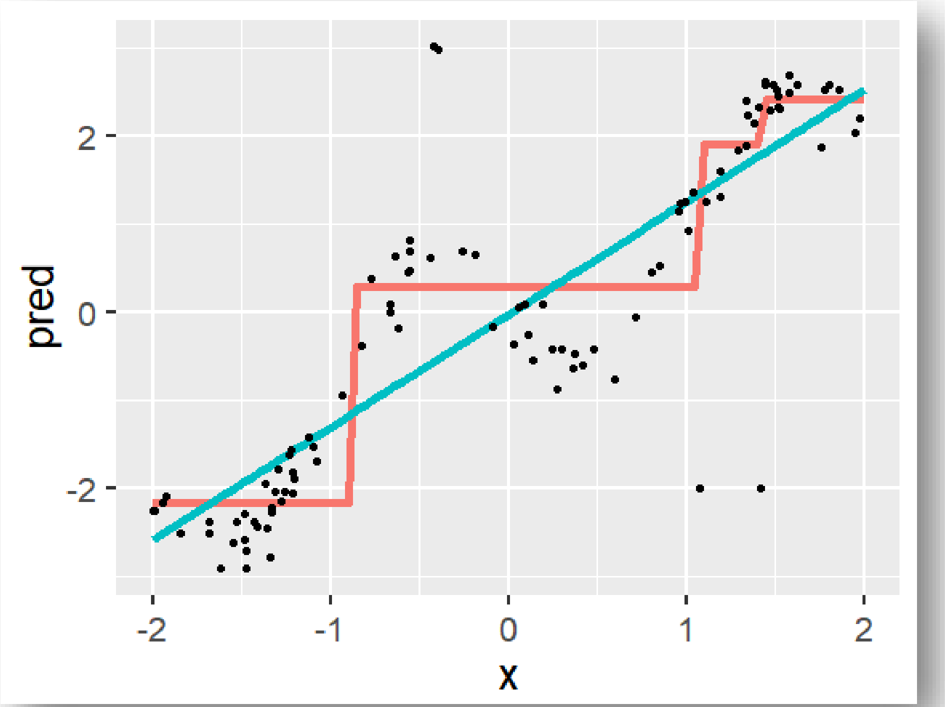
→

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Dose de remédio	Efetividade
2	-6
8	4
12	5
16	-5

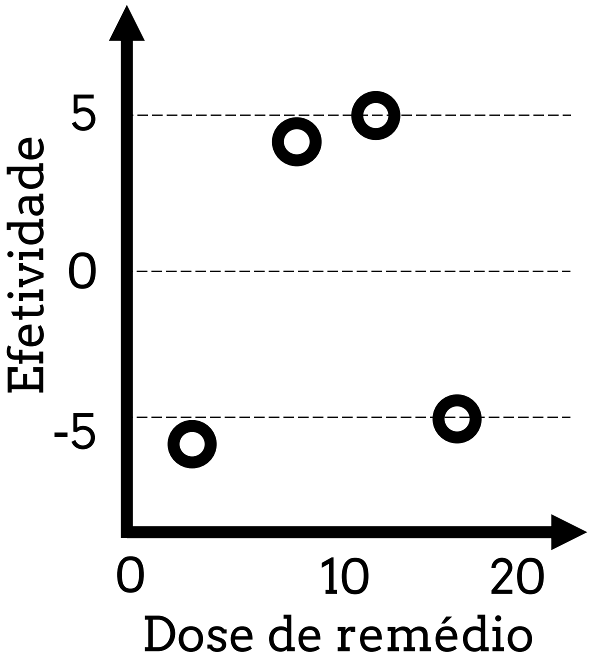


Y

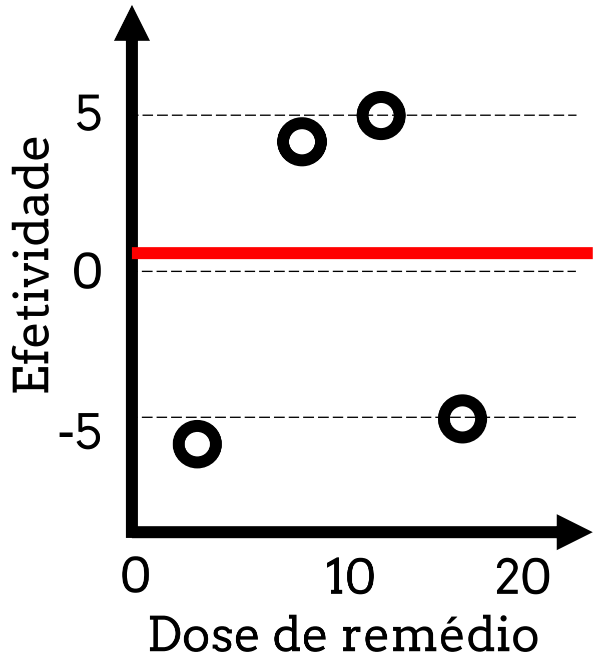


$$Y \approx f(X)$$

Hiperparam	valor
λ	
γ	
ε	
Tree Depth	
Trees	

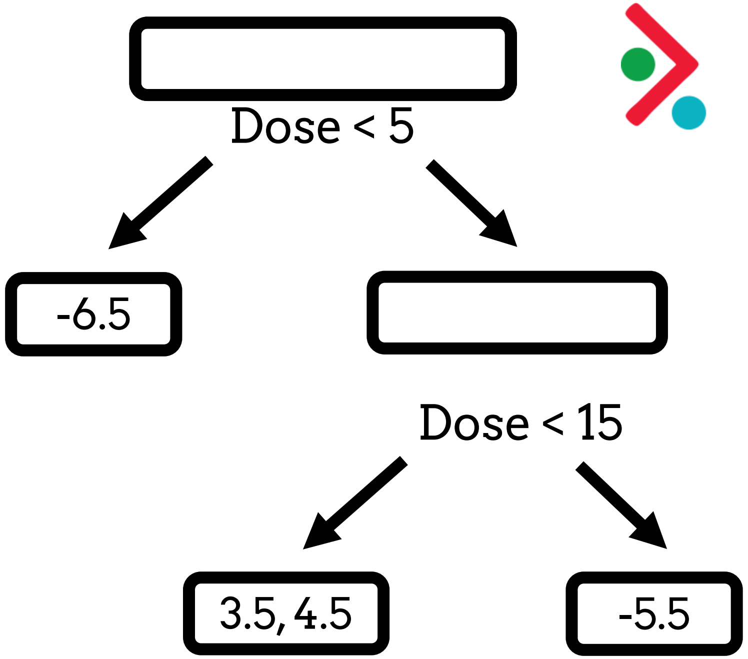


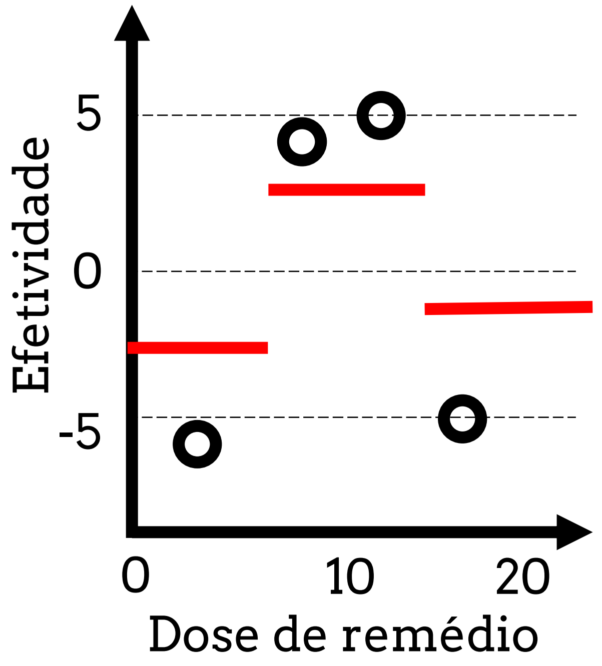
Hiperparam	valor
λ	0
γ	0
ε	0.3
Tree Depth	2
Trees	2



$$f(x) = 0.5$$

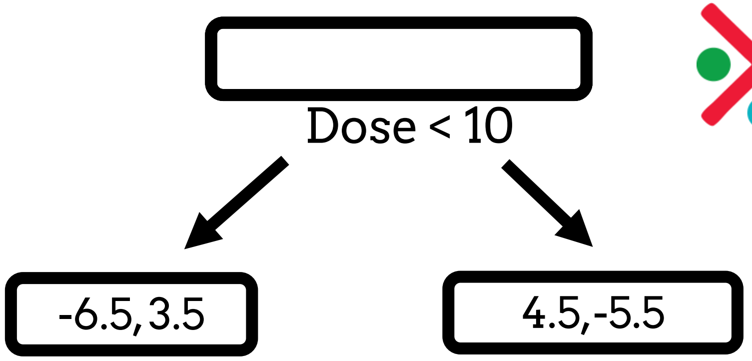
Hiperparam	valor
λ	0
γ	0
ε	0.3
Tree Depth	2
Trees	2

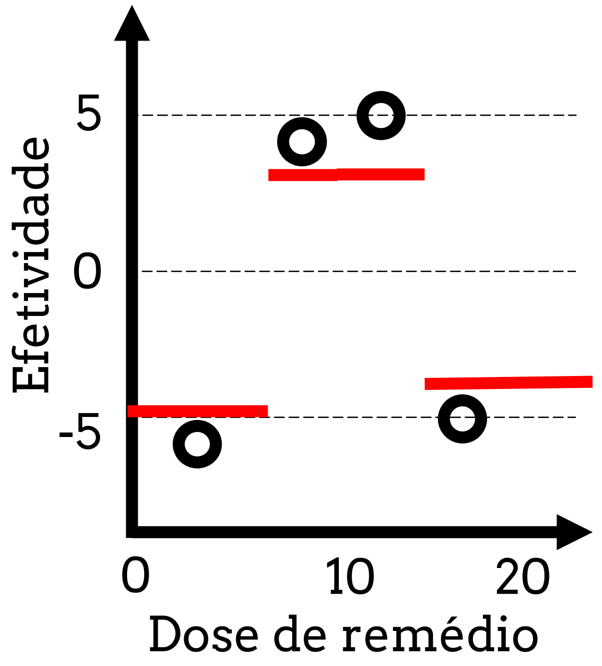


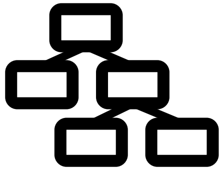
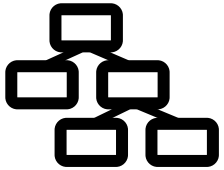


$$f(x) = 0.5 + \varepsilon \times$$

Hiperparam	valor
λ	0
γ	0
ε	0.3
Tree Depth	2
Trees	2





$f(x) = 0.5 + \varepsilon \times$  $+ \varepsilon \times$ 

Exemplo 1: 01-exemplo-hiperparemtro-sql.R

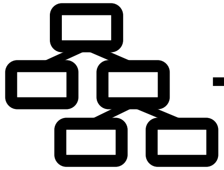
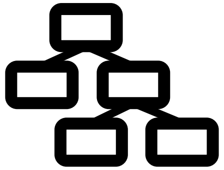
Hiperparam	valor
λ (regularization)	0
γ (loss_reduction)	0
ε (learn_rate)	0.3
tree_depth	2
trees	2

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

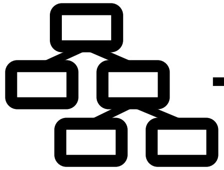
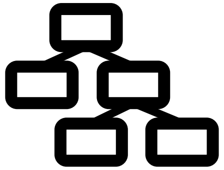
Hora da Função De Custo
“loss function”
“objective”

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

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

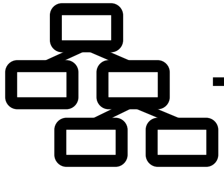
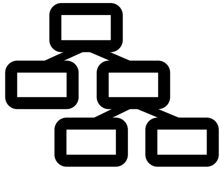
$$\sum L(y_i, f(x_i))$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

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

$$\sum L(y_i, f(x_i))$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

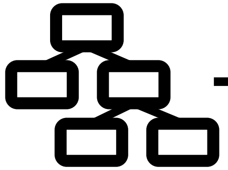
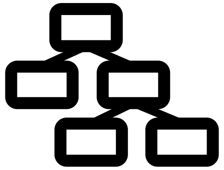
$f(x) = 0.5 + 0.3 \times$  $+ 0.3 \times$ 

$$\sum L(y_i, f(x_i))$$

$$\sum (y_i - f(x_i))^2$$

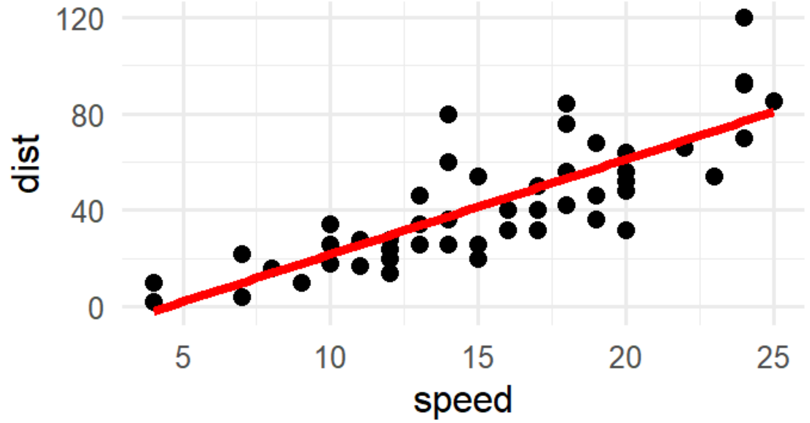
RMSE
Regressão Normal
Mínimos quadrados

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

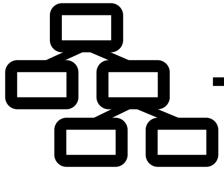
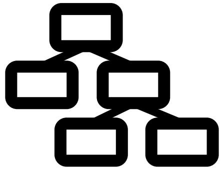
$f(x) = 0.5 + 0.3 \times$  $+ 0.3 \times$ 

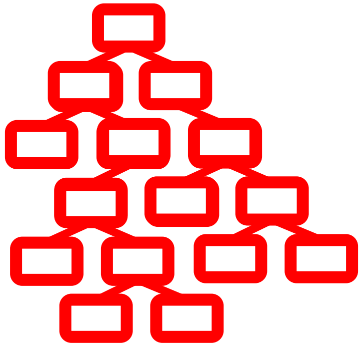
$$\sum L(y_i, f(x_i))$$

$$\sum (y_i - (\beta_o + \beta_1 x_i))^2$$



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

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

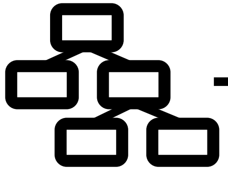
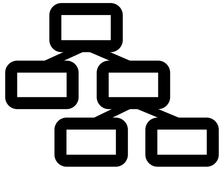
$\sum (y_i - ($  $))^2$

$$\sum L(y_i, f(x_i))$$

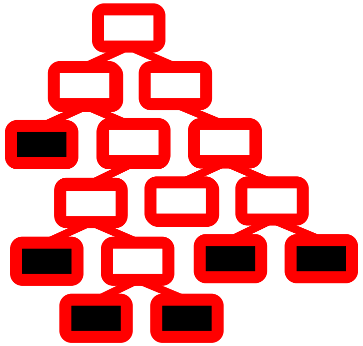
UMA árvore de decisão

?

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

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

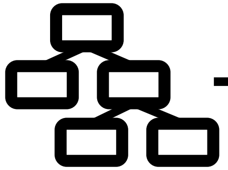
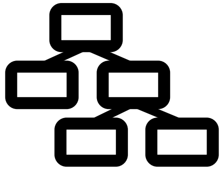
$$\sum L(y_i, f(x_i))$$

$\sum (y_i - ($  $))^2$

UMA árvore de decisão

?

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

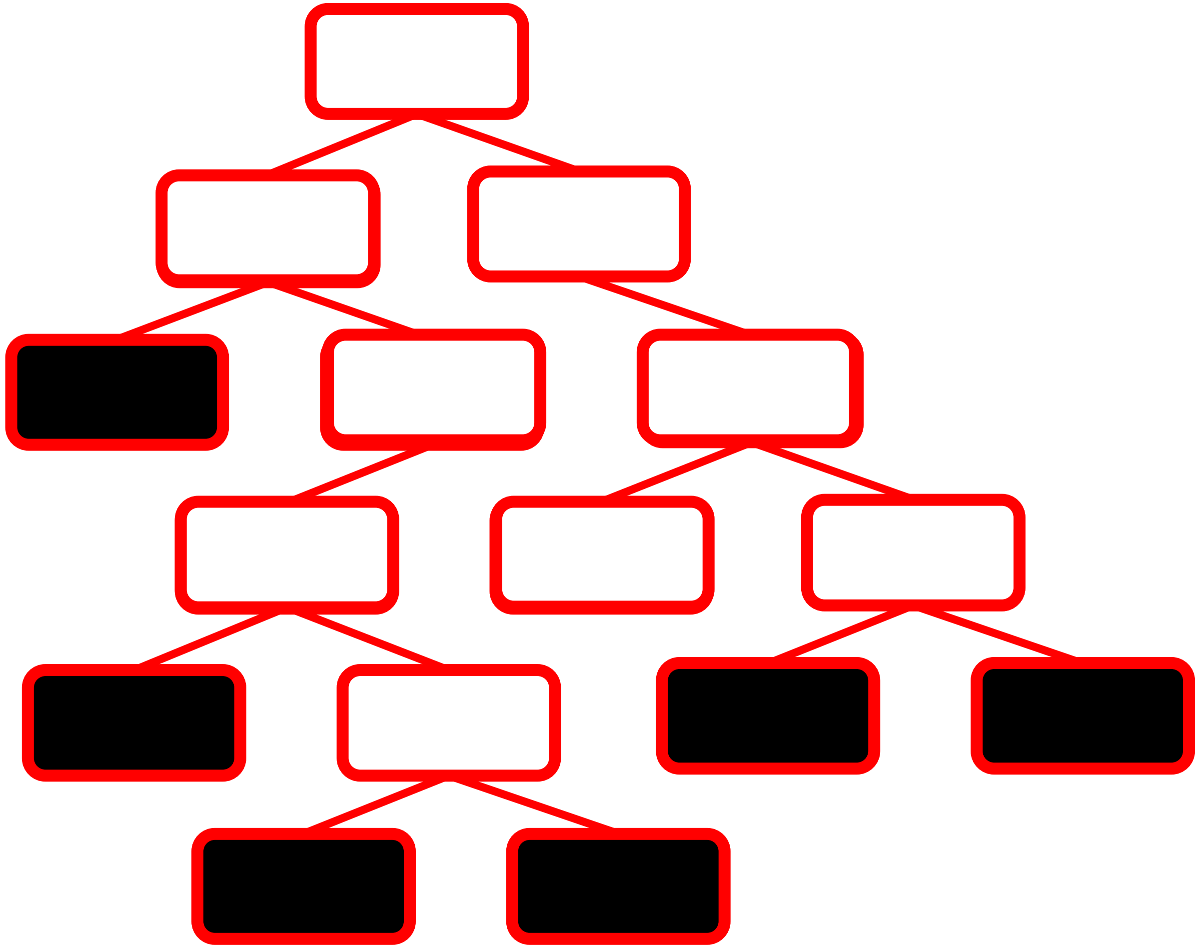
$f(x) = 0.5 + 0.3 \times$  $+ 0.3 \times$ 

UMA árvore de decisão

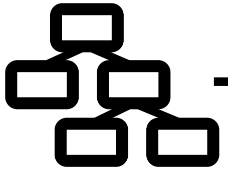
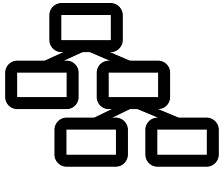
Minimiza a variância de Y dentro de cada folha (bloco preto).

$$\frac{1}{n}\sum(y_i - \overline{y_i})^2$$

Variância



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

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

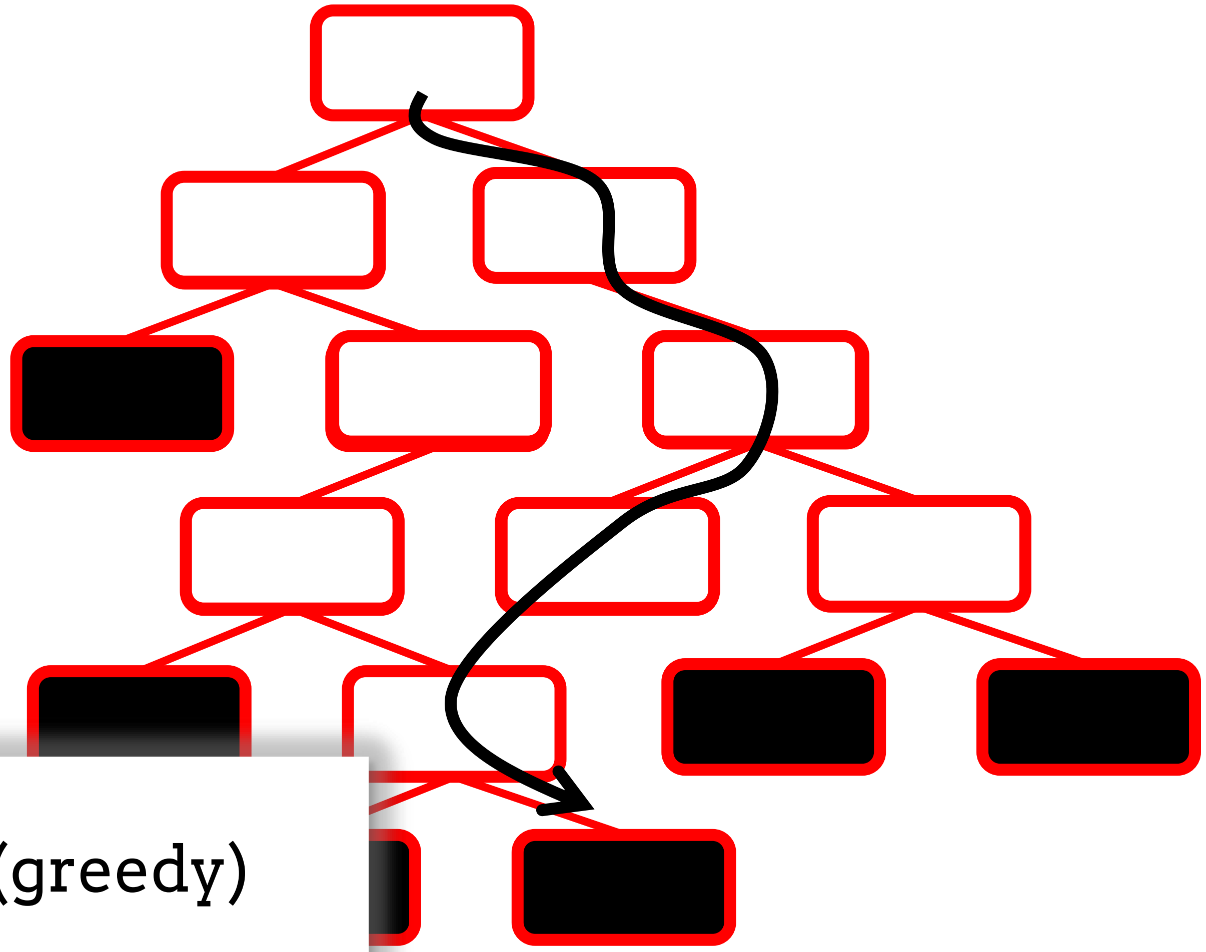
UMA árvore de decisão

Minimiza a variância de Y dentro de cada folha (bloco preto).

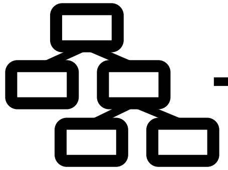
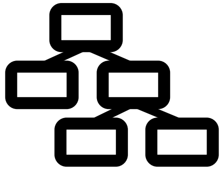
$$\frac{1}{\ }\sum (y_i - \bar{y}_i)^2$$

Algoritmo “Ganancioso” (greedy)

Variância



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

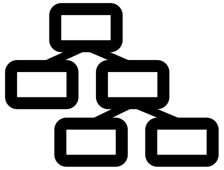
f(x) = 0.5 + 0.3 ×  + 0.3 × 

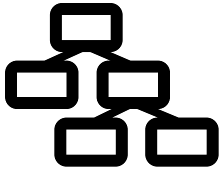
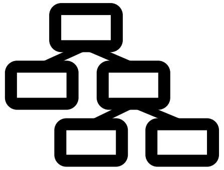
$$\sum (y_i - f(x_i))^2$$

$$\sum L(y_i, f(x_i))$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

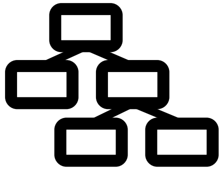
$f1(x) = 0.5$

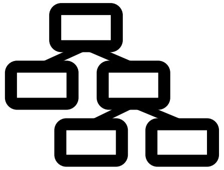
$f2(x) = 0.5 + 0.3 \times$ 

$f3(x) = 0.5 + 0.3 \times$  $+ 0.3 \times$ 

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$ 

$$f3(x) = f2(x) + 0.3 \times$$ 

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

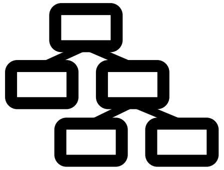
$$f1(x) = 0.5$$

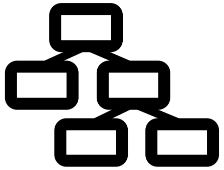
$$f2(x) = f1(x) + 0.3 \times$$

$$f3(x) = f2(x) + 0.3 \times$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$ 

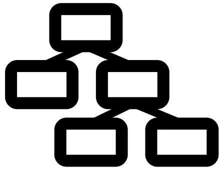
$$f3(x) = f2(x) + 0.3 \times$$ 

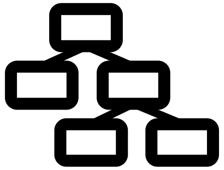
Passo 1:

$$\sum (y_i - f_1(x_i))^2$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

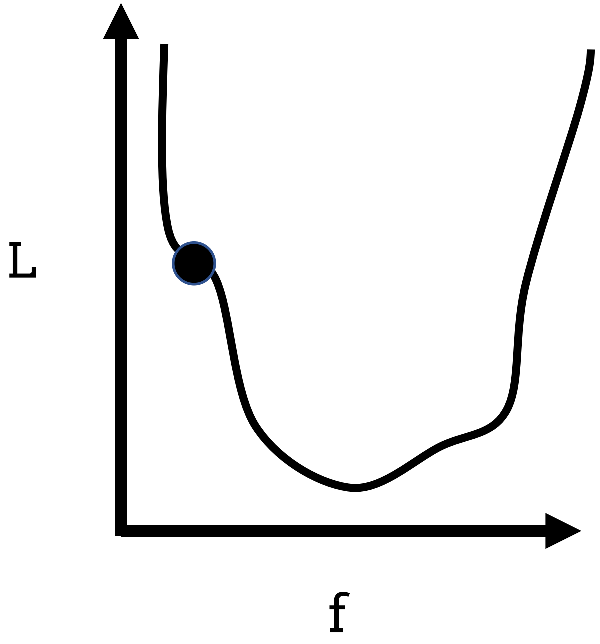
$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$ 

$$f3(x) = f2(x) + 0.3 \times$$ 

Passo 1:

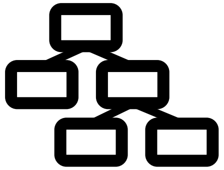
$$L = \sum (y_i - 0.5)^2$$

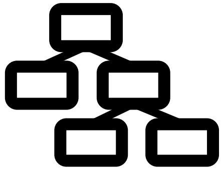


...nada para otimizar! Próximo...

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

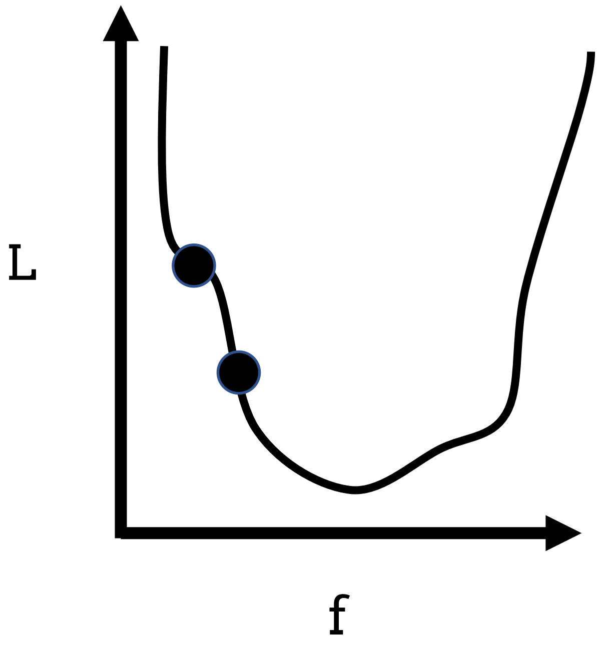
$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$ 

$$f3(x) = f2(x) + 0.3 \times$$ 

Passo 2:

$$L = \sum (y_i - f_2(x_i))^2$$



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

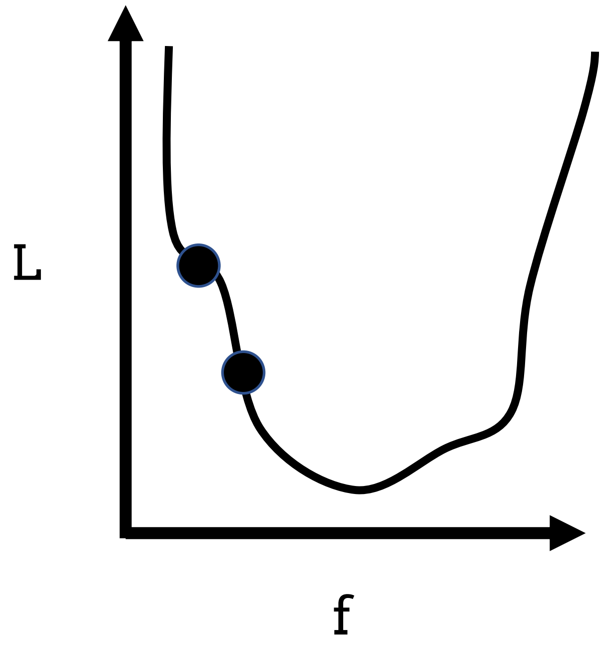
$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$

$$f3(x) = f2(x) + 0.3 \times$$

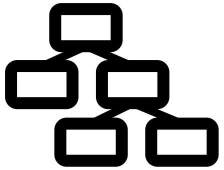
Passo 2:

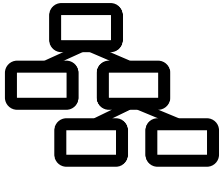
$$L = \sum (y_i - f_1(x_i) - 0.3 \quad)^2$$



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

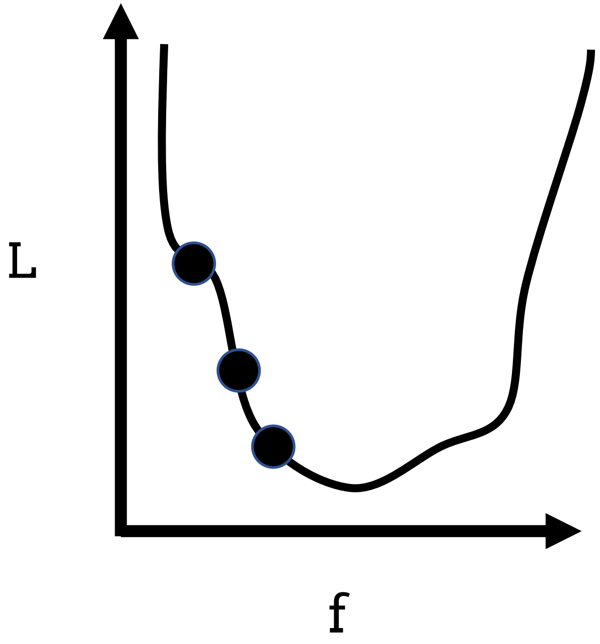
$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$ 

$$f3(x) = f2(x) + 0.3 \times$$ 

Passo 2:

$$L = \sum (y_i - f_3(x_i))^2$$



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

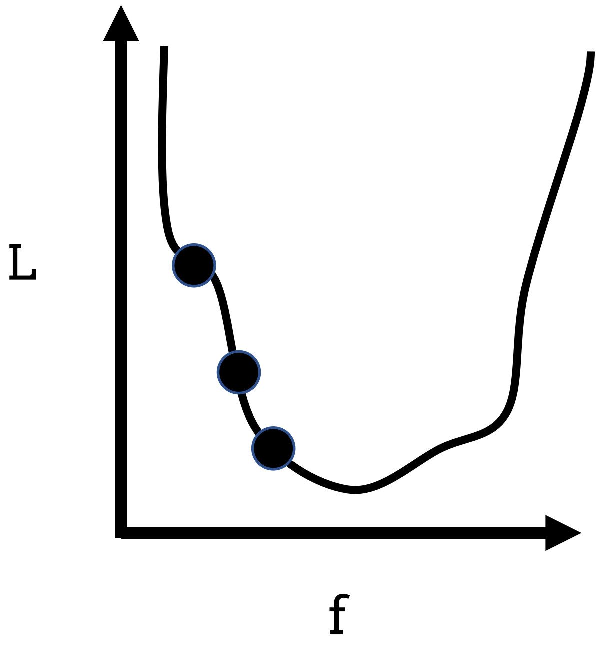
$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$

$$f3(x) = f2(x) + 0.3 \times$$

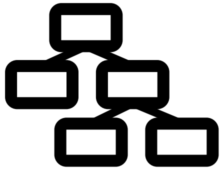
Passo 2:

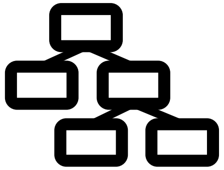
$$L = \sum (y_i - f_2(x_i) - 0.3 \qquad)^2$$



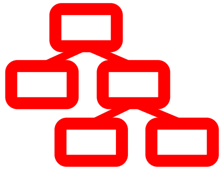
Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

$f1(x) = 0.5$

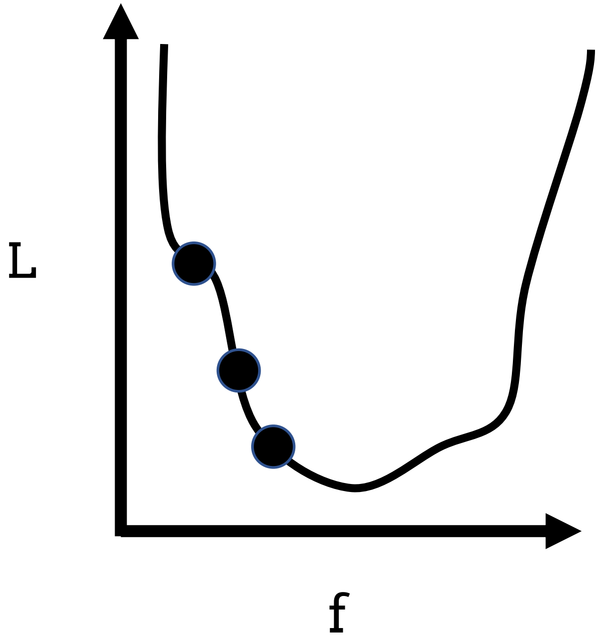
$f2(x) = f1(x) + 0.3 \times$ 

$f3(x) = f2(x) + 0.3 \times$ 

Passo 2:

$L = \sum (y_i - f_2(x_i) - 0.3$  $)^2$

Erro
Resíduo



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$

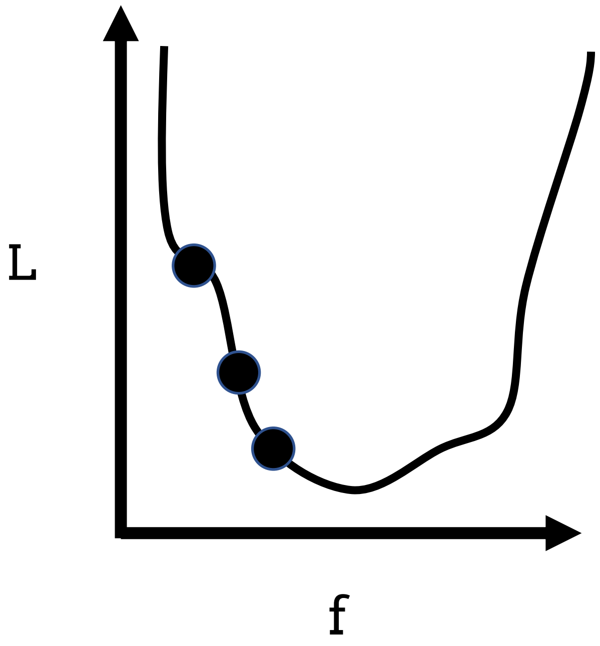
$$f3(x) = f2(x) + 0.3 \times$$

Passo 2:

$$L = \sum \left(r_i - 0.3 \right)^2$$

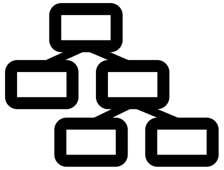
Erro

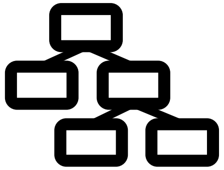
Resíduo



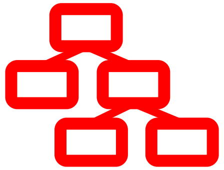
Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$ 

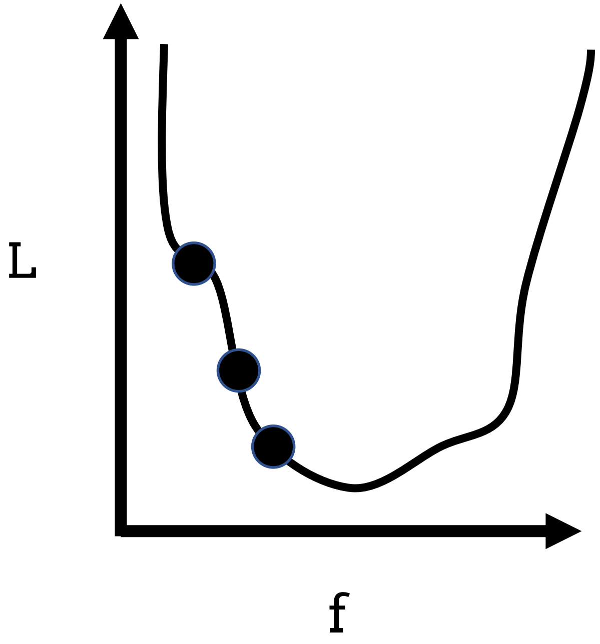
$$f3(x) = f2(x) + 0.3 \times$$ 

Passo 2:

$$L = \sum(r_i - 0.3$$  $$)^2$$

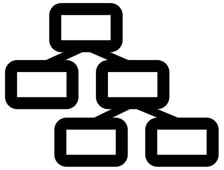
$$\sum(y_i - f(x_i))^2$$

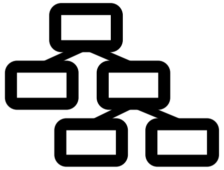
Erro
Resíduo



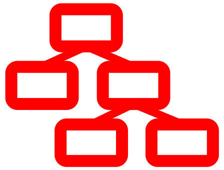
Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$ 

$$f3(x) = f2(x) + 0.3 \times$$ 

Passo m:

$$L = \sum (r_i^{m-1} - 0.3$$  $$)^2$$

Erro

Resíduo

$$fm(x) = \sum_{b=1}^{m} fb(x)$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

$$f1(x) = 0.5$$

$$f2(x) = f1(x) + 0.3 \times$$

$$f3(x) = f2(x) + 0.3 \times$$

Passo m:

$$L = \sum (r_i^{m-1} - 0.3 \quad)^2$$

Erro
Resíduo

$$fm(x) = \sum_{b=1}^{m} fb(x)$$

Modelo final:
Gigantesca soma
de case_whens

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore

$$minimizar\ L(y, f_{-1} + f)$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore

Por exemplo...

$$\sum (y_i - (f_{-1}(x_i) + f\ (x_i)))^2$$

$$minimizar\ L(y, f_{-1} + f)$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore

$$minimizar\ L(y, f_{-1} + f) \approx L(y, f_{-1}) + L'(y, f_{-1})f + \frac{1}{2}L''(y, f_{-1})\ f^2$$

Expansão de Taylor (!!!)

de segunda ordem

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore

$$minimizar\ L(y, f_{-1} + f) \approx \cancel{L(y, f_{-1})} + \underbrace{L'(y, f_{-1})}_{G} f + \frac{1}{2} \underbrace{L''(y, f_{-1})}_{H} f^2$$

Expansão de Taylor (!!!)

de segunda ordem

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore

$$\frac{\mathrm{d}}{df}\left[Gf + \frac{1}{2}Hf^2\right] = 0 \longrightarrow f = -\frac{G}{H}$$

Derivar e igualar a zero

Estratégia consagrada de achar o mínimo

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore

Se...

$$L(y, f_{-1} + f) = \sum(y_i - (f_{-1}(x_i) + f(x_i)))^2$$

Então...

$$G = -2\sum(y_i - f_{-1}(x_i))$$

$$f = \frac{\sum(y_i - f_{-1}(x_i))}{n}$$

E...

$$H = n$$

$$f = -\frac{G}{H}$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore

Se...

$$L(y, f_{-1} + f) = \sum(y_i - (f_{-1}(x_i) + f(x_i)))^2$$

Então...

$$G = -2\sum(y_i - f_{-1}(x_i))$$

$$f = \frac{\sum resíduos}{\# resíduos}$$

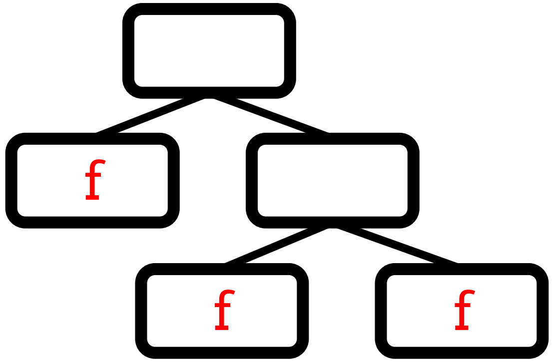
E...

$$H = n$$

$$f = -\frac{G}{H}$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore



Se...

$$L(y, f_{-1} + f) = \sum(y_i - (f_{-1}(x_i) + f\ (x_i)))^2$$

Então...

$$G = -\sum(y_i\ - f_{-1}(x_i))$$

$$f = \frac{\sum resíduos}{\#resíduos}$$

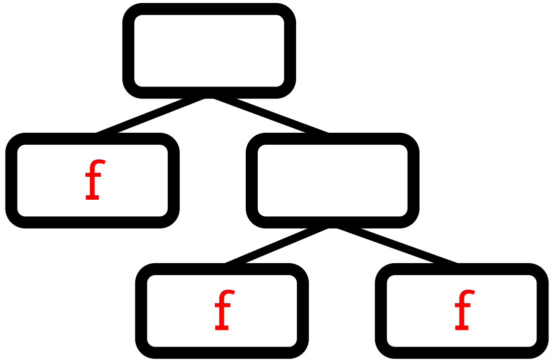
E...

$$H = n$$

$$f = -\frac{G}{H}$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

O valor ótimo de cada árvore



Se...

$$L(y, f_{-1} + f) = \sum(y_i - (f_{-1}(x_i) + f\ (x_i)))^2$$

Então...

$$G = -\sum(y_i\ - f_{-1}(x_i))$$

$$f = \frac{\sum resíduos}{\# resíduos}$$

E...

$$H = n$$

Predição
"output"

$$f = -\frac{G}{H}$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

A loss de cada árvore

Se...

$$f = \frac{\sum resíduos}{\# resíduos}$$

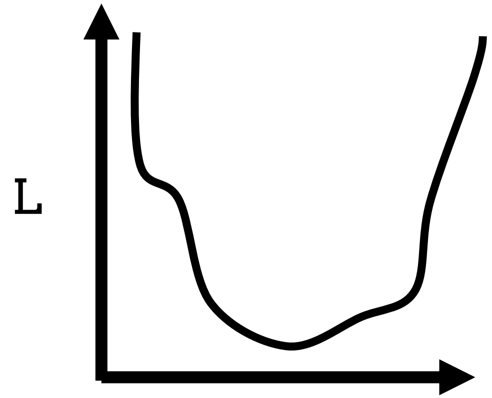
Então...

$$L(y, f_{-1} + f) \approx -\frac{1}{2}\frac{(\sum resíduos)^2}{\# resíduos}$$

$$f = -\frac{G}{H}$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

A loss de cada árvore

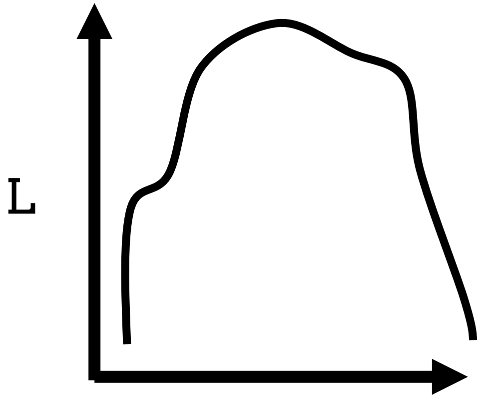


Se...

$$f = \frac{\sum resíduos}{\# resíduos}$$

Então...

$$L(y, f_{-1} + f) \approx \cancel{\frac{1}{2}} \frac{(\sum resíduos)^2}{\# resíduos}$$



$$f = -\frac{G}{H}$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

A loss de cada árvore

Similaridade
"similarity score"

Se...

$$f = \frac{\sum resíduos}{\# resíduos}$$

Então...

$$L(y, f_{-1} + f) \approx \frac{(\sum resíduos)^2}{\# resíduos}$$

$$f = -\frac{G}{H}$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Um Parênteses

$$minimizar\ L(y, f_{-1} + f)$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Um Parênteses

$$minimizar\ L(y, f_{-1} + f) + \lambda f^2 + \gamma T$$

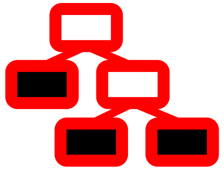
Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Um Parênteses

$$minimizar\ L(y, f_{-1} + f) + \lambda f^2 \cancel{+ \gamma} T$$

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Um Parênteses

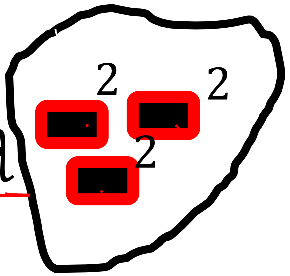
f = 

$$minimizar\ L(y, f_{-1} + f) + \lambda f^2$$

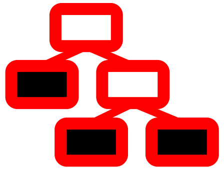
Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Um Parênteses

$$minimizar\ L(y, f_{-1} + f) + \lambda$$



$f =$



Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Um Parênteses

$$minimizar\ L(y, f_{-1} + f) + \lambda\,$$ [redacted] ^{2} [redacted] ^{2} [redacted] ^{2}

$f =$ [redacted]

Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

Um Parênteses

$f =$ [redacted]

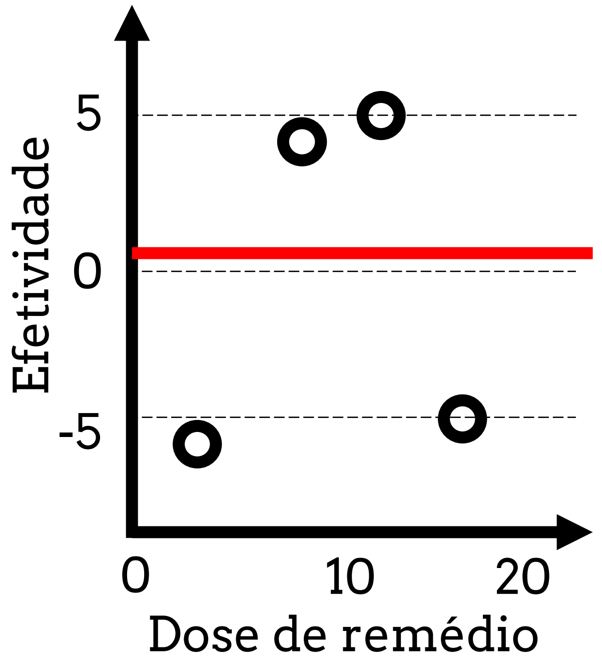
$minimizar\ L(y, f_{-1} + f) + \lambda$ [redacted] ^{2} [redacted] ^{2} [redacted] ^{2}

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

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

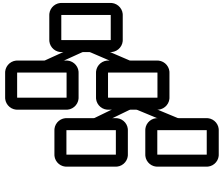
Dose de remédio	Efetividade	Pred
2	-6	0.5
8	4	0.5
12	5	0.5
16	-5	0.5

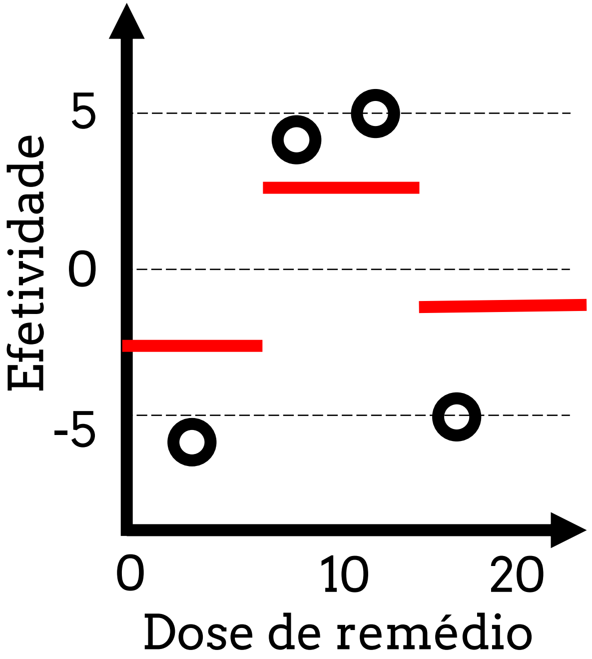
$$f(x) = 0.5$$



$$\sum (y_i - f(x_i))^2 = (0.5 - (-6))^2 + (0.5 - 4)^2 + (0.5 - 5)^2 + (0.5 - (-5))^2$$

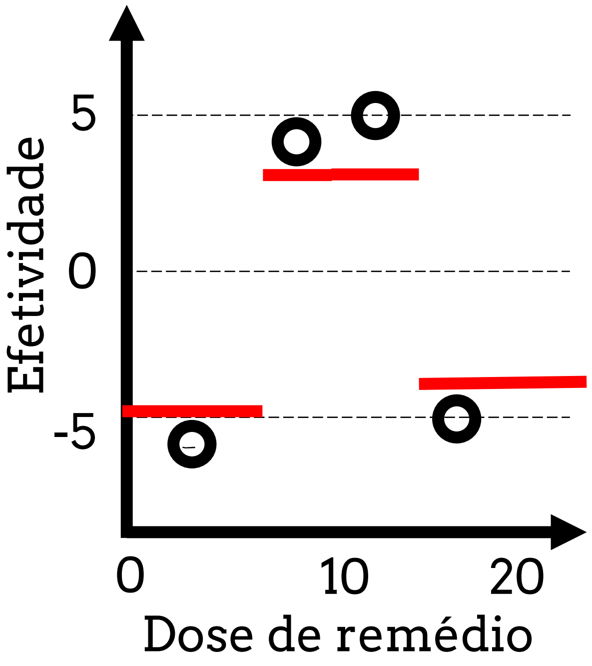
Dose de remédio	Efetividade	Pred
2	-6	-1.45
8	4	1.70
12	5	1.70
16	-5	-1.15

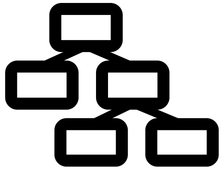
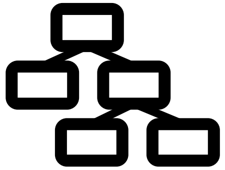
$$f(x) = 0.5 + \varepsilon \times$$ 



$$\sum (y_i - f(x_i))^2 = (-1.45 - (-6))^2 + (1.7 - 4)^2 + (1.7 - 5)^2 + (-1.15 - (-5))^2$$

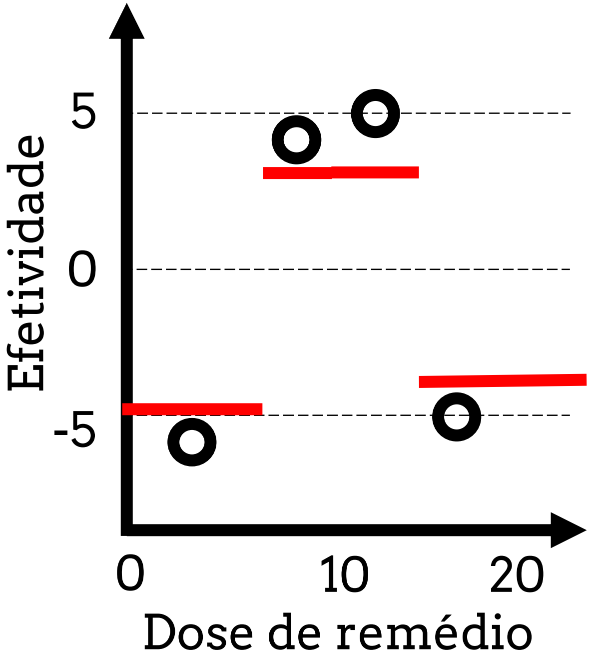
Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

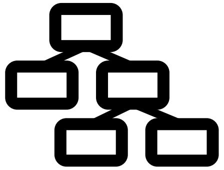
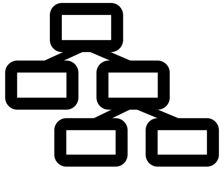


$f(x) = 0.5 + \varepsilon \times$  $+ \varepsilon \times$ 

$$\sum (y_i - f(x_i))^2 = (-2.82 - (-6))^2 + (2.54 - 4)^2 + (2.54 - 5)^2 + (-2.31 - (-5))^2$$

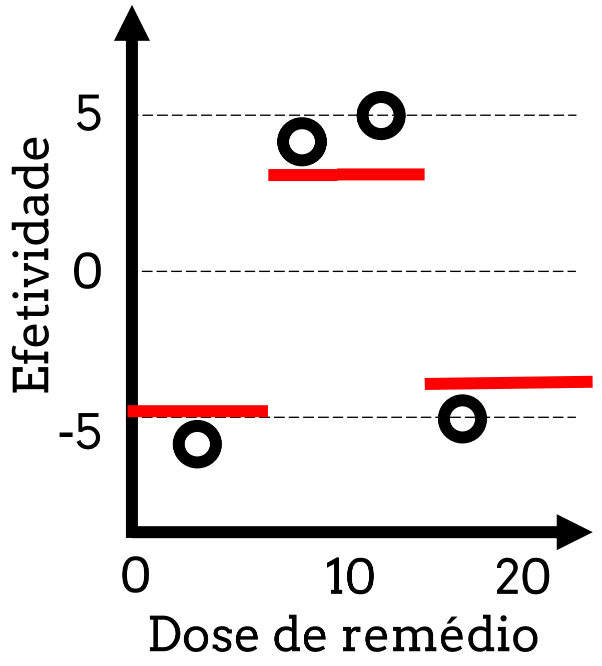
Dose de remédio	Efetividade	Pred
2	-6	-2.82
8	4	2.54
12	5	2.54
16	-5	-2.31

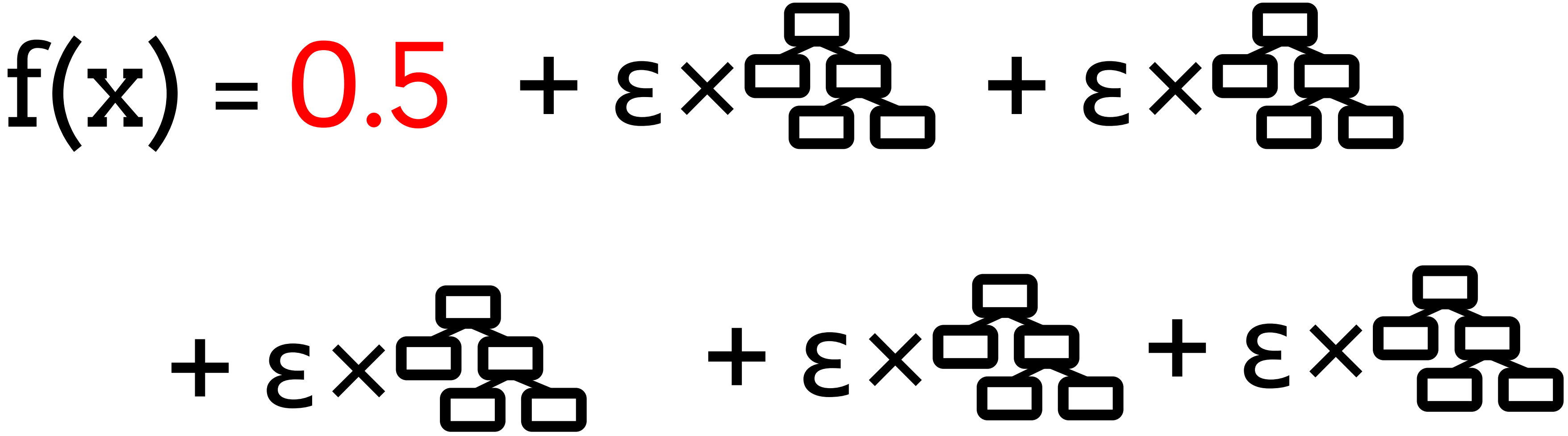


$f(x) = 0.5 + \varepsilon \times$  $+ \varepsilon \times$ 

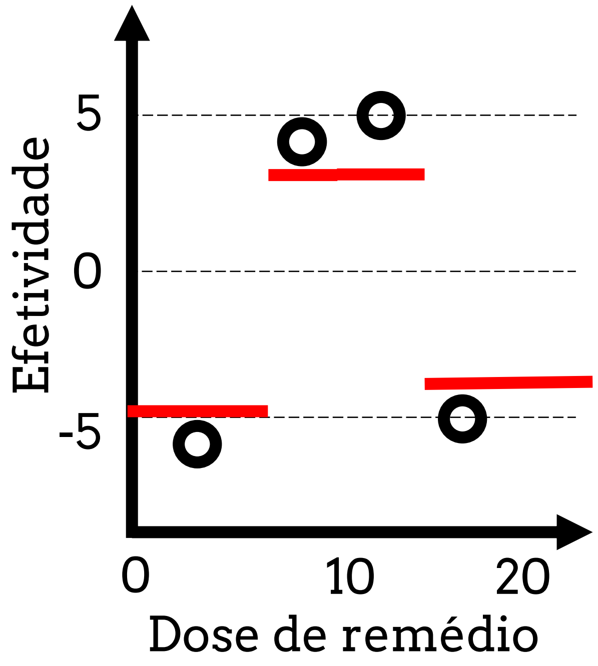
Ao R!

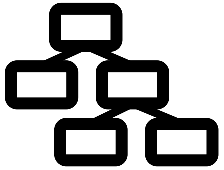
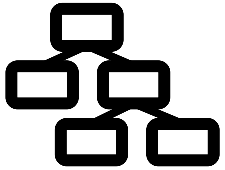
Hiperparam	valor
λ	
γ	
ε	
Tree Depth	
Trees	2





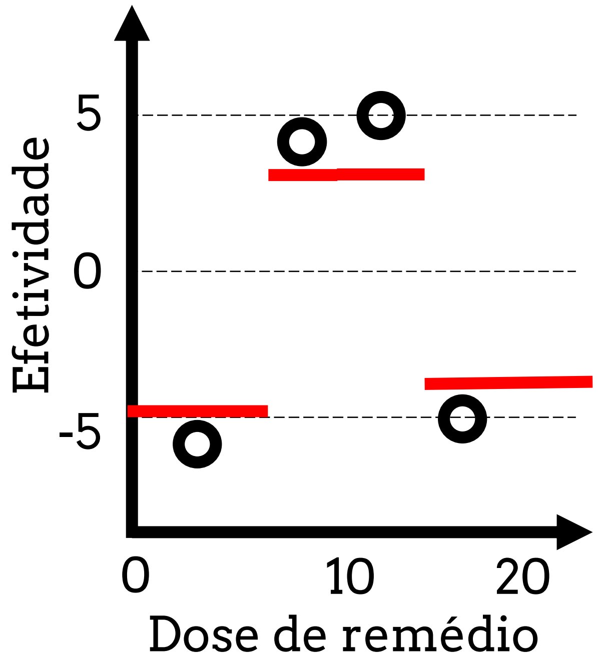
Hiperparam	valor
λ	
γ	
ε	
Tree Depth	
Trees	2

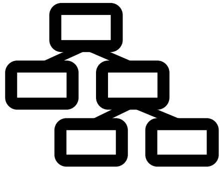
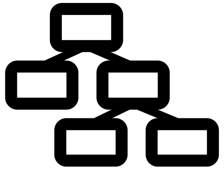


$f(x) = 0.5 + \varepsilon \times$  $+ \varepsilon \times$ 

"Learning Rate"

Hiperparam	valor
λ	
γ	
ε	0.3
Tree Depth	
Trees	2

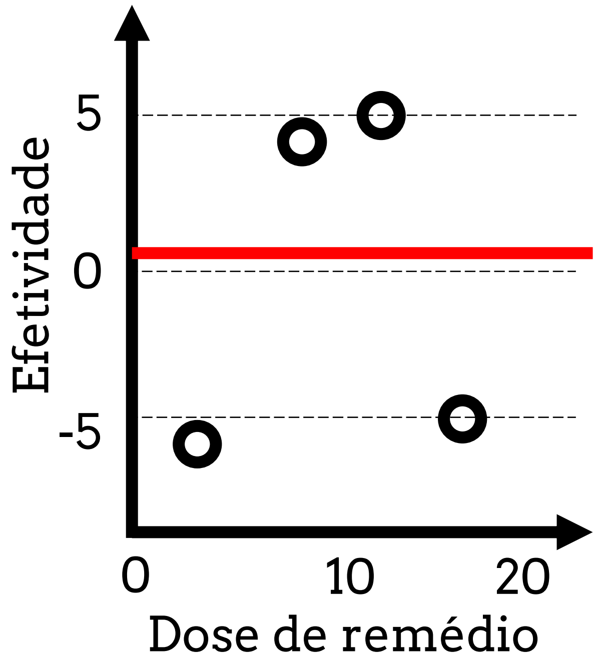


$f(x) = 0.5 + 0.3 \times$  $+ 0.3 \times$ 

"Learning Rate"

Hiperparam	valor
λ	
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

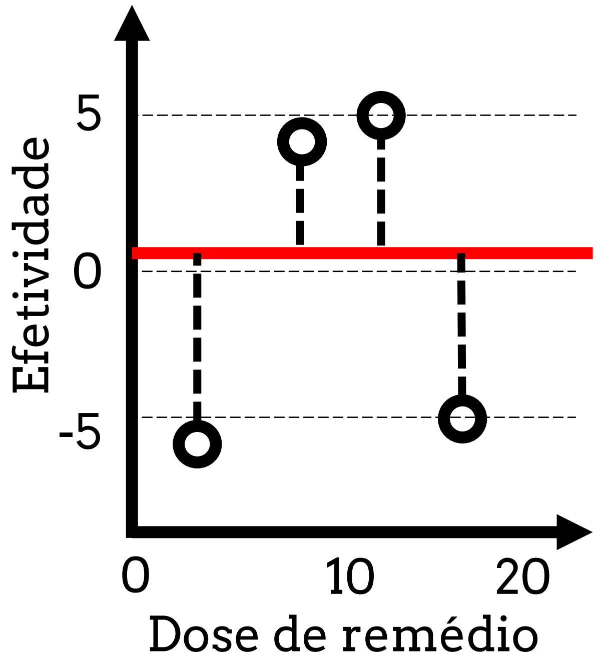


Hora da primeira árvore

Hiperparam	valor
λ	
γ	
ε	0.3
Tree Depth	
Trees	2

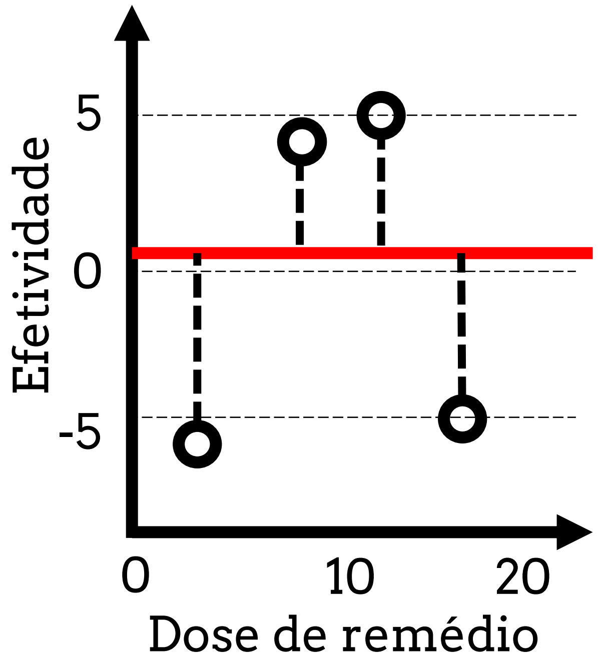
f(x) = 0.5

$$resíduo_i = y_i - f(x_i)$$



Hiperparam	valor
λ	
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$



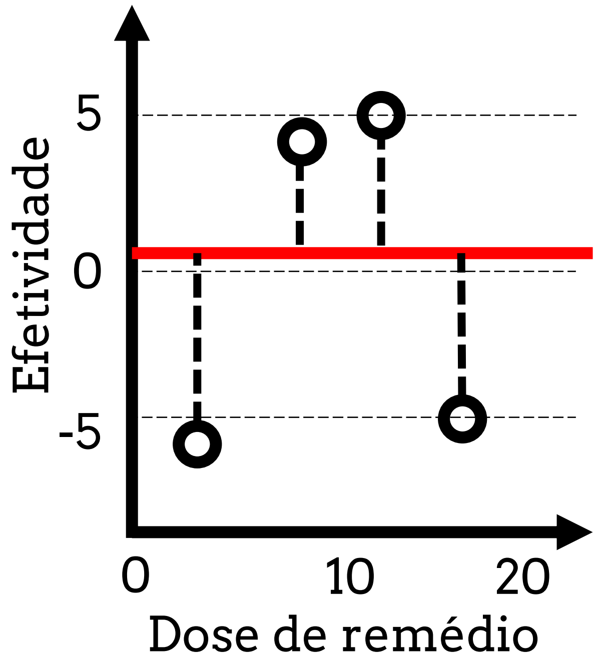
$$resíduo_i = y_i - f(x_i)$$

$$resíduo_1 = -6 - 0.5 = -6.5$$
$$resíduo_2 = 4 - 0.5 = 3.5$$
$$resíduo_3 = 5 - 0.5 = 4.5$$
$$resíduo_4 = -5 - 0.5 = -5.5$$

Hiperparam	valor
λ	
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

-6.5, 3.5, 4.5, -5.5

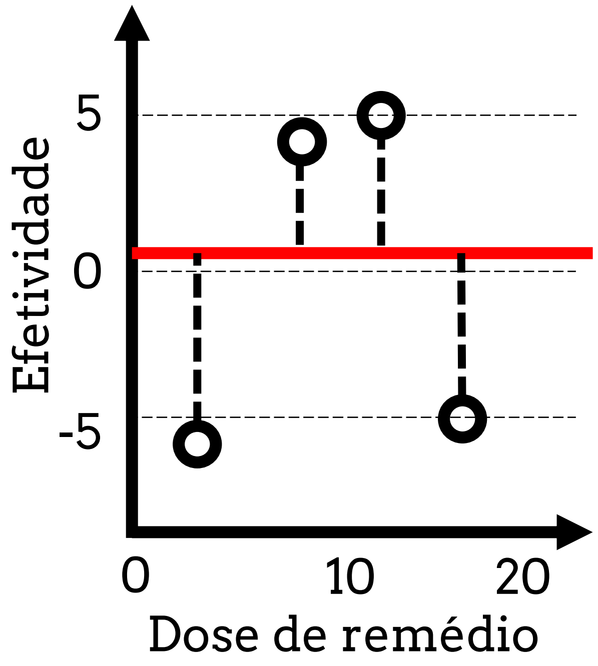


$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

-6.5, 3.5, 4.5, -5.5



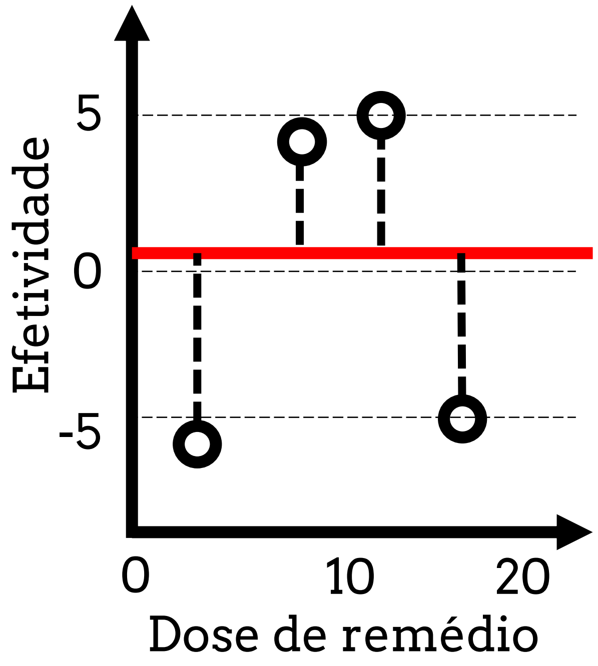
$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

"Regularization Parameter"

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

-6.5, 3.5, 4.5, -5.5



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

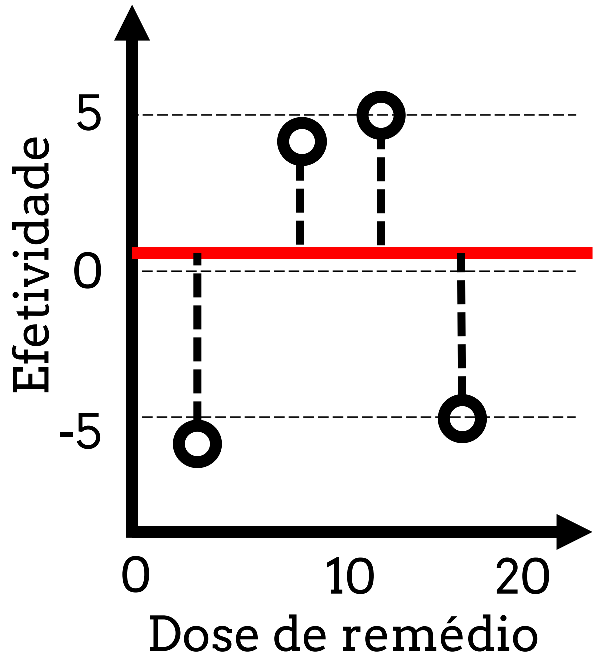
$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

$Similaridade = 4$

-6.5, 3.5, 4.5, -5.5

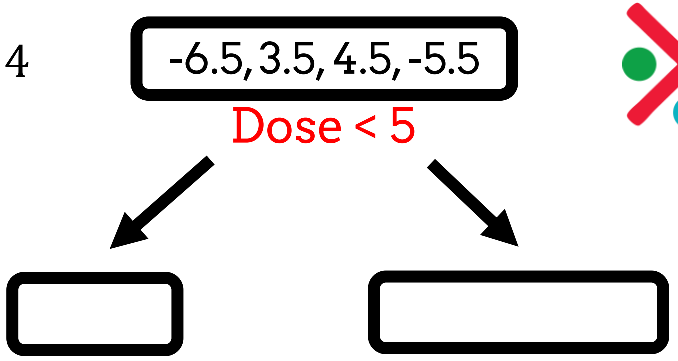


$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

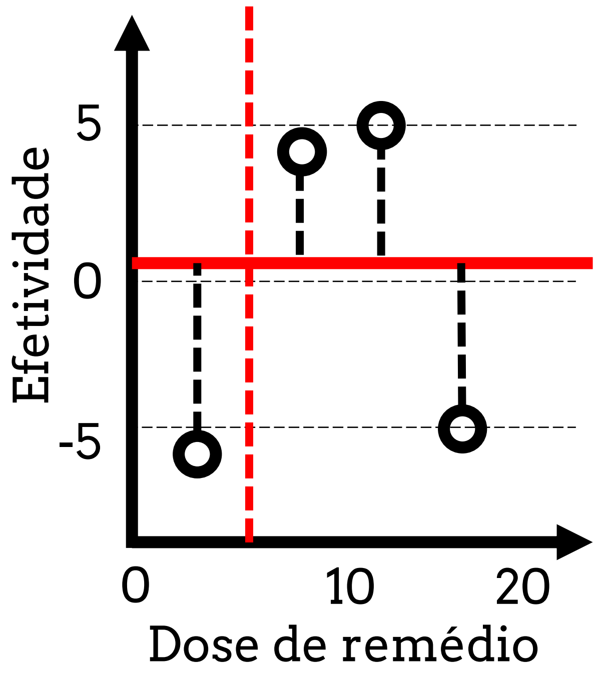
$f(x) = 0.5$

$Similaridade = 4$



$Similaridade_{esq} = \frac{\quad}{\quad} =$

$Similaridade_{dir} = \frac{\quad}{\quad} =$

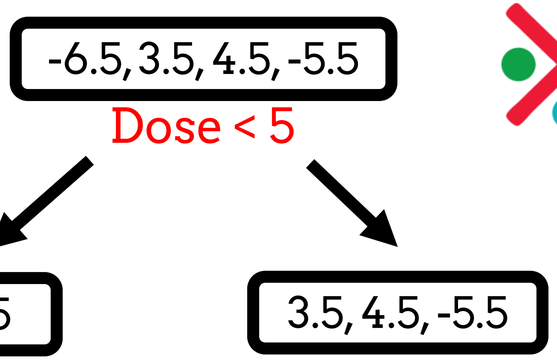


$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

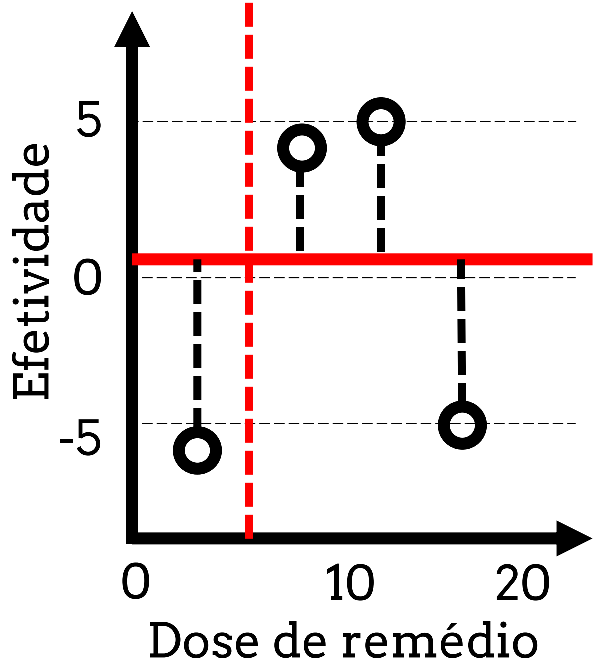
f(x) = 0.5

$Similaridade = 4$



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

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



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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

f(x) = 0.5

$Similaridade = 4$

-6.5, 3.5, 4.5, -5.5

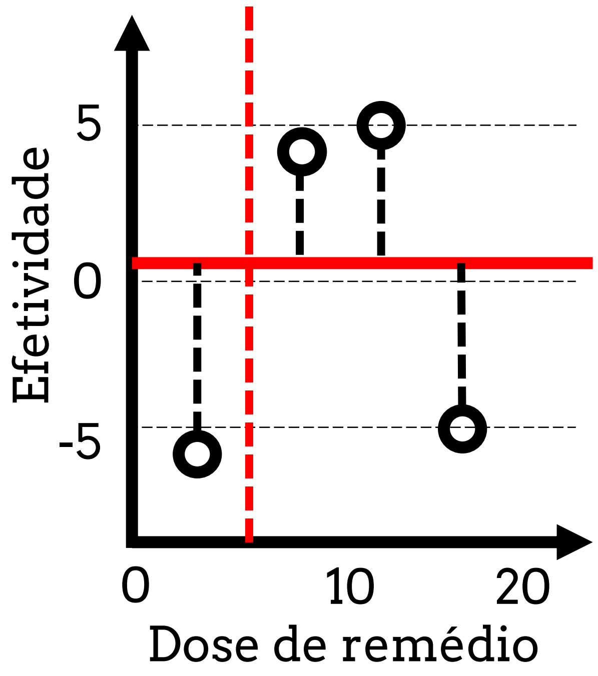
Dose < 5

-6.5

$Similaridade = 42.25$

3.5, 4.5, -5.5

$Similaridade = 2.08$

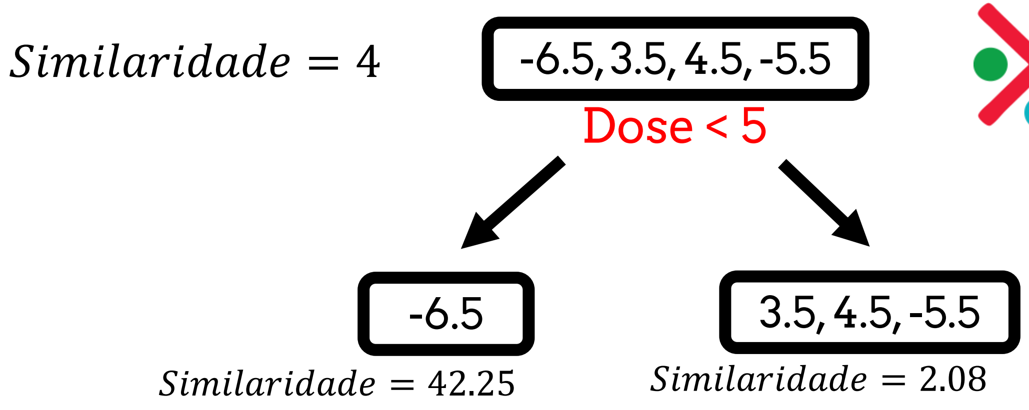


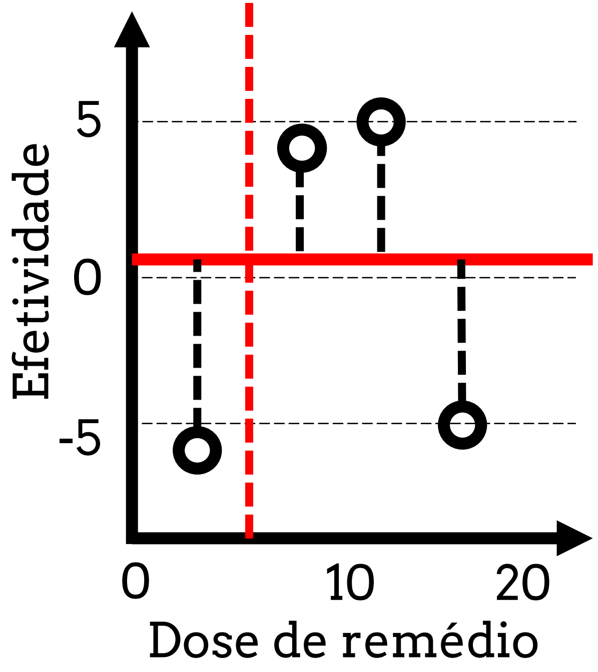
$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$





$$Gain =$$

Pergunta	Gain
Dose < 5	

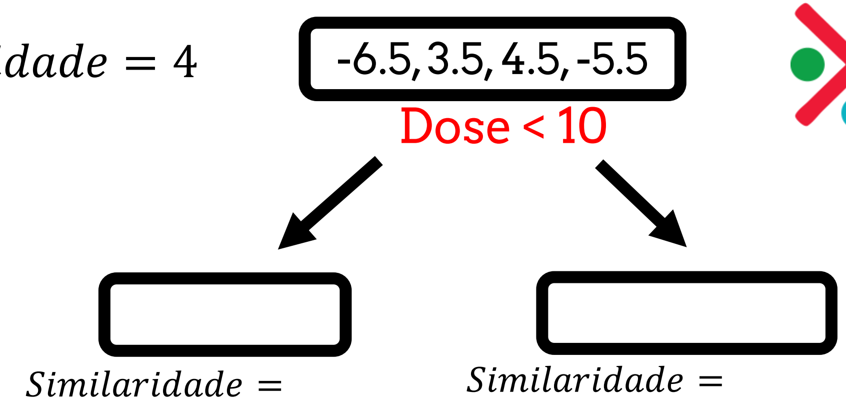
$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

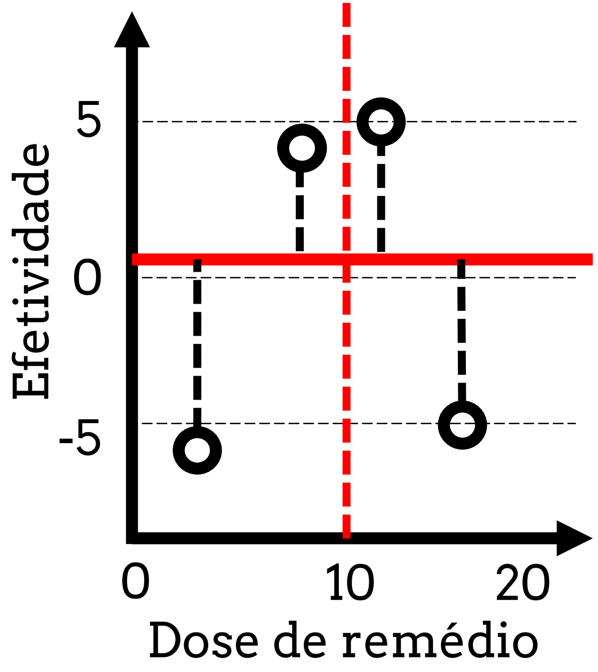
$f(x) = 0.5$

$Similaridade = 4$



$$Similaridade_{esq} = \frac{}{} =$$

$$Similaridade_{dir} = \frac{}{} =$$



Pergunta	Gain
Dose < 5	40.33
Dose < 10	

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

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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

$Similaridade = 4$

-6.5, 3.5, 4.5, -5.5

Dose < 10

-6.5, 3.5

$Similaridade = 4.5$

4.5, -5.5

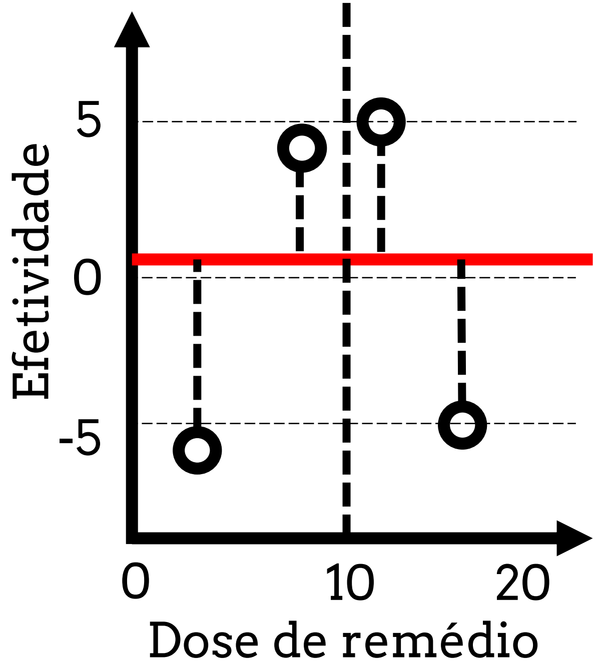
$Similaridade = 0.5$

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

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

$$Gain =$$

Pergunta	Gain
Dose < 5	40.33
Dose < 10	



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

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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

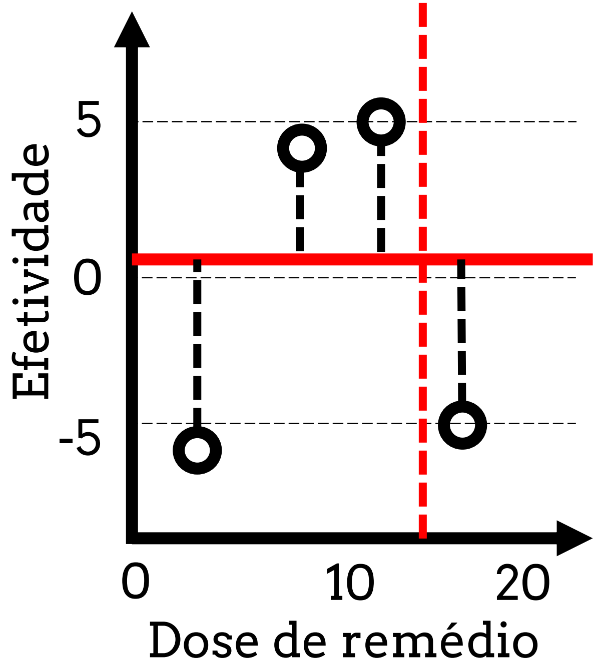
$Similaridade = 4$

-6.5, 3.5, 4.5, -5.5

Dose < 15

-6.5, 3.5, 4.5

-5.5



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
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

$f(x) = 0.5$

$Similaridade = 4$

-6.5, 3.5, 4.5, -5.5

Dose < 15

-6.5, 3.5, 4.5

$Similaridade = 0.75$

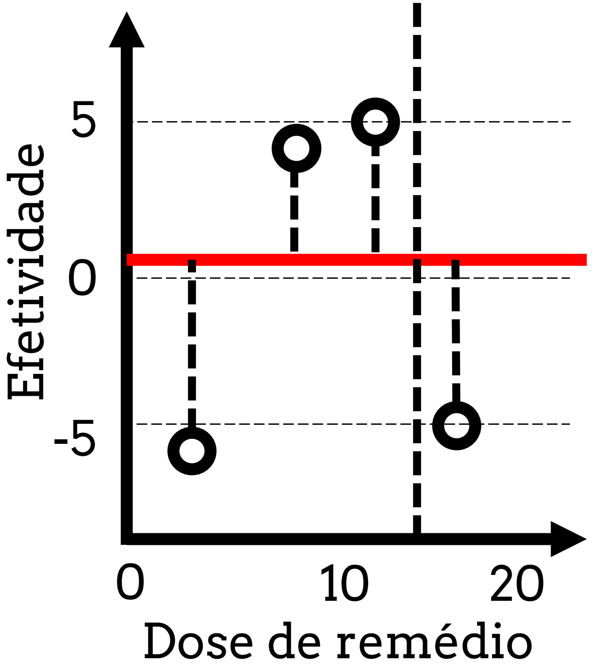
-5.5

$Similaridade = 30.25$

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

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

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



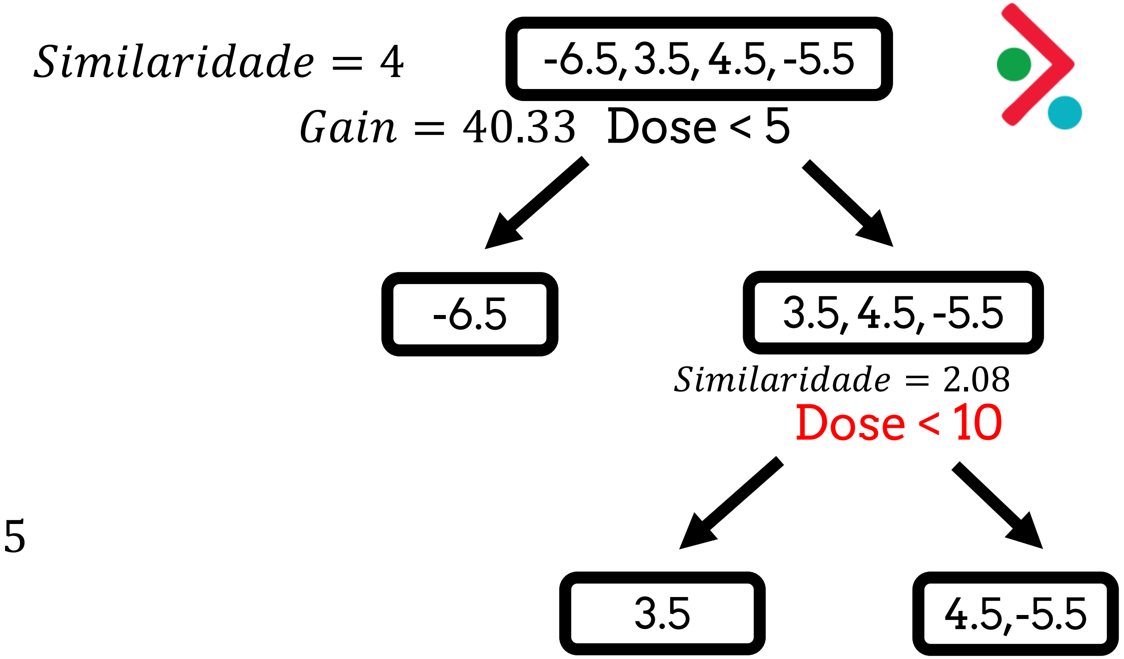
Pergunta	Gain
Dose < 5	40.33
Dose < 10	1
Dose < 15	27

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

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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2

f(x) = 0.5

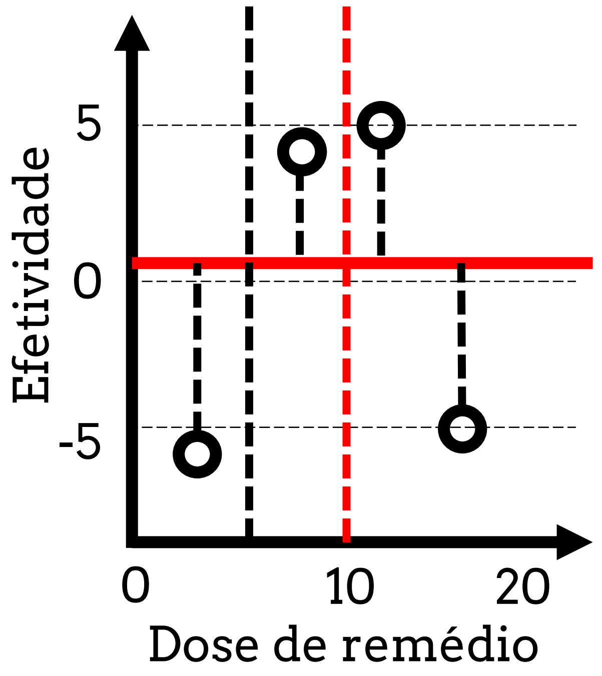


$$Similaridade_{esq} = \frac{(3.5)^2}{1+0} = 12.25$$

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

$$Gain = 12.25 + 0.5 - 2.08 = 10.67$$

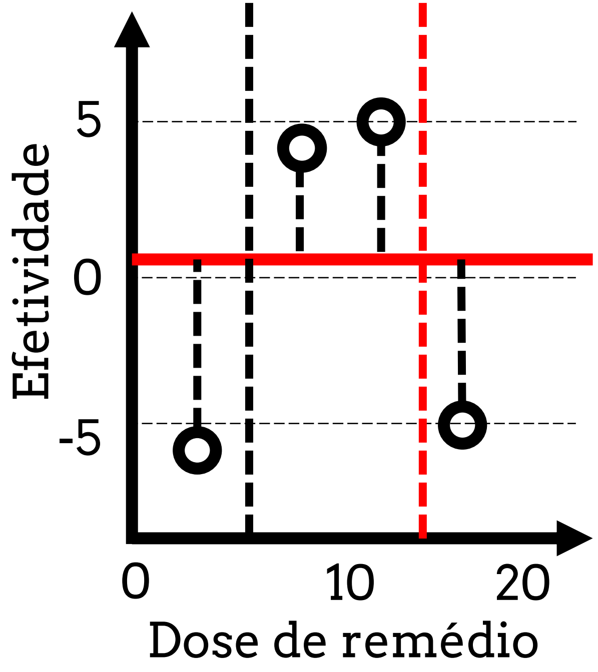
Pergunta	Gain
Dose < 10	10.67
Dose < 15	



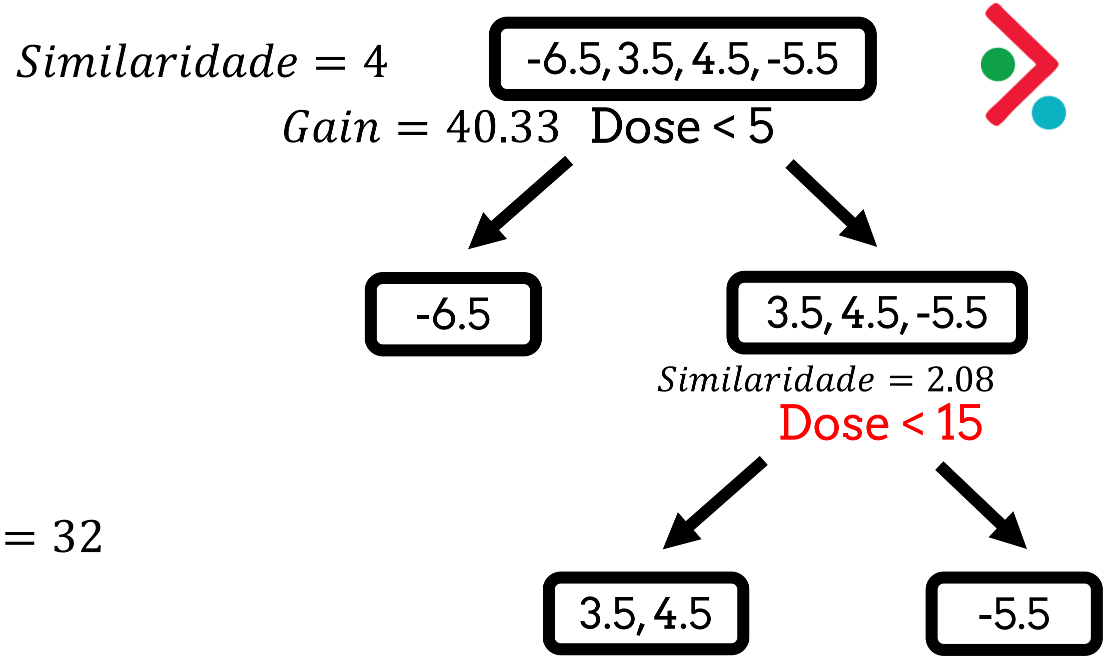
$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	
Trees	2



$f(x) = 0.5$



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

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

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

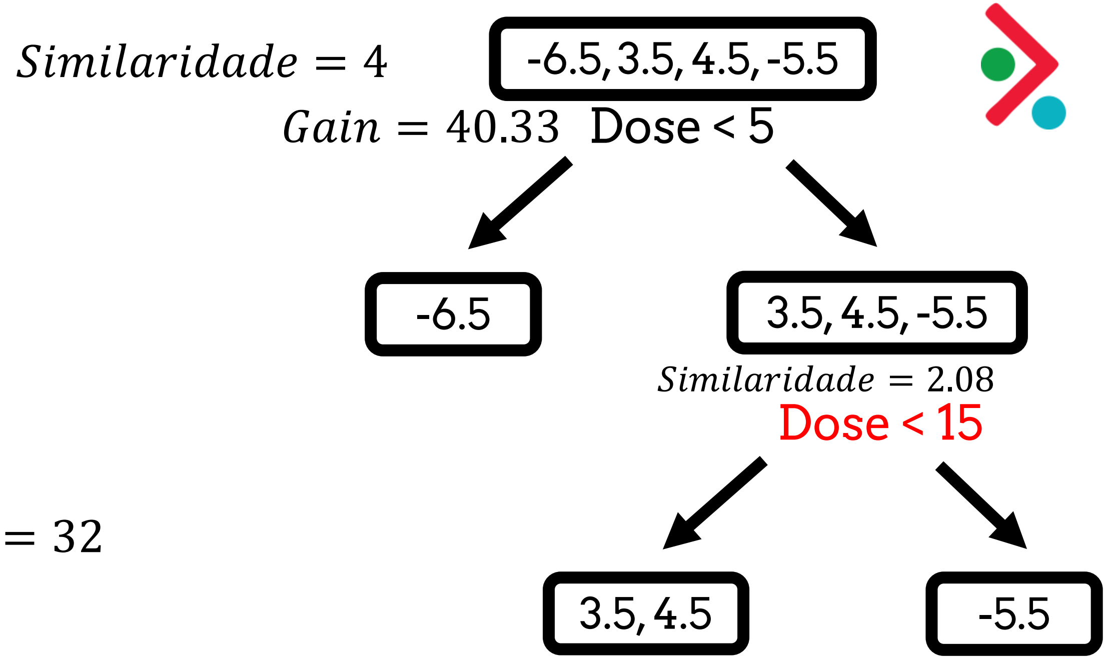
Pergunta	Gain
Dose < 10	10.67
Dose < 15	60.17

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

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

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	2
Trees	2

$f(x) = 0.5$

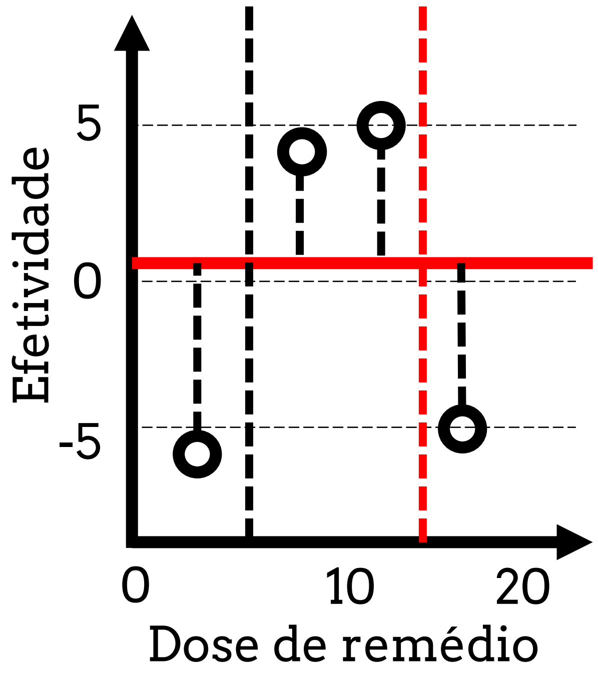


$$Similaridade_{esq} = \frac{(+3.5 + 4.5)^2}{2 + 0} = 32$$

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

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

Pergunta	Gain
Dose < 10	10.67
Dose < 15	60.17



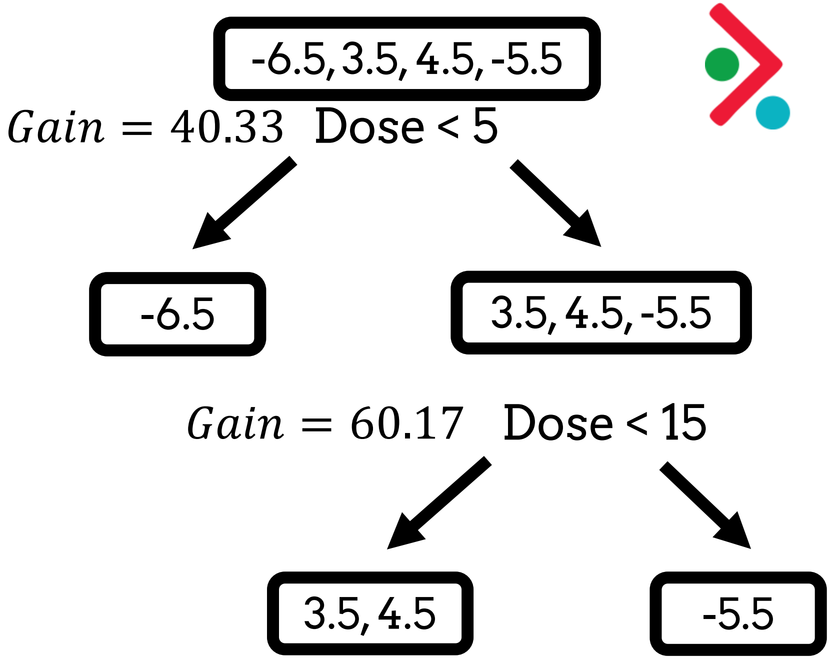
$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

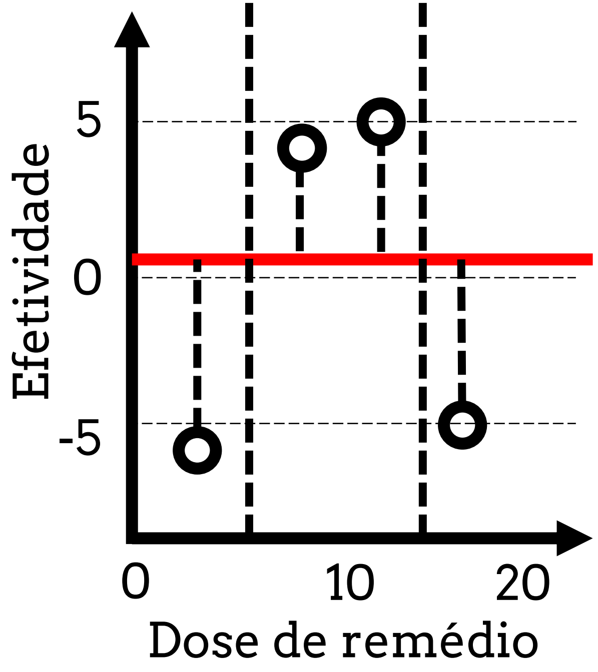
$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	2
Trees	2

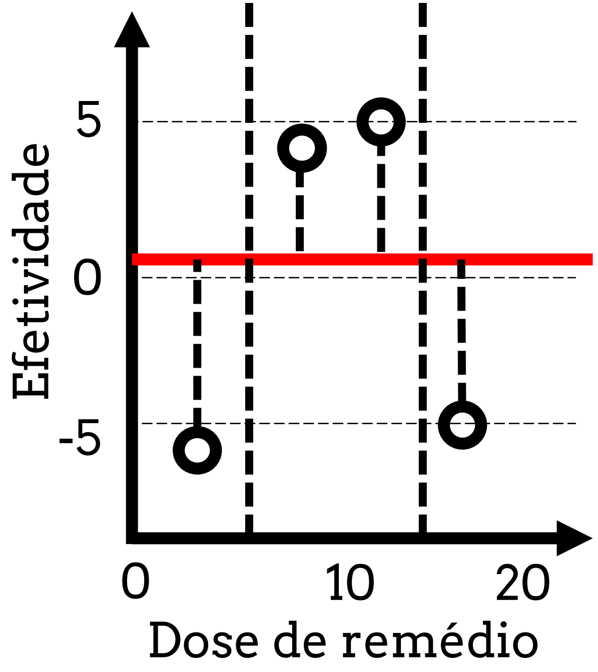
$f(x) = 0.5$

Hora da poda

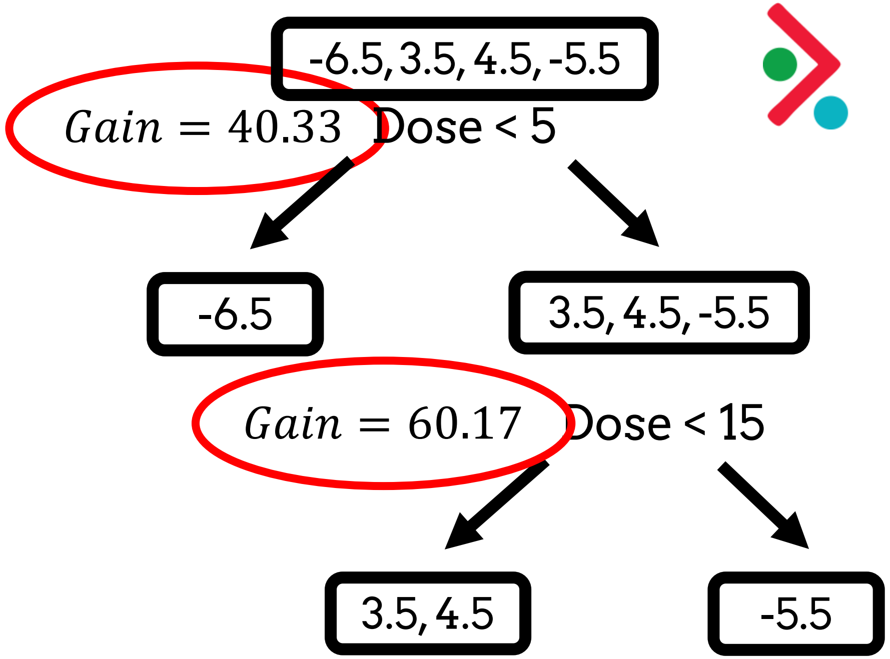




Hiperparam	valor
λ	0
γ	
ε	0.3
Tree Depth	2
Trees	2



$f(x) = 0.5$



Hora da poda

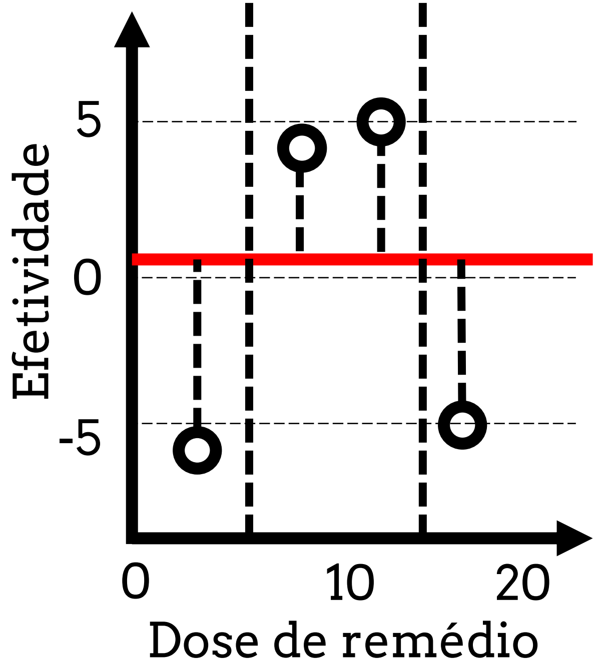
XGBoost usa o Gain para fazer a poda das árvores.

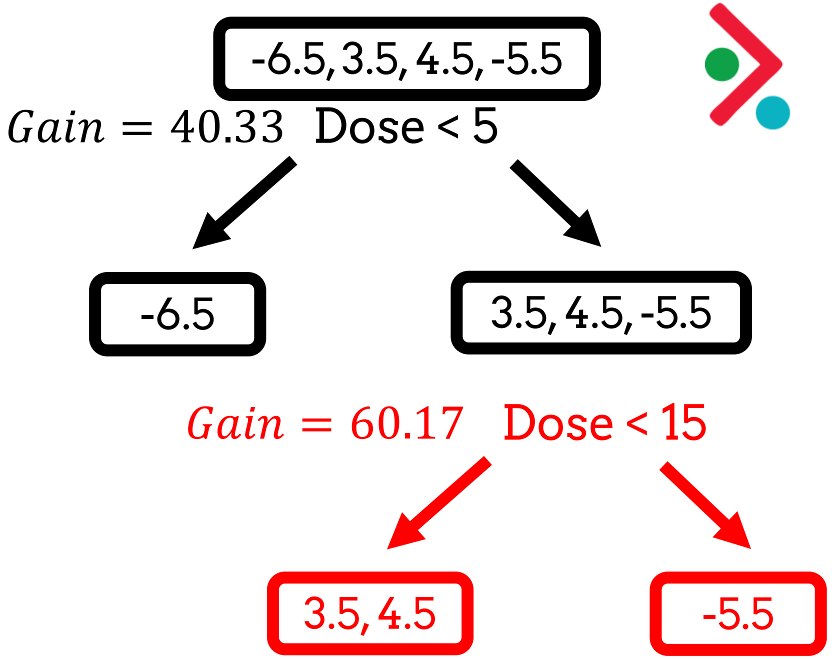
γ

"gamma": nota de corte para o Gain.
Se gain – γ for positivo, então não poda!

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

$f(x) = 0.5$

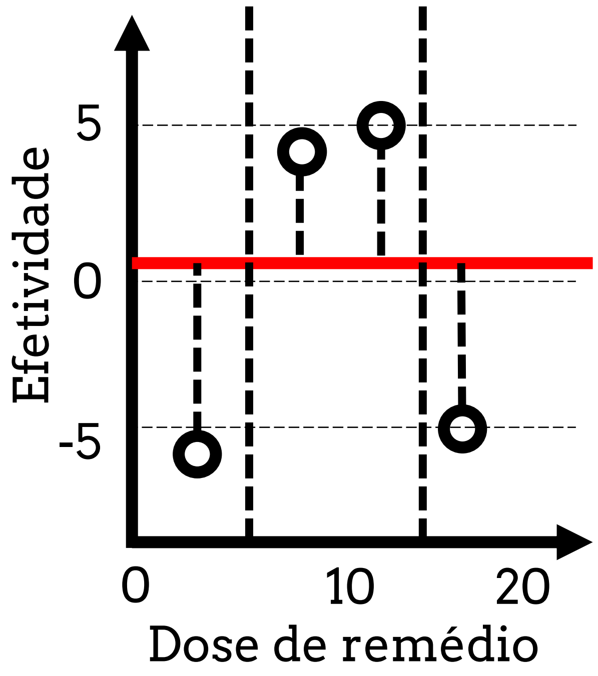


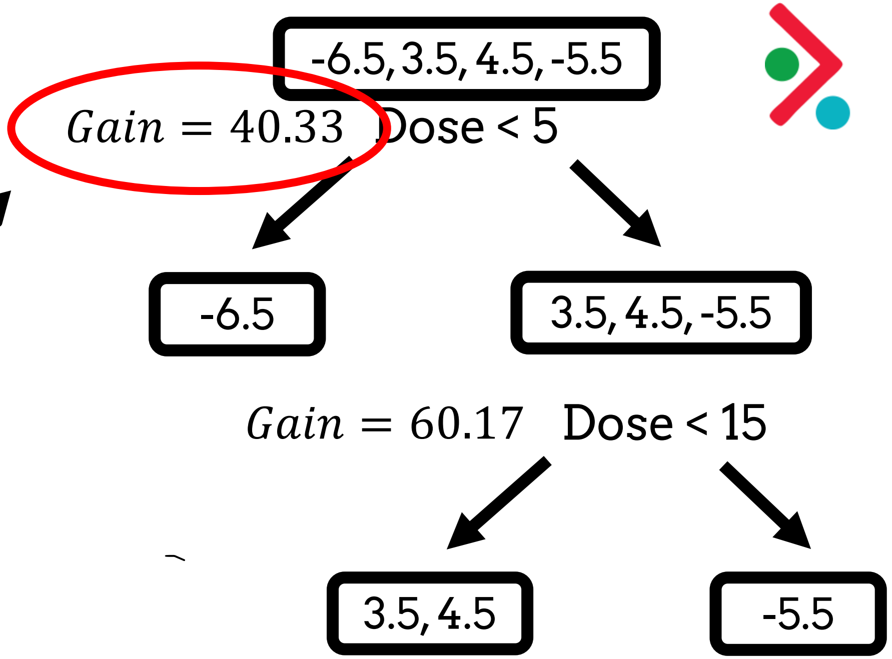


Se gain – γ for positivo,
então não poda!

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

$f(x) = 0.5$



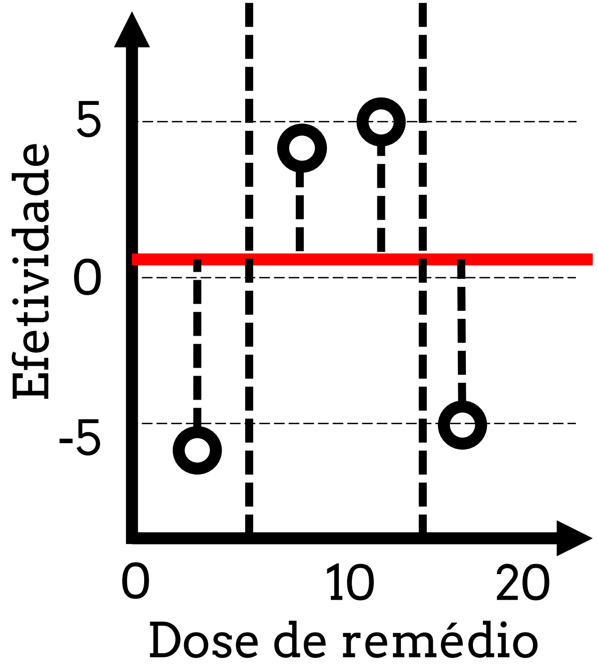


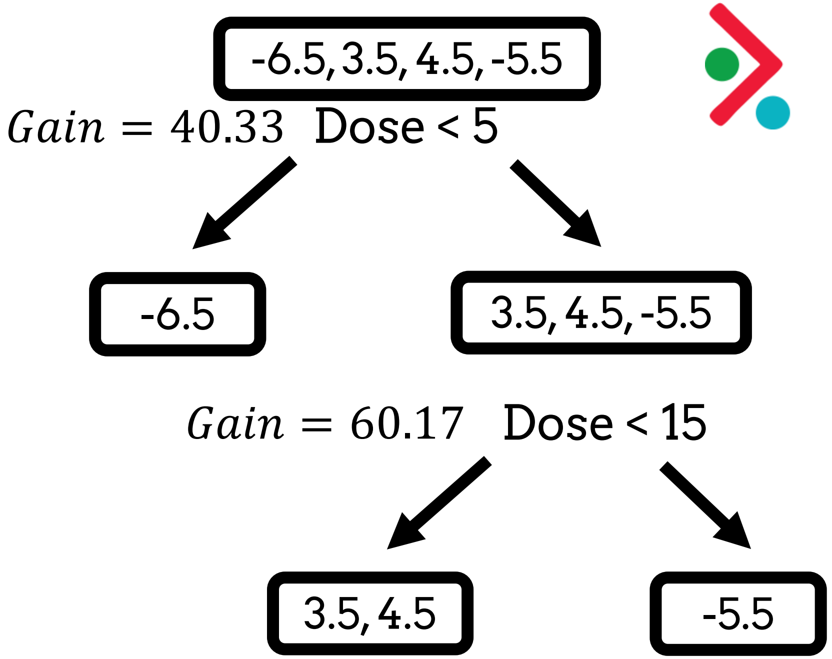
OBSERVAÇÃO: o gain do primeiro nó é menor que 50, indicando para podar. Porém o ramo filho não foi podado, por isso não podamos o pai também.

Se gain – γ for positivo, então não poda!

Hiperparam	valor
λ	0
γ	70
ε	0.3
Tree Depth	2
Trees	2

$f(x) = 0.5$



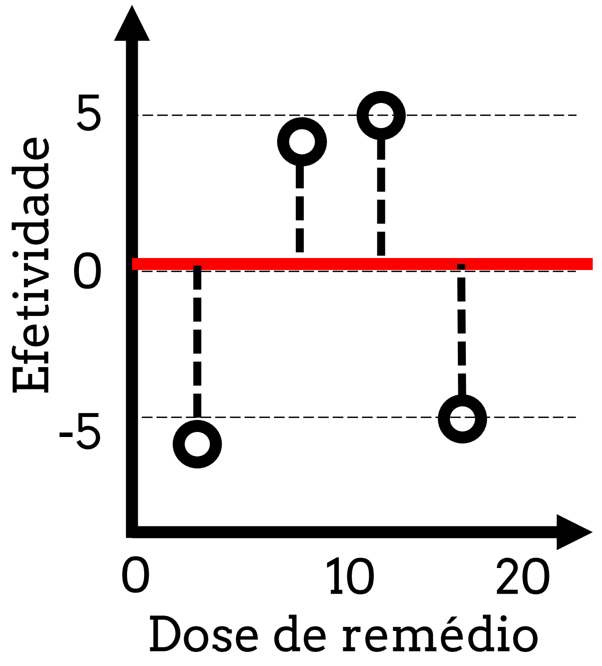


Se gain – γ for positivo, então não poda!

Hiperparam	valor
λ	0
γ	70
ε	0.3
Tree Depth	2
Trees	2

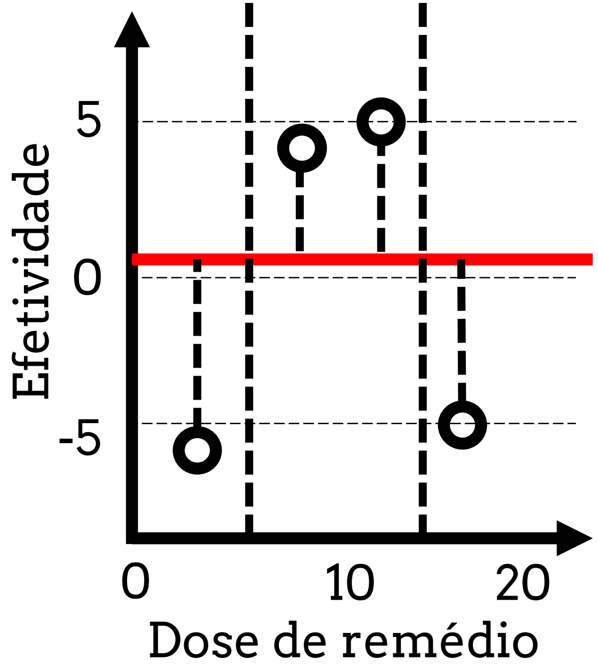
$$f(x) = 0.5 + 0.3 \times (-1)$$

-6.5, 3.5, 4.5, -5.5

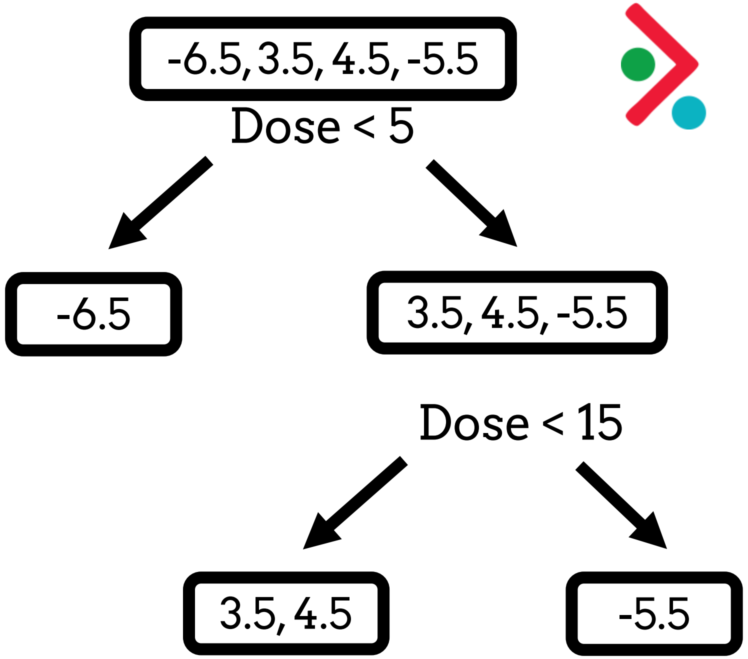


Se gain – γ for positivo, então não poda!

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2



$f(x) = 0.5$



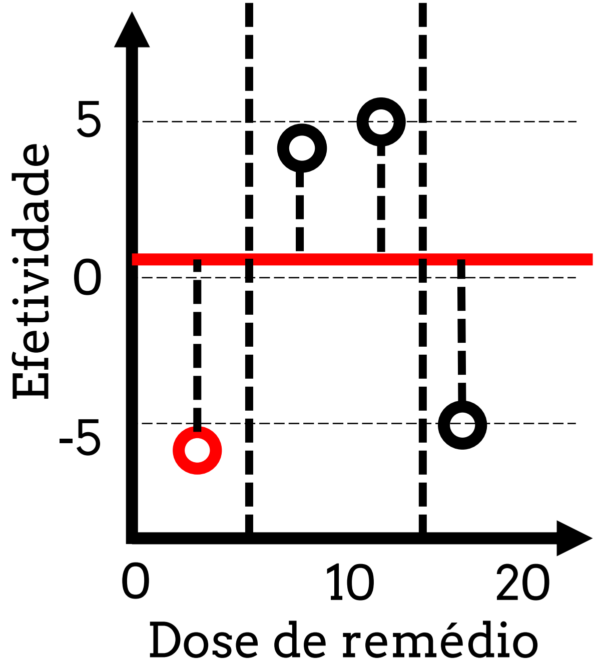
Hora das predições

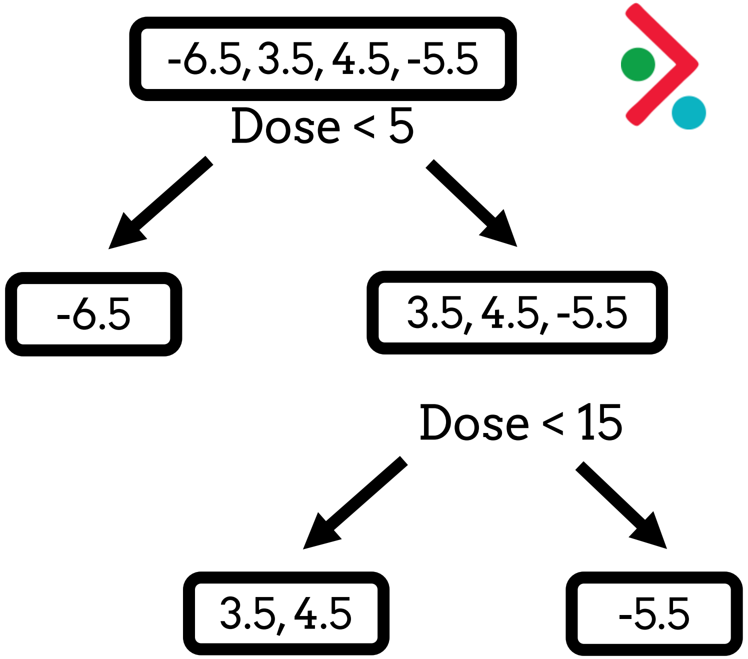
Ou "escoragem"

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

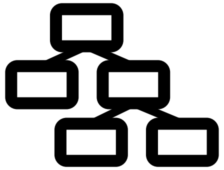
$f(x) = 0.5 + 0.3 \times$

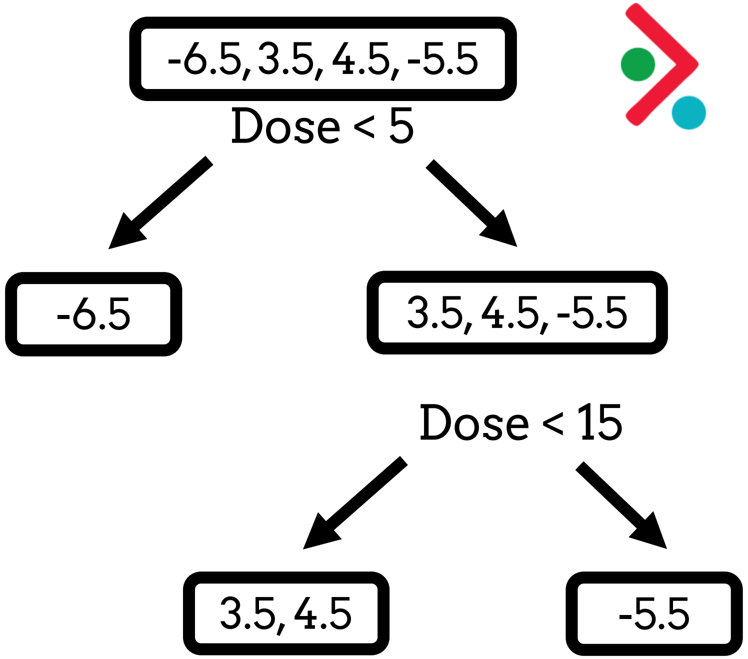
$f(x_1)$



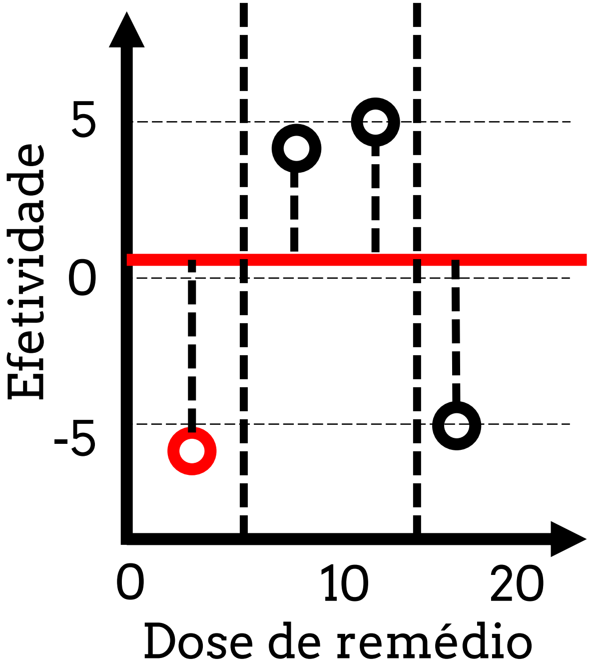


Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

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

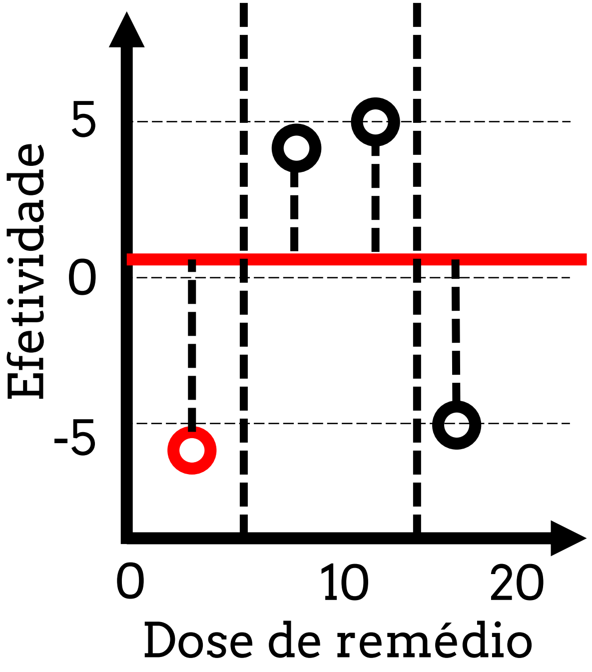


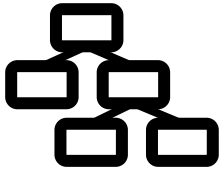
$f(2) = 0.5 + 0.3 \times$

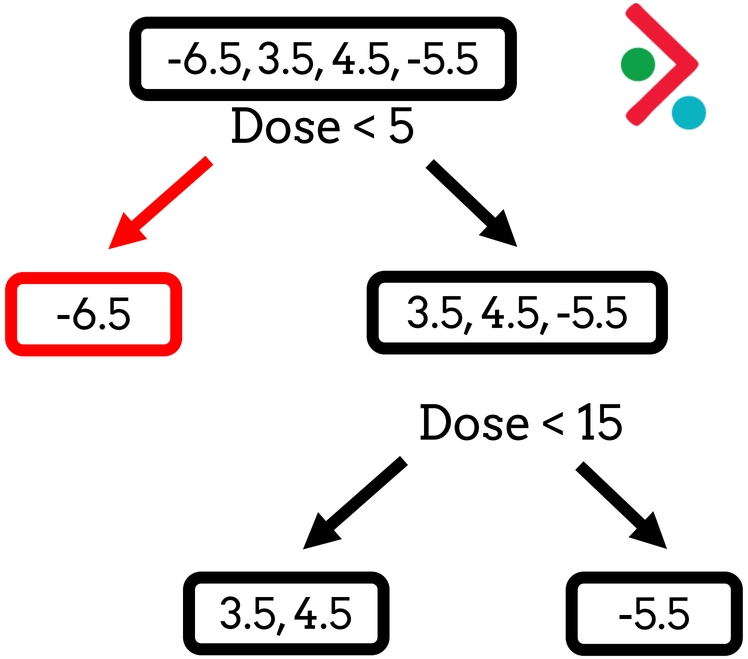


$$predição = \frac{\sum resíduos}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2



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

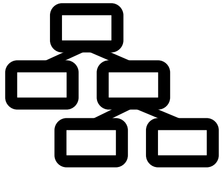


$$f(2) = 0.5 + 0.3 \times -6.5$$

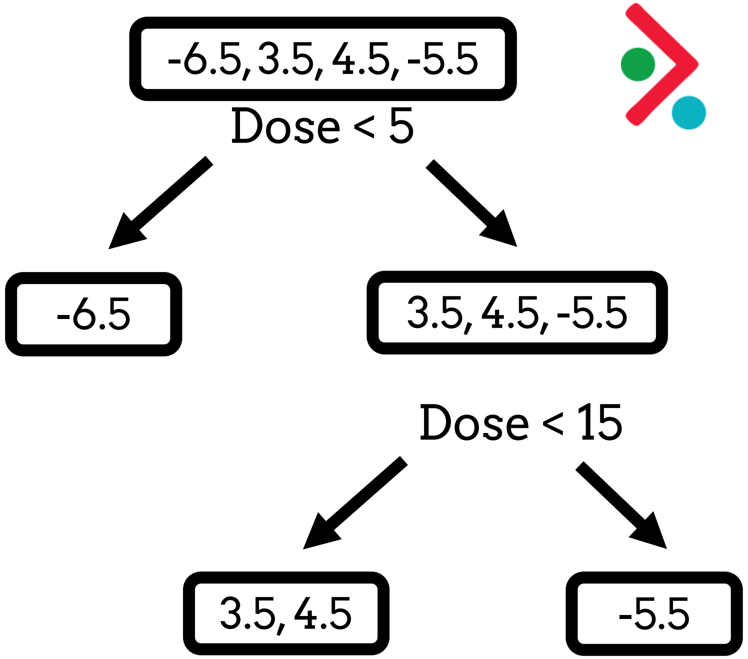
$$predição = \frac{-6.5}{1+0} = -6.5$$

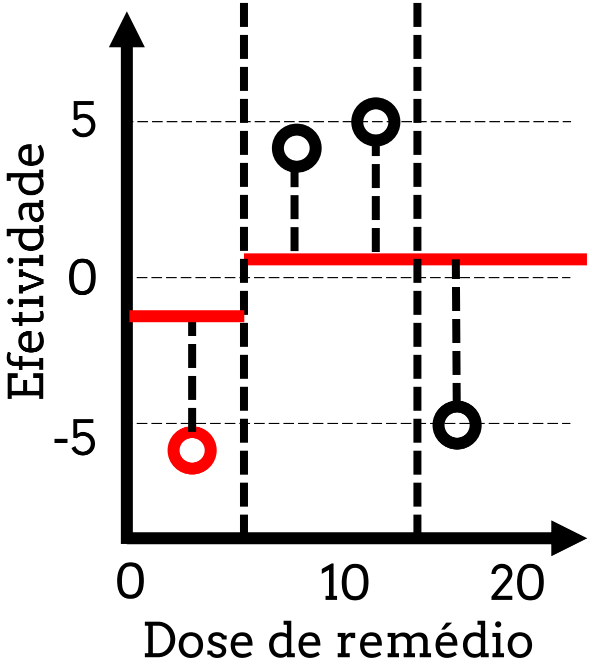
$$predição = \frac{\sum resíduos}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

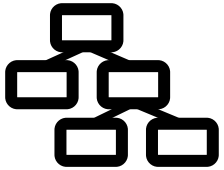
$f(x) = 0.5 + 0.3 \times$ 

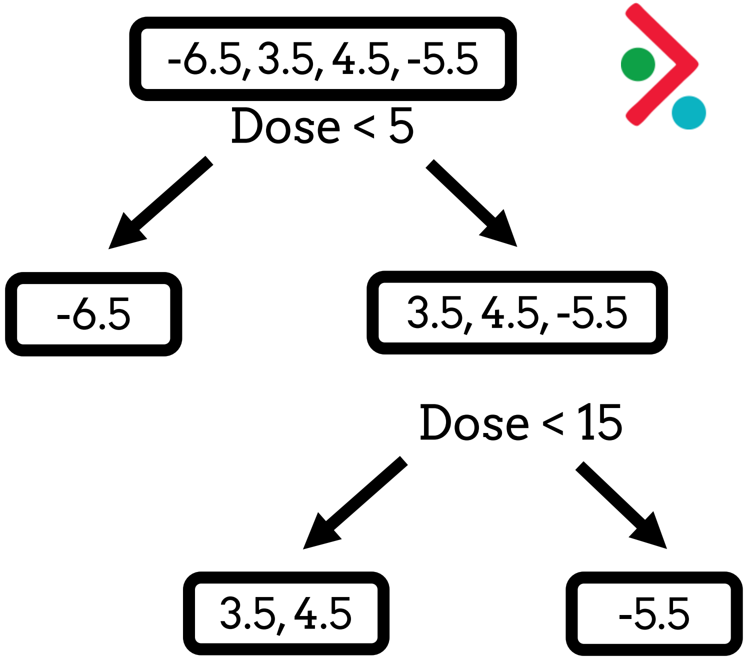
$f(2) = 0.5 + 0.3 \times -6.5 = -1.45$





Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

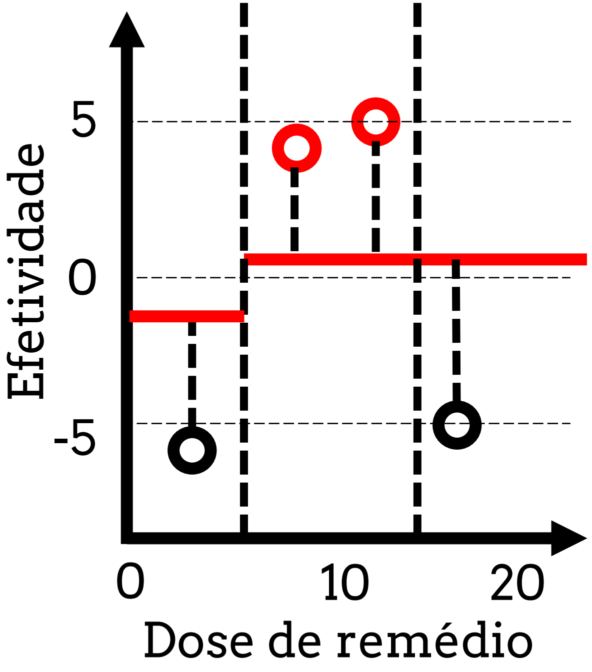
$f(x) = 0.5 + 0.3 \times$ 



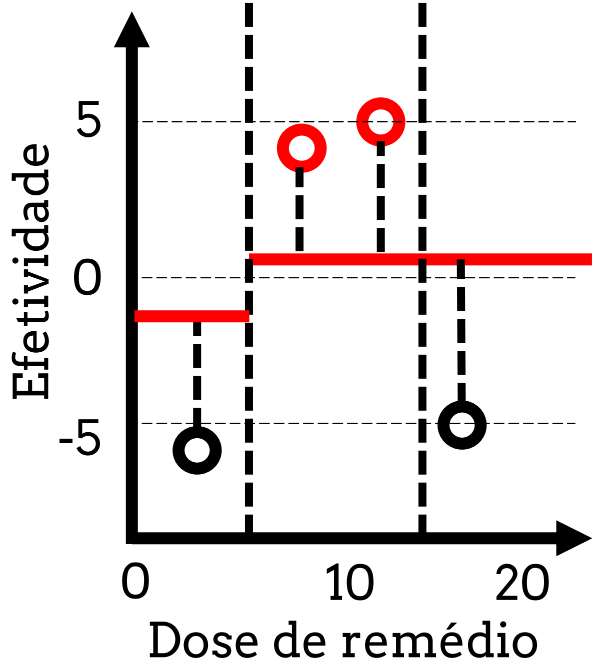
$f(2) = 0.5 + 0.3 \times -6.5 = -1.45$

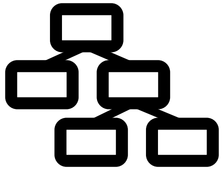
$f(x_2) = 0.5 + 0.3 \times$

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



Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

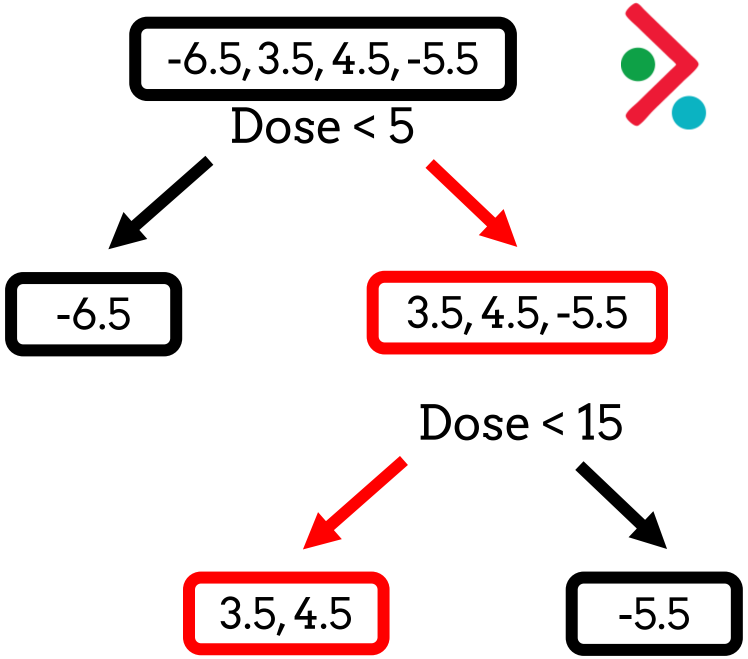


$f(x) = 0.5 + 0.3 \times$ 

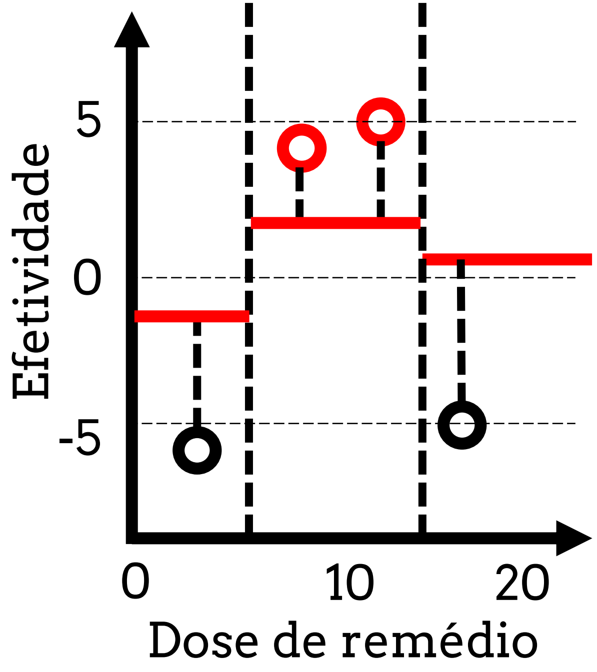
$f(2) = 0.5 + 0.3 \times -6.5 = -1.45$

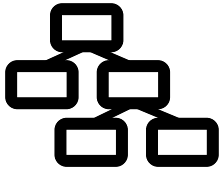
$f(8) = 0.5 + 0.3 \times$

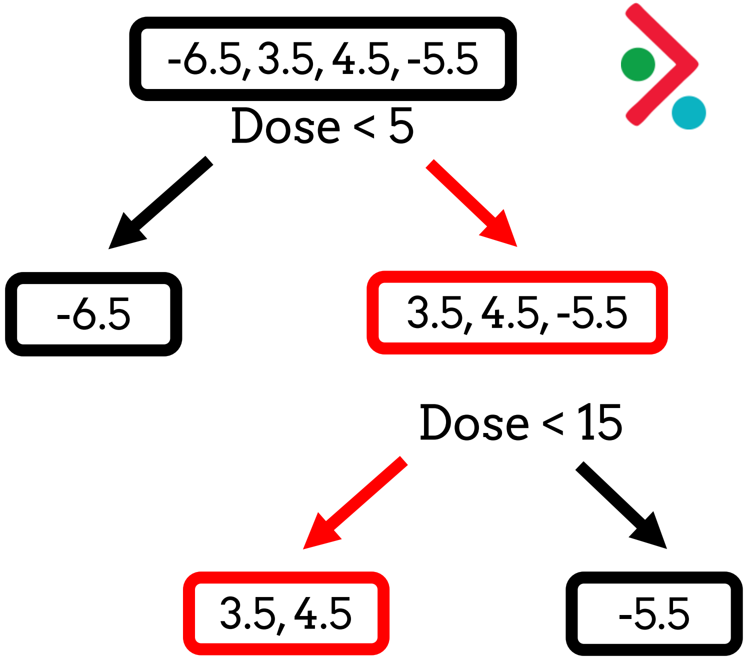
$f(12) = 0.5 + 0.3 \times$



Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2



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



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

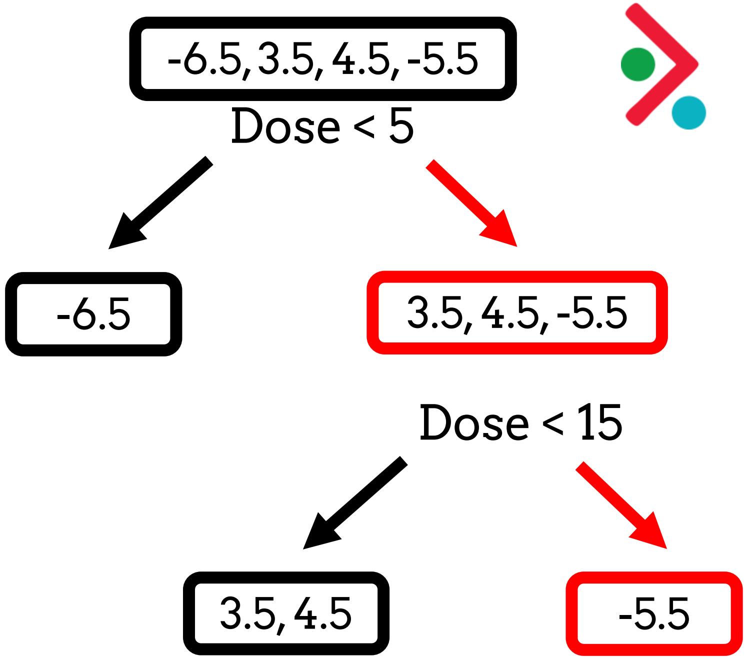
$f(8) = 0.5 + 0.3 \times 4 = 1.7$

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

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

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

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

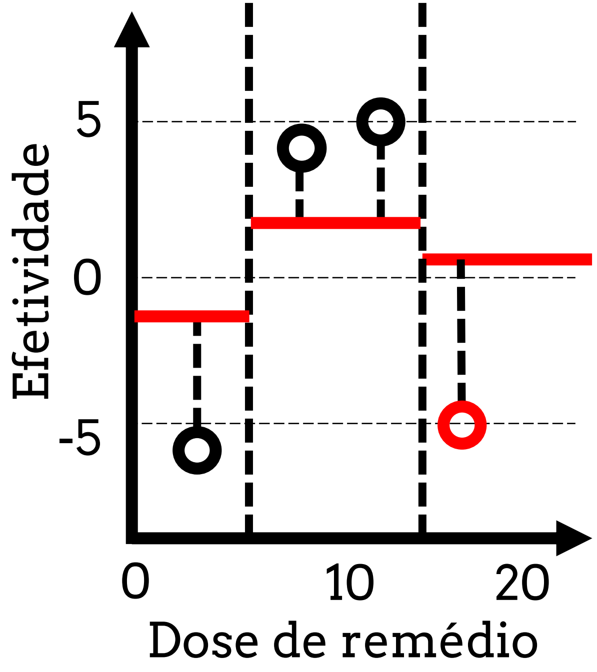


$f(2) = 0.5 + 0.3 \times -6.5 = -1.45$

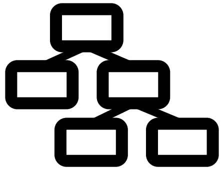
$f(8) = 0.5 + 0.3 \times 4 = 1.7$

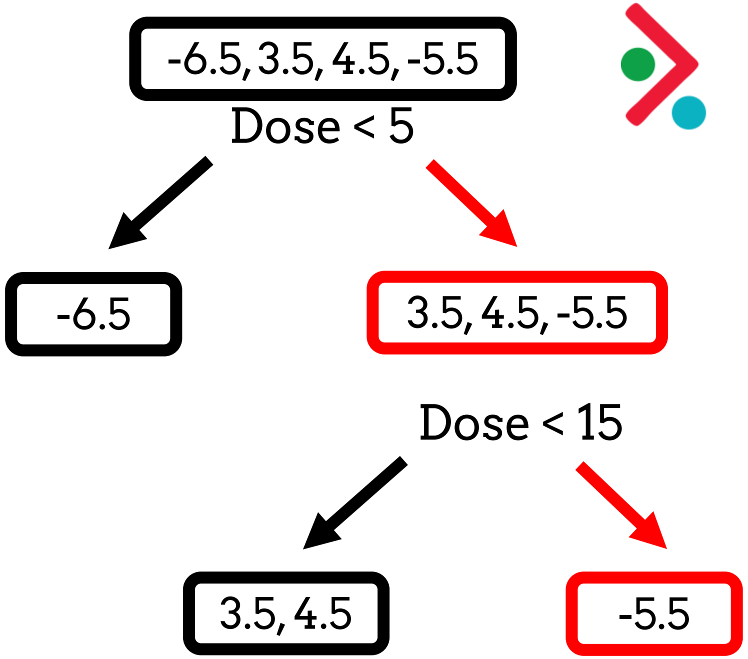
$f(12) = 0.5 + 0.3 \times 4 = 1.7$

$f(16) = 0.5 + 0.3 \times$



Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

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



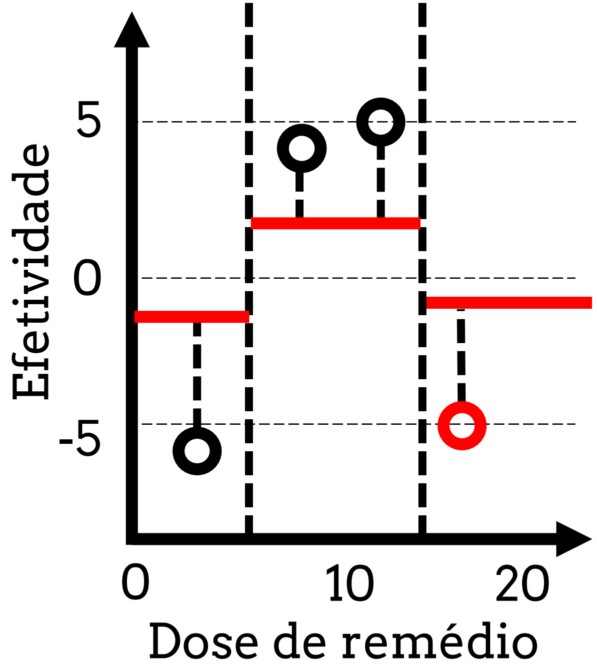
$f(2) = 0.5 + 0.3 \times -6.5 = -1.45$

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

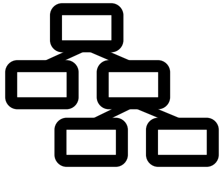
$f(12) = 0.5 + 0.3 \times 4 = 1.7$

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

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

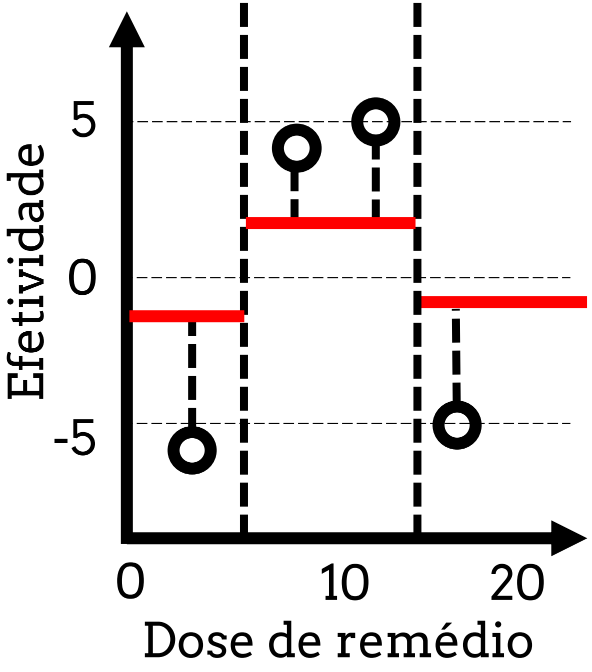


Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

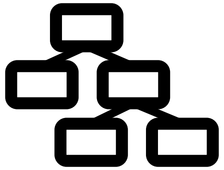
$$f(x) = {\color{red}0.5} + 0.3 \times$$ 

$f(2) = {\color{red}0.5} + 0.3 \times -6.5 = -1.45$
$f(8) = {\color{red}0.5} + 0.3 \times 4 = 1.7$
$f(12) = {\color{red}0.5} + 0.3 \times 4 = 1.7$
$f(16) = {\color{red}0.5} + 0.3 \times -5.5 = -1.15$

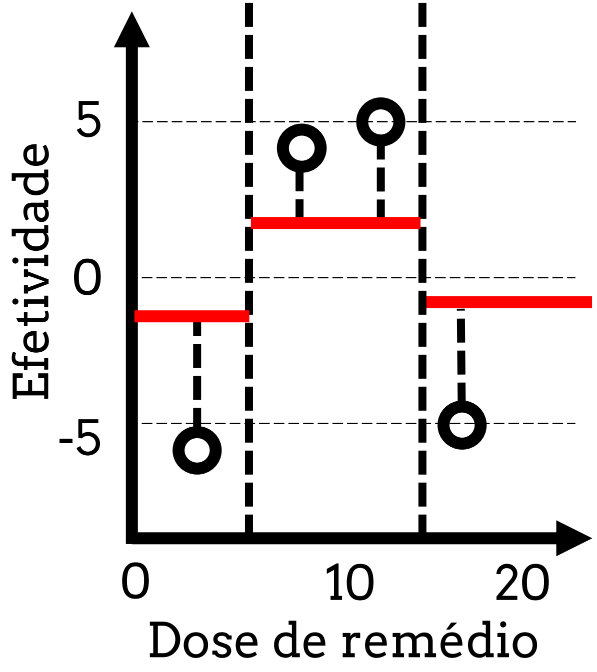
resíduo1 = $-6 - (-1.56) = -4.44$
resíduo2 = $4 - 1.7 = 2.3$
resíduo3 = $5 - 1.7 = 3.3$
resíduo4 = $-5 - (-1.15) = -3.85$



Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

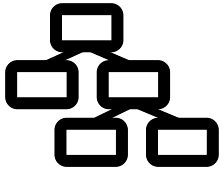
$f(x) = 0.5 + 0.3 \times$ 

-4.44,2.3,3.3,-3.85

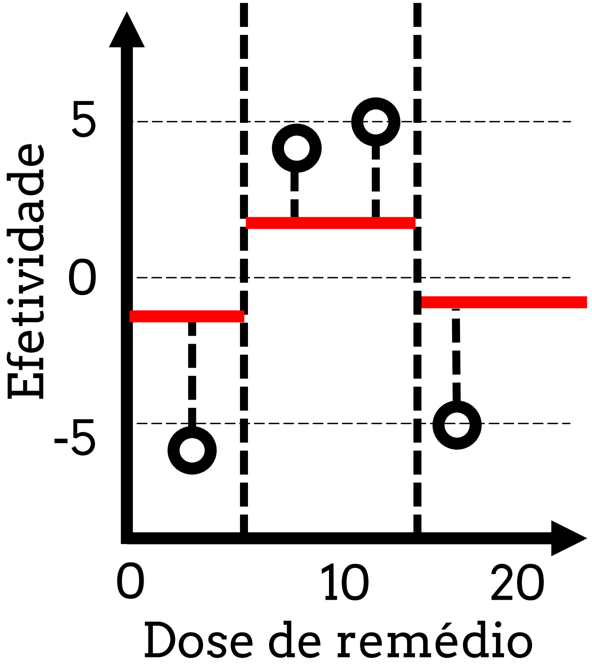


Hora da segunda árvore

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

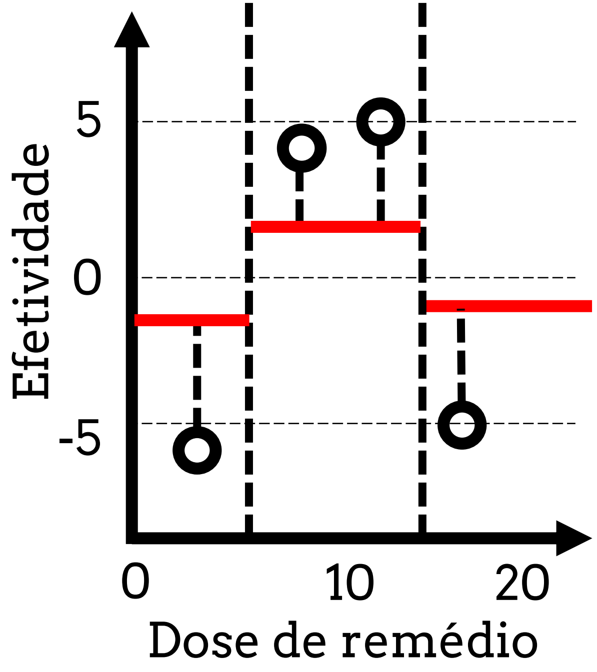
$f(x) = 0.5 + 0.3 \times$ 

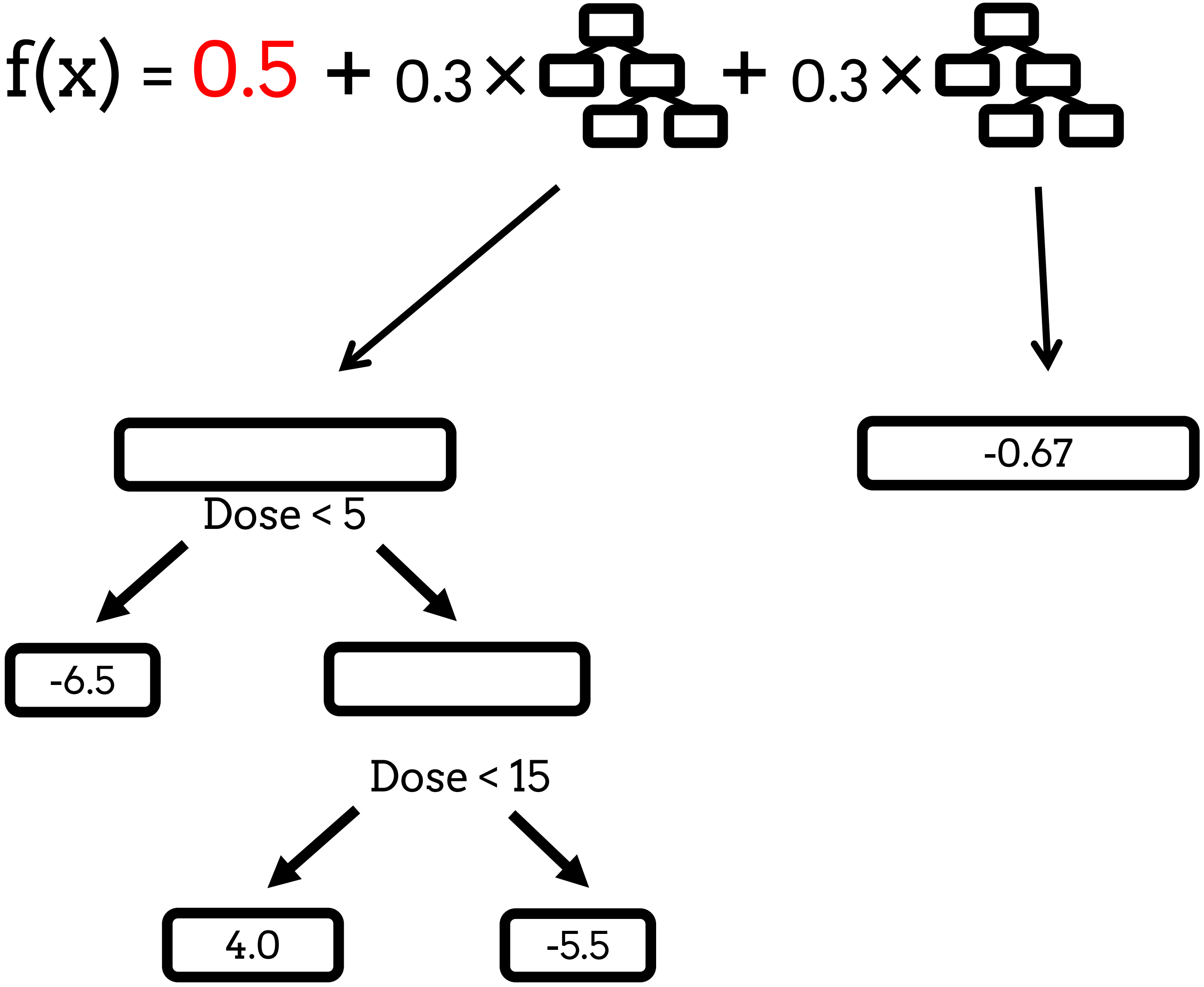
-4.44,2.3,3.3,-3.85



(sim salamin...)

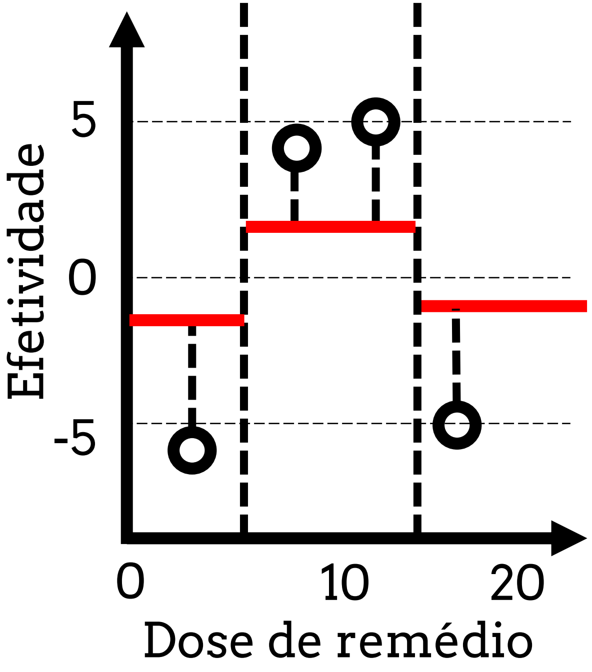
Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2



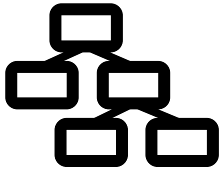


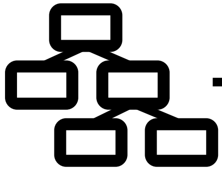
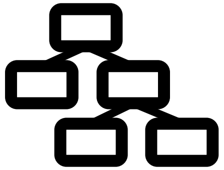
Modelo Final!

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2

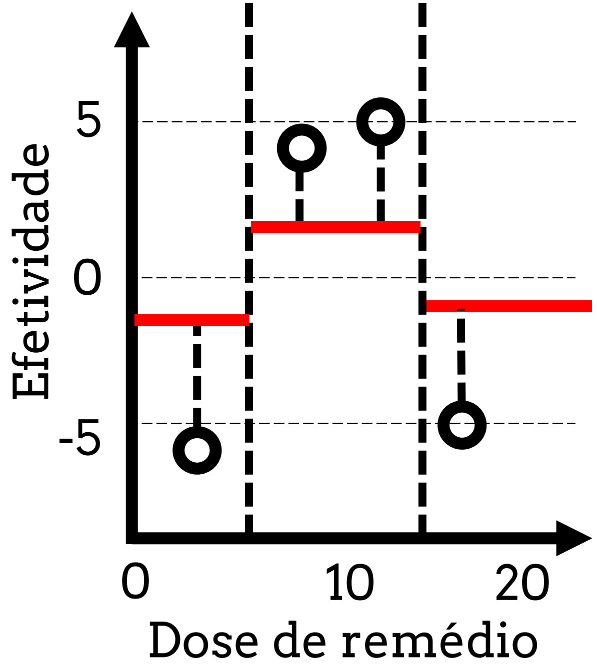


$f1(x) = 0.5$

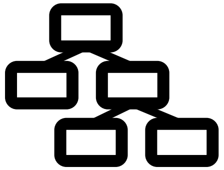
$f2(x) = 0.5 + 0.3 \times$ 

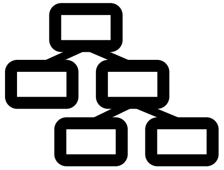
$f3(x) = 0.5 + 0.3 \times$  $+ 0.3 \times$ 

Hiperparam	valor
λ	0
γ	50
ε	0.3
Tree Depth	2
Trees	2



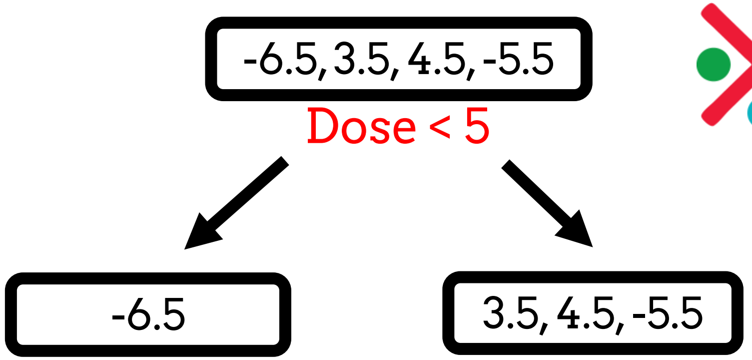
$f1(x) = 0.5$

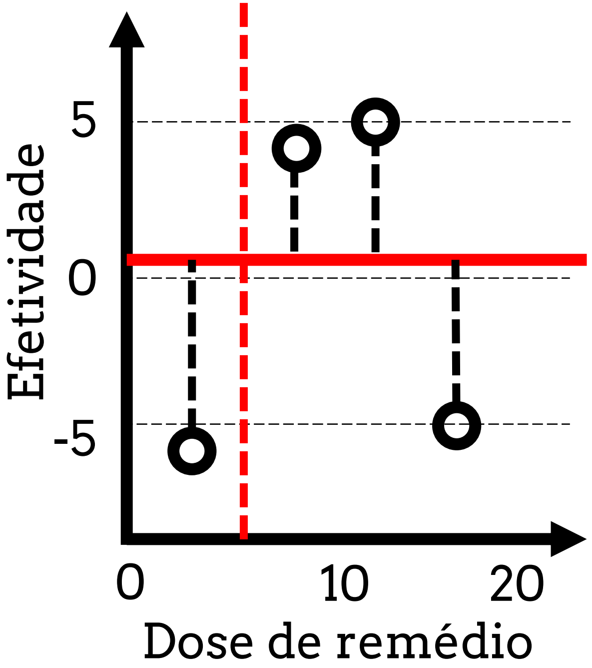
$f2(x) = f1(x) + 0.3 \times$ 

$f3(x) = f2(x) + 0.3 \times$ 

Hiperparam	valor
λ	1
γ	20
ε	0.3
Tree Depth	2
Trees	2

$f(x) = 0.5$





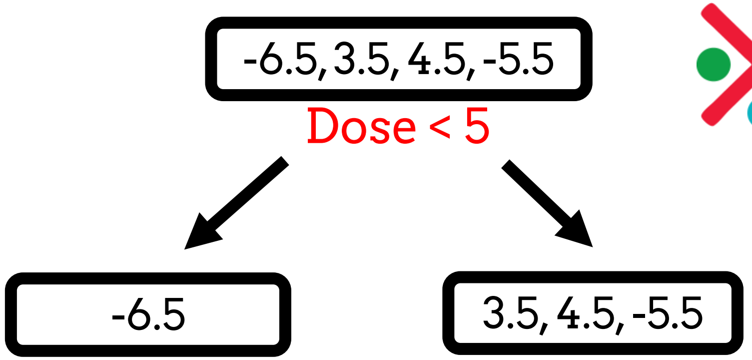
Pergunta	λ=0	λ=1
Dose < 5	40.33	
Dose < 10	1	
Dose < 15	27	

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

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

Hiperparam	valor
λ	1
γ	20
ε	0.3
Tree Depth	2
Trees	2

$$f(x) = 0.5$$

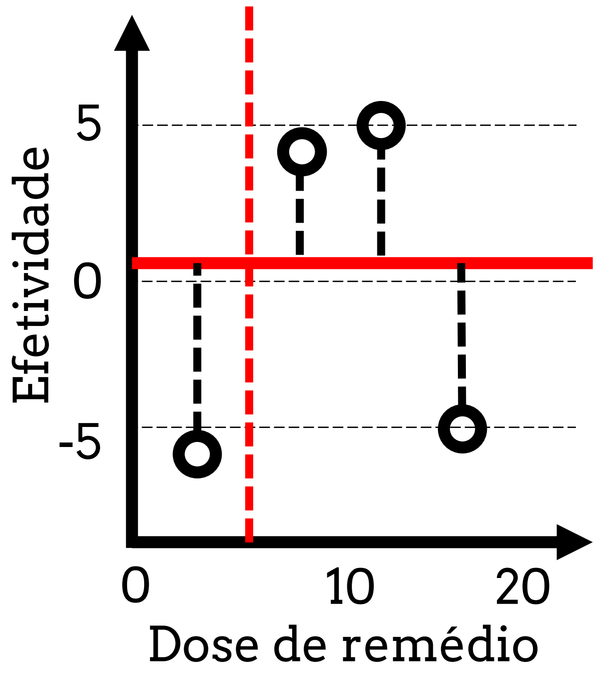


$$Similaridade_{pai} = \frac{(-6.5 + 3.5 + 4.5 - 5.5)^2}{4 + 1} = 3.2$$

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

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

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



Pergunta	λ=0	λ=1
Dose < 5	40.33	21.05
Dose < 10	1	
Dose < 15	27	

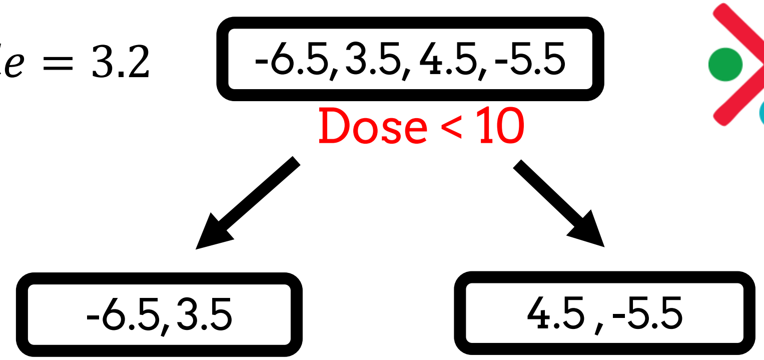
$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

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

Hiperparam	valor
λ	1
γ	20
ε	0.3
Tree Depth	2
Trees	2

$$f(x) = 0.5$$

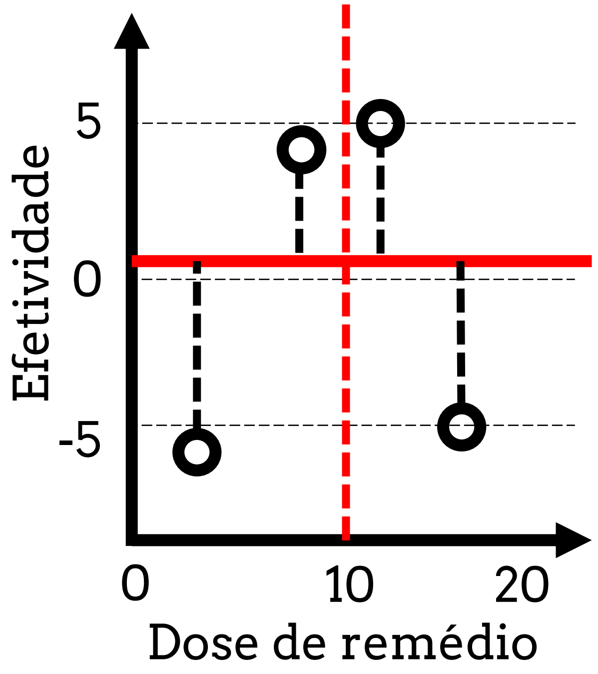
$$Similaridade = 3.2$$



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

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

$$Gain = 3 + 0.33\ - 3.2$$



Pergunta	λ=0	λ=1
Dose < 5	40.33	21.05
Dose < 10	1	0.13
Dose < 15	27	

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

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

Hiperparam	valor
λ	1
γ	20
ε	0.3
Tree Depth	2
Trees	2

$f(x) = 0.5$

$Similaridade = 3.2$ -6.5, 3.5, 4.5, -5.5

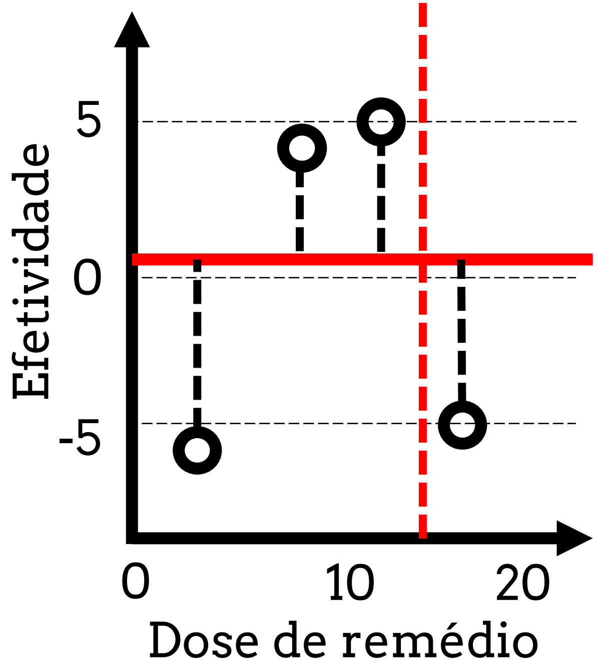
Dose < 15

-6.5, 3.5, 4.5 | -5.5

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

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

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



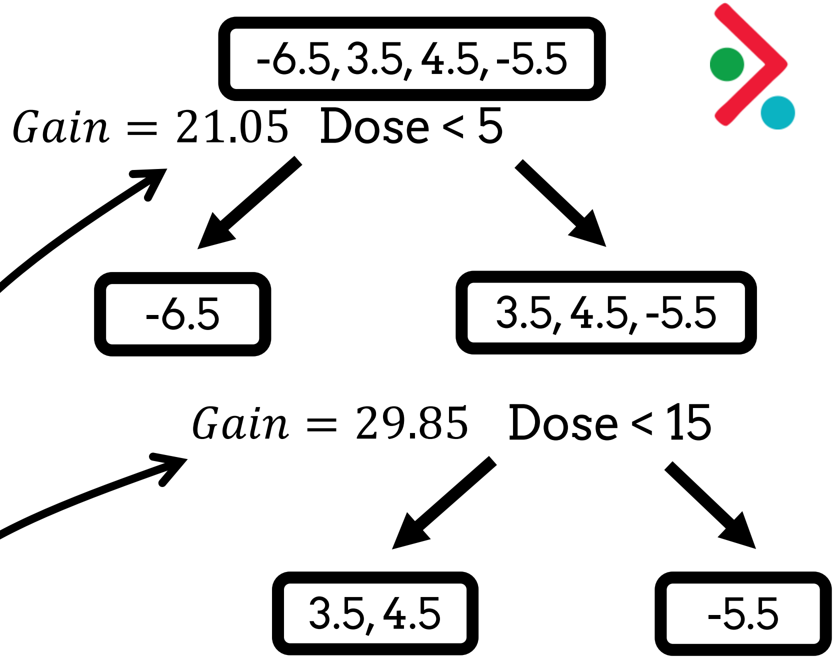
Pergunta	λ=0	λ=1
Dose < 5	40.33	21.05
Dose < 10	1	0.13
Dose < 15	27	12.48

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

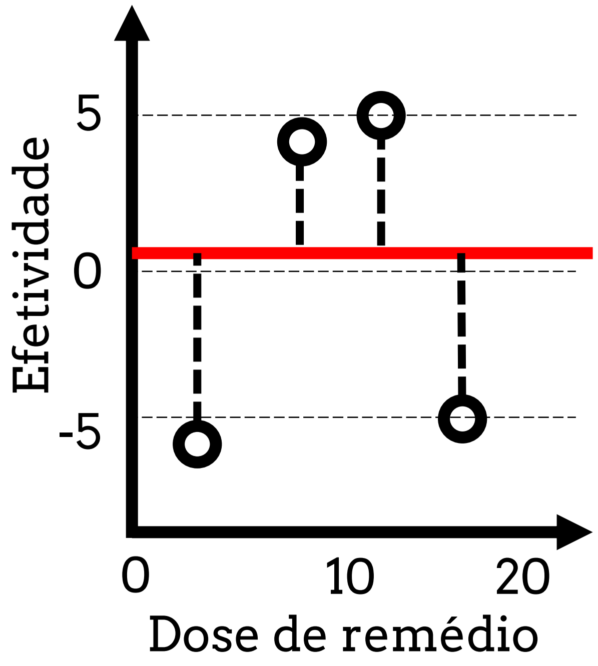
$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	1
γ	20
ε	0.3
Tree Depth	2
Trees	2

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



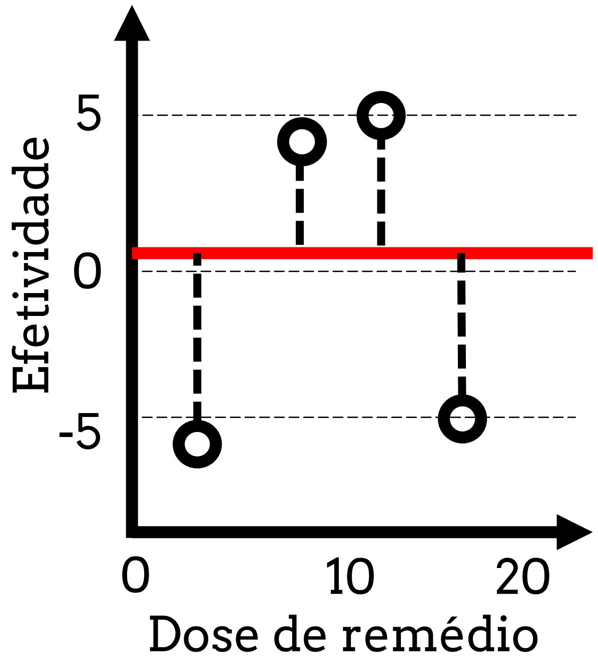
Gains menores,
mais fáceis de podar!

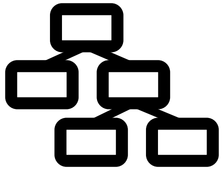


$$Gain = Sim_{esq} + Sim_{dir} - Sim_{pai}$$

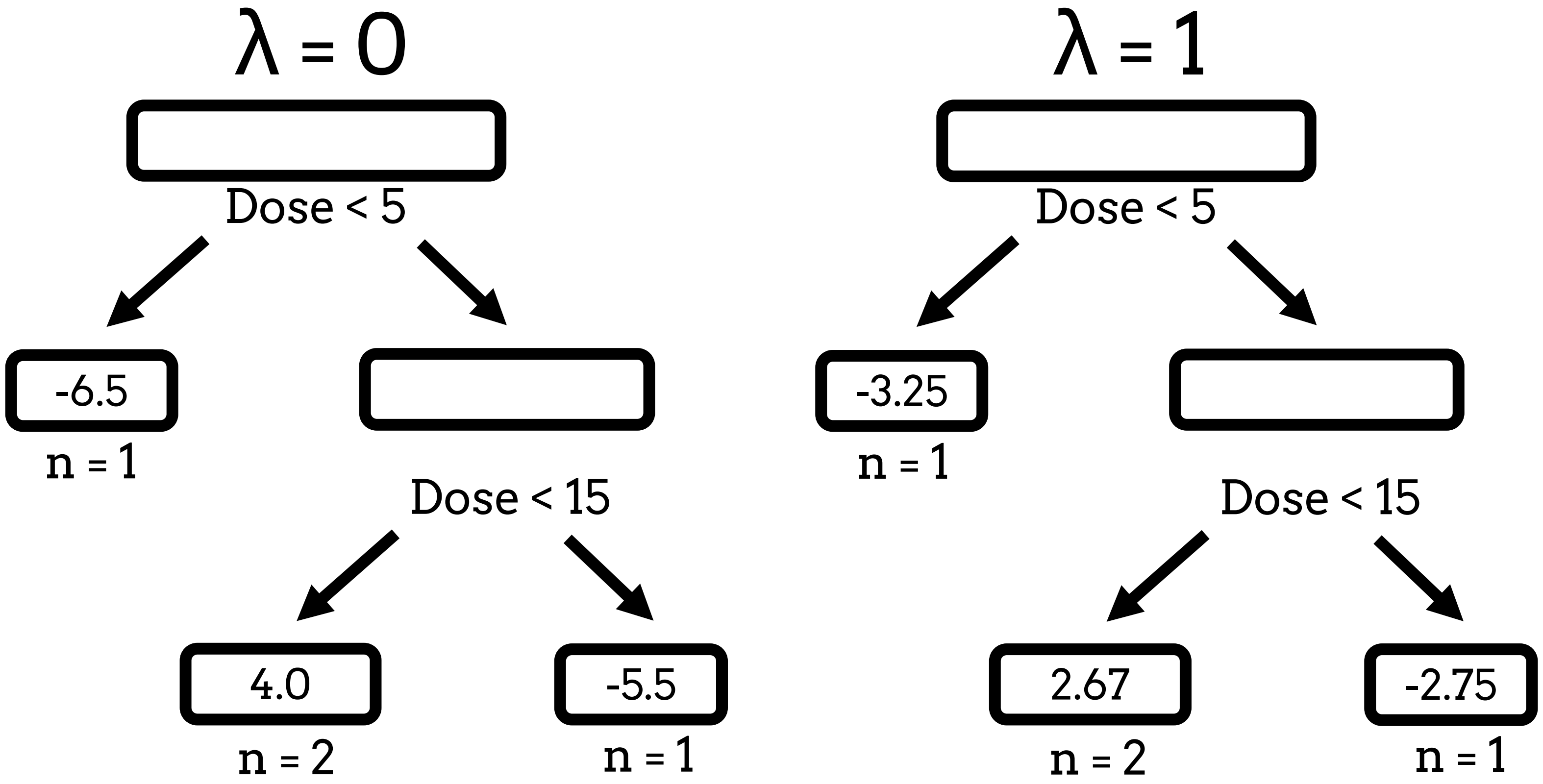
$$Similaridade = \frac{(\sum resíduos)^2}{\#resíduos + \lambda}$$

Hiperparam	valor
λ	1
γ	20
ε	0.3
Tree Depth	2
Trees	2



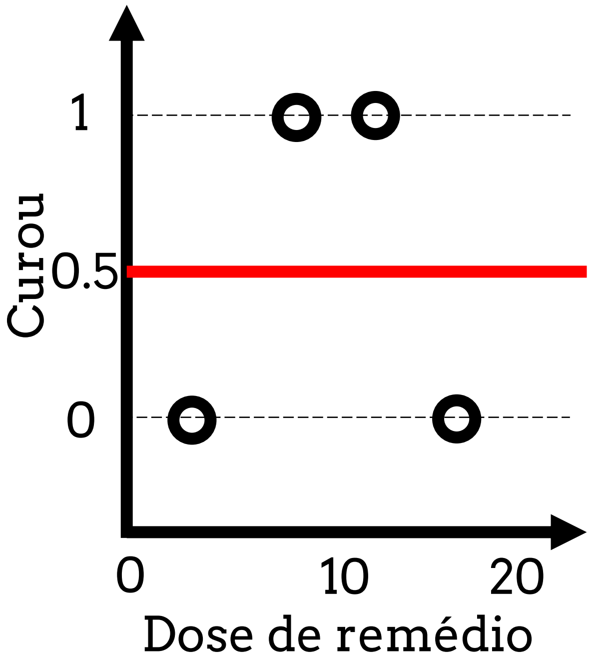
$f(x) = 0.5 + 0.3 \times$ 

Além disso, os scores também diminuíram...



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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2

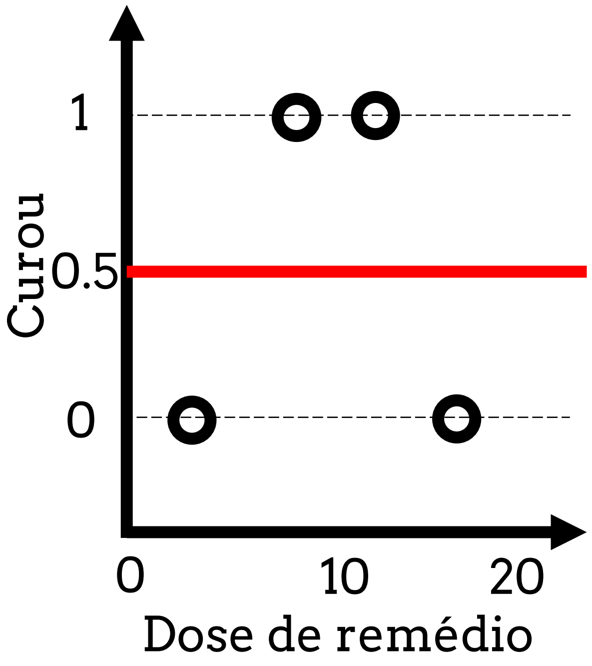


Regressão	Classificação
$\frac{(\sum resíduos)^2}{\#resíduos + \lambda}$	$\frac{(\sum resíduos)^2}{\sum p(1-p) + \lambda}$
$\frac{\sum resíduos}{\#resíduos + \lambda}$	$\frac{\sum resíduos}{\sum p(1-p) + \lambda}$
$f(x) = 0.5 +$	$\log\left(\frac{f(x)}{1 - f(x)}\right) = 0.0 +$

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2

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

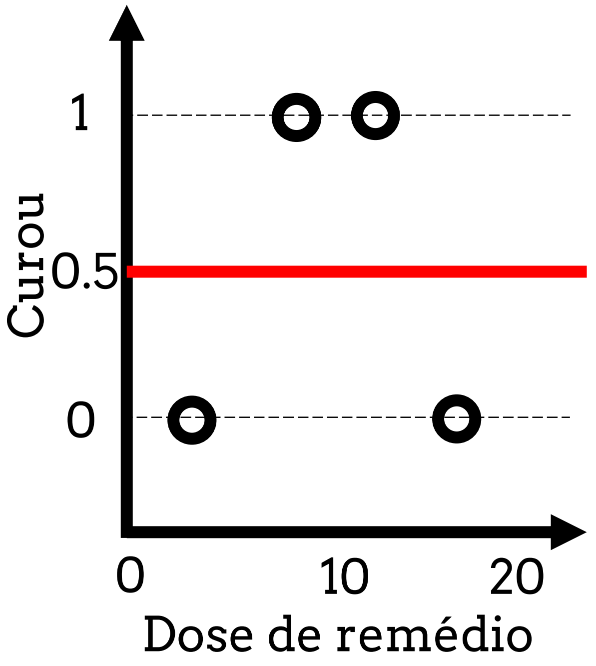
No caso de classificação, vamos trocar f() por p() para relacionar com o fato de que estamos calculando probabilidades.



Hiperparam	valor
λ	0
γ	20
ε	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.



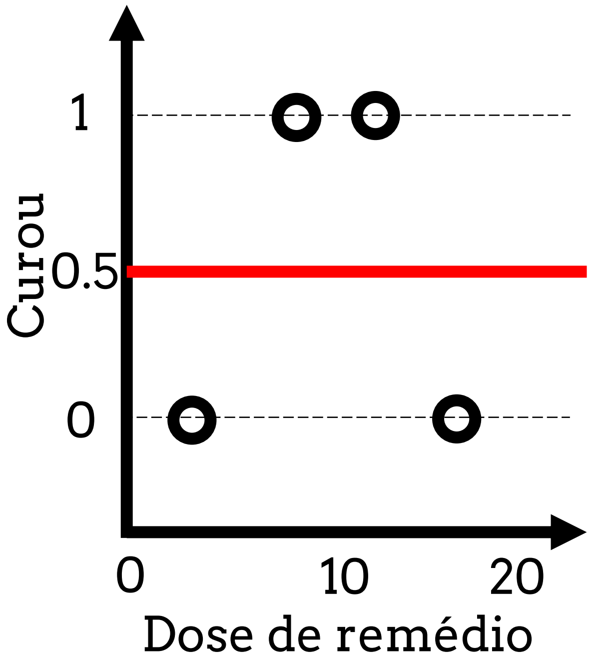
E uma rápida revisão sobre as função logística:

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

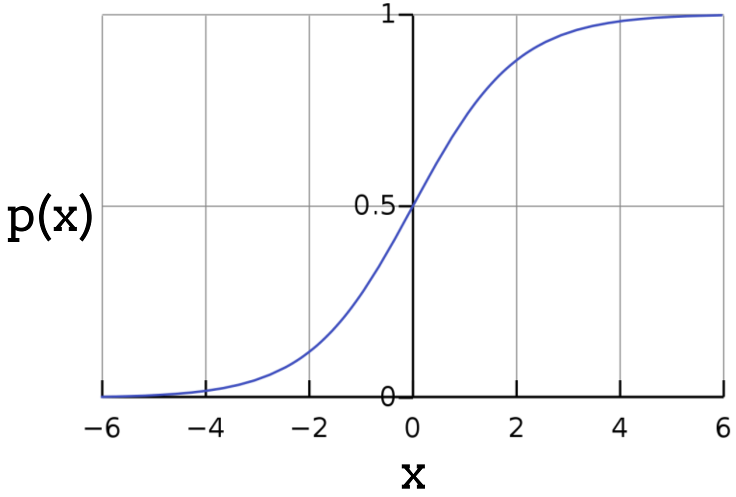
Logaritmo da chance,
ou log-odds,
ou logit

Hiperparam	valor
λ	0
γ	20
ε	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 as função logística:

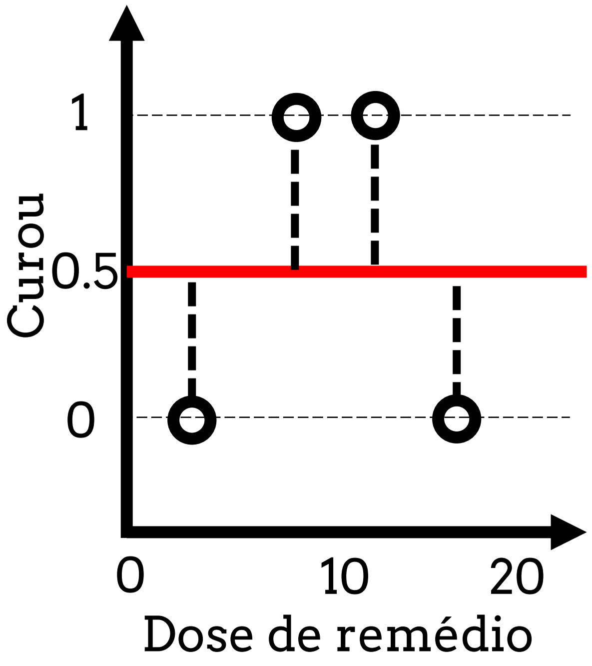
$$\log\left(\frac{p(x)}{1 - p(x)}\right) = x \quad \xrightarrow{\text{inversa}} \quad p(x) = \frac{1}{1 + e^{-x}}$$

Logaritmo da chance,
ou log-odds,
ou logit

Função logística,
ou sigmoide

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2

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



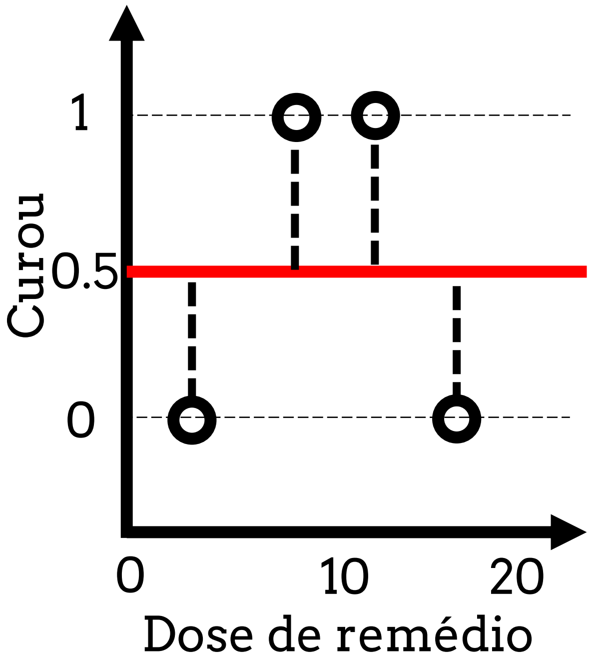
$$\text{resíduo} = y - p(x)$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2

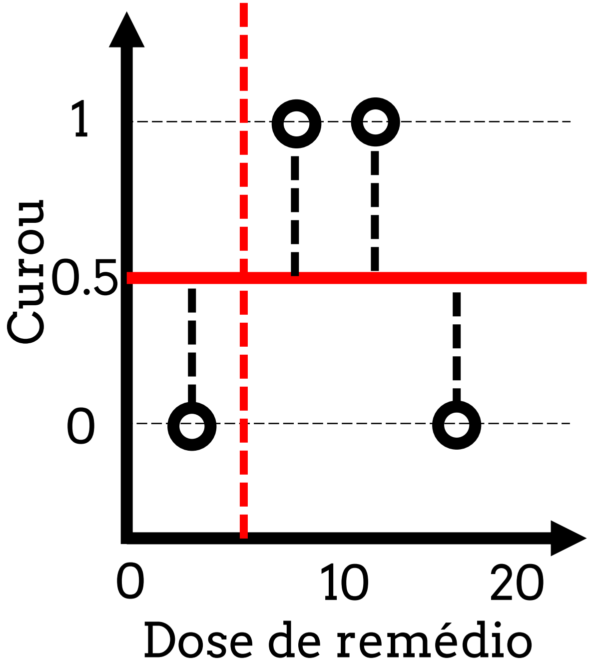
$$\log\left(\frac{p(x)}{1 - p(x)}\right) = 0.0$$

-0.5,0.5,0.5,-0.5



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

Hiperparam	valor
λ	0
γ	20
ε	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

Dose < 5

-0.5

0.5,0.5,-0.5

$$Similaridade_{pai} = \underline{\qquad\qquad\qquad} =$$

$$Similaridade_{esq} = \underline{\qquad\qquad\qquad} =$$

$$Similaridade_{dir} = \underline{\qquad\qquad\qquad} =$$

$$Gain =$$

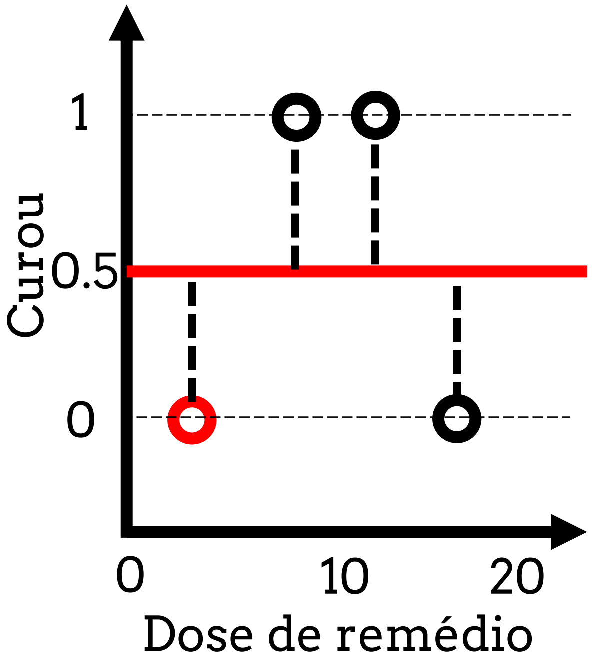
$$Similaridade = \frac{(\sum resíduos)^2}{\sum p(1-p) + \lambda}$$

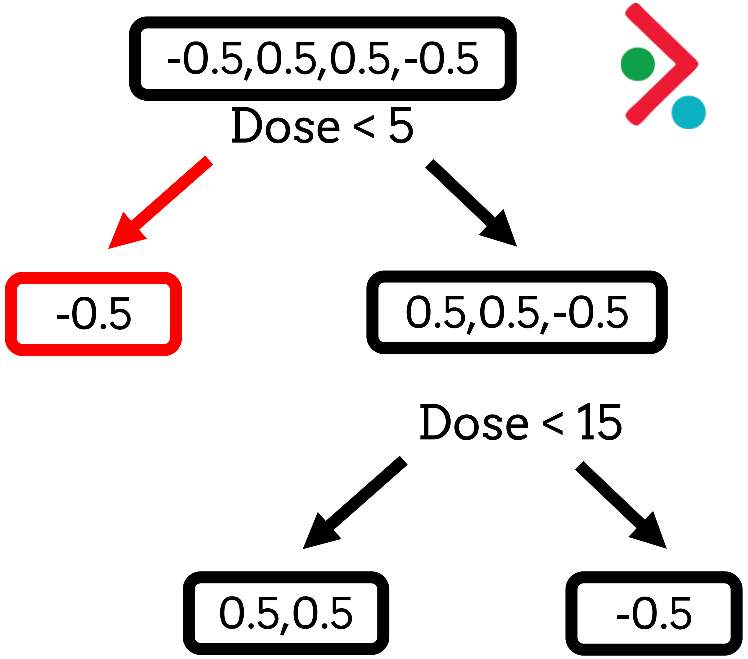
$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2

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

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

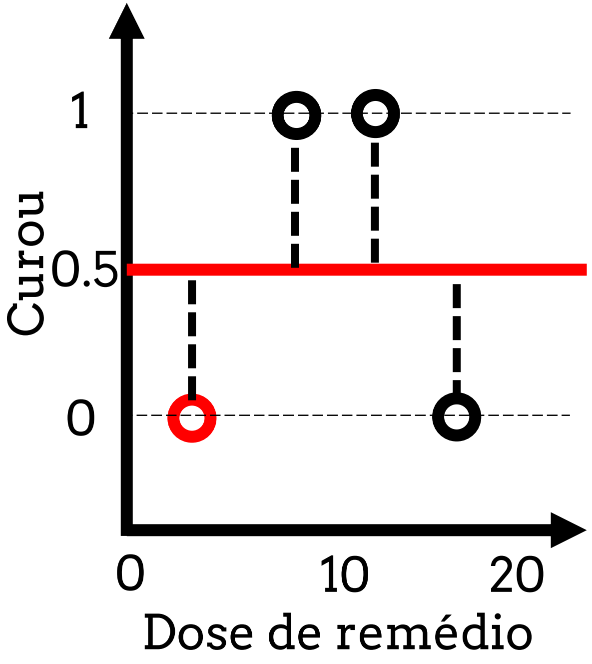




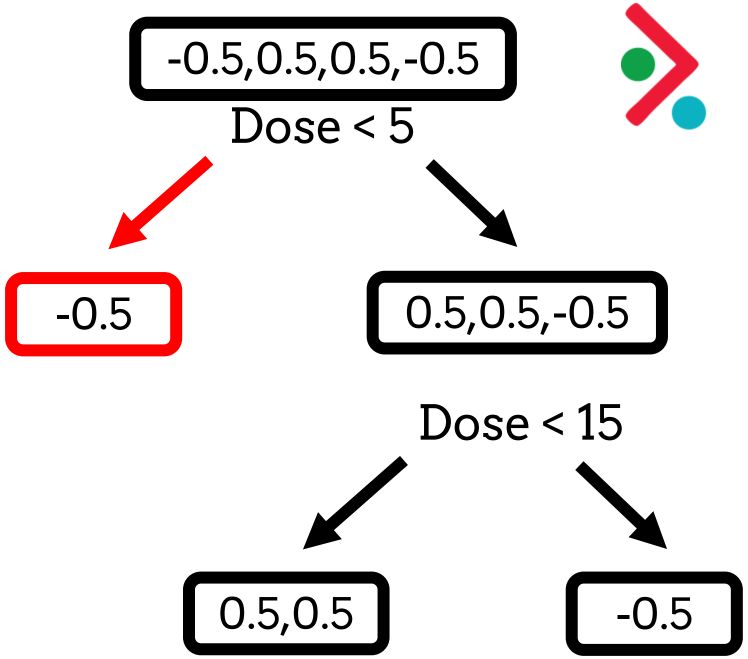
$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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



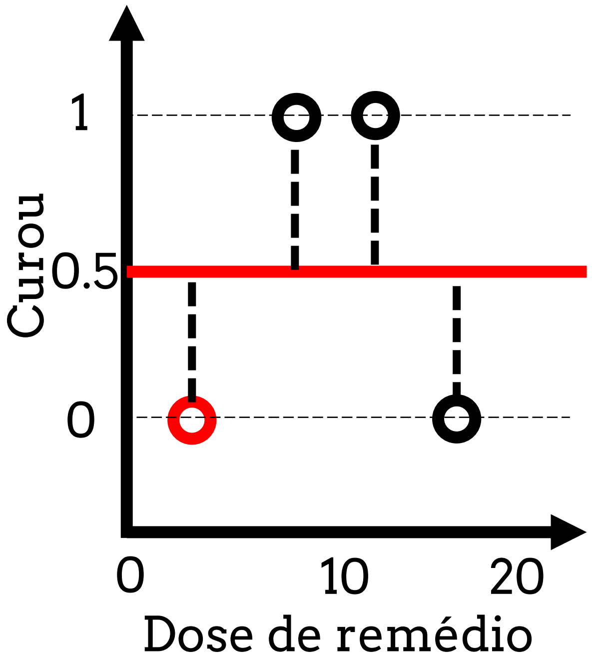
$$\log\left(\frac{p(2)}{1 - p(2)}\right) = 0.0 + 0.3\times(-2)$$

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

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

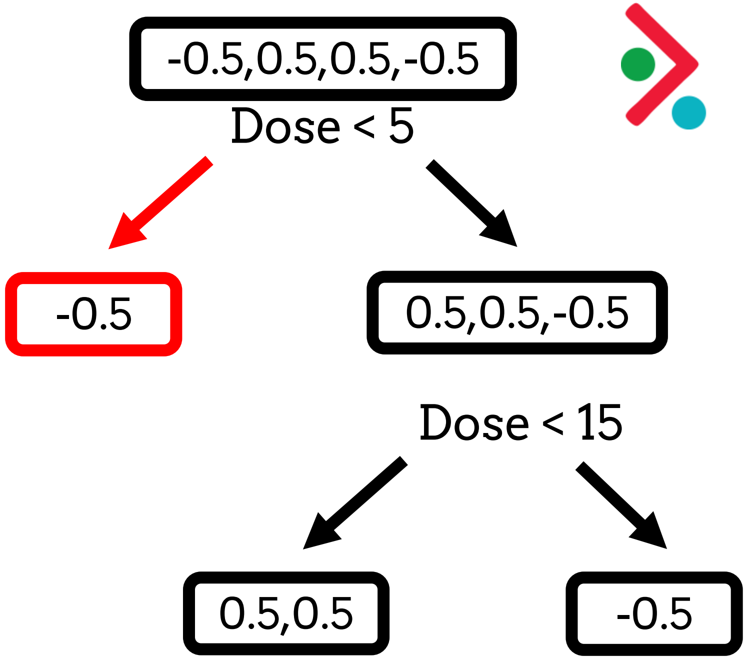
$$p(x) = \frac{1}{1 + e^{-x}}$$

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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

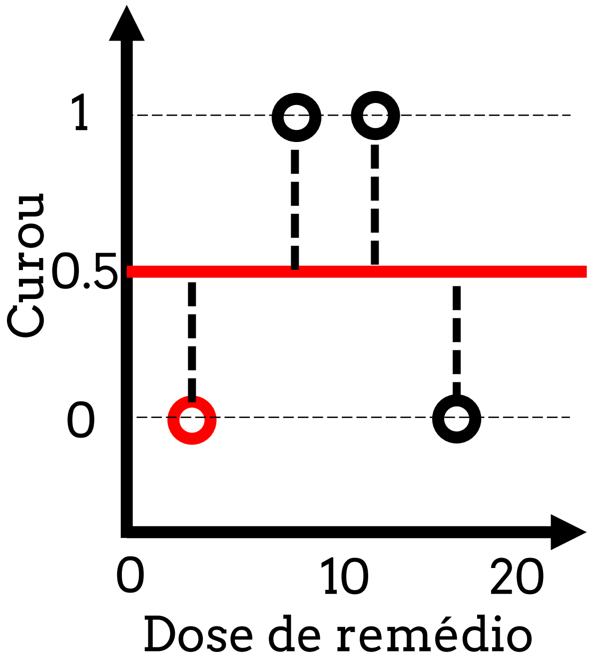
$$\log\left(\frac{p(2)}{1 - p(2)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$



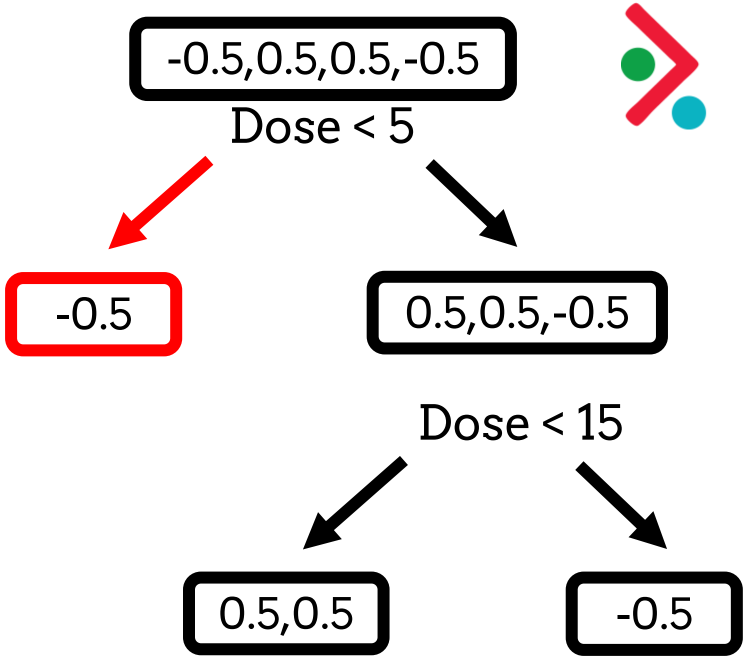
$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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



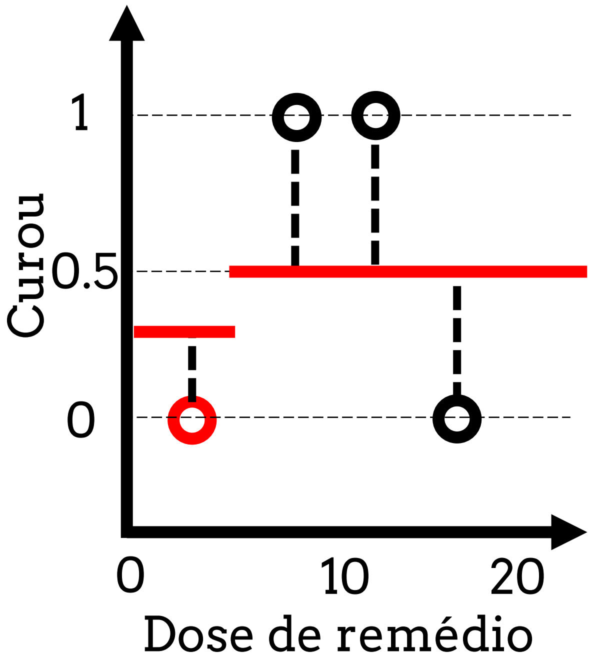
$$\log\left(\frac{p(2)}{1 - p(2)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$

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

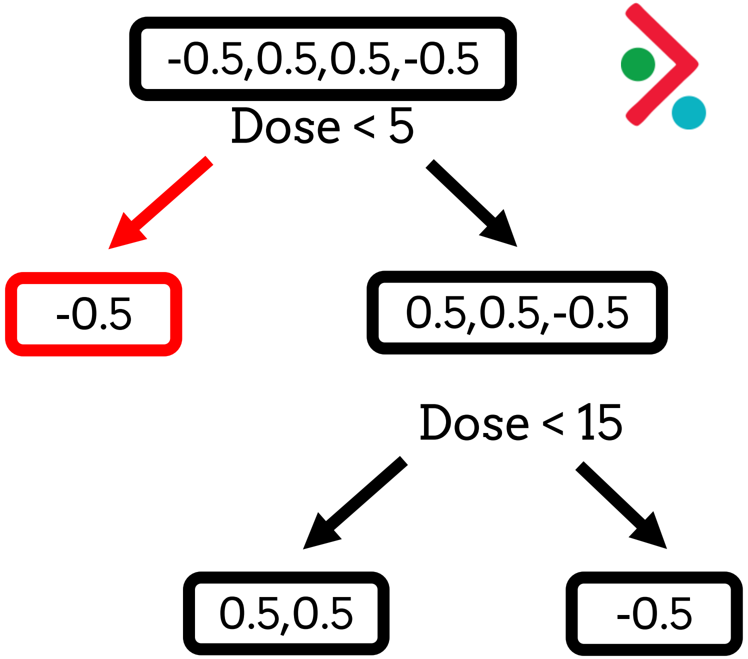
$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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



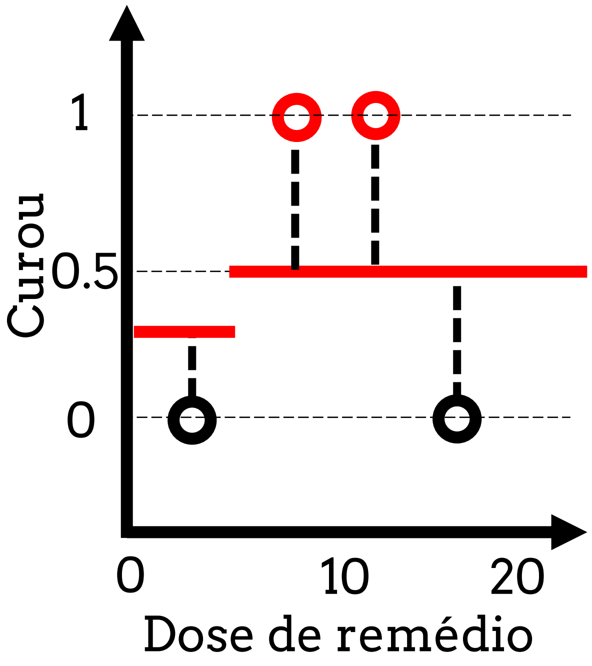
$$\log\left(\frac{p(2)}{1 - p(2)}\right) = 0.0 + 0.3 \times (-2) = -0.6$$

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

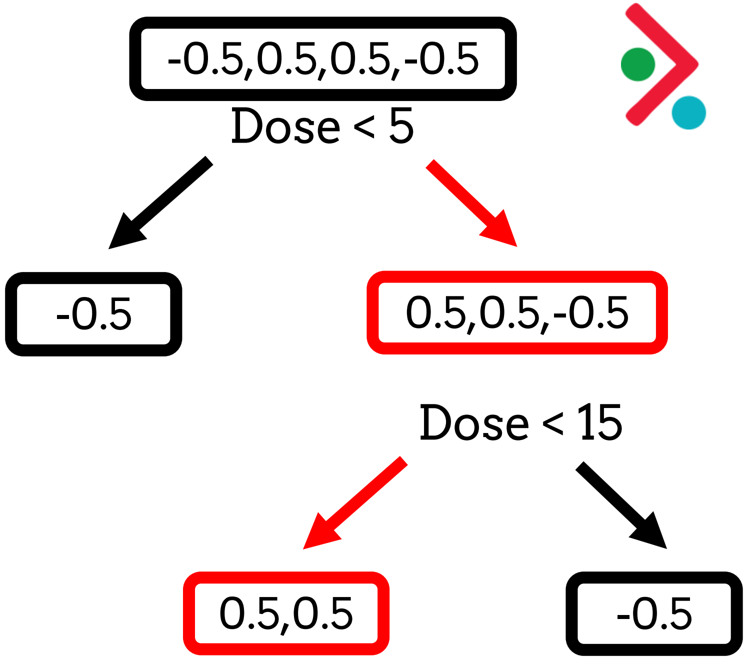
$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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



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

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

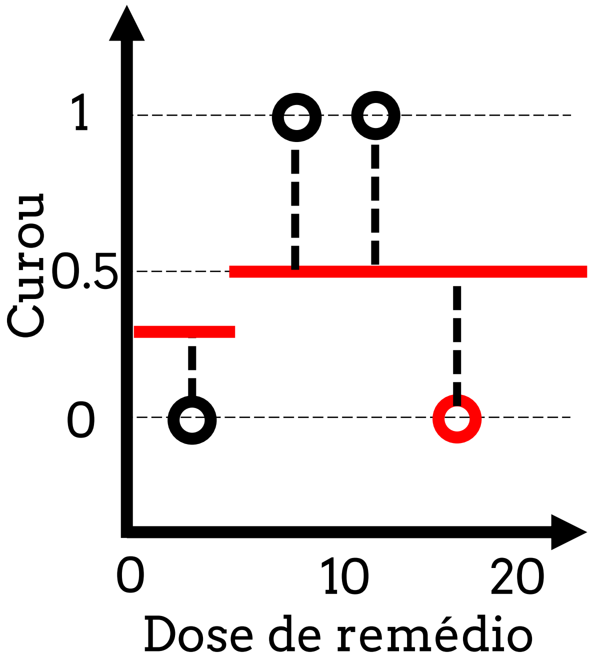
$$\log\left(\frac{p(12)}{1 - p(12)}\right) = 0.0 + 0.3\times 2$$

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

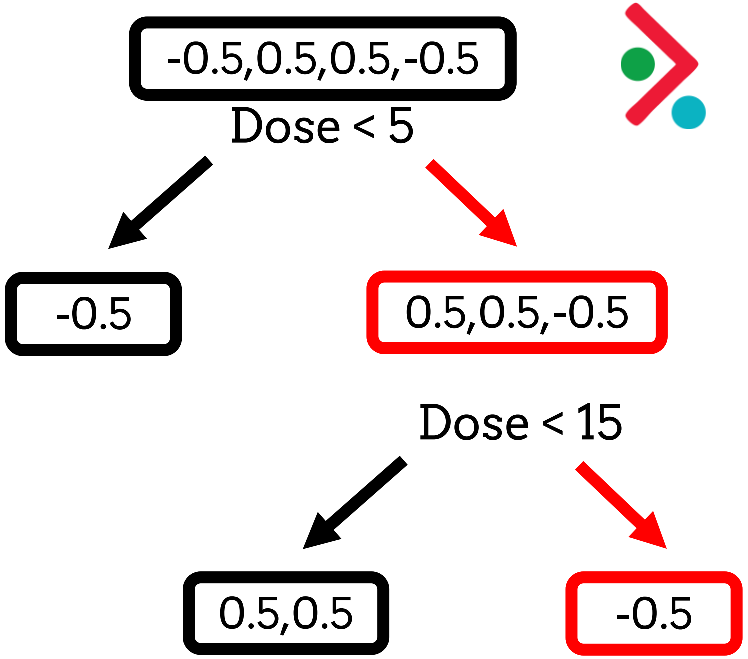
$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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



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

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

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

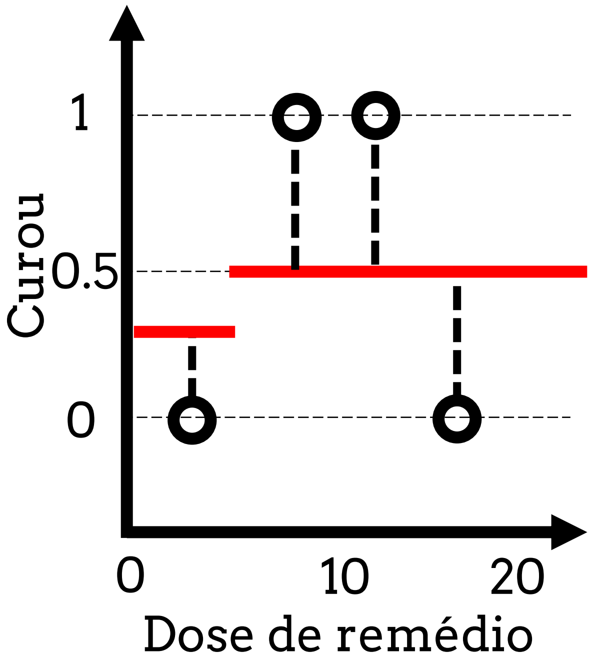
$$\log\left(\frac{p(16)}{1-p(16)}\right) = 0.0 + 0.3\times(-2)$$

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

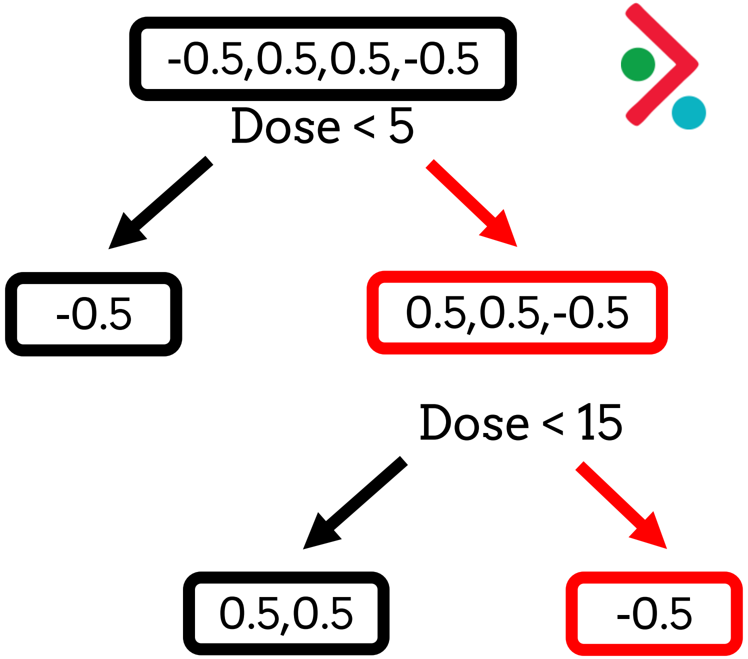
$$predição = \frac{\sum resíduos}{\sum p(1-p)+\lambda}$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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



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

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

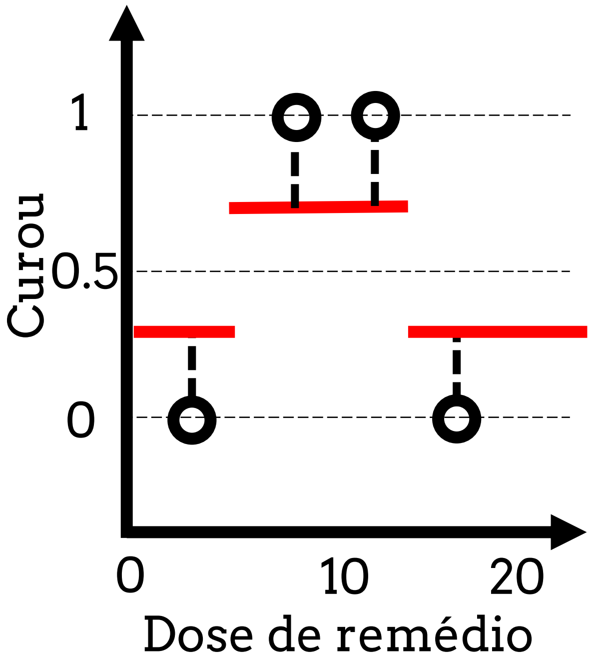
$$\log\left(\frac{p(12)}{1 - p(12)}\right) = 0.0 + 0.3 \times 2 = 0.6$$

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

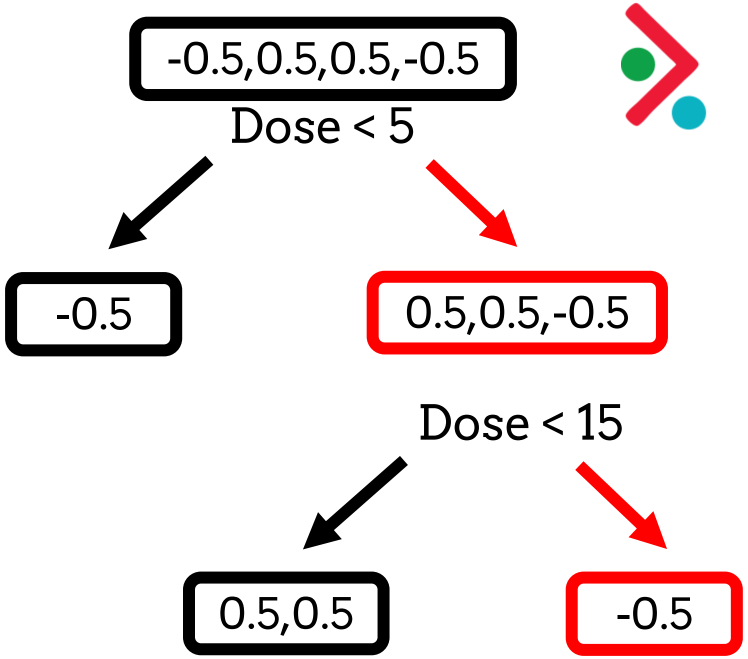
$$predição = \frac{\sum resíduos}{\sum p(1-p) + \lambda}$$

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

Hiperparam	valor
λ	0
γ	20
ε	0.3
Tree Depth	2
Trees	2



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



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

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

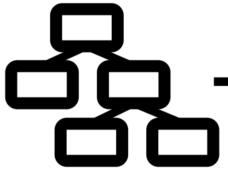
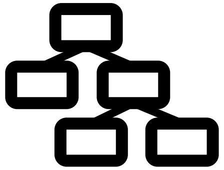
$$p(12) = \frac{1}{1 + e^{-(0.6)}} = 0.65$$

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

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

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

Dose de remédio	Curou	Pred
2	Não	0.256
8	Sim	0.744
12	Sim	0.744
16	Não	0.256

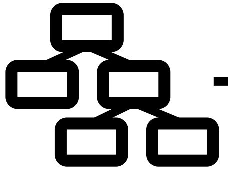
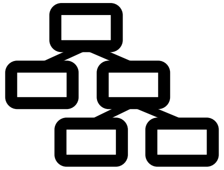
$f(x) = 0.5 + 0.3\times$  $+ 0.3\times$ 

$$\sum L(y_i, f(x_i))$$

$$\sum(y_i \log(f(x_i)) + (1 - y_i)\log(1 - f(x_i)))$$

Deviance
Regressão Logística
Binary Cross-entropy

Dose de remédio	Curou	Pred
2	Não	0.256
8	Sim	0.744
12	Sim	0.744
16	Não	0.256

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

Learning Task Parameters

Specify the learning task and the corresponding learning objective. The objective options are below:

- `objective` [default=reg:squarederror]
    - `reg:squarederror` : regression with squared loss.
    - `reg:squaredlogerror` : regression with squared log loss $\frac{1}{2}[log(pred+1) - log(label+1)]^2$. All input labels are required to be greater than -1. Also, see metric `rmsle` for possible issue with this objective.
    - `reg:logistic` : logistic regression
    - `binary:logistic` : logistic regression for binary classification, output probability
    - `binary:logitraw` : logistic regression for binary classification, output score before logistic transformation
    - `binary:hinge` : hinge loss for binary classification. This makes predictions of 0 or 1, rather than producing probabilities.
    - `count:poisson` –poisson regression for count data, output mean of poisson distribution
        - `max_delta_step` is set to 0.7 by default in poisson regression (used to safeguard optimization)
    - `survival:cox` : Cox regression for right censored survival time data (negative values are

$\sum L(y_i, f($

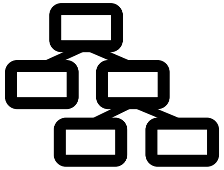
$\sum (y_i \log($

Últimos dois hiperparâmetros

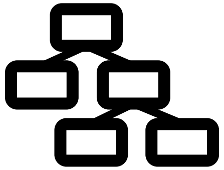
Hiperparam	valor
λ (regularization)	0
γ (loss_reduction)	0
ε (learn_rate)	0.3
tree depth	2
trees	2
sample_size	0.5
mtry	

sample_size: proporção de linhas sorteadas para cada árvore

Dose de remédio	Curou	Pred
2	Não	0.256
16	Não	0.256



Dose de remédio	Curou	Pred
2	Não	0.256
8	Sim	0.744

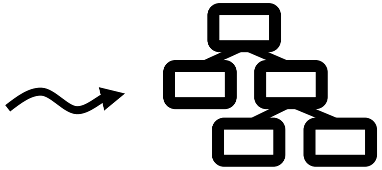


Últimos dois hiperparâmetros

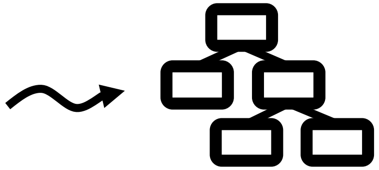
Hiperparam	valor
λ (regularization)	0
γ (loss_reduction)	0
ε (learn_rate)	0.3
tree depth	2
trees	2
sample_size	0.5
mtry	2

mtry: número de colunas sorteadas para cada árvore

x1	x3	Curou	Pred
2	3	Não	0.256
16	5	Não	0.256



x3	x16	Curou	Pred
2	3	Não	0.256
8	2	Sim	0.744



Exercício: 03-exercício-classif.R

Hiperparam	valor
λ (regularization)	0
γ (loss_reduction)	0
ε (learn_rate)	0.3
tree depth	2
trees	2

Dose de remédio	Curou	Pred
2	Não	0.256
8	Sim	0.744
12	Sim	0.744
16	Não	0.256