

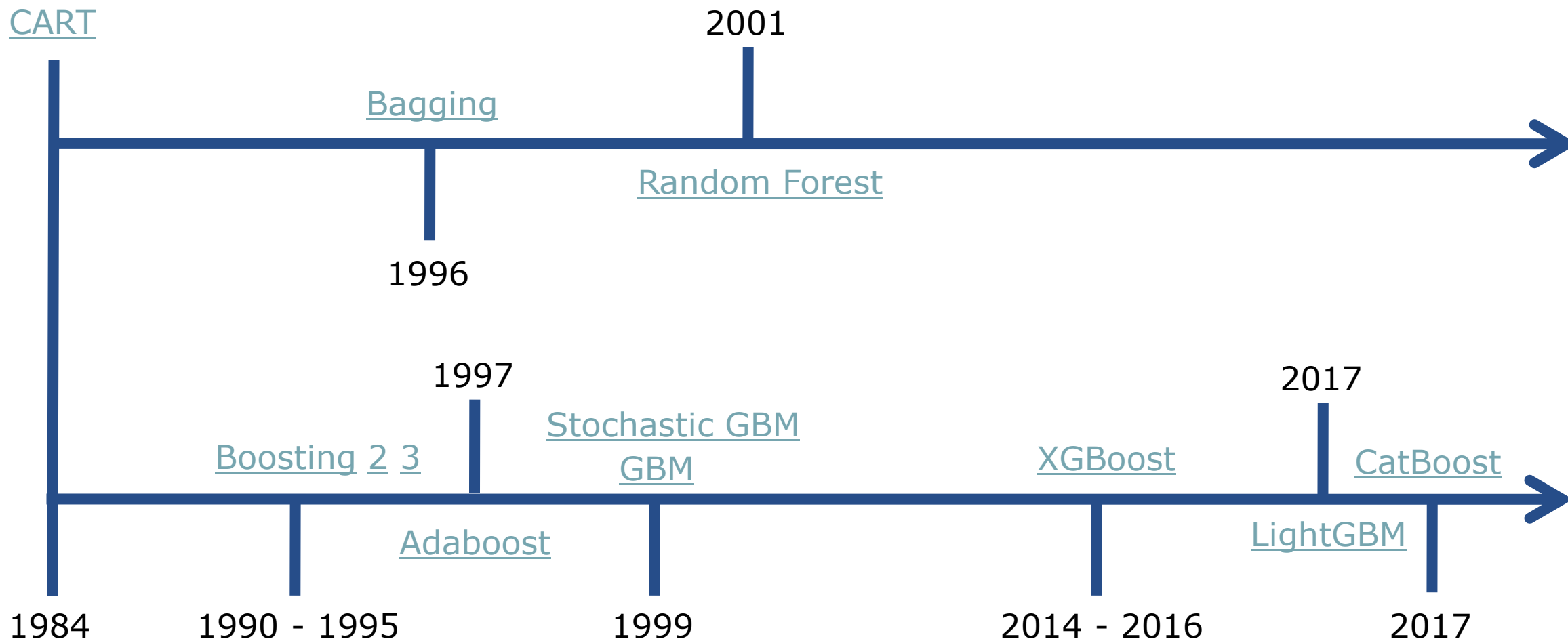


escola
britânica de
artes criativas
& tecnologia

Profissão Cientista de Dados
AdaBoost

Por que estamos aprendendo nessa ordem?

Pois essa foi a ordem das invenções.



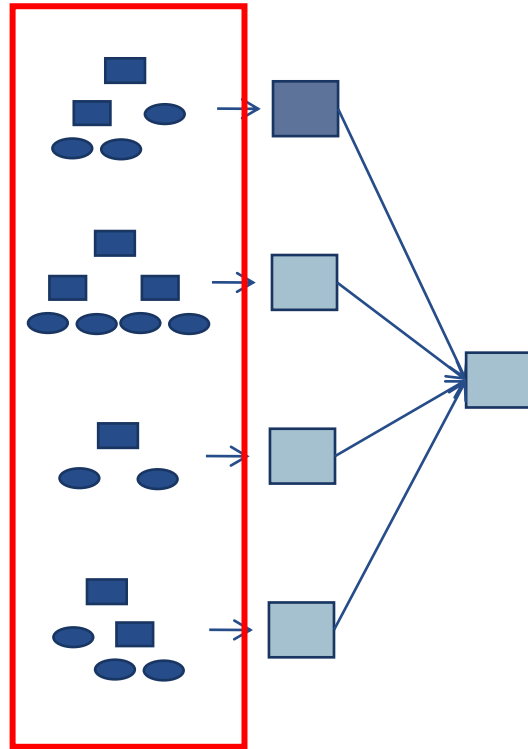
Adaptative Boosting - AdaBoost

[Artigo do AdaBoost \(Freund e Shapire\)](#)
[Scikit-learn - adaboost](#)

RF vs AdaBoost

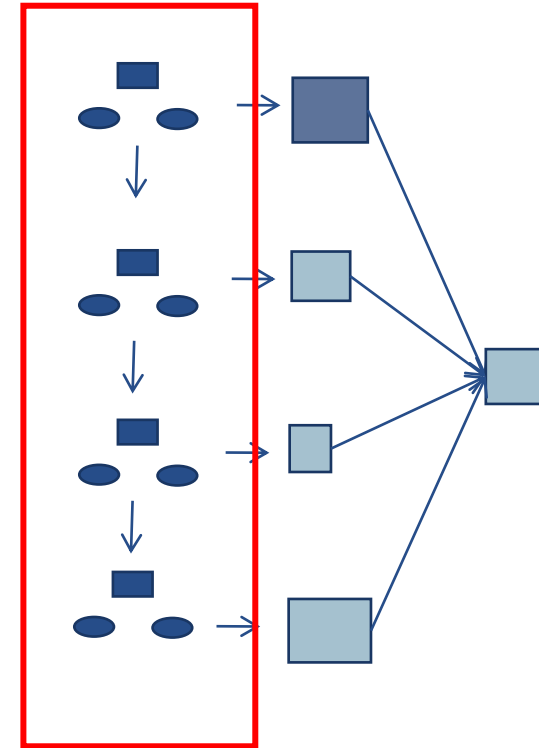
3 grandes diferenças

Random forest



Floresta de árvores

AdaBoost

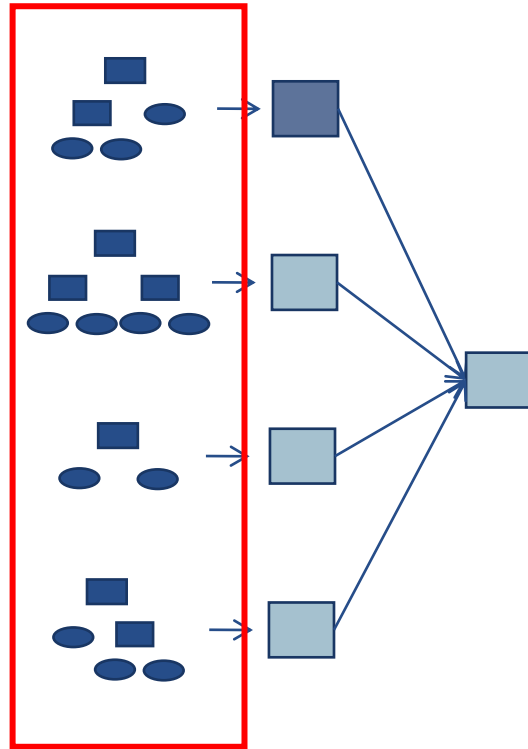


Floresta de Stumps

RF vs AdaBoost

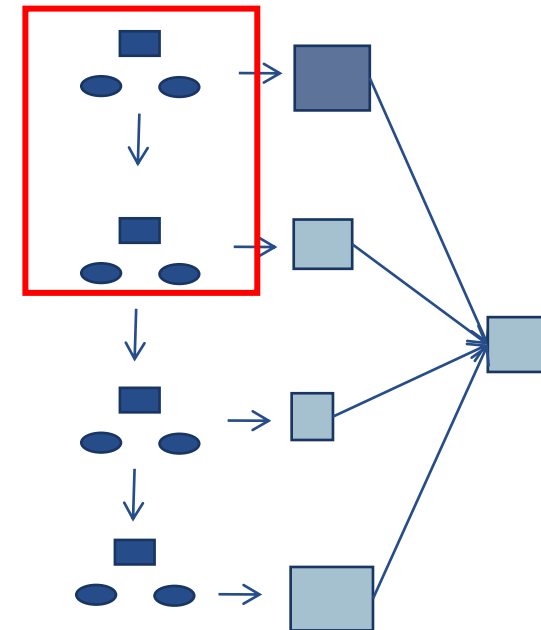
3 grandes diferenças

Random forest



Árvores independentes

AdaBoost

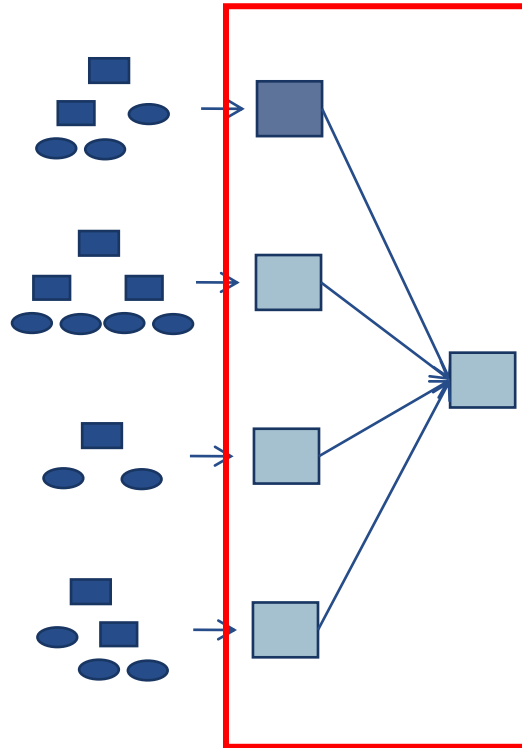


Uma árvore influencia na seguinte

RF vs AdaBoost

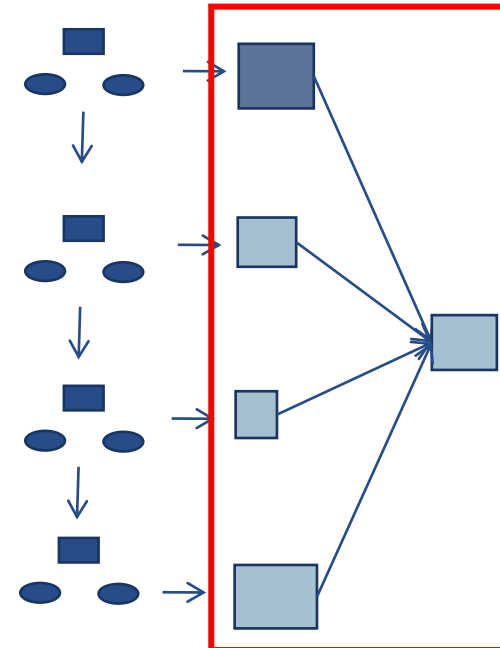
3 grandes diferenças

Random forest



Respostas das árvores
tem o mesmo peso

AdaBoost



Respostas tem pesos
diferentes

AdaBoost

Real AdaBoost

1. Start with weights $w_i = 1/N$, $i = 1, 2, \dots, N$.
 2. Repeat for $m = 1, 2, \dots, M$:
 - (a) Fit the classifier to obtain a class probability estimate $p_m(x) = \hat{P}_w(y = 1|x) \in [0, 1]$, using weights w_i on the training data.
 - (b) Set $f_m(x) \leftarrow \frac{1}{2} \log p_m(x)/(1 - p_m(x)) \in R$.
 - (c) Set $w_i \leftarrow w_i \exp[-y_i f_m(x_i)]$, $i = 1, 2, \dots, N$, and renormalize so that $\sum_i w_i = 1$.
 3. Output the classifier $\text{sign}[\sum_{m=1}^M f_m(x)]$.
-

AdaBoost

1. Inicie com os pesos $w_i = 1/N$, $i = 1, 2, \dots, N$

N é o número de linhas, logo, $w = 1/6 = 0,1666$

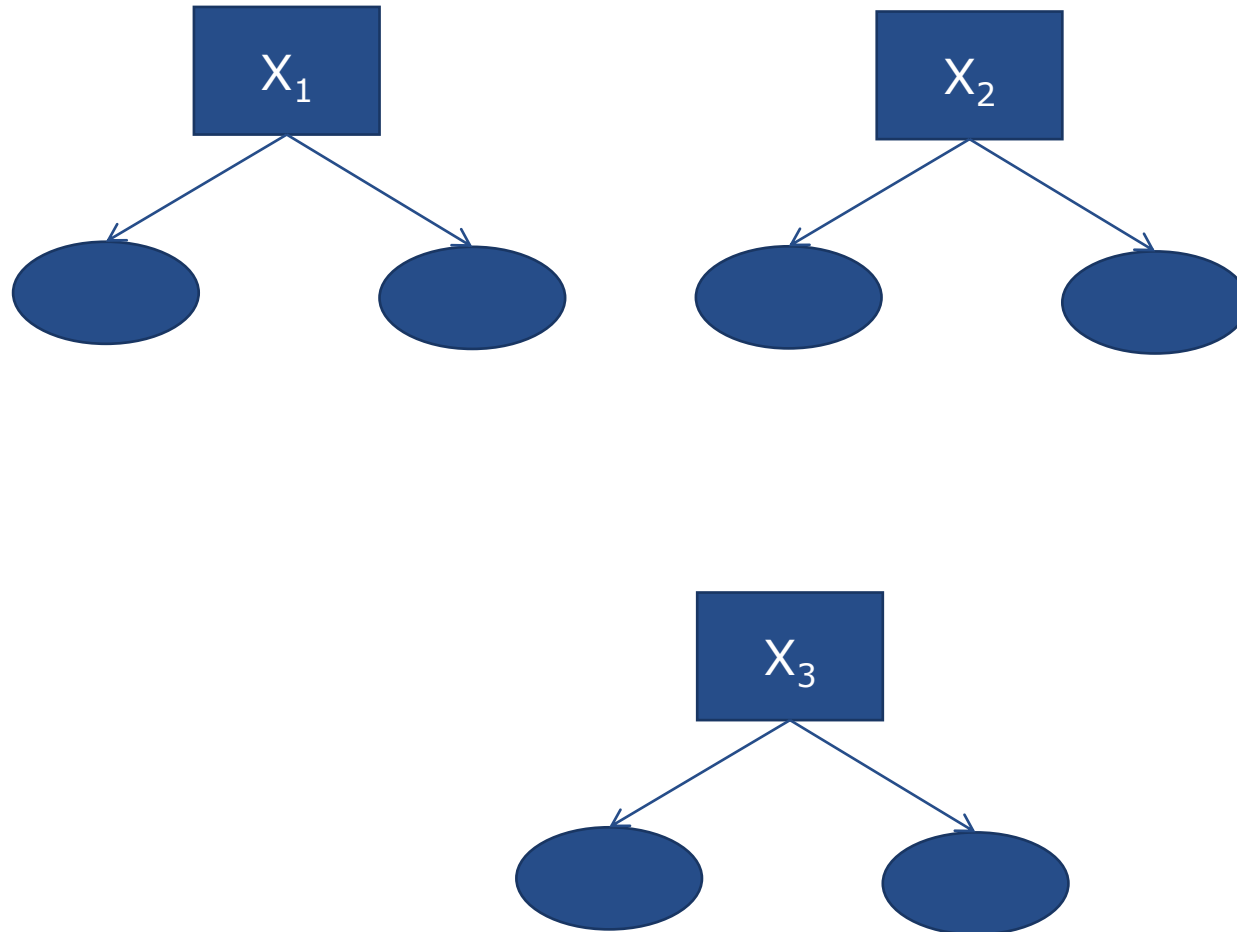
	X₁	X₂	X₃	Y	w
1					0,17
2					0,17
3					0,17
4					0,17
5					0,17
6					0,17

AdaBoost

$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
 - b. Crie **um Stump para cada variável explicativa**

	X_1	X_2	X_3	Y	w
1					0,17
2					0,17
3					0,17
4					0,17
5					0,17
6					0,17



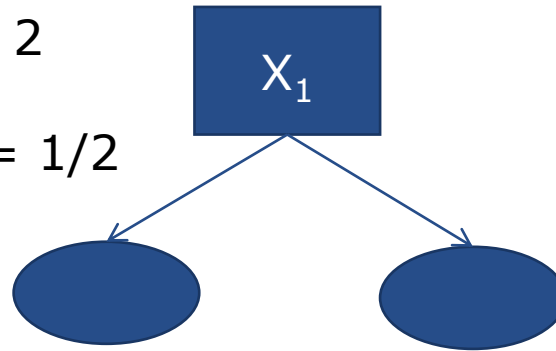
AdaBoost

$m = 1$

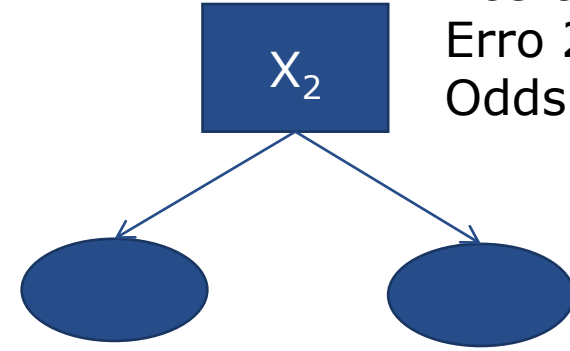
2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
c. Calcule o **odds (QntAcertos/QntErros)** para cada Stump

	X_1	X_2	X_3	Y	w
1					0,17
2					0,17
3					0,17
4					0,17
5					0,17
6					0,17

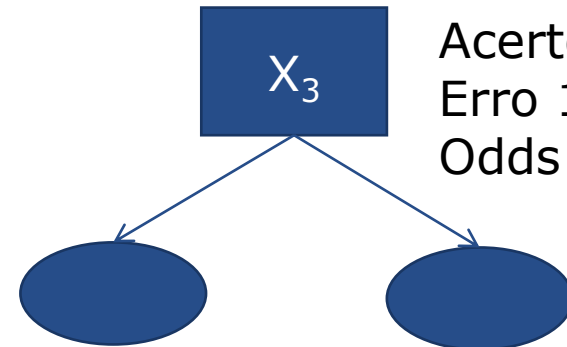
Acerto 2
Erro 4
Odds = 1/2



Acerto 4
Erro 2
Odds = 2



Acerto 5
Erro 1
Odds = 5



AdaBoost

$m = 1$

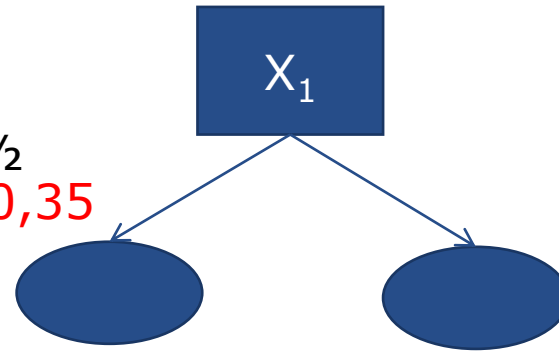
2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

d. Calcule a **performance de cada Stump** = $\frac{1}{2} \ln(\text{QntAcertos}/\text{QntErros})$

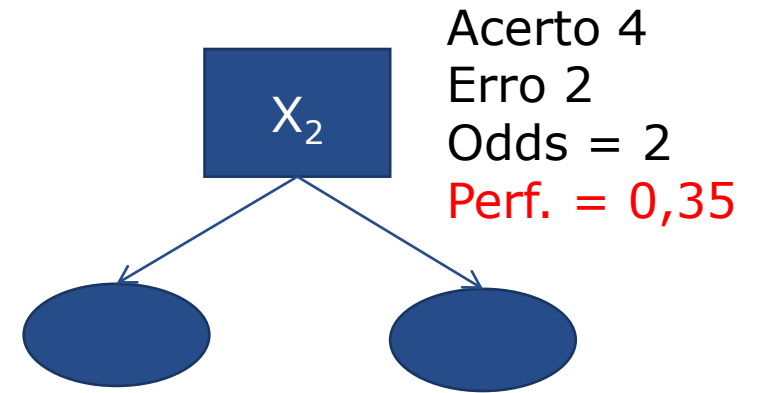
	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,17
0,17
0,17
0,17
0,17
0,17

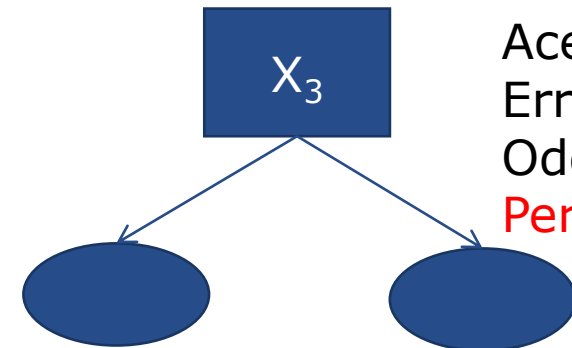
Acerto 2
Erro 4
Odds = $\frac{1}{2}$
Perf. = -0,35



Perf. = $\frac{1}{2} * \ln(\text{odds})$
Perf. = $\frac{1}{2} * \ln(\frac{1}{2})$
Perf. = $\frac{1}{2} * \ln(0,5)$
Perf. = $\frac{1}{2} * (-0,69)$
Perf. = -0,35



Acerto 4
Erro 2
Odds = 2
Perf. = 0,35



Acerto 5
Erro 1
Odds = 5
Perf. = 0,8

AdaBoost

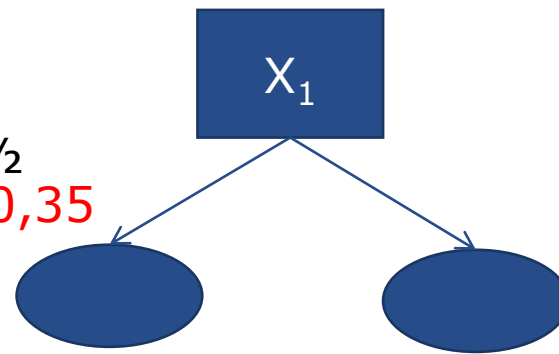
$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
e. Selecione o Stump com **maior performance**

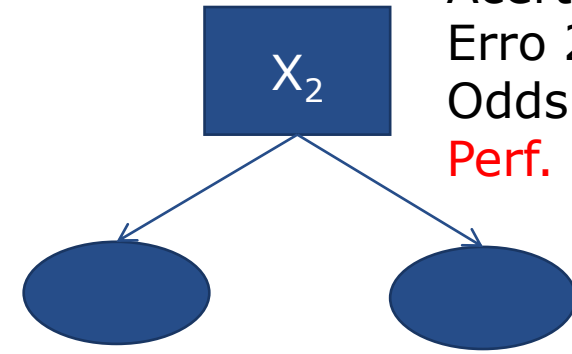
	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,17
0,17
0,17
0,17
0,17
0,17

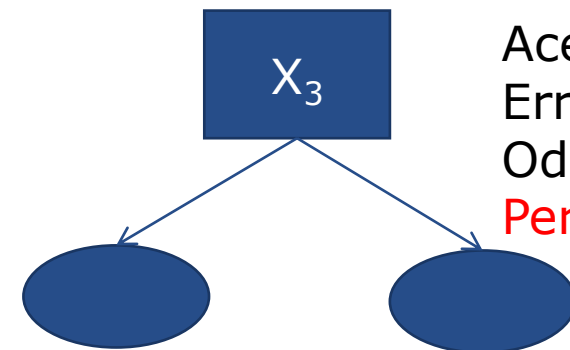
Acerto 2
Erro 4
Odds = $\frac{1}{2}$
Perf. = -0,35



Acerto 4
Erro 2
Odds = 2
Perf. = 0,35



Acerto 5
Erro 1
Odds = 5
Perf. = 0,8

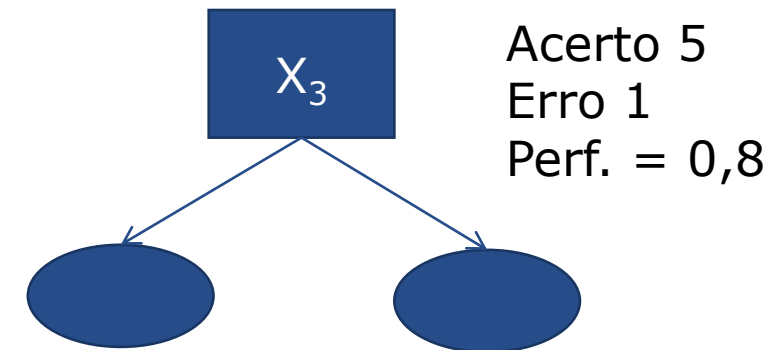


AdaBoost

$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w
1					0,17
2					0,17
3					0,17
4					0,17
5					0,17
6					0,17



AdaBoost

$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- e. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,17
0,17
0,17
0,17
0,17
0,17

w_n

$$w_1 \leftarrow w_1 * e^{(\pm \text{Performance})}$$

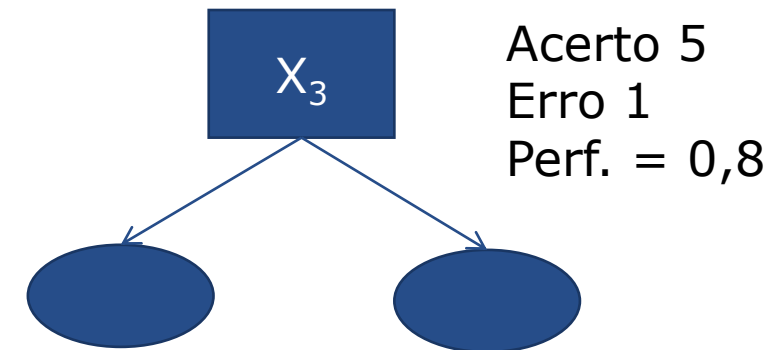
$$w_2 \leftarrow w_2 * e^{(\pm \text{Performance})}$$

$$w_3 \leftarrow w_3 * e^{(\pm \text{Performance})}$$

$$w_4 \leftarrow w_4 * e^{(\pm \text{Performance})}$$

$$w_5 \leftarrow w_5 * e^{(\pm \text{Performance})}$$

$$w_6 \leftarrow w_6 * e^{(\pm \text{Performance})}$$



AdaBoost

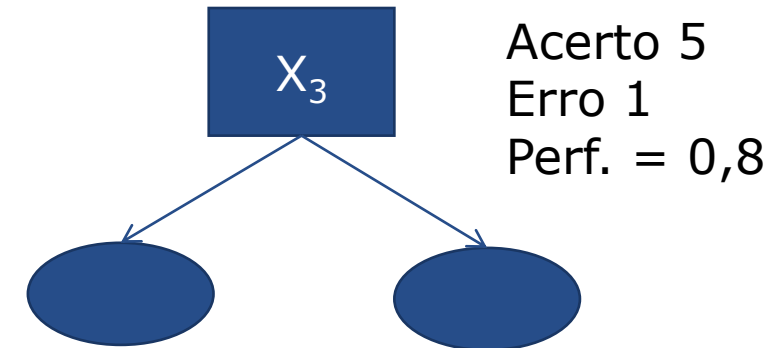
$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w	w_n
1					0,17	
2					0,17	
3					0,17	
4					0,17	
5					0,17	
6					0,17	

$$\begin{aligned}w_1 &\leftarrow w_1 * e^{(\pm \text{Performance})} \\&0,17 * e^{(\pm \text{Performance})} \\&0,17 * e^{(\pm 0,8)} \\&0,17 * e^{(- 0,8)} \\&0,17 * 0,45 \\&0,08\end{aligned}$$



AdaBoost

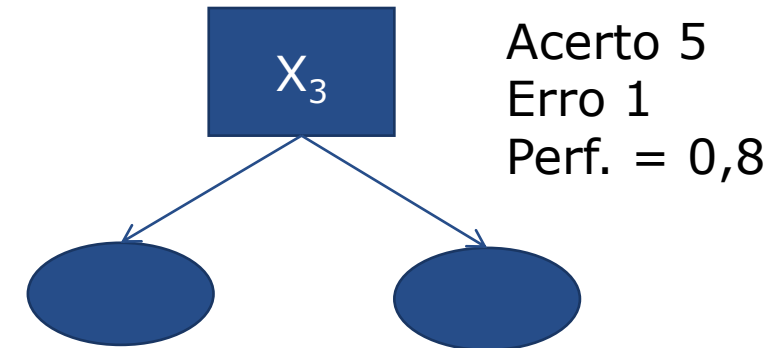
$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w	w_n
1					0,17	0,08
2					0,17	
3					0,17	
4					0,17	
5					0,17	
6					0,17	

$$\begin{aligned}w_1 &\leftarrow w_1 * e^{(\pm \text{Performance})} \\&0,17 * e^{(\pm \text{Performance})} \\&0,17 * e^{(\pm 0,8)} \\&0,17 * e^{(- 0,8)} \\&0,17 * 0,45 \\&0,08\end{aligned}$$



AdaBoost

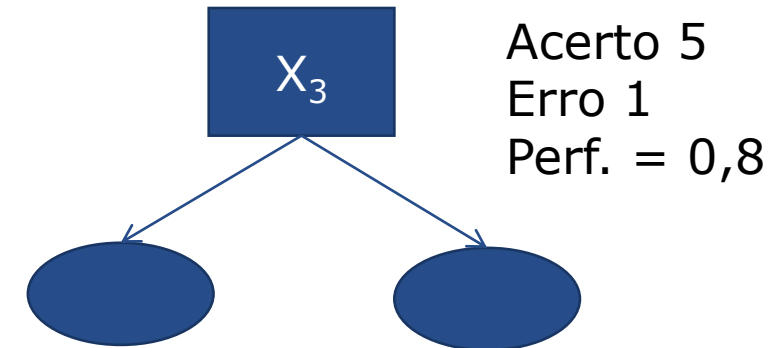
$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w	w_n
1					0,17	0,08
2					0,17	
3					0,17	
4					0,17	
5					0,17	
6					0,17	

$$\begin{aligned}w_2 &\leftarrow w_2 * e^{(\pm \text{Performance})} \\&0,17 * e^{(\pm \text{Performance})} \\&0,17 * e^{(\pm 0,8)} \\&0,17 * e^{(- 0,8)} \\&0,17 * 0,45 \\&0,08\end{aligned}$$



AdaBoost

$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w	w_n
1					0,17	0,08
2					0,17	0,08
3					0,17	
4					0,17	
5					0,17	
6					0,17	

$$w_2 \leftarrow w_2 * e^{(\pm \text{Performance})}$$

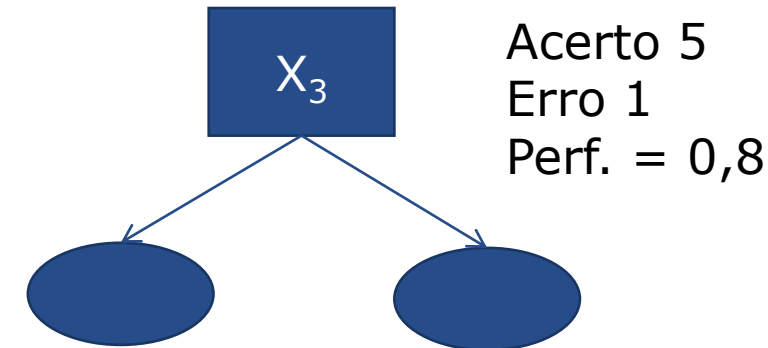
$$0,17 * e^{(\pm \text{Performance})}$$

$$0,17 * e^{(\pm 0,8)}$$

$$0,17 * e^{(- 0,8)}$$

$$0,17 * 0,45$$

$$0,08$$



AdaBoost

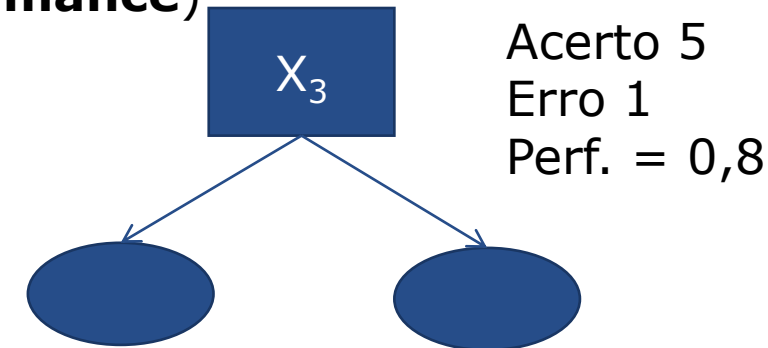
$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w	w_n
1					0,17	0,08
2					0,17	0,08
3					0,17	0,08
4					0,17	0,08
5					0,17	0,08
6					0,17	

$$\begin{aligned} w_6 &\leftarrow w_6 * e^{(\pm \text{Performance})} \\ &0,17 * e^{(\pm \text{Performance})} \\ &0,17 * e^{(\pm 0,8)} \\ &0,17 * e^{(+ 0,8)} \\ &0,17 * 2,23 \\ &0,38 \end{aligned}$$



AdaBoost

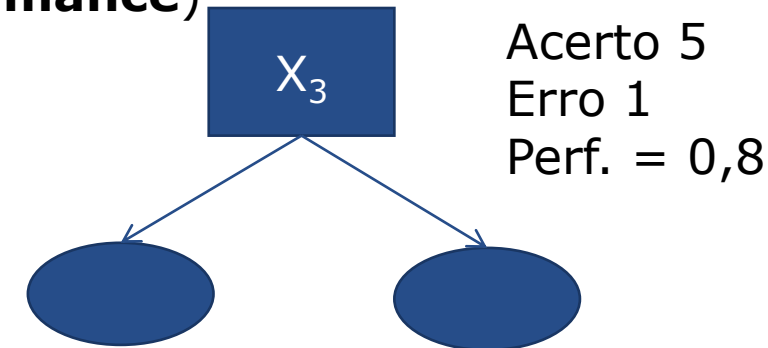
$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w	w_n
1					0,17	0,08
2					0,17	0,08
3					0,17	0,08
4					0,17	0,08
5					0,17	0,08
6					0,17	0,38

$$\begin{aligned} w_6 &\leftarrow w_6 * e^{(\pm \text{Performance})} \\ &0,17 * e^{(\pm \text{Performance})} \\ &0,17 * e^{(\pm 0,8)} \\ &0,17 * e^{(+ 0,8)} \\ &0,17 * 2,23 \\ &0,38 \end{aligned}$$



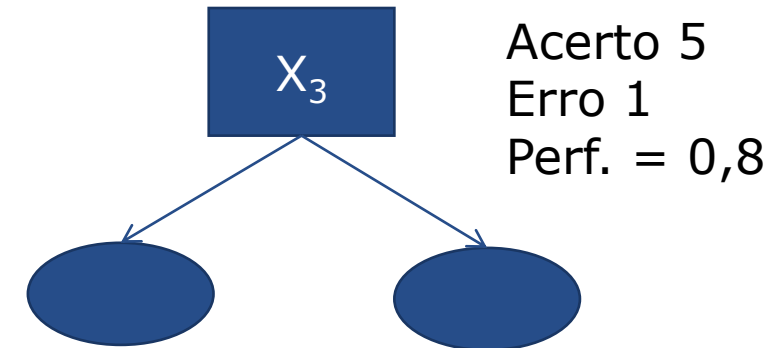
AdaBoost

$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w	w_n
1					0,17	0,08
2					0,17	0,08
3					0,17	0,08
4					0,17	0,08
5					0,17	0,08
6					0,17	0,38

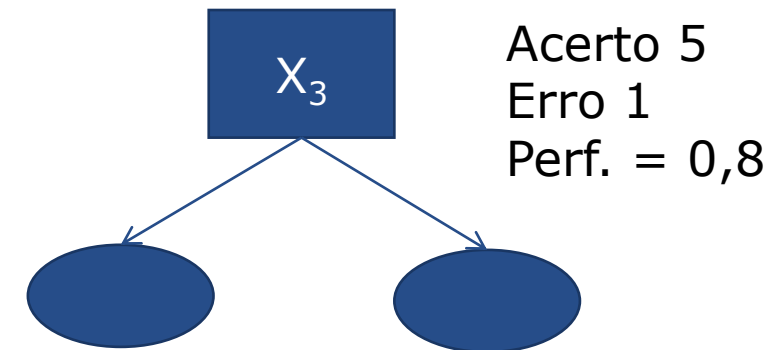


AdaBoost

$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo $+$ quando erro na linha e $-$ quando acerto na linha e $i = 1, 2, \dots, N$

	X_1	X_2	X_3	Y	w
1					0,08
2					0,08
3					0,08
4					0,08
5					0,08
6					0,38



AdaBoost

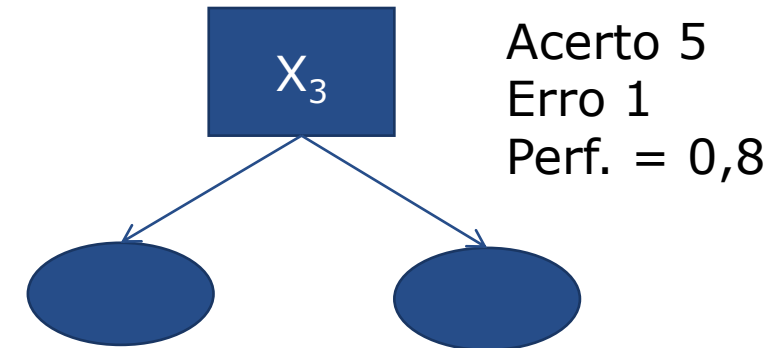
$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

g. **Normalize os pesos** para que a soma deles seja **igual a 1**. ($\sum_i w_i = 1$)

	X_1	X_2	X_3	Y	w	$\div 0,78$	w
1					0,08		0,10
2					0,08		0,10
3					0,08		0,10
4					0,08		0,10
5					0,08		0,10
6					0,38		0,50

$\sum_i w_i = 0,78$ $\sum_i w_i = 1$

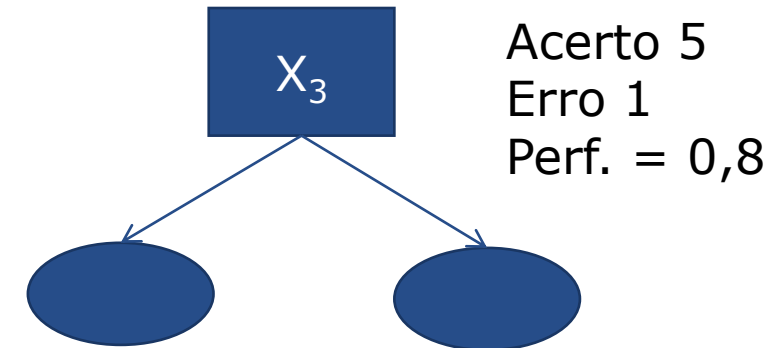


AdaBoost

$m = 1$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- g. **Normalize os pesos** para que a soma deles seja **igual a 1**. ($\sum_i w_i = 1$)

	X_1	X_2	X_3	Y	w
1					0,10
2					0,10
3					0,10
4					0,10
5					0,10
6					0,50



AdaBoost

$$m = 2$$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)

	X_1	X_2	X_3	Y	w
1					0,10
2					0,10
3					0,10
4					0,10
5					0,10
6					0,50

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w) crie um novo conjunto de dados de treino**

	X_1	X_2	X_3	Y	w	Range
1					0,10	
2					0,10	
3					0,10	
4					0,10	
5					0,10	
6					0,50	

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
 - a. Usando os **pesos (w) crie um novo conjunto de dados de treino**

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w) crie um novo conjunto de dados de treino**

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20

AdaBoost

$$m = 2$$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w) crie um novo conjunto de dados de treino**

	X_1	X_2	X_3	Y	w	Range
1					0,10	0 - 0,10
2					0,10	0,10 - 0,20
3					0,10	0,20 - 0,30
4					0,10	0,30 - 0,40
5					0,10	0,40 - 0,50
6					0,50	0,50 - 1

AdaBoost

$$m = 2$$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



0,08

	X_1	X_2	X_3	Y
1				

AdaBoost

$$m = 2$$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



0,65

	X_1	X_2	X_3	Y
1				

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



0,65

	X_1	X_2	X_3	Y
1				
6				

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



0,44

	X_1	X_2	X_3	Y
1				
6				

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



0,44

	X_1	X_2	X_3	Y
1				
6				
5				

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



0,77

	X_1	X_2	X_3	Y
1				
6				
5				

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
 - a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



0,77

	X_1	X_2	X_3	Y
1				
6				
5				
6				

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino

	X_1	X_2	X_3	Y
1				
2				
3				
4				
5				
6				

w
0,10
0,10
0,10
0,10
0,10
0,50

Range
0 - 0,10
0,10 - 0,20
0,20 - 0,30
0,30 - 0,40
0,40 - 0,50
0,50 - 1

Bootstrap
com pesos



	X_1	X_2	X_3	Y
1				
6				
5				
6				
3				
6				

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- a. Usando os **pesos (w) crie um novo conjunto de dados de treino**

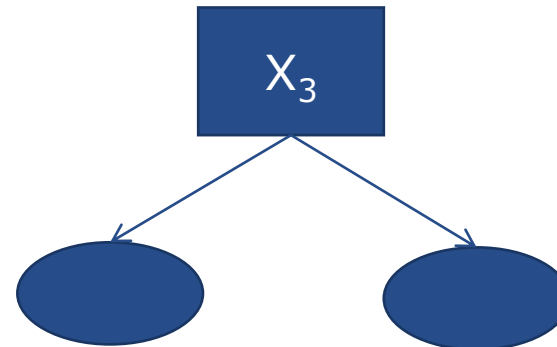
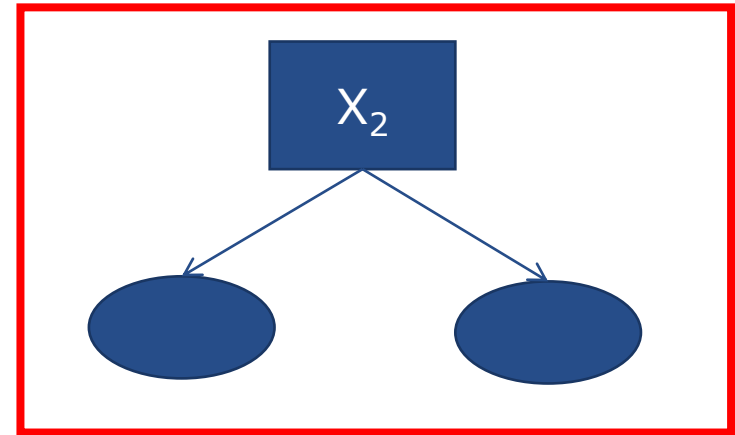
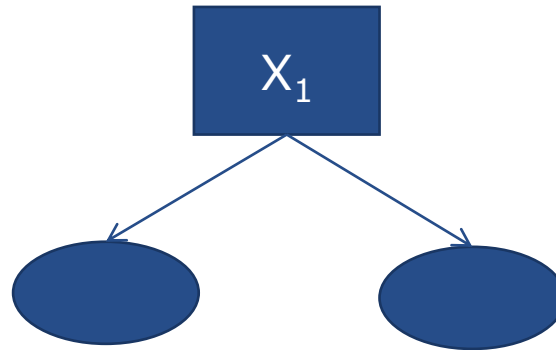
	X_1	X_2	X_3	Y	w
1					0,10
6					0,50
5					0,10
6					0,50
3					0,10
6					0,50

AdaBoost

$m = 2$

2. Repita para $m = 1$ até M (sendo M a quantidade de Stumps/Árvores)
- b. Crie **um Stump para cada variável explicativa**

	X_1	X_2	X_3	Y	w
1					0,10
6					0,50
5					0,10
6					0,50
3					0,10
6					0,50



AdaBoost

M = 20

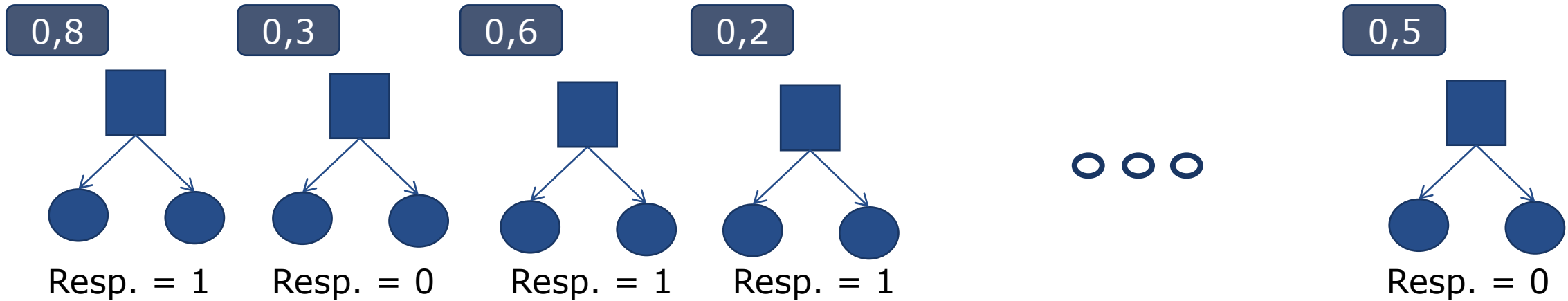
2. Repita para $m = 1$ até M .

- a. Usando os **pesos (w)** crie um novo conjunto de dados de treino.
- b. Crie **um Stump** para cada variável explicativa.
- c. Calcule o **odds ($\text{QntAcertos}/\text{QntErros}$)** para cada Stump
- d. Calcule a **performance de cada Stump** = $\frac{1}{2} \ln(\text{QntAcertos}/\text{QntErros})$
- e. Selecione o Stump com **maior performance**
- f. Atualize o peso de cada linha $w_i \leftarrow w_i * e^{(\pm \text{Performance})}$, sendo + quando erro na linha e – quando acerto na linha e $i = 1, 2, \dots, N$
- g. **Normalize os pesos** para que a soma deles seja **igual a 1**.
 $\sum_i w_i = 1$

AdaBoost

$m = 20$

3. Some a **performance** de cada Stump pra cada classe e a maior será a resposta.

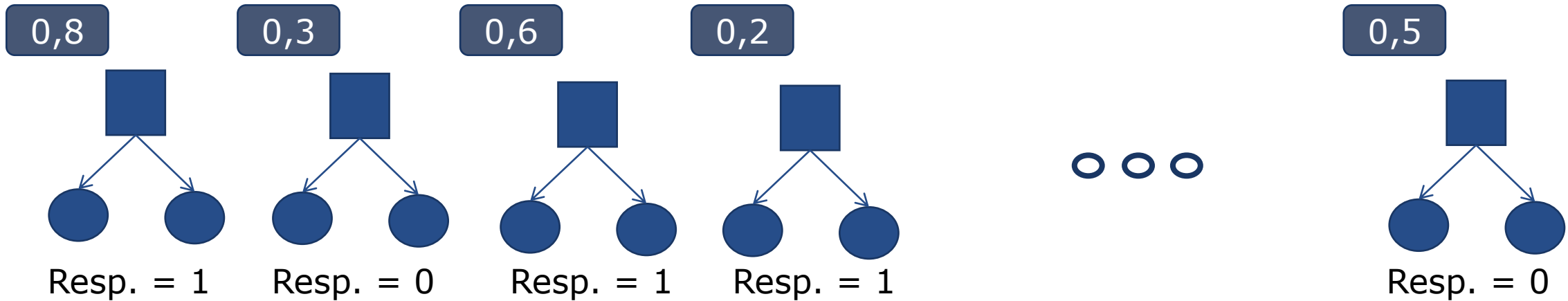


0	1

AdaBoost

$m = 20$

3. Some a **performance** de cada Stump pra cada classe e a maior será a resposta.

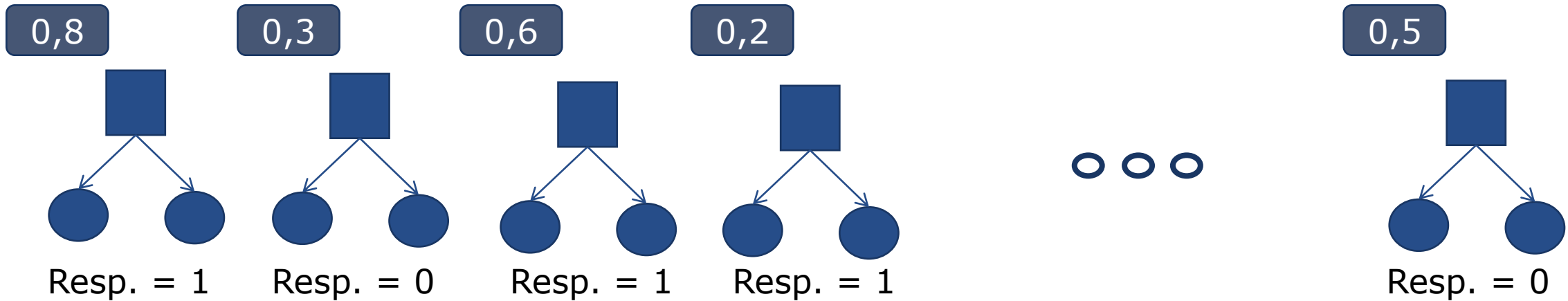


0	1
	0,8

AdaBoost

$m = 20$

3. Some a **performance** de cada Stump pra cada classe e a maior será a resposta.



0	1
0,3	0,8
0,5	0,6
...	0,2
	...

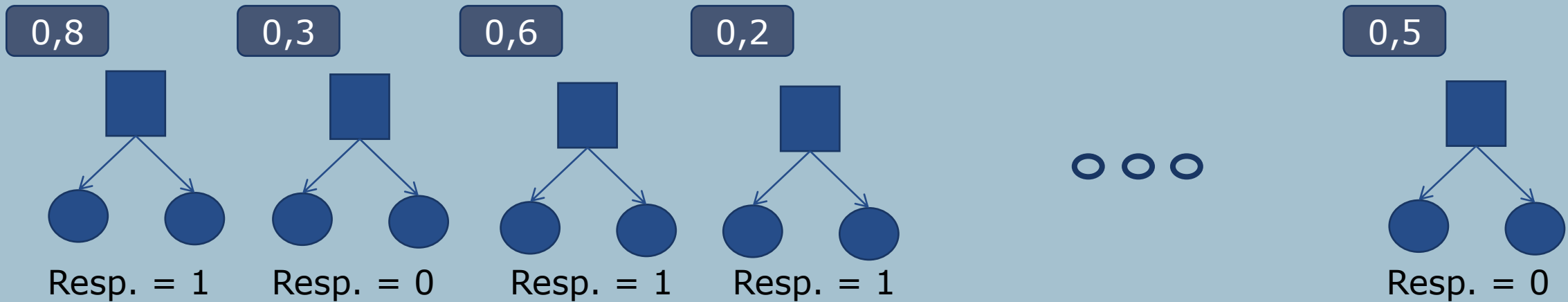
Soma

2,3 3,6

AdaBoost

$m = 20$

3. Some a **performance** de cada Stump pra cada classe e a maior será a resposta.



0	1
0,3	0,8
0,5	0,6
...	0,2
	...
Soma	2,3
	3,6

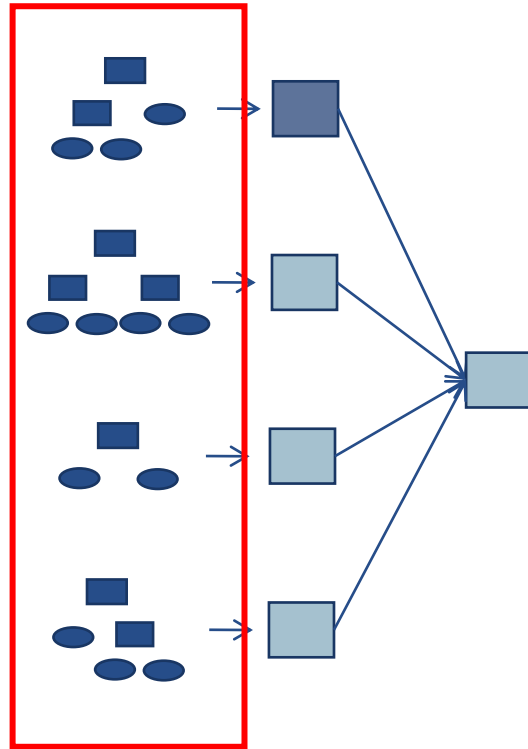


Resposta final da votação ponderada = 1
(*weighted majority vote*)

RF vs AdaBoost

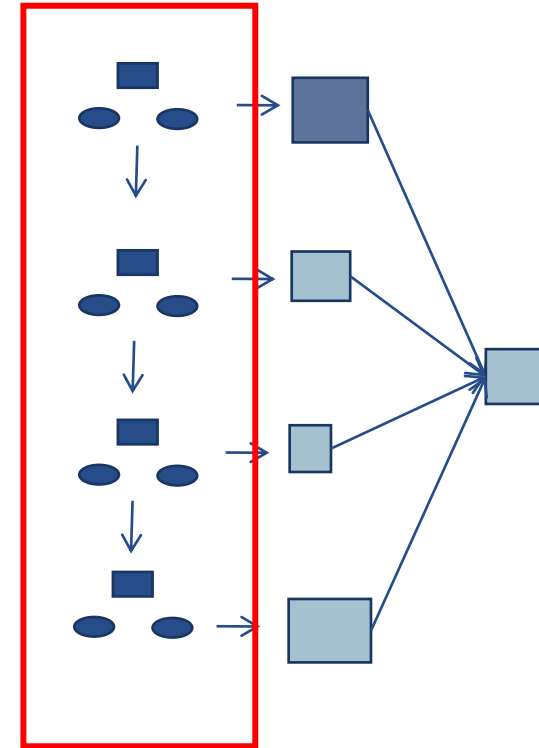
3 grandes diferenças

Random forest



Floresta de árvores

AdaBoost

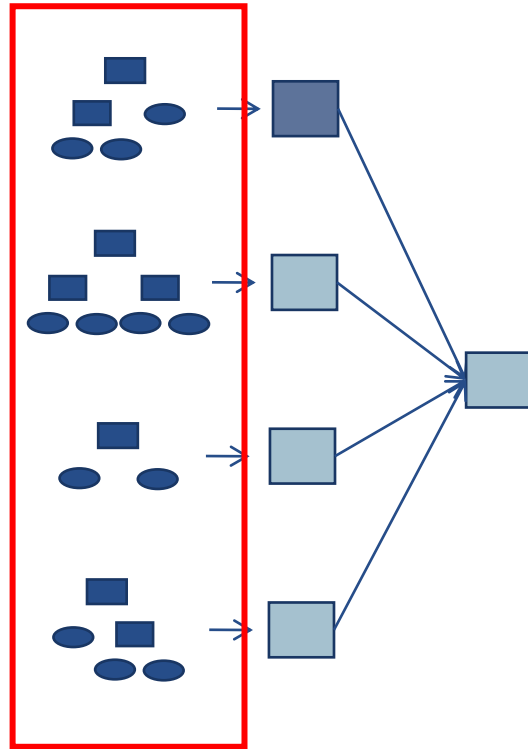


Floresta de Stumps

RF vs AdaBoost

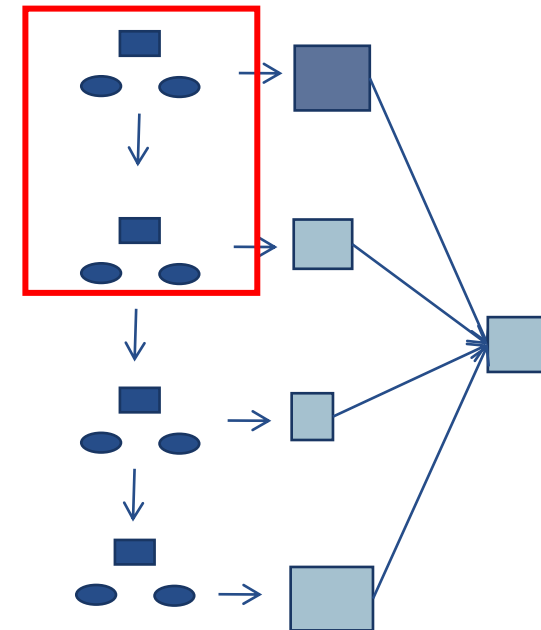
3 grandes diferenças

Random forest



Árvores independentes

AdaBoost

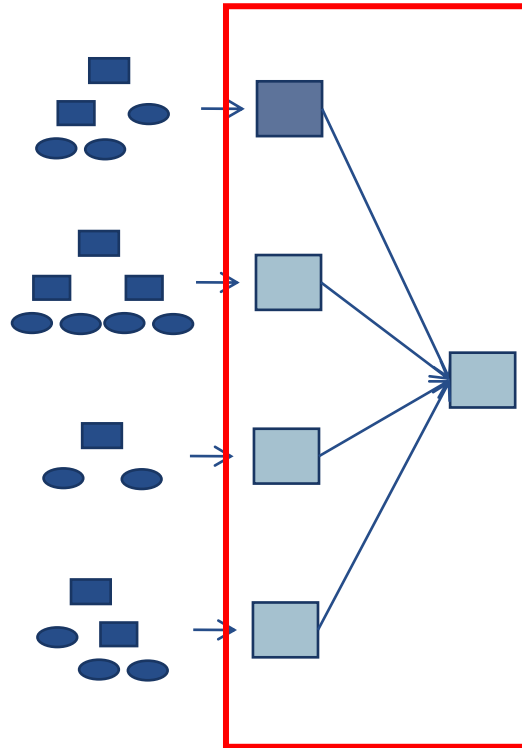


Uma árvore influencia na seguinte

RF vs AdaBoost

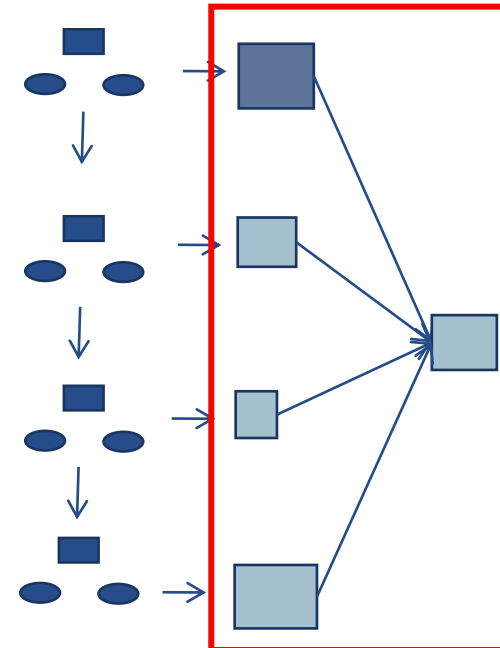
3 grandes diferenças

Random forest



Respostas das árvores
tem o mesmo peso

AdaBoost



Respostas tem pesos
diferentes

AdaBoost

AdaBoost almost never overfits the data no matter how many iterations it is run.
Leo Breiman, 1998