



Data Science  
Academy

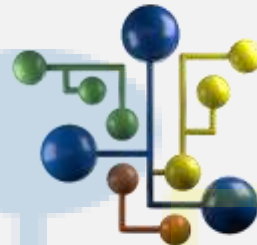
Data Science Academy [thaylathais1@gmail.com](mailto:thaylathais1@gmail.com) 611665e6e32fc3dbb86d0db5

# Machine Learning



Data Science  
Academy

Data Science Academy [thaylathais1@gmail.com](mailto:thaylathais1@gmail.com) 611665e6e32fc3dbb86d0db5



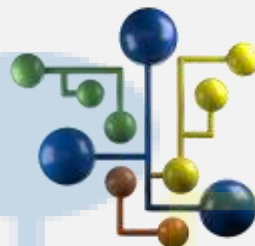
**Data Science  
Academy**

Seja muito bem-vindo(a)!



Data Science  
Academy

Data Science Academy [thaylathais1@gmail.com](mailto:thaylathais1@gmail.com) 611665e6e32fc3dbb86d0db5



**Data Science  
Academy**

# Métodos Ensemble



Métodos Ensemble são uma categoria de Algoritmos de Machine Learning

Aprendizagem  
Supervisionada

Aprendizagem Não  
Supervisionada



## Métodos Ensemble

Random Forest

Bagging

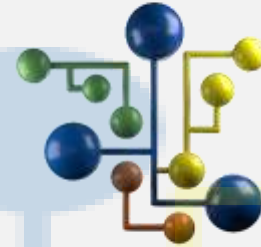
AdaBoost

Gradient Boosting



Data Science  
Academy

Data Science Academy thaylathais1@gmail.com 611665e6e32fc3dbb86d0db5



**Data Science  
Academy**

# Decision Tree, Random Forest e Métodos Ensemble



## Métodos Ensemble

Bagging

Boosting

Voting



## Construção de Ensembles

Construir Vários Modelos

Combinar suas Estimativas





## Estado da Arte em Machine Learning



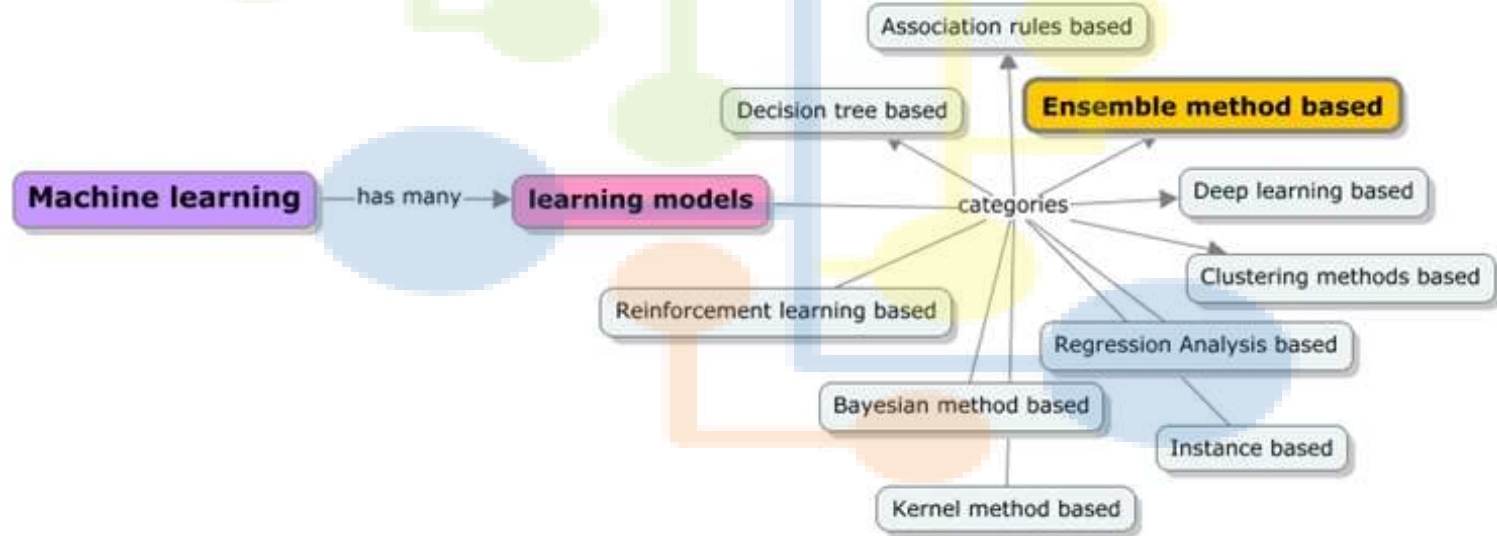


## Estado da Arte em Machine Learning

Acurácia e Simplicidade

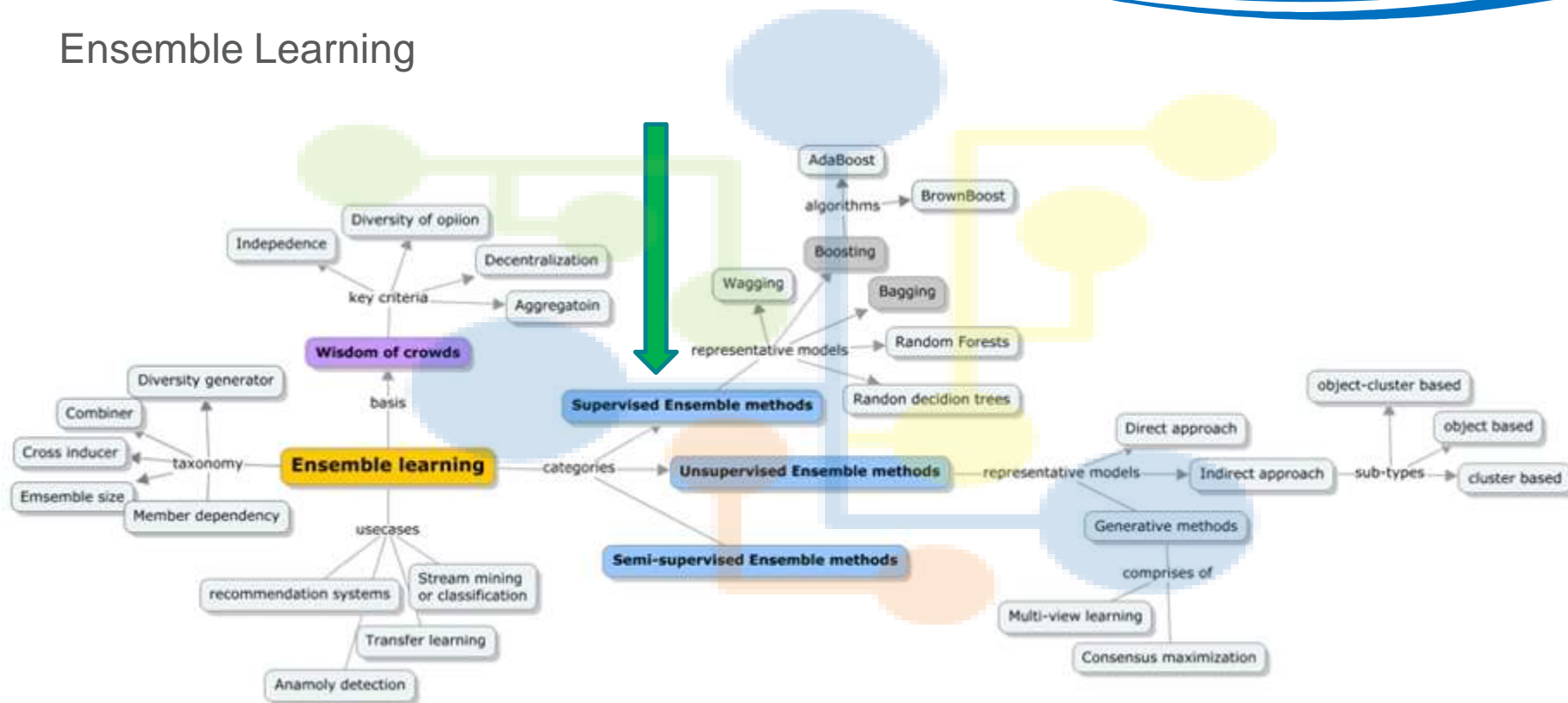


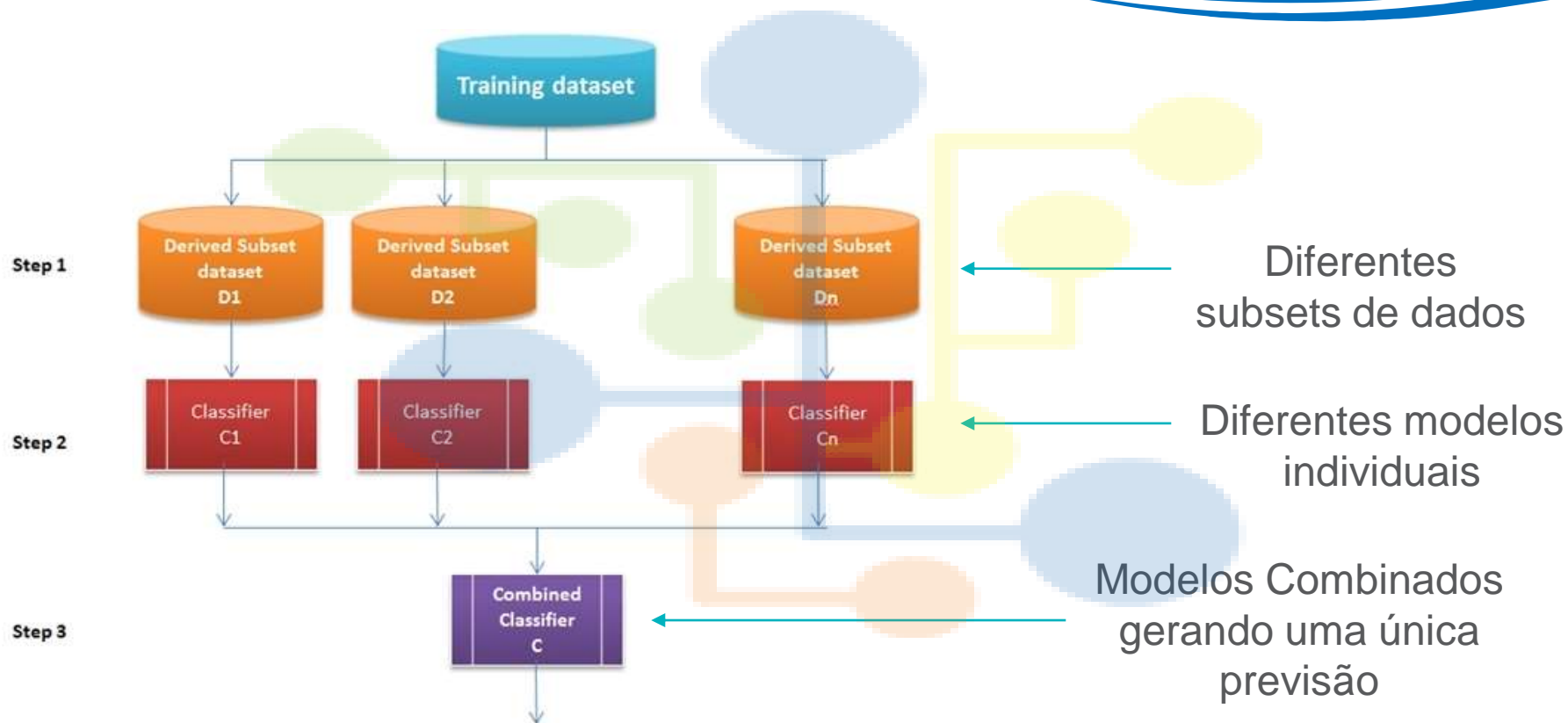
## Métodos Ensemble são uma categoria de Algoritmos de Machine Learning





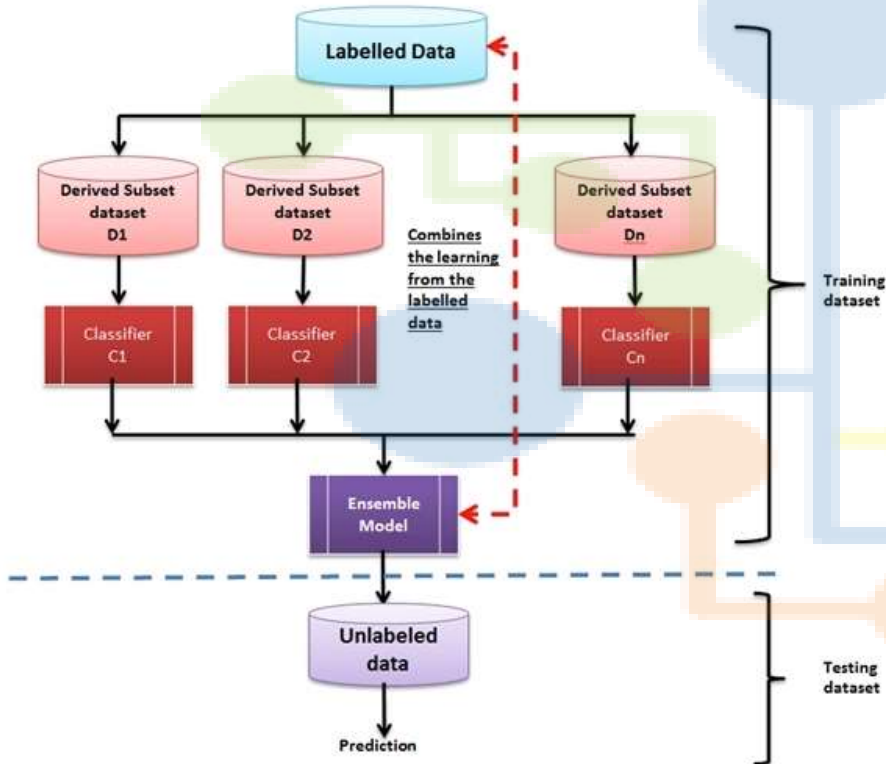
## Ensemble Learning







## Ensemble method for Supervised learning Combining the "learning" technique



Aqui ocorre a criação de diferentes modelos

Aqui ocorre a validação do modelo final



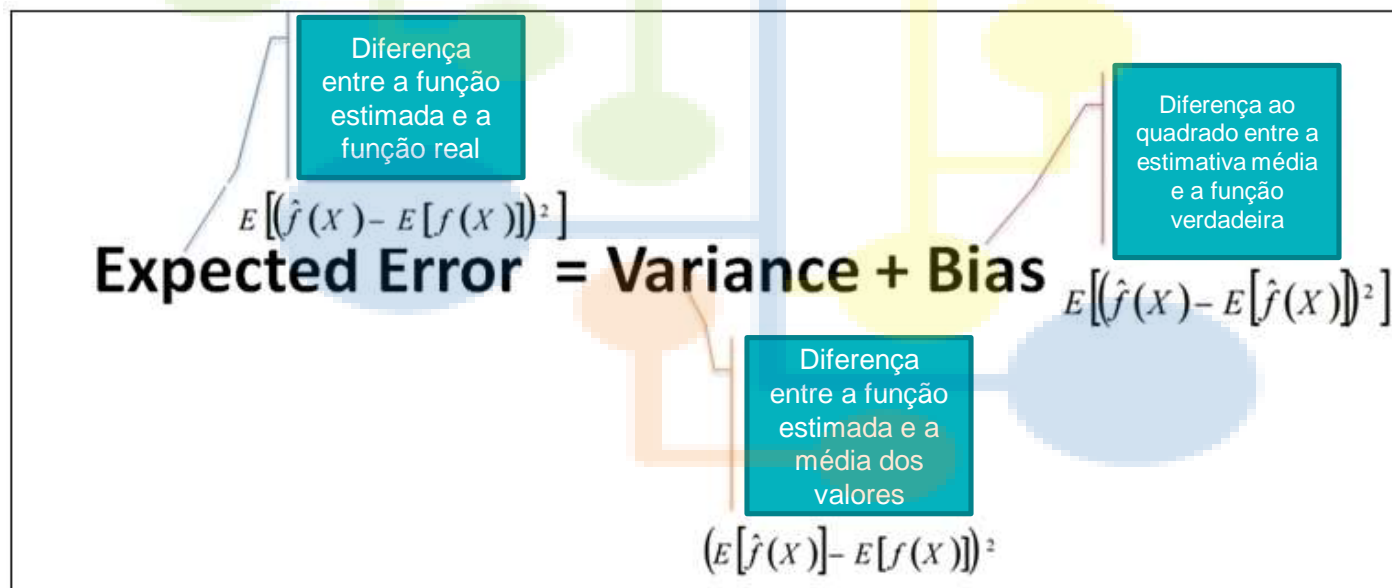
## Bagging (Bootstrap Aggregation)

- 1- É construída uma amostra baseada em Bootstrap (reamostragem) que contém aproximadamente 60% dos registros originais no dataset.
- 2- Os classificadores são treinados usando cada amostra.
- 3- É usada votação da maioria para identificar a classe final do classificador ensemble.





## Bagging (Bootstrap Aggregation)





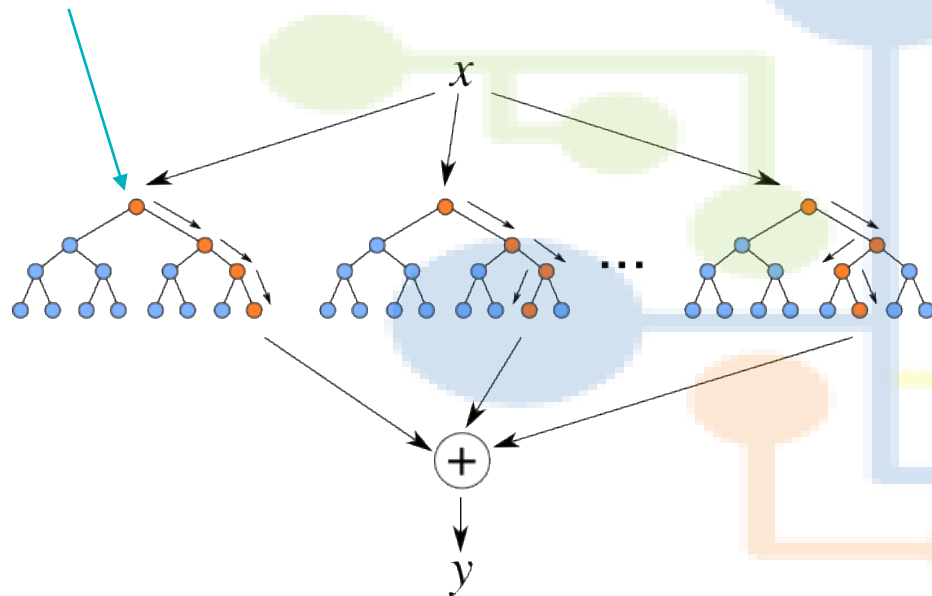


## Random Forest

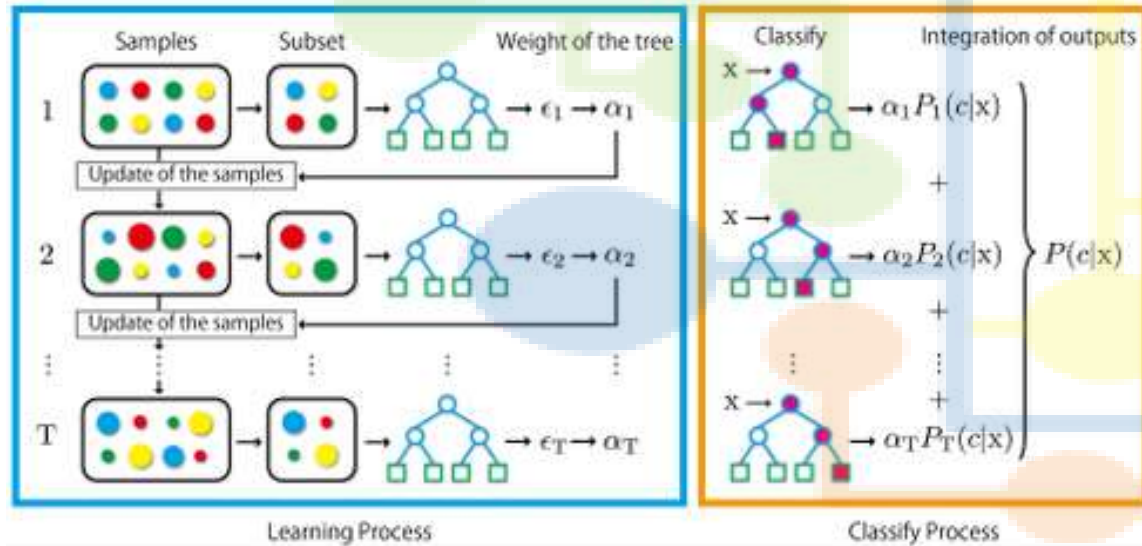




Maior Entropia



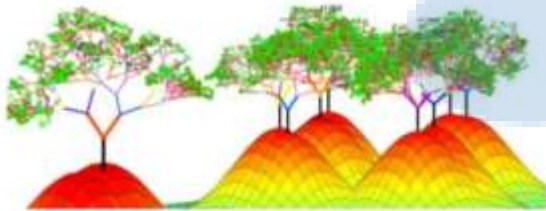
Random Forest



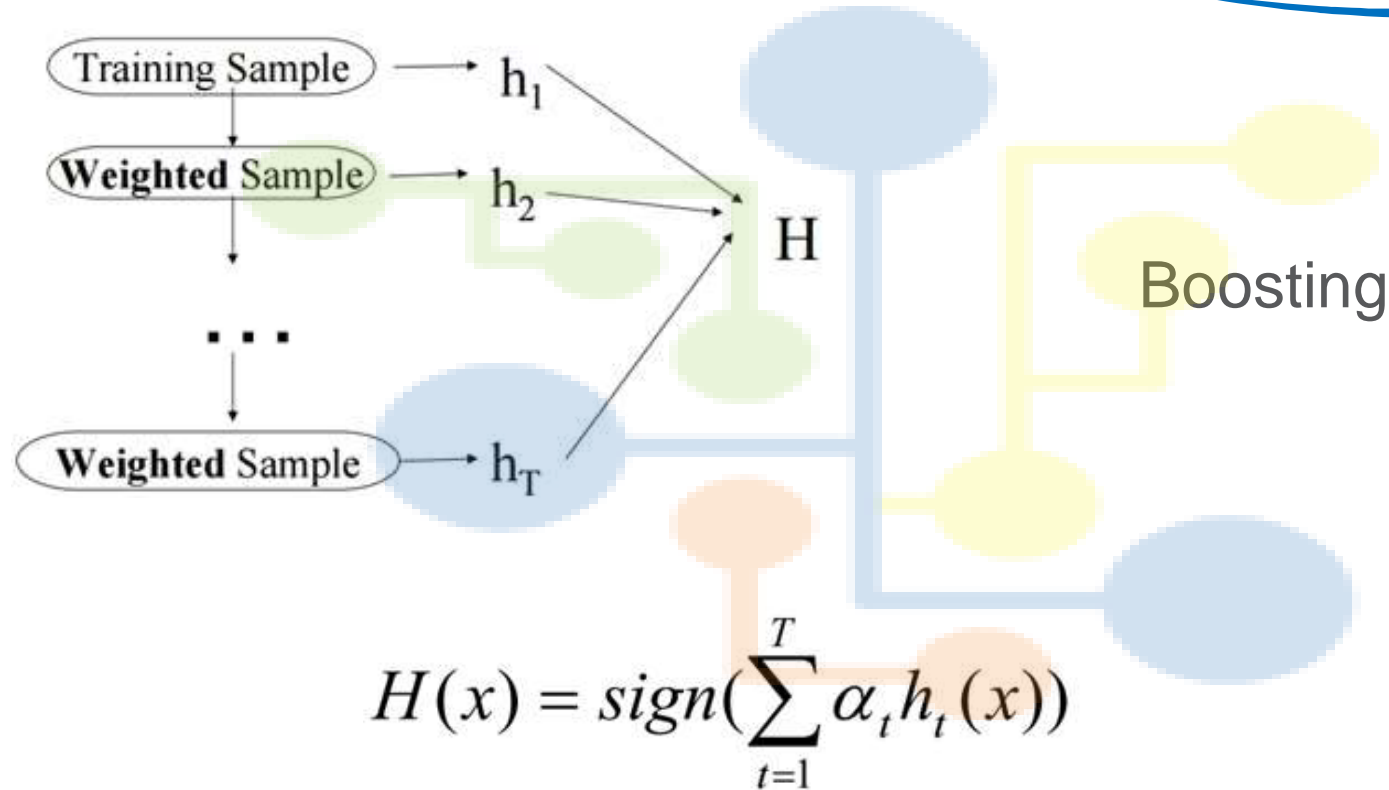
## Random Forest



## Boosting



- Os modelos fracos são treinados sequencialmente.
- Os dados usados para treinar cada classificador são baseados na performance do classificador anterior.
- Cada classificador vota e contribui para o resultado final.





## AdaBoost



$$H(x) = \text{sign}\left(\sum_{t=1}^T \alpha_t h_t(x)\right)$$



Data Science  
Academy

Data Science Academy [thaylathais1@gmail.com](mailto:thaylathais1@gmail.com) 611665e6e32fc3dbb86d0db5



Continue Trilhando uma Excelente Jornada de Aprendizagem!

**Muito Obrigado!**