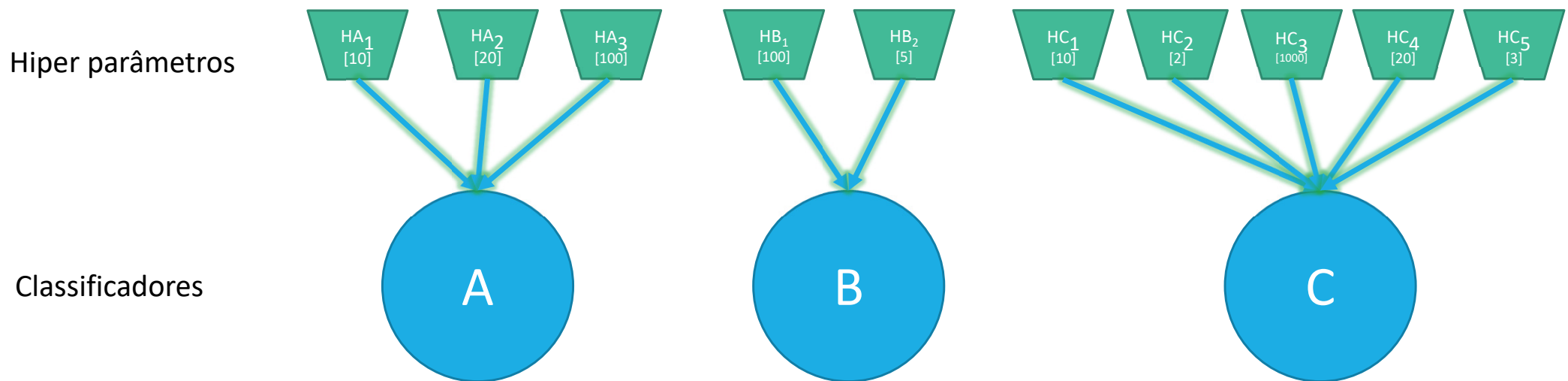


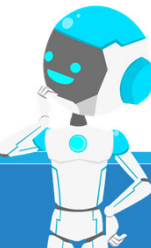
Usando mais Classificadores



Serão necessários 1.220.500 treinos diferentes para testar todas as opções
 $(10 \cdot 20 \cdot 100) + (100 \cdot 5) + (10 \cdot 2 \cdot 1000 \cdot 20 \cdot 3)$

Supondo que cada configuração de cada classificador leve 1 minuto para treino

2,3 Anos!

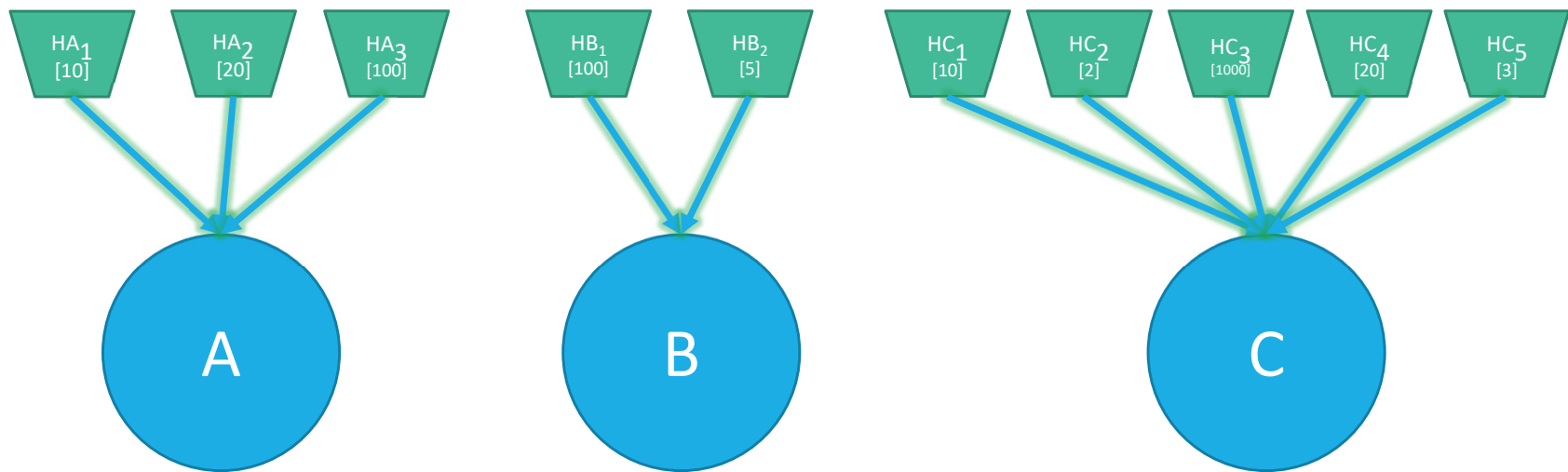


Auto ML

Uma técnica de Auto ML deve conter:

1. Espaço de busca
2. Estratégia de busca
3. Medida de Performance

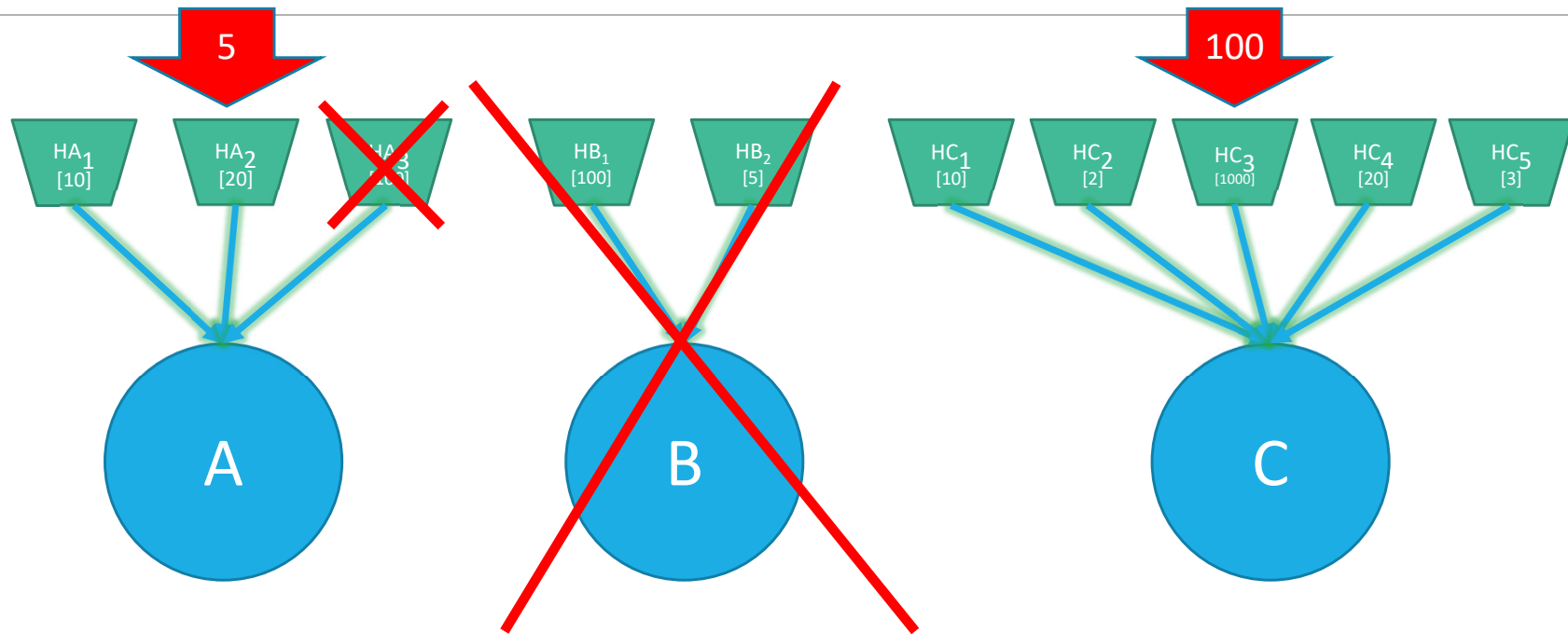
1. Espaço de busca



Espaço de Busca: 1.220.500



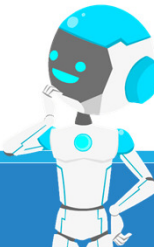
2. Espaço de busca



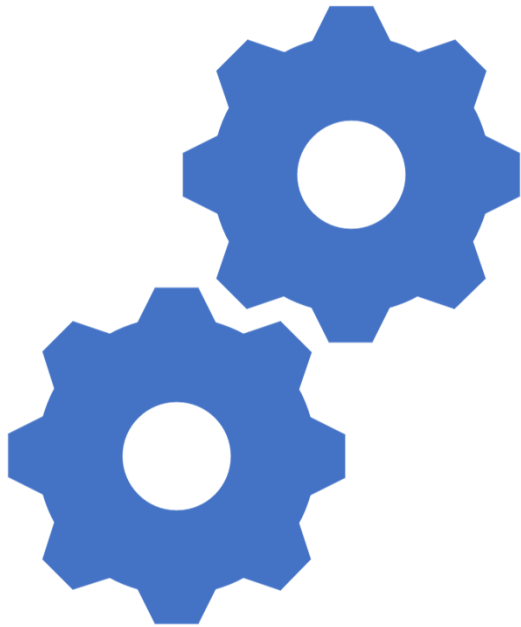
Espaço de Busca: 1.220.500

Novo Espaço de Busca: 120.050

90%



2. Estratégias de Busca



Auto ML busca otimizar a busca pela melhor combinação de hiperparâmetros

Grid Search

Especifica um subconjunto de valores para cada hiper parâmetro

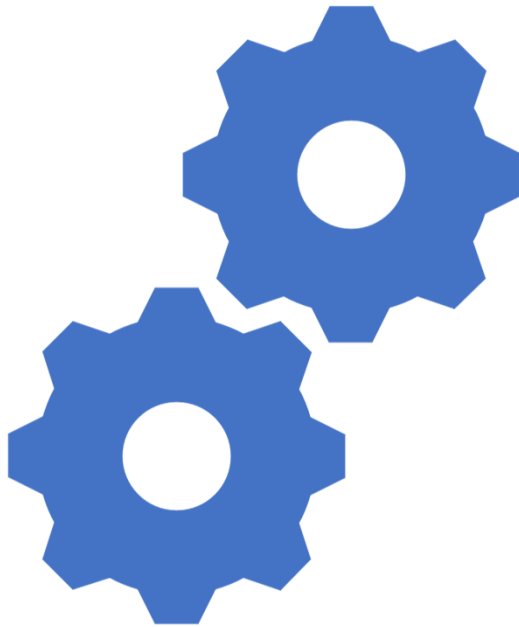
Todas as combinações são testadas

Random Search



- Especifica um subconjunto de valores para cada hiper parâmetro
- Algumas combinações são testadas até um limite de tempo

Otimizador Bayesiano



- Entre os melhores modelos de otimização
- Iterativo
- Utilizam modelo substituto, que tem um custo menor de otimização
- Avaliam resultados dos modelos resultantes antes de escolher os novos hiper parâmetros para os próximos modelos

Algoritmos Genéticos



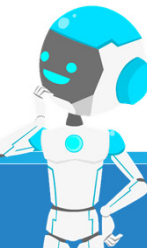
Gerações de hiper parâmetros são testadas



Depois do treino, passam pelos processos de elitismo, mutação e crossover

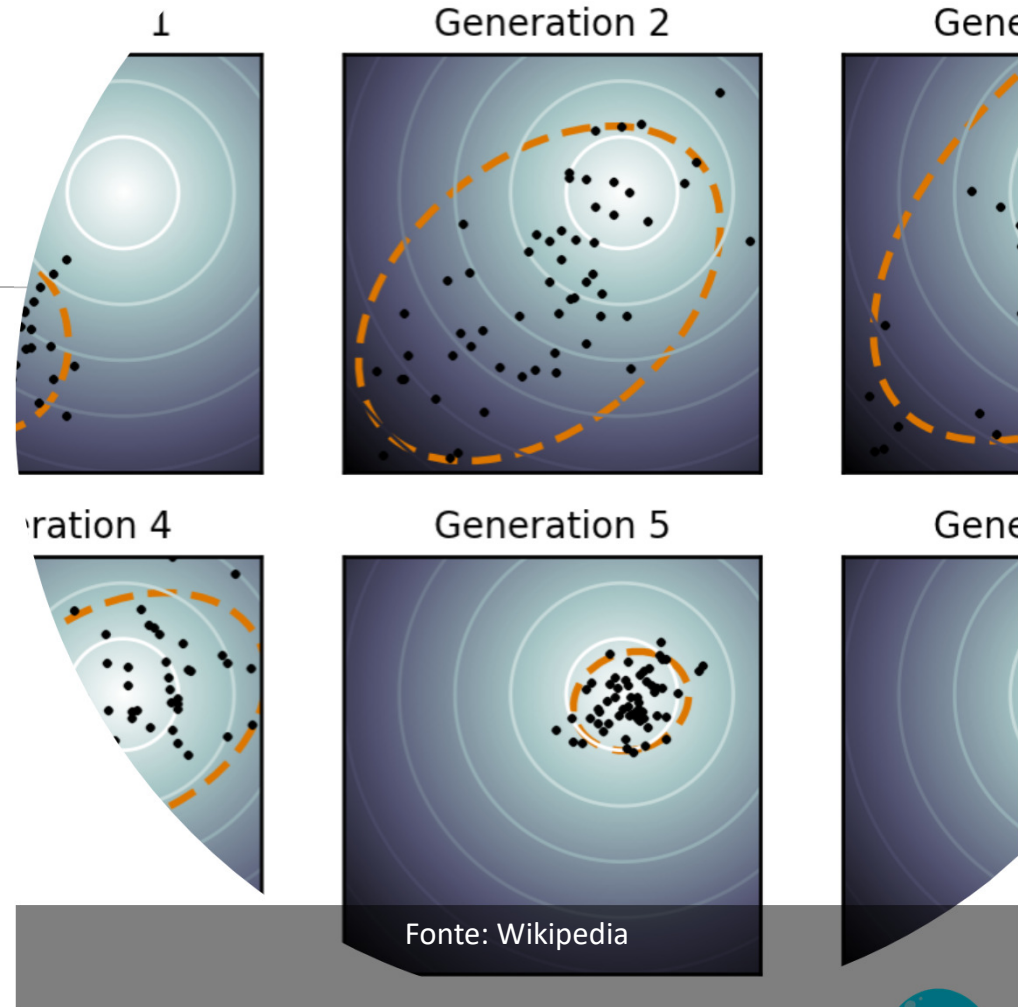


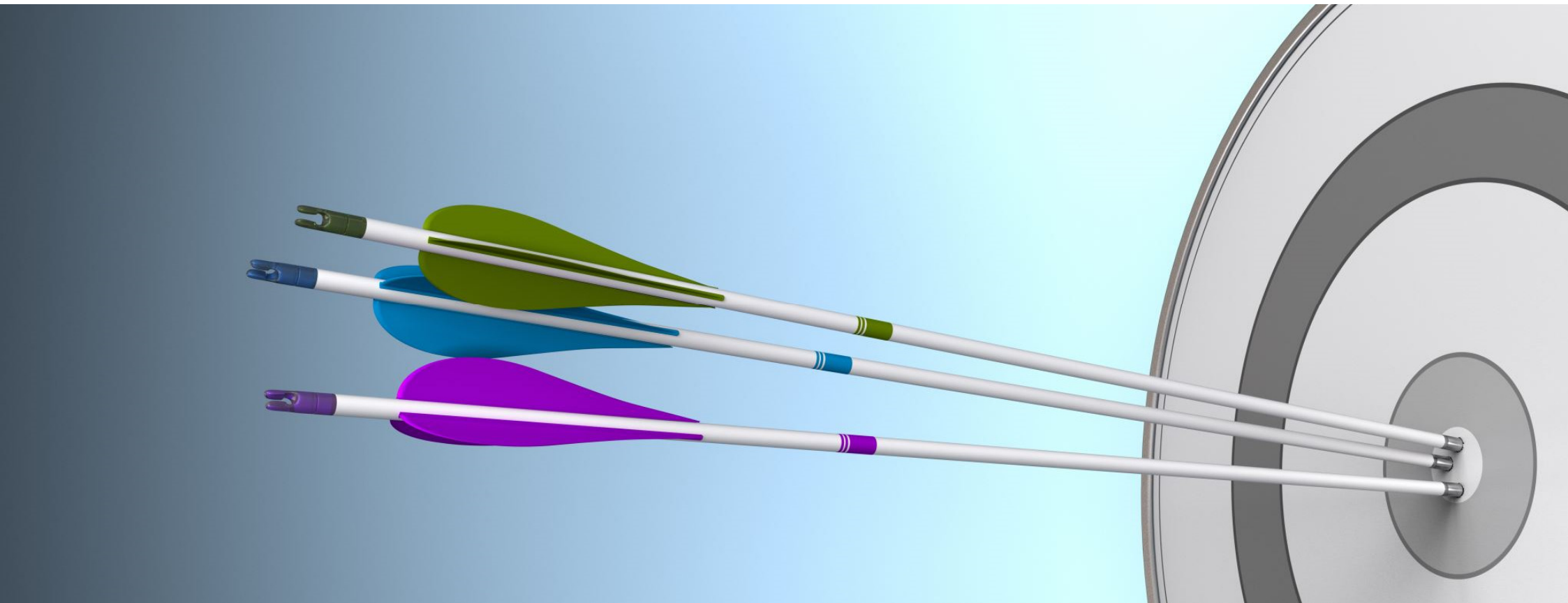
Novas gerações tendem a ter melhor performance



CMA-ES

Covariance matrix
evolutionary strategy

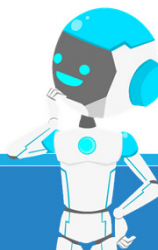




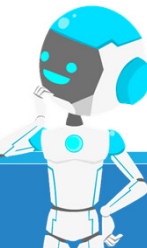
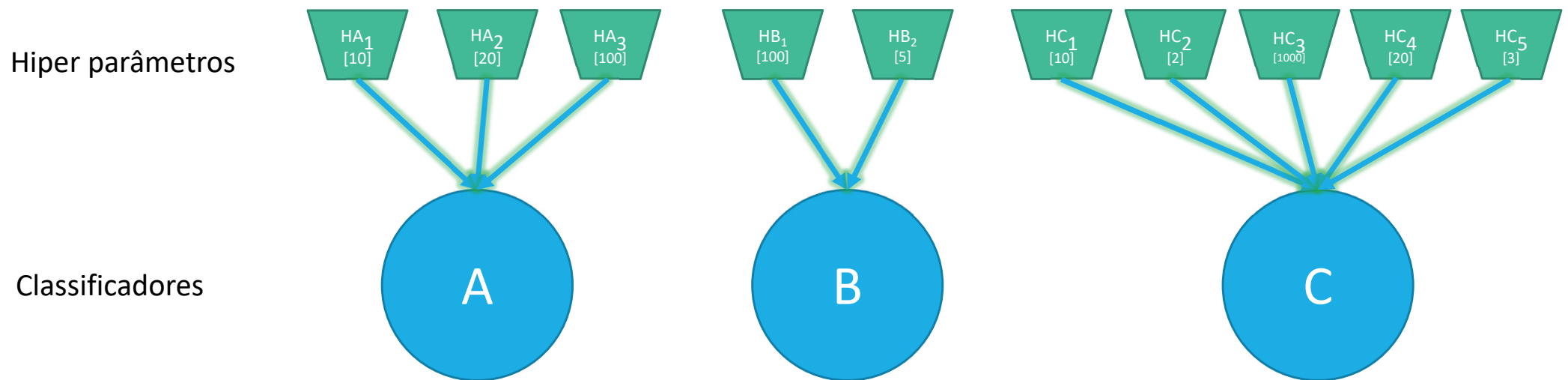
3. Medida de Performance

Como avaliar se estamos melhorando?

- Uma métrica de performance do modelo:
 - MPE
 - RMSE
 - MSE



Multi-Fidelity



Cross-Validation



Partições: 10



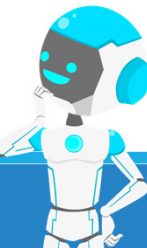
Na prática, os dados passaram pela RNA



EPOCHS x INSTÂNCIAS X PARTICOES



E se eu treinar o modelo com um subconjunto de partições?



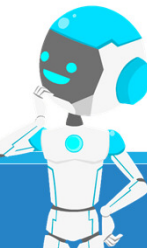
Dados



Cada vez mais dados



E se eu treinar o modelo com um subconjunto de dados?



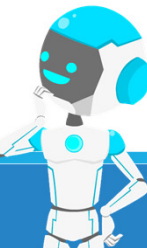
Epochs



Epochs, ou Iterações



E se usarmos poucas iterações?



Atributos

E SE TREINARMOS O MODELO COM UM SUBCONJUNTO DE ATRIBUTOS?

A solid blue horizontal bar spanning the width of the slide, located at the bottom.

Hiper parâmetro

E SE TREINARMOS O MODELO NO DOMÍNIO DE APENAS UM HIPER PARÂMETRO?

Multi-Fidelity



Usar subconjunto de partições



Parte dos dados, ou dados simplificados (imagens)



Poucas iterações



Poucos atributos



Testar domínio de apenas um hiper parâmetro



Curva de Aprendizado

Em vez de testar todos os hiper parâmetros possíveis...

- Ele vai até um limite tempo/aprendizado e congela
- Posteriormente pode-se decidir por quais configurações continuar
- Freeze-Thaw Bayesian Configuration

Hyper Band: Seleciona configurações aleatórias

BOBH: Bayesian Optimization and Hyper Band

Curva de Aprendizizado

