



**INSTITUTO FEDERAL**  
Ceará  
Campus Fortaleza



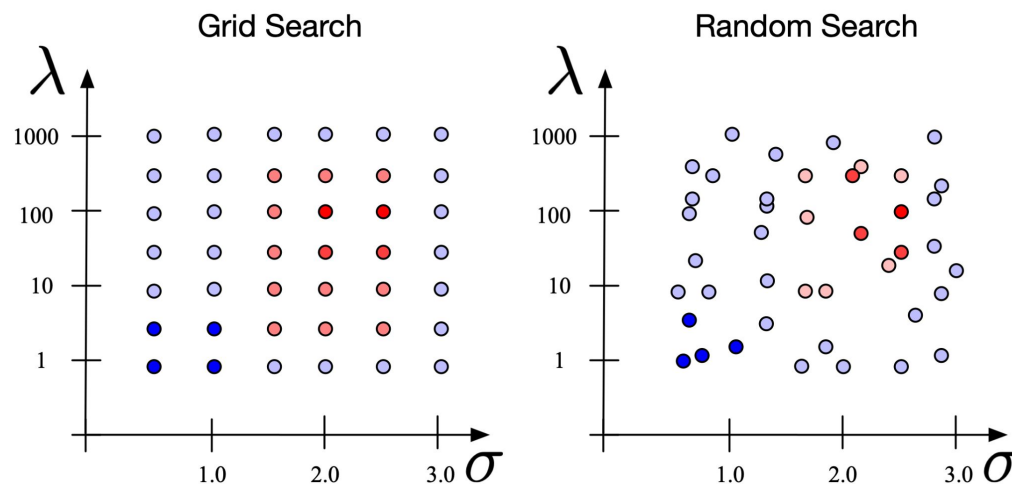
# Boas Práticas de *Machine Learning*

Prof. Matheus Santos  
[matheus.santos@lapisco.ifce.edu.br](mailto:matheus.santos@lapisco.ifce.edu.br)

1. *Grid Search X Random Search*
2. Divisão do *Dataset*
3. *Cross-Validation* e cálculo de métricas
4. Superfície de decisão

## Grid Search X Random Search

- Combinação em grade pré-definida X Combinação aleatória
- Eficiência computacional com grande espaço de hiperparâmetros
- Chance de encontrar hiperparâmetros ótimos



## ***Grid Search X Random Search***

- Exemplo para MLP com Dataset Iris
- Número de neurônios da camada oculta
- Grid Search [2, 500]
  - Tempo: 23 segundos
  - Resultado: 98% (Acc)
- Random Search 498 iterações [2, 500]
  - Tempo: 24 segundos
  - Resultado: 98% (Acc)

## Divisão do *Dataset*

- *Hold-Out*
- *Leave-One-Out (LOO)*
- *Bootstrapping*
- *Cross-Validation (K-Fold)*
- *Cross-Validation (Stratified K-Fold)*
- Etc.

## ***Cross-Validation*** e cálculo de métricas

- *K-fold* e *Stratified K-fold*
- Treino, teste, validação
- Métricas de validação - média e desvio
- Métricas de teste
- Teste estatístico

## ***Cross-Validation*** e cálculo de métricas

### **Métricas de Avaliação**

- Acurácia
- Precisão
- *Recall*
- *F1-Score*
- Acurácia balanceada, Kappa, MCC, etc.
- Matriz de confusão

## Superfície de Decisão

- Separação entre classes
- Linear e não linear
- Plot 2D e 3D



## Referências

1. A. Webb. Statistical Pattern Recognition. John Wiley & Sons, 2nd edition, 2002.
2. HAYKIN, Simon. Redes Neurais: Princípios e prática. Bookman, 2nd edition, 2001.
3. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner. "Gradient-based learning applied to document recognition." Proceedings of the IEEE, 86(11):2278-2324, November 1998.
4. M.-L. Zhang and Z.-H. Zhou, "A review on multi-label learning algorithms", IEEE transactions on knowledge and data engineering, vol. 26, no. 8, pp. 1819–1837, 2013.
5. Hastie, T., Tibshirani, R., Friedman, J., The elements of statistical learning: Data mining, inference, and prediction. Springer, 2009.
6. FISHER, Ronald A.; MARSHALL, Michael. Iris data set. RA Fisher, UC Irvine Machine Learning Repository, v. 440, p. 87, 1936.

# Obrigado pela atenção!

## Dúvidas?

Prof. Matheus Santos  
[matheus.santos@lapisco.ifce.edu.br](mailto:matheus.santos@lapisco.ifce.edu.br)