

# Curso: Aprendizagem de Máquina em Inteligência Artificial

Disciplina: Aprendizado Não Supervisionado

Prof. Marcelo Novaes de Rezende



# Single X Complete Linkage

Partindo de `linkages.ipynb`, explique a formação dos dendrogramas diferentes com `linkages single` e `complete`



Faculdade de Computação e Informática  
**Mackenzie**

# Cluster: ainda o caso prático

Partindo de `big_cluster_2.ipynb`, vamos analisar os cluster formados e encontrar a cidade que mais representa o cluster...



Faculdade de Computação e Informática  
**Mackenzie**

# KNN : K-Nearest-Neighbors



Faculdade de Computação e Informática  
**Mackenzie**

## **KNN : K-Nearest-Neighbors**

Tipicamente utilizado em aprendizado supervisionado (classificação/regressão).

É possível utilizá-lo em aprendizado não supervisionado?  
Vamos ver...



Faculdade de Computação e Informática  
**Mackenzie**

## KNN : K-Nearest-Neighbors

Com tensorflow playground, vamos ver a dificuldade de classificar datasets não separáveis linearmente.

KNN é uma simples e ótima opção para classificações “difíceis”, pois se baseia nos rótulos de “vizinhos próximos”

<https://playground.tensorflow.org/>



Faculdade de Computação e Informática  
**Mackenzie**

## **Implementação “raiz” de KNN (supervisionado)** **Partir de knn\_raiz.ipynb**



Faculdade de Computação e Informática  
**Mackenzie**

**KNN com sklearn (supervisionado)**

**Partindo do knn raiz, use sklearn e gere knn\_sklearn.ipynb**



Faculdade de Computação e Informática  
**Mackenzie**



## **Avaliação dos parâmetros do KNN no sklearn**



Faculdade de Computação e Informática  
**Mackenzie**

## **Como usar KNN no aprendizado não supervisionado?**

É possível conhecer os k-vizinhos mais próximos por algum critério de similaridade e, desta forma, criar uma visão do “mini cluster” com centroide na amostra analisada.



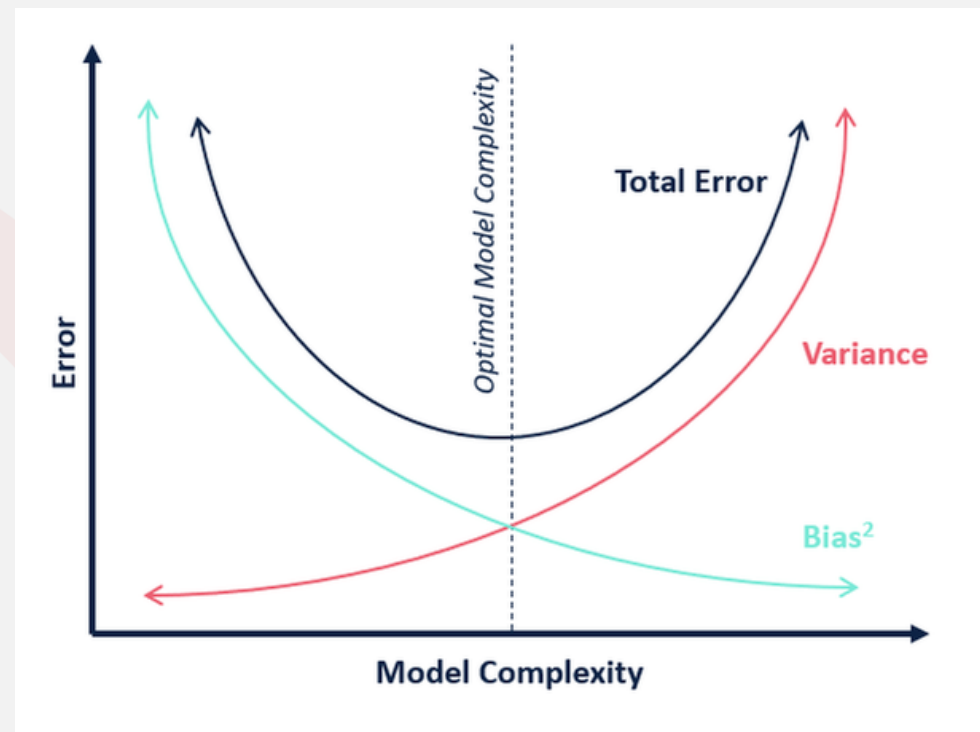
Faculdade de Computação e Informática  
**Mackenzie**

## Regularização...Por que?



Faculdade de Computação e Informática  
**Mackenzie**

Nosso objetivo é minimizar o erro total ajustando a complexidade do modelo. A regularização permite tornar menos complexos os modelos com alguns “artifícios”



Observação de Overfitting e Underfitting

Regressão linear

Usar notebook over\_under.ipynb



Faculdade de Computação e Informática  
**Mackenzie**

## Regularização L1 e L2...penalizando parâmetros altos

L1 Regularization

$$\text{Cost} = \sum_{i=0}^N (y_i - \sum_{j=0}^M x_{ij} W_j)^2 + \lambda \sum_{j=0}^M |W_j|$$

L2 Regularization

$$\text{Cost} = \underbrace{\sum_{i=0}^N (y_i - \sum_{j=0}^M x_{ij} W_j)^2}_{\text{Loss function}} + \lambda \underbrace{\sum_{j=0}^M W_j^2}_{\text{Regularization Term}}$$



# Dropout

O dropout é um método de regularização que simula o ensemble de várias redes via “morte” aleatória de alguns perceptrons ao longo do treinamento.

Isso gera diferentes “arquiteturas” ao longo do treinamento e é uma forma interessante de combater overfitting....

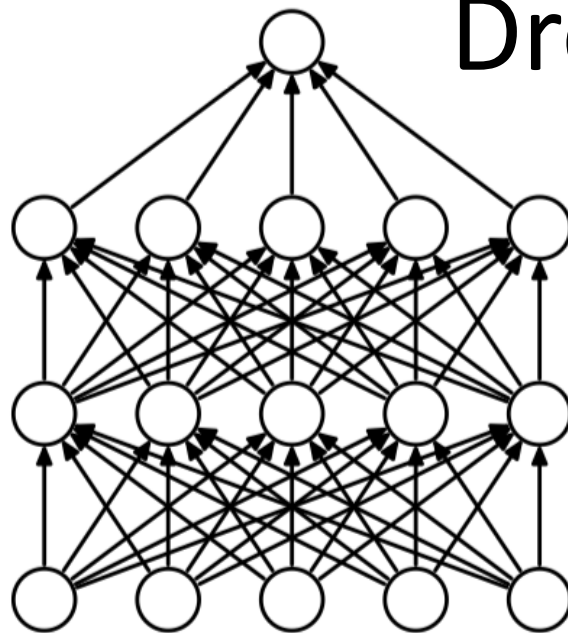
*O livro Deep Learning aprofunda o assunto no capítulo 7. pag 255*

*<https://www.deeplearningbook.org/contents/regularization.html>*

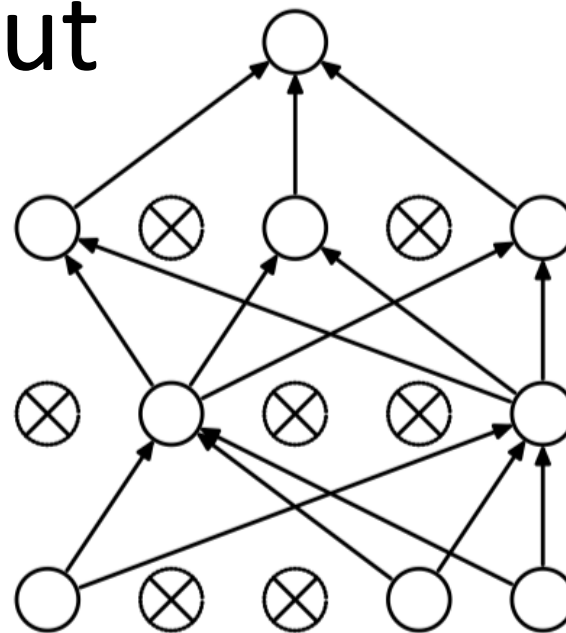


**Faculdade de Computação e Informática  
Mackenzie**

# Dropout



(a) Standard Neural Net



(b) After applying dropout.





Vamos aplicar dropout e normalização L1, L2 em uma rede neural MLP..., observar os efeitos do “remédio” e discutir...

Notebook drop\_reg\_2.ipynb



Faculdade de Computação e Informática  
**Mackenzie**

Sobrou tempo? Vamos usar KNN no Dataset do Estado de São Paulo.

Vamos avaliar a média dos parâmetros do 10 municípios “neighbors”..



Faculdade de Computação e Informática  
**Mackenzie**

Até a próxima aula

---

OBRIGADO!

Prof Marcelo Rezende  
email rezendemn@gmail.com