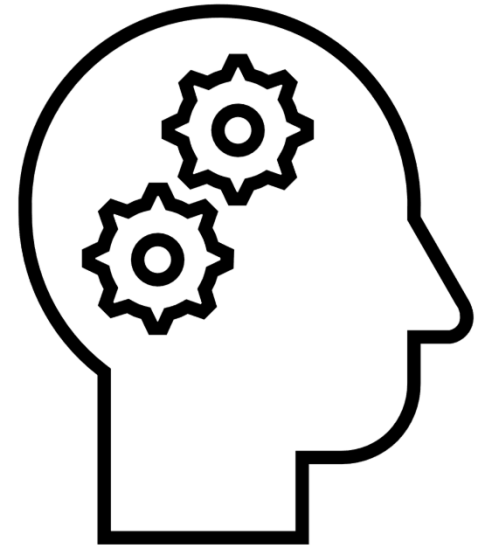


Hiper Parâmetros

- `hidden_layer_sizes`: Lista de neurônios na camada oculta
- `batch_size`: numero de instâncias por batch usadas para atualizar a rede
- `Epochs (max_iter)`: número de vezes que os dados passam pela rede
- `early_stopping`: parar quando não há melhora na performance
- `learning_rate`: o quando os pesos devem ser atualizados a cada iteração



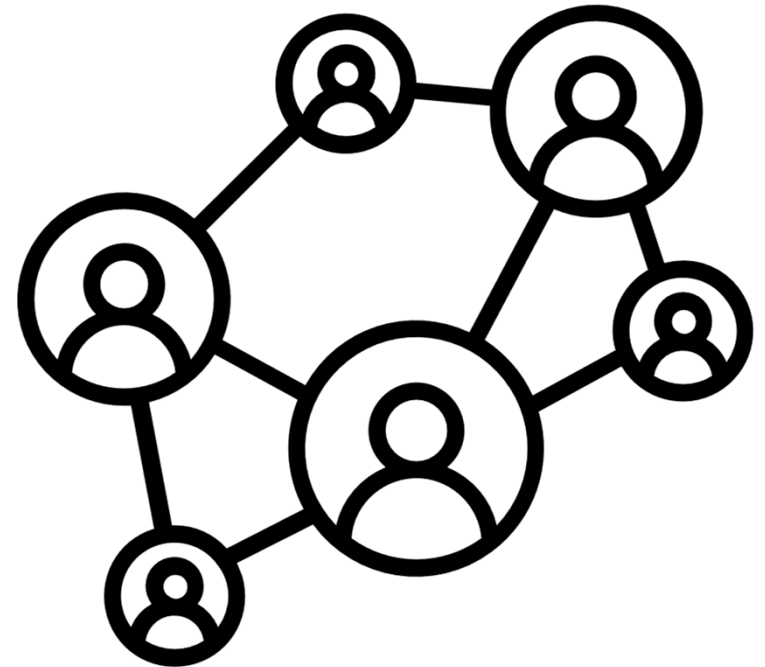
Topologia da Rede:

- `hidden_layer_sizes`
- Adição de camadas



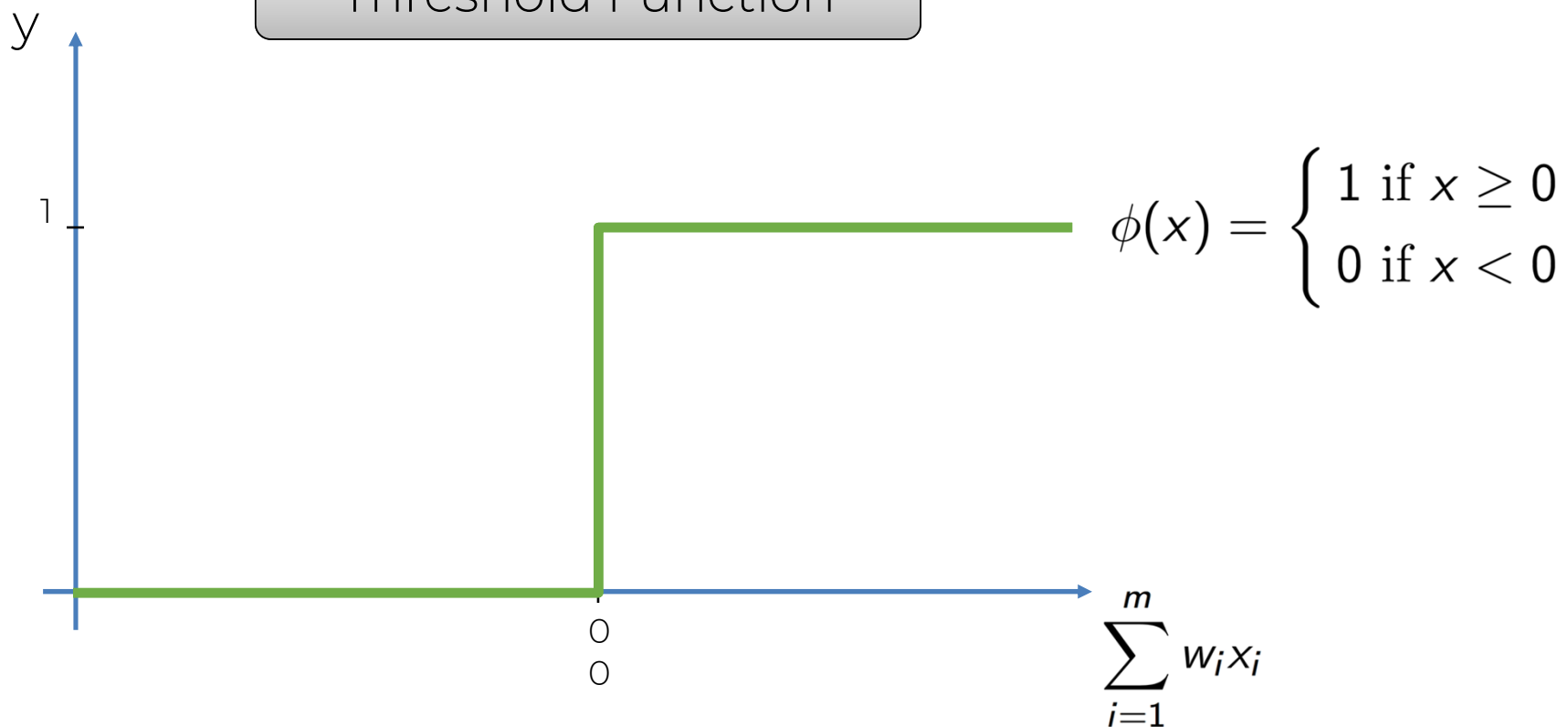
activation

- logistic, tanh, relu etc.

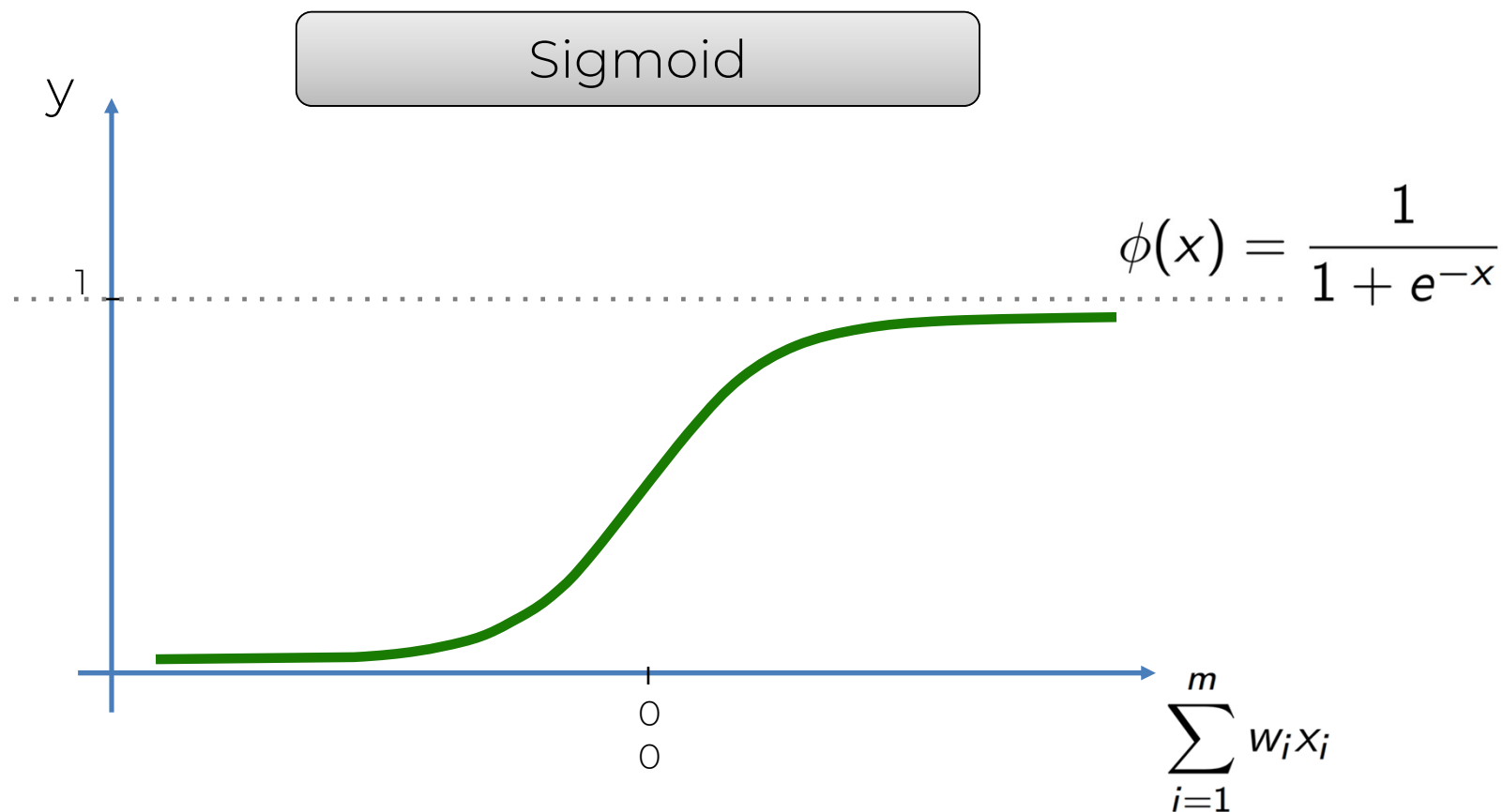


A função de ativação

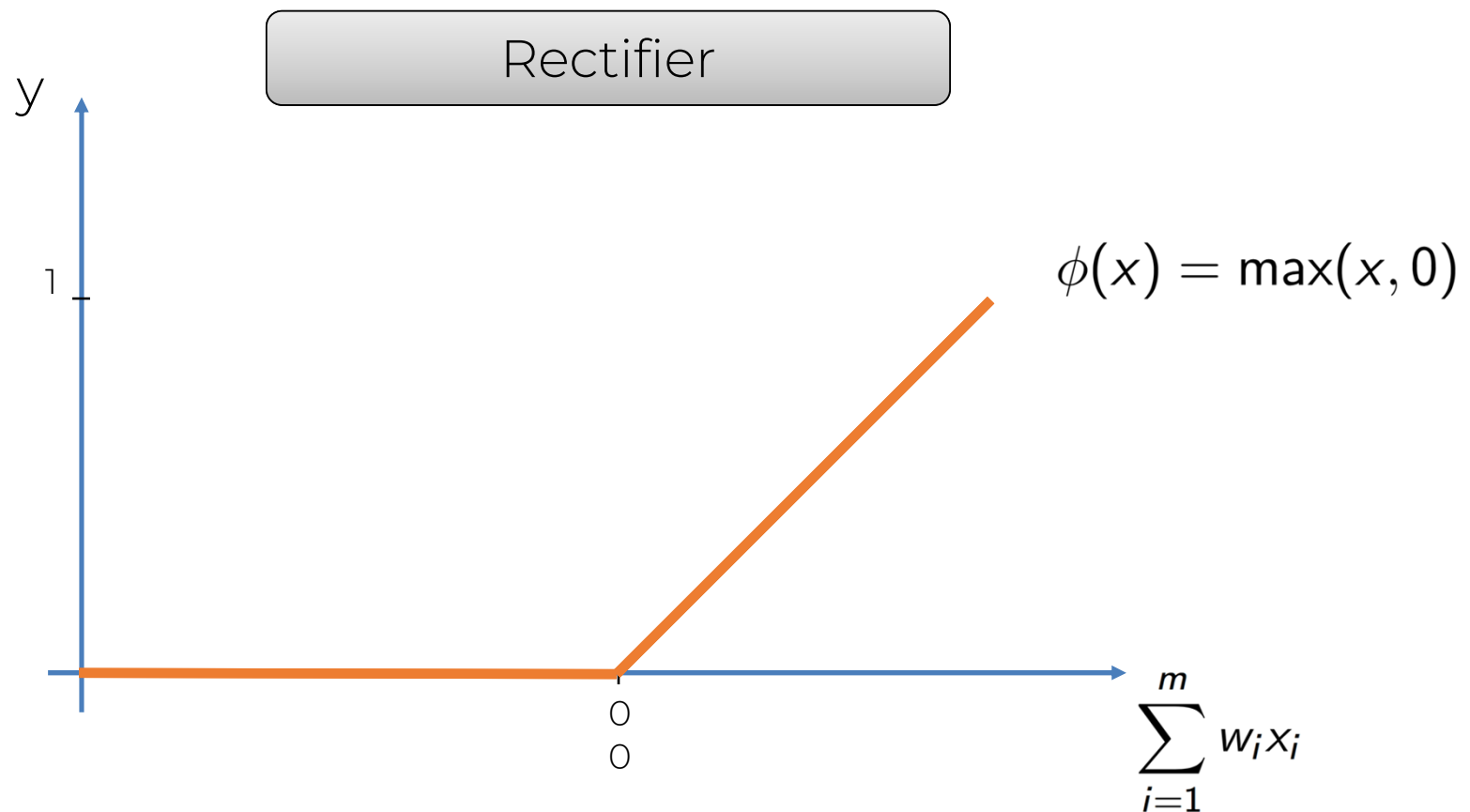
Threshold Function



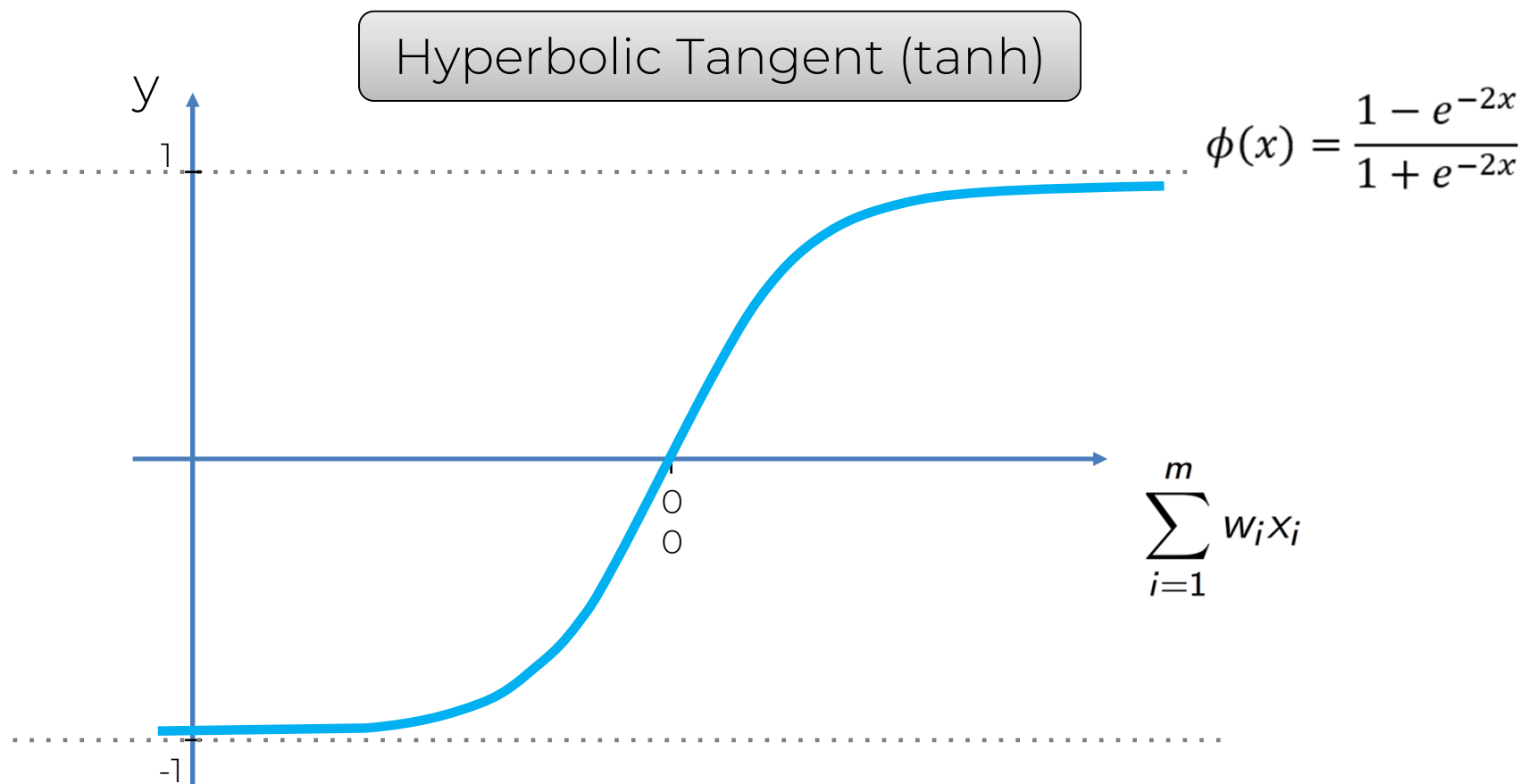
A função de ativação



A função de ativação



A função de ativação



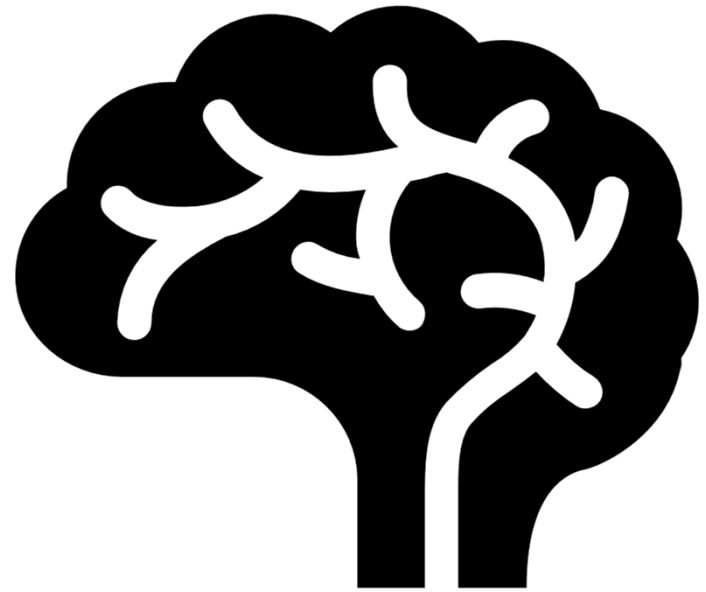
Regularização

- Técnicas para evitar overfitting
 - Early Stopping: interrompe o treinamento quando a performance se mantém estável por muitas epochs ou mesmo piora
 - Data augmentation: gera novos exemplos de dados de treinamento
 - L1 e L2
 - Dropout



Regularização L1 e L2

- Técnica que utiliza um termo de regularização para minimizar o risco de overfitting
- Este termo é multiplicado por um hiper parâmetro alfa
 - L1 (Lasso regularization): Adiciona um termo à loss function proporcional ao valor absoluto dos pesos
 - L2 (Ridge regularization): Adiciona um termo à loss function proporcional ao quadrado dos pesos
- O objetivo é minimizar overfitting, adicionando o termo a Loss function (que o treino busca minimizar)
- Este termo adicional tende a deixar os pesos menores, o que reduz a complexidade do modelo e minimiza overfitting
- Alpha: Usado para adicionar um termo





Dropout

- Técnica de regularização que elimina neurônios de acordo com uma probabilidade
- Isso obriga a rede a aprender de forma robusta e evitar overfitting

Gradient Descent

- Algoritmo de otimização usado para ajustar os pesos de forma a minimizar os erros entre o previsto e os dados de treino.
- Basicamente, ele diz a direção e o quanto estes pesos devem ser ajustados
- Momentum: ajuste do gradiente antes de novo uso, tornando a atualização mais estável

