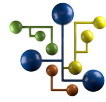




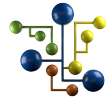
Formação Inteligência Artificial



Deep Learning I



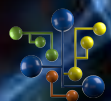
Redes Neurais Artificiais



Redes Neurais Artificiais



A Importância das Redes Neurais Artificiais



Data Science
Academy

Data Science Academy/raphaelbsfontenelle@gmail.com 615c46dde32fc361b30c9ec2



Data Science Academy

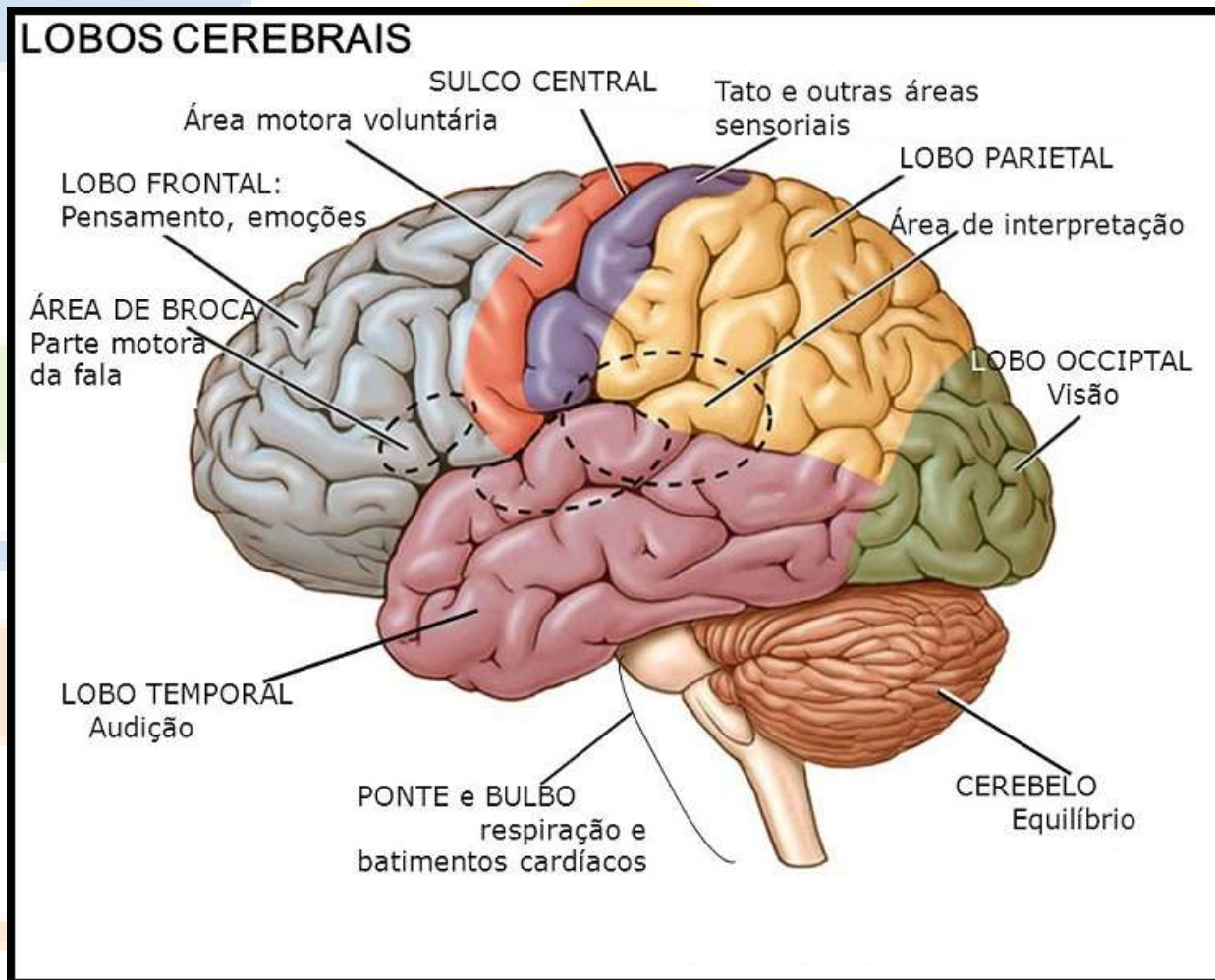


Data Science Academy



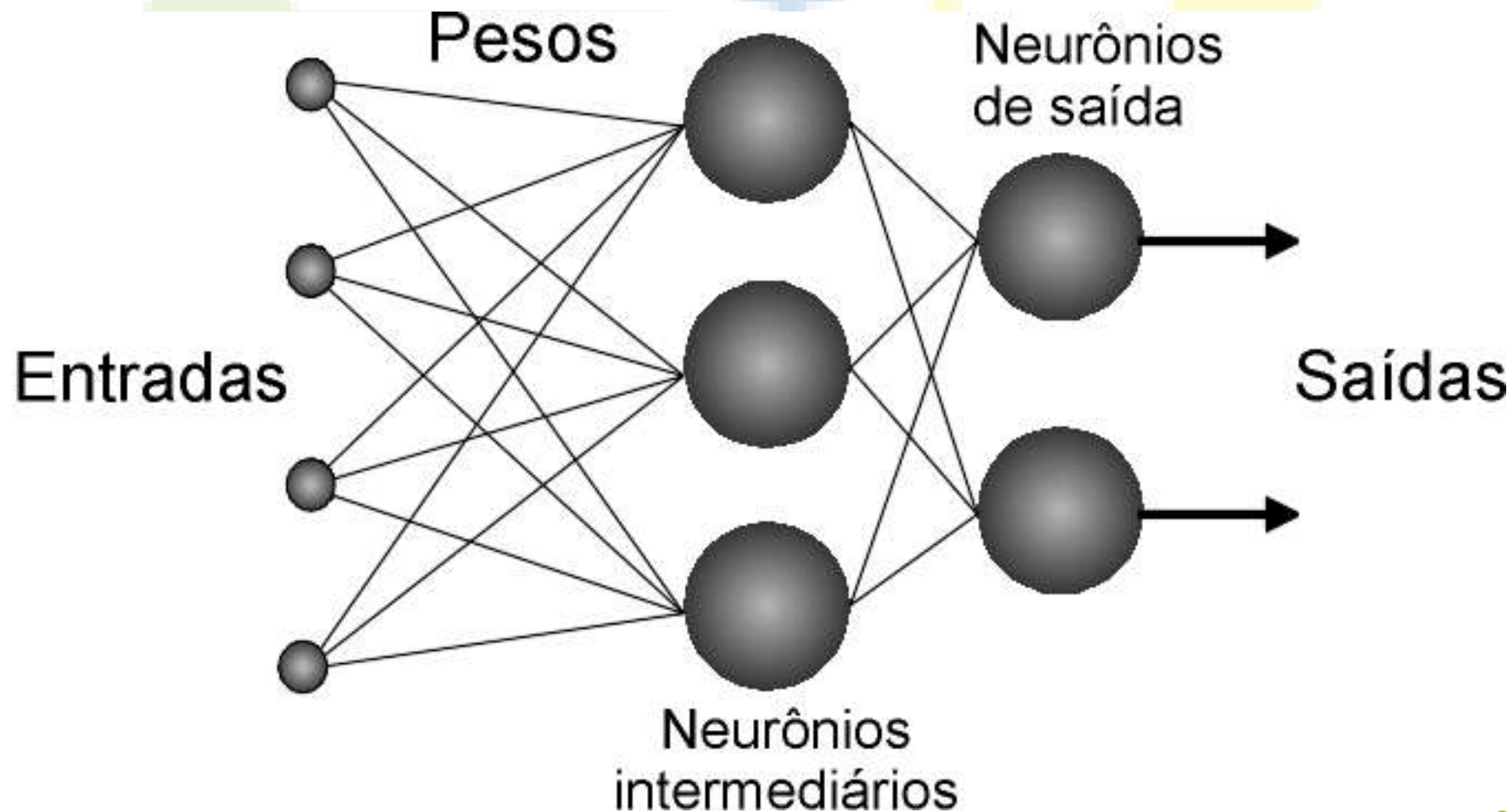
A Importância das Redes Neurais Artificiais

O **cérebro humano** tem sido extensamente estudado, mas ainda não somos capazes de **entender completamente** o seu funcionamento.





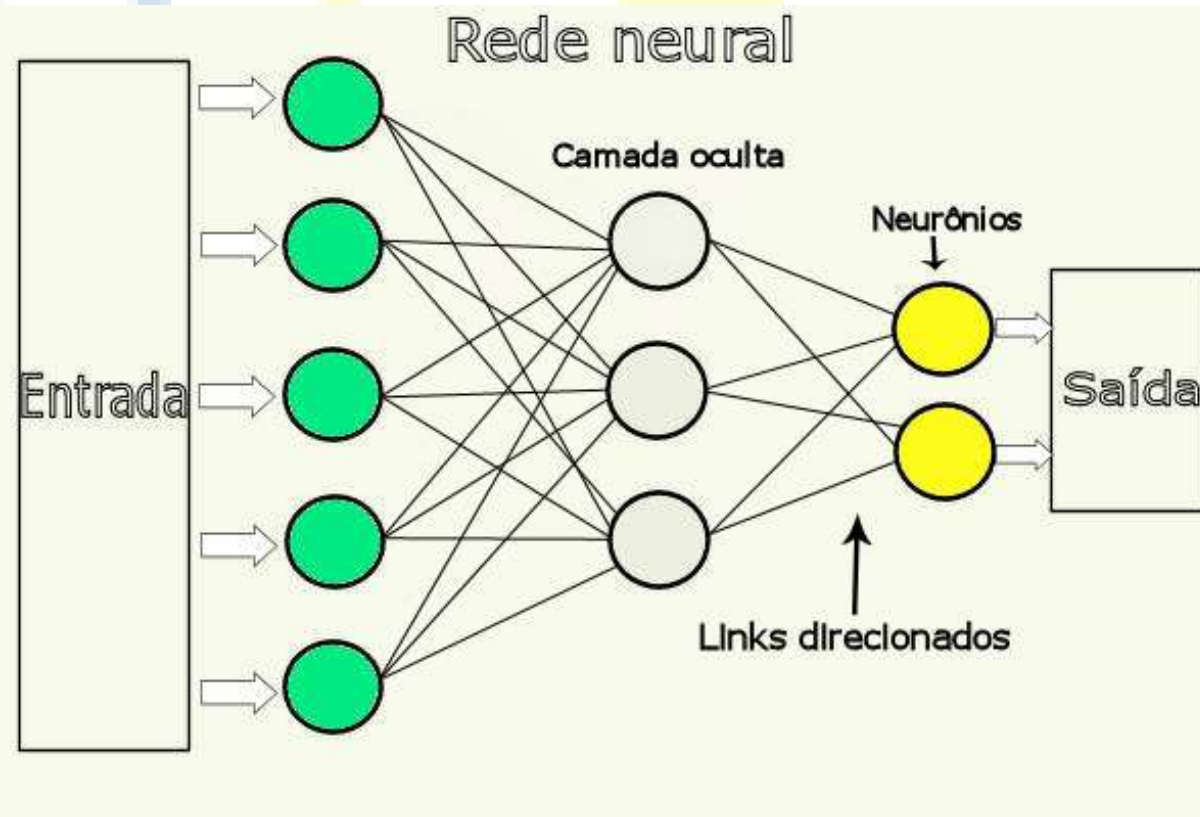
A Importância das Redes Neurais Artificiais





A Importância das Redes Neurais Artificiais

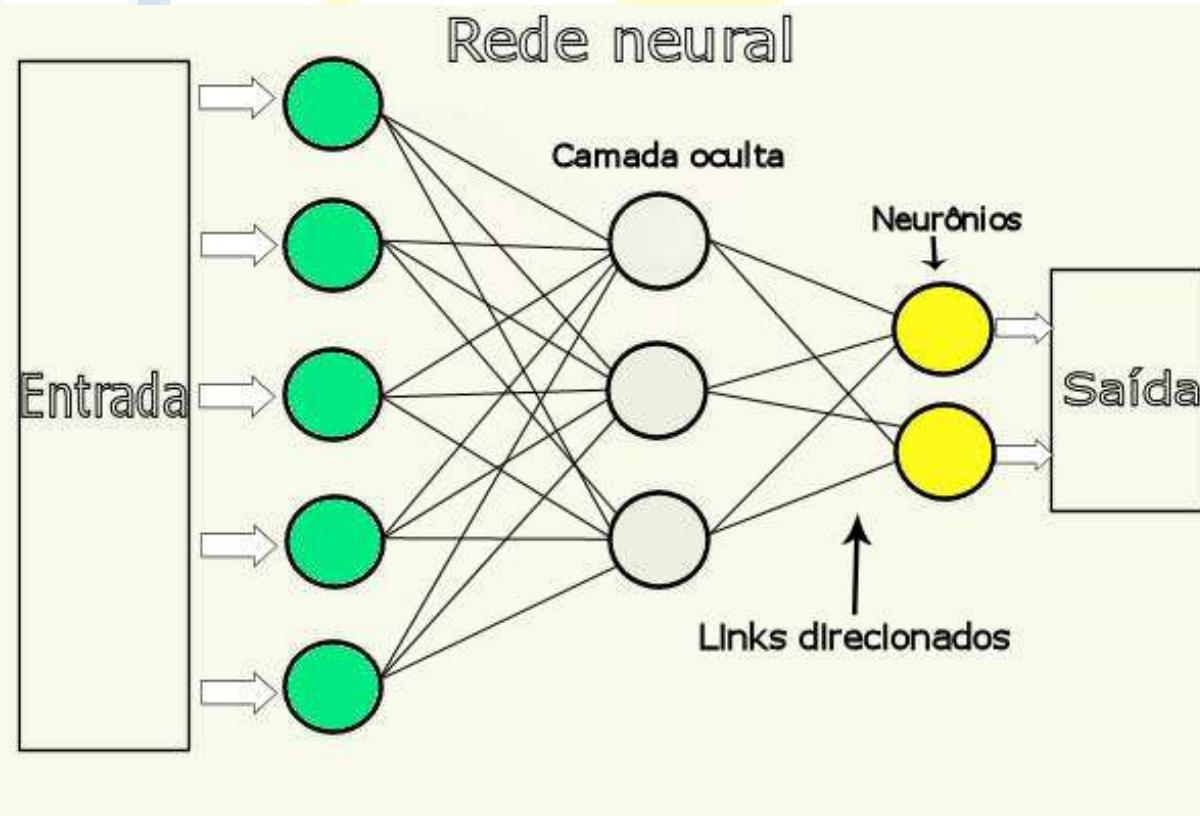
Com Redes Neurais Artificiais, tentamos programar um computador de modo a fazê-lo imitar um comportamento inteligente (saber jogar xadrez, compreender e manter um diálogo, traduzir línguas estrangeiras, resolver problemas de matemática, etc.)





A Importância das Redes Neurais Artificiais

Uma Rede Neural Artificial modela a relação entre um conjunto de sinais de entrada e um sinal de saída usando um modelo derivado de nossa compreensão de como um cérebro biológico responde a estímulos de entradas sensoriais.





A Importância das Redes Neurais Artificiais

- Cérebro humano – 85 bilhões de neurônios
- Cérebro de um gato – 1 bilhão de neurônios
- Cérebro de um rato – 75 milhões de neurônios
- Cérebro de uma barata – 1 milhão de neurônios





A Importância das Redes Neurais Artificiais



Agora você entende porque a computação paralela em GPU's está acelerando o desenvolvimento de sistemas inteligentes, pois somos capazes de processar cada vez mais dados em redes neurais artificiais com cada vez mais neurônios



A Importância das Redes Neurais Artificiais

Programas de reconhecimento de voz e escrita

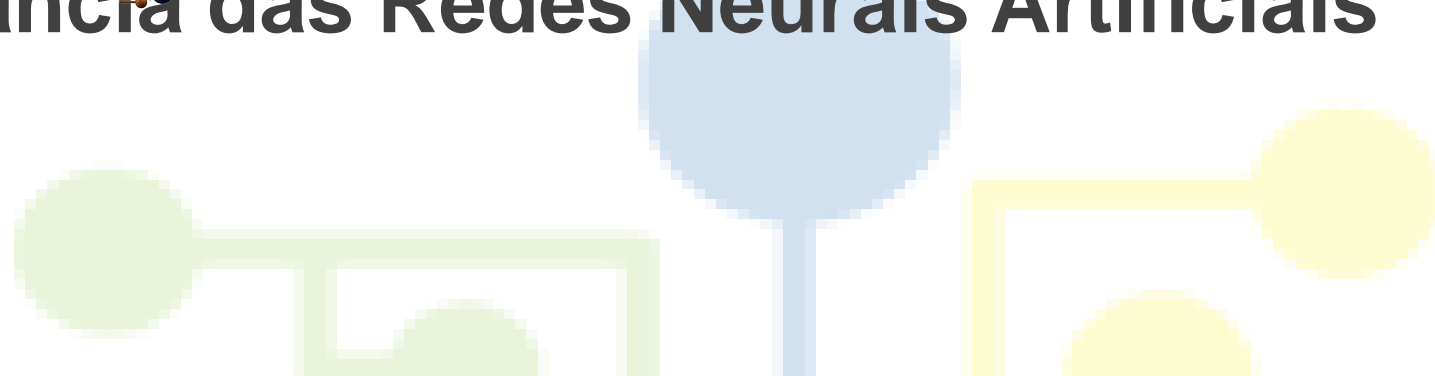
Automação de dispositivos inteligentes

Modelos sofisticados de padrões climáticos





A Importância das Redes Neurais Artificiais

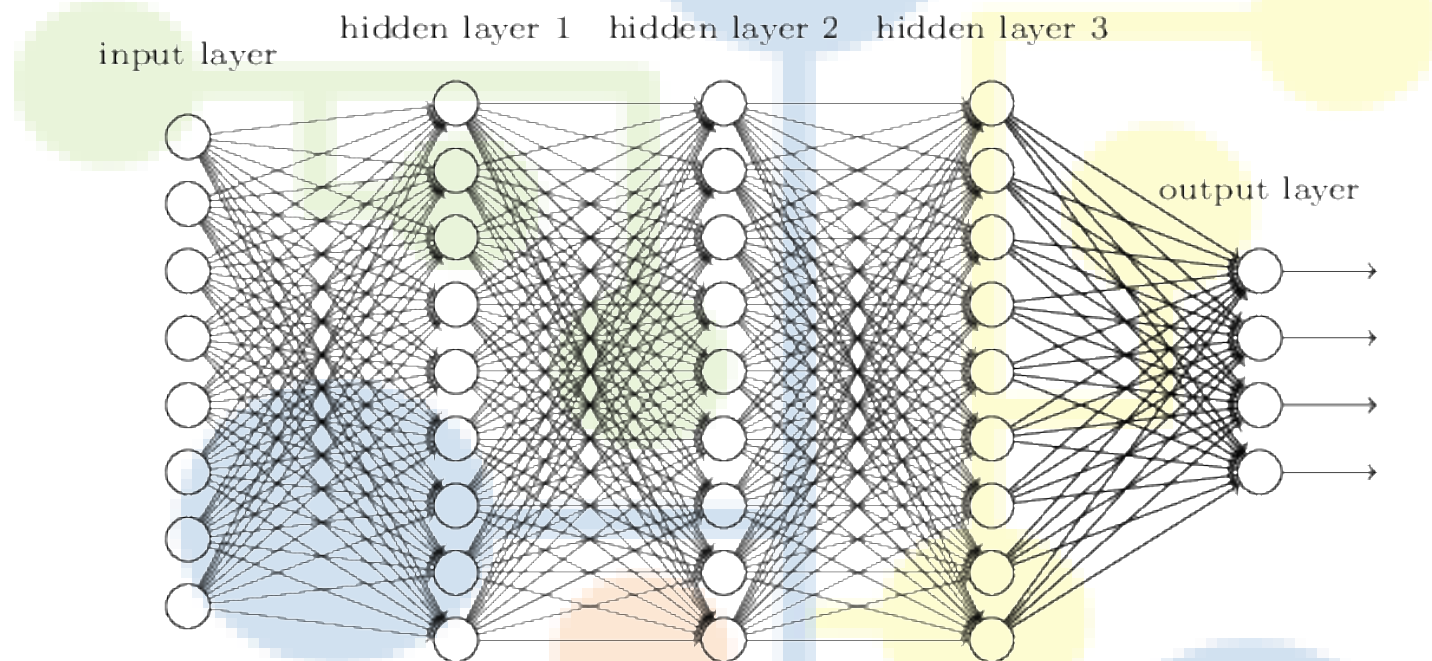


As Redes Neurais Artificiais são modelos versáteis que podem ser aplicadas a quase todas as tarefas de aprendizagem: classificação, previsão numérica e mesmo reconhecimento não supervisionado de padrões





A Importância das Redes Neurais Artificiais



As redes neurais artificiais são melhor aplicadas a problemas onde os dados de entrada e os dados de saída são bem definidos ou, pelo menos, bastante simples, mas o processo que relaciona a entrada com a saída é extremamente complexo



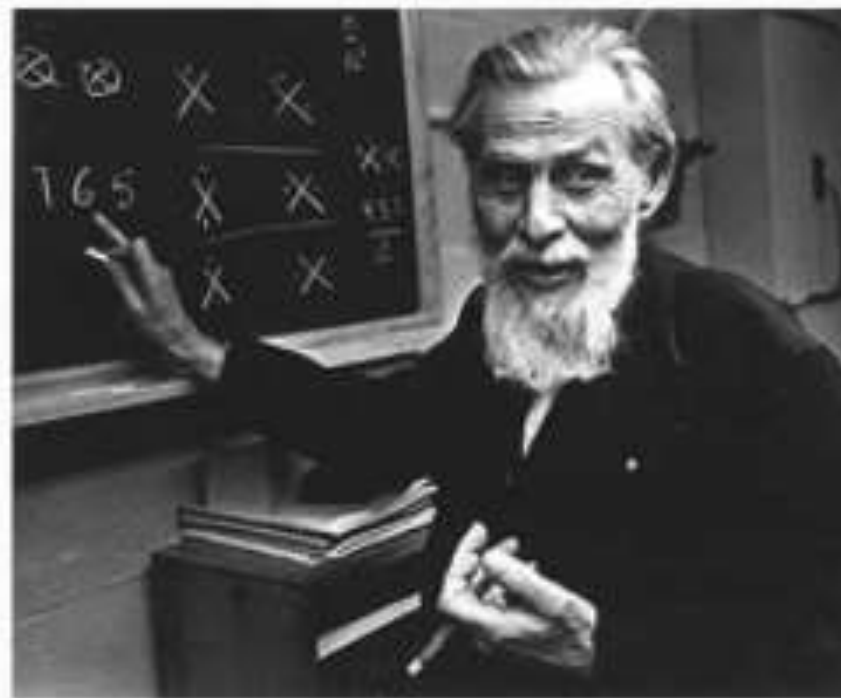
Como Surgiram as Redes Neurais Artificiais?



Como Surgiram as Redes Neurais Artificiais?



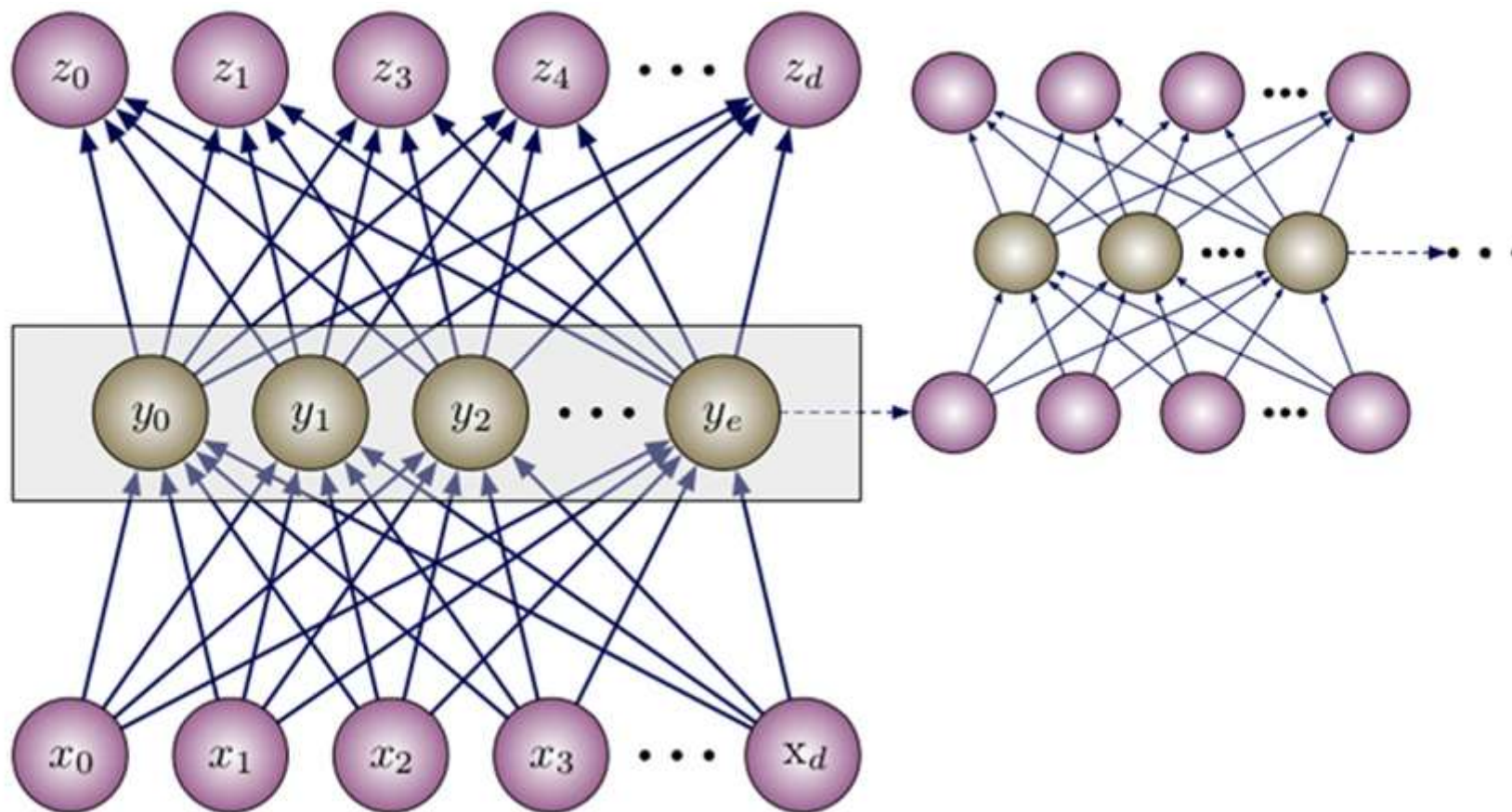
Walter Pitts



Warren McCulloch

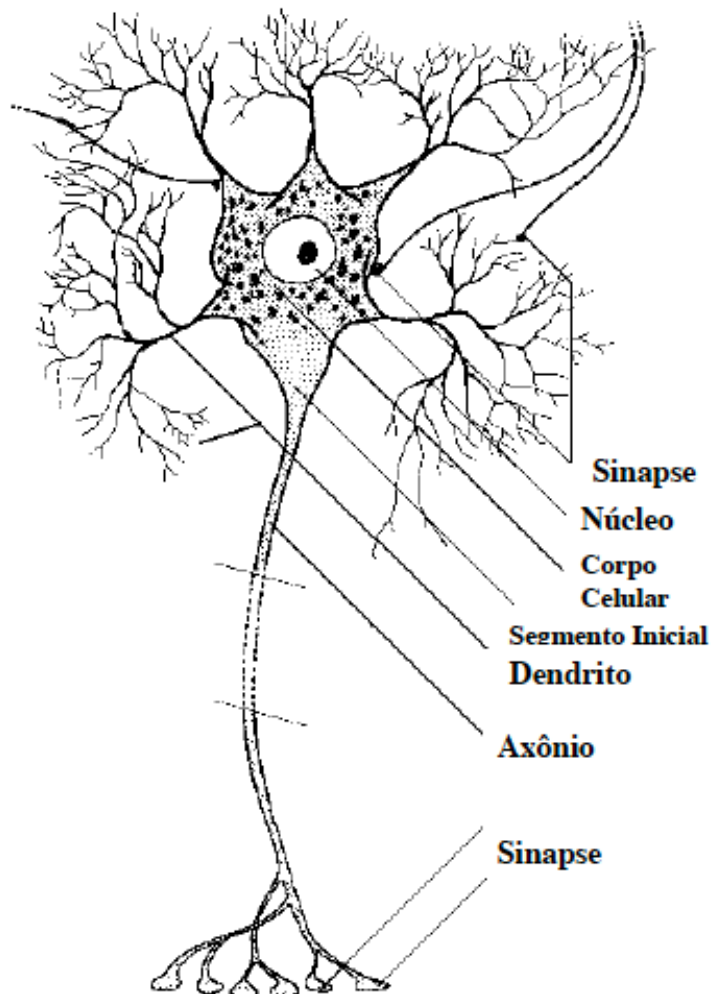


Como Surgiram as Redes Neurais Artificiais?



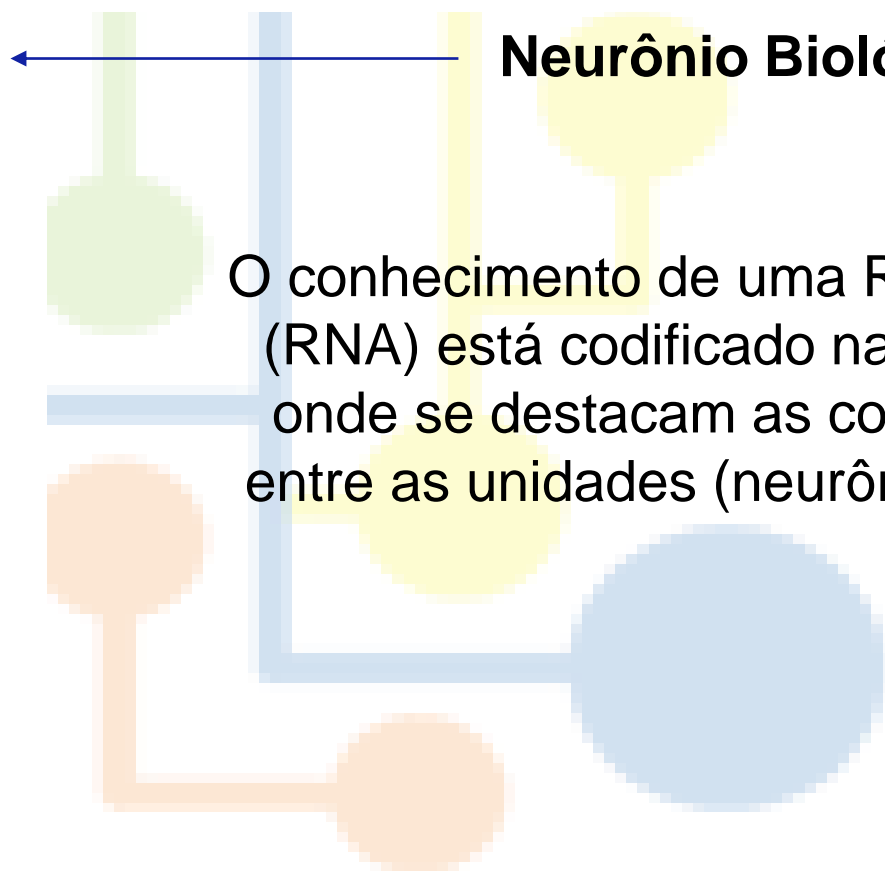


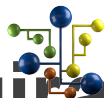
Como Surgiram as Redes Neurais Artificiais?



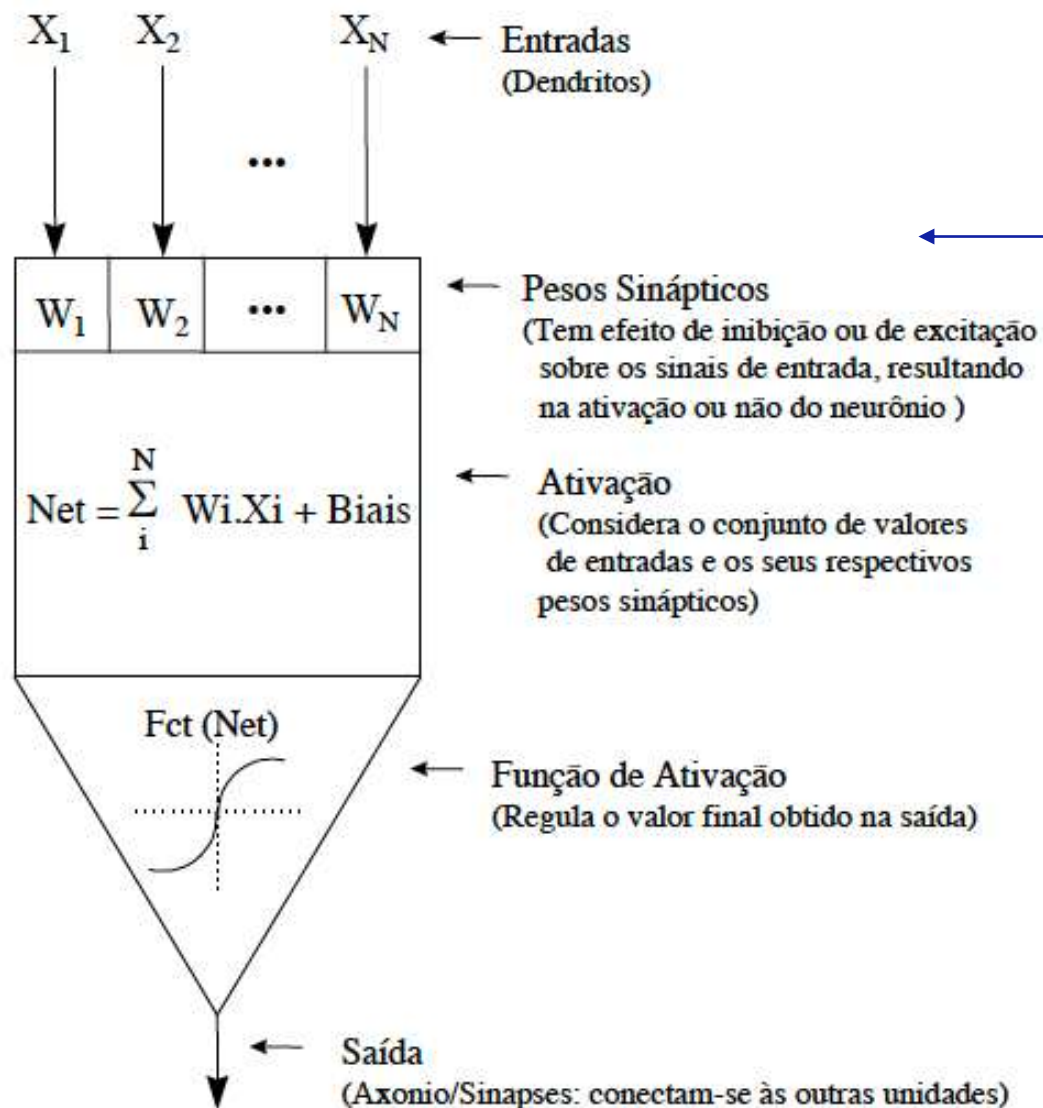
Neurônio Biológico

O conhecimento de uma Rede Neural Artificial (RNA) está codificado na estrutura da rede, onde se destacam as conexões (sinapses) entre as unidades (neurônios) que a compõe





Como Surgiram as Redes Neurais Artificiais?



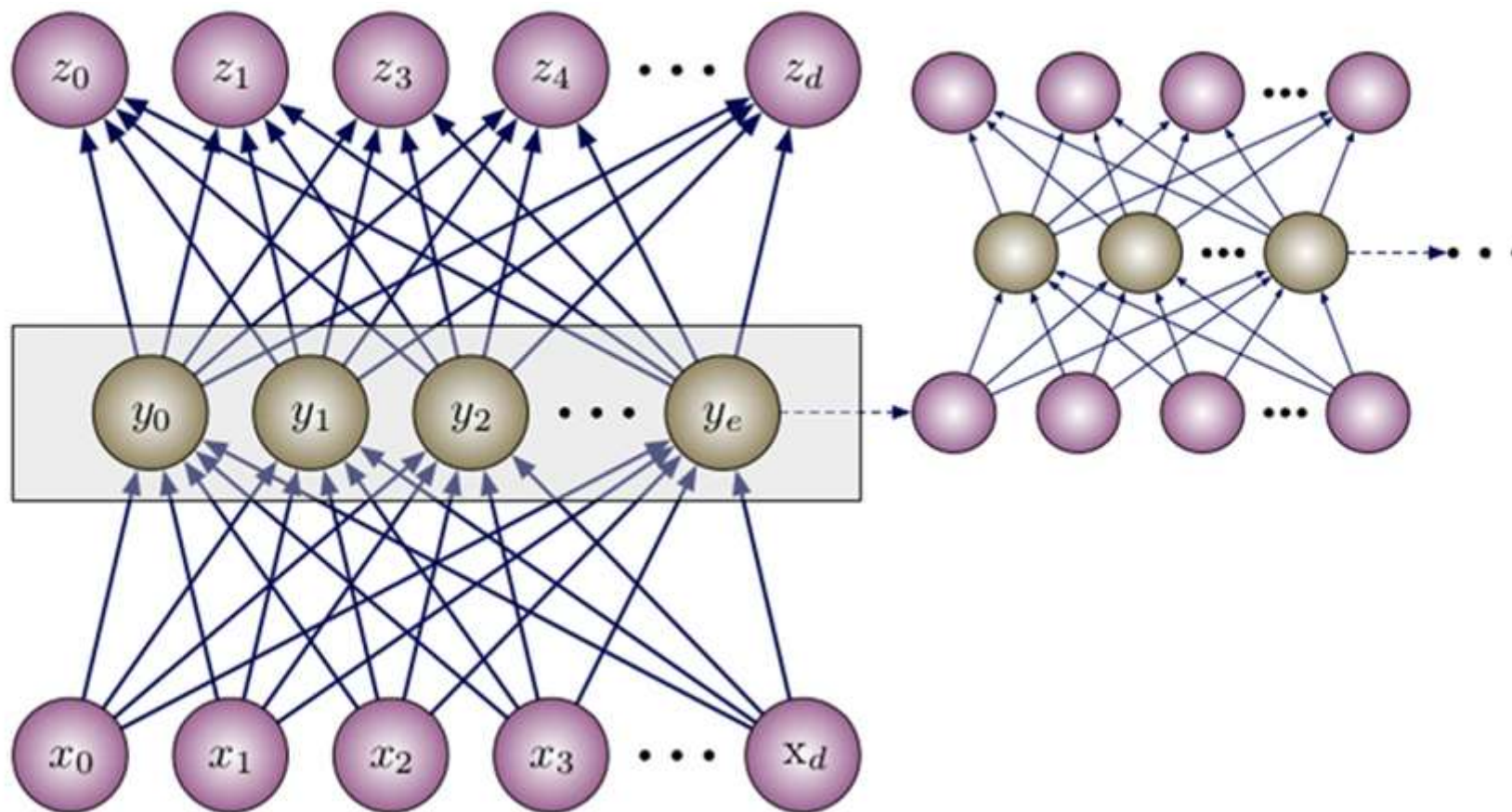
Neurônio Matemático

O conhecimento de uma Rede Neural Artificial (RNA) está codificado na estrutura da rede, onde se destacam as conexões (sinapses) entre as unidades (neurônios) que a compõe



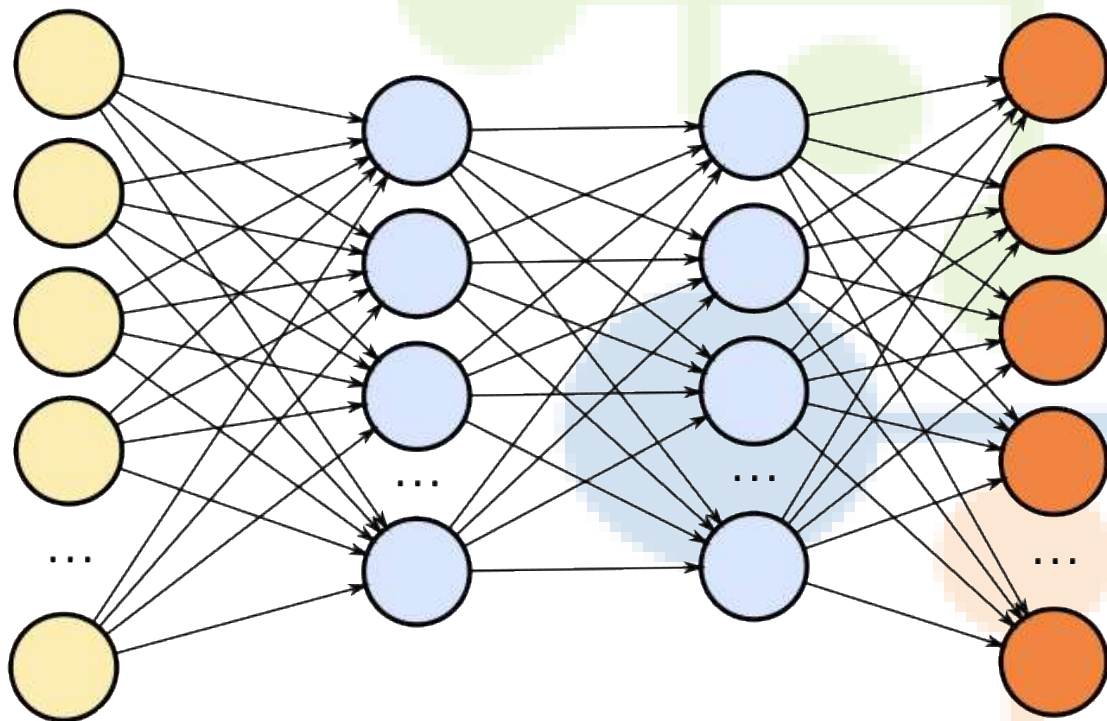


Como Surgiram as Redes Neurais Artificiais?





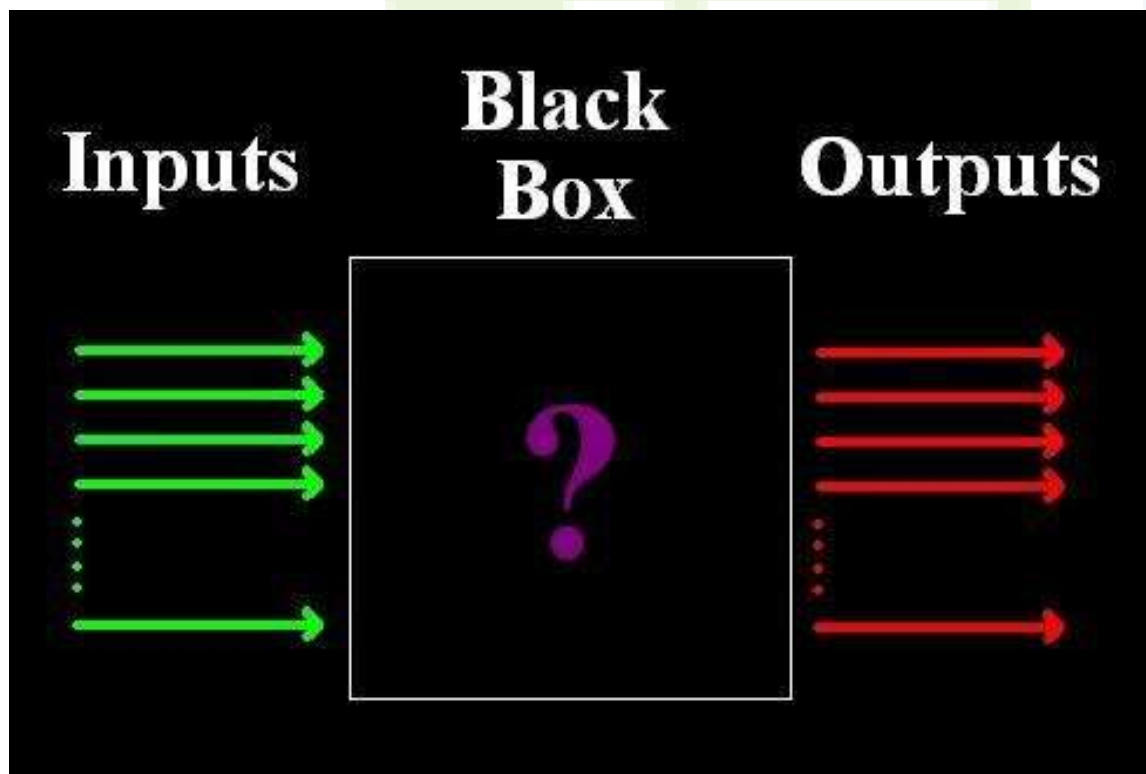
Como Surgiram as Redes Neurais Artificiais?



Um dos benefícios das redes diz respeito ao tratamento de um problema clássico da Inteligência Artificial, que é a representação de um universo não-estacionário (onde as estatísticas mudam com o tempo)



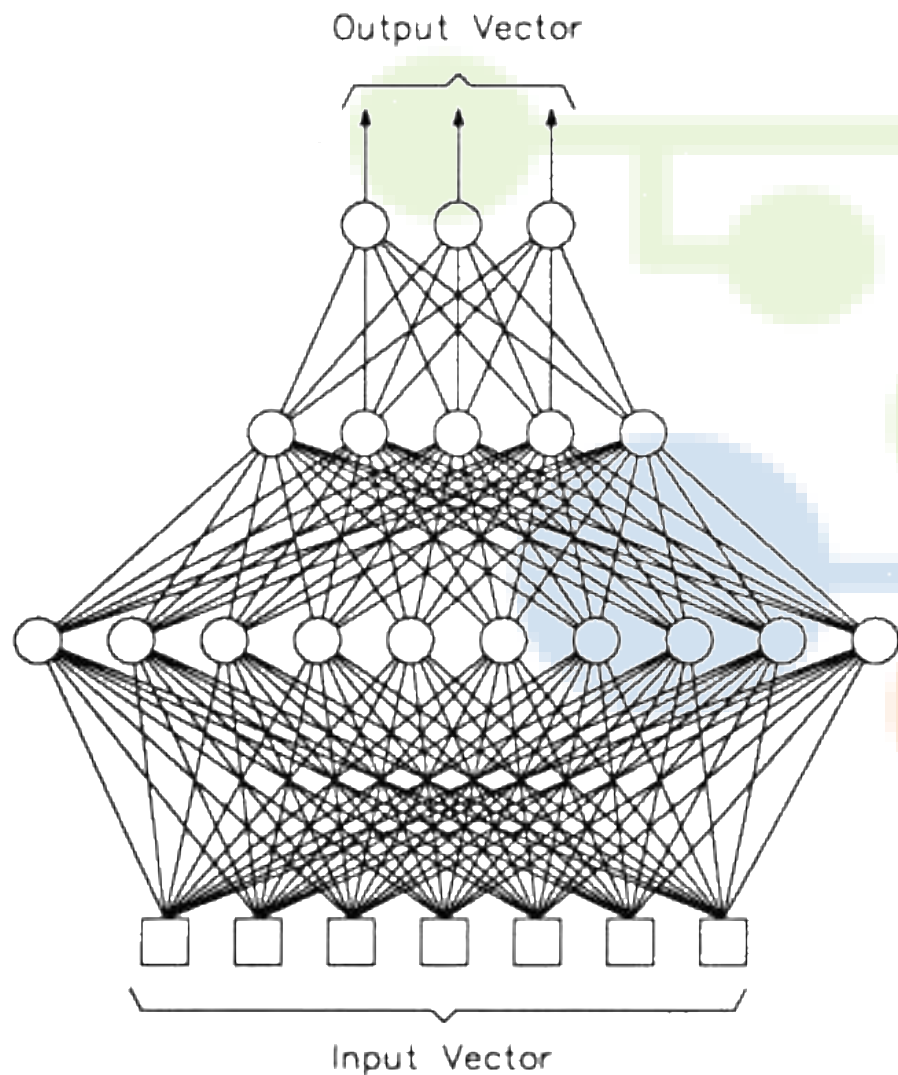
Como Surgiram as Redes Neurais Artificiais?



Uma desvantagem das redes neurais é o fato delas , normalmente, serem uma "caixa preta"



Como Surgiram as Redes Neurais Artificiais?

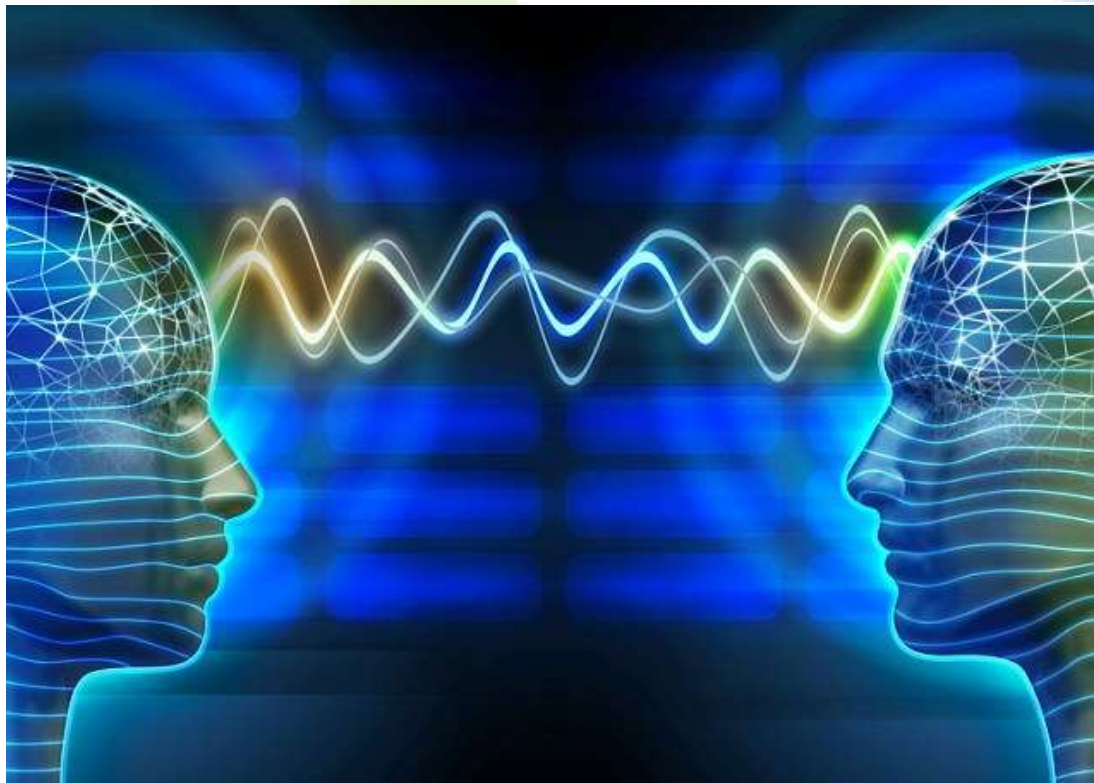


A solução de problemas através das RNAs é bastante atrativa, pois o paralelismo constitui-se na característica principal das RNAs, onde esta cria a possibilidade de um desempenho superior em relação a solução de problemas baseados nos modelos convencionais.





Como Surgiram as Redes Neurais Artificiais?

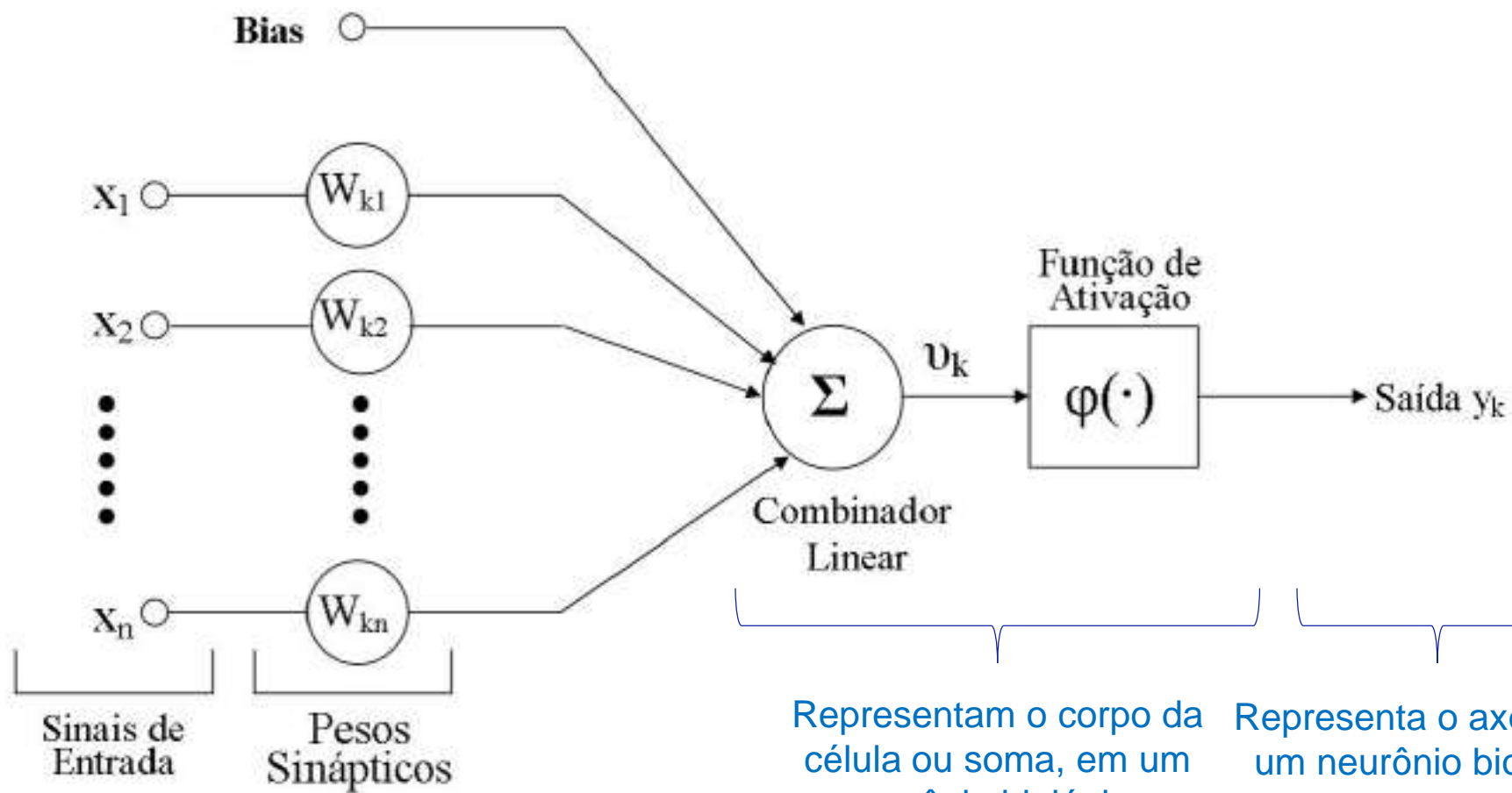


A generalização está associada à capacidade da rede em aprender através de um conjunto reduzido de exemplos, e posteriormente, dar respostas coerentes a dados não apresentados a rede.

O Neurônio Matemático

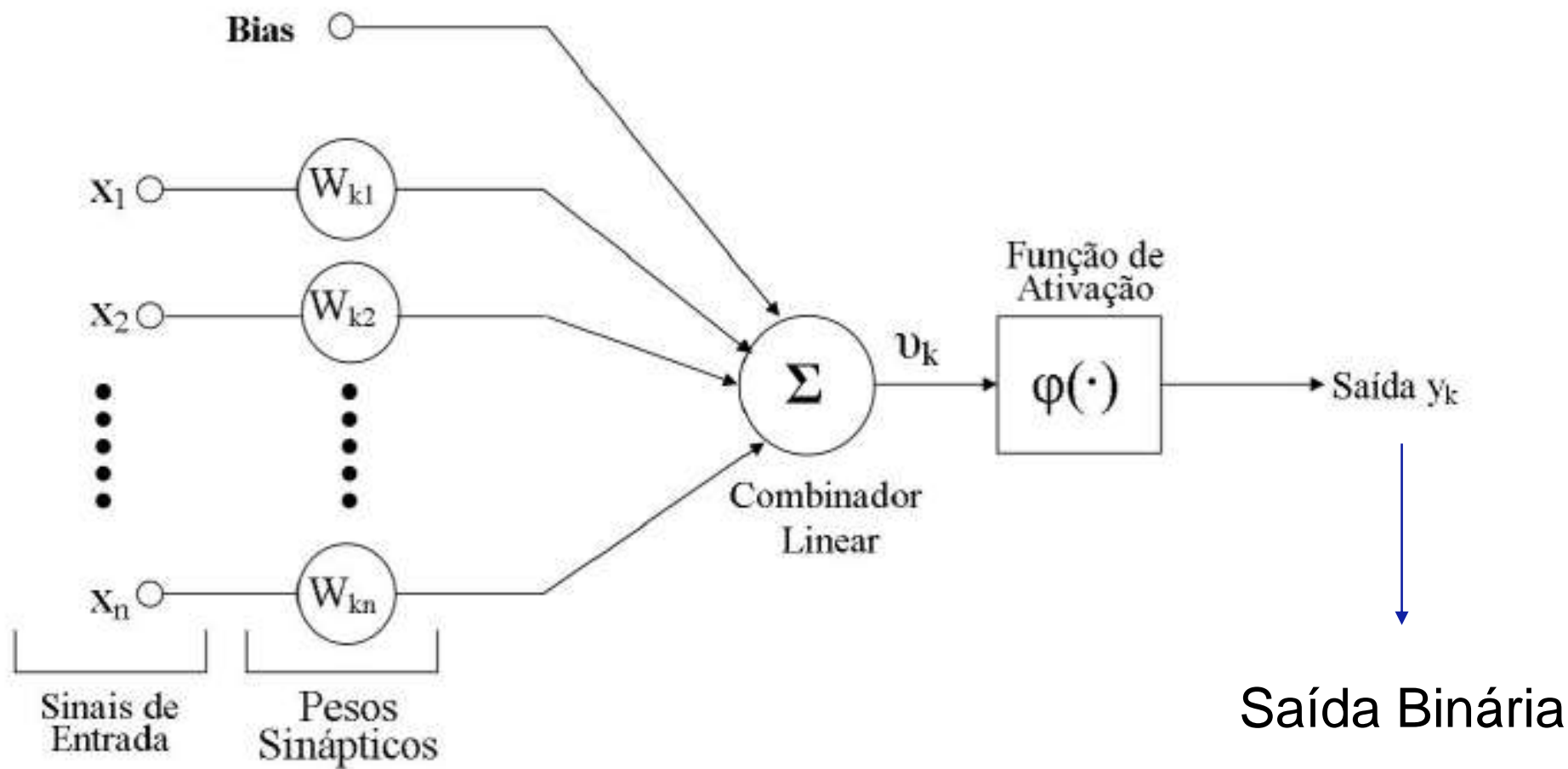


O Neurônio Matemático



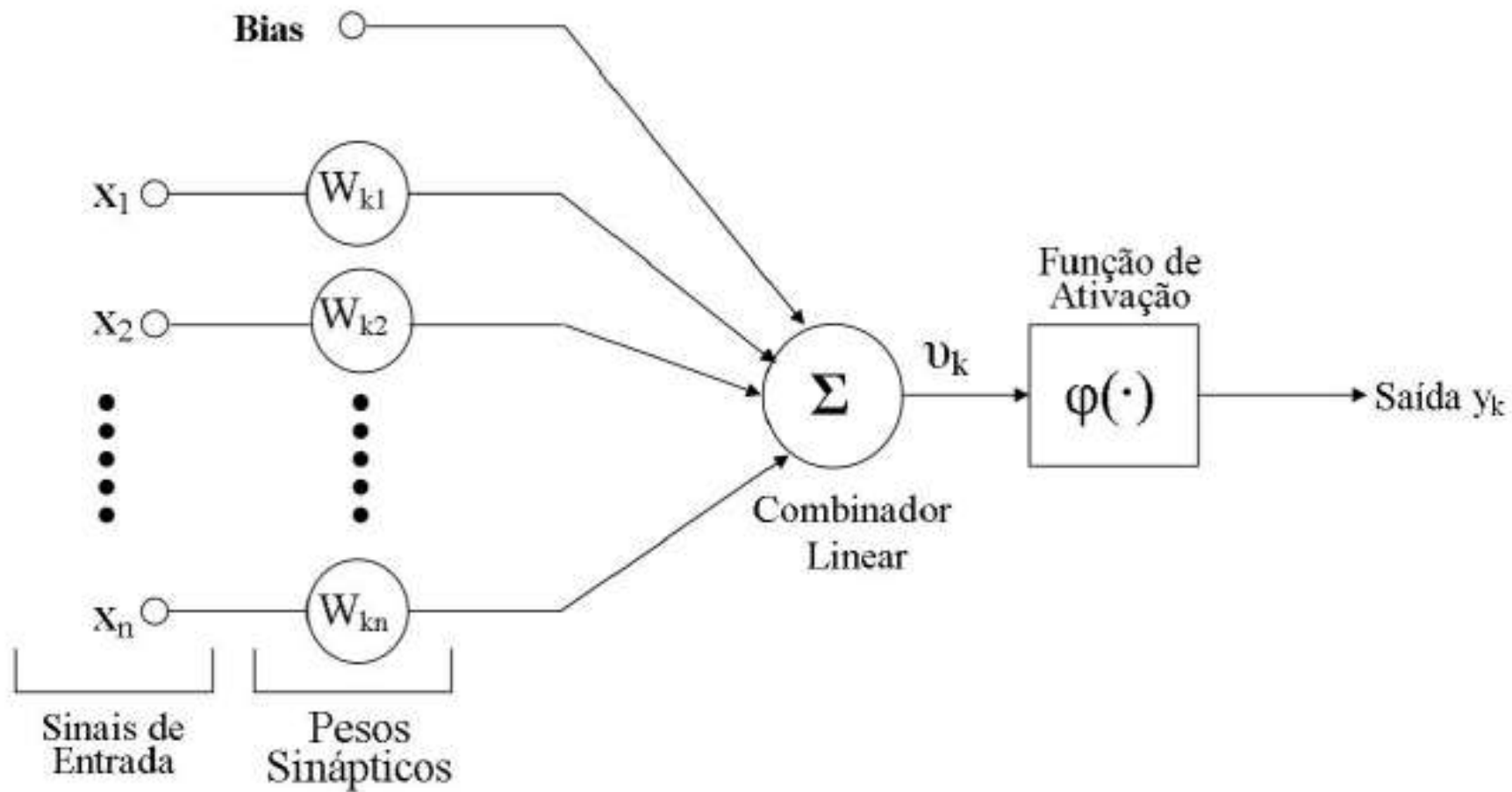


O Neurônio Matemático



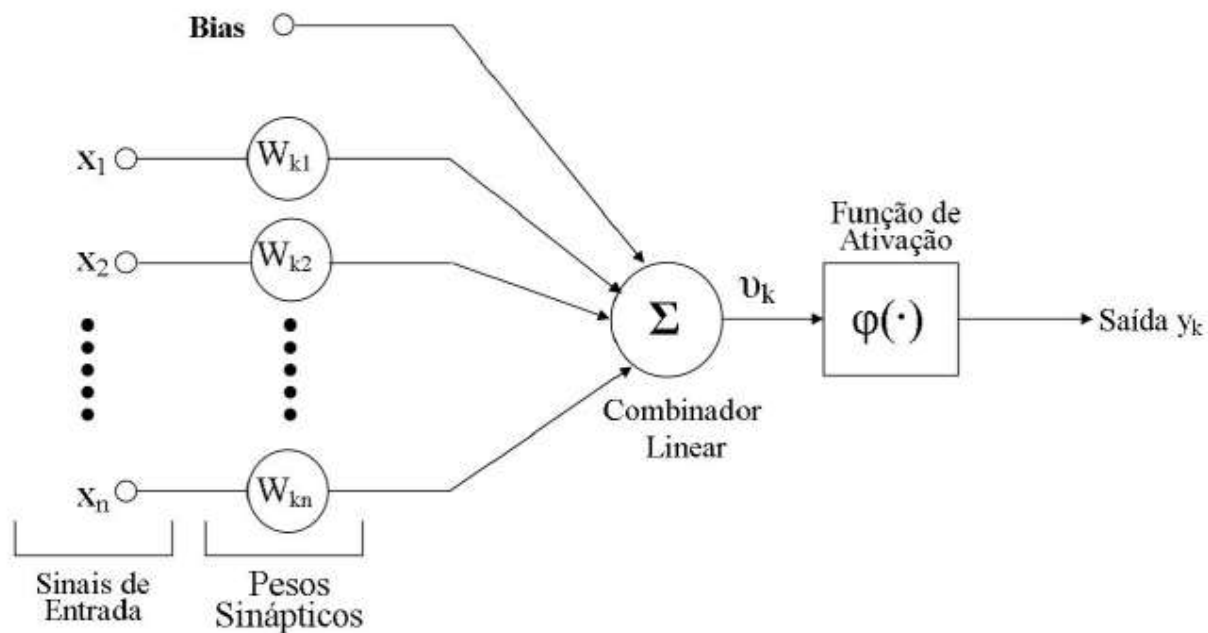


O Neurônio Matemático





O Neurônio Matemático



$$x_1W_1, x_2W_2, \dots, x_nW_n$$

$$v = \sum_{i=1}^n w_i x_i$$

Um neurônio dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação chamado de **threshold**





O Neurônio Matemático

Note que este modelo matemático simplificado de um neurônio é estático, ou seja, não considera a dinâmica do neurônio natural. No neurônio biológico, os sinais são enviados em pulsos e alguns componentes dos neurônios biológicos, a exemplo do axônio, funcionam como filtros de frequência.





O Neurônio Matemático

$$u_k = \sum_{j=1}^m w_{kj} \cdot x_j$$

Fórmula do Neurônio
Artificial

$$y_k = \varphi(u_k)$$

Fórmula da Função de
Ativação

$$\varphi(u) = \begin{cases} 1 & , \text{ se } u \geq 0 \\ 0 & , \text{ se } u < 0 \end{cases}$$





O Neurônio Matemático

Dentre as funções de ativação utilizadas, podemos destacar:

$$\varphi(v) = \frac{1}{1 + e^{-v}}$$

Função Sigmóide

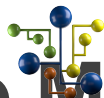
$$\varphi(v) = e^{-v^2}$$

Função Gaussiana

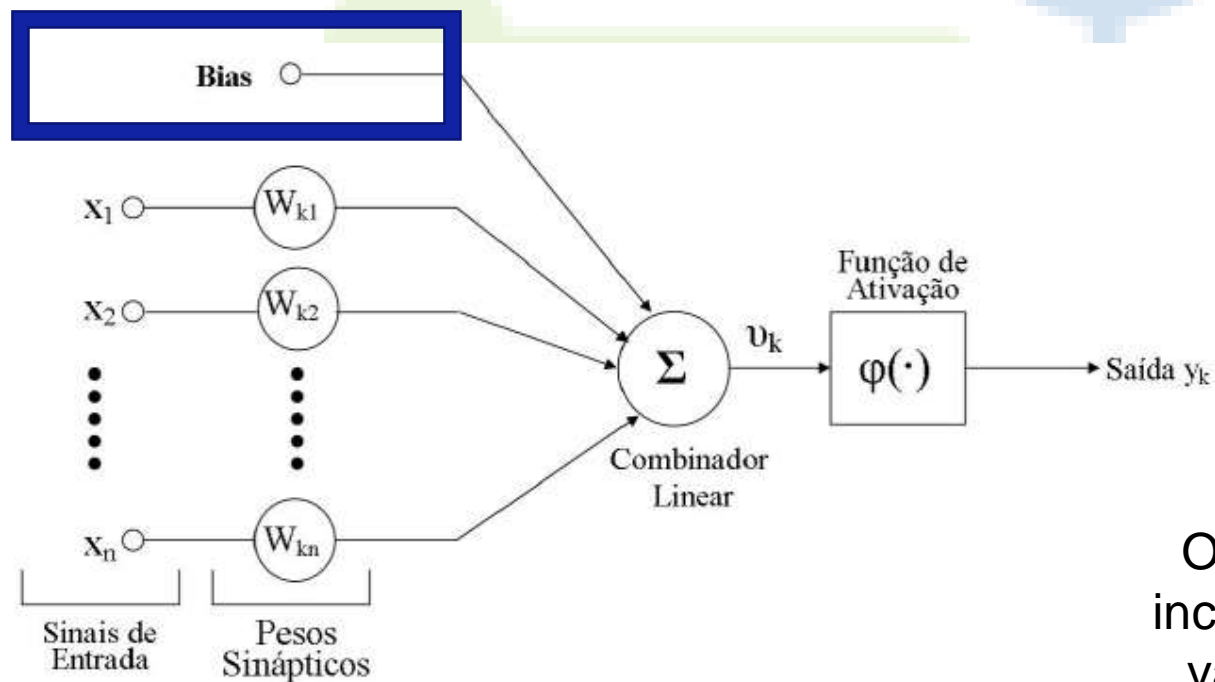
$$\varphi(v) = \tanh(v)$$

Função Tangente
Hiperbólica





O Neurônio Matemático



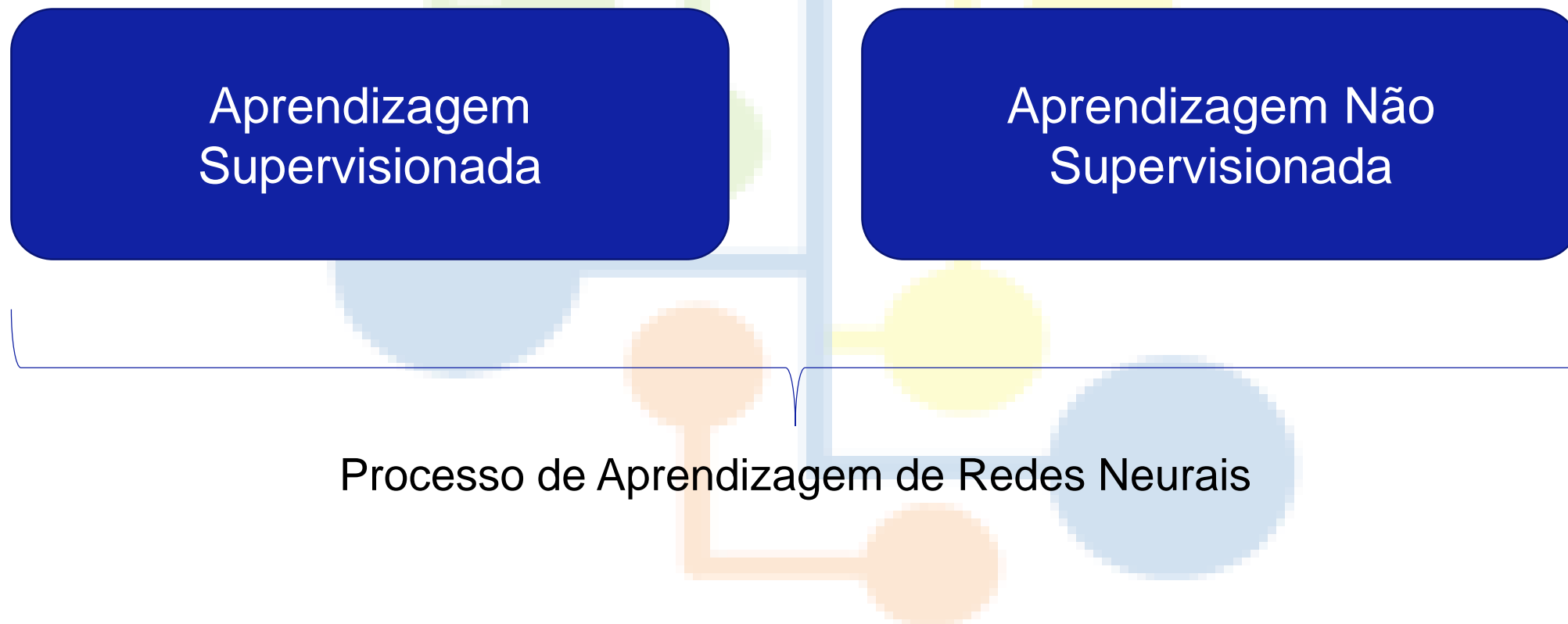
O modelo neuronal matemático também pode incluir uma polarização ou **bias** de entrada. Esta variável é incluída ao somatório da função de ativação, com o intuito de aumentar o grau de liberdade desta função e, conseqüentemente, a capacidade de aproximação da rede



Processo de Aprendizagem

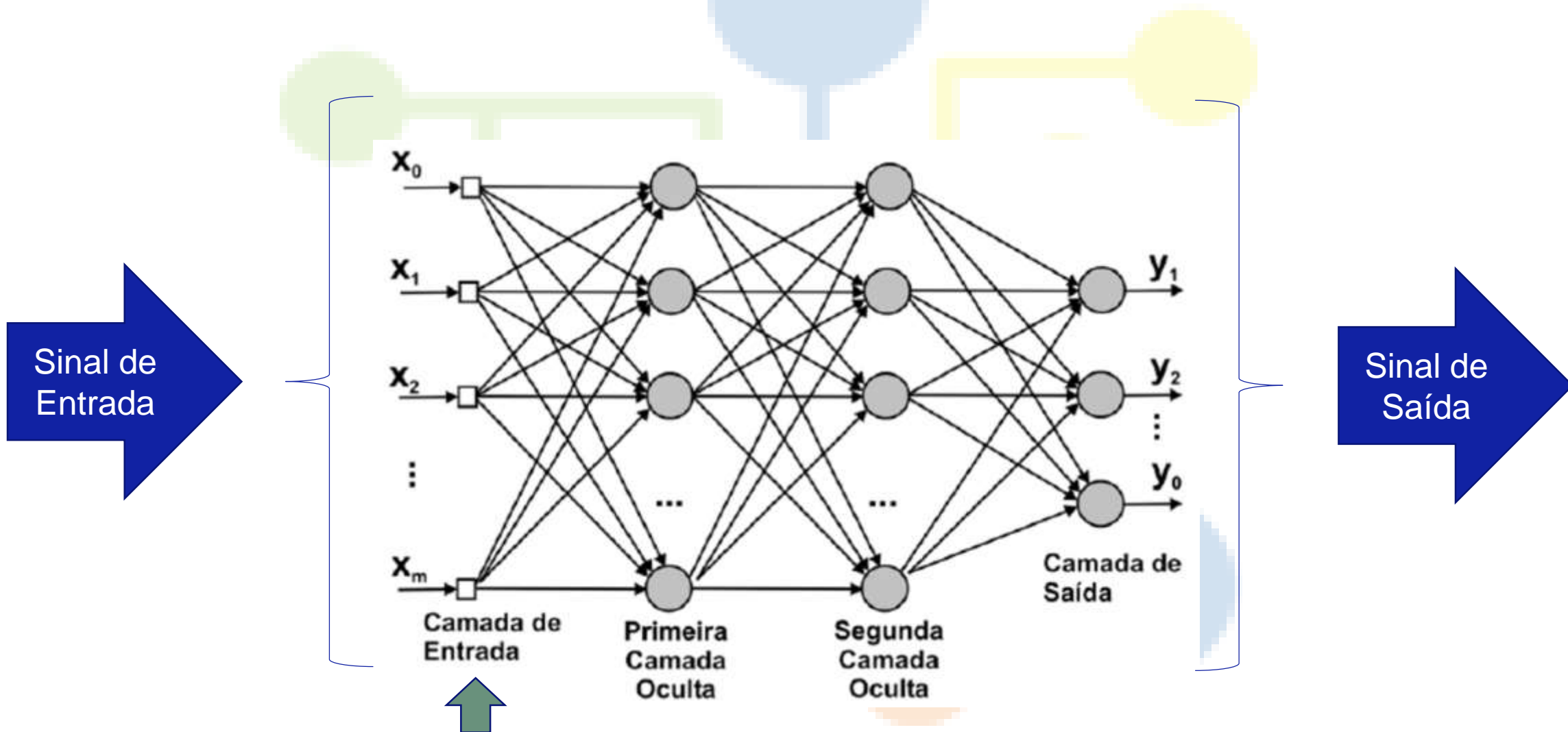


Processo de Aprendizagem





Processo de Aprendizagem





Processo de Aprendizagem

Tipo de Regra de Aprendizagem	Descrição
Aprendizagem por Correção de Erro (Regra Delta)	Utilizado em treinamento supervisionado, esta técnica ajusta os pesos sinápticos por meio do erro, que é obtido através da diferença entre o valor de saída da rede e o valor esperado em um ciclo de treinamento. Com isso gradualmente vai diminuindo o erro geral da rede.
Aprendizagem Hebbiana	Baseado no postulado de aprendizagem de Hebb, que afirma: "se dois neurônios em ambos os lados de uma sinapse são ativados sincronamente e simultaneamente, então a força daquela sinapse é seletivamente aumentada". Este processo de treinamento é feito localmente, ajustando o peso das conexões baseado nas atividades dos neurônios.
Aprendizagem de Boltzmann	Método de aprendizagem estocástico derivado de conceitos da Estatística. Neste modelo os neurônios são estocásticos, podendo residir em dois estados possíveis, ligado (+1) e desligado (-1), e ainda são divididos em dois grupos funcionais, presos e livres, sendo responsáveis pela interação com o ambiente e pela explicação das restrições subjacentes dos padrões de entrada do ambiente, respectivamente. Um ponto importante deste modelo de aprendizagem é que os neurônios possuem conexões bidirecionais.
Aprendizagem Competitiva	Neste modelo de aprendizagem os neurônios são forçados a competir entre si e somente um será ativo, em uma dada iteração, o vencedor, ou seja, o que tiver maior similaridade com o padrão de entrada. Todos os pesos dos neurônios próximos ao neurônio vencedor terão seus valores ajustados.





Processo de Aprendizagem

Aprendizagem Hebbiana

”Quando um axônio da célula A está perto o suficiente para excitar uma célula B e participa do seu disparo repetidamente, então algum processo de crescimento ou modificação metabólica acontece em uma das células ou em ambas, de tal forma que a eficiência de A como uma das células que dispara B é aumentada.”





Processo de Aprendizagem

Aprendizagem Hebbiana

Se dois neurônios em ambos os lados de uma sinapse são ativados **sincronamente**, então a força desta sinapse é aumentada

Se dois neurônios em ambos os lados de uma sinapse são ativados **assincronamente**, então a força desta sinapse é enfraquecida



Deep Neural Networks



Deep Neural Networks

Deep neural network



Deep Learning é na verdade
uma Deep Neural Network
(Rede Neural Profunda)





Deep Neural Networks

Deep neural network



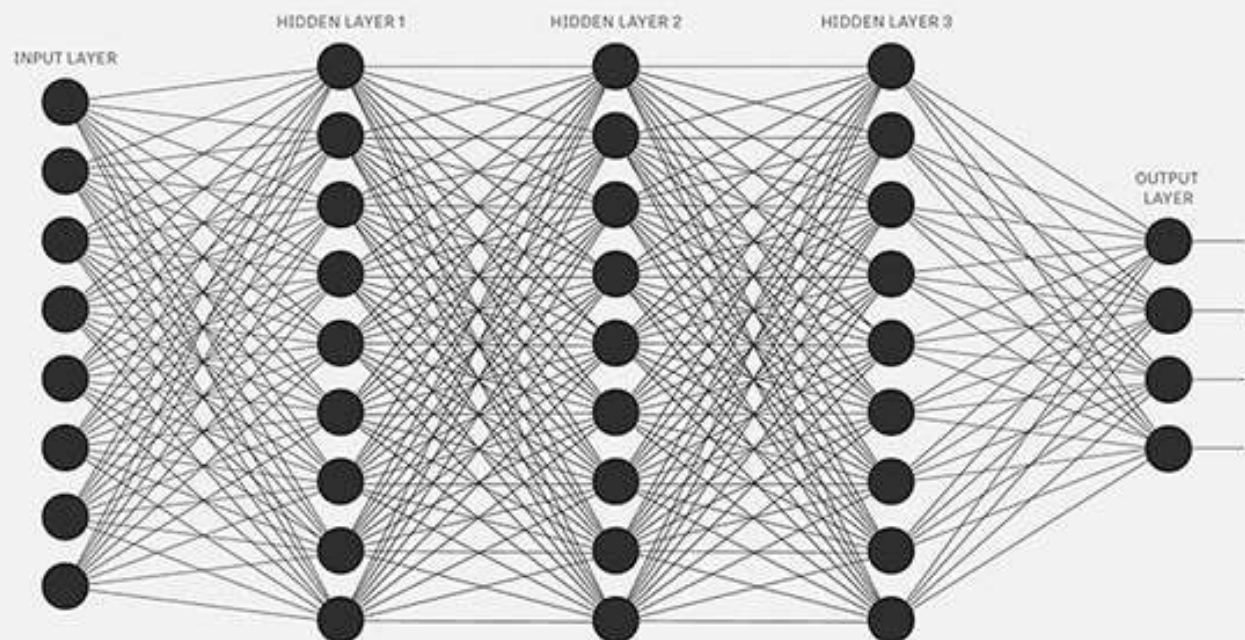
O uso de várias camadas ocultas permite uma acumulação mais sofisticada de elementos simples a outros mais complexos





Deep Neural Networks

Deep neural network



Pode-se considerar dois aspectos de complexidade da arquitetura de um modelo:

- Número de neurônios por camada
- Número de camadas





Deep Neural Networks



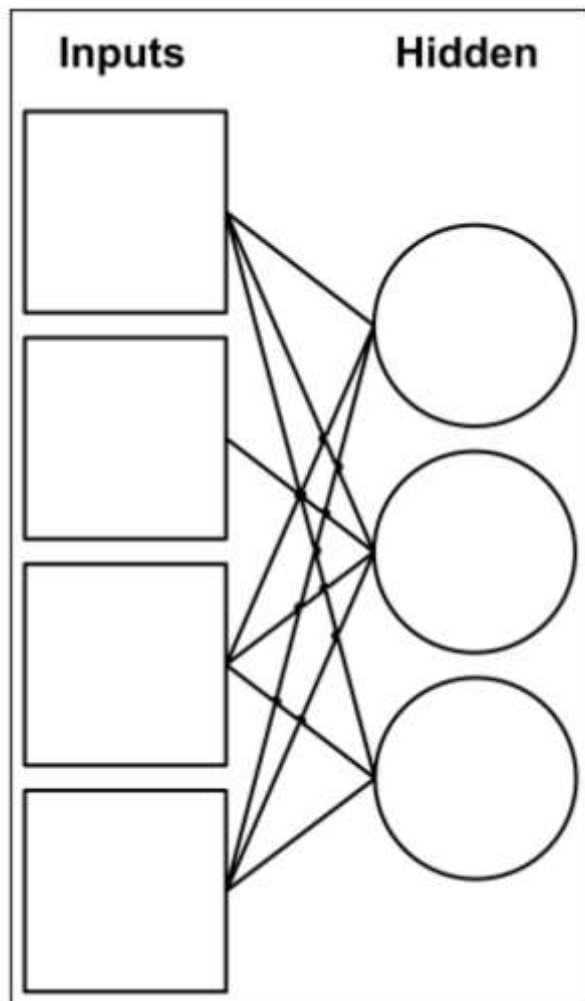
A primeira camada de neurônios poderia ser treinada para capturar diferentes letras do alfabeto e então outra camada poderia reconhecer conjuntos dessas letras como palavras

A vantagem é que a segunda camada não precisa aprender diretamente com os pixels, que são mais complexos.

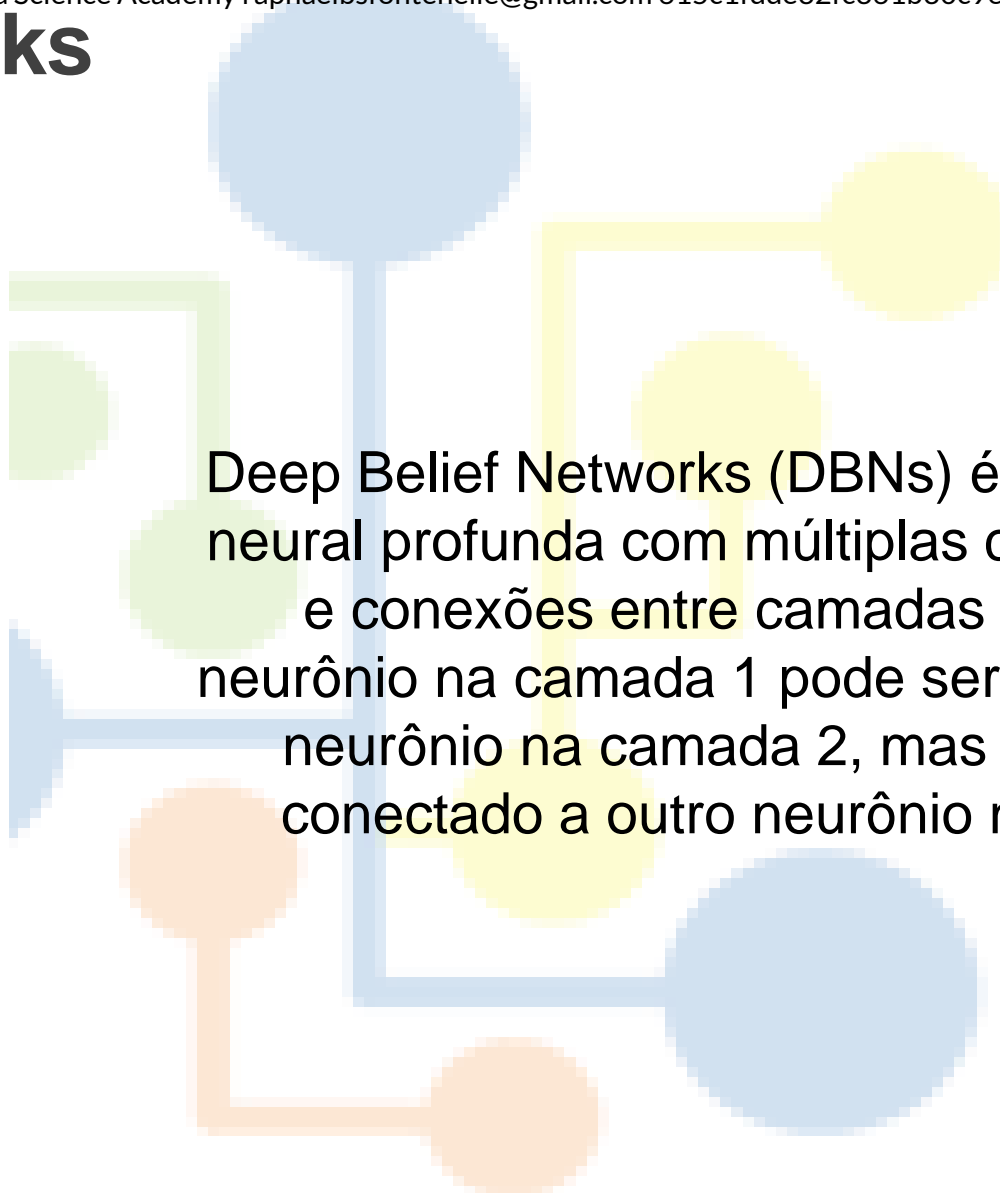




Deep Neural Networks

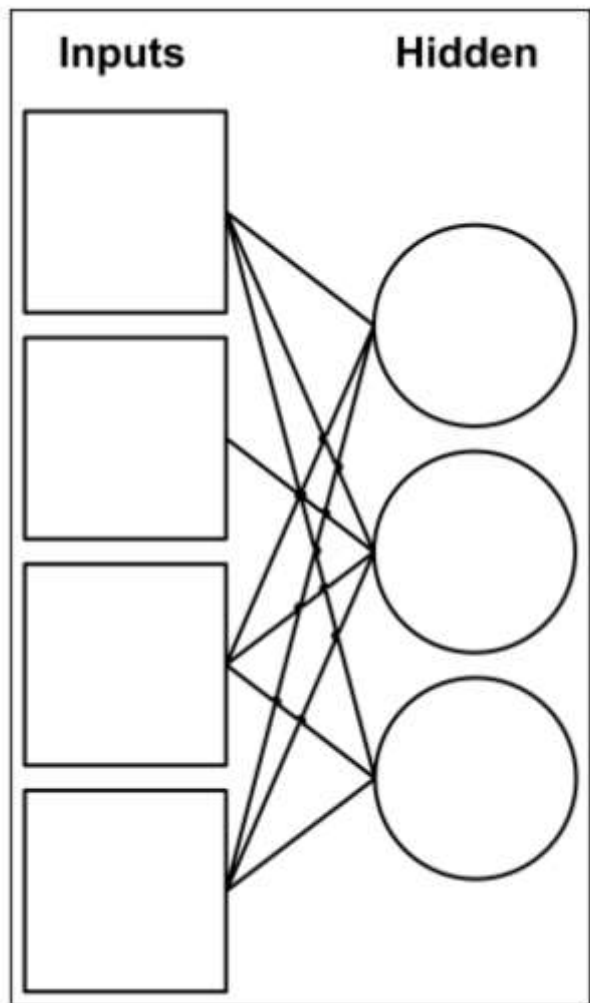


Deep Belief Networks (DBNs) é um tipo de rede neural profunda com múltiplas camadas ocultas e conexões entre camadas (ou seja, um neurônio na camada 1 pode ser conectado a um neurônio na camada 2, mas não pode ser conectado a outro neurônio na camada 1)

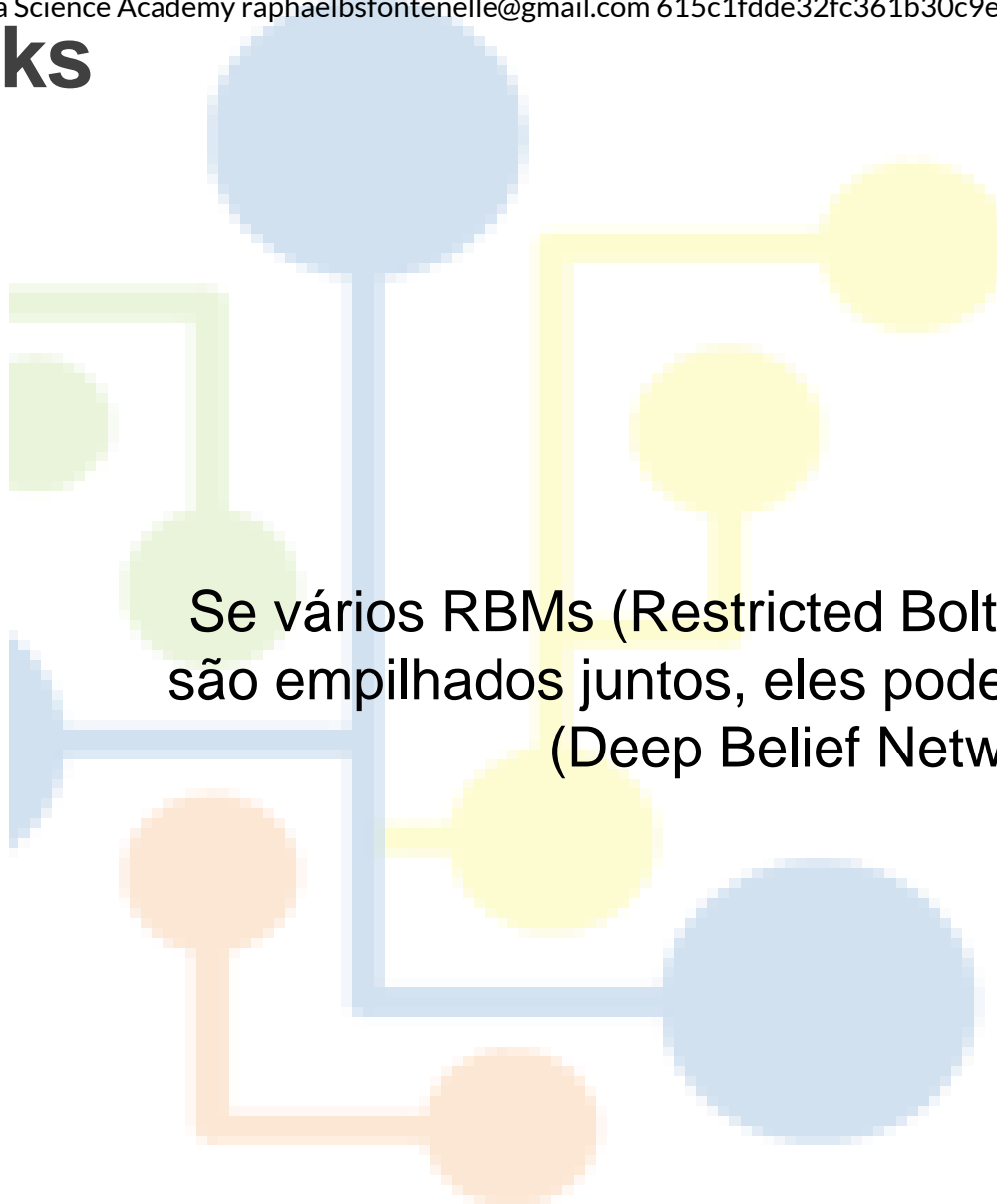




Deep Neural Networks

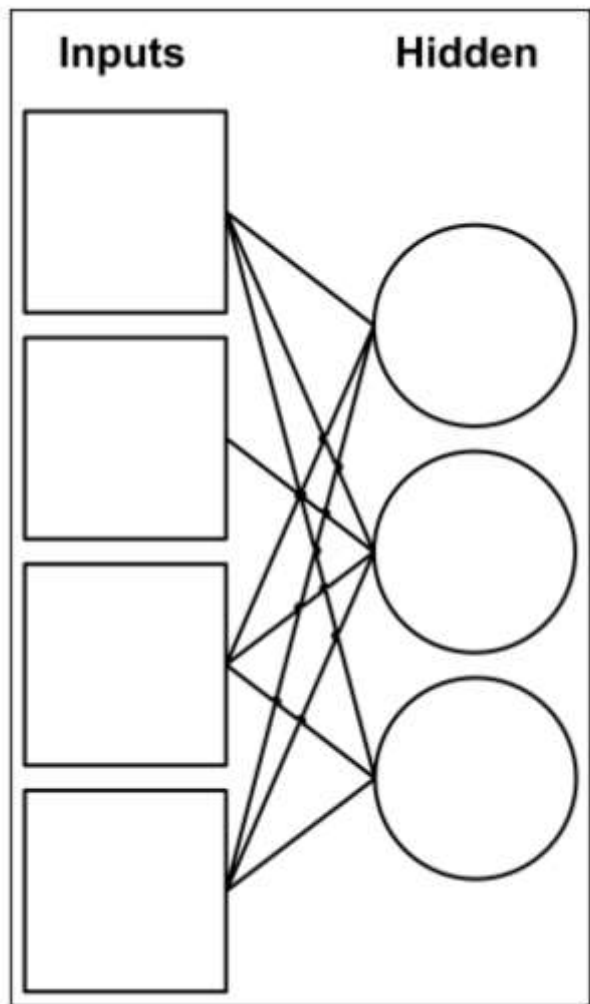


Se vários RBMs (Restricted Boltzmann Machines) são empilhados juntos, eles podem formar um DBN (Deep Belief Network)





Deep Neural Networks



DBNs às vezes são usados como um estágio de pré-treinamento para uma rede neural profunda



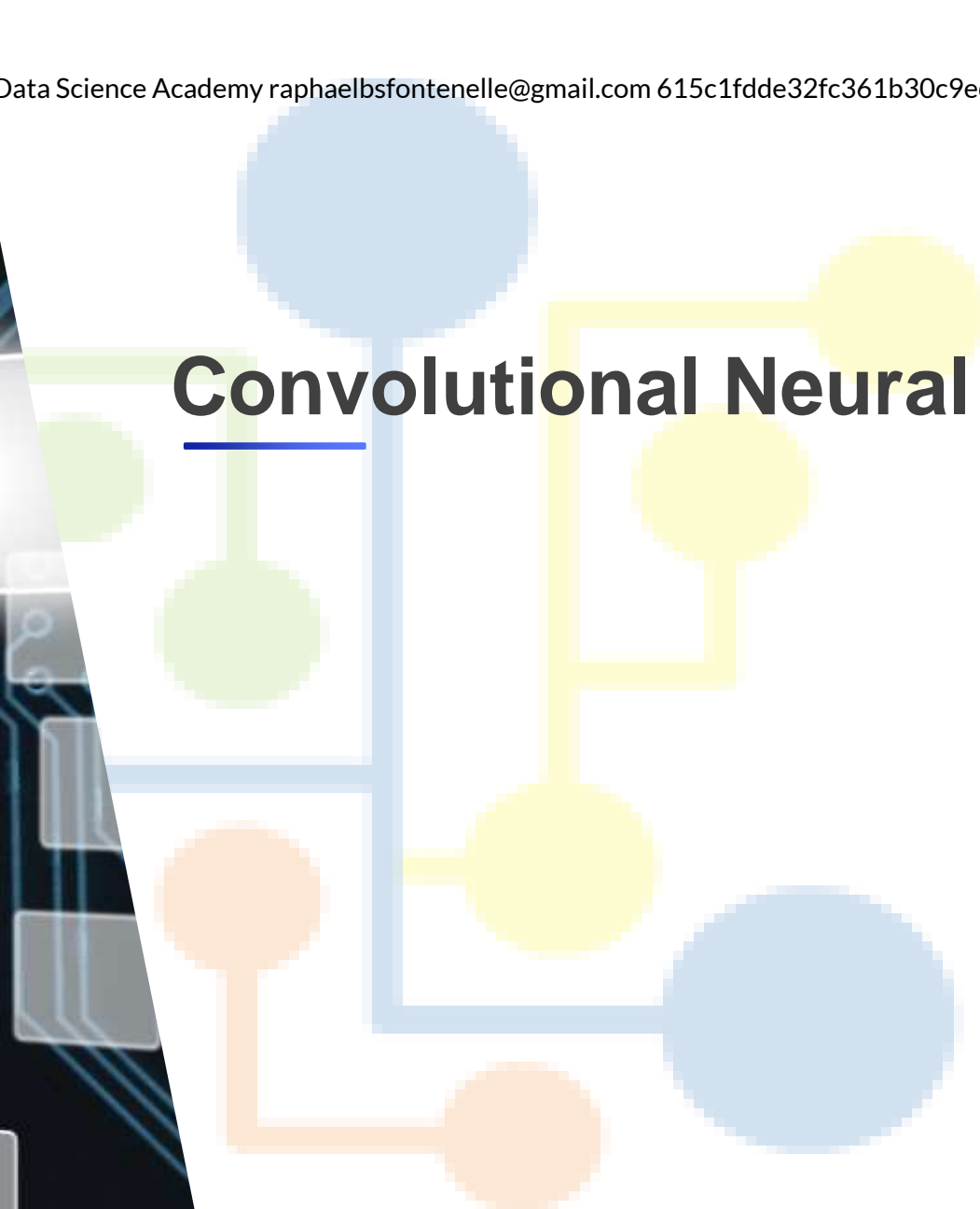
Deep Neural Networks

Mas existem 2 tipos de redes neurais profundas que tem aumentado em popularidade na utilização de Deep Learning:

Convolutional Neural Networks (CNN's)
Recurrent Neural Networks (RNN's)

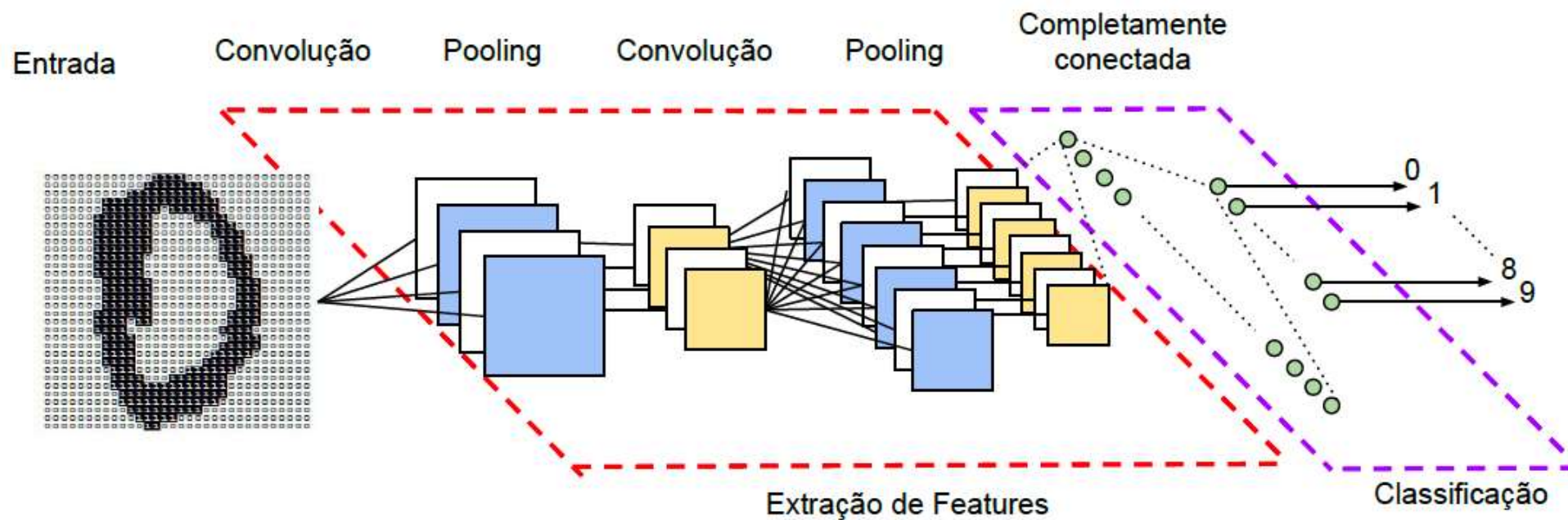


Convolutional Neural Networks



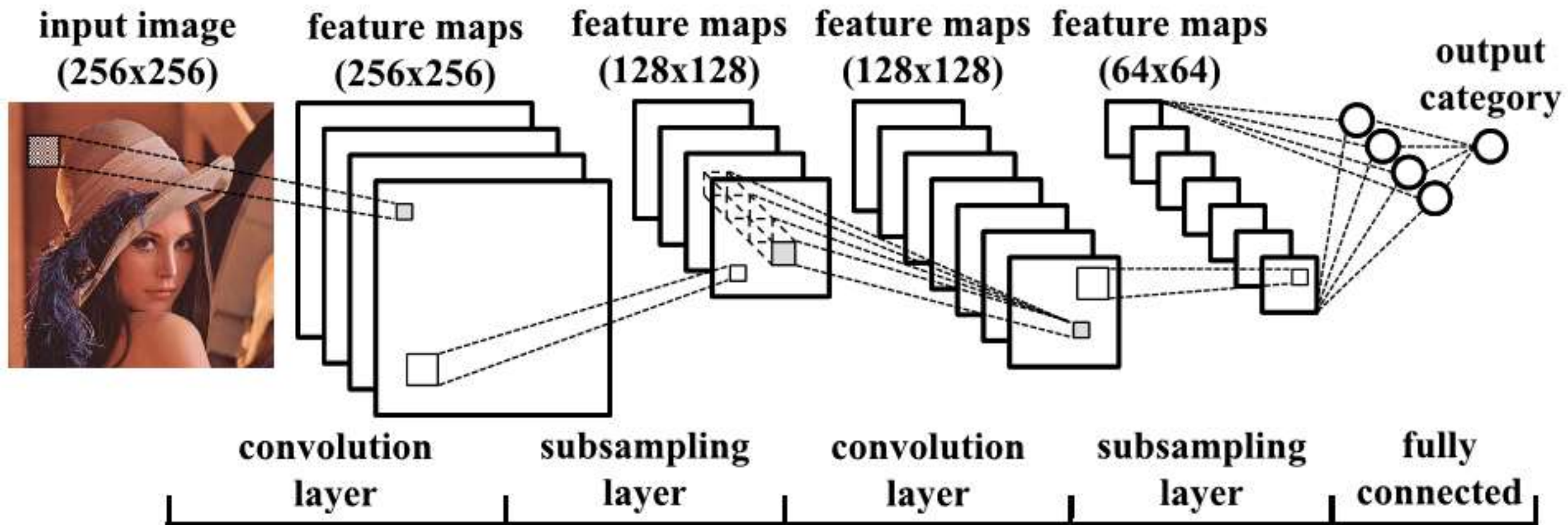


Convolutional Neural Networks



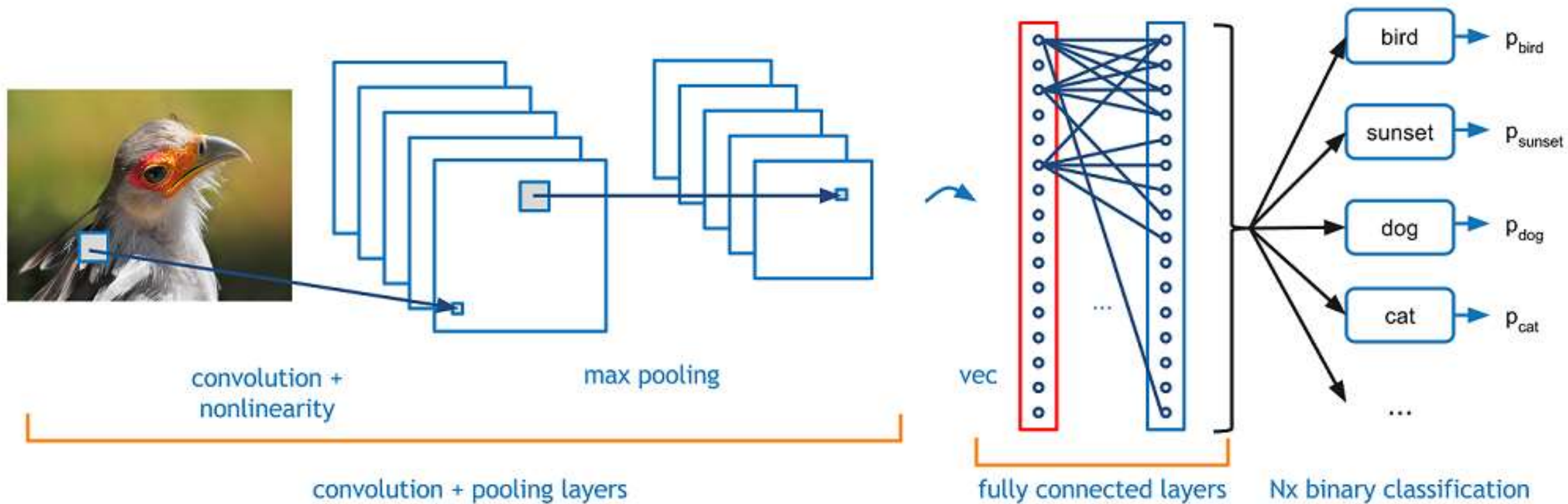


Convolutional Neural Networks



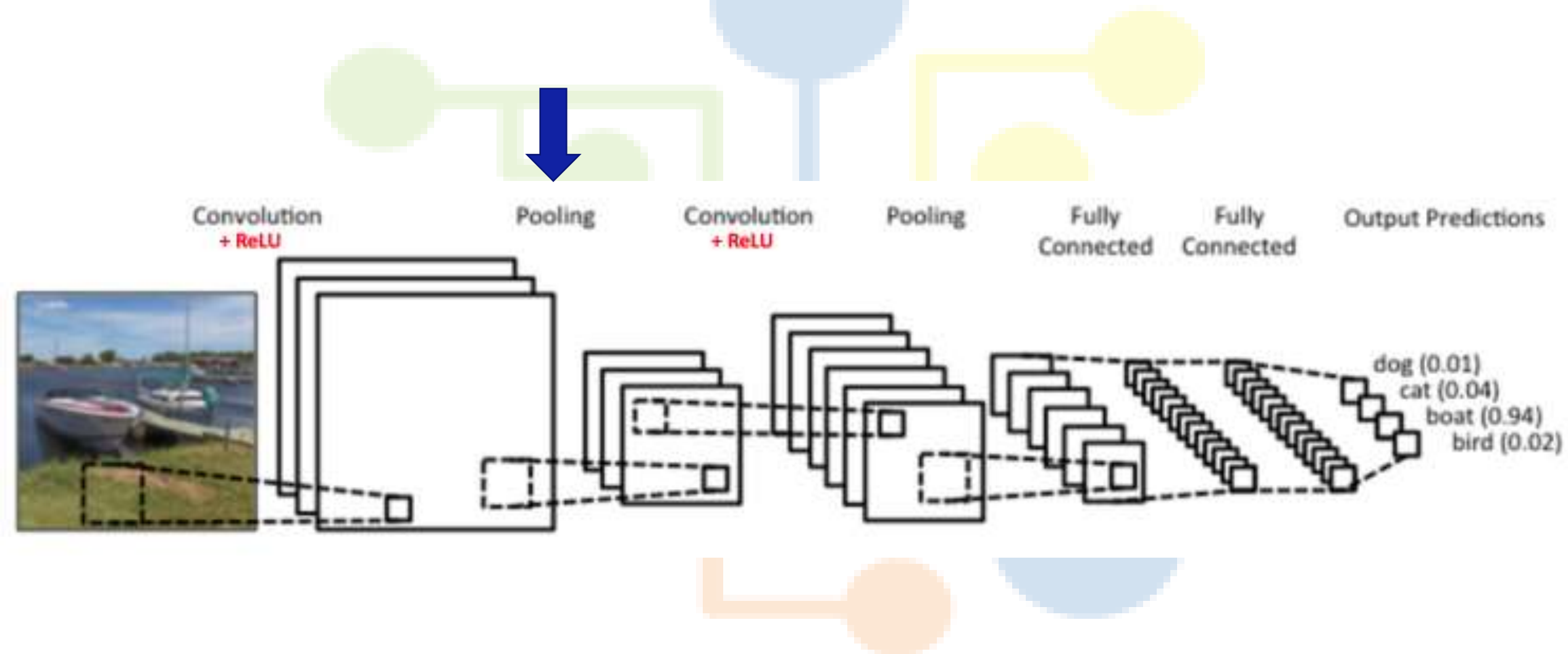


Convolutional Neural Networks



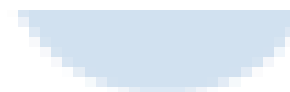
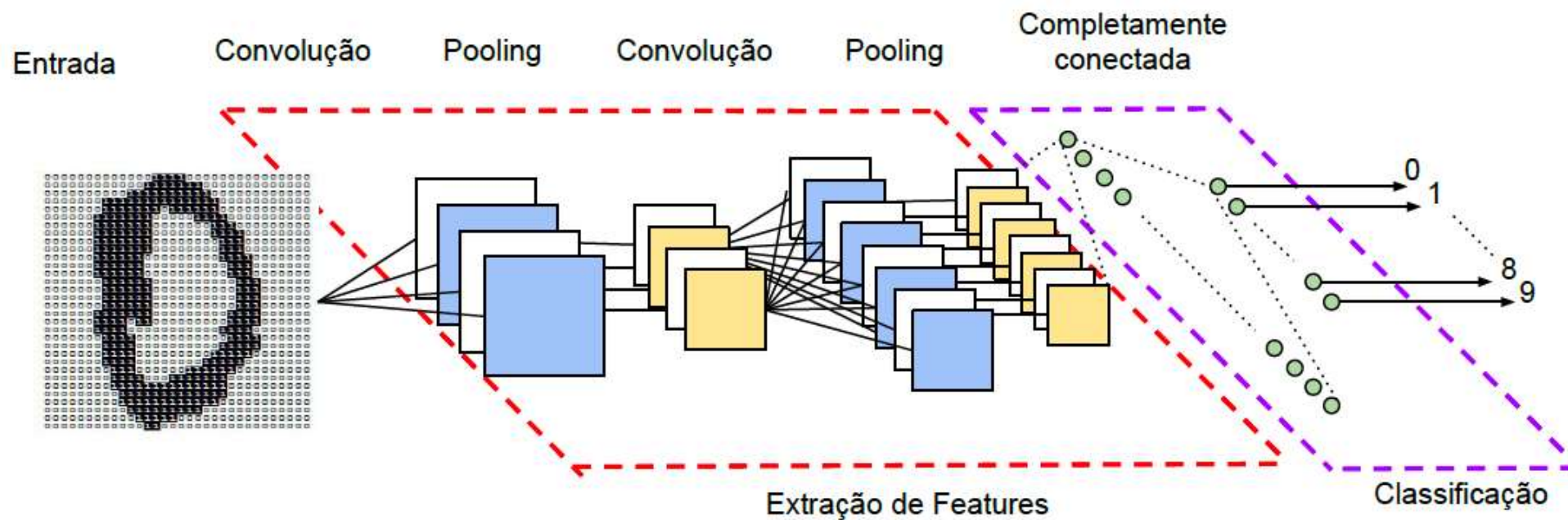


Convolutional Neural Networks





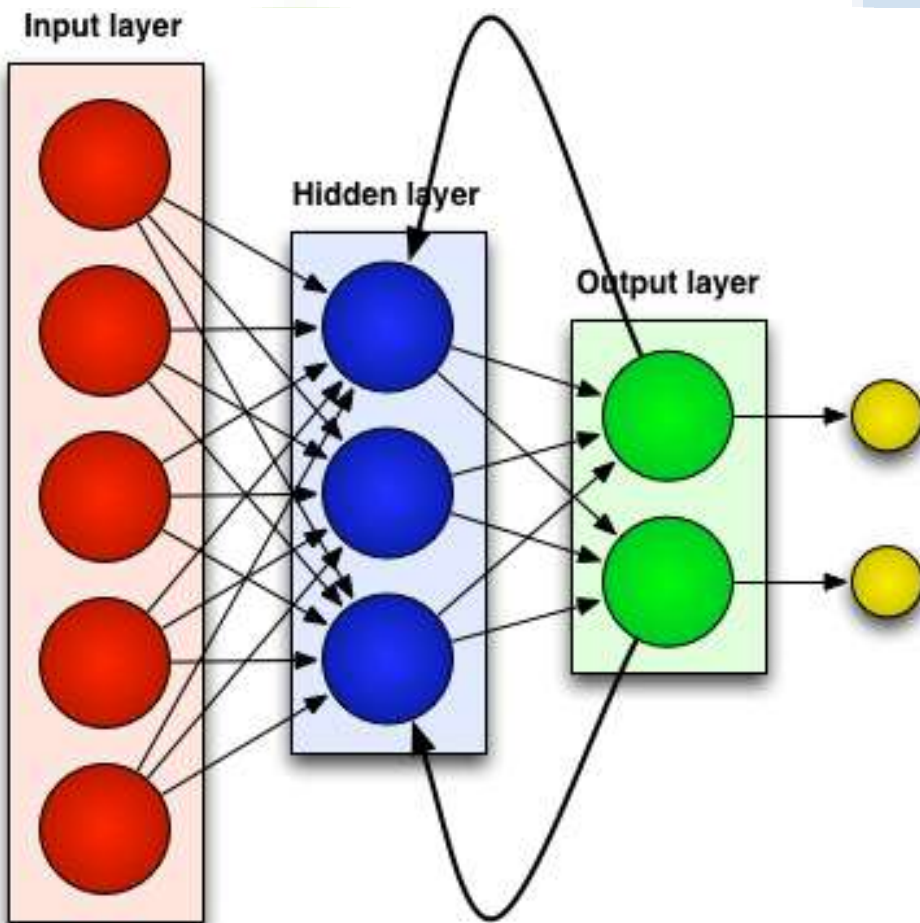
Convolutional Neural Networks



Recurrent Neural Networks



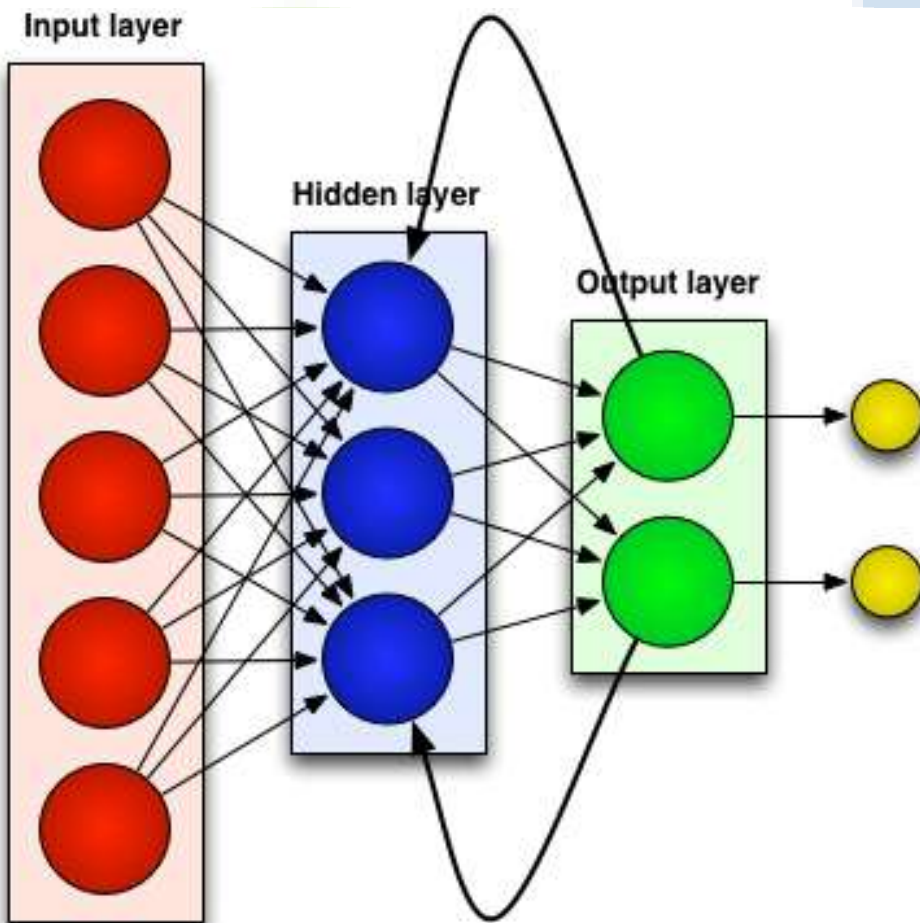
Recurrent Neural Networks



A ideia por trás das RNNs é fazer uso de informações sequenciais



Recurrent Neural Networks

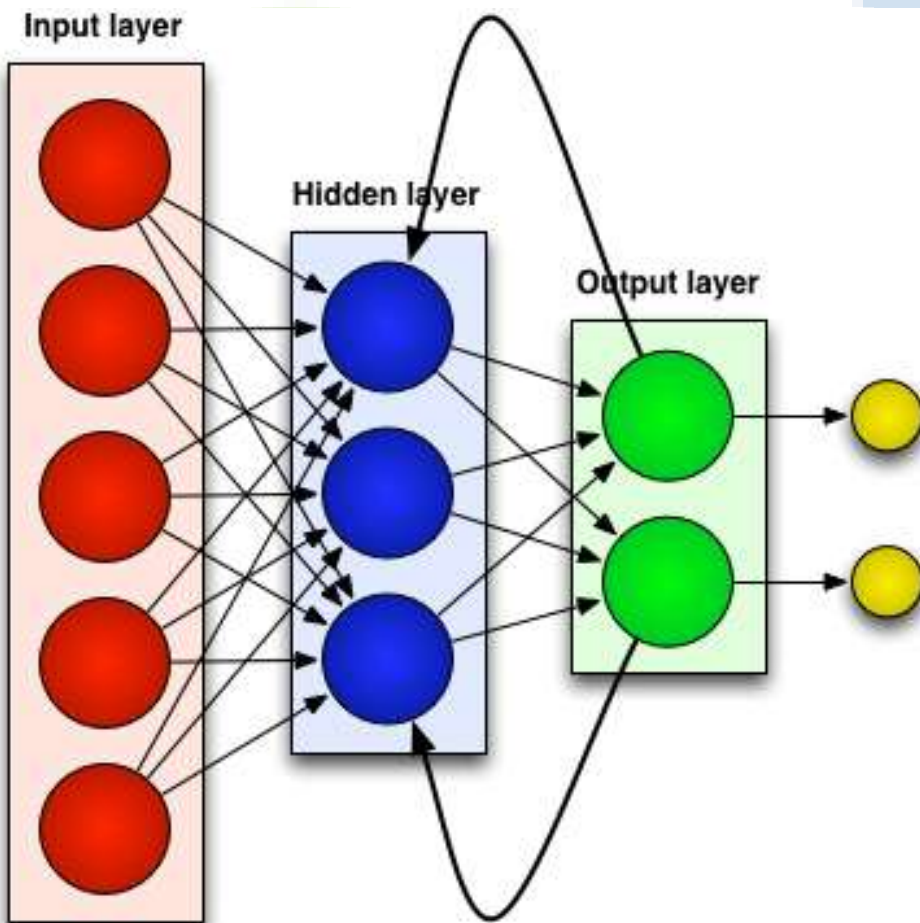


Redes Neurais Recorrentes (Recurrent Neural Networks, RNN's) constituem uma ampla classe de redes cuja evolução do estado depende tanto da entrada corrente quanto do estado atual





Recurrent Neural Networks

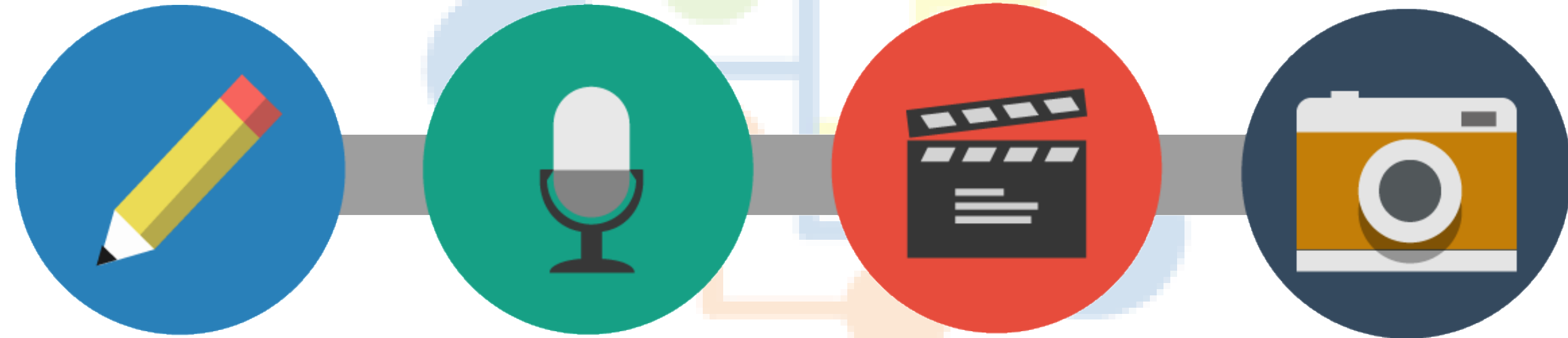


O tipo mais comumente usado de RNNs são as LSTMs (Long Short-Term Memory), que são muito melhores na captura de dependências de longo prazo do que RNNs padrões



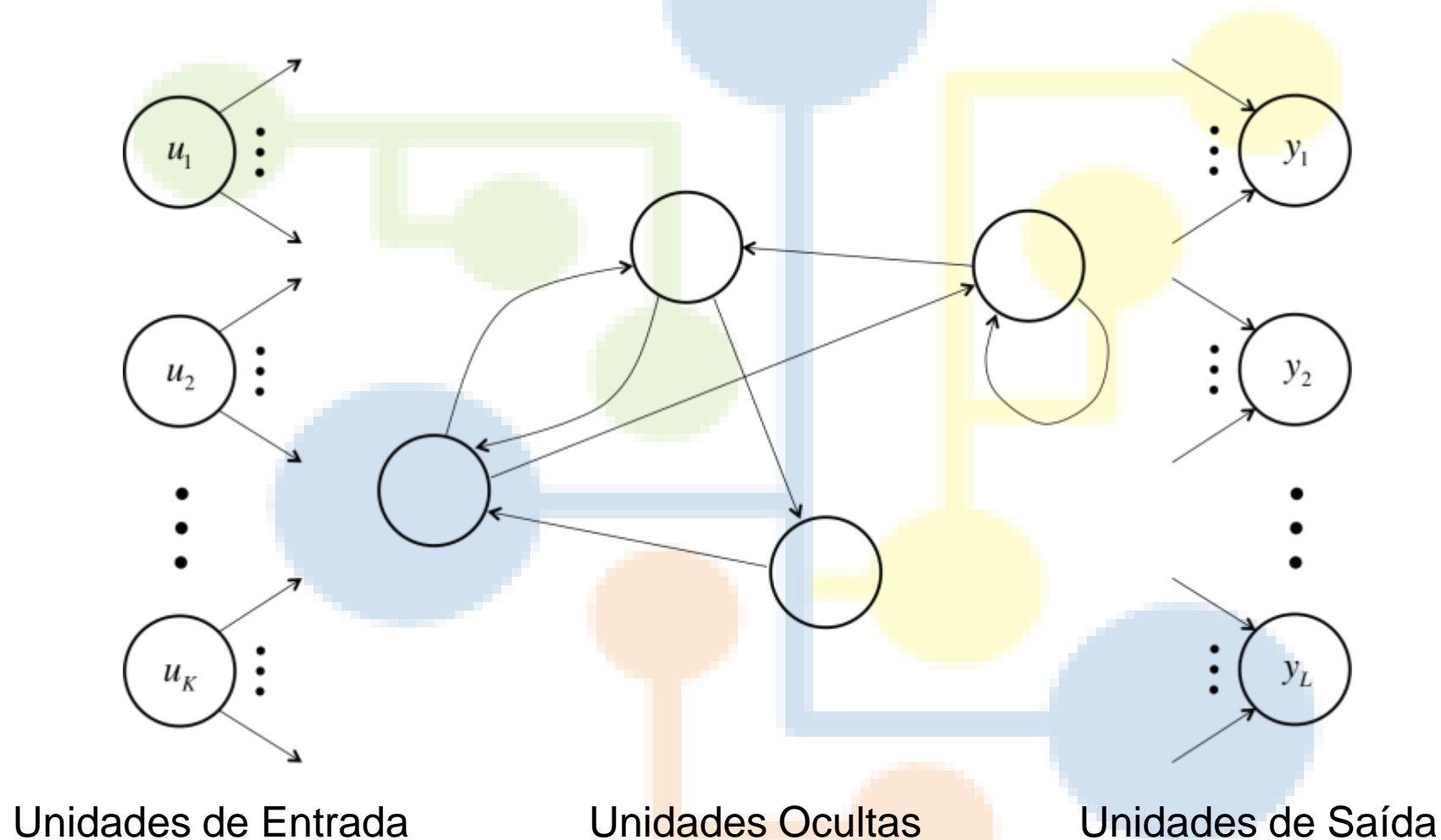
Recurrent Neural Networks

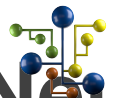
A arquitetura usada em RNNs é adequada para permitir o processamento de informação sequencial (textos, áudio e vídeo)



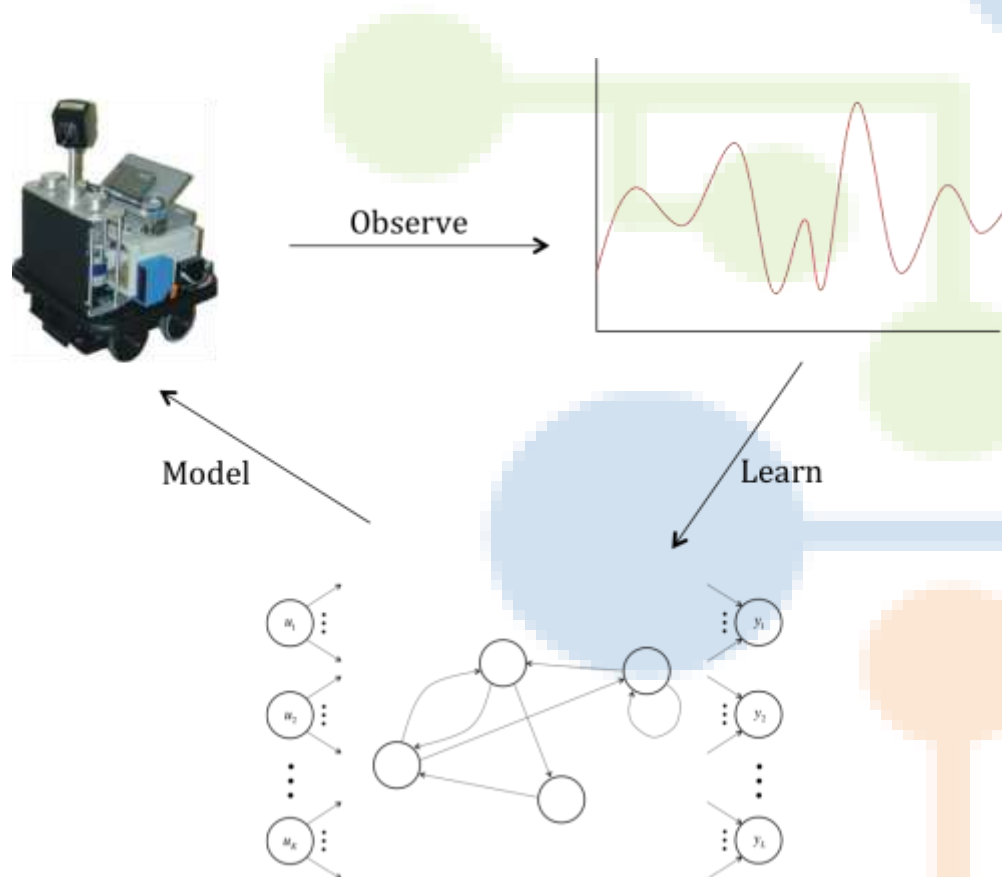


Recurrent Neural Networks





Recurrent Neural Networks

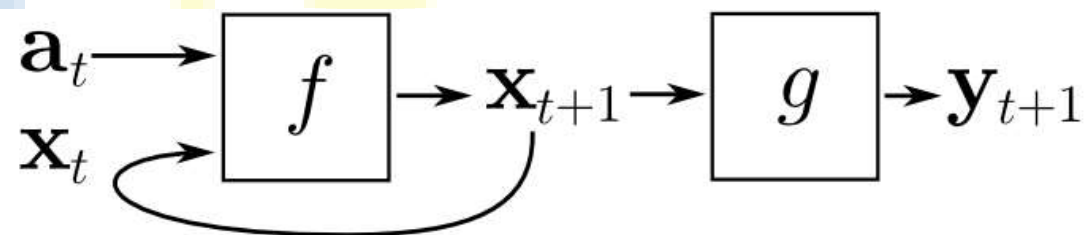


Basicamente, o objetivo em usar RNN's é examinamos os sistemas reais e seus comportamentos ao longo do tempo em resposta aos estímulos

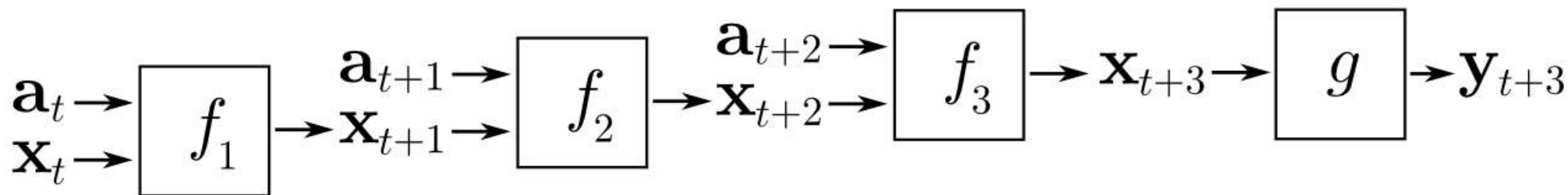


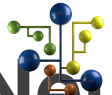
Recurrent Neural Networks

Retropropagação através do tempo
(Backpropagation Through Time, BPTT)



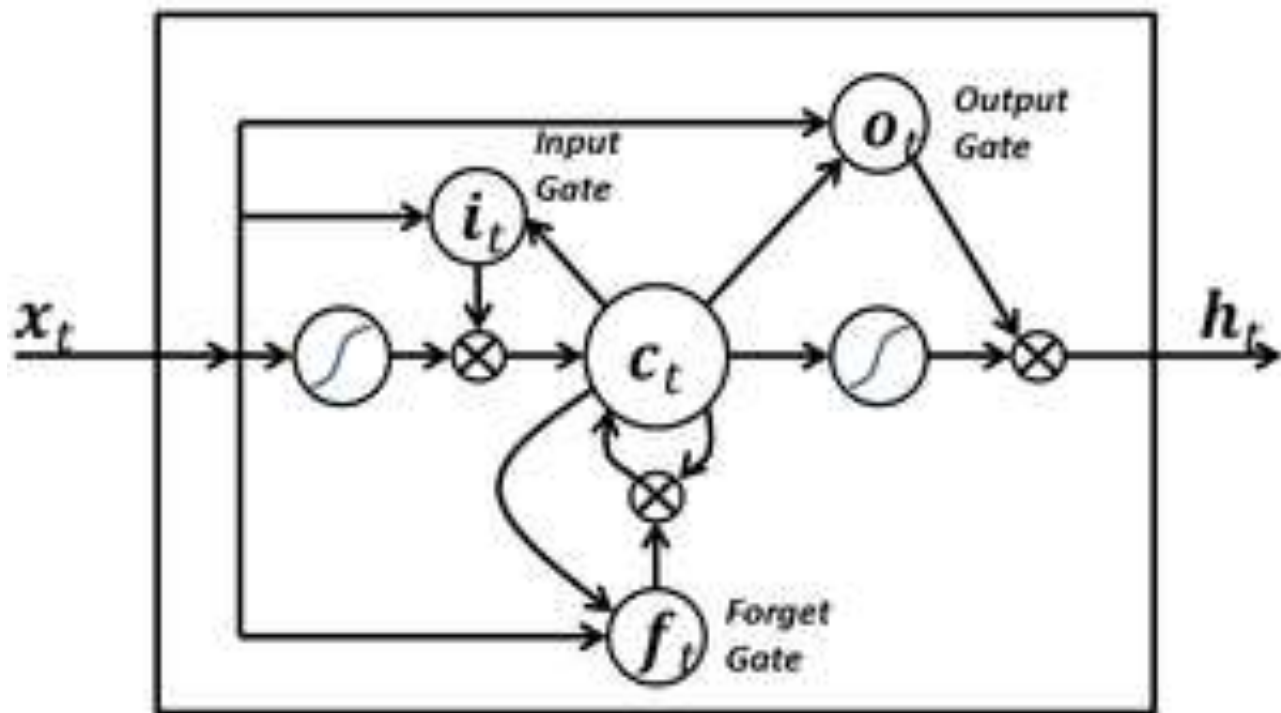
↓ unfold through time ↓





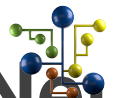
Recurrent Neural Networks

LSTM (Long Short-Term Memory)



GRU (Gated Recurrent Unit)





Recurrent Neural Networks

a b c d e f g h i j k
l m n o p q r s t
u v w x y z . ' ? " " !
A B C D E F G H I
J K L M N O P Q R
S T U V W X Y Z
1 2 3 4 5 6 7 8 9 10

for what is thought
to be best in any relation
point of view.
Translation
another language; in
language foreign to the
written communication
is thought



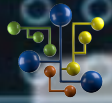


Recurrent Neural Networks

Percebeu porque precisamos de 3 cursos inteiros para falar apenas de Deep Learning??

Deep Learning Frameworks
Deep Learning I
Deep Learning II





Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2

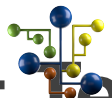
LSTM - Long-Short Term Memory



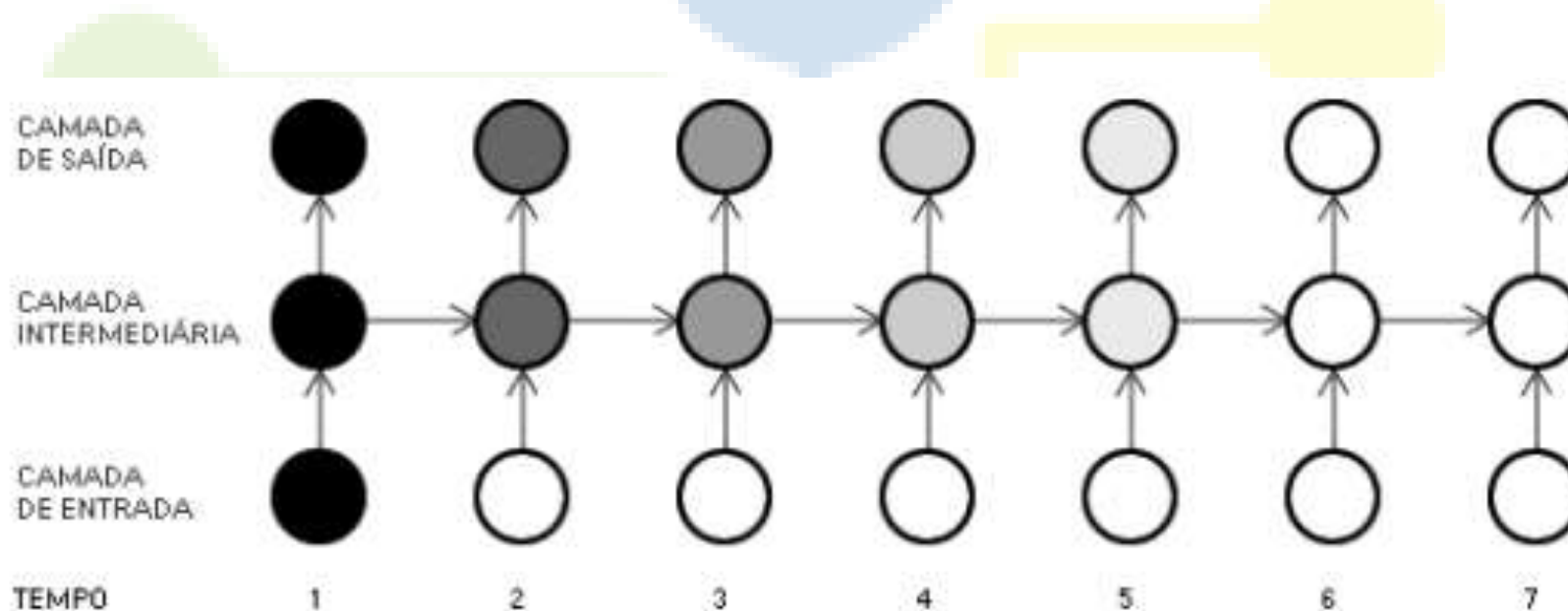
Data Science Academy



Data Science Academy



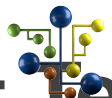
Long-Short Term Memory



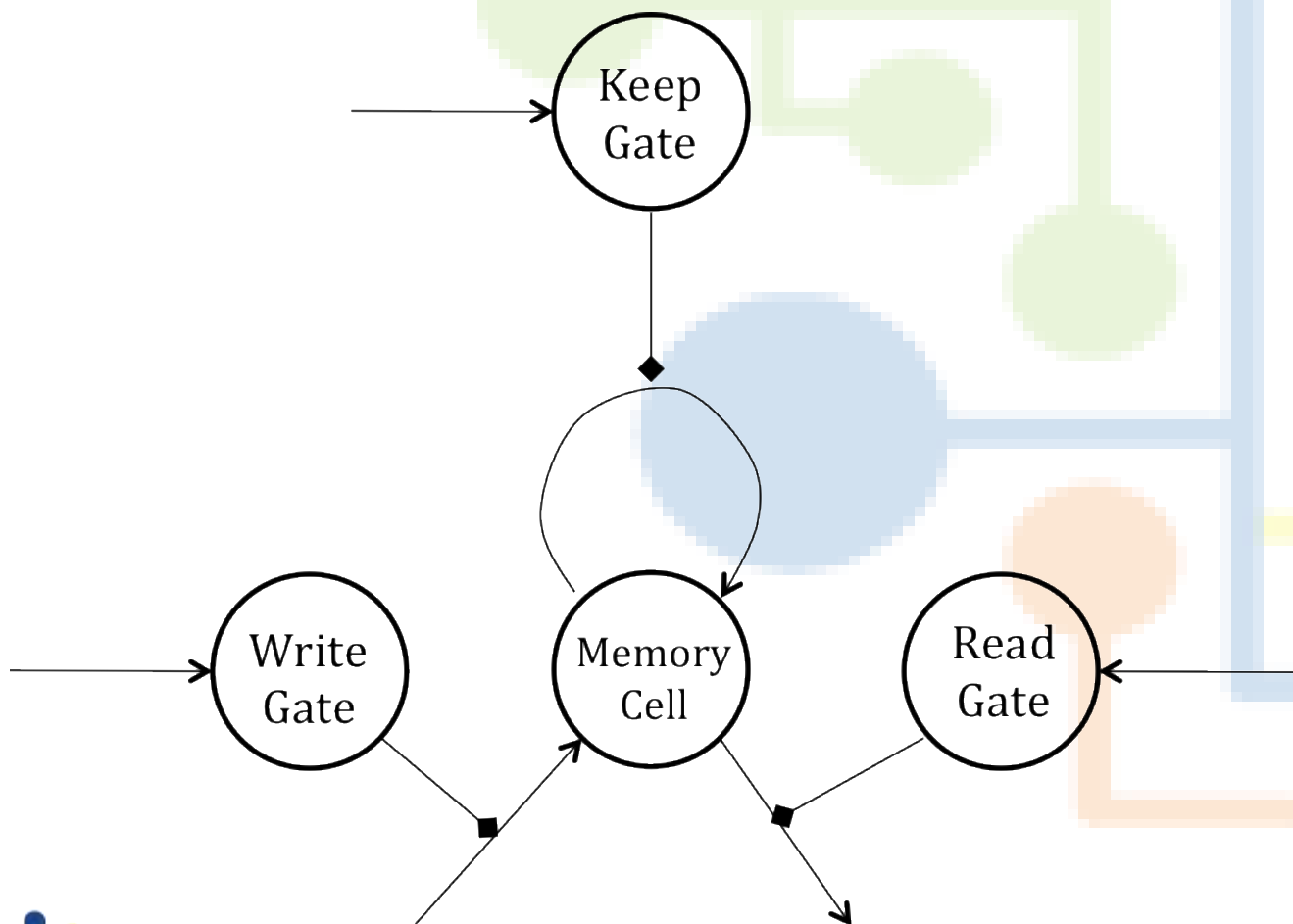
RNN's e o problema da dissipação do gradiente (vanishing gradient problem)

Problema da dissipação do gradiente. Os nós representam neurônios e seu conteúdo representa a sensibilidade para a entrada, quanto mais escuro o nó, mais sensível à entrada preta.

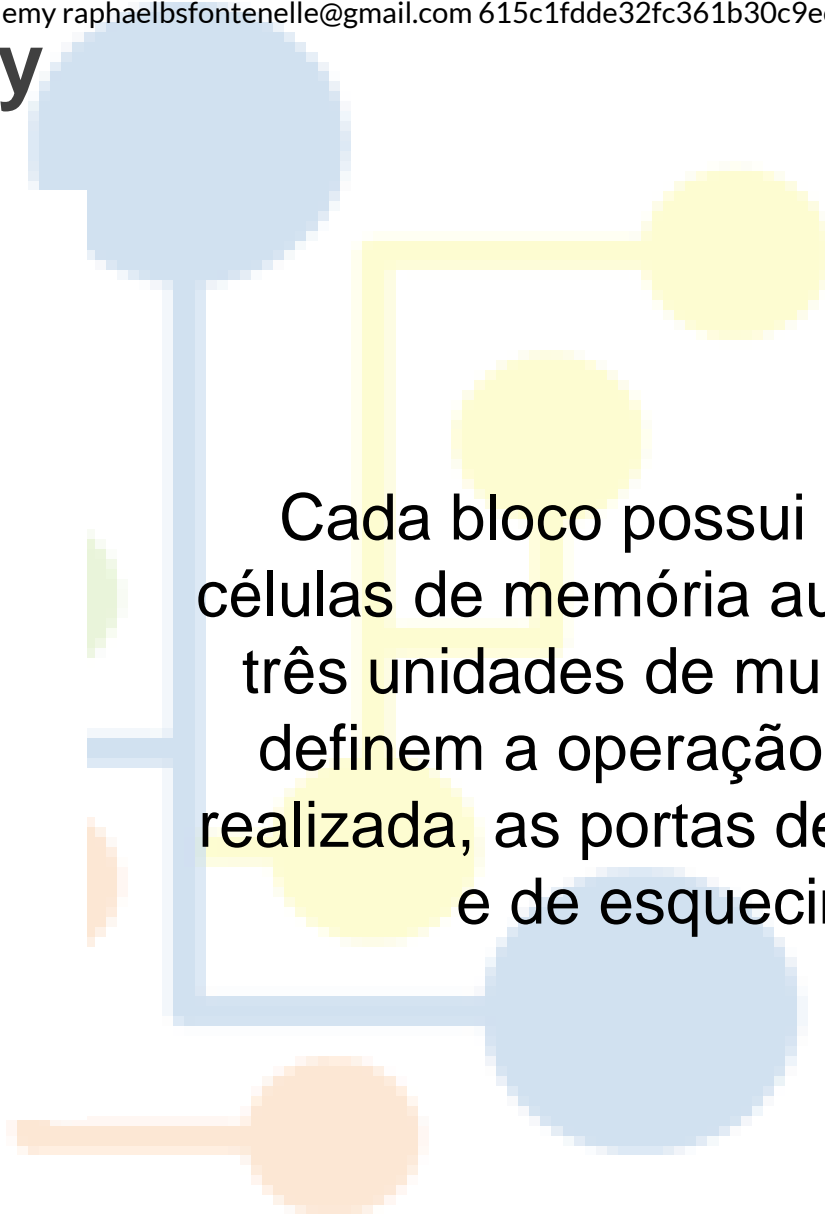




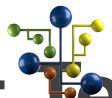
Long-Short Term Memory



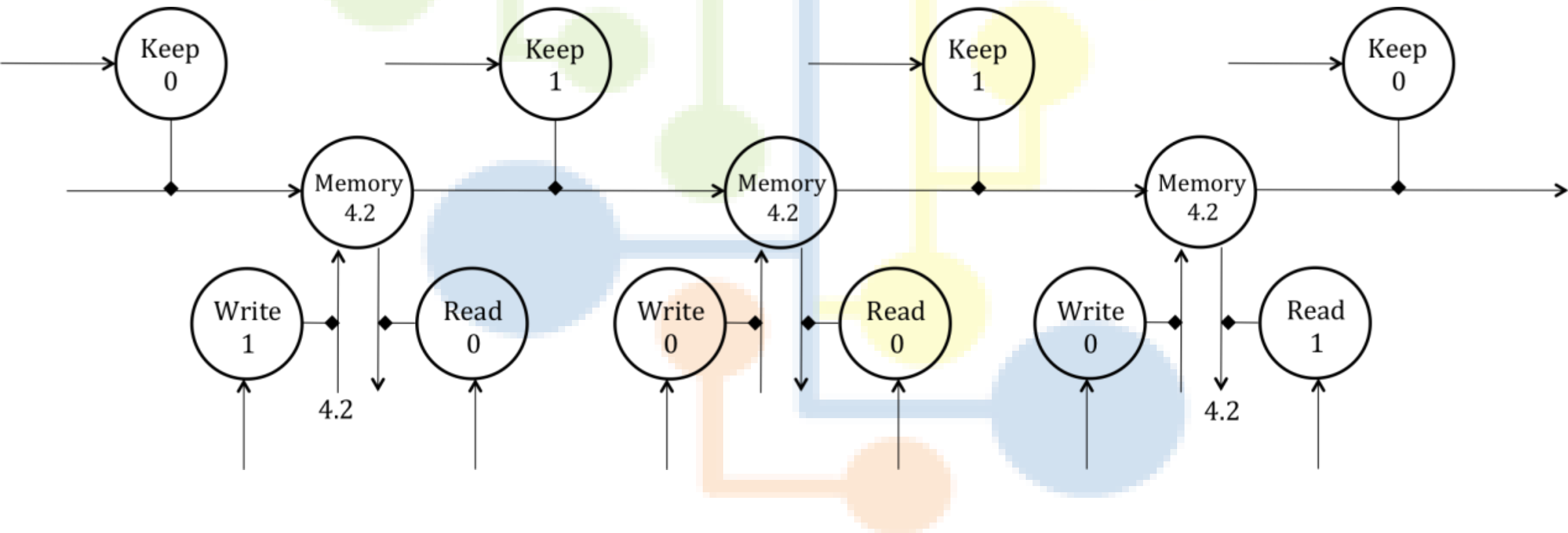
Dentre várias soluções propostas para reduzir os impactos da dissipação do gradiente em RNNs, a solução mais efetiva é a arquitetura Long Short-Term Memory (LSTM)

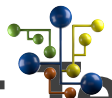


Cada bloco possui uma ou mais células de memória autoconectadas e três unidades de multiplicação que definem a operação que deve ser realizada, as portas de entrada, saída e de esquecimento.

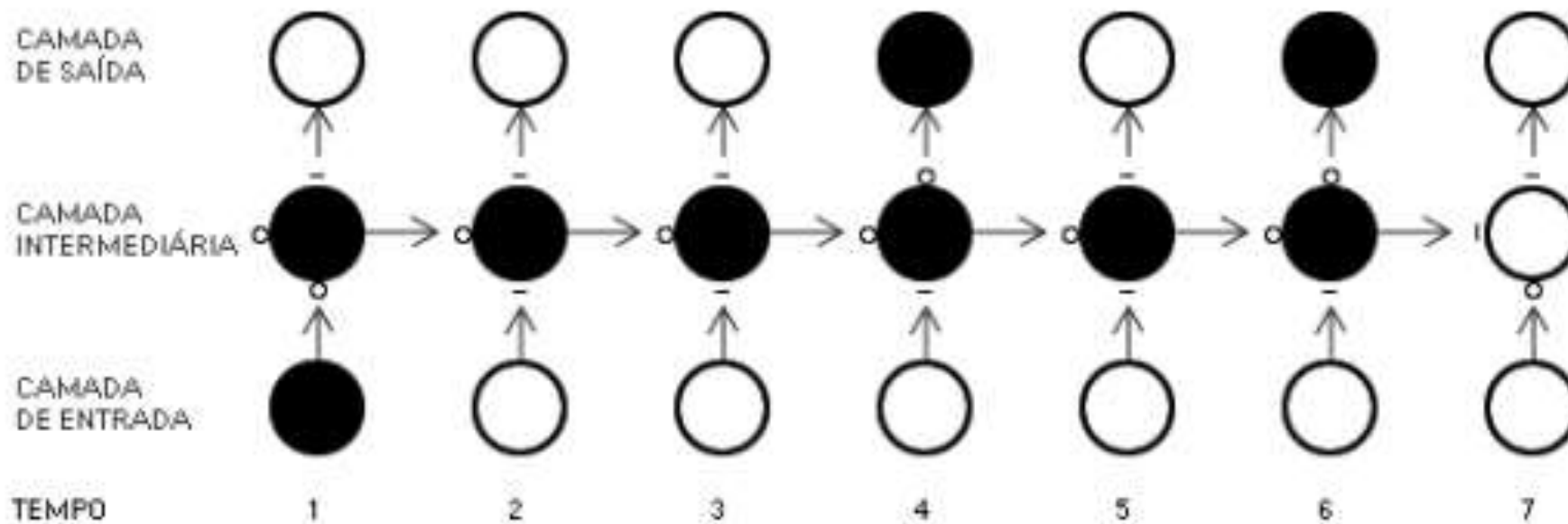


Long-Short Term Memory





Long-Short Term Memory



Função Softmax



Tarefas de regressão, como prever o preço de uma casa



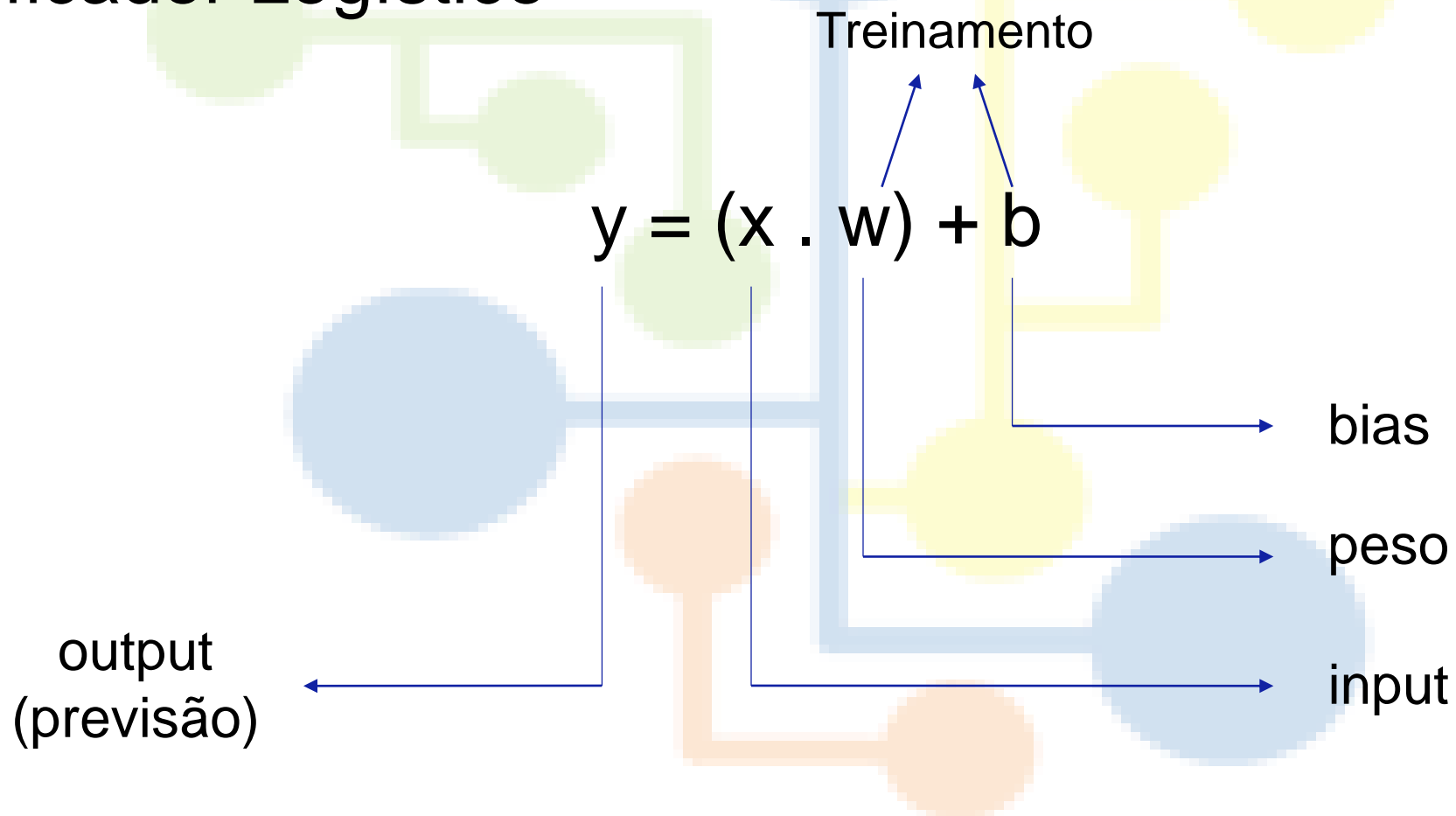
Tarefas de classificação, como prever viabilidade de crédito (sim/não)





Função Softmax

Classificador Logístico





Função Softmax

Classificador Logístico

$$y = (x \cdot w) + b$$

good

2.0

$p = 0.7$

Good

1.0

$p = 0.2$

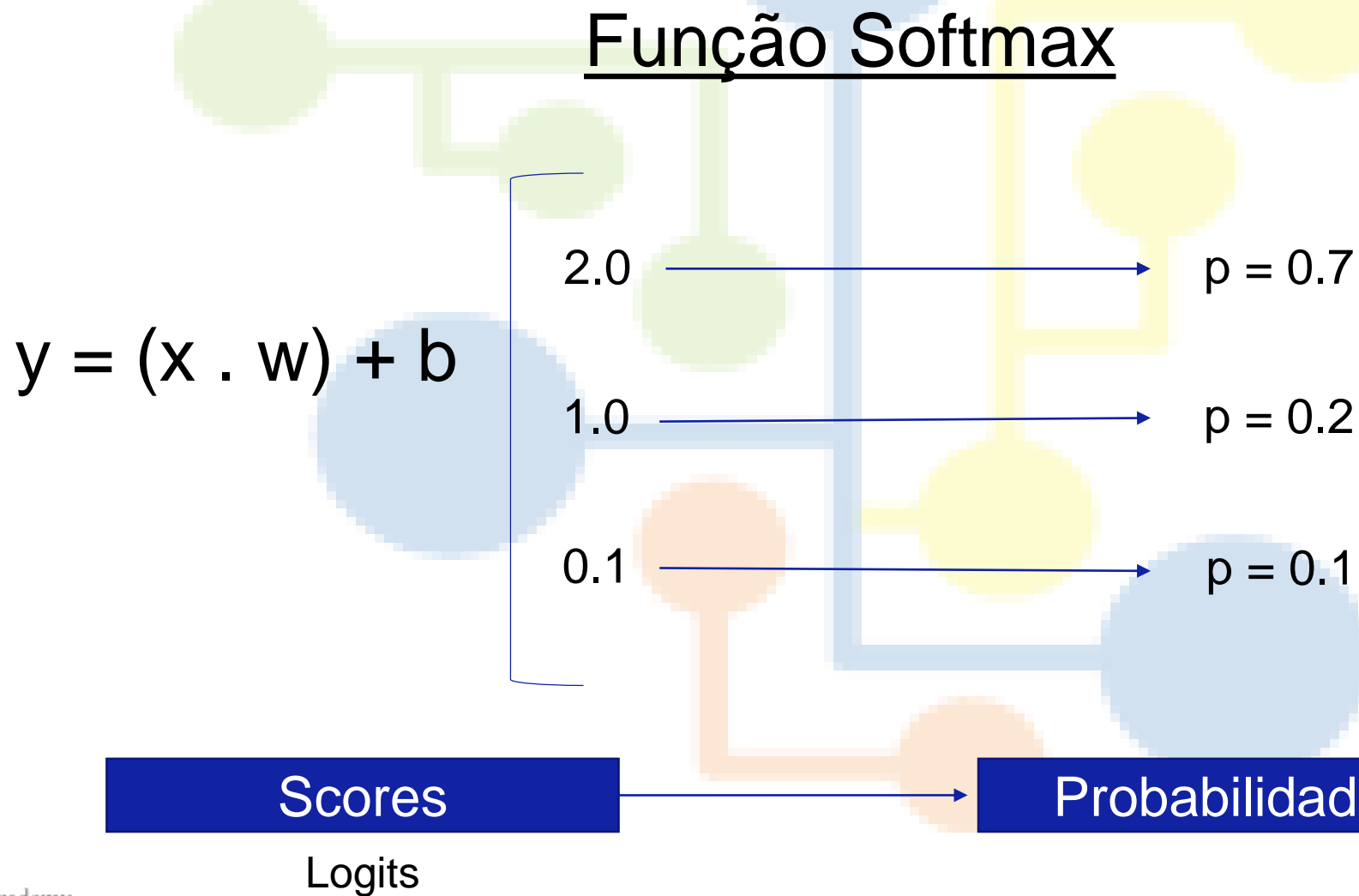
Medium

0.1

$p = 0.1$

Bad







Função Softmax

Função Softmax

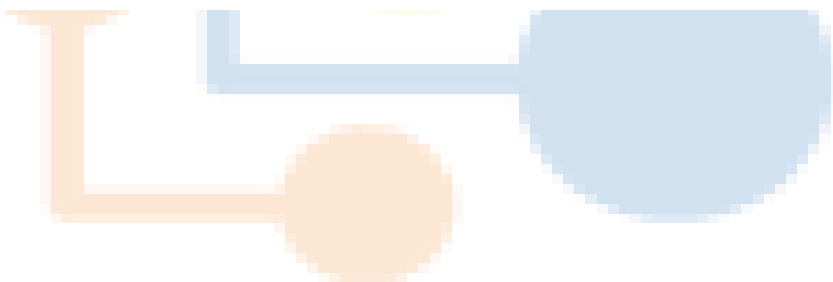
```
import numpy as np

scores = [3.0, 1.0, 0.2]

def softmax(x):
    return np.exp(x) / np.sum(np.exp(x), axis = 0)

print(softmax(scores))
```

```
[ 0.8360188  0.11314284  0.05083836]
```





Função Softmax

$$\phi_i = \frac{e^{z_i}}{\sum_{j \in \text{group}} e^{z_j}}$$



Hot-Encoding

Hot-Encoding



Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2

$$y = (x \cdot w) + b$$

good

2.0

$p = 0.7$

Good

1.0

$p = 0.2$

Medium

0.1

$p = 0.1$

Bad



Data Science Academy



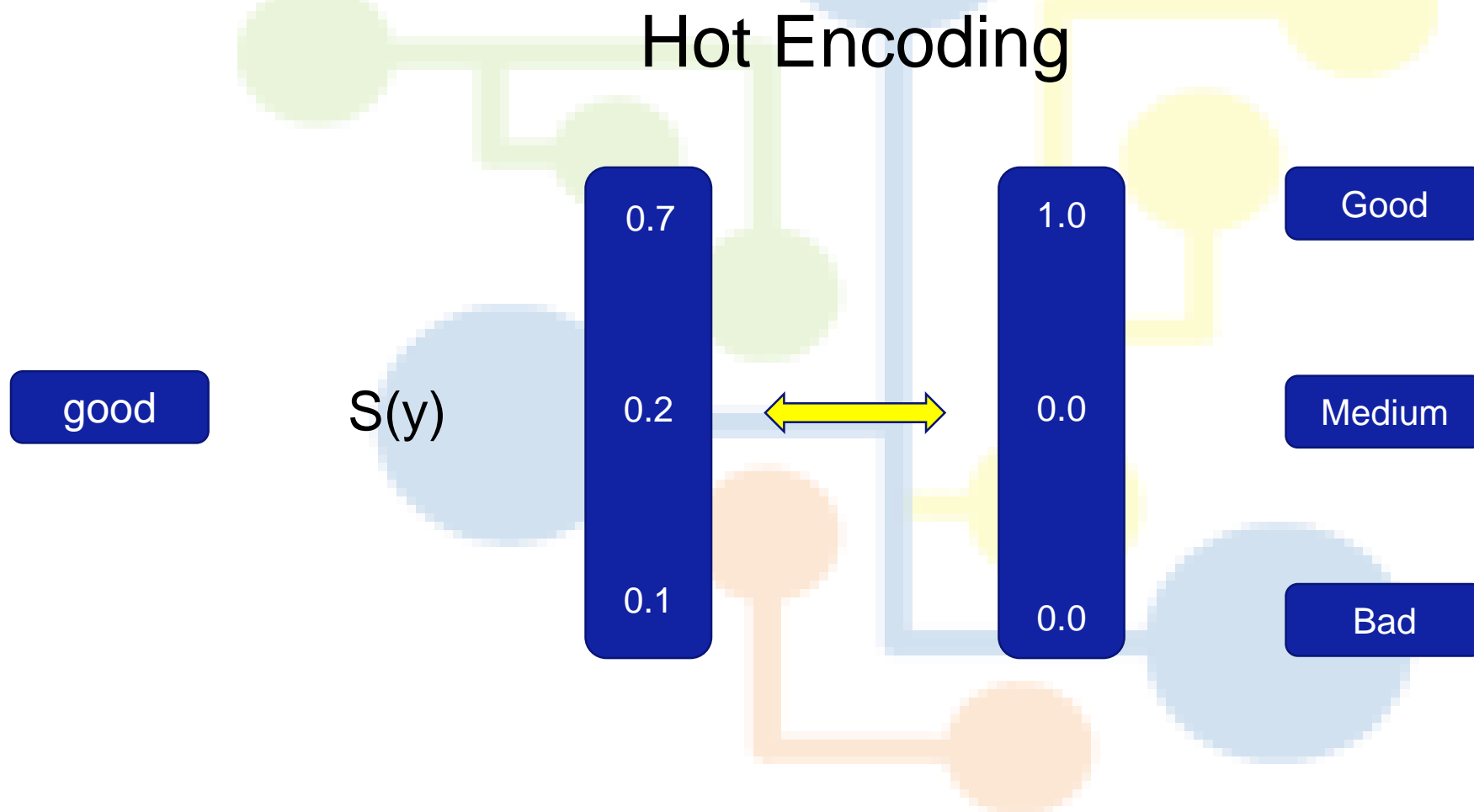
Data Science Academy

Hot-Encoding



Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2



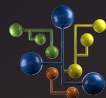
$S(y)$ = Função Softmax que converte os scores para probabilidades



Data Science Academy



Data Science Academy



Data Science
Academy

Data Science Academy raphaelbsfontenelle@gmail.com 615c1fdde32fc361b30c9ec2

Obrigado



Data Science Academy



Data Science Academy