

Mini-Curso: Introdução à Inteligência Artificial com Redes Neurais Artificiais



Parte 1

Professor José Francisco Pessanha

06 de dezembro de 2024

15:30 – 18:00, sala RAV62, 6º andar, bloco F

https://homeprojextransicaoenergetica.netlify.app/_site/eventos

Projeto de Extensão

TRANSIÇÃO ENERGÉTICA: vantagens e desafios técnicos das energias renováveis para o equilíbrio entre custos, segurança e mudanças climáticas



Departamento de
Estatística



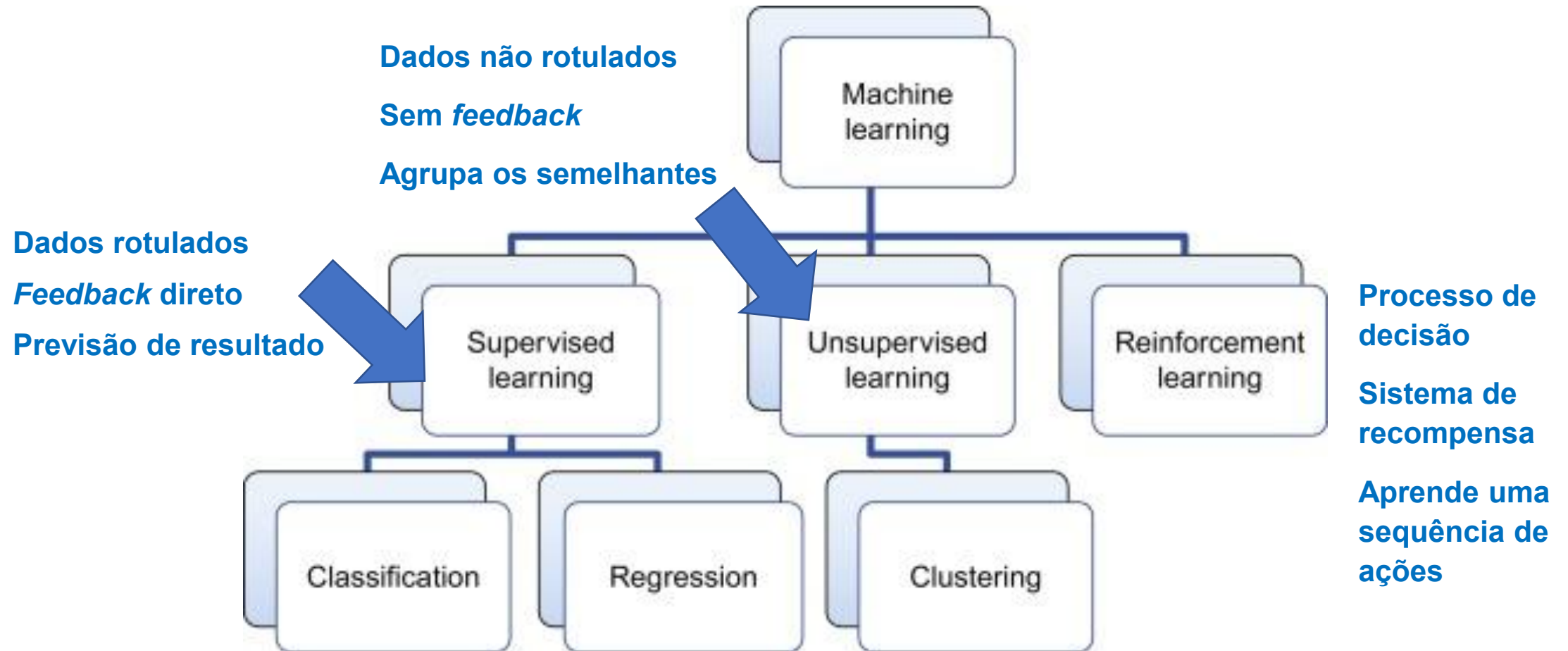
ELE
Depto. de Eng. Elétrica



Redes Neurais Artificiais - RNA

- Modelo de aprendizado de máquina inspirado na estrutura neural de organismos inteligentes e que adquirem conhecimento através da experiência.
- As redes neurais artificiais estão no centro nevrálgico do aprendizado de máquina (GÉRON, 2021)
- Permite construir sistemas inteligentes, i.e., com capacidade de aprender, assimilar, errar e aprender com seus erros.
- Diversas arquiteturas de redes
- Redes Neurais Supervisionadas: *redes multilayer perceptron* MLP em problemas de regressão e classificação
- Redes Neurais Não Supervisionadas: *self organizing map* SOM na análise de agrupamentos (clustering)

Tipos de aprendizado



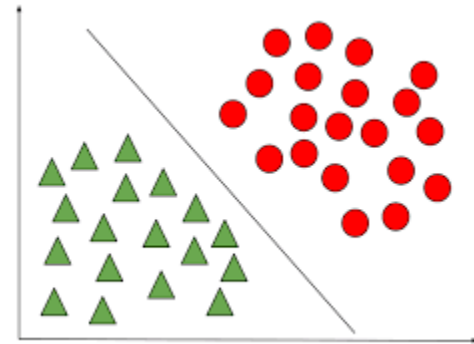
RNA para aprendizado supervisionado

Principais tipos de problemas de aprendizado supervisionado: $Y = f(X_1, X_2, X_3, \dots, X_k)$

1. Classificação: a rede aprende a categorizar entradas em classes (**Y é uma variável qualitativa**).

Exemplos:

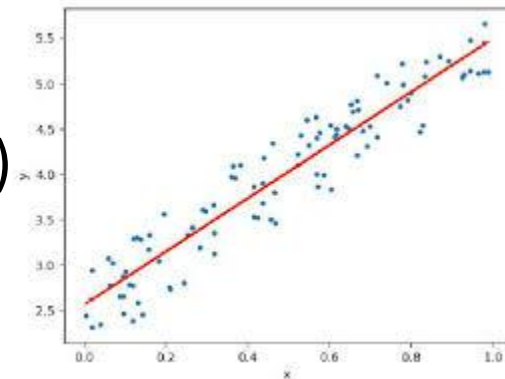
- classificar contribuintes nas categorias { adimplentes , inadimplentes }
- classificar transações nas categorias { normais , suspeitas }
- classificar segurados nas categorias { sinistrados , não sinistrados }
- classificar e-mails nas categorias { spam , não spam }



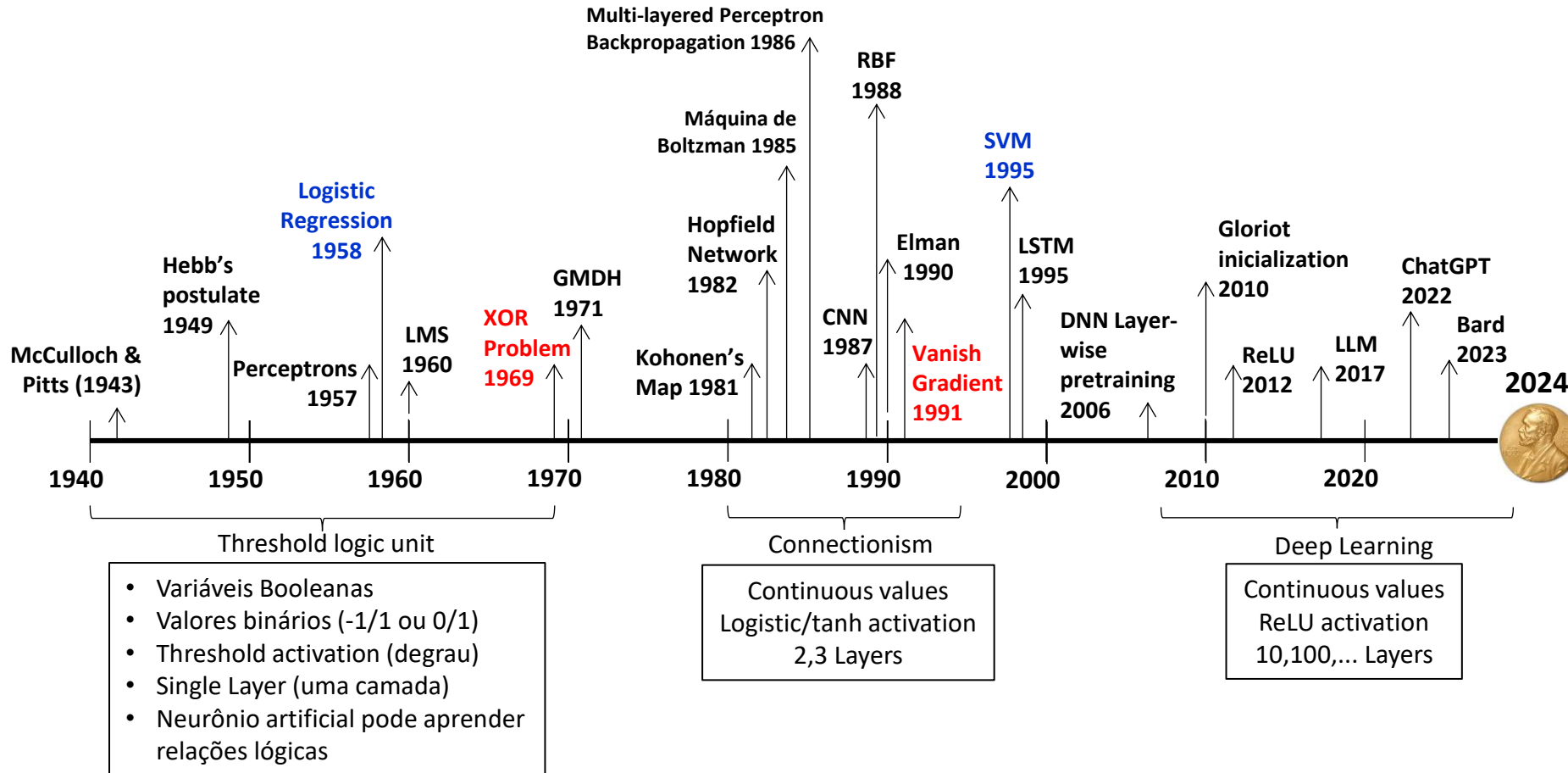
1. Regressão: a rede aprende a prever um valor contínuo (**Y é uma variável quantitativa**).

Exemplos:

- prever o preço de uma casa (Y) com base em suas características (X)
- prever as vendas (Y) em função das características do mercado atendido (X)
- prever a geração eólica (Y) em função da velocidade do vento (X)



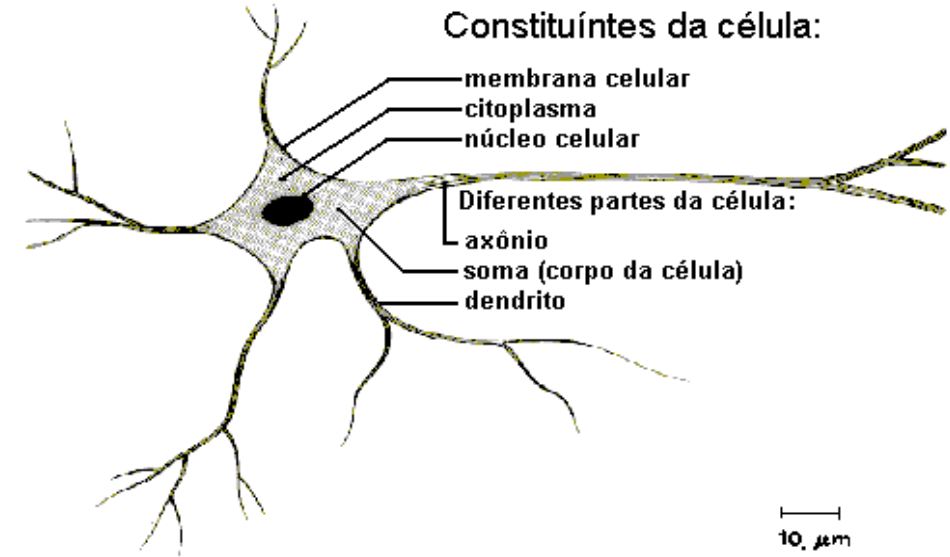
Marcos do desenvolvimento das redes neurais artificiais



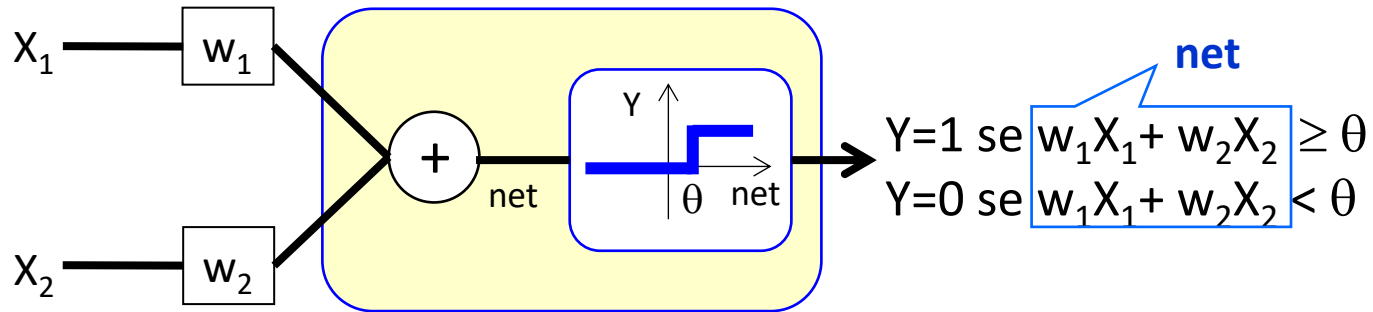
Neurônio Artificial (1943)

Neurônio biológico: Os impulsos recebidos por um neurônio são processados, e atingindo um dado limiar (*threshold* θ), o neurônio dispara, produzindo uma substância neurotransmissora que flui do corpo celular para o axônio que se conecta a um dendrito de outro neurônio.

O objetivo dos neurônios é processar sinais que transportam informações (SEJNOWSKI, 2019)



Neurônio artificial de McCulloch e Pitts



Primeiro modelo computacional do neurônio
(*Threshold Logic Unit – TLU*)

Caso com n variáveis de entrada

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

$$y = g(net) = \begin{cases} 1 & \text{se } g(net) \geq \theta \\ 0 & \text{se } g(net) < \theta \end{cases}$$

Retropropagação do erro (1974)

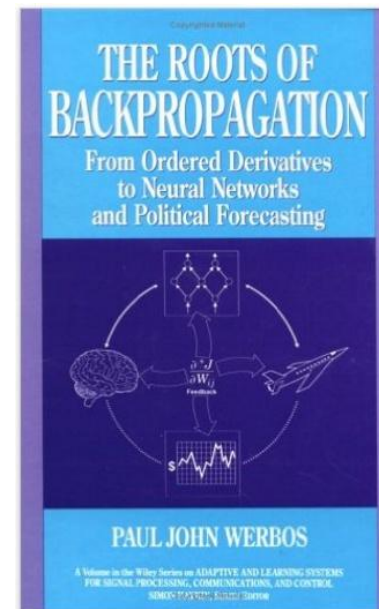
Paul Werbos Introduziu o método de retropropagação do erro (*error backpropagation*).

Tese de doutorado em Harvard: *Beyond regression: new tools for prediction and analysis in the behavioral sciences* [Beyond regression: new tools for prediction and analysis in the behavioral sciences](#)

Simon Haykin, the editor of this series, offered to publish my 1974 Ph.D. thesis, *Beyond Regression*, because this thesis has become something of a classic reference in the neural network and engineering world. The thesis is now recognized as the original source of *backpropagation*, which is now the most widely used algorithm by far in the neural network world. Also, the thesis *communicates* that algorithm to an audience with *no prior understanding* of neural networks;



Paul J. Werbos
1947



Rumelhart, Hinton e Williams (1986)

Redescobriram e desenvolveram o método de retropropagação do erro para o ajuste dos pesos sinápticos.

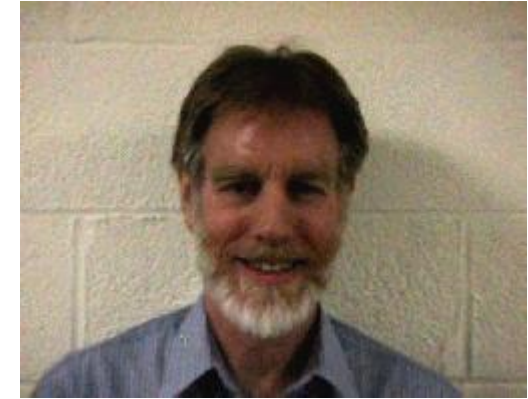
O método de retropropagação do erro permite treinar eficientemente redes com camadas intermediárias (*Multilayer Perceptron MLP*).



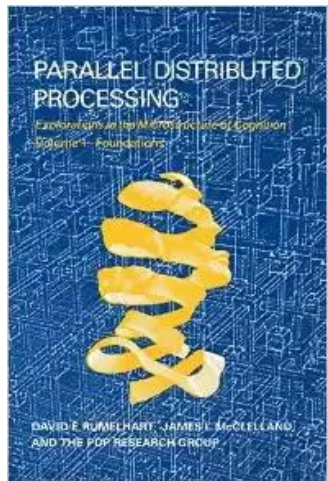
David Rumelhart
1942 - 2011



Geoffrey E. Hinton
1947



Ronald J. Williams
1945-2024

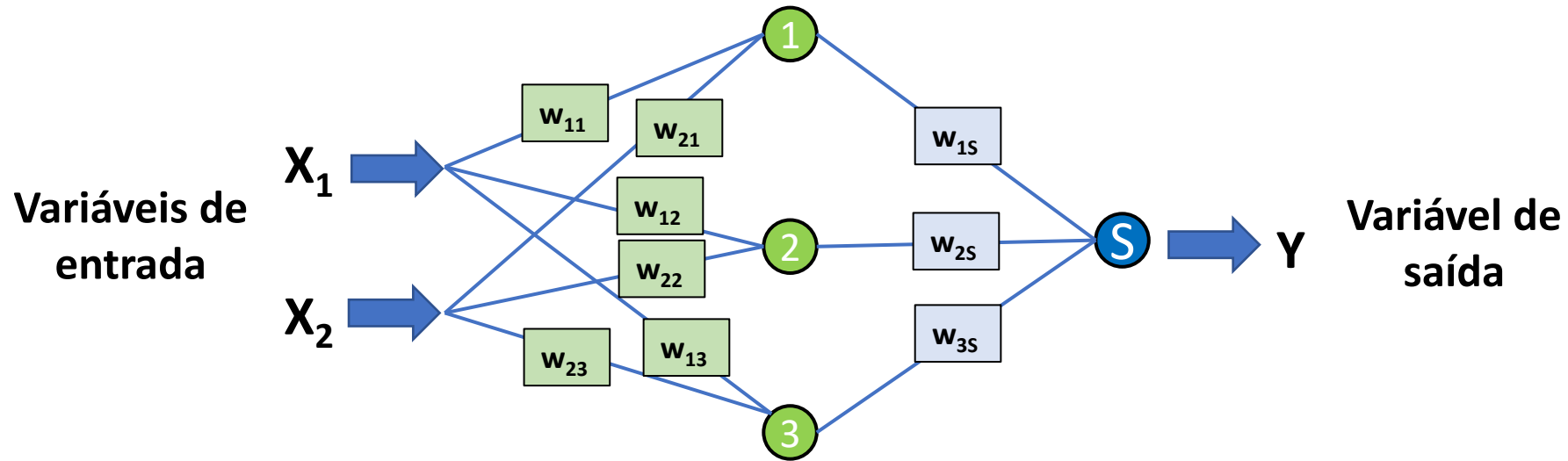


Rumelhart, D. E., Hinton, G. E. & Williams, R. J. in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition. Vol. 1: Foundations* (eds Rumelhart, D. E. & McClelland, J. L.) 318–362 (MIT, Cambridge, 1986)

All the knowledge is in the connections

David Rumelhart

Multilayer Perceptron - MLP



$$Y = g[w_{1S}f_1(w_{11}X_1 + w_{21}X_2) + w_{2S}f_2(w_{12}X_1 + w_{22}X_2) + w_{3S}f_3(w_{13}X_1 + w_{23}X_2)]$$

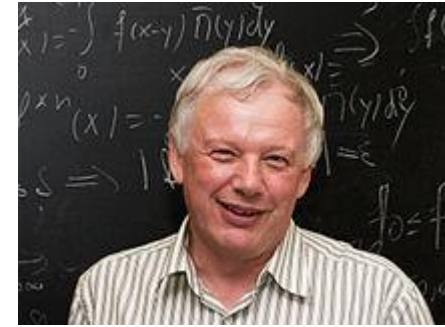
$$Y = g \left(\sum_{k=1}^3 w_{kS} f_k \left(\sum_{j=1}^2 w_{jk} x_j \right) \right)$$

Uma RNA é uma regressão não linear



Redes RBF – Radial Basis Function (1988)

- Proposta por Broomhead e Lowe em 1988
- Função de ativação gaussiana nos neurônios da camada oculta
- Função de ativação linear no neurônio da camada de saída, a resposta da rede é uma combinação linear de funções de base radial das entradas.
- Aprendizagem rápida
- Utilizada nas mesmas tarefas da MLP



David S. Broomhead

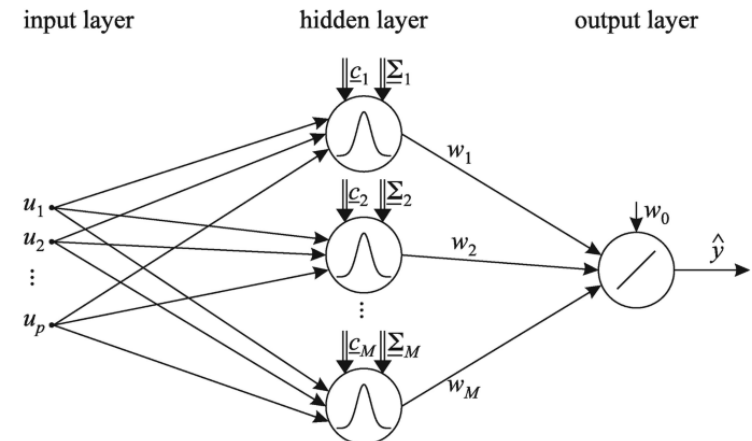
1950 - 2014



David Lowe

RADIAL BASIS FUNCTIONS, MULTI-VARIABLE FUNCTIONAL INTERPOLATION AND ADAPTIVE NETWORKS

Authors: D S Broomhead and D Lowe



Redes Neurais Recorrentes (RNN)

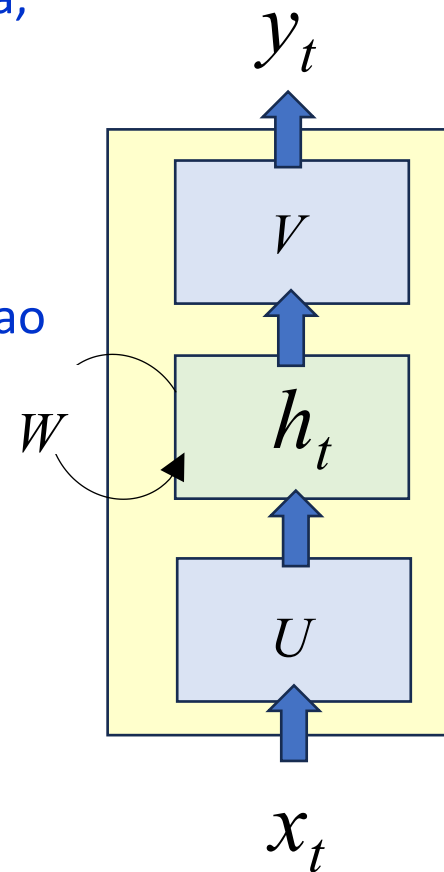
- O tipo mais comum de rede neural artificial projetado para lidar com dados de séries temporais é a Rede Neural Recorrente (RNN) (GERRISH, 2019).
- Em uma RNN, a saída em um dado instante depende de elementos anteriores na sequência, refletindo as dependências temporais dos dados.
- Unidades recorrentes mantêm um estado oculto (h) que armazena informações sobre entradas anteriores em uma sequência. Essas unidades podem "lembrar" informações de etapas anteriores ao retroalimentar seu estado oculto, permitindo capturar dependências ao longo do tempo
- A RNN considera a dependência sequencial entre os elementos ao processar informações.

$$h_t = f(W \cdot h_{t-1} + U \cdot x_t)$$

$$y_t = g(V \cdot h_t)$$

A recorrência dos estados latentes h permite que a RNN capture a estrutura de dependência temporal presente em séries temporais

- Em uma RNN, o compartilhamento de pesos (U, V, W) garante que os mesmos parâmetros sejam aplicados a cada passo da sequência, preservando a estrutura temporal dos dados

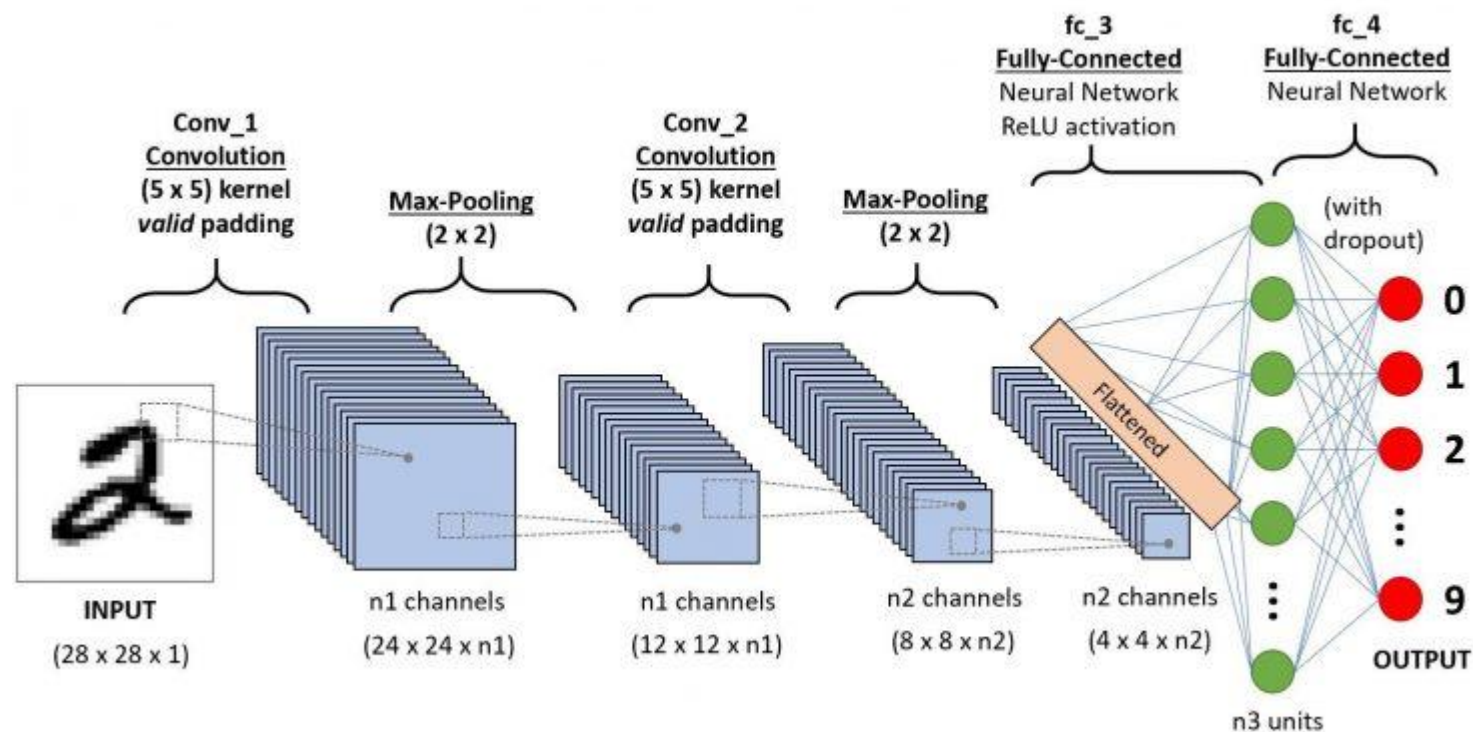


Redes convolucionais – CNN (1999)



Yann LeCun
1960

Conjunto de treinamento



<https://datasciencepr.com/convolutional-neural-network/>

Redes LSTM (1997)



Sepp Hochreiter
1967



Jürgen Schmidhuber
1963

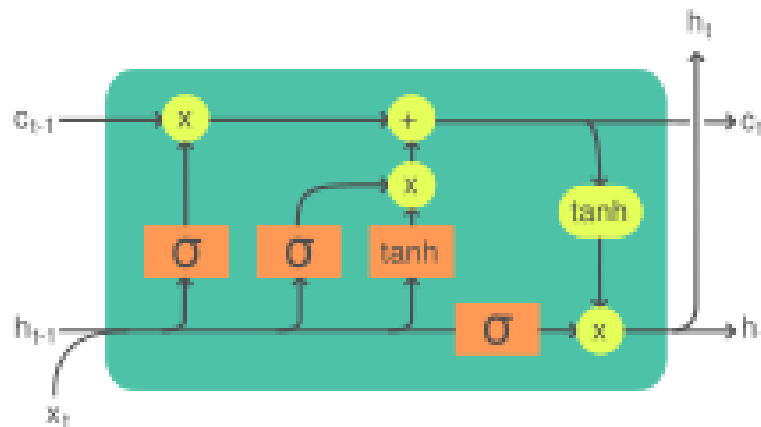
LONG SHORT-TERM MEMORY

NEURAL COMPUTATION 9(8):1735–1780, 1997

Sepp Hochreiter
Fakultät für Informatik
Technische Universität München
80290 München, Germany
hochreit@informatik.tu-muenchen.de
<http://www7.informatik.tu-muenchen.de/~hochreit>

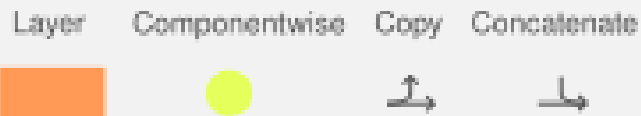
Jürgen Schmidhuber
IDSIA
Corso Elvezia 36
6900 Lugano, Switzerland
juergen@idsia.ch
<http://www.idsia.ch/~juergen>

<https://www.bioinf.jku.at/publications/older/2604.pdf>



A célula LSTM pode
processar dados
sequencialmente e
manter seu estado
oculto ao longo do
tempo

Legend:



- Evolução das redes neurais recorrentes (RNN)
- Processa dados sequenciais
- Processamento de linguagem natural
- Tradução
- Legendagem de imagens
- Reconhecimento de escrita
- Conversão de vídeo para texto

Considerações finais

Foram apresentadas algumas arquiteturas de redes neurais artificiais para treinamento supervisionado: MLP, RBF, GMDH, RNN e LSTM

Aplicações apresentadas: ajuste de funções e previsão

As redes neurais artificiais ocupam um lugar de destaque na ciência de dados.

As redes neurais artificiais são bastante flexíveis e funcionam como uma ferramenta genérica para modelagem em diferentes tipos de problemas em *business intelligence* e mineração de dados.

Os recursos disponibilizados pelo R e Python facilitam a implementação das redes neurais artificiais

Os códigos disponibilizados na apresentação podem ser facilmente adaptados para outras aplicações.