

CURSO SUPERIOR DE ADS

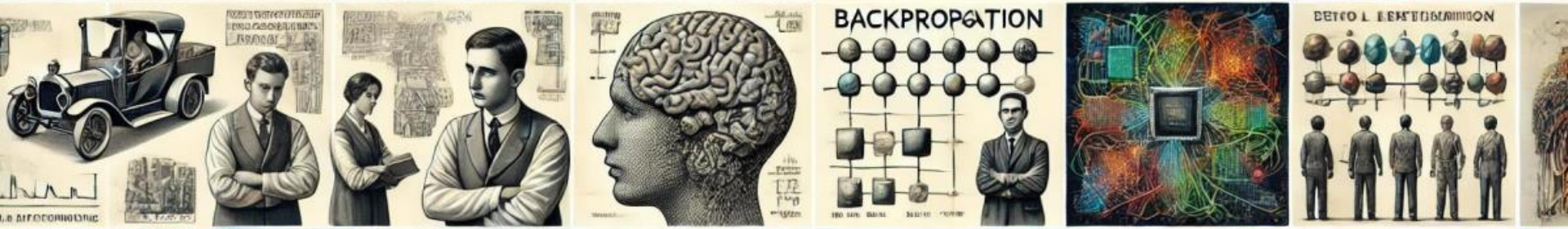
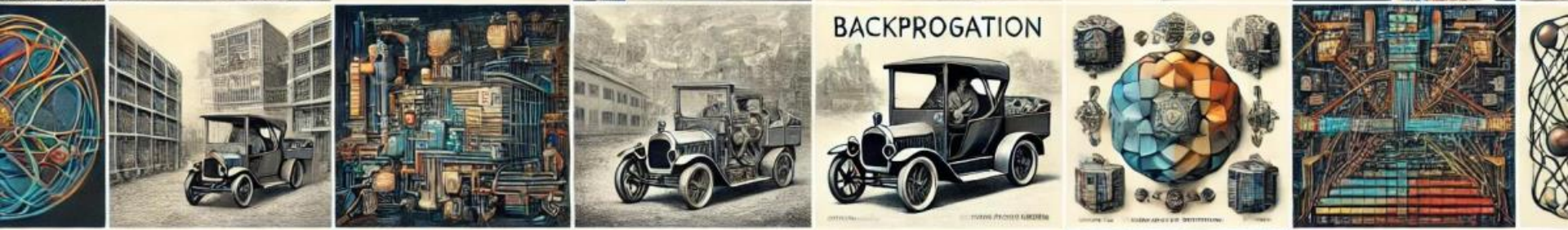
Deep Learning



Prof. Fernando Marlon Soares Figueiredo

Disciplina: Ciência de Dados e Bigdata





DEEP LEARNING



A **inteligência artificial** encontra a **inteligência humana**

"Se quiser entender a IA, você precisa ler A Revolução do Aprendizado Profundo."

—ERIK BRYNJOLFSSON, professor do MIT e coautor de *A Segunda Era Das Máquinas*

A REVOLUÇÃO DO APRENDIZADO PROFUNDO

TERRENCE J. SEJNOWSKI



- No livro *A Revolução do Aprendizado Profundo*, Terrence J. Sejnowski apresenta uma visão fascinante da trajetória das redes neurais e do aprendizado profundo, desde suas primeiras teorias até seu impacto transformador na sociedade atual. A obra explora como o avanço da tecnologia e a disponibilidade massiva de dados permitiram que o deep learning revolucionasse áreas como visão computacional, tradução de linguagem e saúde, mudando nossa relação com a inteligência artificial. Além de traçar a história dos grandes pioneiros do campo, Sejnowski aborda as implicações éticas e sociais dessa tecnologia, discutindo como o aprendizado profundo poderá influenciar o futuro da humanidade e levantar questões sobre o potencial e os riscos da IA.

O que é o aprendizado profundo?

- O aprendizado profundo é um método de inteligência artificial (IA) que ensina computadores a processar dados de uma forma inspirada pelo cérebro humano.
- Os modelos de aprendizado profundo conseguem reconhecer padrões complexos em imagens, texto, sons e outros dados para produzir insights e previsões precisas.
- É possível usar métodos de aprendizado profundo para automatizar tarefas que normalmente exigem inteligência humana, como descrever imagens ou transcrever um arquivo de som em texto.

Por que o aprendizado profundo é importante?

- A inteligência artificial (IA) tenta treinar computadores para pensar e aprender como os humanos. A tecnologia de aprendizado profundo impulsiona muitas aplicações de IA usadas em produtos cotidianos, como:
 - Assistentes digitais
 - Controles remotos de televisão ativados por voz
 - Detecção de fraudes
 - Reconhecimento facial automático

Por que o aprendizado profundo é importante?

- Também é um componente essencial das tecnologias emergentes, como carros autônomos, realidade virtual e muito mais.
- Os modelos de aprendizado profundo são arquivos de computador treinados por cientistas de dados para executar tarefas usando um algoritmo ou um conjunto predefinido de etapas.
- As empresas usam modelos de aprendizado profundo para analisar dados e fazer previsões em várias aplicações.

Quais são os usos do aprendizado profundo?

- O aprendizado profundo tem vários casos de uso, como no setor automotivo, aeroespacial, de manufatura, eletrônico, de pesquisa médica, entre outros. Estes são alguns exemplos de aprendizado profundo:

Os carros autônomos usam modelos de aprendizado profundo para detectar sinais de trânsito e pedestres automaticamente.

Os sistemas de defesa usam aprendizado profundo para sinalizar automaticamente áreas de interesse em imagens de satélite.

A análise de imagens médicas usa aprendizado profundo para detectar automaticamente células cancerosas para diagnóstico médico.

As fábricas usam aplicações de aprendizado profundo para detectar automaticamente quando pessoas ou objetos estão a uma distância de risco das máquinas.

Visão computacional

- A [visão computacional](#) é a capacidade dos computadores de extrair informações e insights de imagens e vídeos.
- Os computadores podem usar as técnicas de aprendizado profundo para compreender imagens da mesma forma que os humanos. A visão computacional tem várias aplicações, como:
- Moderação de conteúdo para remover automaticamente conteúdos inapropriados ou perigosos de arquivos de vídeos e imagens
- Reconhecimento facial para identificar rostos e reconhecer atributos como olhos abertos, óculos e barba
- Rótulo de imagens para classificar logotipos de marcas, roupas, equipamentos de segurança e outros detalhes de imagens

Visão computacional

Moderação de conteúdo para remover automaticamente conteúdos inapropriados ou perigosos de arquivos de vídeos e imagens

Reconhecimento facial para identificar rostos e reconhecer atributos como olhos abertos, óculos e barba

Rótulo de imagens para classificar logotipos de marcas, roupas, equipamentos de segurança e outros detalhes de imagens

Reconhecimento de voz

- Os modelos de aprendizado profundo podem analisar a fala humana mesmo existindo diversos padrões de fala, tom, altura, idioma e sotaque. Assistentes virtuais como a Amazon Alexa e software de transcrição automática usam o reconhecimento de fala para realizar estas tarefas:
- Ajudar atendentes de centrais de atendimento e classificar chamadas automaticamente.
- Converter conversas clínicas em documentação em tempo real.
- Criar legendas de vídeos e gravações de reuniões com precisão para permitir um maior alcance do conteúdo

Processamento de linguagem natural

- Os computadores usam algoritmos de aprendizado profundo para reunir insights e significados de dados e documentos de texto. A capacidade de processar texto natural criado por humanos tem vários casos de uso, inclusive nestas funções:

Chatbots e atendentes virtuais automatizados

Resumo automático de documentos ou artigos jornalísticos

Análise de business intelligence de documentos longos, como e-mails e formulários

Indexação de expressões-chave que indicam sentimentos, como comentários positivos e negativos em mídias sociais

Mecanismos de recomendação

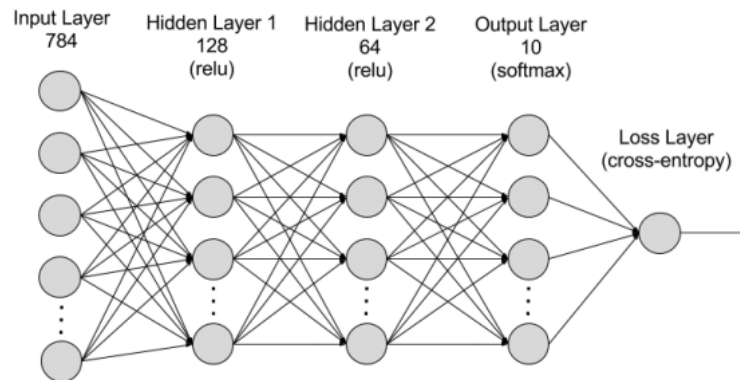
- As aplicações podem usar métodos de aprendizado profundo para rastrear a atividade do usuário e desenvolver recomendações personalizadas. Podem analisar o comportamento de vários usuários e ajudá-los a descobrir novos produtos ou serviços.
- Por exemplo, muitas empresas de mídia e entretenimento, como Netflix, Fox e Peacock, usam aprendizado profundo para oferecer recomendações personalizadas de vídeo.

Como funciona o aprendizado profundo?

- Os algoritmos de aprendizado profundo são redes neurais modeladas conforme o cérebro humano.
- Por exemplo, o cérebro humano contém milhões de neurônios interconectados que trabalham juntos para aprender e processar informações.
- Da mesma forma, redes neurais de aprendizado profundo, ou redes neurais artificiais, são compostas por muitas camadas de neurônios artificiais que trabalham juntas dentro do computador.

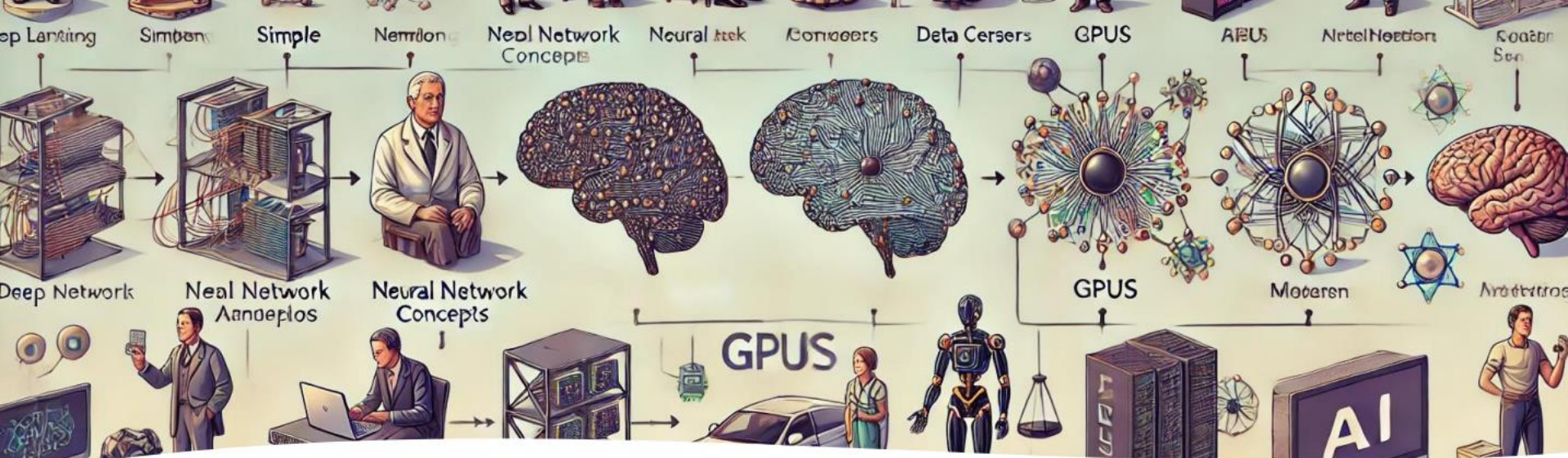
Como funciona o aprendizado profundo?

- Os neurônios artificiais são módulos de software chamados nós, que utilizam cálculos matemáticos para processar dados.
- As redes neurais artificiais são algoritmos de aprendizado profundo que usam esses nós para solucionar problemas complexos.



Quais são os componentes de uma rede de aprendizado profundo?

- Os componentes de uma rede neural profunda estão listados abaixo.
- **Camada de entrada**
- Uma rede neural artificial conta com vários nós que inserem dados nela. Esses nós compõem a camada de entrada do sistema.
- **Camada oculta**
- A camada de entrada processa e transmite os dados para camadas mais distantes na rede neural. Essas camadas ocultas processam informações em diferentes níveis, adaptando o comportamento à medida que recebem novas informações. As redes de aprendizado profundo têm centenas de camadas ocultas que podem ser usadas para analisar um problema por vários ângulos diferentes.



Componentes de Deep Learning

- Por exemplo, se recebesse a imagem de um animal desconhecido e precisasse classificá-lo, você o compararia aos animais que já conhece. Por exemplo, observaria o formato dos olhos e das orelhas, tamanho, número de pernas e padrão de pelo. Você tentaria identificar padrões, tais como:
 - O animal tem cascos, então pode ser uma vaca ou um cervo.
 - O animal tem olhos de gato, então pode ser algum tipo de gato selvagem.



Componentes de Deep Learning

- As camadas ocultas em redes neurais profundas funcionam da mesma forma. Se um algoritmo de aprendizado profundo está tentando classificar uma imagem de animal, cada uma de suas camadas ocultas processa uma característica diferente do animal e tenta categorizá-la com precisão.



Camada de saída

- A camada de saída consiste nos nós que emitem dados. Modelos de aprendizado profundo que geram respostas “sim” ou “não” têm apenas dois nós na camada de saída. Por sua vez, aqueles que produzem uma gama maior de respostas têm mais nós.



O que é aprendizado profundo no contexto do machine learning?

Rotular manualmente centenas de milhares de imagens de animais.

Fazer com que os algoritmos de machine learning processem essas imagens.

Testar esses algoritmos em um conjunto de imagens desconhecidas.

Identificar por que alguns resultados são imprecisos.

Aprimorar o conjunto de dados rotulando novas imagens para melhorar a precisão dos resultados.



- Esse processo é chamado de **aprendizado supervisionado**.
- No **aprendizado supervisionado**, a precisão dos resultados só melhora quando há um conjunto de dados amplo e bastante variado.
- Por exemplo, o algoritmo pode identificar gatos pretos com precisão, mas não gatos brancos, pois o conjunto de dados de treinamento continha mais imagens de gatos pretos.
- Nesse caso, seria necessário rotular mais imagens de gatos brancos e treinar os modelos de machine learning mais uma vez.



Quais são os benefícios do aprendizado profundo em relação ao machine learning?

- **Processamento eficiente de dados não estruturados**
- Os métodos de machine learning consideram dados não estruturados, como documentos de texto, difíceis de processar porque o conjunto de dados de treinamento pode ter infinitas variações.
- Por sua vez, os modelos de aprendizado profundo podem compreender dados não estruturados e fazer observações gerais sem extrair recursos manualmente. Por exemplo, uma rede neural pode reconhecer que essas duas frases de entrada diferentes têm o mesmo significado:

Você pode me explicar como fazer o pagamento?

Como eu transfiro o dinheiro?



Relacionamentos ocultos e descoberta de padrões

- Uma aplicação de aprendizado profundo pode analisar grandes quantidades de dados mais profundamente e revelar novos insights para os quais talvez não tenha sido treinada. Por exemplo, pense em um modelo de aprendizado profundo treinado para analisar as compras do consumidor. O modelo tem dados somente de produtos que você já comprou. Porém, a rede neural artificial pode sugerir novos produtos que você não comprou, comparando seus padrões de compra aos de outros clientes semelhantes.

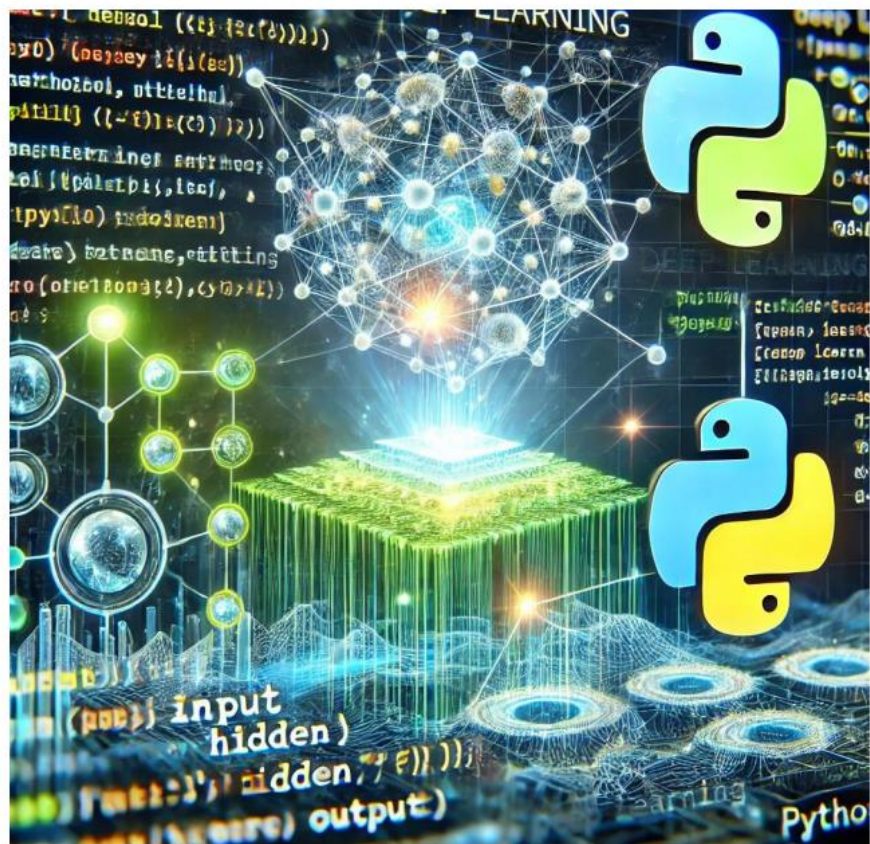
Aprendizado sem supervisão

- Os modelos de aprendizado profundo podem aprender e melhorar ao longo do tempo com base no comportamento do usuário.
- Não necessitam de grandes variações de conjuntos de dados rotulados. Por exemplo, considere uma rede neural que corrija ou sugira automaticamente palavras por meio da análise do comportamento de digitação de um usuário.
- Digamos que tenha sido treinado na língua inglesa e possa corrigir erros ortográficos de palavras em inglês. Porém, se você digita frequentemente palavras em outros idiomas, como *danke*, a rede neural aprende automaticamente e corrige essas palavras também.
- **Processamento de dados voláteis**
- Conjuntos de dados voláteis têm grandes variações. Um exemplo são os valores de reembolso de empréstimos de um banco. Uma rede neural de aprendizado profundo também pode categorizar e classificar esses dados, por exemplo, analisando transações financeiras e sinalizando algumas delas para detecção de fraudes.

Quais são os desafios do aprendizado profundo?

1. Desafios do aprendizado profundo:

Como tecnologia emergente, o aprendizado profundo depende de grandes quantidades de dados de alta qualidade e de um processamento robusto. Qualquer erro nos dados pode afetar o desempenho, exigindo uma etapa rigorosa de pré-processamento e armazenamento de dados.



Quais são os benefícios do aprendizado profundo na nuvem?

1. Benefícios do aprendizado profundo na nuvem:

1. **Velocidade:** A nuvem permite o uso de clusters de GPUs e CPUs para treinar modelos rapidamente e processar grandes volumes de dados de forma eficiente.
2. **Escalabilidade:** Recursos de hardware na nuvem são praticamente ilimitados, possibilitando a distribuição de workloads entre diferentes processadores, atendendo a redes neurais de diferentes tamanhos.

O que é o aprendizado profundo na Nuvem?

- Os serviços de aprendizado profundo em nuvem utilizam computação escalável, permitindo a execução de redes neurais com alto desempenho e baixo custo. Abaixo estão alguns dos principais serviços de diferentes provedores de nuvem para aplicações de aprendizado profundo:



O que é o aprendizado profundo na Nuvem?

- **Google Cloud Platform (GCP)**
- **AutoML Vision:** Treinamento e implementação de modelos de visão personalizados sem necessidade de conhecimento avançado em machine learning.
- **Speech-to-Text:** Transcrição de áudio com reconhecimento de fala preciso e suporte para vários idiomas.
- **Dialogflow:** Plataforma para criação de chatbots avançados, utilizando processamento de linguagem natural.



Microsoft Azure

- **Azure Custom Vision:** Criação de modelos de visão computacional personalizados com integração simples.
- **Azure Speech:** API de reconhecimento e síntese de fala, com suporte para tradução em tempo real.
- **Azure Bot Service:** Desenvolvimento e hospedagem de chatbots que podem se conectar a diversas plataformas de mensagens.



Aprofundando no Aprendizado Profundo...



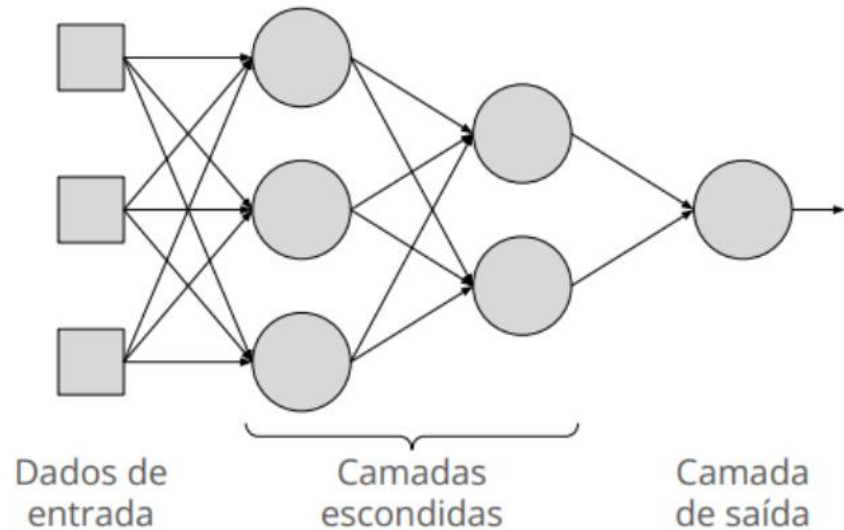
Arquitetura

Entrada: linha da base de dados

Saída: previsão

Neurônios organizados em camadas:

- Conectam com a próxima camada
- Geram uma resposta só (copiada)



REDES NEURAIS

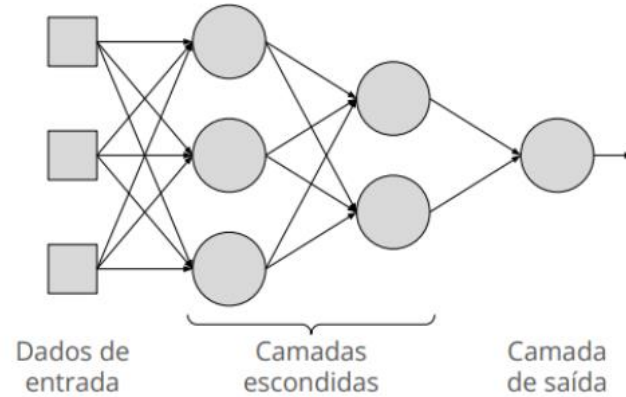
Arquitetura

Entrada: linha da base de dados

Saída: previsão

Neurônios organizados em camadas:

- Conectam com a próxima camada
- Geram uma resposta só (copiada)



Neurônio artificial

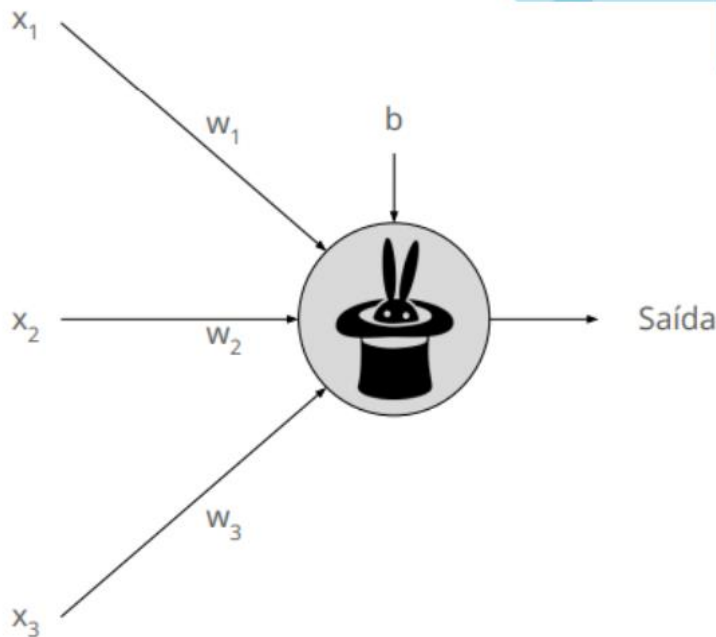
Entrada: linha da base de dados ou saída dos neurônios da camada anterior (x_i)

Saída: uma combinação das entradas

- Cada ligação tem um peso (w_i)
- Cada neurônio tem um bias (b)

A saída do neurônio é igual a:

$$f\left(\sum_i x_i w_i + b\right) = f(x_1 w_1 + x_2 w_2 + x_3 w_3 + b)$$



Função de ativação

É a função que se aplica a soma ponderada feita pelo neurônio:

$$\sum_i x_i w_i + b$$

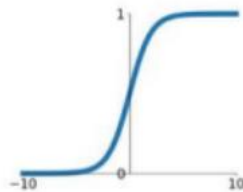
Para classificação e bastante comum usar:

- ReLU para as camadas escondidas
- Softmax para a última camada

$$\sigma(x_j) = \frac{e^{x_j}}{\sum_i e^{x_i}}$$

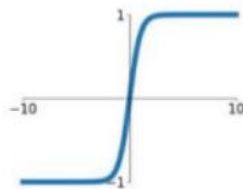
Sigmoid

$$\sigma(x) = \frac{1}{1+e^{-x}}$$



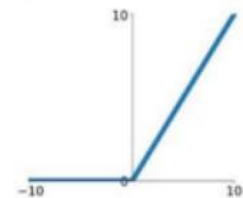
tanh

$$\tanh(x)$$



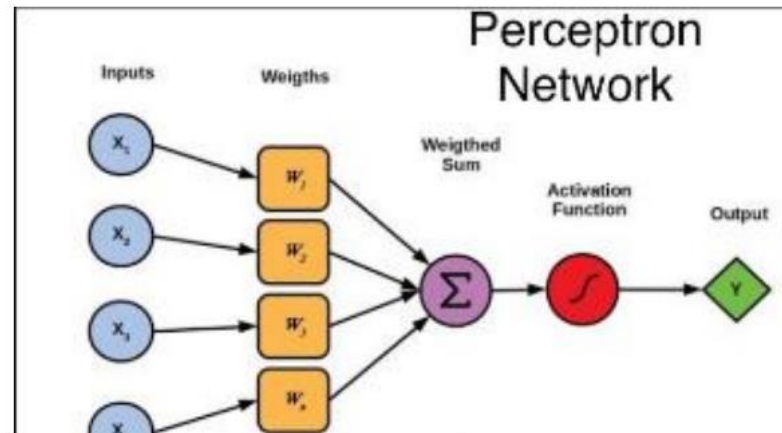
ReLU

$$\max(0, x)$$



Conceito de Perceptron

- O **perceptron** é o neurônio artificial mais básico e representa uma das estruturas iniciais das redes neurais. Ele funciona de maneira simples:
- Recebe uma ou mais entradas (x_1, x_2, \dots, x_n).
- Multiplica cada entrada por um peso (w_1, w_2, \dots, w_n).
- Soma esses produtos e aplica uma função chamada **função de ativação**.
- Se o valor resultante da soma ultrapassar um certo **limiar**, o perceptron gera uma saída (ativado). Caso contrário, não gera. Esse comportamento imita a resposta “ligado/desligado” dos neurônios biológicos.



Função de Ativação

- A função de ativação decide se o neurônio será ativado com base no valor da soma ponderada das entradas. Alguns exemplos comuns de funções de ativação são:

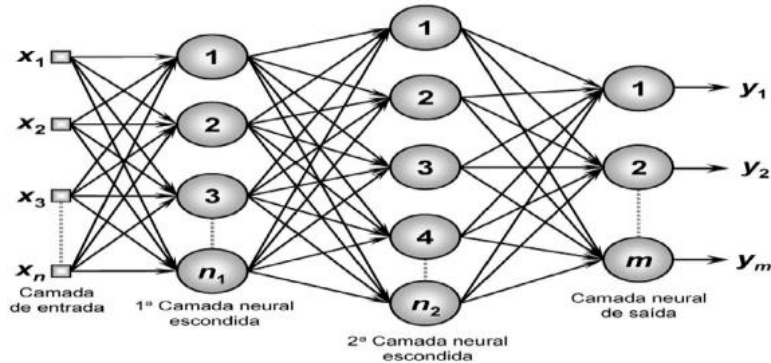
1. Função Degrau - Ativa se a soma for maior que um limiar específico.

2. Sigmoid - É uma função mais suave, útil em redes com várias camadas.

Problema da Linearidade e a Necessidade de Camadas Ocultas

- Um perceptron isolado é limitado; ele só pode resolver problemas que são **linearmente separáveis** (como "AND" e "OR").
- Para resolver problemas mais complexos (não linearmente separáveis), como o "XOR", é necessário introduzir **camadas ocultas** na rede.
- Cada camada adiciona complexidade e capacidade de aprendizado, permitindo à rede representar funções mais complexas.

Redes Neurais de Múltiplas Camadas (Multilayer Networks)

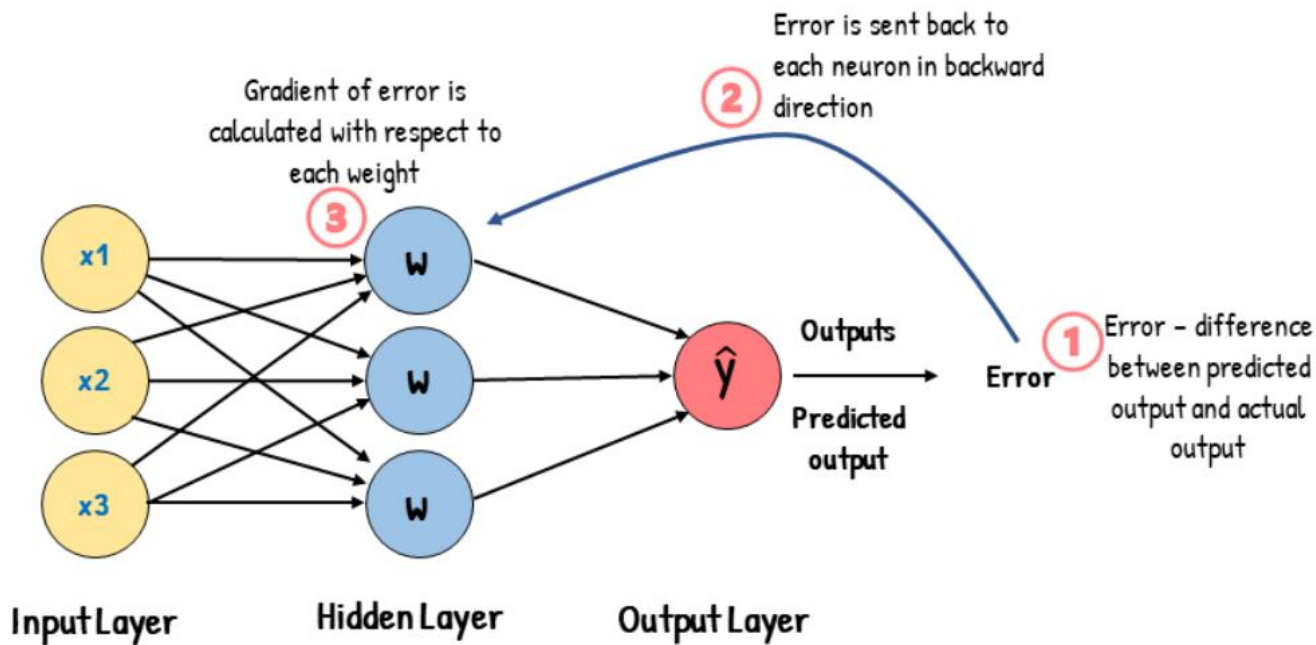


- As redes neurais podem ter várias camadas de perceptrons conectados, formando uma **rede neural de múltiplas camadas**. Cada camada contribui para o aprendizado de padrões mais abstratos, construindo representações mais sofisticadas dos dados. A partir disso, a rede pode resolver problemas complexos que exigem combinações não-lineares dos dados de entrada.

Algoritmo de Retropropagação (Backpropagation)

- Para que uma rede neural aprenda e melhore, utilizamos o **algoritmo de retropropagação**. Esse algoritmo ajusta os pesos de cada conexão com base nos erros cometidos pela rede durante a fase de treinamento. A retropropagação funciona assim:
 1. A rede faz uma previsão com um conjunto de pesos inicializados aleatoriamente.
 2. Compara a previsão com a resposta correta (erro).
 3. Propaga o erro para trás pela rede e ajusta os pesos para reduzir o erro.

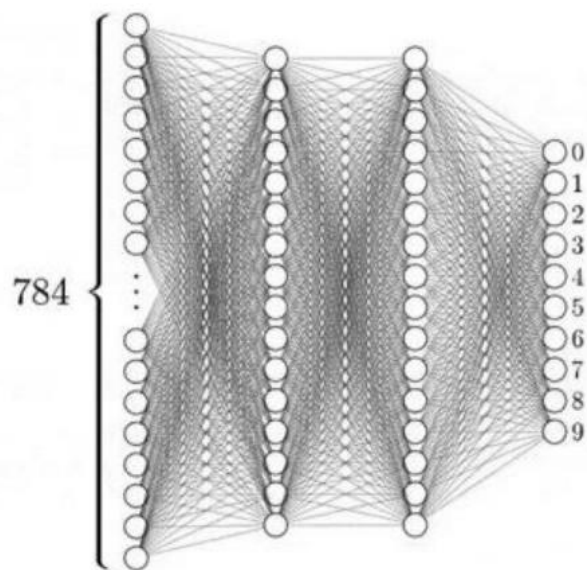
Backpropagation



Importância da Função Custo (Loss function)

- A **função custo** é uma medida de erro da rede. No treinamento, a rede tenta minimizar essa função custo, que mede a diferença entre as previsões da rede e os valores reais.
- No contexto matemático, uma função custo comum é o **erro quadrático médio**, onde o objetivo é minimizar a média dos quadrados dos erros para ajustar os pesos da rede de modo que ela melhore a cada iteração.

Quando uma rede neural é boa?



Quando a rede produz o resultado correto

Ou seja: **quando os pesos da rede geram as respostas corretas**

Como obter bons pesos para uma rede?



Algoritmo:

1. Iniciar com pesos aleatórios
2. Classificar dados de treinamento
3. Ver onde erramos (medir o erro)
4. Modificar os pesos para melhorar

Medir o erro do modelo

Para medir o erro usamos uma função de custo:

- MSE para regressão:

$$(y_i - \hat{y}_i)^2$$

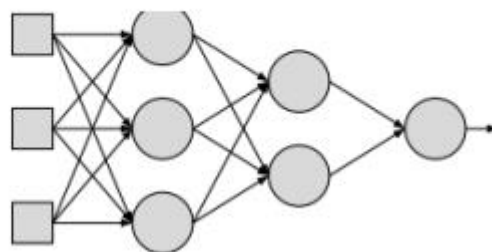
- Cross entropy para classificação:

$$-\sum_{i=0}^C y_i \cdot \log(\hat{y}_i)$$

Fazer a média na amostra



Resumo - Treinamento do modelo



Entrada:

- Rede neural com parâmetros \mathbf{W} e \mathbf{b} , para otimizar
- Um conjunto de treinamento $\mathbf{D}_{\text{train}}$
- Um valor para o learning rate α

Repetir n vezes:

- Computar os erros da rede neural para todos os elementos de $\mathbf{D}_{\text{train}}$
- Atualizar os parâmetros:

$$w = w - \alpha \frac{\delta E}{\delta w} \quad b = b - \alpha \frac{\delta E}{\delta b}$$

Saída:

- Rede treinada