

Sistemas de Visão e Percepção Industrial

6-Complementos sobre Percepção e Representação

Introdução ao Deep Learning

Universidade de Aveiro

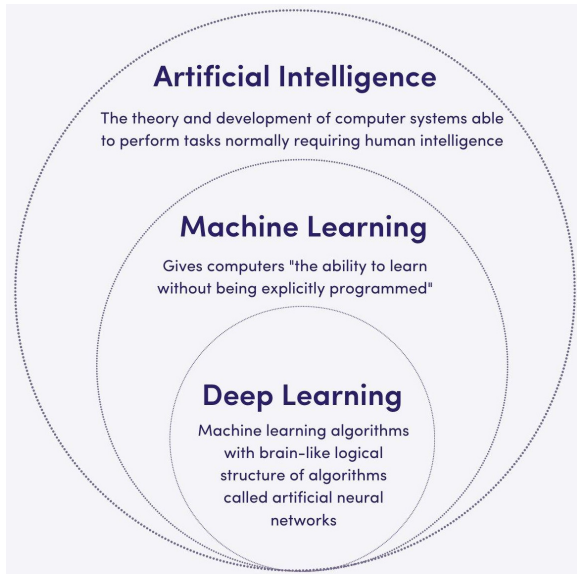
Joel Baptista, Vitor Santos, 2022-2023

- 1 Conceitos de Aprendizagem Profunda
- 2 Redes Neurais Convolucionais

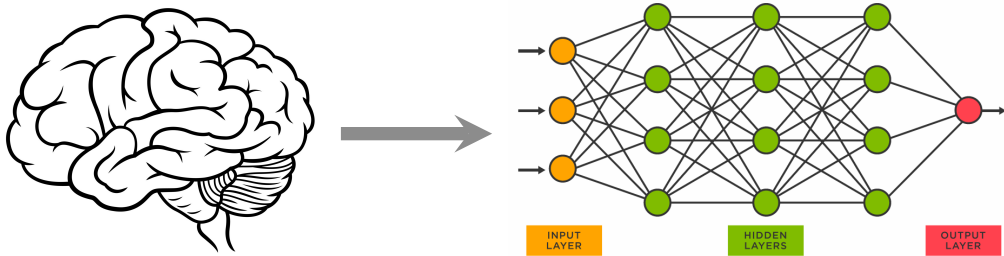
Conceitos de Aprendizagem Profunda

Aprendizagem Profunda (Deep Learning)

- A aprendizagem de máquinas (Machine Learning) é composta por técnicas e algoritmos que permitem que os computadores aprendam a partir de dados sem serem especificamente programados.
- A aprendizagem profunda (Deep Learning) é um subcampo do Machine Learning que usa redes neurais de várias camadas que permitem aprender representações complexas a partir de dados.
 - Isso é útil em problemas em que o modelo do sistema não é conhecido ou de difícil definição.



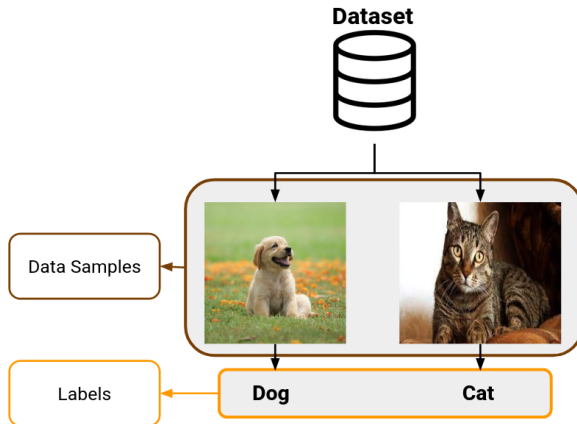
Redes neurais



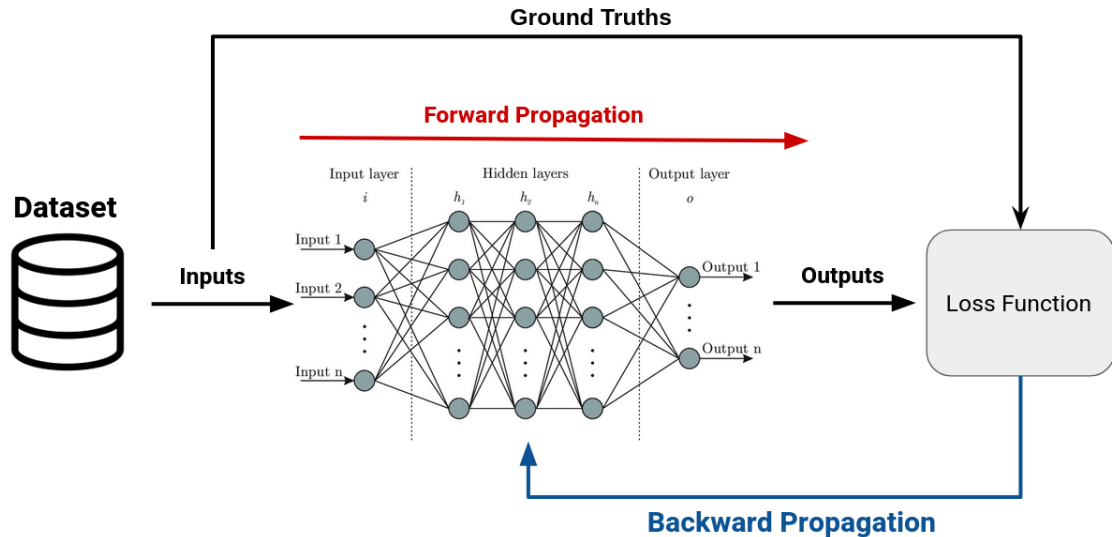
- Uma rede neuronal é um modelo de aprendizagem de máquinas inspirado na estrutura e função do cérebro humano.
- Consiste em neurões interligados organizados em camadas, que processam e transmitem informações por meio de ligações ponderadas.
- A rede é treinada com um conjunto de dados para aprender padrões e relações, e pode ser usada para fazer previsões sobre novos dados.

Conjunto de dados (Dataset)

- Um conjunto de dados é uma coleção de dados usados para treinar e testar modelos de aprendizagem de máquinas.
- Um conjunto de dados consiste em amostras de dados (entradas) e rótulos/"Ground truth" (saídas).
- Os conjuntos de dados podem ter uma variedade de fontes, como sensores ou câmaras, e podem incluir diferentes tipos de dados, como imagens, texto ou séries temporais.
- A qualidade do conjunto de dados é crucial em problemas de Deep Learning, influenciando fortemente a qualidade final do modelo.

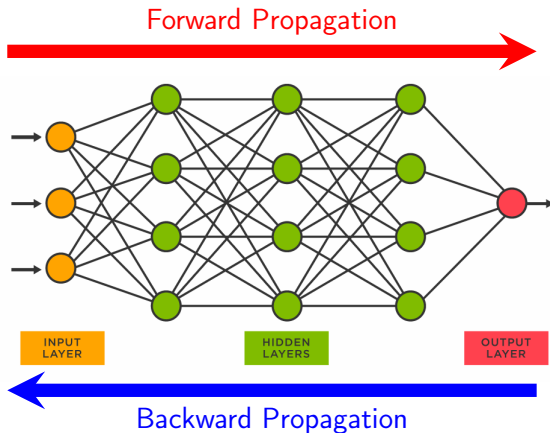


Processo de treino



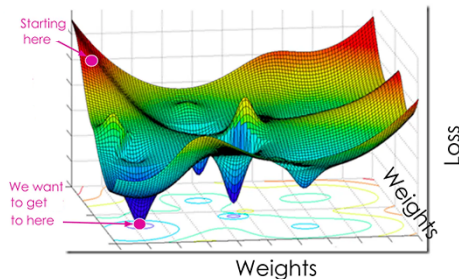
Propagação para frente e para trás

- **Propagação direta (Forward propagation):** o processo no qual a entrada é alimentada na camada de entrada, os cálculos são executados nas camadas ocultas e a saída é apresentada na camada de saída.
- **Retropropagação (Backward Propagation):** Este processo calcula o erro obtido no processo de propagação direta e atualiza o valor das ligações ponderadas em conformidade.
- Esse processo acontece de forma iterativa até a rede neuronal aprender a ter um bom desempenho nos dados disponíveis.



Funções de Perda (Loss function)

- Uma função de perda (Loss function) é uma função matemática que mede a diferença entre a saída prevista de uma rede neuronal artificial e a saída real (ground truth/rótulo).
- O objetivo da função de perda é quantificar o erro entre a saída prevista e a saída real, para que a rede aprenda a minimizar esse erro ao longo do tempo por meio do processo de retropropagação.
- A escolha da função de perda a ser usada depende da formulação do problema.



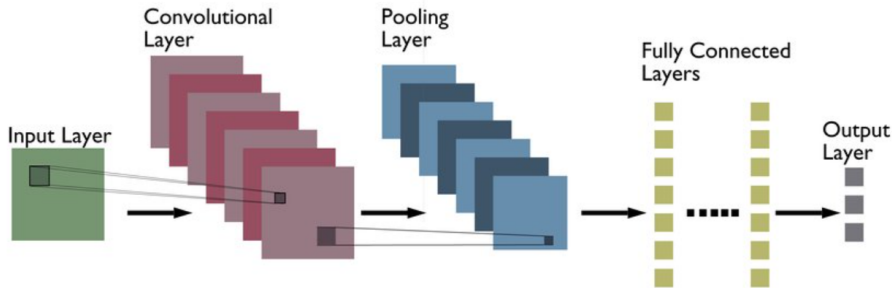
Noções Importantes sobre Redes Neurais

- Podem ser treinadas em conjuntos de dados para aprender padrões.
- Podem ser usadas para fazer previsões em novos dados.
- O processo de treino é iterativo, e opera em duas fases: propagação para frente e para trás.
- O treino conta com uma função de perda que calcula a diferença entre o valor previsto e o valor real.

Redes Neuronais Convolucionais

Redes Neurais Convolucionais

- As Redes Neurais Convolucionais são um tipo de redes neuronais artificiais projetadas principalmente para processar e interpretar imagens.
- Essas redes são compostas principalmente por:
 - Camadas Convolucionais
 - Camadas de "pooling" (agrupamento)
 - Camadas totalmente ligadas (fully connected)



Camadas de Convolução

- As camadas de convolução são usadas para extrair características (features) da imagem.
 - Isso faz-se aplicando filtros de convolução que podem ser aprendidos em toda a imagem.
- A saída dessas camadas é chamada de mapas de características (Feature Maps) e geralmente são matrizes aproximadamente do mesmo tamanho da entrada.
 - Mas, o tamanho da saída pode depender do preenchimento (padding) e do passo (stride).

Diagram illustrating a 1D convolution operation:

Input Matrix (8x8):

0	0	0	0	0	0	0	0
0	3	1	1	2	8	4	0
0	1	0	7	3	2	6	0
0	2	3	5	1	1	3	0
0	1	4	1	2	6	5	0
0	3	2	1	3	7	2	0
0	9	2	6	2	4	1	0
0	0	0	0	0	0	0	0

Filter (3x3):

1	0	-1
1	0	-1
1	0	-1

Resulting Feature Map (6x6):

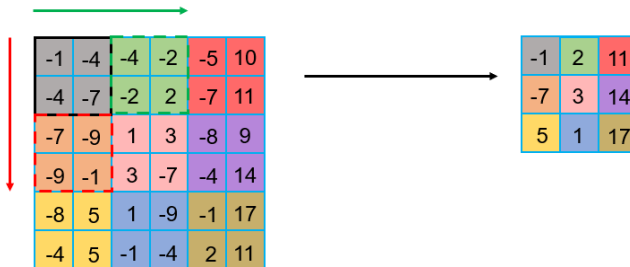
-1	-4	-4	-2	-5	10
-4	-7	-2	2	-7	11
-7	-9	1	3	-8	9
-9	-1	3	-7	-4	14
-8	5	1	-9	-1	17
-4	5	-1	-4	2	11

- Tamanho do filtro: 3x3 pixels
- Passo (stride): 1 pixel
- Preenchimento (padding): 1 pixel
- Tamanho da matriz de entrada: 6x6 pixels
- Tamanho da matriz de saída: 6x6 pixels
- Tamanho da matriz preenchida: 8x8 pixels

Imagem de entrada — Filtro — Feature map

Camadas de agrupamento (Pooling)

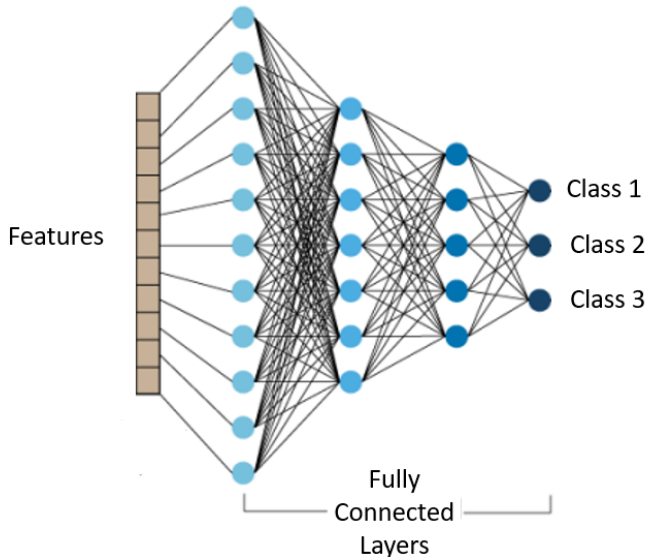
- As camadas de agrupamento (Pooling) são usadas para reduzir as dimensões espaciais dos Feature Maps, mantendo a informação importante.
- Normalmente, existem dois tipos de camadas de pooling:
 - Max Pooling: Somente o valor máximo é transmitido.
 - Average Pooling: Os valores de saída são iguais à média dos valores selecionados.



Max pooling: tamanho do filtro=2x2 e passo=2

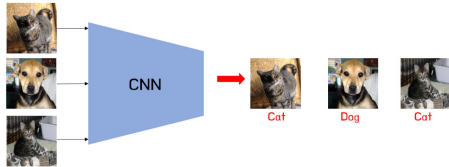
Camadas totalmente ligadas

- As camadas totalmente ligadas são usadas para classificar a imagem de entrada com base nos feature maps extraídos.
- Uma camada totalmente ligada é uma camada onde cada neurônio de entrada está ligado a cada neurônio de saída.
- A saída da última camada totalmente ligada é usada para fazer a previsão.



- As Redes Neurais Convolucionais são atualmente utilizadas em três formulações de problemas principais:
 - Reconhecimento de imagem: usado para classificar imagens em diferentes categorias.
 - Detecção de objetos: usado para detectar a localização de um objeto dentro de uma imagem, usando uma caixa delimitadora.
 - Segmentação Semântica: Usada para segmentar diferentes regiões de uma imagem.
- Esses problemas podem ser encontrados em diversos campos de estudo:
 - MLOps industriais
 - Condução Autônoma
 - Robótica
 - Imagem médica
 - etc.

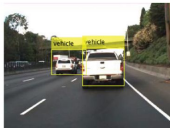
Ilustração de Aplicações



Reconhecimento de imagem



OBJECT DETECTION
ALGORITHM



Deteção de objetos



Segmentação de imagem