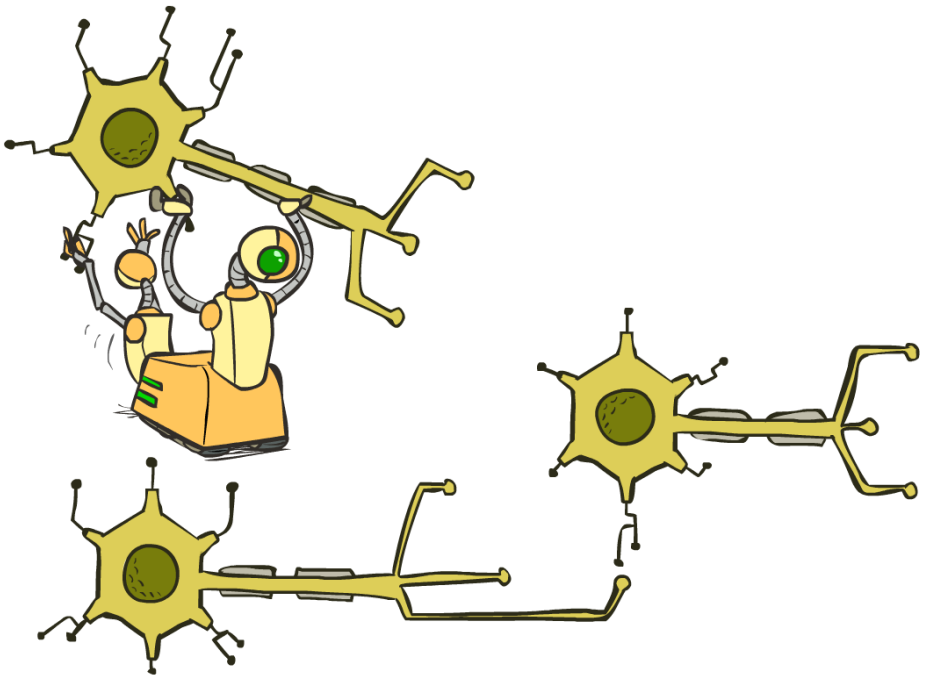
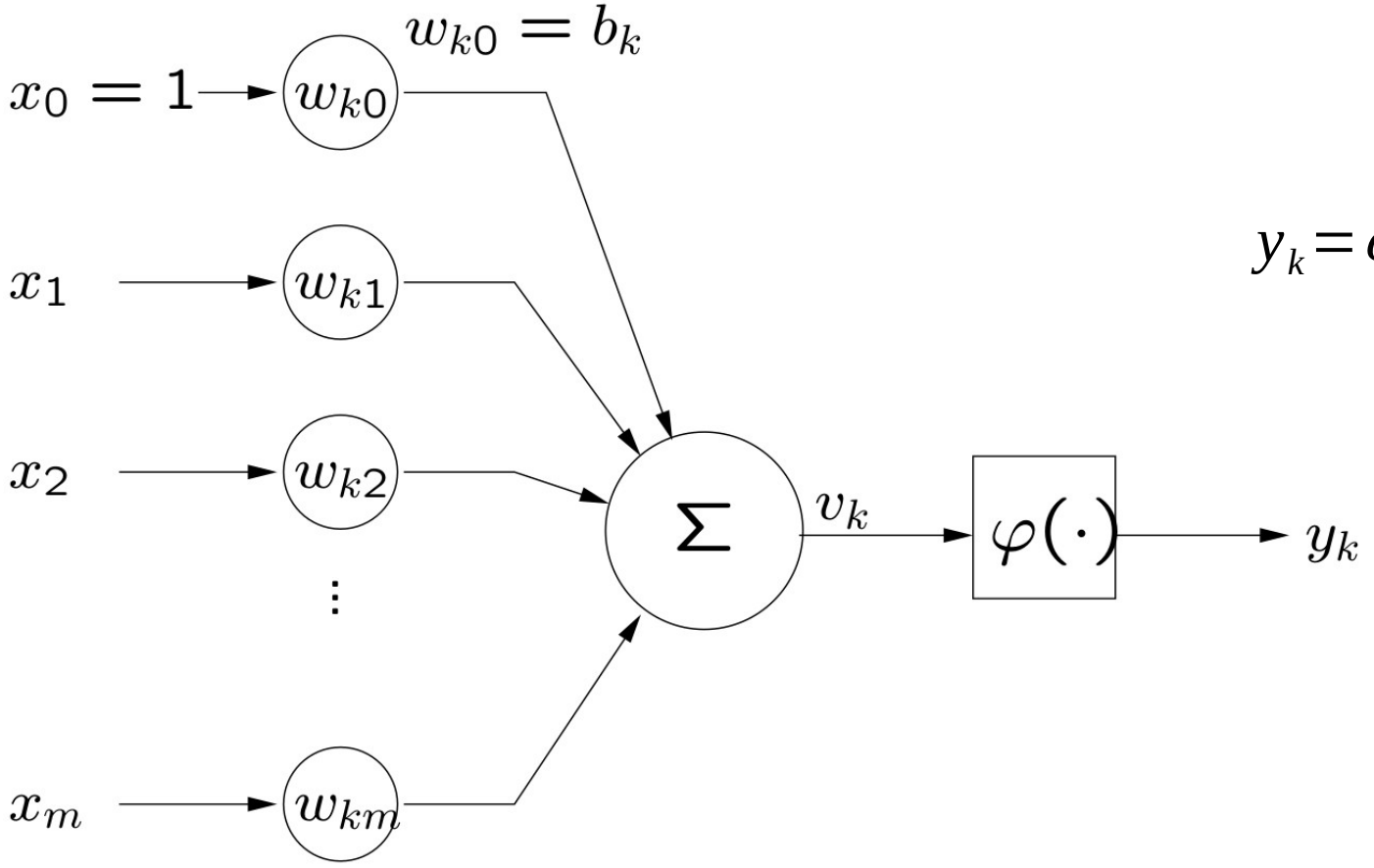
redes neuronais artificiais (RN | *NN*)



fonte: Berkley CS188

Ciências ULisboa

o modelo aditivo de neurónio



$$y_k = \varphi\left(\sum_{i=0}^{m} w_{ki}\, x_i\right)$$

Ciências ULisboa

caraterísticas

pesos das ligações (sináticas) são os locais onde o conhecimento é armazenado

- alteram-se ao longo do processo de aprendizagem

v é a soma ponderada das entradas

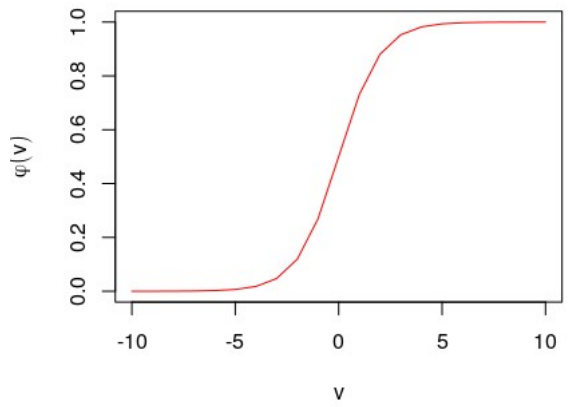
- cada entrada multiplicada pelo peso respetivo e todas somadas
- uma entrada especial, viés (***bias***) para estabelecer o nível quando nenhuma entrada ativa

Ciências ULisboa

função de ativação não linear

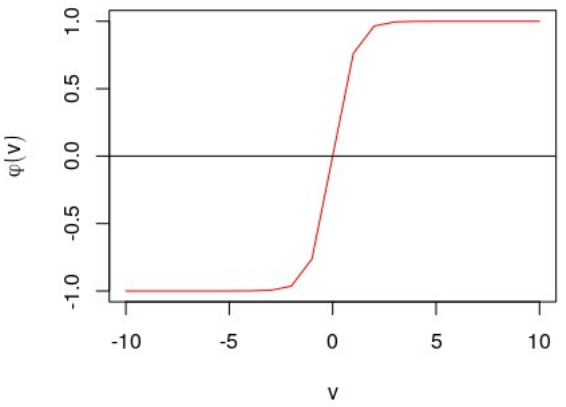
funções de ativação mais típicas

fundamental para que a RN não seja um simples modelo linear!



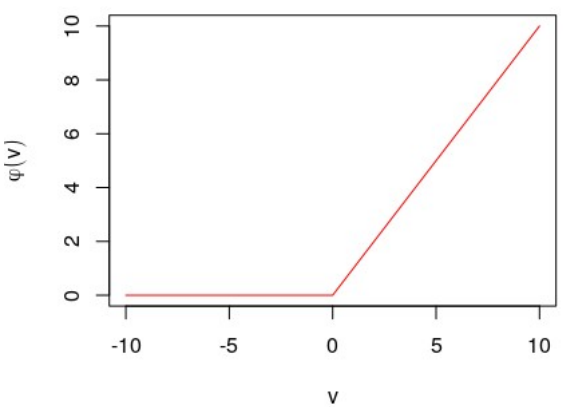
$$\varphi(v)=\frac{1}{1+e^{-av}}$$

logística



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

tangente hiperbólica



$$\varphi(v)=max(0,v)$$

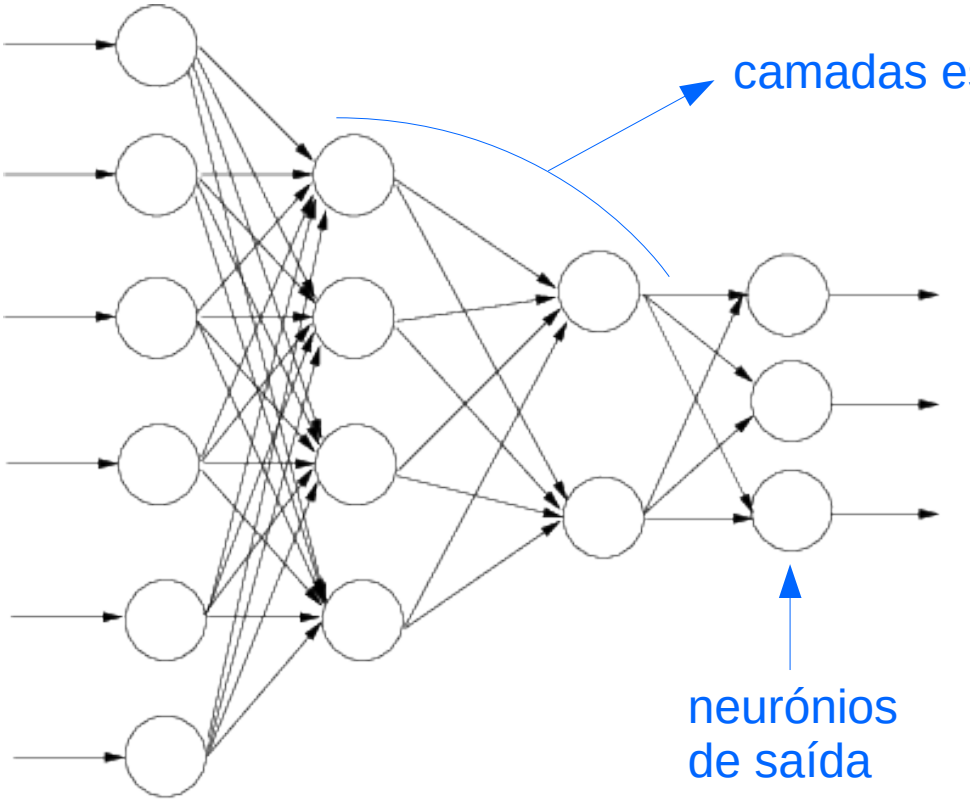
relu (rectified linear unit)

Ciências ULisboa

redes de alimentação progressiva

percetrão multi-camada

***multi-layer perceptron* (*MLP*)**



Ciências ULisboa

MLP

é a designação habitual de uma RN progressiva multi-camada

- **com menos saídas do que entradas**

propagação progresssiva dos sinais, camada a camada

- da entrada para a saída

tem 1 ou mais camadas escondidas

- tipicamente de ligação completa – todos os nós do nível *i* estão ligados aos neurónios do nível *i*+1

pode usar diversos métodos de **aprendizagem supervisionada**

- baseada em dados

Ciências ULisboa

MLP aproximador universal

mostra-se que o MLP devidamente dimensionado é um

aproximador universal de funções

i.e.

com um nº adequado de neurónios escondidos o MLP pode aproximar, tão proximamente quanto se queira, qualquer função limitada e não constante

nº de neurónios escondidos…

começar por experimentar com 50% # de entrada + # de saída

Ciências ULisboa

aprendizagem do MLP

o algoritmo de aprendizagem mais comum é o

retropropagação do erro (*error* ***backpropagation***)

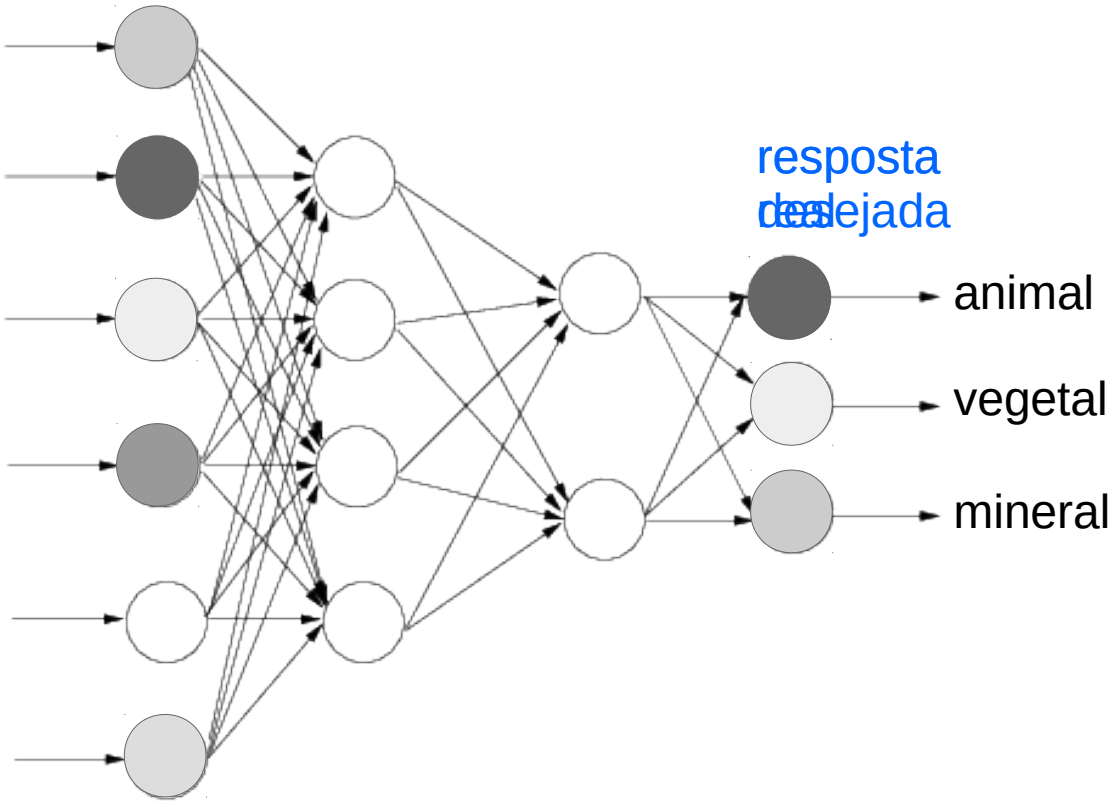
é um algoritmo supervisionado baseado numa descida de gradiente dos pesos em relação ao erro

- o erro é calculado à saída do MLP

Ciências ULisboa

resposta do MLP - exemplo

após algumas épocas de aprendizagem



Ciências ULisboa

backprop em duas penadas

passo progressivo:

apresenta-se um exemplo à entrada e calcula-se a saída

por propagação adiante dos sinais, camada a camada

mantendo os pesos sináticos constantes

é um algoritmo computacionalmente eficiente

passo regressivo:

os pesos são ajustados a partir do sinal de erro à saída

erro = diferença entre a saída real e a desejada

o erro é propagado para trás ajustando os pesos, camada a camada

Ciências ULisboa

descida do gradiente

η: rácio de aprendizagem

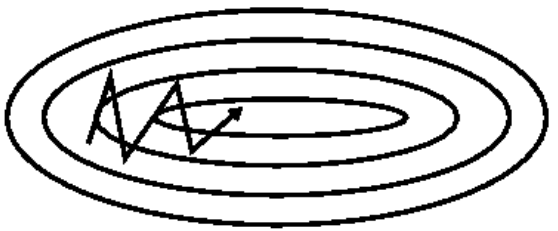
$E(w)$: erro ou função de custo

simples é mais lenta e pode ficar presa em mínimos locais

$$\Delta w_t = -\eta \nabla E(w)$$

com **inércia** (*momentum*)

$$\Delta w_t = \alpha \Delta w_{t-1} - \eta \nabla E(w)$$



- acelera com sinais consecutivos do gradiente idênticos
- estabiliza com sinais consecutivos do gradiente contrários

ADAM: versão eficiente (disponível no scikit-learn)

fonte: https://ruder.io/optimizing-gradient-descent/index.html#shufflingandcurriculumlearning

Ciências ULisboa

convergência

parâmetro de aprendizagem

se η <<

trajetória de aprendizagem suave, mas lenta

pode ficar preso em mínimos locais

se η >>

trajetória de aprendizagem rápida, mas instável

pode não convergir para o ótimo

Ciências ULisboa

aspetos importantes no backprop

baralhar os exemplos entre épocas (*shuffling*)

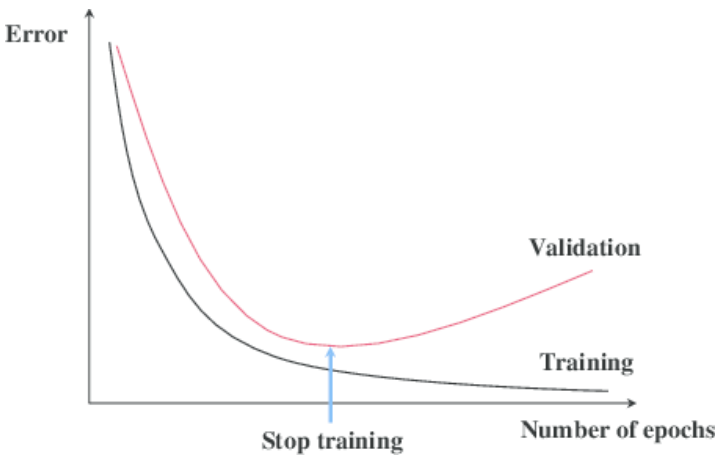
época = apresentação de todos os exemplos de treino

evita “vícios” no processo de aprendizagem por sequências de exemplos repetidas

paragem prematura

treino e validação

com paragem quando o erro de validação aumenta



fonte: http://fouryears.eu/2017/12/06/the-mystery-of-early-stopping/comment-page-1/

aprendizagem profunda
deep learning (*DL*)

em resumo: RN com muitas camadas

adequada para problemas complexos

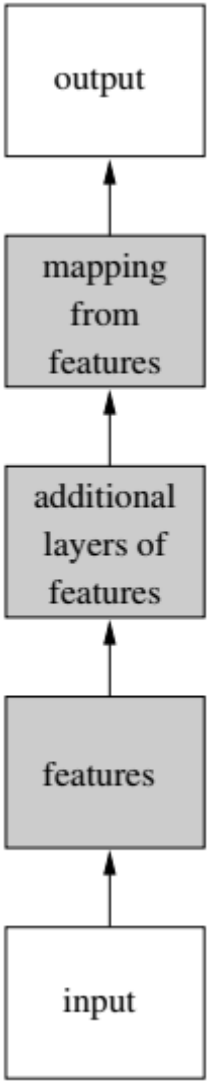
- com detalhes “decomponíveis”

aprende hierarquias de conceitos

- dos mais simples (menos abstratos) até
- aos mais complexos (mais abstratos)

usa muitos parâmetros ⇒ muitos exemplos

MLP – é componente essencial



aspetos relevantes

em MLP para DL

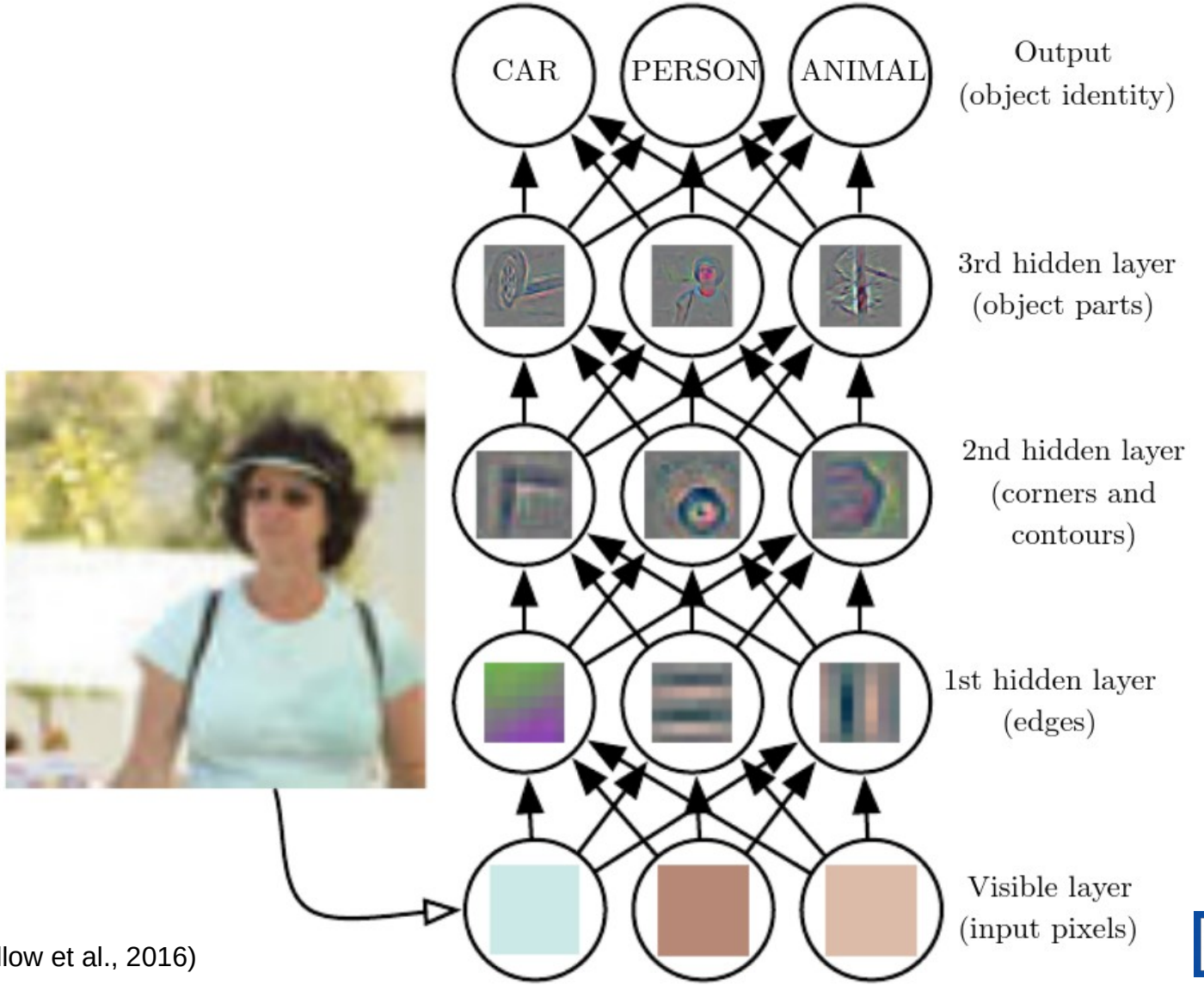
- função de custo: entropia cruzada
- função de ativação: retificada (ex. ReLU)

paralelização

- adaptáveis a GPUs – ganho de cerca de 1 ordem de grandeza em tempo de processamento

Ciências ULisboa

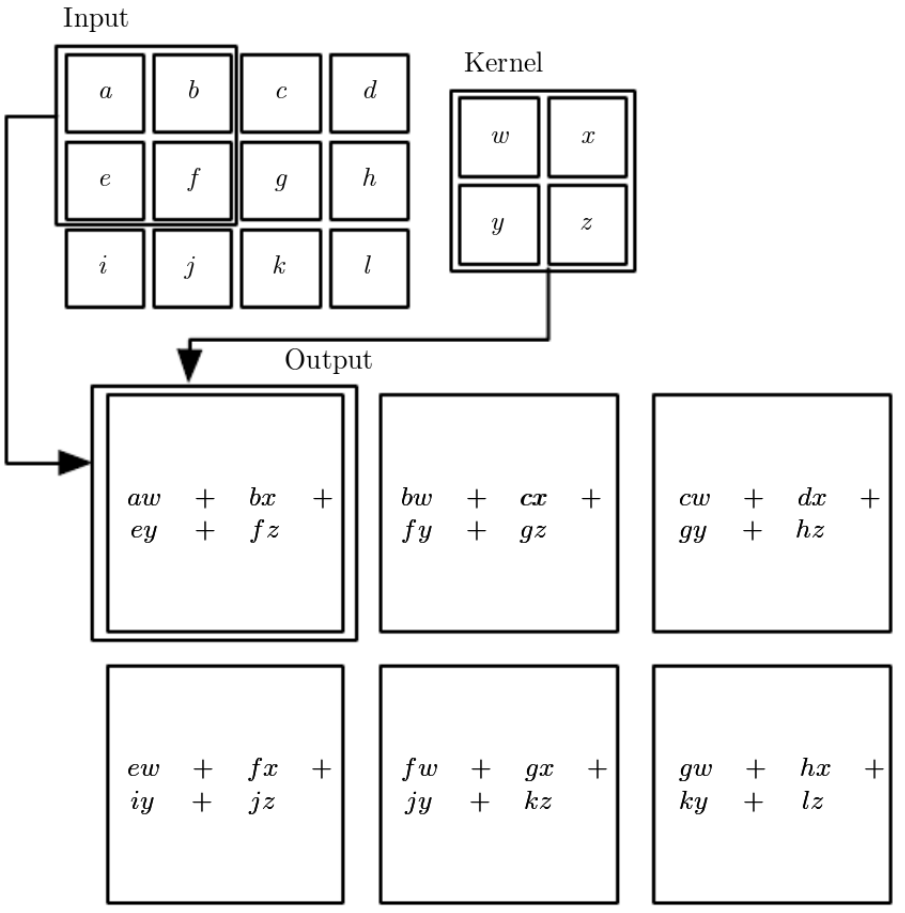
DL para imagens



fonte: (Goodfellow et al., 2016)

Ciências ULisboa

para imagens → redes convolucionais (profundas)



kernel é aplicado (multiplicado) sucessivamente a todo o input

extração de detalhes

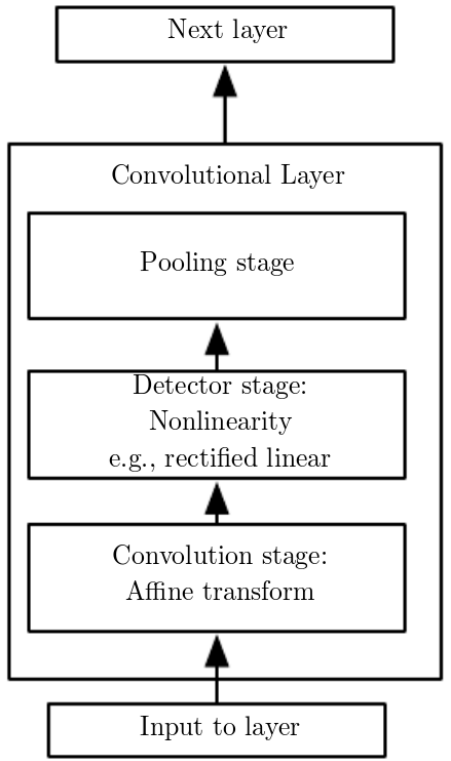
fonte: (Goodfellow et al., 2016)

Ciências ULisboa

CNN – *convolutional NN*

rede profunda

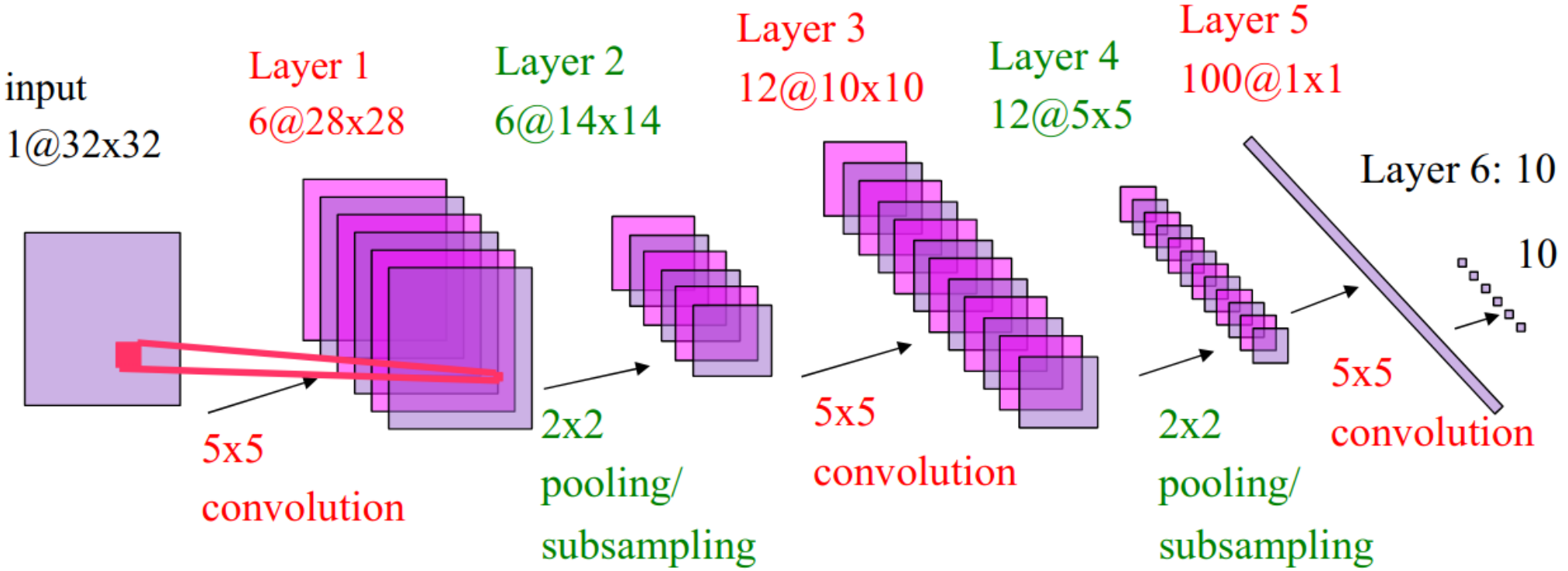
usa várias camadas convolucionais



"pooling" amostragem que aumenta invariância face a translações da entrada

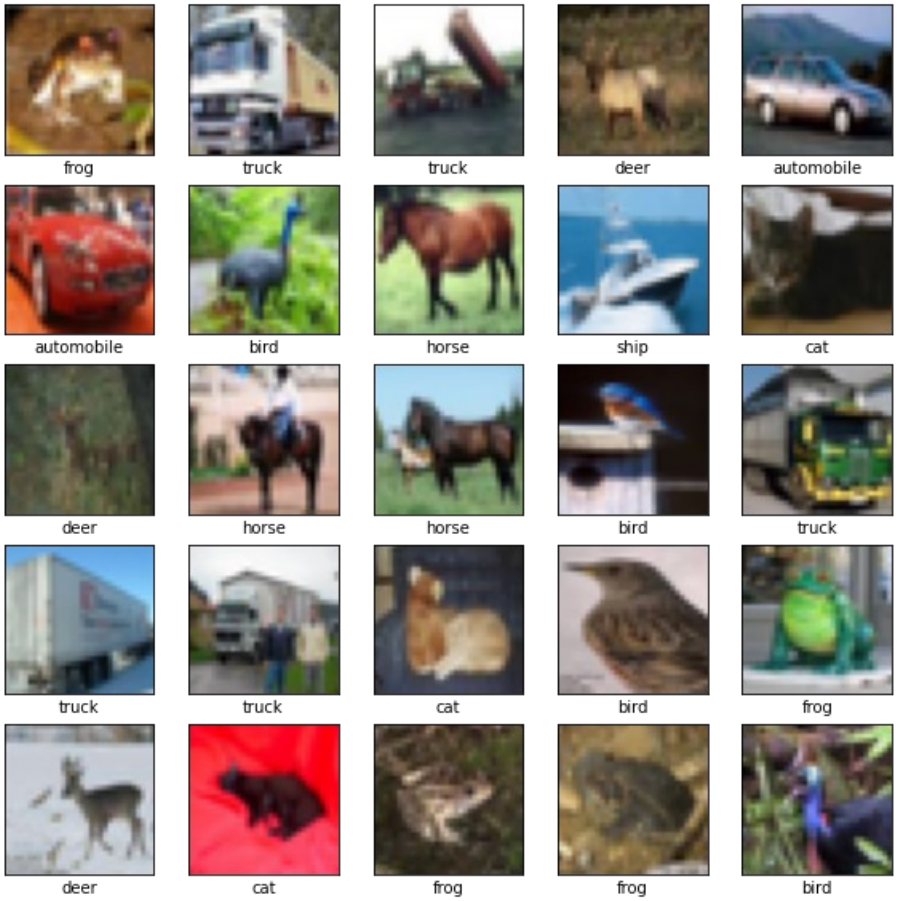
Ciências ULisboa

uma CNN – para o MNIST



Ciências ULisboa

ex. CIFAR10



10 classes: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, truck

60.000 exemplos – imagens a cores 32x32
50.000 de treino
10.000 de teste

6.000 exemplos por classe

Ciências ULisboa

ex. CNN para o CIFAR10

2 níveis convolucionais (convolução, mutualização)

1 nível de linearização *pooling*

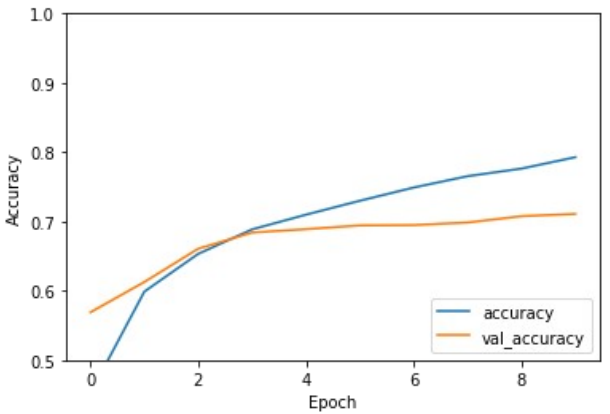
2 níveis MLP

---

total: 122.570 parâmetros (pesos)

10 épocas de aprendizagem

precisão: ~72%



fonte: https://www.tensorflow.org/tutorials/images/cnn

Ciências ULisboa

diversidade de outros modelos NN
(nem todos têm versão profunda)

mapas auto-organizados (não supervisionada)

- *clustering*

auto-codificadores (síntese de caraterísticas)

redes contraditórias generativas (*generative adversarial networks, GAN*)

redes com realimentação

- para problemas temporais e de análise de texto (*LSTM*)
- neurodinâmica – redes de Hopfield (memórias endereçáveis pelo conteúdo)

redes estocásticas – redes de Boltzmann

Ciências ULisboa

refª adicional (DL)

Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning*. Cambridge: MIT press. http://www.deeplearningbook.org

Ciências ULisboa