

Teste - Versão 1

29/11/2024

Aprendizagem Profunda

90 minutos

Nome: _____

ID: _____

Problema	Valores	Classificação
1	5	
2	2.5	
3	2.5	
4	5	
5	3	
6	2.5	
Total	20.5	

Problema 1 (Escolha Múltipla, 5 valores)

Circula a opção correta.

1.1 Em que situação o uso de dropout é mais recomendado?

- a) Durante a inferência para aumentar a precisão do modelo
- b) Durante o treino para reduzir o risco de overfitting
- c) Durante o pré-processamento para normalizar os dados
- d) Durante a validação para estimar o erro do modelo

Resposta correta: b

1.2 Qual das seguintes técnica é usada para mitigar o problema de desaparecimento do gradiente (vanishing gradients) em redes profundas?

- a) Funções de ativação sigmoide
- b) Inicialização Xavier ou He
- c) Aumento do número de épocas (epochs)
- d) Redução do tamanho do batch (batch size)

Resposta correta: b

1.3 Qual é o papel da camada de pooling numa rede neuronal convolucional?

- a) Reduzir a dimensionalidade espacial enquanto preserva informações importantes
- b) Aumentar a quantidade de parâmetros na rede
- c) Normalizar os dados de entrada
- d) Gerar uma previsão final baseada nos pesos da rede

Resposta correta: a

1.4 Porque é que as funções de ativação não-lineares são essenciais em redes neurais?

- a) Para permitir que a rede aprenda relações complexas entre os dados
- b) Para garantir que o output da rede seja normalizado
- c) Para reduzir o número de parâmetros da rede
- d) Para manter os gradientes constantes durante o treino

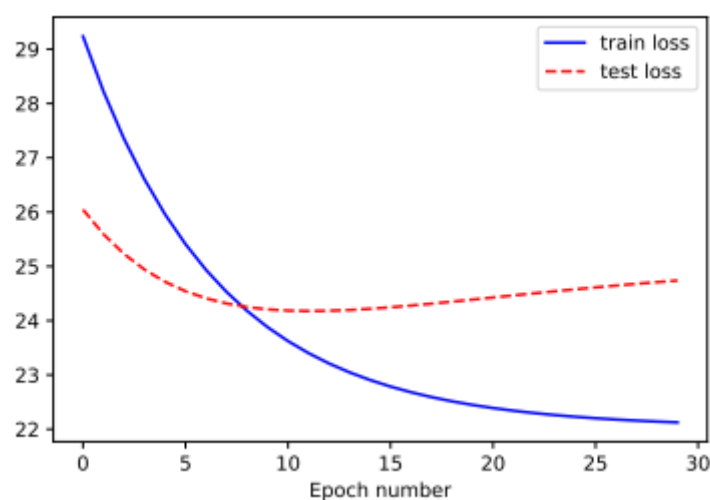
Resposta correta: a

1.5 Qual é a principal vantagem das redes neurais recorrentes (RNNs) em relação às redes feedforward (como MLPs)?

- a) Menor tempo de treino em grandes conjuntos de dados
- b) Capacidade de modelar dados sequenciais e com dependências temporais
- c) Melhor desempenho em dados de alta dimensionalidade, como imagens
- d) Maior facilidade na escolha de hiperparâmetros

Resposta correta: b

1.6 Um modelo foi treinado por 30 épocas (epochs) com gradiente descendente, resultando no seguinte gráfico de perdas (loss functions) de treino e teste:



Qual das seguintes afirmações é uma explicação plausível para o que pode estar a acontecer?

- a) O modelo está em underfitting.

- b) O modelo está em overfitting.
- c) O modelo generaliza bem para os exemplos de teste.
- d) Nenhuma das anteriores.

Resposta correta: b

1.7 Uma rede neural está a entrar em *overfitting*. Que estratégias podem ajudar a mitigar este problema?

- a) Aumentar a probabilidade de *dropout*.
- b) Diminuir a quantidade de dados de treino.
- c) Aumentar o número de unidades das camadas intermédias.
- d) Todas as anteriores.

Resposta correta: a

1.8) Supõe que uma camada de max pooling com um kernel de 2×2 e um stride de 1 recebe o seguinte input:

$$\mathbf{x}_{\text{in}} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}.$$

Qual será o output da camada de pooling?

- a) $\mathbf{x}_{\text{out}} = \begin{bmatrix} 3 & 4 \end{bmatrix}.$
- b) $\mathbf{x}_{\text{out}} = \begin{bmatrix} 1.5 & 2.5 \\ 4.5 & 5.5 \end{bmatrix}.$
- c) $\mathbf{x}_{\text{out}} = \begin{bmatrix} 5 & 6 \end{bmatrix}.$
- d) $\mathbf{x}_{\text{out}} = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}.$

Solução: c)

1.9 A técnica de early stopping monitoriza qual das seguintes losses?

- a) treino

- b) validação
- c) teste
- d) nenhuma das anteriores

Solução: b

1.10 Que tipo de rede tem loops de feedback?

- a) RNN
- b) CNN
- c) Perceptron
- d) nenhuma das anteriores

Solução: a

Problema 2 (Verdadeiro e Falso, 2.5 valores)

2.1 Redes neuronais feedforward são usadas principalmente para dados sequenciais devido à sua capacidade de capturar dependências temporais. _____

2.2 Redes neuronais recorrentes (RNNs) utilizam camadas de pooling para reduzir a dimensionalidade espacial dos dados. _____

2.3 Em redes neuronais convolucionais (CNNs), as conexões entre neurônios formam ciclos, permitindo a propagação de informações ao longo do tempo. _____

2.4 Redes neuronais convolucionais (CNNs) utilizam sempre funções de ativação sigmoide nas suas camadas convolucionais. _____

2.5 Redes feedforward precisam obrigatoriamente de funções de ativação não-lineares para produzir previsões.

Todas Falsas.

Problema 3 (Perceptron, 2.5 valores)

3) Suponha que lhe é dado o seguinte conjunto de dados:

x_1	x_2	y
0	0	0
0	1	0
1	0	0
1	1	1

Sabendo que queremos treinar um perceptron com os dados fornecidos e que:

- Os pesos iniciais foram aleatoriamente definidos: $w_1=0.9$ e $w_2=0.9$.
- O limiar de ativação (*activation threshold*) foi definido como $\theta=0.5$.
- O *learning rate* ficou definido como $\alpha=0.5$.

3.1) Qual será o vetor de pesos atualizado (w_1, w_2) depois de passarmos o exemplo 1 pelo algoritmo do perceptron? Apresenta todos os cálculos efetuados.

3.2) Qual será o vetor de pesos atualizado (w_1, w_2) depois de passarmos o exemplo 2 pelo algoritmo do perceptron? Apresenta todos os cálculos efetuados.

Problema 4 (Resposta curta, 5 valores)

4.1 Como é que a divisão de um conjunto de dados em conjuntos de treino, validação e teste pode ajudar a identificar overfitting?

Solução: Overfitting: o modelo ajusta-se tanto ao conjunto de treino que não generaliza bem. Erro baixo no treino e erro alto no conjunto de validação podem ser usados para identificar este problema.

4.2 Imagina que estás a projetar um sistema de aprendizagem profunda para detetar fadiga em condutores de carros. É crucial que o modelo detete a fadiga para prevenir acidentes. Qual das seguintes métricas de avaliação é a mais apropriada: Accuracy, Precision ou Recall? Explica a tua escolha.

Solução: Recall. É importante que não falhemos nenhum caso em que o condutor esteja cansado.

4.3 Explica porque precisamos de funções de ativação.

Solução: Sem não-linearidades, a nossa rede só consegue aprender funções lineares, porque a composição de funções lineares continua a ser linear.

4.4 Qual é o papel de um otimizador como o Adam ou SGD em redes neuronais?

Solução: Atualizar os pesos da rede durante o treino, minimizando a função de perda

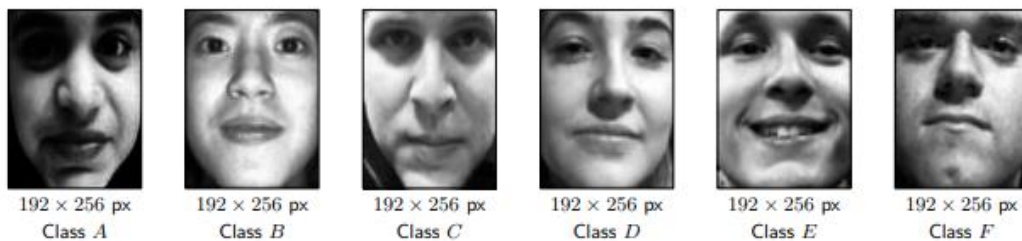
4.5 Dá um exemplo de quando se deve usar uma RNN em vez de uma rede neuronal feedforward?

Solução: Qualquer task que envolva dados sequenciais (por exemplo texto).

Problema 5 (CNNs, 3 valores)

5. Na sua loja física da Primark, os 6 funcionários utilizam atualmente um sistema de cartões para registrar os horários de entrada e saída. No entanto, os donos da loja ouviram falar dos avanços recentes em sistemas de visão por computador e decidiram substituir esse sistema por reconhecimento facial utilizando CNNs.

Para treinar o modelo, recolheram um grande conjunto de dados de fotografias dos seus 6 funcionários: Alice, Bruna, Carlos, Diana, Eduardo e Francisco. Cada fotografia no conjunto de dados é uma imagem em tons de cinza de 192×256 pixels, semelhante às ilustradas na figura abaixo, e está etiquetada de acordo com o funcionário correspondente (label).

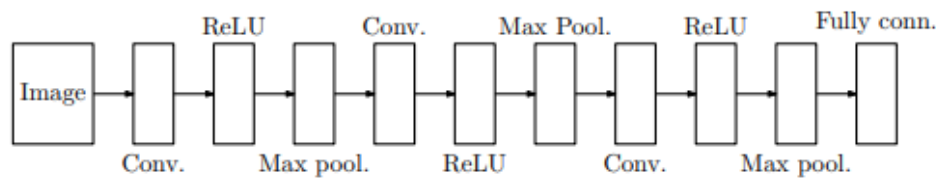


Imagens exemplo das 6 classes que a CNN deve reconhecer. Alice corresponde à classe A, Bruna à classe B, e assim por diante.

5.1 Explica brevemente, em 2-3 frases, por que razão uma CNN é uma escolha adequada de arquitetura para a tarefa descrita acima (classificação de imagens).

Solução: Solution: CNNs take advantage of the spacial structure of the image, unlike standard feedforward networks. Moreover, convolutional and pooling layers exploit the fact that the same feature may appear in different parts of the image, enabling the network to process those occurrences in a similarly way.

5.2 Supõe que foi usada a seguinte arquitetura para o problema mencionado:



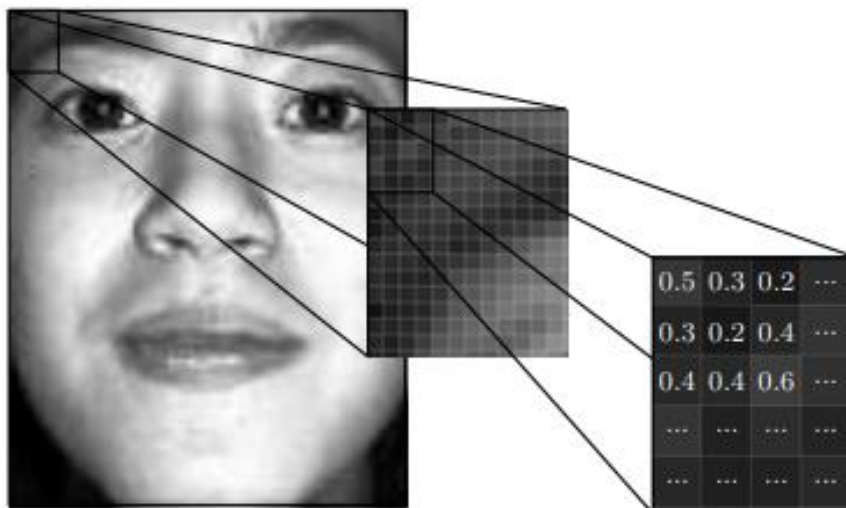
Com a seguinte implementação em PyTorch:

```
nn.Sequential(
    nn.Conv2d(1, 5, kernel_size=3, stride=1, padding=1),
    nn.ReLU(),
    nn.MaxPool2d(kernel_size=2, stride=2),
    nn.Conv2d(5, 10, kernel_size=5, stride=1, padding=0),
    nn.ReLU(),
    nn.MaxPool2d(kernel_size=2, stride=2),
    nn.Conv2d(10, 20, kernel_size=2, stride=2, padding=0),
    nn.ReLU(),
    nn.MaxPool2d(kernel_size=5, stride=2),
    nn.Flatten(),
    nn.Linear(2800, 6))
```

Preenche a seguinte tabela com os valores adequados:

Layer	Output size	N. weights	N. biases
Input	$192 \times 256 \times 1$	0	0
1st conv. layer	$192 \times 256 \times 5$	45	5
1st pooling layer	$96 \times 128 \times 5$	0	0
2nd conv. layer	$92 \times 124 \times 10$	1250	10
2nd pooling layer	$46 \times 62 \times 10$	0	0
3rd conv. layer	$23 \times 31 \times 20$	800	20
3rd pooling layer	$10 \times 14 \times 20$	0	0
Output layer	6×1	16,800	6

5.3 Considere o diagrama seguinte, que contém os valores para a primeira janela de pixels numa das imagens no conjunto de dados.



Supõe que, após o treino, um dos filtros na primeira camada convolucional é definido pelos seguintes parâmetros:

$$\mathbf{K} = \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} \quad b = 0.5.$$

Para o filtro fornecido, calcula o valor mais à esquerda no canto superior após a camada de pooling. Não esquecer de que a primeira camada convolucional inclui um padding de tamanho 1 (utiliza zeros como valor de padding).

Solution: The input of the max-pool layer corresponding to the provided pixels is given by

$$\begin{aligned} \mathbf{h}_{\text{conv}} &= \text{ReLU} \left(\begin{bmatrix} -0.5 & 0.2 \\ -0.9 & 0.0 \end{bmatrix} + 0.5 \right) \\ &= \begin{bmatrix} 0.0 & 0.7 \\ 0.0 & 0.5 \end{bmatrix}. \end{aligned}$$

At the output of the pooling layer, we thus have

$$h_{\text{pool}} = 0.7.$$

Problema 6 (RNNs, 2.5 valores)

6.1 Associa o tipo de tarefa com os exemplos correspondentes:

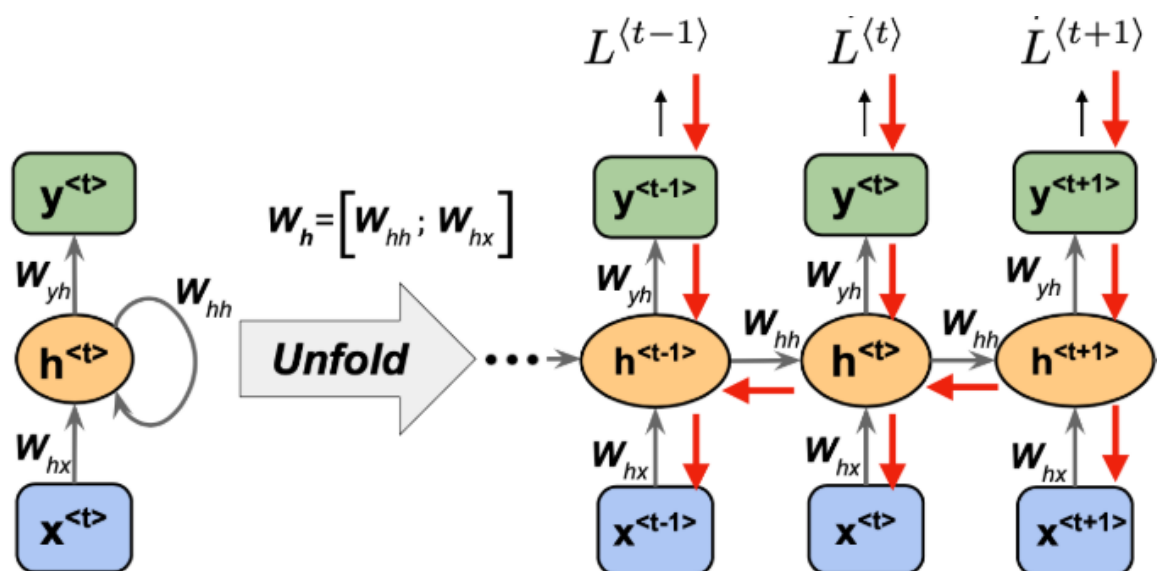
Exemplo de Tarefa	Tipo de Correspondência
Tradução de texto.	c
Análise de sentimentos.	b

Geração de música.	a
Legendagem automática de imagens.	a
Geração de texto.	a/c
Deteção de spam.	b
Sumarização de texto.	c
Chatbots.	c

Tipos de correspondência:

- A) One-to-Many
- B) Many-to-One
- C) Many-to-Many

6.2 Considere a figura seguinte:



6.2.1 Quantos conjuntos distintos de pesos estão representados na imagem?

6.2.2 Como a loss final é calculada na rede representada na imagem?

6.2.3 Considerando que a tarefa a ser realizada é tradução de texto (seq2seq), qual função de ativação seria mais adequada para a última camada e qual função de perda (loss function) deve ser utilizada? Justifica.

6.2.3 Considerando que a tarefa a ser realizada é tradução de texto (seq2seq), em que informação te basearias para definir o número de unidades da última camada?