

**Iniciado em** sexta, 16 out 2020, 18:42

**Estado** Finalizada

**Concluída em** sexta, 16 out 2020, 19:54

**Tempo empregado** 1 hora 12 minutos

**Avaliar** 8,00 de um máximo de 10,00(80%)

Questão 1

Correto

Atingiu 2,00 de 2,00

Quais das alternativas abaixo compõe os pontos necessários na prática para o projeto viável de treinamento de redes profundas?

Escolha uma:

- ☒ a. Conhecer os valores possíveis da função de perda utilizada, buscando por taxas de aprendizado e batchsizes adequados ao otimizador utilizando pequenos conjuntos de treinamento e validação. ✓
- ☐ b. Inicializar sempre os pesos com valores aleatórios é suficiente para garantir a viabilidade do treinamento.
- ☐ c. Utilizar o maior número de instancias disponível para executar experimentos exaustivos com relação a função de custo, taxa de aprendizado, outros hiperparâmetros e escolhas como: momentum, decaimento de taxa de aprendizado, tamanho do batch, em como o otimizador utilizado.
- ☐ d. Utilizar os valores padrão disponíveis nos pacotes de software para aprendizado de máquina, investigando diferentes arquiteturas de rede neural populares na literatura, até encontrar aquela que gera o melhor resultado no conjunto de testes.

Sua resposta está correta.

A resposta correta é: Conhecer os valores possíveis da função de perda utilizada, buscando por taxas de aprendizado e batchsizes adequados ao otimizador utilizando pequenos conjuntos de treinamento e validação..

Questão 2

Incorreto

Atingiu 0,00 de 2,00

Funções de perda diferentes geram intervalos de valores diferentes para um mesmo problema. Qual a consequência de uma função escolhida gerar valores muito pequenos durante o treinamento?

Escolha uma:

- ☐ a. Podem ocorrer problemas de overfitting devido à convergência do modelo acontecer muito rapidamente.
- ☐ b. Problemas de precisão numérica e relacionados à baixa magnitude do gradiente.
- ☒ c. Problemas relacionados à alta magnitude do gradiente e baixa generalização do modelo. ✗
- ☐ d. Facilita o treinamento gerando maior generalização do modelo final.

Sua resposta está incorreta.

A resposta correta é: Problemas de precisão numérica e relacionados à baixa magnitude do gradiente..

Questão **3**

Correto

Atingiu 2,00 de 2,00

Considere um problema cuja saída desejada seja um valor contínuo entre -20 e 10. Considerando que não há possibilidade de normalizar ou transformar esse intervalo, qual par função de ativação e qual função de perda são as mais adequadas para resolver esse problema?

Escolha uma:

- ☒ a. Ativação Linear e Perda pelo Erro Absoluto Médio. ✓
- ☐ b. Ativação LeakyRelu e Perda Entropia Cruzada.
- ☐ c. Ativação Relu e Perda Entropia Cruzada.
- ☐ d. Ativação Tangente Hiperbólica e Perda Quadrática.

Sua resposta está correta.

A resposta correta é: Ativação Linear e Perda pelo Erro Absoluto Médio..

Questão **4**

Correto

Atingiu 2,00 de 2,00

Carregue a base de dados Fashion-MNIST e projete uma rede do tipo ResNetInception utilizando os módulos Inception e Residuais conforme vistos em aula. A arquitetura deve ter, nessa ordem:

- Bloco Residual com 64 filtros
- Maxpooling com pool=2, stride=2
- Módulo Inception V1 com número de filtros: 32, 64, 64, 64, 16
- Maxpooling com pool=2, stride=2
- GlobalAveragePooling2D

Além disso utilizar: otimizador Adam, com batchsize 32, e 5 épocas.

Investigue os seguintes valores de taxa de aprendizado: 0.001, 0.005, 0.01 e 0.05, todos com decaimento exponencial com valor 0.05

Treine por 5 épocas, utilizando as primeiras 1000 imagens (:1000) para treinamento e as próximas 1000 para validação (1000:2000).

Qual taxa de aprendizado teve maior acurácia na validação?

Escolha uma:

- ☐ a. 0.005
- ☐ b. 0.05
- ☒ c. 0.01 ✓
- ☐ d. 0.001

Sua resposta está correta.

As respostas corretas são: 0.001, 0.005, 0.01, 0.05

Considere as situações abaixo listadas, observadas de forma independente após treinar modelos de rede profunda por 50 épocas para um problema de classificação com 20 classes usando a perda quadrática:

I - O valor da perda convergiu para um valor próximo a zero

II - O valor da perda após a primeira época foi de 2.61

III - A acurácia final medida no treinamento foi de 65% e na validação de 68%

IV - O valor da perda caiu de 0.90 na primeira época para 0.72 na última época

OBS: para interpretar os valores acima, calcule a perda quadrática e a acurácia para os seguintes cenários de vetores de probabilidade de saída (para 20 classes): todas as classes equiprováveis (aleatório uniformemente distribuído), e de um vetor de saída com a classe correta, porém com baixa probabilidade, com 0.2 para a classe correta, mas que possui outros 8 valores com 0.1.

Quais das situações acima estão tipicamente ligadas a problemas no cálculo do gradiente (causado muitas vezes por escolhas erradas na arquitetura e função e custo) e/ou na convergência do modelo causados por escolhas erradas do otimizador e seus hiper-parâmetros?

Escolha uma:

- ☐ a. Todos
- ☒ b. II e IV ✓
- ☐ c. I e II
- ☐ d. III e IV

Sua resposta está correta.

A resposta correta é: II e IV.

[◀ Exercícios - soluções](#)

Seguir para...



[Avaliação ▶](#)