

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CAMPUS FLORIANÓPOLIS

Trabalho Prático 1 - Deep Learning

Alunos Guilherme dos Santos (202402018)

Professor Prof. Dr. Eric Aislan Antonelo

Florianópolis, 24 de Abril de 2024

Conteúdo

1	Questão 1	1
2	Questão 2	2
3	Questão 3	3
4	Questão 4	4
5	Questão 5	5
6	Questão 6	6

1 Questão 1

2 Questão 2

As redes neurais convolucionais (CNNs) são mais eficientes do que as redes multicamadas totalmente conectadas em vários problemas, especialmente no processamento de imagens e visão computacional, além de outros problemas que envolvem dados em grade, como séries temporais. A eficiência das CNNs se deve à sua capacidade de processar dados de forma estruturada. Por exemplo, uma imagem pode ser representada por uma matriz de dimensões $N \times M$ e J camadas. As CNNs são capazes de realizar operações convolutivas nesses dados, o que permite a extração de características relevantes de forma mais eficaz. Além disso, essa arquitetura é mais adequada para capturar padrões locais e invariantes a deslocamentos, características essenciais em problemas de processamento de imagens e séries temporais. Em comparação, as redes totalmente conectadas exigem uma quantidade muito maior de parâmetros para lidar com a estrutura complexa dos dados em grade, tornando-as menos eficientes e mais suscetíveis ao overfitting.

3 Questão 3

Sim, há um erro no algoritmo. As redes neurais com retropropagação de erro atualizam os pesos ao final do processo de treinamento. Esse processo consiste em etapas sequenciais: primeiro, a passagem direta (forward pass), em que os cálculos são feitos em cada camada com os pesos atuais; em seguida, o cálculo do erro, que é a diferença entre a saída prevista e a desejada; depois, a passagem retroativa (backward pass), na qual o gradiente do erro é calculado retroativamente da camada de saída para as camadas ocultas; e, por fim, a atualização dos pesos, na qual os pesos em cada camada são ajustados com base no gradiente calculado, utilizando um algoritmo de otimização. Se os pesos da camada $n - 1$ estão sendo atualizados conjuntamente com os pesos da camada n durante o processo de treinamento, isso indica um desvio do procedimento padrão e pode comprometer a convergência do algoritmo ou levar a resultados imprecisos.

4 Questão 4

O pequeno incremento na saída da hipótese após aplicar o descenso do gradiente pode ser esperado dentro do contexto do método iterativo do descenso do gradiente. Esse método consiste em ajustar os parâmetros da hipótese com múltiplas iterações. O incremento pode indicar que o algoritmo está se aproximando de um mínimo local ou que a taxa de aprendizado é muito pequena, o que pode resultar em uma convergência mais lenta para o ponto desejado ou em uma estagnação prematura. Assim, esse resultado não necessariamente indica um erro de implementação, mas sim um comportamento típico do processo de otimização.

5 Questão 5

A remoção de um peso da rede neural pode resultar na redução do overfitting, o que possibilitaria a obtenção de um modelo mais generalizado. Este modelo simplificado teria a capacidade de lidar com uma variedade mais ampla de sinais, o que poderia explicar o aumento no desempenho.

6 Questão 6