



Aprendizagem Profunda – 2024

Prof. Dr. Saulo Oliveira

Nome: _____

Matrícula: _____ Data: _____

A pontuação da prova é em função de escores. Não existe meio escore. Expresse bem a sua resposta. Respostas secas/curtas sem muito rigor e detalhes de descrição podem prejudicá-los no acúmulo de escores.

Questão 1. (3 escores)

Cite e caracterize as variantes do Gradiente Descendente.

Questão 2. (5 escores)

Você pode representar a seguinte função booleana com um Perceptron simples? Se sim, mostre os pesos e desenhe a rede. Se não, explique por que não em 1-2 frases.

| A | B | $f(A, B)$ |
|---|---|-----------|
| 1 | 1 | 0 |
| 0 | 0 | 0 |
| 1 | 0 | 1 |
| 0 | 1 | 0 |

Questão 3. (3 escores)

Suponha que tenhamos um conjunto de dados de pacientes que visitaram o Hospital Geral de Fortaleza durante o ano de 2023. Um conjunto de características (por exemplo, temperatura, altura) também foi extraído para cada paciente. Nosso objetivo é detectar se um novo paciente visitante tem Zica, Dengue ou Chikungunya (um paciente pode ter um ou mais destas doenças). Pois bem, foi decidido que usaremos uma MLP para resolver este problema. Temos duas opções:

- I. treinar uma MLP separada para cada uma das doenças; ou
- II. para treinar uma única MLP com um neurônio de saída para cada doença, mas com uma camada oculta compartilhada.

Qual método você prefere? Justifique sua resposta.



Questão 4. (2 escores)

Por que as duas funções a seguir não são úteis nas camadas ocultas de uma MLP com retropropagação? Por quê?

a) $f(x) = x$ b) $f(x) = \text{sign}(x) = \begin{cases} +1 & \text{se } x \geq 0; \\ -1 & \text{caso contrário.} \end{cases}$

Questão 5. (3 escores)

Nesta questão, você implementará o algoritmo do AdaGrad para otimizar uma função de custo simples, preenchendo uma tabela de valores, até se alcançar o mínimo da função. Você começará com o valor para o parâmetro $w^{(0)} = -2$ e $\eta = 0,5$, determinará o valor da função de custo e atualizará os parâmetros usando a informação do gradiente.

$$\text{ADAGRAD: } w^{(t+1)} = w^{(t)} - \frac{\eta}{\sqrt{G^{(t)}}} g^{(t)}$$

$$\text{FUNÇÃO DE CUSTO: } \ell(w) = \frac{1}{2}(w - 4)^2$$

$$\text{DERIVADA DE } \ell(w): \nabla_w \ell(w) = w - 4$$

$w^{(t+1)}$ é o peso após o ajuste;

$w^{(t)}$ é o peso antes do ajuste;

η é a taxa de aprendizagem;

$g^{(t)}$ é o gradiente (derivada) da função de custo $\ell(w^{(t)})$;

$G^{(t)}$ é a soma dos quadrados dos subgradientes anteriores, isto é, $G^{(t)} = \sum_{i=1}^t (g^{(i)})^2$.

Adicionalmente, adote $g^{(0)} = 1$.

| | Inicial | Iteração 1 | Iteração 2 | Iteração 3 | Iteração 4 |
|-----------|---------|------------|------------|------------|------------|
| w | -2 | | | | |
| $\ell(w)$ | 18 | | | | |

Se garantam 🙏. Tô torcendo por vocês, daqui do meu lugar.