

Aula 2: derivada e gradiente descendente

Sexta, 02/10/2026. O que não couber hoje entra na aula 3.

1. Revisão da aula 1

de memória: o que o neurônio calcula; a regra quando erra; por que o E termina; e o XOR
vídeos: o E e a tabela das épocas; o XOR, 4 erros por época, a regra em círculo

2. Seminário

os vídeos: neurônio entre 0 e 1; custo = perda; o paradoxo da derivada
o perceptron de E e OU, linha a linha; uma pergunta de cada um

3. O neurônio sem o degrau e a perda

$y = w x + b$: ajustar o neurônio é ajustar uma reta
EQM: $L(w, b) = (1/N) \sum (y_i - t_i)^2$; L de loss; o E é a função lógica
vídeo: a reta gira com w, e a perda desenha uma curva com um fundo
para que lado mexer w, e quanto?

4. A derivada

vídeo: a secante vira tangente; de perto, a curva vira reta; $f'(x)$ como limite
 x^2 pela definição: $2x + h \rightarrow 2x$; tabela 3; 2,5; 2,1; 2,01
três regras: constante, $c \cdot w$, soma
 $L(w) = (w - 2)^2$: o passo tem o sinal contrário ao da derivada

5. Derivada parcial e gradiente

vídeo: dois cortes na superfície; $\partial L / \partial w$ e $\partial L / \partial b$; conferir pela derivada numérica
 ∇L aponta para onde L mais cresce; $\Delta L \approx \nabla L \cdot \text{passo}$; $\vec{w}$ da aula 1
gradiente e degrau: a mesma raiz, gradus, passo

6. Gradiente descendente e a taxa de aprendizado

vídeo 3D: $(w, b) \leftarrow (w, b) - \eta \nabla L$; a tigela e, ao lado, a reta chegando aos pontos
vídeo da taxa: fator $1 - 2\eta$; 0,1 devagar, 0,9 oscila, 1,05 diverge
o limiar da reta: 0,513; o limiar do projeto de dezembro

7. De onde vem

Cauchy, 18/10/1847: órbitas; faire décroître la fonction u
Widrow e Hoff, 1960: a Adaline, a lancheira, 17 botões; a regra com z

8. O código

no laptop: pontos com semente, perda, gradiente, laço, várias taxas
a entrega: o gráfico da perda por passo e a taxa que diverge

9. Fechamento

duas semanas: 3Blue1Brown, álgebra linear 3 e 4; Nielsen, capítulo 1
entrega até 16/10; seminário da aula 3: matrizes como transformações