1 a) No escala con el tamaño de redes neuronales. de propagación y gran cantidad de datos, 10/10

2 Supervisado 10/10

Este tipo de aprendizaje tiene como característica principal el uso de datos que se encuentran etiquetados para el entrenamiento

No-Supervisado

Este tipo de aprendizaje no necesita ningun tipo de etiquetado en la etapa de entrenamiento, ya que ~~sea~~ a partir de los datos va creando sus propios resultado

3 Sen(x) $-x^2$ 10/10

Para encontrar un punto critico debemos derivar la funcion e igualarla a 0 eso significa que la pendiente de la tangente se encuentra horizontal. La derivada del sen es cos(x)

Buscamos la x que de como resultado 0.

y la remplazamos en la funcion original

~~sen~~ cos(90) = 0 y el sen(90) = 1, asi hallamos un maximo local

Para saber si es máximo o mínimo requerimos de la 2da derivada

4 ~~Prox~~

10/20

Sustitumas x por $x_0 - \epsilon g$

$F(x_0-\epsilon g) = f(x_0) + ((x_0 - \epsilon g) - x_0)^T g + \frac{1}{2}((x_0-\epsilon g) - x_0)^T H((x_0-\epsilon g)-x_0)$

se simplfica

$F(x_0) - \epsilon g^T g + \frac{1}{2}\epsilon^2 g^T H g$

Derivamos a partir de ϵ y encotramos aquella que de 0

falto un signo e indicar que esto es la derivada

$g^T g + \epsilon g^T \cdot H \cdot g = 0$

sin el signo correcto no puedes llegar a esto

Despejamos ϵ $\qquad \epsilon = \frac{g^T}{g^T H g}$

5 $\quad ax^2 + y(ay - bx)$

Derivamos $\quad \frac{d}{dx} ax^2 + y(ay - bx)$

$=$

$2ax - yb$

$\frac{d}{dy} ax^2 + y(ay - bx) = 2ay - bx$

$\Delta f = \begin{bmatrix} 2ax - yb \\ 2ay - bx \end{bmatrix} = \begin{bmatrix} .52x - .48b \\ .52y - .48b \end{bmatrix}$

0/15

20/20

Derivamos

$H = \begin{bmatrix} 2a & -b \\ -b & 2a \end{bmatrix} = \begin{bmatrix} .52 & -.48 \\ -.48 & .52 \end{bmatrix}$

5

Dada $0.26x^2 + 9(0.20 - 0.48x)$

1 Definimos un punto x_0

2 Iteramos n veces lo siguiente: 15/15

3 Usamos el valor de x_0 para reemplazarla en el gradiente, eso nos da un vector

4 Ese vector lo multiplicamos por una tasa (ϵ), que puede ser definida o una tasa cambiante definida por

$$e = \frac{g^T}{g^T H g}$$

5 ~~La tasa~~

Al valor x_0 le restamos el gradiente por la tasa, y ese valor es nuestro nuevo x_0

6 Inicia de nuevo la iteración