

Actividad 3

Fecha de entrega: Viernes 10 de septiembre, 23:59

Profesor: Pablo Estévez V.
Auxiliar: Ignacio Reyes J.
Semestre: Primavera 2021

Instrucciones generales

- La tarea es de carácter **individual**. No está permitido compartir desarrollos matemáticos ni líneas de código. Sí se puede conversar con otros estudiantes respecto a cómo resolver los problemas.
- La entrega final debe corresponder a un solo PDF. Además de éste, debe entregar los códigos de Python respectivos para ser testeados por el equipo docente.

Considere una unidad con un peso w , un bias b , y una función de activación sigmoide logística como se indica en la figura. La entrada y la salida son escalares.

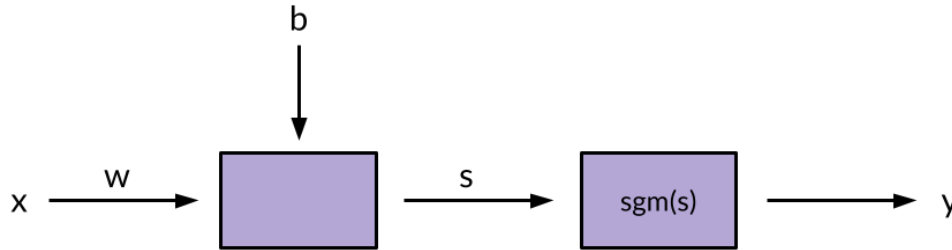


Figura 1: Modelo.

Nota: $y = \text{sgm}(s) = \frac{1}{1+\exp(-s)}$, donde $s = w \cdot x + b$.

Parte a

Derive una regla de aprendizaje por gradiente descendente para el ajuste del peso w y del bias b .

1. usando el error cuadrático $E = (t - y)^2$.
2. usando la entropía cruzada $E = -t \ln(y) - (1 - t) \ln(1 - y)$.

Parte b

Implemente el entrenamiento del modelo anterior en Python usando las librerías Numpy y Matplotlib (sin Tensorflow). Para ello debe construir un dataset con dos clases:

- Clase A: 100 muestras obtenidas de una distribución $\mathcal{N}(-1, 1)$. Salida deseada $t = 0$.
- Clase B: 100 muestras obtenidas de una distribución $\mathcal{N}(+1, 1)$. Salida deseada $t = 1$

Utilice la regla de aprendizaje asociada a la entropía cruzada, derivada en la parte anterior. Nota: La regla de la parte anterior se calculó sobre un solo ejemplo. Para ajustar los parámetros con varias observaciones, calcule el ajuste asociado a cada uno de los ejemplos y tome el promedio de las correcciones. Esto equivale a promediar la función de costos sobre todos los ejemplos.

Use como parámetros iniciales $w = -0,1$ y $b = 0,4$. Use una tasa de aprendizaje que entregue un entrenamiento estable pero no demasiado lento. Itere 100 veces.

Grafique la evolución de la función de costos durante el entrenamiento (curva de aprendizaje). Grafique también la evolución de los parámetros durante el entrenamiento mediante un scatter 2D. Por último grafique el modelo (recordar que se trata de una función de \mathbb{R} en \mathbb{R}).

Analice lo obtenido. ¿Por qué w no se va a infinito? Imagine qué efecto tendría esto sobre la función de costos en el caso de una muestra mal clasificada.