

MDS7104 Aprendizaje de Máquinas**Profesor:** Francisco Vásquez L.**Auxiliares:** Catalina Lizana G., Álvaro Márquez S., Diego Olguín W.**Fecha:** 19 de Noviembre de 2024.

Auxiliar 13: XGBoost y Deep Learning

P1. Considere el *dataset de hongos de UCI*, este contiene diferentes *features* e indica si un hongo es venenoso o no (el *target*). El objetivo es implementar XGBoost para poder predecir, en base a los *features* del *dataset* si un hongo es venenoso o no.

- Implemente XGBoost para la tarea de clasificación, estudie métricas en un conjunto de entrenamiento y otro de clasificación ¿cómo cambian los resultados de acuerdo a los parámetros elegidos?
- Realice una búsqueda de grilla para decidir los mejores parámetros a utilizar. ¿Qué interpretación tienen los parámetros obtenidos?
- Estudie con el modelo cuáles atributos tienen mayor importancia a la hora de predecir la toxicidad de un hongo.

P2. Considere una neurona sigmoide con una entrada unidimensional x , peso w , sesgo b , y salida $y = \sigma(wx + b)$. El objetivo es la variable unidimensional z . Considere la función de costo $C(w, b) = \frac{1}{2}(y - z)^2$;

- Encuentre $\nabla C(w, b)$ y muestre que

$$\|\nabla C\| \leq \frac{1}{4} \sqrt{1 + x^2} (1 + z^2)$$

- Escriba la iteración de descenso por gradiente para la secuencia de pesos (w_n, b_n) .
- Considere una red neuronal con entrada de dimensión 2, una sola capa oculta con 2 neuronas y salida de dimensión 1, con función de activación sigmoide. Escriba *backpropagation* para el caso de esta red.

P3. [Sobre el aprendizaje exacto]

- Considere una red neuronal de una sola capa oculta con función de activación ReLU $\mathcal{N} : \mathbb{R} \rightarrow \mathbb{R}$. Pruebe, indicando el ancho necesario N de la red, que esta puede aprender de manera exacta la función identidad, esto es $\mathcal{N}(x) = x$.
- Considere una red neuronal de una sola capa oculta con función de activación ReLU $\mathcal{N} : \mathbb{R}^2 \rightarrow \mathbb{R}$. Pruebe, indicando el ancho necesario N de la red, que esta puede aprender de manera exacta
 - $f(x_1, x_2) = x_1$.
 - $f(x_1, x_2) = \max\{x_1, x_2\}$.
 - $f(x_1, x_2) = \min\{x_1, x_2\}$.
- Pruebe que no existe una red neuronal de una sola capa oculta con función de activación sigmoide que pueda aprender de manera exacta la función

$$f(x) = \begin{cases} 0, & \text{si } 0 \leq x \leq 0.5, \\ x - 0.5, & \text{si } 0.5 < x \leq 1. \end{cases}$$

¿Contradice esto el teorema de aproximación universal?