**Hugo Córdoba Leal**

**Sprint 2. Introducción al Deep learning**

**• Cuestión 1.** Se ha visto que existen diferentes métricas para evaluar un modelo de DL, ¿sería correcto emplear la misma métrica para un modelo de detección de cáncer de mama que para uno de clasificación de tipos de música? Justifica tu respuesta.

* No sería del todo correcto emplear la misma métrica. Para la elección de la métrica de evolución lo más optimo seria tener en cuenta los objetivos específicos y las características del problema que se está abordando. Para ello, se tendría que tener en cuenta la naturaleza del problema, la importancia de los errores.

1. Naturaleza del problema. Para el caso de detección de cancer de mama, se suele gestionar como un problema de clasificación binaria. Sin embargo, por parte del problema de clasificación de tipos de música, este problema conlleva la clasificación de múltiples clases.
2. Importancia de errores. En el primer caso, los errores con llevarían poder dar falsos testimonios a los pacientes, por lo que es crucial minimizar falsos negativos. Sin embargo, en el segundo caso, los errores no conllevan consecuencias tan importantes.

Por ello, la elección de la métrica de evaluación debe basarse en los requisitos específicos del modelo a evaluar.

• **Cuestión 2.** Los modelos de deep learning necesitan de grandes cantidades de datos para aprender, sin embargo, muchos de los datos que podemos encontrar no son de uso público. ¿Sería ético entrenar a tus modelos con estos datos?, ¿qué alternativas podrías encontrar?

* El uso de datos para entrenar modelos, se enfrenta a una consideración ética crucial. Mi reflexión hacia este tema es que se debe primar siempre el consentimiento y la privacidad. Si los datos no son de uso público y no existe consentimiento de los individuos hacia la utilización de sus datos, que estos sean utilizados para entrenar modelos, bajo mi punto de vista no sería ético o sería muy cuestionable.

Ante esta situación, se me ocurre la posible solución de usar conjuntos de datos generados sintéticamente. Se mantendría las características originales de los datos pero no estarían vinculados a ningún individuo específico. Se podrían mezclar datos entre sí para generar nuevos conjuntos para entrenar nuestros modelos.

**• Cuestión 3.** Cuando se trata de algoritmos de deep learning, la mayoría de las librerías tienen soporte para GPU. Es más, podemos encontrar un mayor rendimiento empleando GPU que CPU, ¿a qué se debe este comportamiento?

* La mejora de rendimiento al utilizar GPU para deep learning se debe, principalmente, a la arquitectura y capacidad de procesamiento paralelo que las GPU pueden ofrecer. Esto se debe a varias razones, entre ellas, arquitectura paralela (Las GPU están diseñadas para manejar grandes cantidades de operaciones en paralelo) o un mayor ancho de banda de memoria (Las GPU suelen tener un mayor ancho de banda de memoria en comparación con las CP.) o la capacidad de escalabilidad (Las GPU permiten escalar el rendimiento de manera eficiente al agregar múltiples unidades).