

Caso práctico | Modalidades y técnicas de Deep Learning

Introducción

RiseData no sólo se centra en mejorar el rendimiento deportivo de todos sus clientes, sino que también aporta su granito de arena en el entrenamiento y desarrollo de herramientas que permitan diagnósticos médicos más rápidos y precisos. Por ello, una de tus tareas es lograr una clasificación correcta de radiografías torácicas mediante el entrenamiento de redes neuronales convolucionales.

Para que puedas proceder con estas tareas, hay un trabajo previo:

- Se han seleccionado imágenes de rayos X de tórax (anterior y posterior) de cohortes retrospectivas de pacientes pediátricos. Estos pacientes son niños de 1 a 5 años que acudieron a sus citas de atención clínica rutinaria en el Centro Médico de Mujeres y Niños de Guangzhou, Guangzhou.
- Para poder pasar al análisis de estas imágenes de rayos X de tórax, todas las radiografías se revisaron inicialmente a modo de control de calidad. Aquellos escaneos que se consideraron de baja calidad o ilegibles se eliminaron.
- Dos médicos con amplia experiencia calificaron los diagnósticos de las imágenes antes de autorizarlos para entrenar el sistema de IA.
- Para tener en cuenta cualquier error de calificación por parte de los médicos, el conjunto de evaluación también fue verificado por un tercer experto.

Planteamiento del problema

Con todos los datos preparados, te han asignado el entrenamiento de una red CNN que sea capaz de clasificar correctamente las imágenes, con un acierto superior al 80%. La finalidad de este proceso es ayudar a los médicos cuando hay un gran volumen de radiografías y que puedan priorizar los casos más urgentes o que necesitan un mayor nivel de atención médica.

Desarrollo

El conjunto de datos, que puedes encontrar [aquí](#), está organizado de la siguiente manera:

- Tres carpetas (train, test y val) con sus correspondientes subcarpetas para cada categoría de imagen.
- En total hay 5.863 imágenes de rayos X en formato JPEG, divididas en dos categorías, neumonía y normal.

Ejercicio 1

Carga los datos y realiza una primera exploración. Para rematar, visualiza los jpg.

Ejercicio 2

Construye, compila y entrena una red neuronal convolucional (CNN) que clasifique correctamente las imágenes.

Ejercicio 3

Crea las curvas correspondientes para identificar si el modelo generaliza correctamente y ver si hay overfitting.

Ejercicio 4

Evalúa el modelo, haciendo la predicción sobre los jpg de la carpeta *val*.

Requisitos de la entrega

Debes entregar un notebook ejecutado que incluya:

- Las decisiones que vas tomando a lo largo del desarrollo del ejercicio.
- Las conclusiones a las que has llegado tras la realización del ejercicio.