



Modelos preentrenados

Modelos preentrenados

La utilización de modelos preentrenados en aprendizaje profundo es una estrategia valiosa que aprovecha el conocimiento adquirido por modelos entrenados en conjuntos de datos masivos y generalizados. Este enfoque, se considera una forma de transferencia de aprendizaje, donde un modelo previamente entrenado en una tarea puede adaptarse eficientemente a una tarea específica relacionada, incluso si los conjuntos de datos son diferentes.

Esta estrategia se asemeja a una optimización en dos fases:

Fase 1:

El modelo preentrenado ha pasado por un proceso de optimización en un conjunto de datos amplio y diverso, lo que le permite aprender patrones y representaciones generales. Este preentrenamiento actúa como una optimización genérica, capturando características fundamentales y complejas del dominio.



Fase 2:

Se realiza la optimización específica para la tarea de interés con un conjunto de datos más específico. Durante este ajuste fino o "fine-tuning", el modelo se adapta a las características particulares y sutilezas de los nuevos datos, permitiendo que el conocimiento previo se transfiera de manera efectiva. Dado que gran parte de la información valiosa ya ha sido capturada durante la fase de preentrenamiento, se requiere menos tiempo y datos para adaptar el modelo a la tarea específica.





Esta estrategia es especialmente útil cuando se enfrenta a conjuntos de datos pequeños o cuando el tiempo y los recursos de entrenamiento son limitados. Los modelos preentrenados, a menudo entrenados en hardware y datos masivos, sirven como bloques de construcción sólidos que aceleran el proceso de optimización específica, esto ayuda a:

La eficiencia del
entrenamiento

Apuntar a modelos
más robustos y
generalizables

Ambos elementos se benefician de la riqueza de la información aprendida durante el preentrenamiento.

