Aumentación de Datos,

Transferencia de Aprendizaje, Batch Normalization y Dropout.

Para esta práctica requerirá recursos computacionales mayores que con todos los modelos vistos. Recomendamos encarecidamente que utilice algún servicio online que provea GPUs para realizar las pruebas.

**Ejercicio 1.**  CIFAR10 con aumentación de datos.

Entrene el mejor modelo que haya obtenido de las prácticas anteriores en CIFAR10 utilizando varios tipos de aumentación de datos al mismo tiempo, incluyendo cambios de brillo y color, rotaciones, flips, shifts, y zoom.

Deberá entrenar el modelo durante más tiempo para que converja (posiblemente hasta 4 o 5 veces más).

**Ejercicio 2**. Batch Normalization (BN)

A partir del mejor modelo que haya encontrado en el ejercicio 1:

1. Utilice capas Batch Normalization (keras.layers.normalization.BatchNormalization) entre las capas de la red, excepto luego de la capa Flatten y las capas de MaxPooling (si las usa). Compare la cantidad de iteraciones necesarias para converger con respecto al modelo sin BN. En la mayoría de los casos, debería disminuir ligeramente. ¿Se mantuvo igual el accuracy? En algunos casos, puede que disminuya.
2. No está del todo claro si las capas de BN deben utilizarse antes o después de la función de activación. Separe las activaciones de las capas convolucionales y densas, y agregue las capas BN entre medio de estas. ¿Cambió significativamente el accuracy o el tiempo de entrenamiento?
3. Si logró una disminución de las iteraciones, pruebe quitar varias capas de BN de la red. En muchos casos puede mantenerse la mejora agregando BN en algunas capas clave.

**Ejercicio 3**. Dropout

La capa Dropout introduce un factor estocástico (ruido) durante el entrenamiento de la red, poniendo en 0 de forma aleatoria algunas activaciones o valores intermedios. De esta forma, se puede regularizar ligeramente a la red, ya que funciona como una forma de aumentación de datos que se genera dentro de la red. A partir del mejor modelo que haya encontrado en el ejercicio 1:

1. Utilice capas Dropout (keras.layers.Dropout) entre las capas Dense de la red. Modifique el valor de **p**, la probabilidad de poner en 0 un valor de salida, probando los valores 0.1, 0.5 y 0.9. ¿Cómo afectan al accuracy final obtenido en el conjunto de prueba?
2. En general, la capa Dropout no ayuda a las capas convolucionales, ya que al tener estas tan pocos parámetros (en comparación con las capas Dense), no requieren regularización. Verificar esto agregando Dropout a las capas convolucionales, reentrenando y verificando el accuracy del conjunto de prueba.

**Ejercicio 4**. Transferencia de Aprendizaje en CIFAR10.

Entrene una red para CIFAR10 utilizando transferencia de aprendizaje a partir de una red MobileNet pre-entrenada en ImageNet.

1. Congelando las capas ya entrenadas, pruebe reentrenar al menos dos tipos de capas finales para reemplazar las originales.
2. Con la mejor configuración de capas finales de a), repita el proceso pero ahora sin congelar los pesos de la red.
3. Compare el tiempo de entrenamiento de a), b) y un modelo que haya entrenado previamente y haya obtenido similar accuracy.