

## Actividad 5

En este ejemplo, se va a observar el modelo LSTM convolucional en una aplicación para la predicción del un fotograma. El proceso de predecir qué fotogramas de vídeo vienen después dada una serie de fotogramas anteriores. Es decir que se entrenará con videos de números para después hacer predicciones de video. De esta manera serán una “animaciones” predictoras.

Para este trabajo, se utilizará el conjunto de datos Moving MNIST.

Se descargará el conjunto de datos y luego se construirá y se procesarán previamente los conjuntos de formación y validación.

Para poder hacer la predicción del fotograma, se usará un modelo el cual para hacer que funcione se utilizará un fotograma anterior, el cual se llamará  $f_n$ , para de esta manera poder predecir un nuevo fotograma, llamado  $f_{(n + 1)}$ . Para poder lograr que el modelo realice y cree las predicciones, se tendrán que procesar los datos de tal manera que se va a "desplazar" las entradas y salidas, donde los datos de entrada son el marco  $x_n$ , datos los cuales se usarán para predecir el marco  $y_{(n + 1)}$

Los datos que se descargaron son secuencias de marcos, cada uno de los cuales se utiliza para predecir el próximo marco. Se imprimen algunos marcos de estos para poder ver qué es con lo que se trabajará.

Para crear y construir el modelo LSTM convolucional se usarán capas ConvLSTM2D, las cuales aceptan entradas como los son (batch\_size, num\_frames, width, height, channels), para de esta manera regresar una predicción en video del mismo tamaño.

Con nuestro modelo y datos contruidos, ahora podemos entrenar el modelo. Para el entrenamiento del modelo se usó un entorno de ejecución GPU ya que si no se usa tarda mucho tiempo. Usando GPU aproximadamente cada época de entrenamiento duró aproximadamente 2 minutos.

Ya que el modelo está entrenado construido se pueden generar varias predicciones de las animaciones basadas en los videos que se analizaron, es decir los datos de entrenamiento.

Para este paso se elegirá un conjunto aleatorio de validación para después elegir los primeros diez fotogramas de ellos. De esta manera se puede permitir que el modelo prediga 10 nuevas animaciones y así, comparar las predicciones con las animaciones o fotogramas reales.

Al final se elegirán varios ejemplos del conjunto de validación y se construirán animaciones para poder ver tanto los videos predichos como los videos reales.

