

TP2: RNNs

Fecha de entrega límite: 1 de septiembre de 2022

Objetivo: Entender el funcionamiento de una red neuronal recurrente. Poder modificar sus parámetros comprendiendo las implicancias de esas modificaciones.

Requerimientos:

- Conocimientos de Python básico (y tener Python instalado o utilizar alguna plataforma online de Python en notebooks: Google Collaboratory, Kaggle, etc)
- Teoría de ML: redes neuronales feed-forwards, redes neuronales recurrentes
- Uso básico de Git y GitHub.

Ejercicio 1: Predicción del precio de Bitcoin en USD

Descargar el precio del BTC histórico en USD:

<https://finance.yahoo.com/quote/BTC-USD/history?p=BTC-USD>

Utilizando como base el script **stock_prediction.ipynb** cree una RNN y LSTM que sea capaz de predecir el precio futuro del BTC en USD.

- A. Hacer split del dataset en entrenamiento/test.
- B. Modificar con diferentes valores la dimensión del vector de la secuencia de entrada, la cantidad de capas, numero de features en el estado oculto. Explicar que sucede (en qué caso es conveniente un valor mayor o menor).
- C. Dibujar la curva de la función costo en función de los pasos de entrenamiento.
- D. Comparar el entrenamiento de una RNN y LSTM.
- E. ¿Qué métricas utilizaría para hacer una comparación objetiva entre la performance de cada modelo (RNN vs LSTM)?

Ejercicio 2: Predicción de dominio a partir de un título de un ítem de un e-commerce

Utilizar el **dataset** que se encuentra dentro del repositorio:

"modulo-2-rnn/rnn/datasets/items_meli_proccess.csv"

Utilizando como base el script **text_class_prediction.ipynb** cree una RNN o una LSTM que sea capaz de clasificar "correctamente" un dominio de otro.

- A. Hacer split del dataset en entrenamiento/test/valid.

-
- B. Elegir 2 clases y desarrollar un modelo de clasificación binaria.
 - C. Dibujar la curva de la función costo en función de los pasos de entrenamiento.
 - D. Pruebe cambiar al menos dos parámetros de la red. Explique.
 - E. ¿Qué métricas utilizaría para medir objetivamente la performance del modelo?
 - F. Repita estos pasos pero esta vez seleccionando 4 clases del dataset dado.