

Introducción a las Redes Long Short-Term Memory (*LSTM*)

Rodrigo Trejo

'Jesús les habló otra vez, diciendo: «Yo soy la Luz del mundo; el que me sigue no andará en tinieblas, sino que tendrá la Luz de la vida».'

Juan 8:12

April 5, 2024

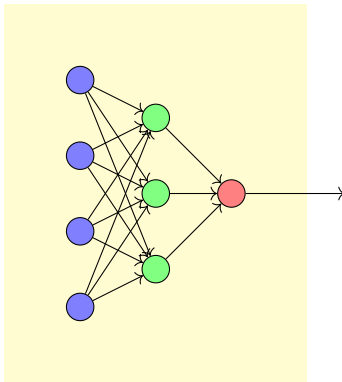
Presentation Overview

- ① Limitaciones de las RNA's
- ② Redes Recurrentes
- ③ La Celda LSTM

Predicción

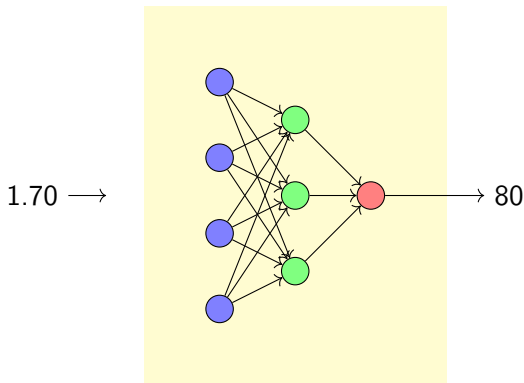
De acuerdo a las estadísticas de la Ciudad de México, si una persona mide 1.70 m, ¿Cuál será su peso en kg?

1.70 →



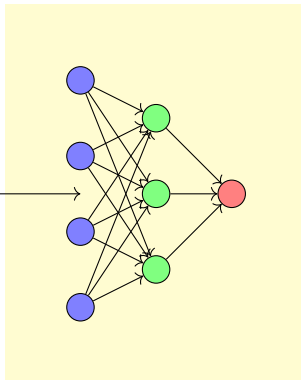
Predicción

De acuerdo a las estadísticas de la Ciudad de México, si una persona mide 1.70 m, ¿Cuál será su peso en kg?



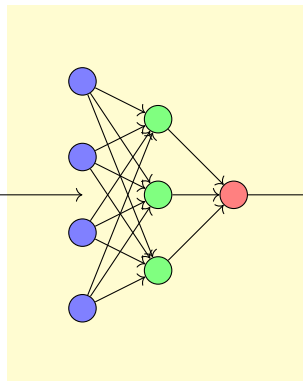
Clasificación

¿Cuál es el nombre del cantante?



Clasificación

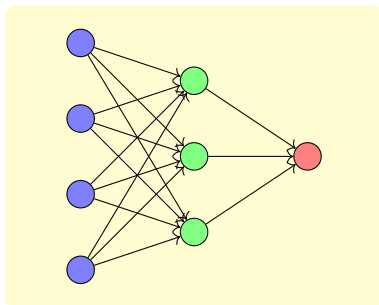
¿Cuál es el nombre del cantante?



Maluma

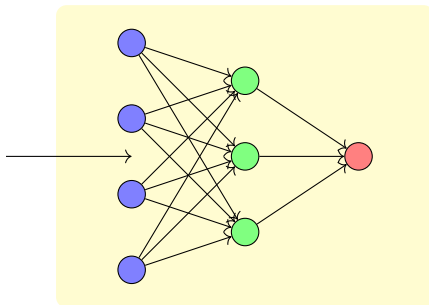
¿Secuencias?

Angelito va en el metro camino a su casa y se encuentra en esta estación...
¿Cuál es la siguiente?



¿Secuencias?

Angelito va en el metro camino a su casa y se encuentra en esta estación...
¿Cuál es la siguiente?

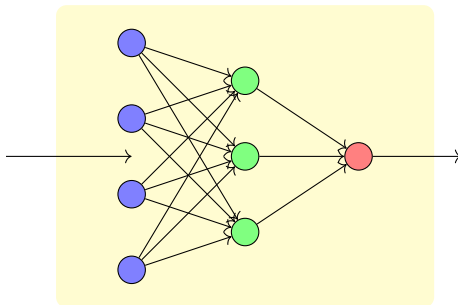


¿Secuencias?

Angelito va en el metro camino a su casa y se encuentra en esta estación...
¿Cuál es la siguiente?



PINO SUÁREZ



¿Secuencias?

¿Qué debería conocer la Red Neuronal para poder decir qué estación sigue?



¿Secuencias?

¿Qué debería conocer la Red Neuronal para poder decir qué estación sigue?



- Saber en qué línea está.

¿Secuencias?

¿Qué debería conocer la Red Neuronal para poder decir qué estación sigue?



- Saber en qué línea está.
- **Recordar** en qué línea está.

¿Secuencias?

¿Qué debería conocer la Red Neuronal para poder decir qué estación sigue?



- Saber en qué línea está.
- **Recordar** en qué línea está.
- **Recordar** en qué estaciones estuvo antes (para deducir la línea).

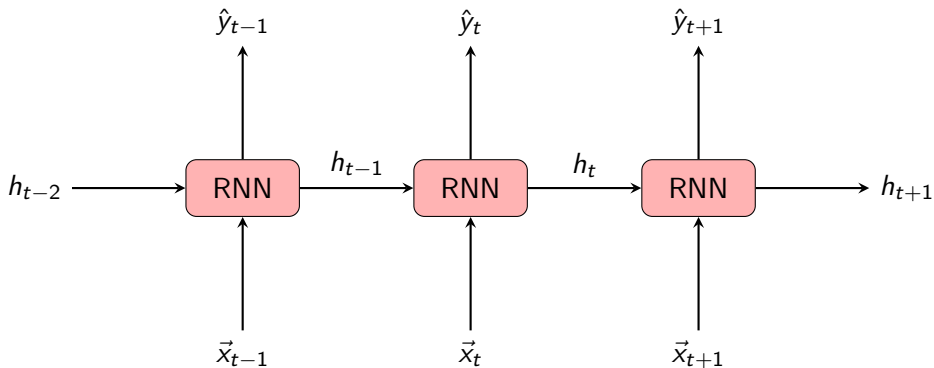
¿Secuencias?

¿Qué debería conocer la Red Neuronal para poder decir qué estación sigue?



- Saber en qué línea está.
- **Recordar** en qué línea está.
- **Recordar** en qué estaciones estuvo antes (para deducir la línea).
- TENER MEMORIA.

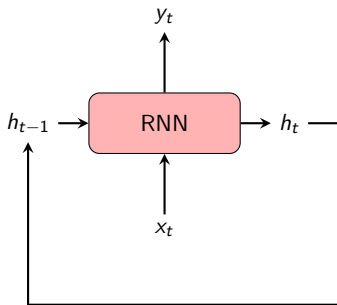
Red Recurrente (RNN): Una RNA con memoria



Observaciones sobre las RNN

- Cada 'rectángulo' RNN representa exactamente la misma Red Neuronal (con los mismos parámetros, capas, neuronas, tasa de aprendizaje) *solo que en diferentes tiempos*.
- Cada salida \hat{y}_{ti} es la predicción que el modelo arroja en un tiempo ti
- Cada salida que retroalimenta a la misma red h_{ti} se conoce como *estado oculto* y representa la memoria a corto plazo de la Red Neuronal, pues envía información de la activación de la misma red pero del estado anterior.

Representación de una RNN



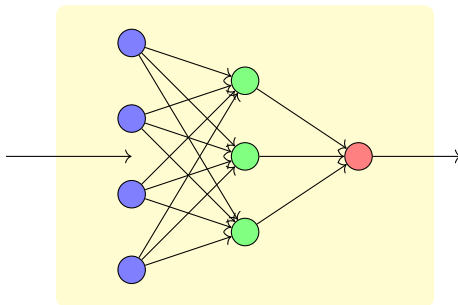
Al observar el diagrama anterior se suele pensar que se cuenta con una Red Neuronal *diferente* para cada tiempo t , por ello es mejor usar su representación con lazo de retroalimentación.

RNN y las Secuencias

Angelito va en el metro camino a su casa y se encuentra en esta estación...
¿Cuál es la siguiente?



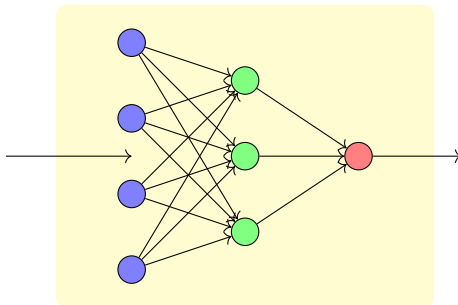
PINO SUAREZ



MERCED

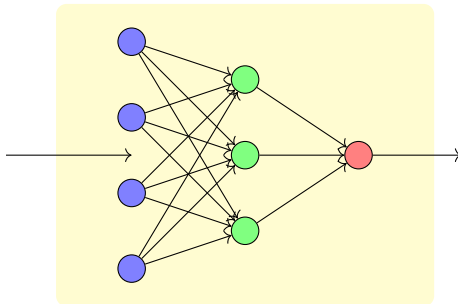
RNN y las Secuencias

Ahora, *como todos los días...* Angelito llega al metro Pantitlán... ¿Qué estación sigue?



Problema: Memoria a corto plazo

Ahora, como todos los días... Angelito llega al metro Pantitlán... ¿Qué estación sigue?



Problema: Memoria a corto plazo

- Prácticamente, la Red Neuronal Recurrente solo obtiene información del *estado inmediato anterior*.
- No se acuerda de lo que paso en tiempos muy alejados al presente.
- No puede hacer predicciones certeras a futuro

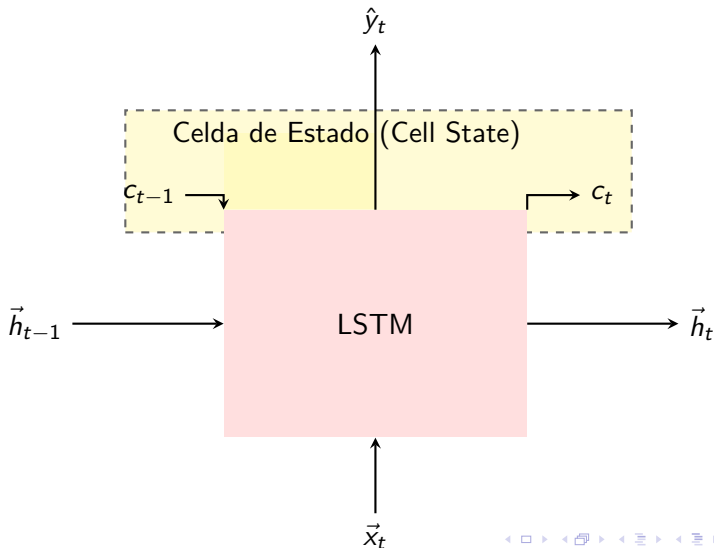


Problema: Memoria a corto plazo



¿La solución?
LA CELDA LSTM
(*Long Short-Term Memory*)

La Celda LSTM



Elementos de la Celda LSTM

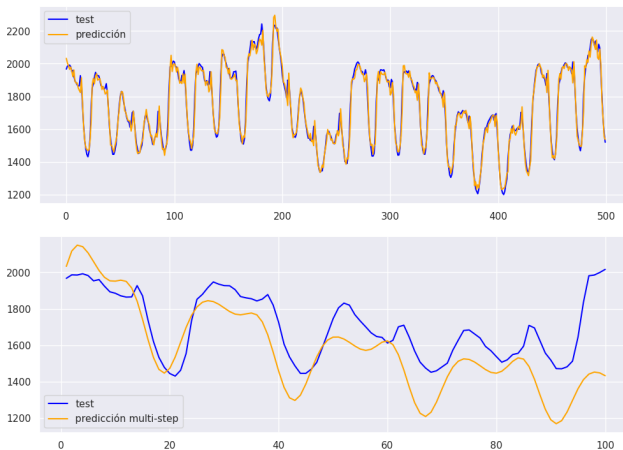
- Los elementos \vec{h}_{t-1} , \vec{h}_t , \vec{y}_t , \vec{x}_t tienen la misma función que en una RNN.
- La Red LSTM incorpora una nueva funcionalidad a diferencia de una RNN convencional: la celda de estado (Memoria a largo plazo).
- La entrada c_{t-1} almacena información de la memoria a largo plazo, que puede ser de utilidad para la Red Neuronal.
- La salida c_t contiene información actualizada para la memoria a largo plazo.

Elementos de la Celda LSTM

Una Red LSTM cuenta con *cuatro redes neuronales* diferentes en su interior, que en la literatura se asocian cuatro compuertas:

- **Forget Gate:** Decide qué información de la memoria a largo plazo se recuerda o se olvida (c_{t-1}).
- **Candidate memory:** Presenta elementos candidatos de la memoria a corto plazo para formar parte de la memoria a largo plazo (c_t).
- **Update Gate:** Actualiza la memoria a largo plazo, seleccionando los mejores candidatos de los presentados por la **Candidate memory**.
- **Output Gate:** Decide qué información del estado anterior (h_{t-1}) y de las entradas actuales vale la pena llevar a la memoria a corto plazo (h_t).

Ejemplito: Predictor de consumo de energía eléctrica



¡Gracias!

