

Predicción de series temporales

Redes LSTM (*Long Short Term Memory Networks*)

Manuel Carlevaro

Grupo de Materiales Granulares

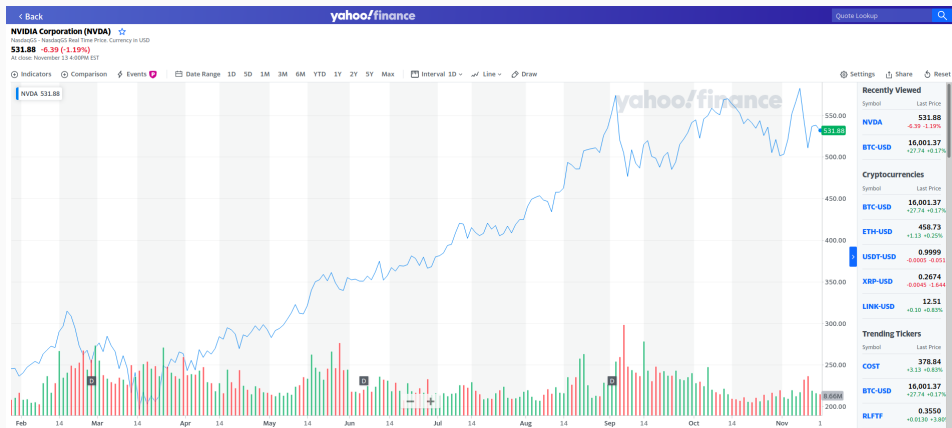
Dpto. Ing. Mecánica, UTN FRLP.

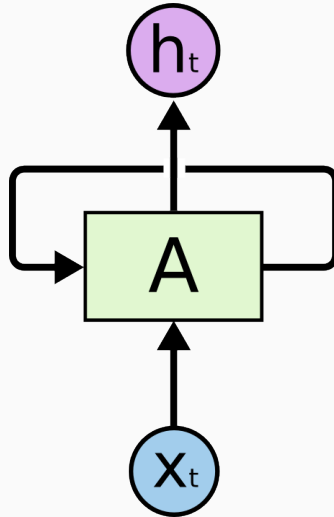
Taller de aprendizaje GMG - ML

17 de noviembre, 2020

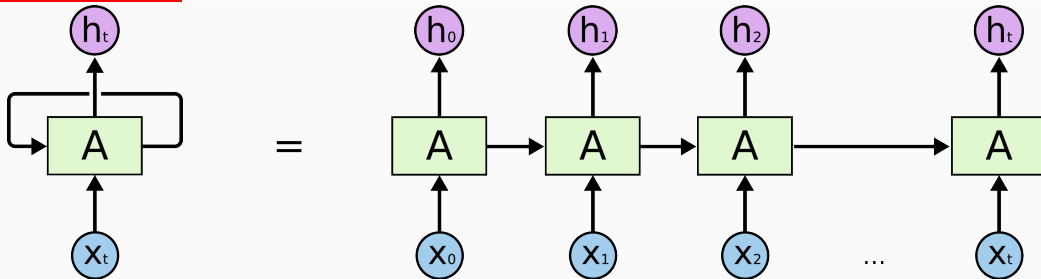
Secuencia temporal

$$x_0, x_1, x_2, \dots, x_{t-2}, x_{t-1}, x_t \rightarrow x_{t+1}$$





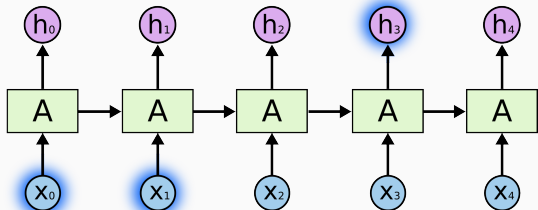
Redes neuronales recurrentes (RNN)



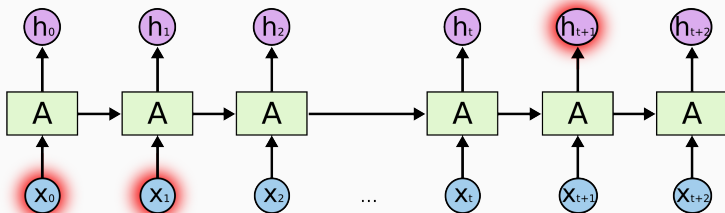
- John Hopfield: redes de Hopfield (1982)
- Aplicaciones: reconocimiento de voz, modelado del lenguaje, traducciones, descripción de imágenes ...
- *The Unreasonable Effectiveness of Recurrent Neural Networks*, [Andrej Karpathy blog](#)

Dependencia de largo plazo

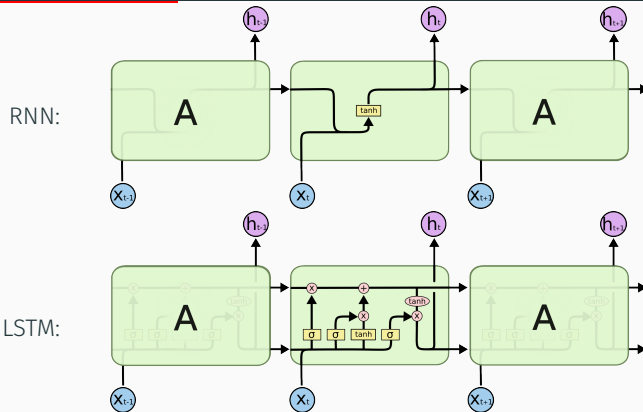
Las nubes están en el **cielo**.



Crecí en Córdoba ..., hablo fuido **cordobés**.

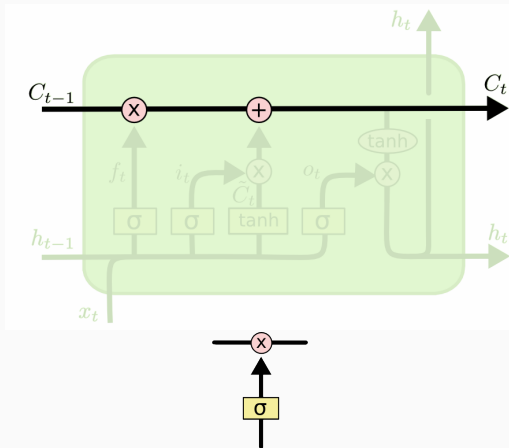


Redes LSTM (Long Short Term Memory Networks)

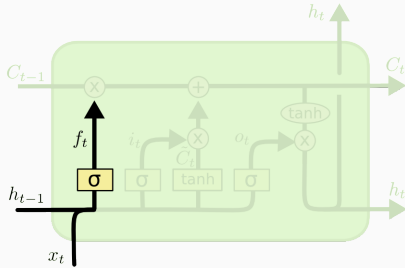


- Sepp Hochreiter y Jürgen Schmidhuber. "Long Short-Term Memory". En: *Neural Computation* 9.8 (1997), págs. 1735-1780.
- Aplicaciones: control de robots, predicción de series temporales, reconocimiento de voz, composición musical, aprendizaje de gramática, administración de pasajeros en aeropuertos, predicción de tránsito de corto plazo, diseño de drogas, etc.

Idea central de las LSTM

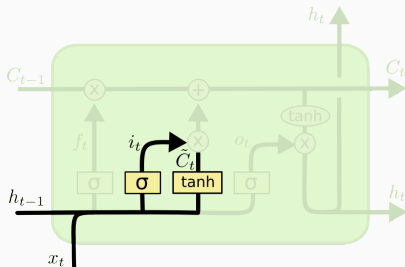


- El estado de la celda tiene solo dos interacciones lineales
- LSTM puede olvidar o agregar información mediante *gates*
- Las *gates* son NN sigmoideas y productos elemento por elemento:
 - $\sigma \rightarrow 1$: todo pasa
 - $\sigma \rightarrow 0$: nada pasa



$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f)$$

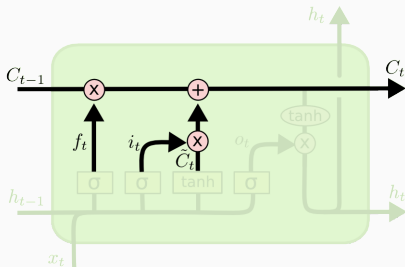
Input gate layer y actualización de estado



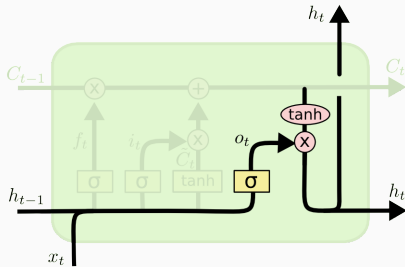
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Input gate layer y actualización de estado



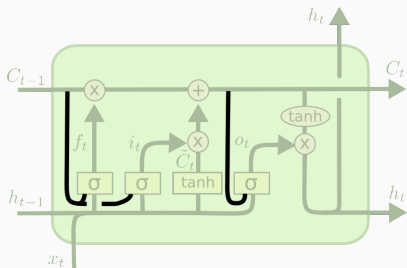
$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$



$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

Variantes de LSTM: “peephole connections”



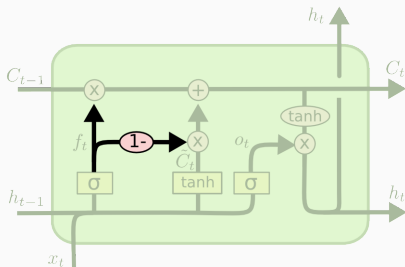
$$f_t = \sigma (W_f \cdot [C_{t-1}, h_{t-1}, x_t] + b_f)$$

$$i_t = \sigma (W_i \cdot [C_{t-1}, h_{t-1}, x_t] + b_i)$$

$$o_t = \sigma (W_o \cdot [C_t, h_{t-1}, x_t] + b_o)$$

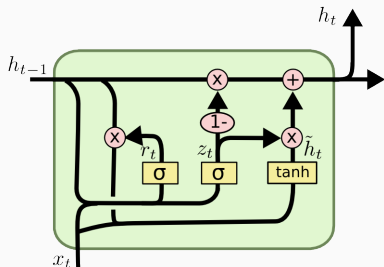
Felix A Gers y Jürgen Schmidhuber. “Recurrent nets that time and count”. En: *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*. Vol. 3. IEEE. 2000, págs. 189-194

Variantes de LSTM: *forget/input gates* acopladas



$$C_t = f_t * C_{t-1} + (1 - f_t) * \tilde{C}_t$$

Variantes de LSTM: GRU (Gated Recurrent Unit)



$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Kyunghyun Cho y col. "Learning phrase representations using RNN encoder-decoder for statistical machine translation". En: *arXiv preprint arXiv:1406.1078* (2014)

Esta presentación está adaptada de *Understanding LSTM Networks*, de [Christopher Olah](#). Las imágenes son de ese post.

- [1] Sepp Hochreiter y Jürgen Schmidhuber. “Long Short-Term Memory”. En: *Neural Computation* 9.8 (1997), págs. 1735-1780.
- [2] Felix A Gers y Jürgen Schmidhuber. “Recurrent nets that time and count”. En: *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks. IJCNN 2000. Neural Computing: New Challenges and Perspectives for the New Millennium*. Vol. 3. IEEE. 2000, págs. 189-194.
- [3] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk y Yoshua Bengio. “Learning phrase representations using RNN encoder-decoder for statistical machine translation”. En: *arXiv preprint arXiv:1406.1078* (2014).
- [4] Ian Goodfellow, Yoshua Bengio y Aaron Courville. *Deep Learning*. <http://www.deeplearningbook.org>. MIT Press, 2016.
- [5] Jason Brownlee. *Deep Learning for Time Series Forecasting*. <https://machinelearningmastery.com/deep-learning-for-time-series-forecasting/>. Jason Brownlee, 2018.

¿Habrá próxima temporada?

✉ manuel.carlevaro@gmail.com

 • Lua^ATeX

