

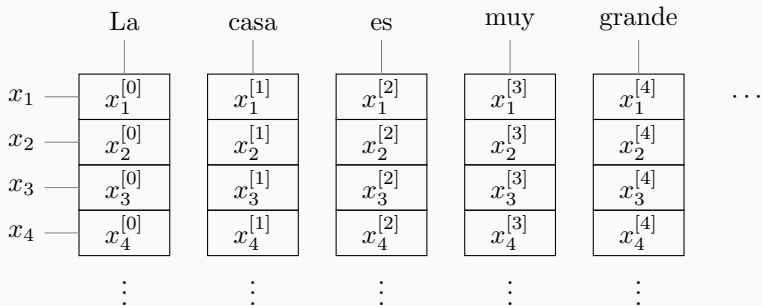
Aprendizaje profundo

CELDA RECURRENTE BÁSICA

Gibran Fuentes Pineda

3 de octubre de 2023

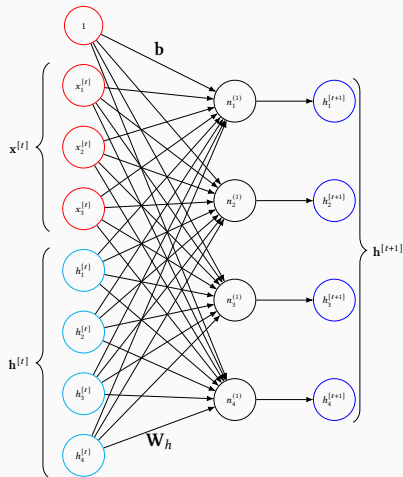
Motivación: secuencias de palabras



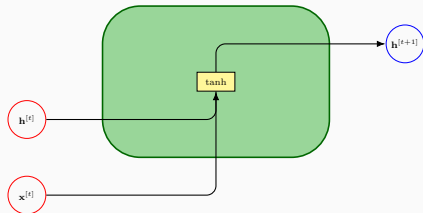
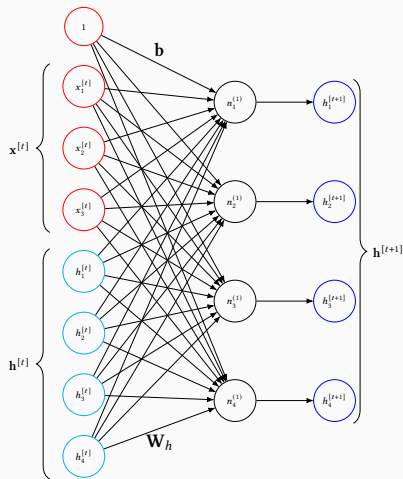
Unidad recurrente básica

- Capas con retro-alimentación en sus conexiones
 1. Entradas en tiempo t ($\mathbf{x}^{[t]}$)
 2. Estado en tiempo t ($\mathbf{h}^{[t]}$)

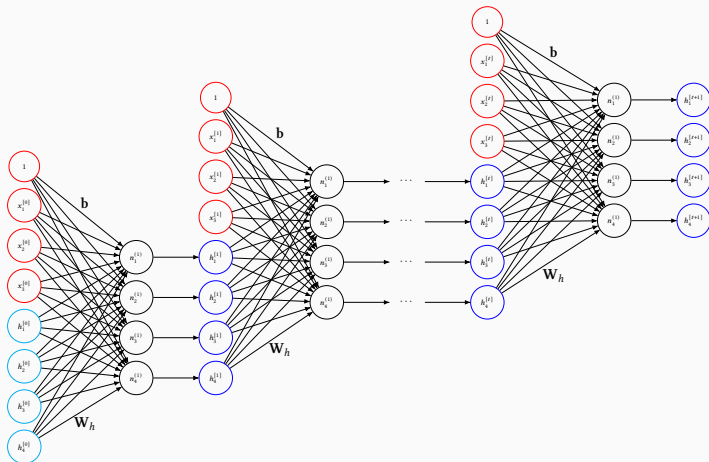
$$\begin{aligned} \mathbf{h}^{[t+1]} &= \phi \left(\mathbf{w}_h \cdot \underbrace{\begin{bmatrix} \mathbf{h}^{[t]} \\ \mathbf{x}^{[t]} \end{bmatrix}}_{\text{Concatenación}} + \mathbf{b}_h \right) \\ &= \phi \left(\mathbf{w}_{hh} \cdot \mathbf{h}^{[t]} + \mathbf{w}_{hx} \cdot \mathbf{x}^{[t]} + \mathbf{b}_h \right) \end{aligned}$$



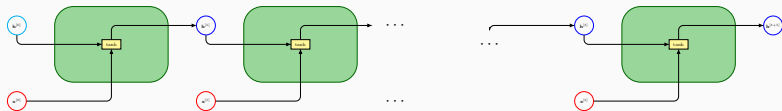
Unidad recurrente básica: diagrama de celda



Unidad recurrente básica: despliegue

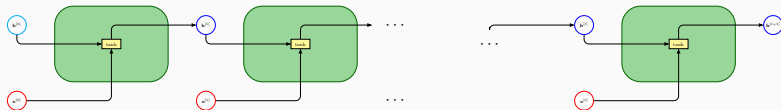


Unidad recurrente básica: despliegue de celdas



Modelando dependencias a corto plazo

- En teoría una red recurrente básica puede modelar dependencias a corto y largo plazo
 - Siegelmann y Sontag mostraron que las redes recurrentes son Turing completas¹



¹Siegelmann and Sontag. On The Computational Power Of Neural Nets, 1995.

El problema de la memoria a largo plazo

- En práctica es muy difícil entrenarlas para tareas con dependencias a largo plazo

