

Redes recurrentes

Capa de neuronas recurrentes desplegada en el tiempo

Hay varias posibilidades...

Problemas de entrenamiento:

Redes Recurrentes vs. Redes hacia adelante

Redes Neuronales de Tensores Recursivas

Ejemplo:

Útiles cuando los patrones en los datos cambian con el tiempo

Pueden recibir una **secuencia** de valores como **entrada** y pueden también devolver una **secuencia** de valores de **salida**



Estas redes tienen bucles de retroalimentación

El primero es un perceptrón recurrente → una sola neurona que recibe una entrada, esa entrada se suma, pasa por la función de activación, pero una salida vuelve a ingresar

Para entenderla, la desplegamos en el tiempo

Tengo los mismos problemas que con las redes profundas.

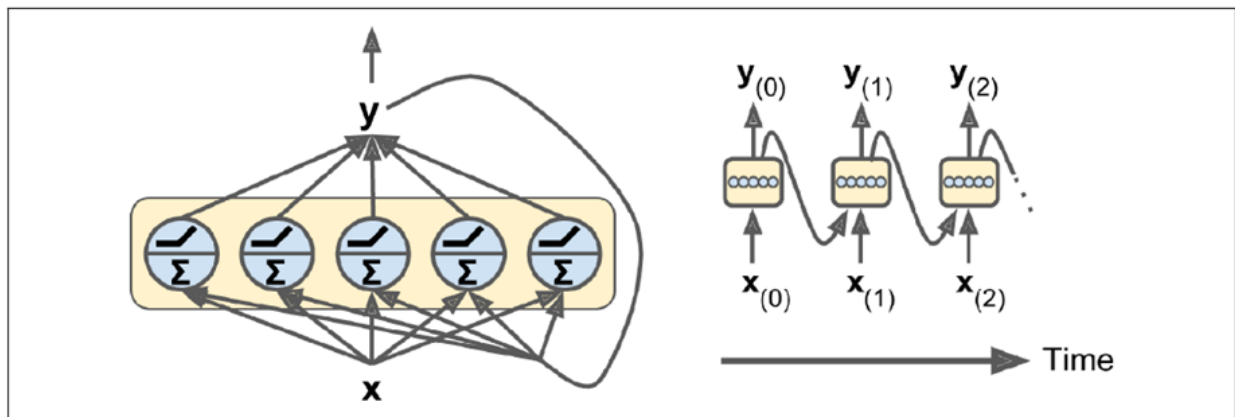
Si tengo 10 pasos temporales para una entrada, lo que voy a tener en el tiempo es

como si tuviese una entrada que pasa a otra neurona y así sucesivamente hasta t pasos

Capa de neuronas recurrentes desplegada en el tiempo

Una capa con 5 neuronas recurrentes que convergen en una salida y, que es retroalimentada nuevamente

Si hago un despliegue temporal veo que hay una salida y0 que vuelve a entrar en el paso x1, y así sucesivamente



Ecuación de la salida de una capa recurrente

$$y_{(t)} = \phi\left(\mathbf{W}_x^T \mathbf{x}_{(t)} + \mathbf{W}_y^T y_{(t-1)} + \mathbf{b}\right)$$

phi: funcion de activacion

W: matriz de pesos

x(t): vector de entrada

y(t-1): salida

b: umbral

Celdas de memoria:

La parte de una red neuronal que mantiene algún estado a lo largo de los pasos de tiempo recibe el nombre de “Celda de memoria” o simplemente “celda”.

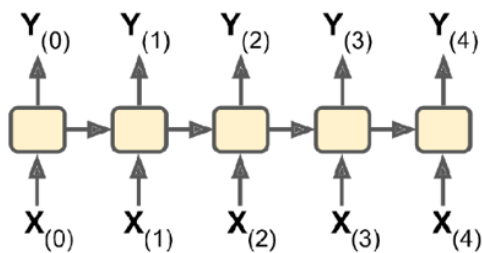
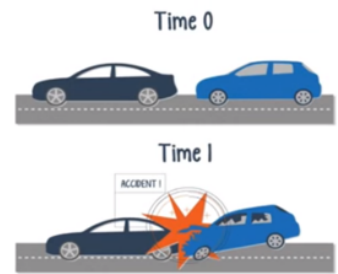
Cómo las redes recurrentes en el paso de tiempo, es una función de todas las salidas anteriores, se podría decir que tiene una especie de memoria....

Hay varias posibilidades...

- **Entrada:** secuencia
- **Salida:** secuencia



clasificar vídeos
cuadro a cuadro



Series temporales:

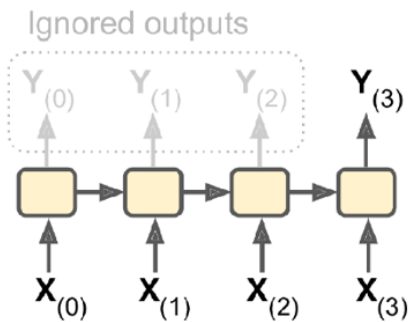
Por ejemplo, precios de acciones, se introducen los precios de los últimos N días y la salida deben ser los precios desplazados un día en el futuro (N-1 hasta mañana)

Los cuadros de video son secuencias de tiempo

- **Entrada:** secuencia
- **Salida:** singular



clasificación de documentos



Entrada:
Secuencia de palabras de una crítica cinematográfica. Cada palabra en un paso de tiempo.

Salida:
un único valor entre -1 y 1 (-1[odio], +1[amor])



Ignoro todas las salidas anteriores y me quedo con la ultima salida

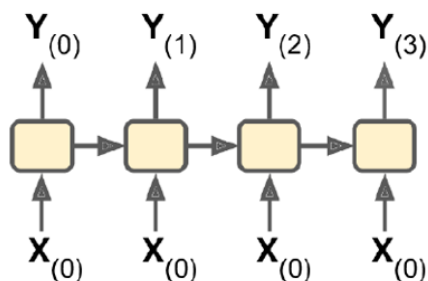
Podria ser un analisis de sentimientos con redes recurrentes

Se entrenan con Backpropagation. Para entrenar las redes hay que hacer n despliegue en el tiempo.

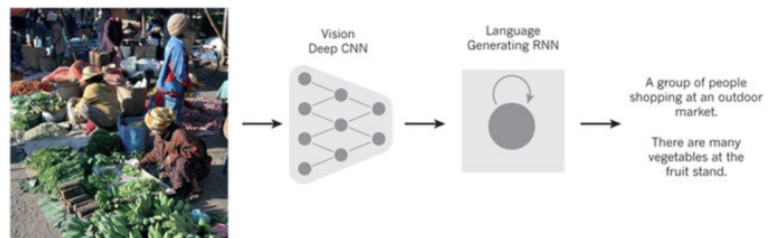
- **Entrada:** singular
- **Salida:** secuencia



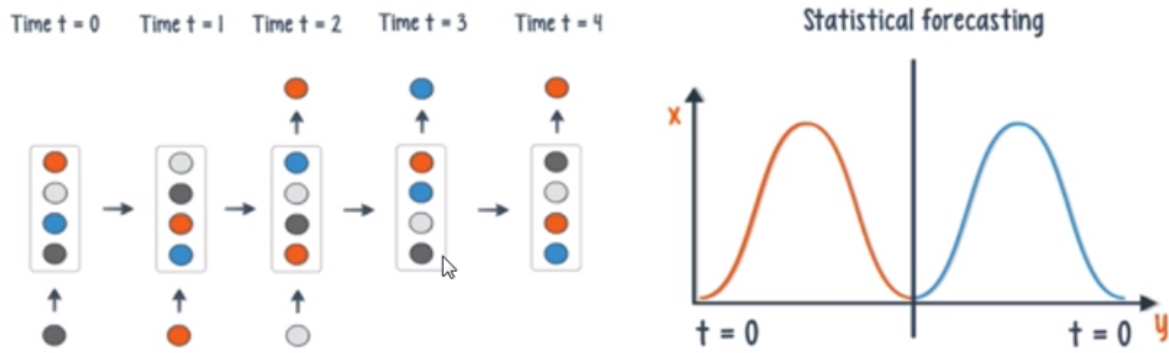
Etiquetamiento



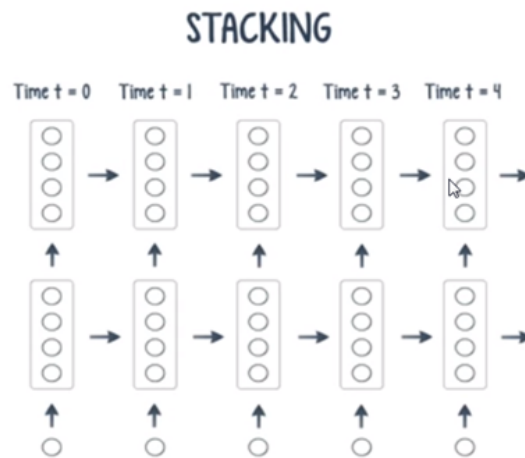
Pasamos una imagen una y otra vez (o la salida de un red Convolutiva) y va generando una secuencia de palabras que puede ser un título para la imagen.



Si se introduce una pequeña demora, la red puede de forma estadística, pronosticar la demanda en una cadena de suministros.



Apilamiento: Se pueden apilar y formar capas

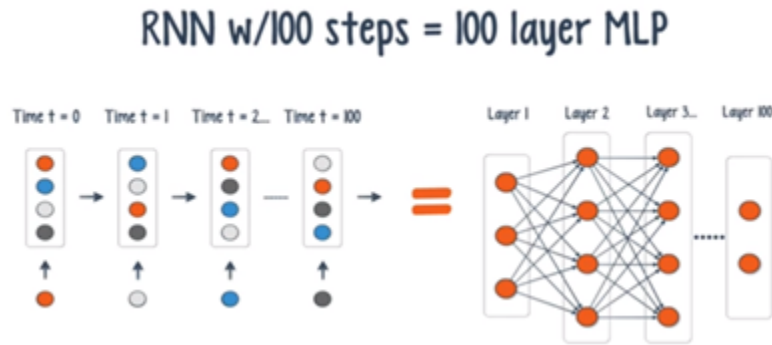


Problemas de entrenamiento:

El problema del desvanecimiento del gradiente es exponencialmente peor

1 sola capa con 100 pasos temporales es como entrenar una red de 100 capas de alimentación hacia adelante

(problema similar a redes profundas)



Soluciones posibles a los problemas de entrenamiento

- **Gating**
 - LSTM
 - GRU
- Gradient Clipping
- Better Optimizers
- Steeper gates

Gating es una técnica que le dice a la red cuando olvidar el input y cuando recordarlo

LSTM y GRU son nombres de celdas de memoria.

Redes neuronales o unidades minimas de memoria que tienen ciertos mecanismos internos que permiten olvidar algunos pasos para facilitar el entrenamiento

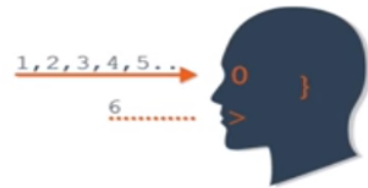
Redes Recurrentes vs. Redes hacia adelante

- **Red hacia adelante:**
Clasificación o regresión
- **Red Recurrente:**
Serie, predicción, pronóstico

Feedforward net = Classifier/Regressor



Recurrent Net = Forecaster

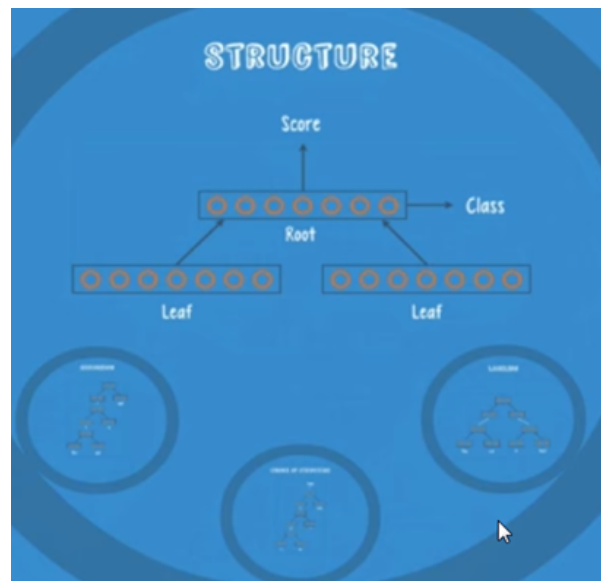


Redes Neuronales de Tensores Recursivas

(Recursive Neural Tensor Nets)

Redes recurrentes pero donde la red es un tensor

- Para realizar tareas de análisis de sentimiento, tiene en cuenta el **orden** y la **agrupación sintáctica**
- Tienen estructura de árbol
 - Neuronas raíz
 - Neuronas Hojas



El nodo raíz genera un “score” que es vuelto a ingresar en los nodos hojas de forma recursiva

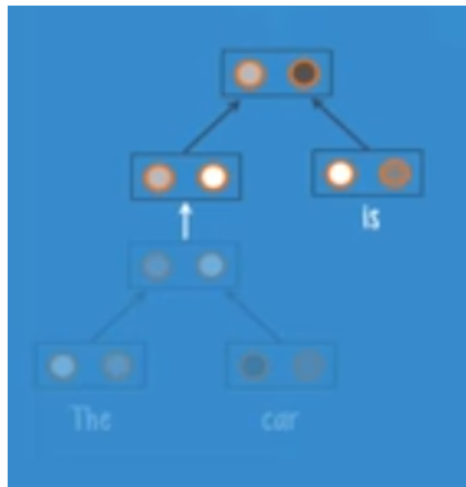
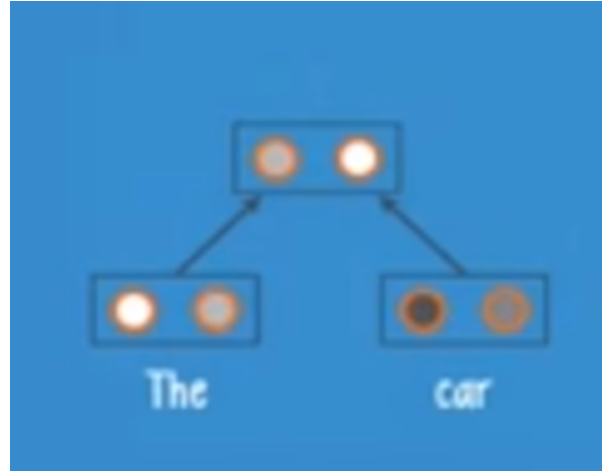
Ejemplo:

“The car is fast”

“Alimentamos” con las dos primeras palabras a los grupos hojas uno y dos respectivamente.

Las palabras las convertimos a vectores (sej con la tecnica Bag of Words) para que puedan ser trabajadas por las redes neuronales

Trabajan con array de numeros, vectores. No con manipulacion de signos



Paso 02

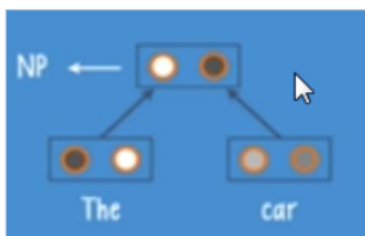


Paso 03

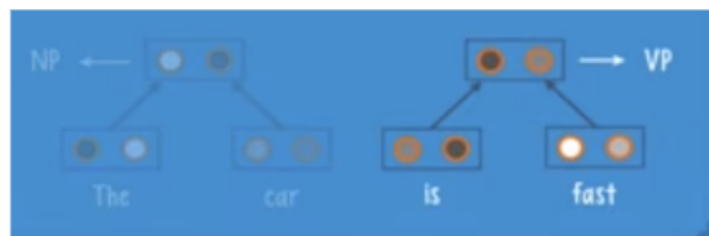
Entonces lo que hace la red es ir aprendiendo la estructura sintáctica

Etiquetamiento gramatical

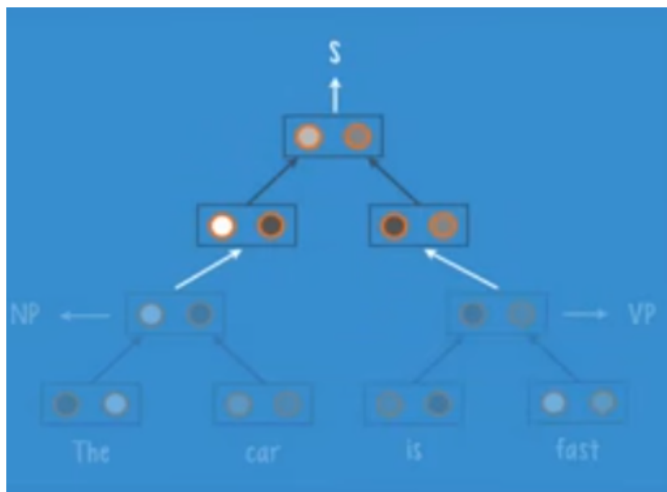
Una vez que la red obtuvo el mejor parser, vuelve sobre sus pasos (*backtracks*) a través del parser para obtener las etiquetas gramaticales:



Etiqueta esto como
frase nominal



Luego vuelve sobre este paso y etiqueta el
árbol como Frase Verbal

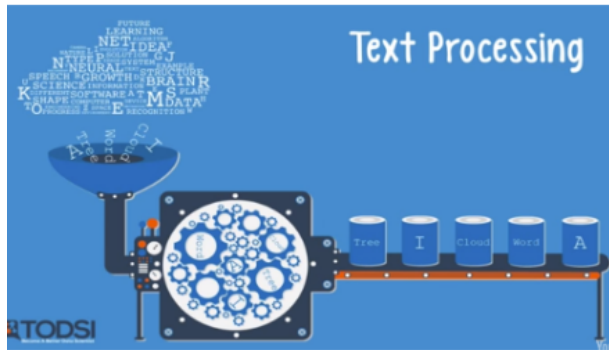


Etiqueta todo como “S” (sentence)

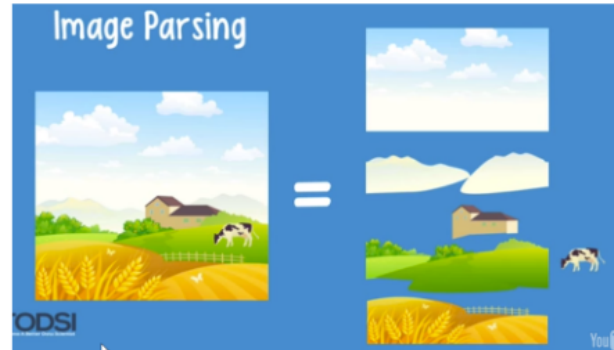
Este tipo de redes utilizan **BackPropagation**. Para ello comparan el árbol de dependencias obtenido con un árbol de dependencias generado a mano para una misma oración de entrada

La función de Loss o error es más compleja, porque hace una diferencia entre el árbol entero sintáctico vs el árbol generado de forma manual

Usos



Especialmente Análisis de Sentimientos



Detección de componentes en imágenes