13. Predicción de secuencias

Ejercicio 13.1. Se te ha dado un conjunto de datos de mensajes etiquetados con sentimientos positivos y negativos. Suponer que las palabras en los mensajes están representadas mediante embeddings de dimensión 100.

1. Dibujar el esquema de la arquitectura de una Red Neuronal Recurrente (RNN) con **dos capas ocultas** que permita resolver este problema. Dibujar dos versiones: una compacta y una desplegada en el tiempo. Indicar la dimensión de la entrada, los estado ocultos y la salida.
2. ¿Qué función de pérdida utilizaría?
3. ¿Sobre qué instantes de tiempo aplicaría la función de pérdida? En caso que haya instantes de tiempo en el cuál no calcula la función de pérdida, ¿qué haría en esos instantes para no tener que cambiar los algoritmos de minimización?

Ejercicio 13.2. Dados los tipos de arquitecturas básicas para RNNs: *uno-a-muchos*, *muchos-a-uno*, *muchos-a-muchos-alineado* y *muchos-a-muchos-no-alineado*, decidir qué arquitectura usaría en cada uno de los siguientes casos. Justifique.

1. Generar la descripción de un producto a partir de una foto del mismo.
2. Predecir si una reseña de un producto es positiva o negativa.
3. Traducir tickets de compra (español a inglés por ejemplo).
4. Pronosticar la serie diaria de ventas del próximo mes a partir de datos del último año.
5. Generar recomendaciones de productos en base a una categoría.
6. Corregir ortografía de una frase introducida por el usuario.
7. Dado un audio largo de una conversación, predecir quién habla en segmento de la conversación.
8. Dado un libro generar el audio libro.

Ejercicio 13.3. Dado el desdoblamiento de los primeros pasos temporales de una RNN (Figura 7), donde tanto la capa oculta como la capa de salida son lineales:

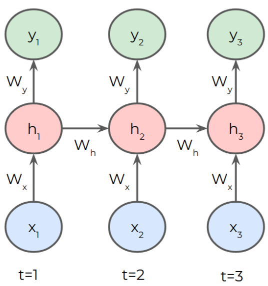

Figura 7: Desdoblamiento de la RNN en los tiempos $t = 1, 2, 3$.

1. Escribe las ecuaciones de la RNN (actualización de la capa oculta y cálculo de la salida) en función de los pesos W_x, W_h y W_y. Suponer que todos los biases son 0.
2. Escribe a la red como una composición de funciones.
3. Calcula los valores de las unidades ocultas h_1, h_2, h_3 y de las salidas y_1, y_2, y_3 para los valores de entrada
$$x_1 = 2, \quad x_2 = -0.5, \quad x_3 = 1,$$
asumiendo que $W_x = W_h = W_y = 1$ y sin funciones de activación adicionales (lineal puro).
4. Explica qué está computando la red en este caso.
5. Dibuja el grafo de cómputo de la RNN para estos tres pasos, incluyendo la función de pérdida $\mathcal{L}$ al final de la salida y_3.

Ejercicio 13.4. Considera una RNN simple con las siguientes ecuaciones:

$$h_t = \tanh\bigl(W_{xh}x_t + W_{hh}h_{t-1} + b_h\bigr), \quad y_t = W_{hy}h_t + b_y,$$

donde:

- x_t es la entrada en el tiempo t,
- h_t es el estado oculto en el tiempo t (con $h_0 = 0$),
- W_{xh}, W_{hh}, W_{hy} son matrices de pesos,
- b_h, b_y son vectores de sesgo.
- $tanh$ de un vector se define como la aplicación punto a punto.

1. Dado los siguientes valores, calcula los estados ocultos h_1 y h_2 para la secuencia de entrada:

$$x_1 = 1, \quad x_2 = 2,$$

$$W_{xh} = \begin{pmatrix} 0.6 & 0.1 \end{pmatrix}, \quad W_{hh} = \begin{pmatrix} 0.5 & 0.2 \\ 0.3 & 0.7 \end{pmatrix}, \quad b_h = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad b_y = 0.1.$$

2. Con los valores anteriores, ¿cuál es la dimensionalidad de h_t, y_t y W_{hy} en esta RNN? Justifica tu respuesta.

Ejercicio 13.5.
Dada una RNN con los siguientes parámetros y una instancia k:

- Entrada $\mathbf{x}_{\langle t\rangle}^{(k)} \in \mathbb{R}^p$
- Salida $\mathbf{o}_{\langle t\rangle}^{(k)} \in \mathbb{R}^c$
- Estado oculto $\mathbf{h}_{\langle t\rangle}^{(k)} \in \mathbb{R}^q$
- Matrices de pesos $W_{xh} \in \mathbb{R}^{q\times p}$, $W_{hh} \in \mathbb{R}^{q\times q}$, y $W_{hy} \in \mathbb{R}^{c\times q}$
- Vectores de sesgo $b_h \in \mathbb{R}^q$ y $b_y \in \mathbb{R}^k$
- Función de activación: tangente hiperbólica: $\tanh(z) = \dfrac{e^z - e^{-z}}{e^z + e^{-z}}$

Se pide:

1. Escribe las ecuaciones para el estado oculto $\mathbf{h}_{\langle t\rangle}^{(k)}$ y la salida $\mathbf{o}_{\langle t\rangle}^{(k)}$ en el paso de tiempo t.
2. Supongamos que tenemos los siguientes valores:
   $p = 3, q = 2, c = 1$

   $$x = \left[\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} \begin{pmatrix} 0 \\ 4 \\ 2 \end{pmatrix} \right], \quad h_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   $$W_{xh} = \begin{pmatrix} 0.1 & 0.2 & 0.3 \\ 0.4 & 0.5 & 0.6 \end{pmatrix}, \quad W_{hh} = \begin{pmatrix} 0.7 & 0.8 \\ 0.9 & 1.0 \end{pmatrix}, \quad W_{hy} = \begin{pmatrix} 1.1 & 1.2 \end{pmatrix}, \quad b_h = \begin{pmatrix} 0.1 \\ 0.2 \end{pmatrix}, \quad b_y = \begin{pmatrix} 0.3 \end{pmatrix}$$

   Calcular los valores resultantes del estado oculto y las salidas en cada instante de tiempo. Tip $\mathbf{h}_{\langle 1\rangle}^{(k)} = \begin{pmatrix} 0.53 \\ 0.92 \end{pmatrix}$
3. Programar la cuenta anterior utilizando el siguiente código base:

```
import numpy as np

# Definir la función de activación tanh
def tanh(z):
    return np.tanh(z)

# Valores iniciales y parámetros
x = [np.array([1, 2, 0]), np.array([0, 4, 2])]
```

```
h_0 = np.array([0, 0])

W_xh = np.array([[0.1, 0.2, 0.3], [0.4, 0.5, 0.6]])
W_hh = np.array([[0.7, 0.8], [0.9, 1.0]])
W_hy = np.array([[1.1, 1.2]])

b_h = np.array([0.1, 0.2])
b_y = np.array([0.3])
```

Ejercicio 13.6. Se extiende a las RNN como muestra la imagen. Dar las fórmulas asociadas a la nueva arquitectura (en donde la línea punteada corresponde a conexiones recurrentes).

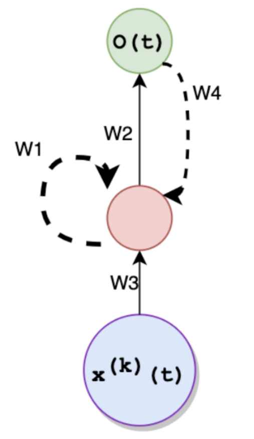

Figura 8: Diagrama de RNN

Ejercicio 13.7. Diseñar la arquitectura de una red neuronal de tipo Encoder - Decoder que permita resumir texto. Contestar luego las siguientes preguntas:

1. ¿Qué tipo de red utilizó en el Encoder?
2. ¿Qué tipo en el Decoder?
3. ¿Cómo conecta el Encoder con el Decoder?
4. ¿Utilizó algún mecanismo de atención?
5. ¿Utilizó redes que leen la entrada de manera secuencial?
6. ¿Utilizó redes que leen la entrada de manera bidireccional?
7. ¿Utilizó redes que leen la entrada todo al mismo tiempo?
8. ¿Hay algún límite en el tamaño de la entrada?