

Trabajo Práctico: Redes recurrentes

Deep Learning

M. Sc. Saúl Calderón Ramírez
Instituto Tecnológico de Costa Rica,
Escuela de Ingeniería en Computación,
PAttern Recognition and MACHine Learning Group (PARMA-Group)

23 de julio de 2019

El presente proyecto introduce las redes recurrentes y su implementación en Pytorch.

- **Modo de trabajo:** Grupo de dos personas.
- **Tipo de entrega:** digital, por medio de la plataforma TEC-digital.

Parte I

(100 puntos) Red de Elman para series temporales

1. (60 puntos) Estimación de la posición en dos dimensiones en una serie temporal

Modifique la red de Elmann para resolver el problema de estimación de la posición $\langle x_t, y_t \rangle$ de un objeto en un tiempo t , dada una secuencia de posiciones k anteriores:

$$[\langle x_{t-1}, y_{t-1} \rangle, \langle x_{t-2}, y_{t-2} \rangle, \dots, \langle x_{t-k}, y_{t-k} \rangle]$$

Lo cual se puede conceptualizar como una serie de muestras $\vec{m}_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix}$, por lo que entonces, la secuencia se puede reescribir como una secuencia de vectores:

$$S_i = [\vec{m}_{t-1}, \vec{m}_{t-2}, \dots, \vec{m}_{t-k}]$$

con lo que el modelo se representa en la siguiente función:

$$\langle x_t, y_t \rangle = f([\vec{m}_{t-1}, \vec{m}_{t-2}, \dots, \vec{m}_{t-k}])$$

Se aconseja usar Pandas para crear los reportes en .csv de los resultados.

1. **(10 puntos)** Explique e ilustre los cambios necesarios en la arquitectura provista e implementelos.
2. **(30 puntos)** Genere tres conjuntos de datos distintos basados en una onda sinusoidal $y_t = \sin(r_t)$, $x_t = r_t$ con un largo de $k = 3000$ muestras, contaminadas con ruido Gaussiano con desviación estándar σ_1 , σ_2 y σ_3 , a su elección.
 - a) Reporte todos los hiperparámetros utilizados.
3. **(20 puntos)** Para cada set de datos, separe el 80 % de la secuencia para entrenamiento, y el 20 % restante para validación, y genere una tabla donde se mida el error medio absoluto para las tres secuencias.
 - a) Comente los resultados y la sensibilidad del modelo al ruido.
 - b) Proponga e implemente una métrica que mida la sensibilidad del ruido del modelo, es decir, el valor de la métrica debe ser mayor entre más sensibilidad al ruido.

Parte II

(100 puntos) Red recurrente para la estimación del idioma por palabra

El sistema implementado para el código adjunto soluciona el problema de estimar el lenguaje de una palabra. El conjunto de datos cuenta con palabras de 18 lenguajes distintos. El modelo tiene entonces como entrada una secuencia de caracteres, con largo variable k , en codificación ASCII:

$$S_i = [\vec{m}_1, \vec{m}_2, \dots, \vec{m}_k]$$

y como salida, un vector \vec{y} con notación $1 - K$ para representar las $K = 18$ clases posibles:

1. **(10 puntos)** Para la red recurrente implementada en el ejemplo provisto, dibuje el grafo sin desarrollar, y detalle sus funciones de activación y función de pérdida.
 - a) Reporte de los resultados promedios del modelo para 10 corridas, usando lo que se conoce como una **matriz de confusión**.
2. **(30 puntos)** Implemente y reporte en el informe el cálculo de falsos positivos, falsos negativos, sensibilidad, exhaustividad y *F1-score*, y calcule el promedio para todas las clases.

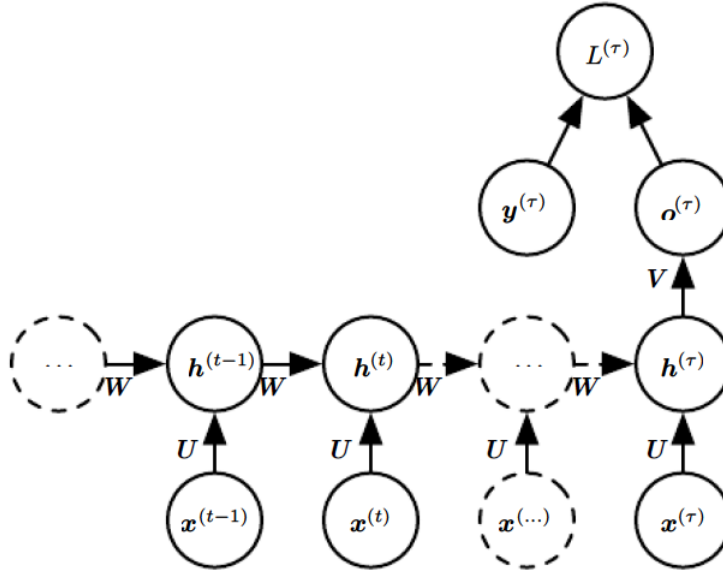


Figura 1: Modificación propuesta a la red recurrente original. Tomado del libro *Deep Learning*, Yoshua Bengio et. al.

- a) Incluya una reseña breve de cada métrica.
3. **(30 puntos)** Explique qué ventajas tiene la arquitectura mostrada en la Figura 1 respecto a la arquitectura original presentada en el código provisto, y explique las modificaciones en el código necesarias para implementar tal arquitectura.
 - a) Pruebe y reporte los resultados de *F1-Score*, exhaustividad y precisión, en el reporte usando el paquete *tikz* y el script de *matplot2tikz*, con una gráfica para el modelo original y el modificado.
4. **(30 puntos)** Investigue en qué consiste el problema de balanceo de clases.
 - a) Explore y explique el balanceo de clases para el problema a resolver en esta sección y para el modelo originalmente implementado en el código
 - b) Proponga e implemente las modificaciones pertinentes para realizar tal balanceo de clases, usando la técnica de su elección.
 - c) Pruebe y reporte los resultados de *F1-Score*, exhaustividad y precisión, en el reporte usando el paquete *tikz* y el script de *matplot2tikz*, con una gráfica para el modelo original y el modificado.