

REDES NEURONALES ARTIFICIALES

Lección 4

REDES NEURONALES RECURRENTES

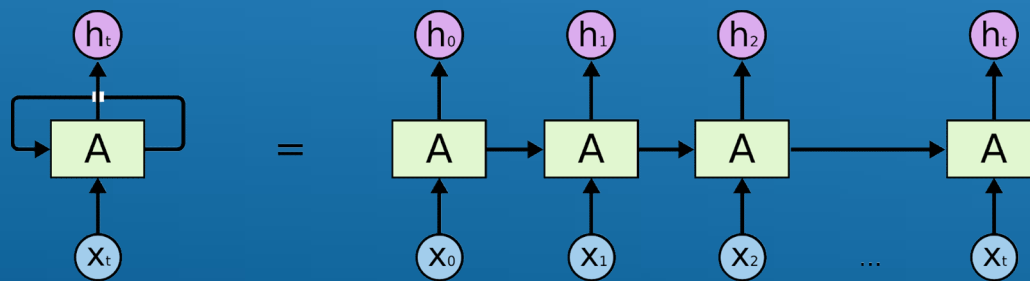
Entrenamiento – Retropropagación a través del tiempo

Con cada red neuronal el entrenamiento se vuelve mas complejo, y esta no es la excepción. Para realizar el entrenamiento de las redes neuronales recurrentes se sigue el mismo principio que en los perceptrones multicapa, aplicando el algoritmo de retropropagación con una variación importante, la retropropagación del error también se aplicará al bucle recursivo de las neuronas.

Para facilitar las cosas se puede visualizar las redes neuronales recurrentes como una secuencia de redes neuronales tradicionales que se entrenan una después de otra con el algoritmo de retropropagación.

Entrenamiento – Retropropagación a través del tiempo

La siguiente imagen muestra a la izquierda del signo igual una red neuronal recurrente, y a la derecha se muestra el bucle recursivo de una manera secuencial para mayor comprensión visual, representando cada paso en el tiempo como si fuera una nueva capa de una red neuronal tradicional. Cuando decimos que se aplicará la retropropagación al bucle recursivo nos referimos a que se tratara como si cada neurona fuera una red neuronal tradicional y cada paso en el tiempo es la siguiente capa, ya que el error de un paso en el tiempo depende del paso anterior.



Long Short Term Memory (LSTM)

LSTM es una arquitectura de las redes recurrentes que permiten construir redes neuronales recurrentes con la capacidad de mantener información más antigua por un mayor tiempo de ser necesario.

Los LSTM están compuesto por 3 compuertas:

- Una encargada de decidir qué información entra
- Una encargada de decidir qué información olvidar
- Una encargada de decidir qué información sale a la siguiente capa.

Long Short Term Memory (LSTM)

Este aumento en el tiempo que la información es recordada por la red permite hacer predicciones de las cuales se necesite mucho más que solo los datos más recientes.

Por ejemplo, en una red recurrente sin LSTM la capacidad de predecir la siguiente palabra en un texto depende de las palabras más recientes, por lo tanto si la palabra clave para predecir se encuentra 2 párrafos atrás resultaría imposible predecir la siguiente palabra, y es acá donde entra la utilidad de las LSTM.

Long Short Term Memory (LSTM)

En una pequeña oración como “El cielo es...” una red sin LSTM no tendría problema en adivinar que la palabra siguiente es “azul”, pero si en un párrafo dice “yo nací en Francia” y 2 párrafos después se encuentra la frase “mi idioma nativo es...” entonces sólo una red que cuente con LSTM podría predecir que la siguiente palabra es “francés”.

Cuando se esta hablando de un genero, ya sea masculino o femenino, la red puede recordar el genero del sujeto para siempre predecir el genero de los verbos o cualquier cosa.