

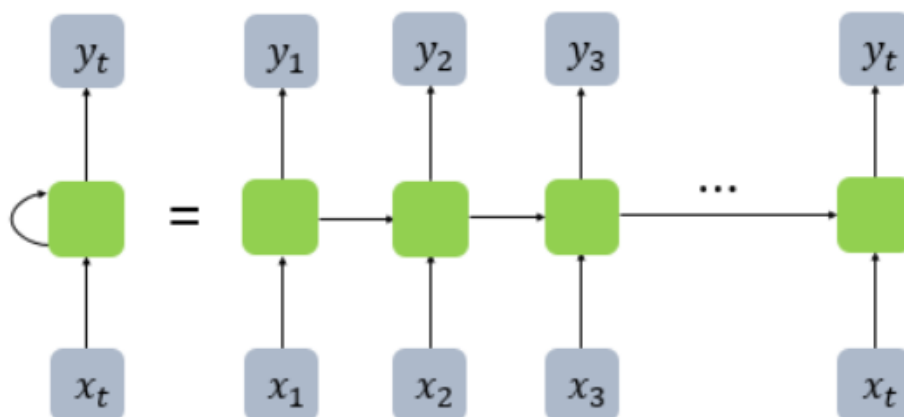
순환 신경망(Recurrent Neural Network)

순환 신경망(RNN)

- 시퀀스 모델 : 입력과 출력을 시퀀스 단위로 처리하는 모델
 - 정의
 - 일반적인 정의 : 일련의 연속적인 사건들 or 사건이나 행동 등의 순서
 - 자연어처리에서의 정의 : **연관된 연속의 데이터**
 - 문장은 단어 시퀀스 → 즉, 이러한 **시퀀스(문장, 단락 등)**를 처리하기 위해 고안된 모델을 시퀀스 모델이라 함
- RNN
 - RNN 특징 : 은닉층에서 활성화 함수를 지난 결과값을 출력층 방향으로도 보내면서, 다시 은닉층 노드의 다음 계산의 입력으로 보냄
 - feed forward NN : 은닉층에서 활성화 함수를 지난 결과값은 오직 출력층 방향으로 함
 - RNN 도식



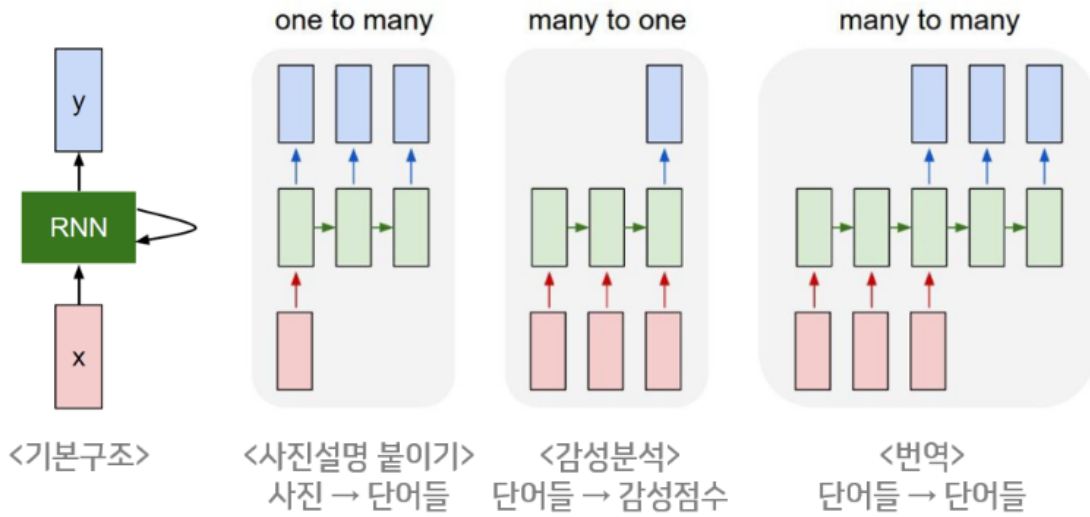
- x : input layer의 입력 벡터, y : output layer의 출력 벡터
- cell : RNN의 은닉층에서 활성화 함수를 통해 결과를 내보내는 역할을 하는 노드. 이전의 값을 기억하려고 하는 일종의 **메모리 역할**을 수행. (**memory cell** or **RNN cell** 이라고도 불림)
- **hidden state** : memory cell이 출력층 방향으로 또는 다음 $t+1$ 자신에게 보내는 값



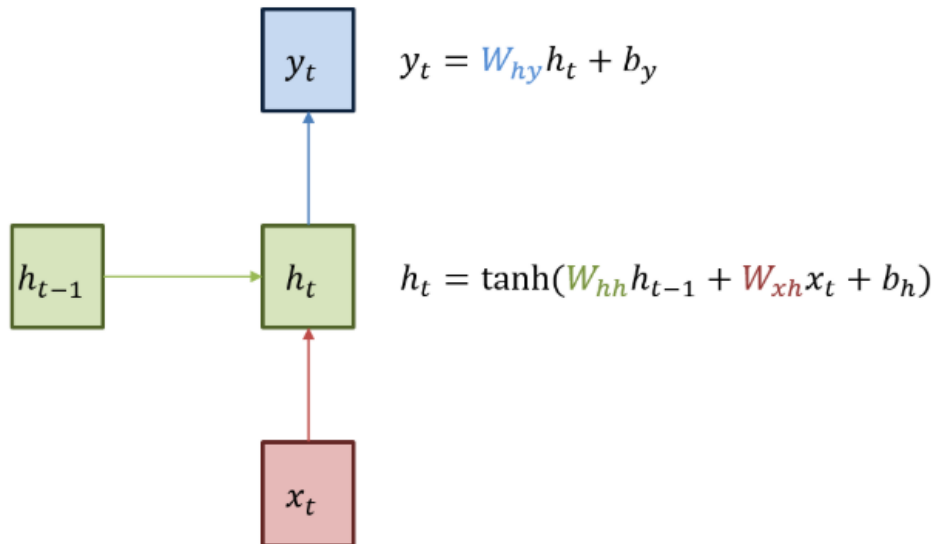
RNN의 다양한 용도

RNN은 입력과 출력의 길이를 다르기 설계 가능 → 다양한 용도로 사용 가능

- one-to-many(일 대 다) : 하나의 입력에 대해서 여러개의 출력
- many-to-one(다 대 일) : 단어 시퀀스에 대해서 하나의 출력
- many-to-many(다 대 다) : 단어 시퀀스에 대해서 여러개의 출력



RNN 연산 과정



- h_t : 현재 시점 t 에서의 hidden state
- W_{xh} : 입력값을 위한 가중치, W_{hh} : 이전 시점 $t-1$ 의 은닉 상태값인 h_{t-1} 을 위한 가중치
- 은닉층 수식 : $h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t + b_h)$
- 출력층 수식 : $y_t = f(W_{hy}h_t + b_y)$
 - f 는 비선형 활성화 함수 중 하나
- 각각의 가중치 W_{hh} , W_{xh} , W_{yh} 의 값은 모든 시점에서 값을 동일하게 공유
- 만약, 은닉층이 2개 이상일 경우 은닉층 2개의 가중치는 서로 다름

RNN 은닉층 연산 → 벡터와 행렬 연산으로 이해하기

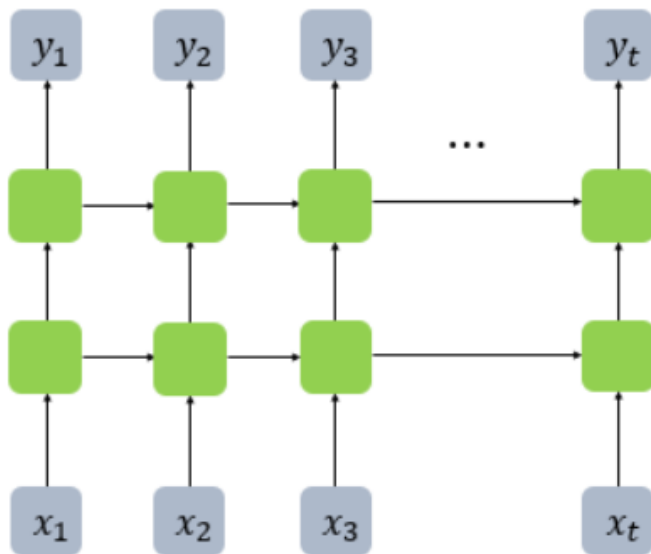
$$\tanh \left(\begin{matrix} W_h \\ D_h \times D_h \end{matrix} \times \begin{matrix} h_{t-1} \\ D_h \times 1 \end{matrix} + \begin{matrix} W_x \\ D_h \times d \end{matrix} \times \begin{matrix} x_t \\ d \times 1 \end{matrix} + \begin{matrix} b \\ D_h \times 1 \end{matrix} \right) = \begin{matrix} h_t \\ D_h \times 1 \end{matrix}$$

- d : 단어 벡터의 차원, D_h : 은닉 상태의 크기
- $x_t : (d \times 1)$
- $W_x : (D_h \times d), W_h : (D_h \times D_h)$
- $h_{t-1} : (D_h \times 1)$
- $b : (D_h \times 1)$

위 그림은 배치 크기가 1이고, d 와 D_h 두 값을 4로 가정하였을 때의 예제 그림

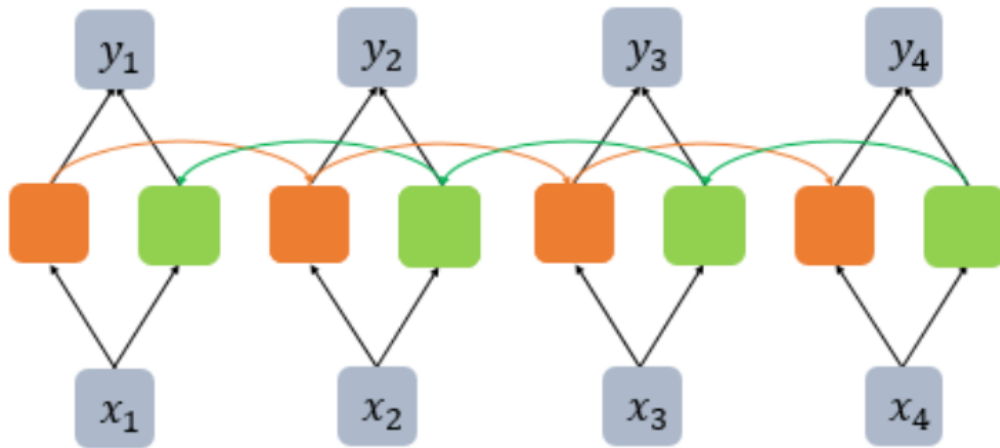
깊은 순환 신경망 (Deep RNN)

- 다수의 은닉층을 가지는 RNN
- example : 2개의 은닉층을 가진 deep RNN



양방향 순환 신경망 (Bidirectional RNN)

- 양방향 순환 신경망 : 시점 t 에서의 출력값을 예측할 때 이전 시점의 데이터뿐만 아니라, **이후 데이터**로도 예측할 수 있다는 아이디어.



- 하나의 출력값을 예측하기 위해 2개의 메모리 셀 사용
 - 주황색 메모리 셀 : 앞 시점의 은닉 상태(forward states)를 전달받아 현재의 은닉 상태 계산 → 이전 시점의 데이터를 이용하여 현재 시점(t)의 hidden state 계산
 - 초록색 메모리 셀 : 뒤 시점의 은닉 상태(backward states)를 전달받아 현재의 은닉 상태 계산 → 향후 시점의 데이터를 이용하여 현재 시점(t)의 hidden state 계산
- 과거 시점의 데이터만 고려하는 것이 아니라 향후 시점의 데이터를 고려하여 문제를 풀 때 **더 많은 힌트**를 얻게 됨

출처

- <https://wikidocs.net/22886>
- <https://ratsgo.github.io/natural%20language%20processing/2017/03/09/rnnlstm/>