순환 신경망 (Recurrent Neural Network, RNN)

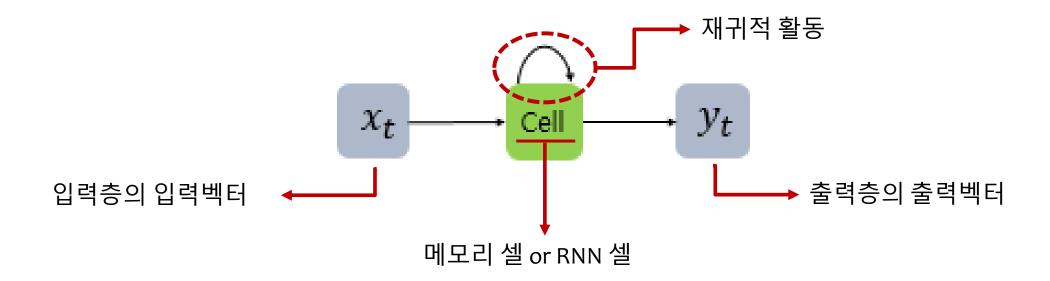
이하은 (20192939)

OO 순환 신경망 (Recurrent Neural Network, RNN)



순환 신경망 (Recurrent Neural Network, RNN)

; 입력과 출력을 시퀀스 단위로 처리하는 모델 (딥러닝에 있어서 가장 기본적인 시퀀스 모델!)

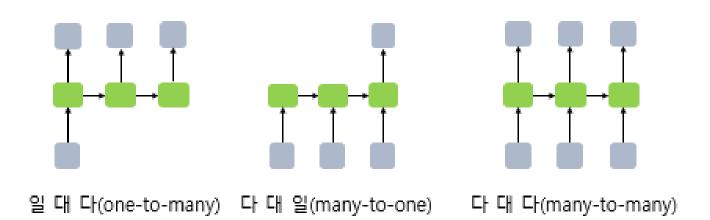


○ ○ 순환 신경망 (Recurrent Neural Network, RNN)



순환 신경망 (Recurrent Neural Network, RNN)

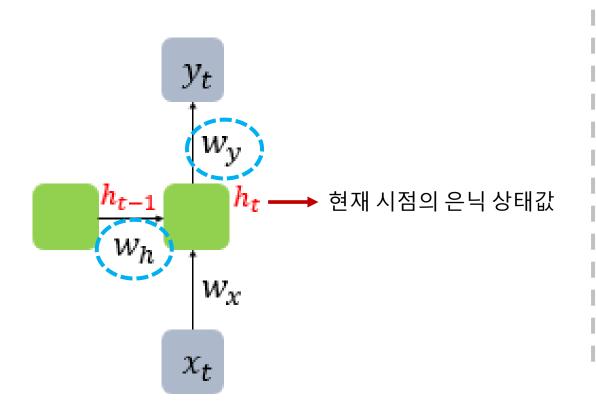
; 입력과 출력을 시퀀스 단위로 처리하는 모델 (딥러닝에 있어서 가장 기본적인 시퀀스 모델!)



OO 순환 신경망 (Recurrent Neural Network, RNN)



RNN 수식 정의



은닉층: $h_t = \tanh(W_x x_t + W_h h_{t-1} + b)$

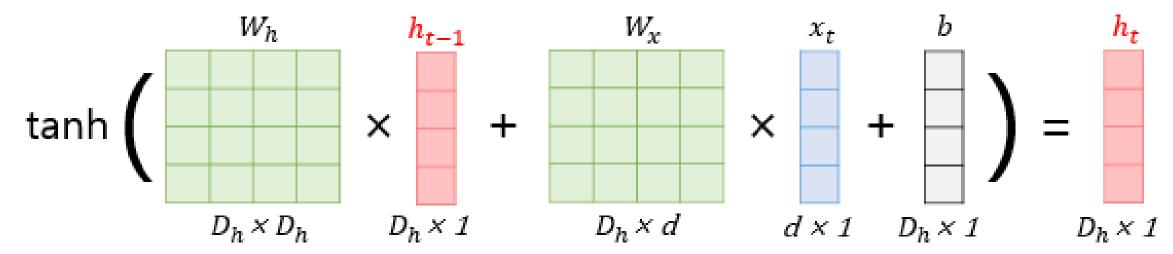
출력층: $y_t = f(W_y h_t + b)$

OO 순환 신경망 (Recurrent Neural Network, RNN)



RNN 수식 정의

$$egin{aligned} x_t : (d imes 1) \ W_x : (D_h imes d) \ W_h : (D_h imes D_h) \ h_{t-1} : (D_h imes 1) \ b : (D_h imes 1) \end{aligned}$$

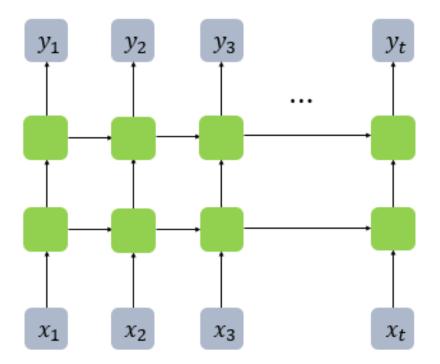


OOQ 깊은 순환 신경망 (Deep Recurrent Neural Network) C



깊은 순환 신경망

; 순환 신경망 RNN에서 은닉층이 다수인 경우

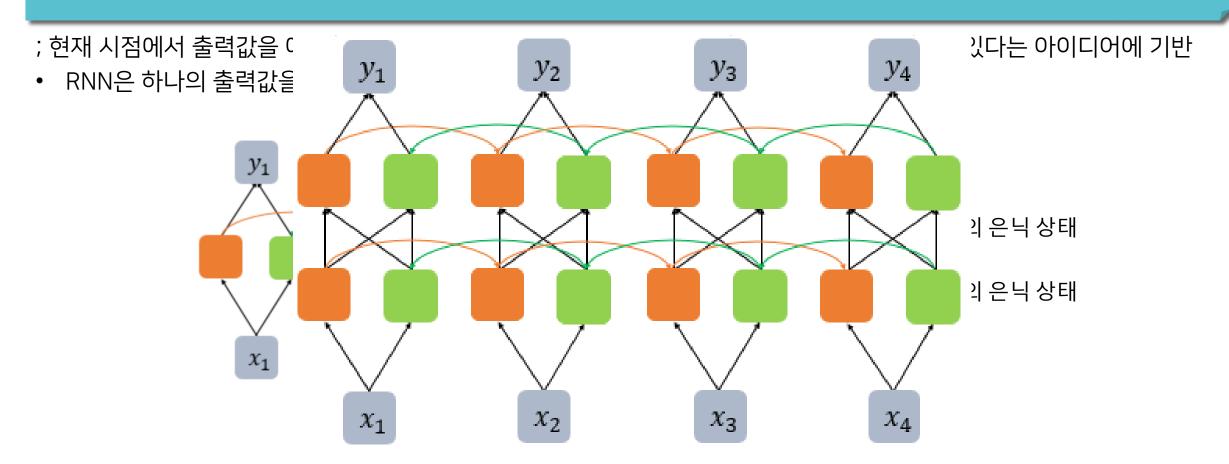




양방향 순환 신경망

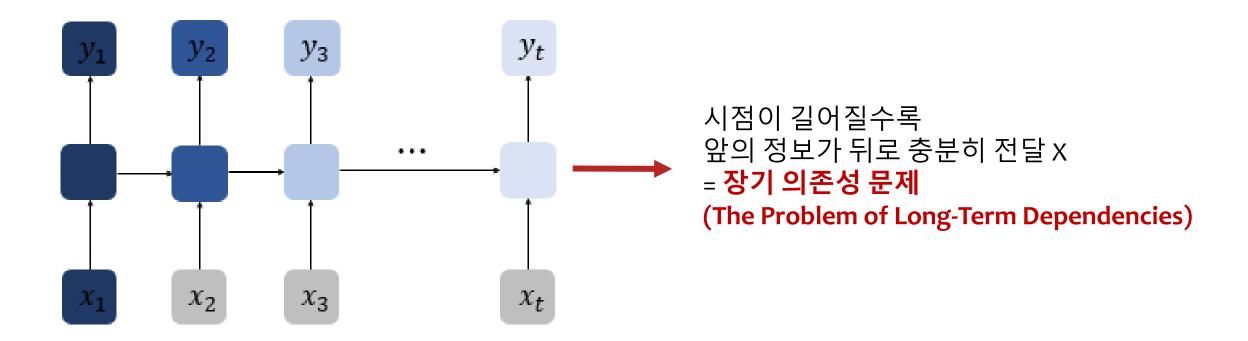


양방향 순환 신경망 (Bidirectional Recurrent Neural Network)





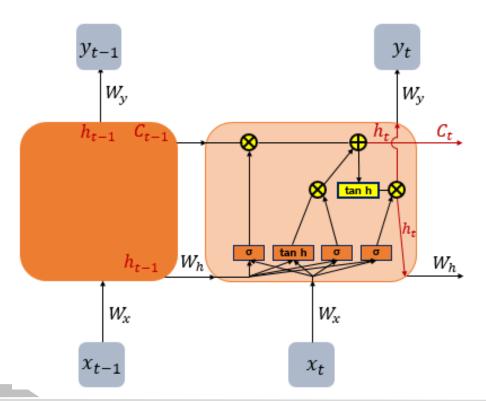
바닐라 RNN의 한계





LSTM (Long Short Term Memory)

; 은닉층에 메모리 셀에 입력 게이트, 망각 게이트, 출력 게이트를 추가하여 불필요한 기억을 지우고 기억해야할 것들을 정함 => 기존 RNN에서 셀 상태 값이 추가됨.



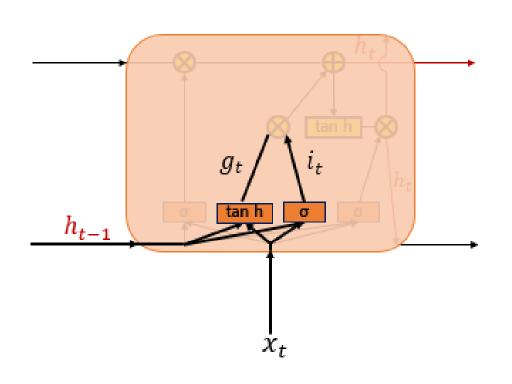
추가된 게이트

- 입력 게이트 시그모이드 함수 존재
- 출력 게이트 -



LSTM (Long Short Term Memory) – 입력 게이트

; 현재 정보를 기억하기 위한 게이트

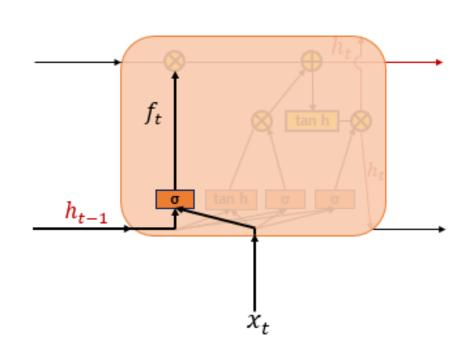


$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + b_i) \ g_t = tanh(W_{xg}x_t + W_{hg}h_{t-1} + b_g)$$



LSTM (Long Short Term Memory) – 삭제 게이트

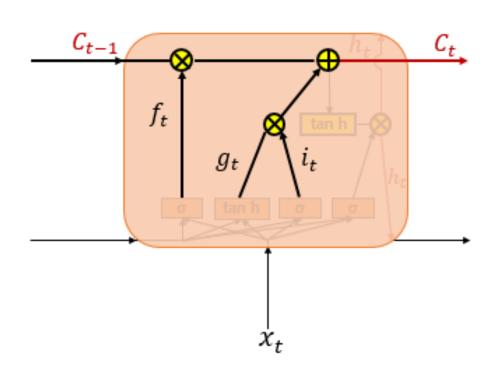
; 기억을 삭제하기 위한 게이트



$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + b_f)$$



LSTM (Long Short Term Memory) – 셀 상태(장기 상태)

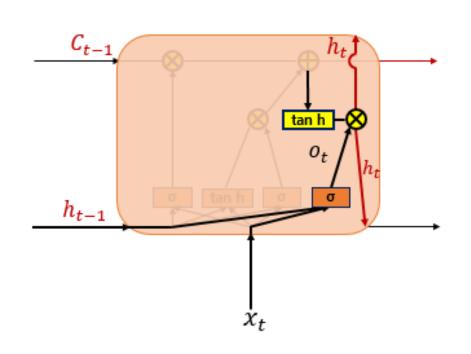


$$C_t = f_t \circ C_{t-1} + i_t \circ g_t$$



LSTM (Long Short Term Memory) – 출력 게이트

; 현재 시점의 은닉 상태를 결정



$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o)$$

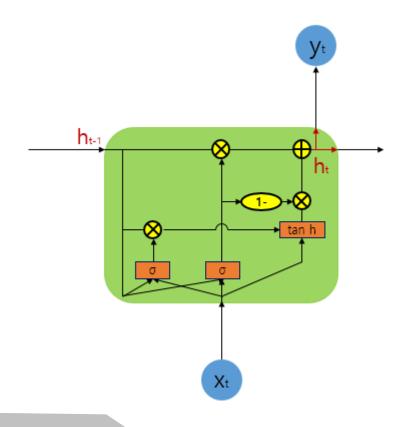
 $h_t = o_t \circ tanh(c_t)$

OOO게이트 순환 유닛 (Gated Recurrent Unit, GRU)



GRU (Gated Recurrent Unit)

; LSTM과 달리 업데이트 게이트와 리셋 게이트 두 가지 게이트만 존재



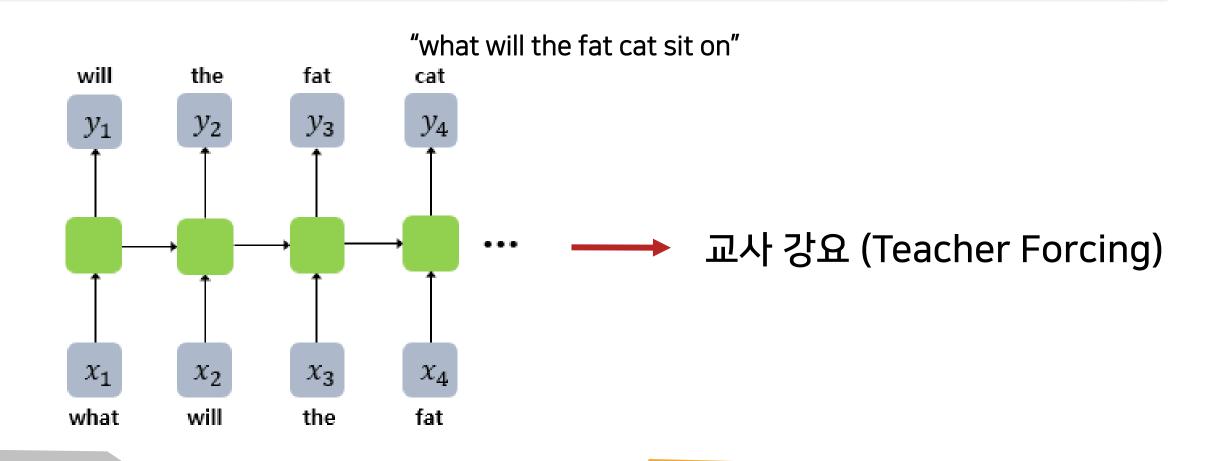
데이터 양이 적은 경우: GRU > LSTM 데이터 양이 많은 경우: GRU < LSTM



RNN 언어모델 (RNNLM)



RNN 언어 모델 (Recurrent Neural Network Language Model, RNNLM)

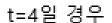


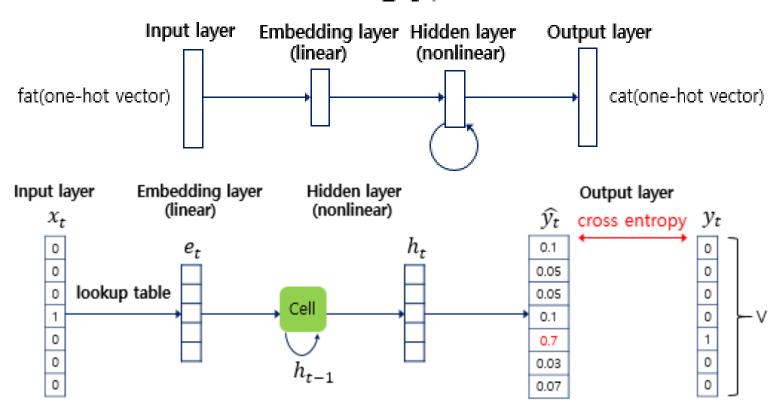


RNN 언어모델 (RNNLM)



RNN 언어 모델 (Recurrent Neural Network Language Model, RNNLM)







RNN을 이용한 텍스트 생성



RNN을 이용한 텍스트 생성 (Text Generation using RNN)

- 경마장에 있는 말이 뛰고 있다
- 그의 말이 법이다
- 가는 말이 고와야 오는 말이 곱다



samples	X	y
1	경마장에	있는
2	경마장에 있는	말이
3	경마장에 있는 말이	뛰고
4	경마장에 있는 말이 뛰고	있다
5	그의	말이
6	그의 말이	법이다
7	가는	말이
8	가는 말이	고와야
9	가는 말이 고와야	오는
10	가는 말이 고와야 오는	말이
11	가는 말이 고와야 오는 말이	곱다



글자 단위 RNN



글자 단위 RNN 언어 모델 (Char RNNLM)

; 기존의 RNN은 단위가 단어 벡터 => 단어 레벨에서 글자 레벨로 변경

