Chapter

5

RNN

1. RNN(Recurrent Neural Network)
2. LSTM(Long-Short Term Memory)
3. LSTM 모델을 사용한 날씨 예측(Keras)

강의에 앞서서..

❖ 본 문서는 아래의 자료들을 활용하여 만들어 졌음을 알립니다

❖ 모두를 위한 딥러닝 강좌

- 네이버 Search & Clova AI 부분 리더 김성훈 교수님
- https://www.youtube.com/playlist?list=PLlMkM4tgfjnLSOjrEJN31gZATbcj_MpUm
- https://www.edwith.org/boostcourse-dl-tensorflow/lecture/43739/

❖ 스탠포드 대학 CNN 강좌

- Fei-Fei Li & Andrej Karpathy & Justin Johnson
- http://cs231n.stanford.edu/slides/2020/

CS231n: Convolutional Neural Networks for Visual Recognition
- This course, Prof. Fei-Fei Li & Justin Johnson & Serena Yeung
- Focusing on applications of deep learning to computer vision

강의에 앞서서..

❖ Hands on Machine Learning

- https://github.com/ExcelsiorCJH/Hands-On-ML/blob/master/Chap14-Recurrent_Neural_Networks/Chap14_1-Recurrent_Neural_Networks.md
- https://excelsior-cjh.tistory.com/185

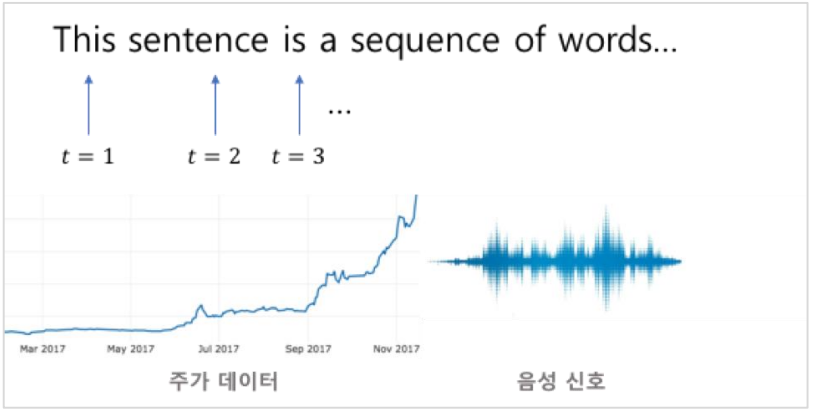
❖ Understanding LSTM Networks

- https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Recurrent Neural Network

❖ 순환 신경망(RNN, Recurrent Neural Network)

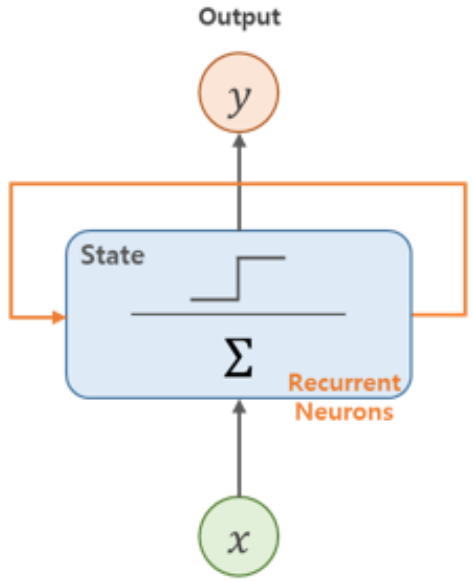
- Sequencial data를 위한 모델
- 음성인식, 자동변역, 주가예측
- Sequencial data
  - 데이터 집합 안에 순서를 가지는 데이터
  - 예 , "Hello" : H,e,l,l,o의 순서로 구성
  - 종류: **x = (x0, x1,...xt, ...) , t : time**
    - 시계열(time serie) : 시간을 기준으로 생성된 연속 데이터
      - » 주가변동, 온도변화, 질병발생현황, 센서신호데이터
    - 자연어(NL, Natural Language)
    - 음성신호
    - DNA염기서열

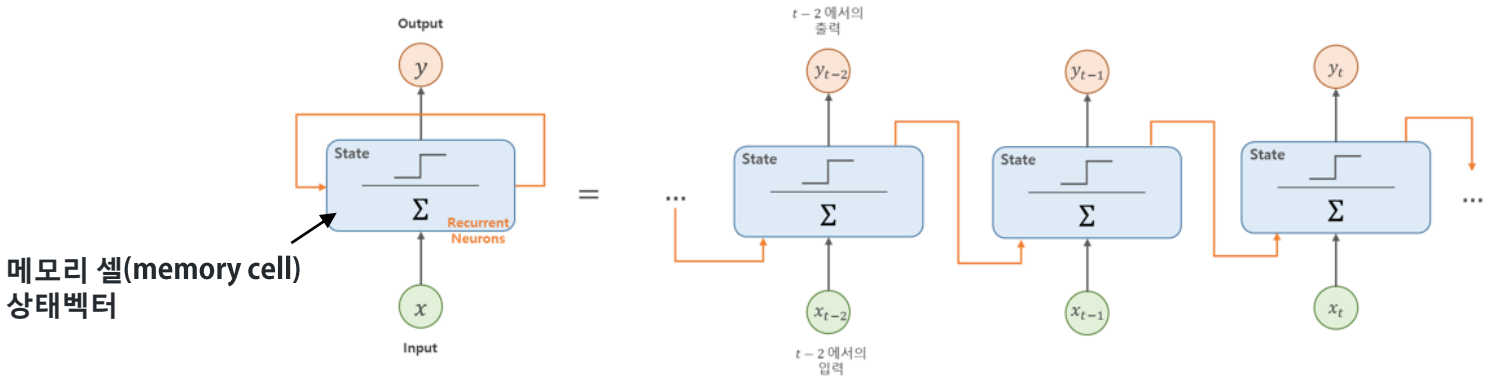


Recurrent Neural Network

❖ 순환 신경망(RNN, Recurrent Neural Network`

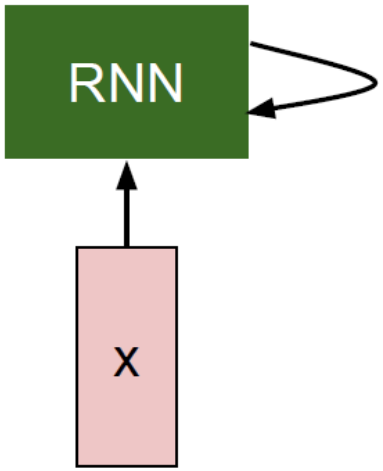
- 순환 뉴런(Recurrent Neurons)으로 Layer 구성
- 출력이 다시 입력으로 순환되는 형태
- 이전 상태가 현재상태에 영향을 주는 체인구조
- 타임스텝(time step)마다 순환 뉴런이 펼쳐짐





Recurrent Neural Network

❖ 순환뉴런의 상태값 계산



$$h_t = f_W(h_{t-1}, x_t)$$

new state

some function with parameters W

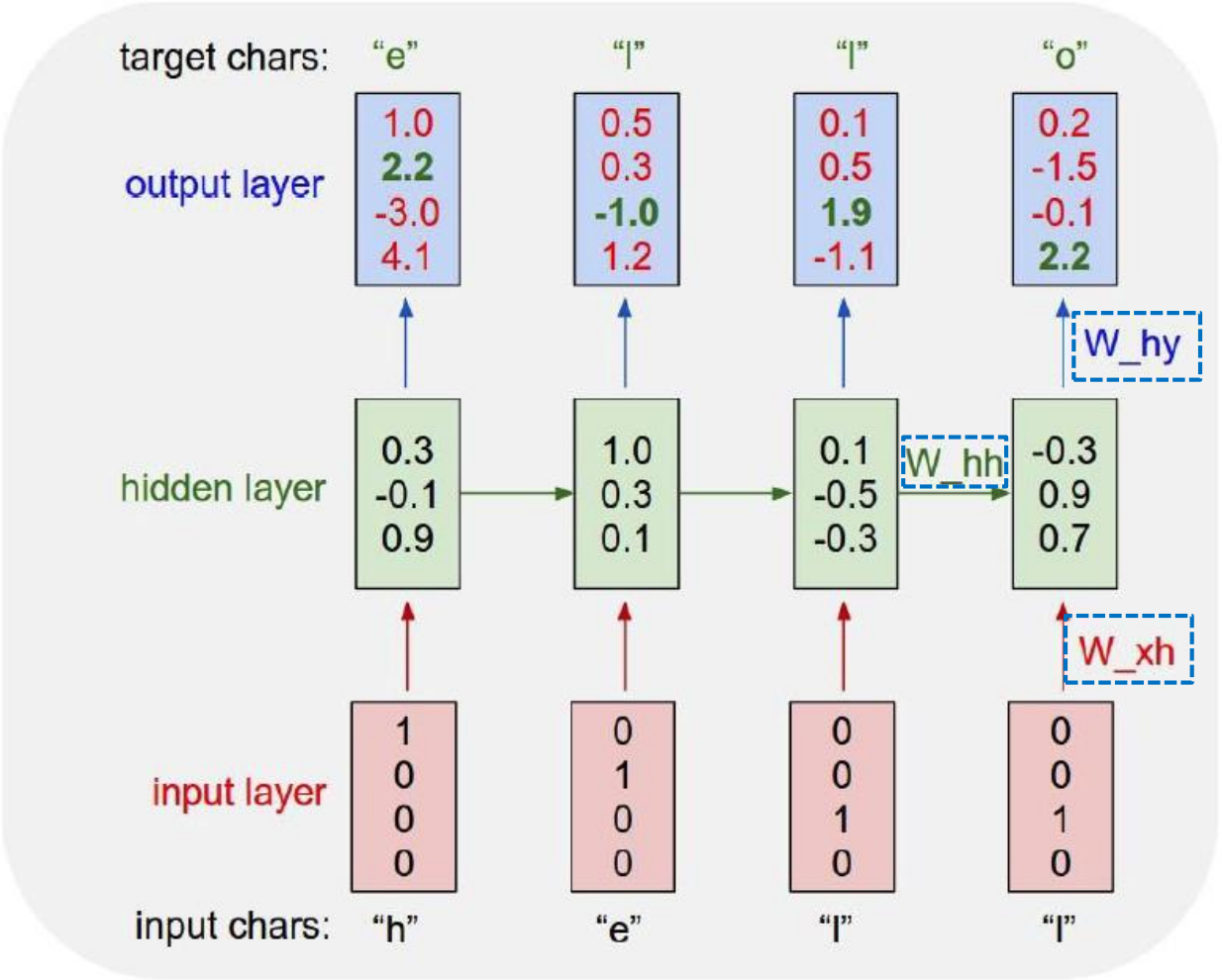
old state

input vector at some time step

$$h_t = \tanh(W_{hh}h_{t-1} + W_{xh}x_t + b_h)$$

Recurrent Neural Network

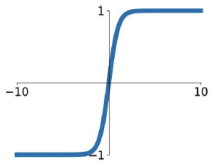
'hello' 라는 단어의 순환뉴런 상태 값 계산 과정



$$y_t = W_{hy} h_t$$

$$h_t = \tanh(W_{hh} h_{t-1} + W_{xh} x_t)$$

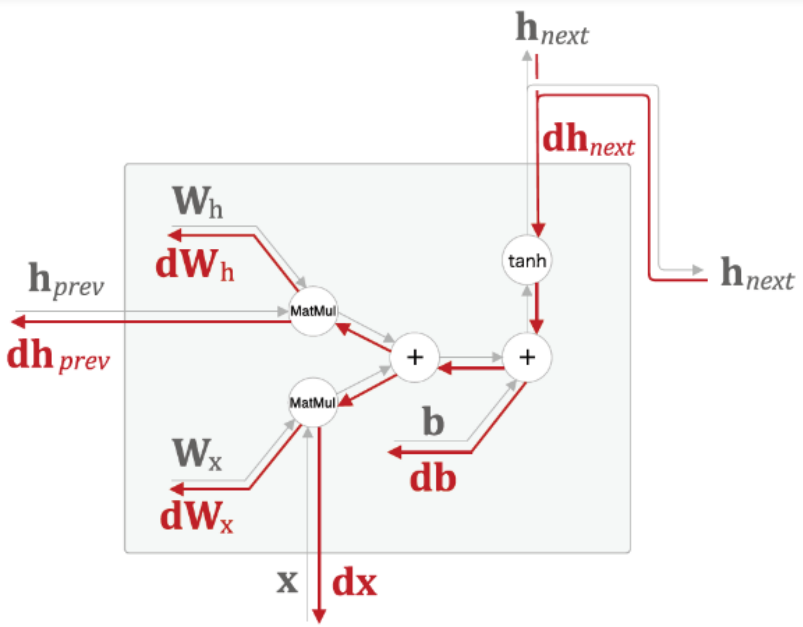
$\tanh(x)$



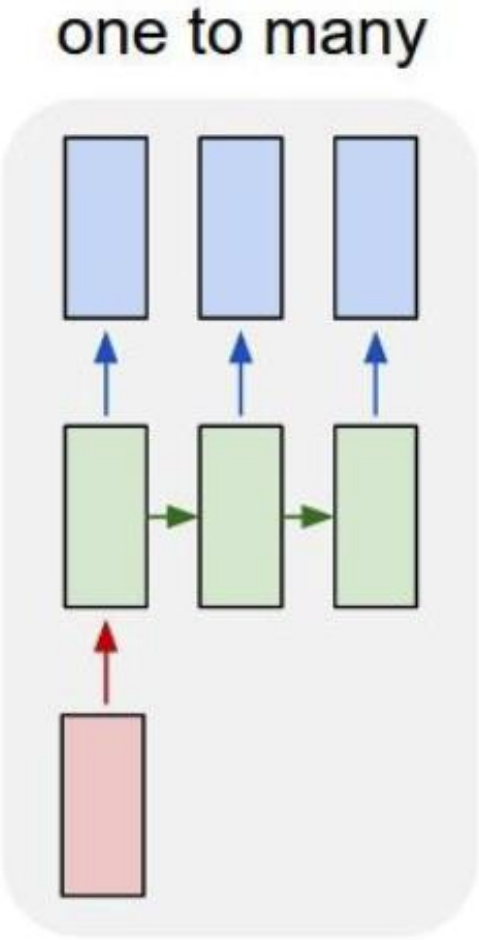
Recurrent Neural Network

❖ RNN 계층 계산 그래프

- Forward propagation
  - 이전 상태 값 (h_{prev})에 가중치(W_h)를 곱하고 입력(x)에 가중치((W_x) 를 곱한 값을 결합하고 바이어스를 결합하여 탄젠트 활성화 함수로 다음 상태값 (h_{next}) 계산
- Back propagation
  - 다음 상태값 (h_{next})과 정답의 오차를 계산하여 역전파하면서 가중치와 상태값을 수정

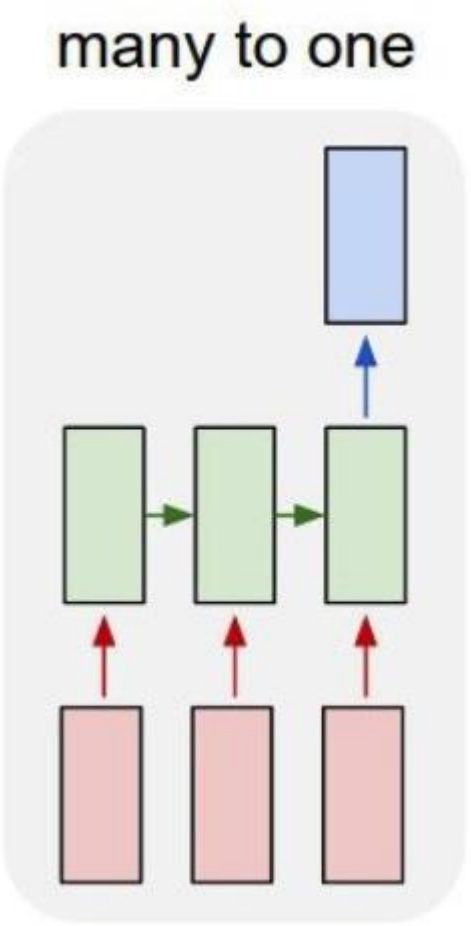


Recurrent Neural Network



e.g. **Image Captioning**
image -> sequence of words

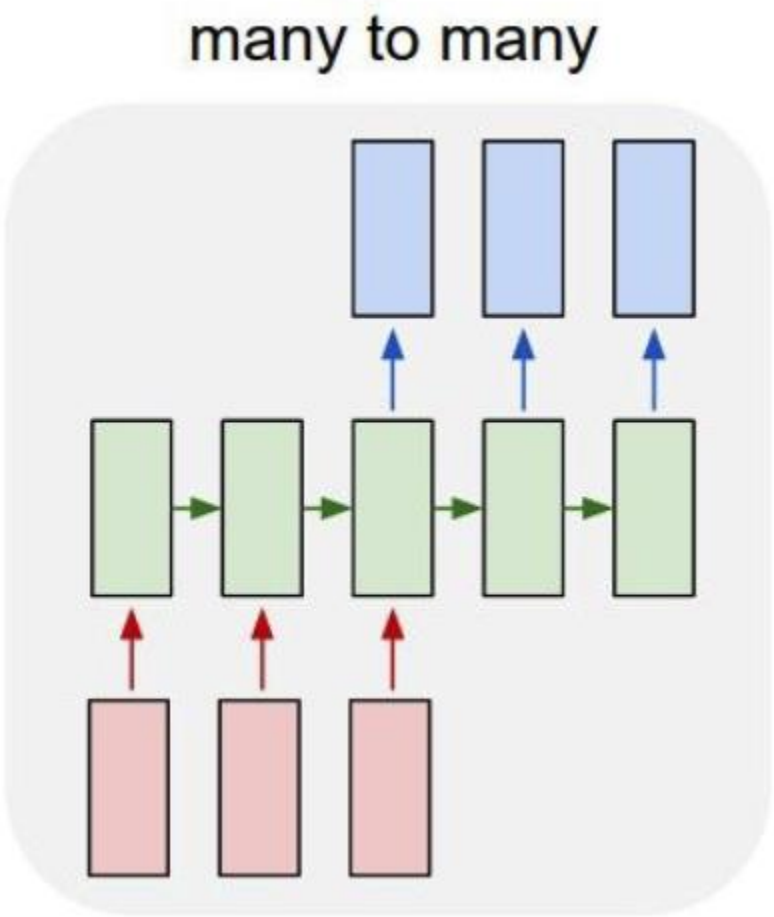
Recurrent Neural Network

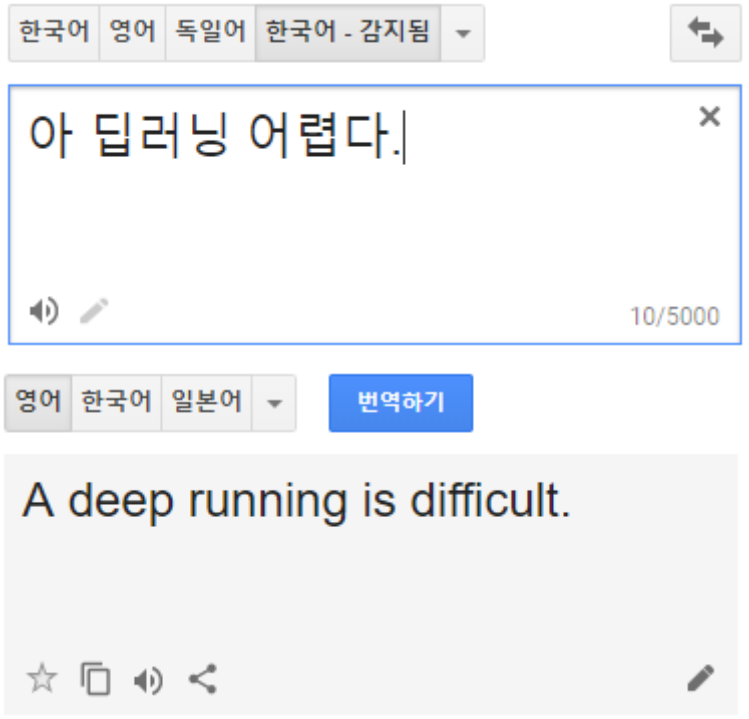


e.g. **Sentiment Classification**
sequence of words -> sentiment

https://blog.webhose.io/2017/02/09/how-to-use-rated-reviews-for-sentiment-classification/

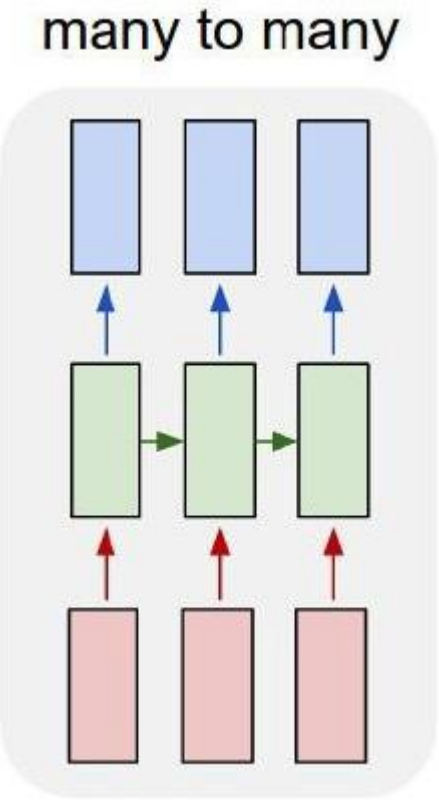
Recurrent Neural Network

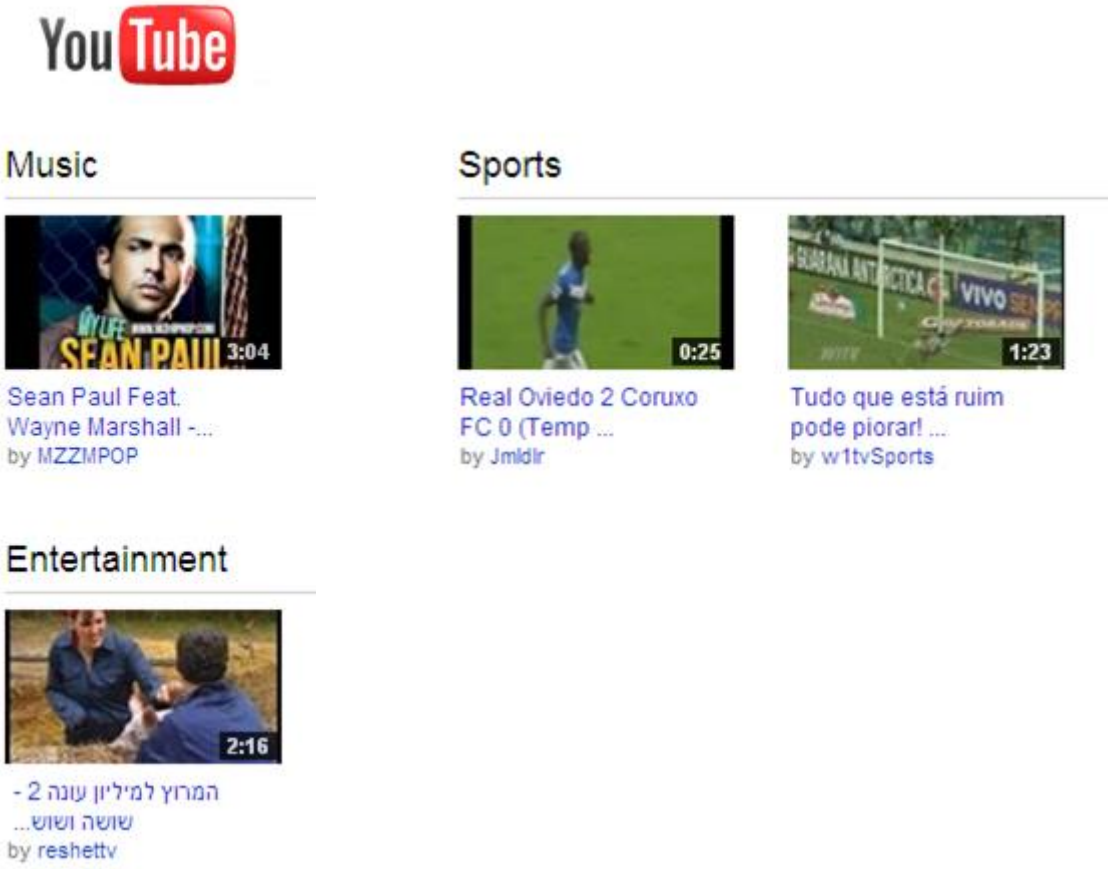




e.g. **Machine Translation**
seq of words -> seq of words

Recurrent Neural Network





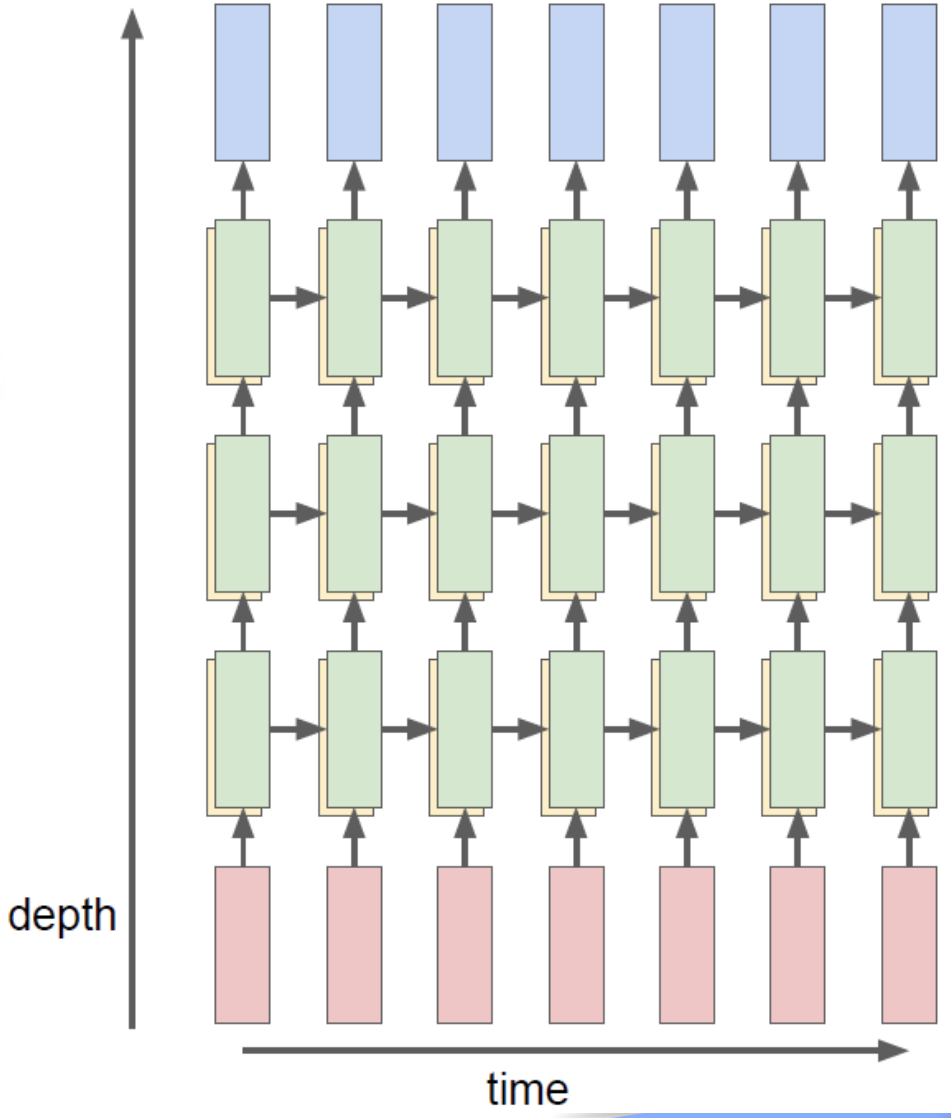
e.g. **Video classification on frame level**

Recurrent Neural Network

Multilayer RNNs

$$h_t^l = \tanh W^l \begin{pmatrix} h_t^{l-1} \\ h_{t-1}^l \end{pmatrix}$$

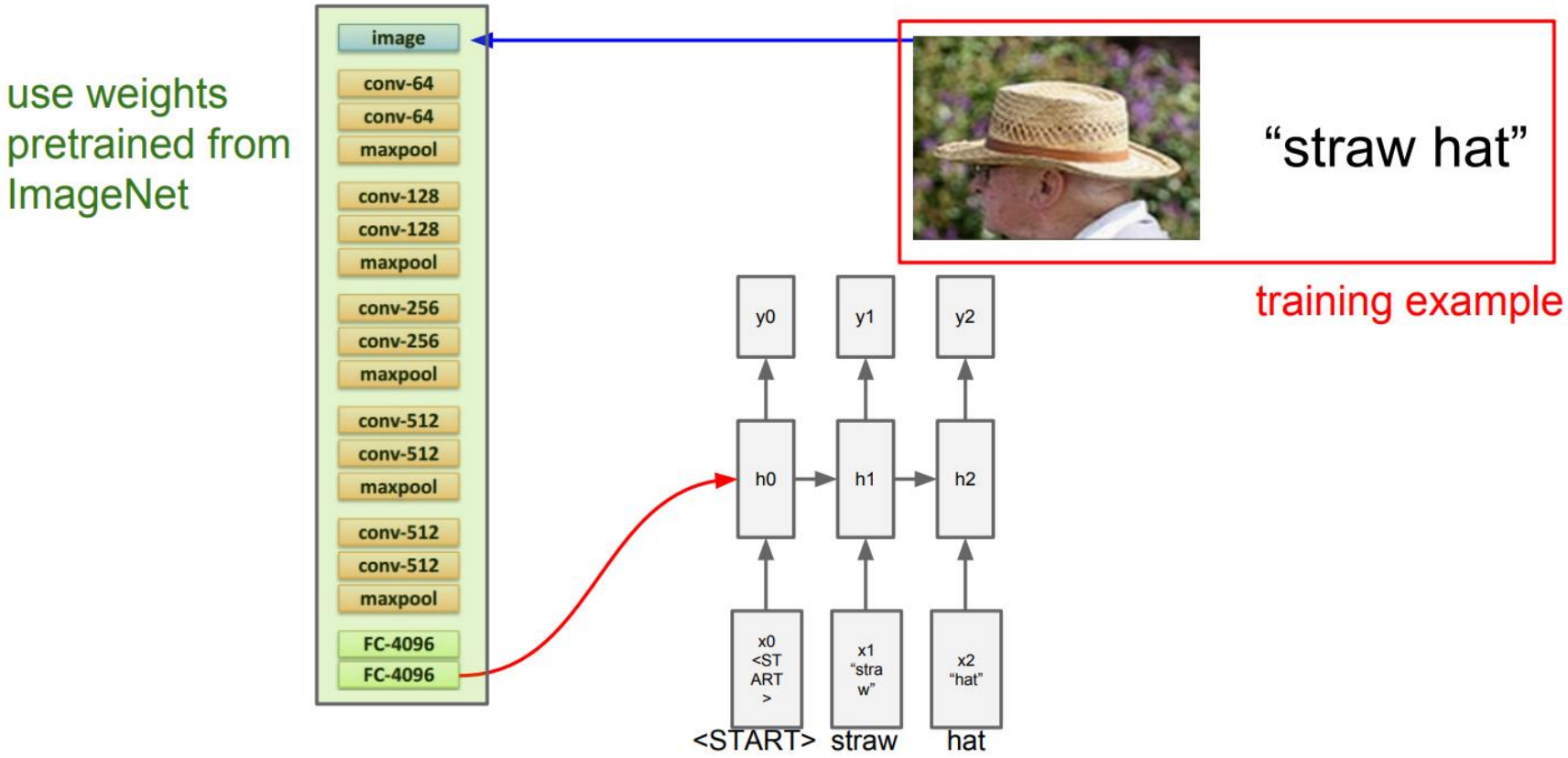
$h \in \mathbb{R}^n$. W^l $[n \times 2n]$



Recurrent Neural Network

❖ Transfer Learning

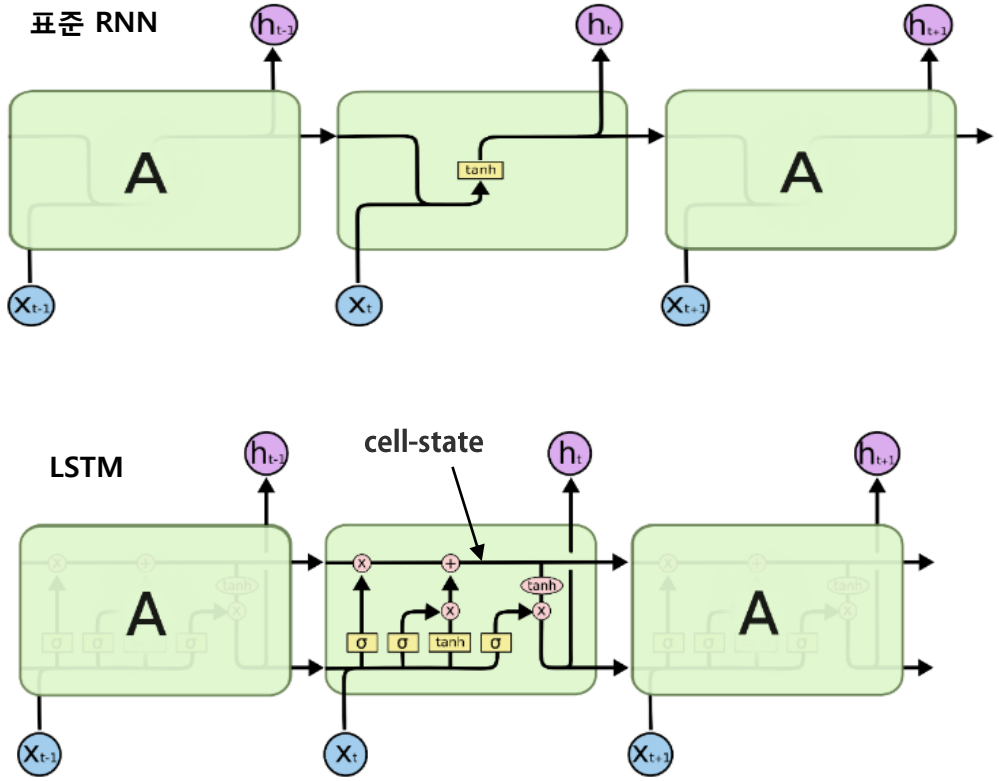
- CNN+RNN



LSTM

❖ LSTM(Long-Short Term Memory)

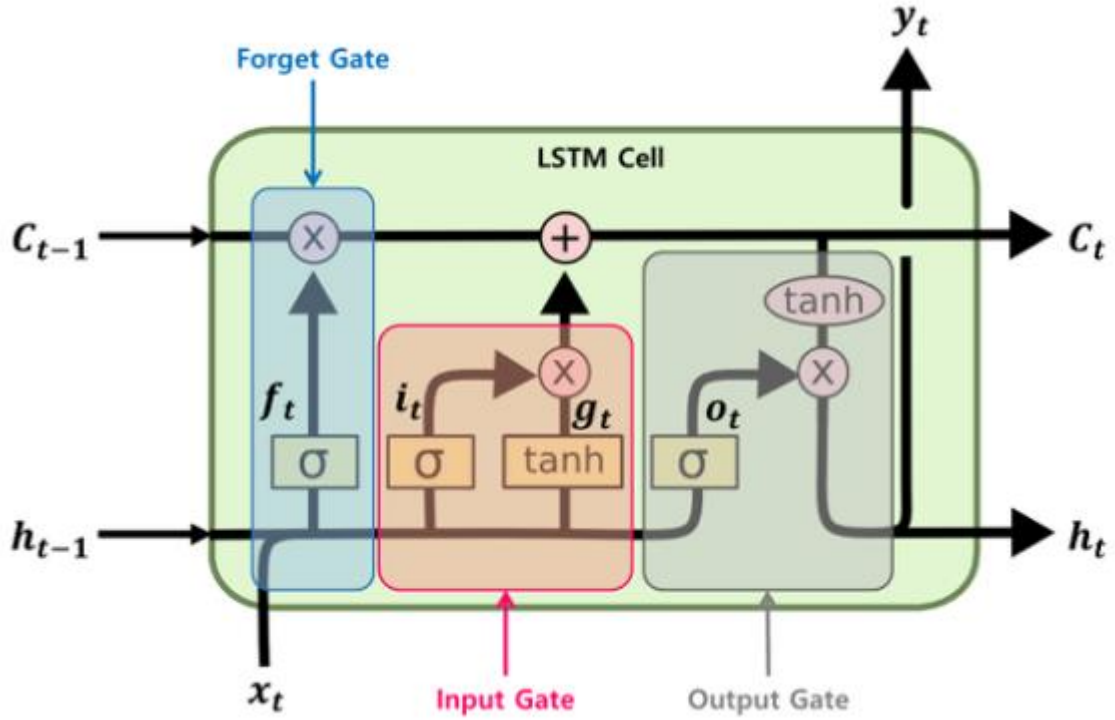
- RNN에서의 기울기 소실(Gradient Vanishing) 문제를 해결하고자 메모리를 도입한 모델



LSTM

❖ 두 개의 상태 벡터를 이용하여 장기, 단기 상태를 저장

- h : 단기 상태(short-term state),
- C : 장기 상태(long-term state)



$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] \; + \; b_f\right)$$

$$i_t = \sigma\left(W_i \cdot [h_{t-1}, x_t] \; + \; b_i\right)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] \; + \; b_C)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$

$$o_t = \sigma\left(W_o \; [h_{t-1}, x_t] \; + \; b_o\right)$$

$$h_t = o_t * \tanh\left(C_t\right)$$

https://excelsior-cjh.tistory.com/185

Keras를 사용한 RNN

- **keras.layers.SimpleRNN**
  - 이전 타임 스텝의 출력이 다음 타임 스텝으로 공급되는 완전 연결된 RNN.
- **keras.layers.LSTM**
  - 1997 년 Hochreiter & Schmidhuber 에서 처음 제안 .
  - RNN의 장기상태 소멸에 대한 문제점 보완 모델

Keras를 사용한 RNN

❖ LSTM 모델을 사용한 날씨 예측

- 날씨 시계열 데이터 사용
- 단일변수(온도)로 온도 예측
  - Lab14_1_time_series_univariate.ipynb)
- 다중변수(온도, 대기압 및 공기 밀도)로 온도 예측
  - Lab14_2_time_series_multivariate.ipynb
- https://www.tensorflow.org/tutorials/structured_data/time_series?hl=ko