

RNN

ML/DL Basic - Week 05

한유진

```
lookup.KeyValue  
f.constant(['em  
=tf.constant([0  
ce = tf.lookup.StaticV  
init,  
num_oov_buckets=5)
```

```
lookup.StaticVocabular  
initializer,  
num_oov_buckets,  
lookup_key_dtype=None  
name=None,  
experimental_is_open
```

Sequence Data

Sequence Data

: 순서(Sequence)가 있는 데이터

-> 특정 순서를 가지므로 독립적이지 않음

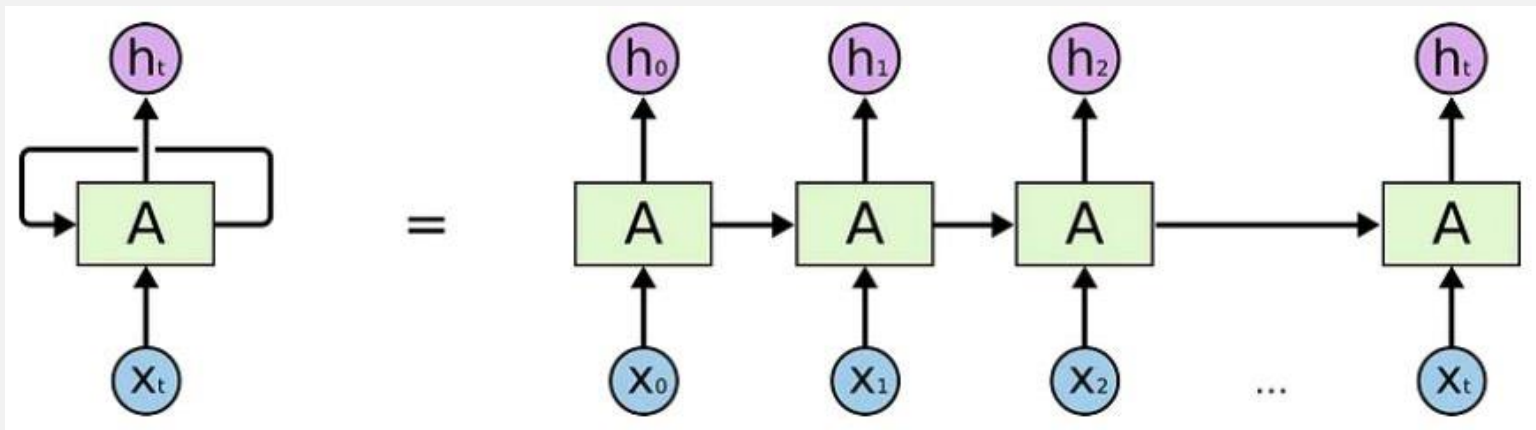
Ex) 시계열 데이터, 텍스트 데이터

일반적인 지도학습의 알고리즘은 입력 데이터가

‘independently identically distributed’라고 가정

-> 이 알고리즘에서 데이터의 순서는 상관 없음

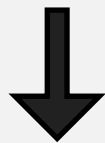
Recurrent Neural Network



**A의 결과가 다시 A로 들어가 루프를 만듦
-> 현재의 state가 다음 state에 영향을 미침**

Recurrent Neural Network

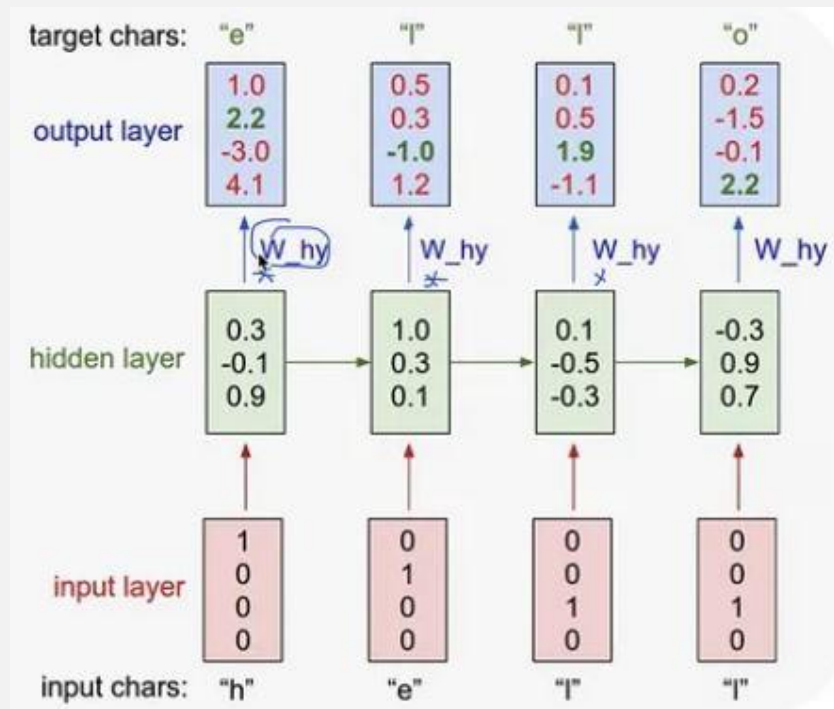
$$h_t = f_w(h_{t-1}, x_t)$$



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

$$y_t = W_{hy}h_t$$

Character - level language model example



Input Layer -> Hidden layer

- "hello": 4가지 종류의 글자
-> 크기가 4인 벡터로 처리
- Input을 위한 가중치와 계산 후 tanh 함수에 전달

Hidden layer -> Output layer

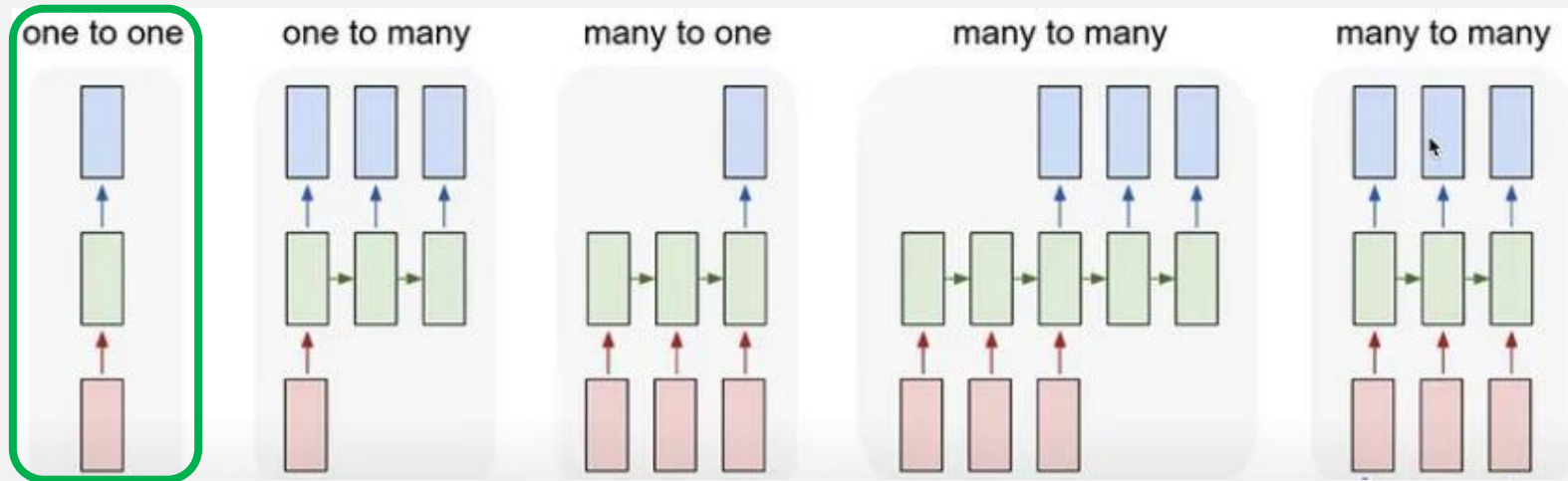
- Hidden 값과 가중치 계산
- One-hot encoding을 거치면 제일 큰 값을 결과로 가짐

Recurrent Neural Network

Applications

- Language Modeling
- Speech Recognition
- Machine Translation
- Conversation Modeling/Question Answering
- Image/Video Captioning
- Image/Music/Dance Generation

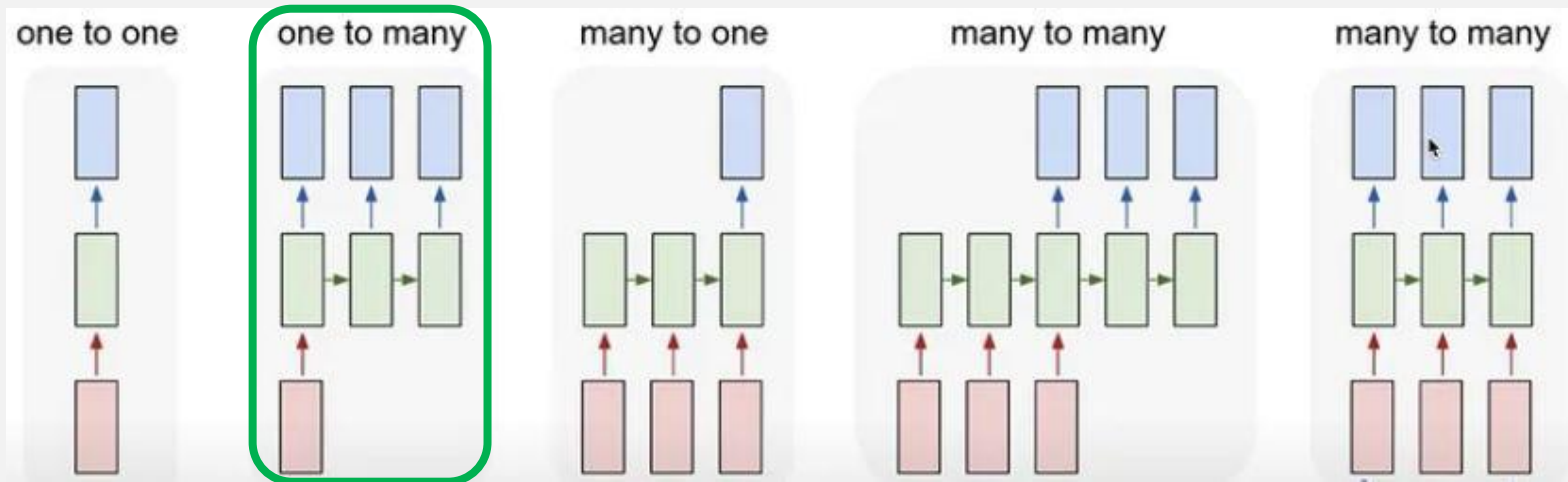
Recurrent Neural Network의 다양한 형태



한 개의 입력에 대해 한 개의 출력 생성

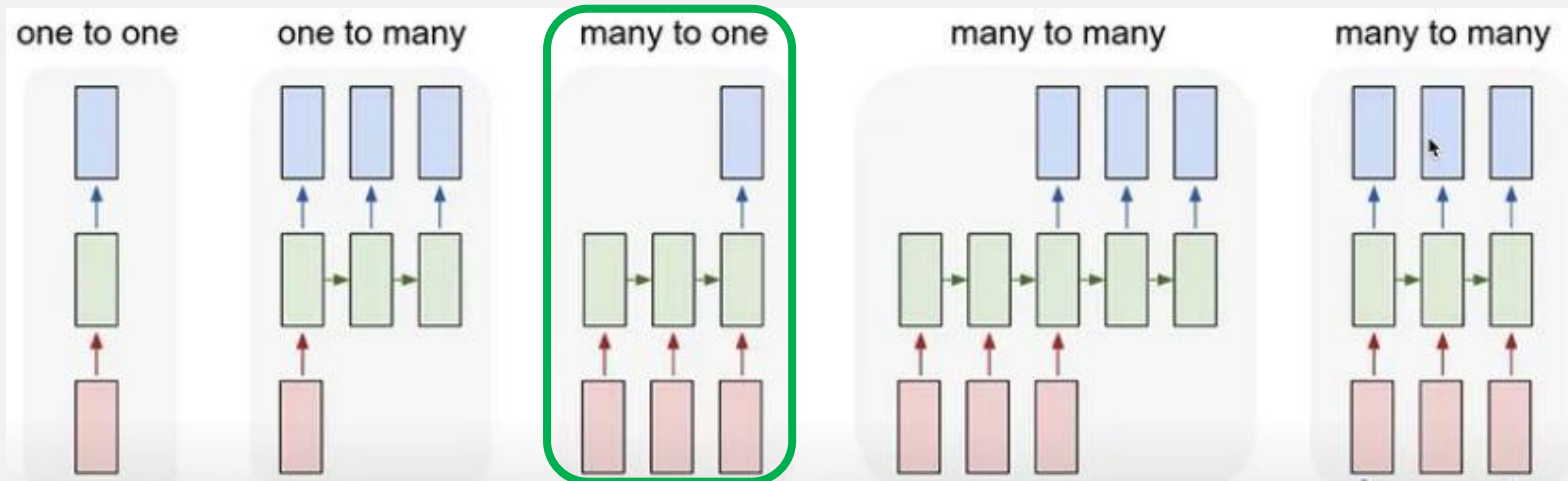
-> 간단한 기계 학습 문제에 사용, Vanilla Neural Network

Recurrent Neural Network의 다양한 형태



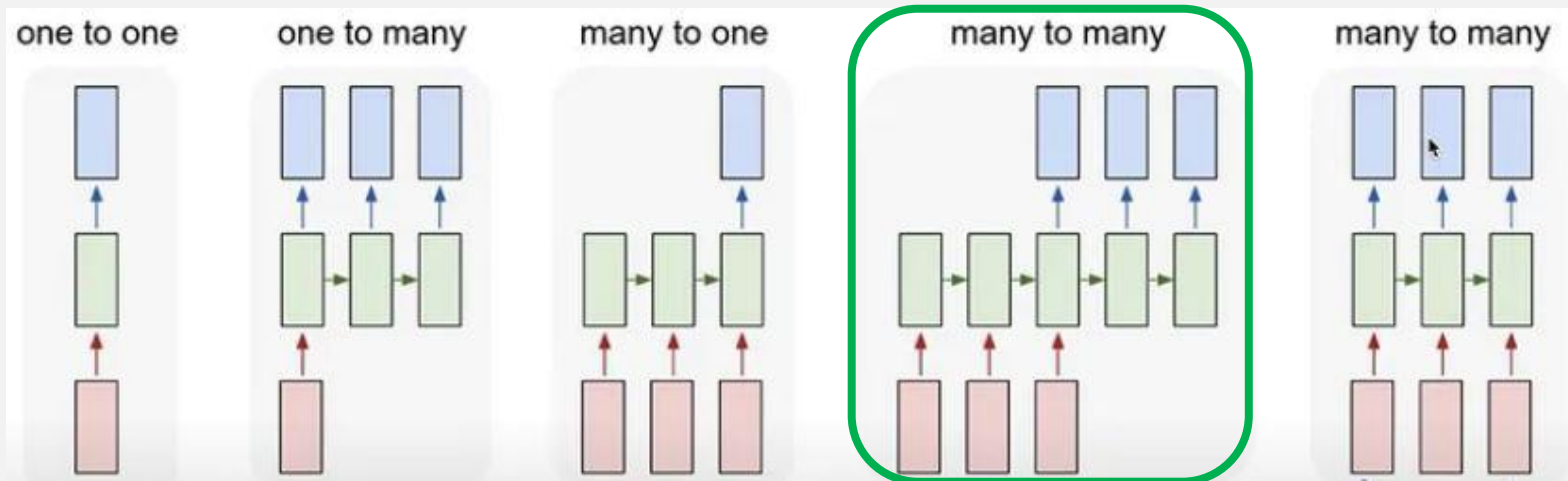
한 개의 입력에 대해 여러 개의 출력 생성
-> Image Captioning에 사용

Recurrent Neural Network의 다양한 형태



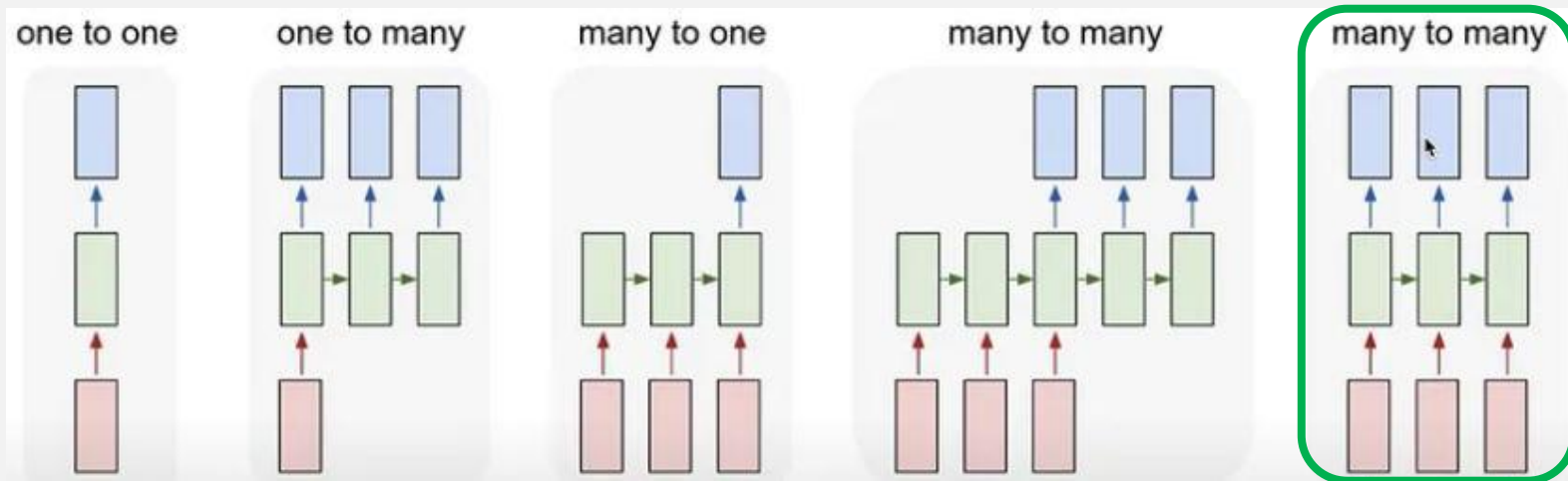
여러 개의 입력에 대해 한 개의 출력 생성
-> Sentiment Classification에 사용

Recurrent Neural Network의 다양한 형태



여러 개의 입력에 대해 여러 개의 출력 생성
모두 입력한 후 출력
-> Machine Translation에 사용

Recurrent Neural Network의 다양한 형태



여러 개의 입력에 대해 여러 개의 출력 생성
입력 받은 즉시 출력
-> Video Classification에 사용

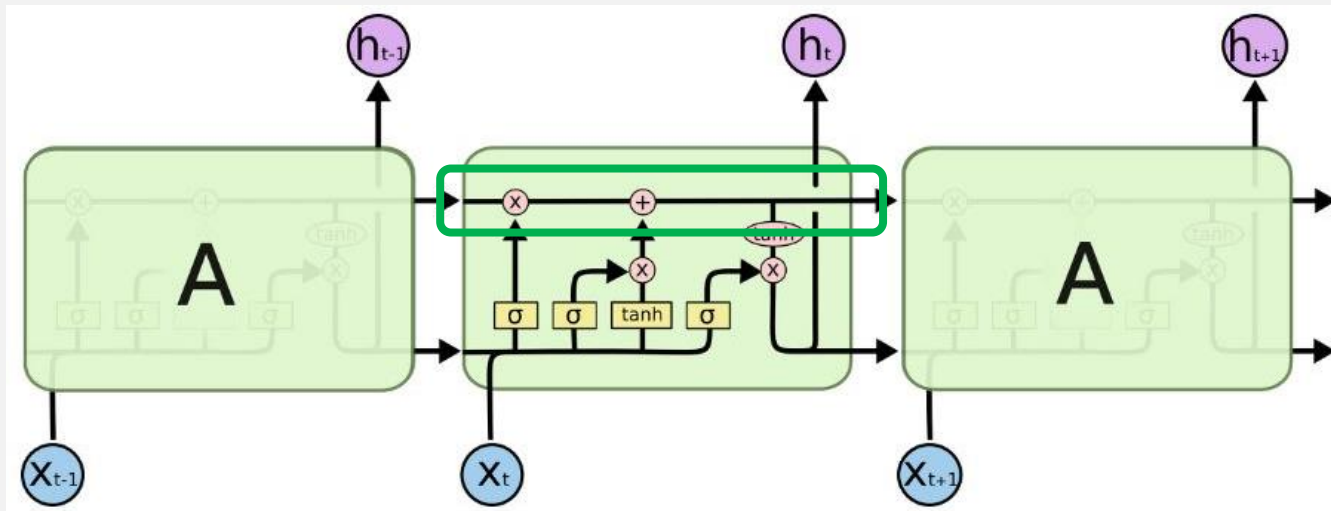
Recurrent Neural Network의 한계

Long Term Dependency Problem

: 시퀀스 데이터의 길이가 길어질수록, 과거의 중요한 정보에 대해 학습이 어려워지는 문제

-> LSTM, GRU와 같은 모델

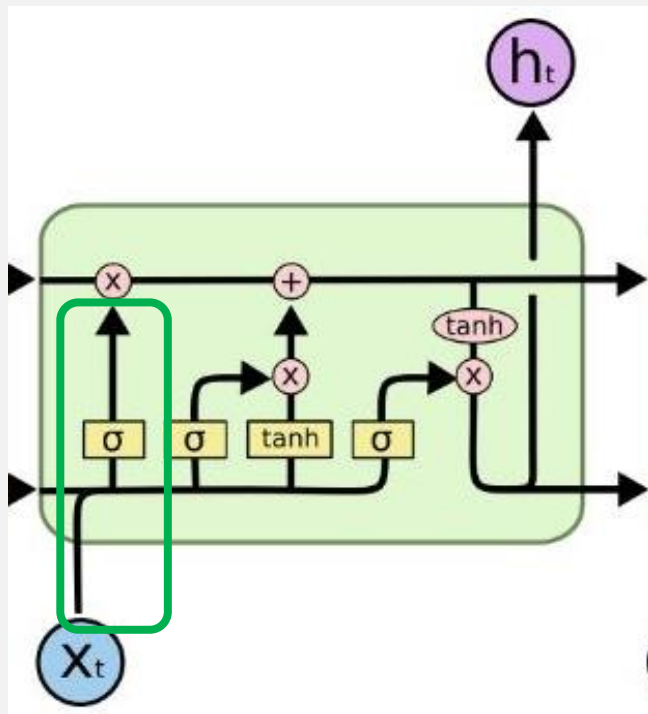
LSTM



Cell state

정보가 전혀 바뀌지 않고 그대로 흐르게 함

LSTM

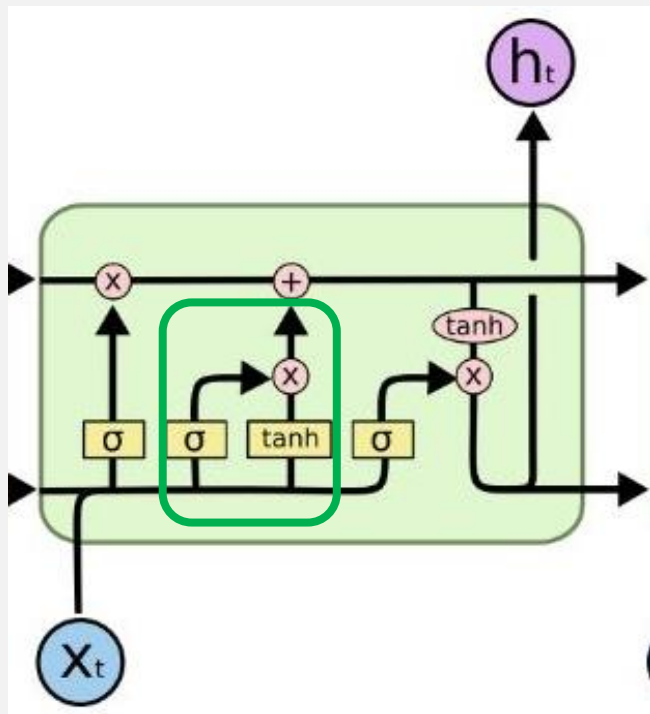


1. Forget gate

: 어떤 정보를 버릴 것인지 정하며
Sigmoid layer에 의해 결정됨

1이면 모든 정보를 보존,
0이면 모든 정보를 제거

LSTM



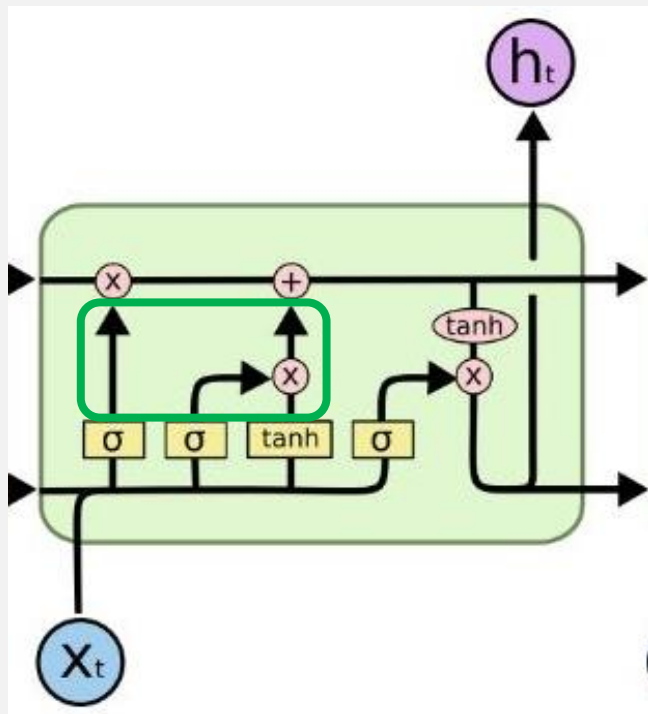
2. Input gate

: 어떤 새로운 정보를 Cell state에 저장/추가할 것인지 정함

1) Input gate layer에서 어떤 값을 업데이트할지 결정

2) tanh 층에서 state에 추가될 수 있는 후보 값 생성

LSTM

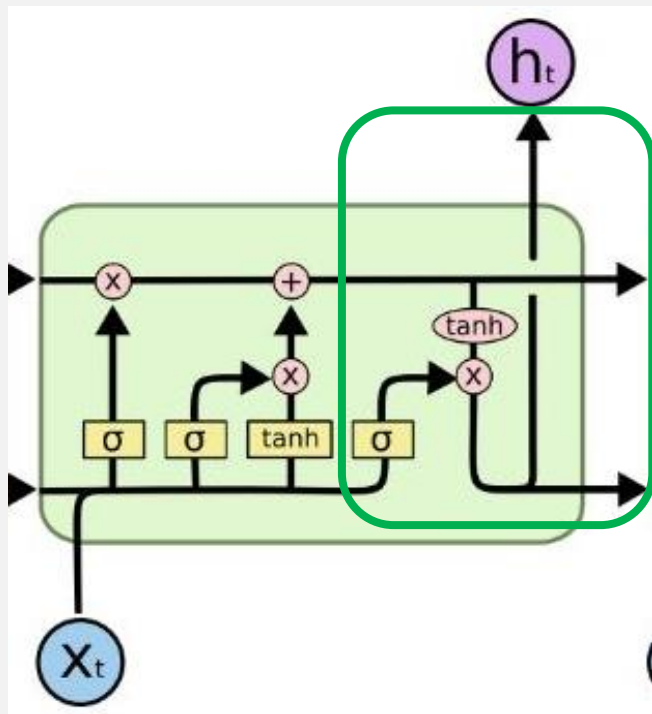


3. Cell State 업데이트

: 오래된 cell state C_{t-1} 에서
새로운 Cell state C_t 로 업데이트

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

LSTM



4. Output gate

: 무엇을 출력할지 결정

- 1) Sigmoid layer를 이용하여 cell state 중 어떤 부분을 출력할지 정함
- 2) Cell state를 \tanh 함수에 입력하고 위에서 구한 값과 곱해 출력할 부분 결정

Neural Weather Forecaster

기온 예측 모델 만들기(Neural Weather Forecaster)

개요	평가기준표	제출	리뷰 결과
----	-------	----	-------

리뷰어 지정 전

리뷰 #1 | -

 프로젝트 평가	 코드 리뷰	 수강생 메모	 리뷰 목록
---	---	--	---

리뷰	결과	리뷰일	리뷰어
리뷰 #1 (현재 리뷰)	리뷰어 지정 전	2023.11.01	-