1. rnn의 사용목적

:순차 데이터를 사용하기 위함

1. 순차데이터

순서에 의미가 있으며, 순서가 달라지면 의미가 손상되는 데이터

(시간적 의미) temporal sequence

1. resampling

temporal sequence -> time series

1. 심층 신경망과 순차데이터

수많은 히든레이어를 거쳐 one-hot 결과값을 내야함

다중입력 단일출력 – 시리

다중입력 다중출력 – 번역기

단일입력 다중출력 – 사진 묘사 ex) 고양이 밥먹는 사진 – ‘고양이가 밥을 먹는다’ 출력

1. 기본 rnn
2. 기억시스템

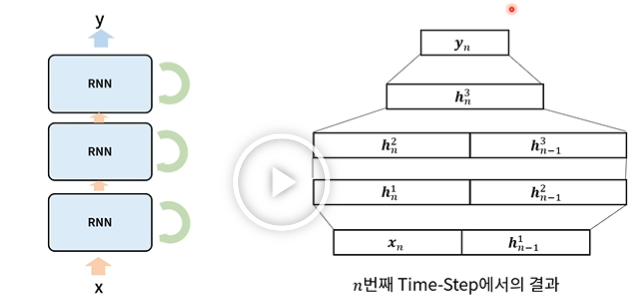
(cf) 무기억시스템

얕은 신경망

이전 스텝의 입/출력에 영향받지 않음

1. 순환 신경망

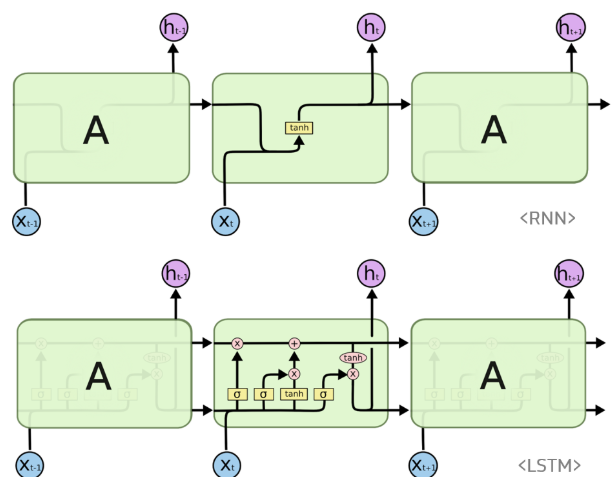
xn + hn-1 -> hn + yn

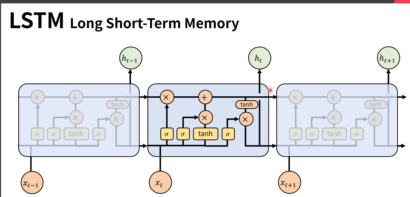
1. 다중계층순환 rnn

입력 시퀀스 길이 \* depth까지 계산해야해서 복잡

1. LSTM

(cf) rnn과 다른 점



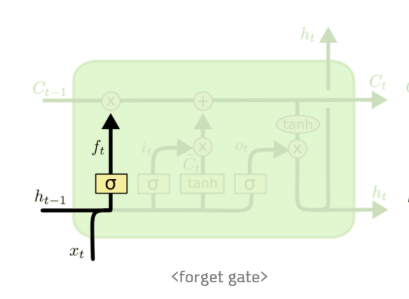


cell state – 기억을 오래 유지할 수 있는 구조

새로운 특징을 덧셈으로 받는 구조

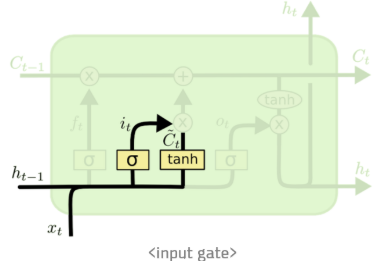
hidden state에서 출력과 다른 계층으로 넘김을 동시에 담당

forget gate



sigmoid 활성함수로서 0~1을 곱해줌 -> 잊고싶은 state는 0을 곱해줌

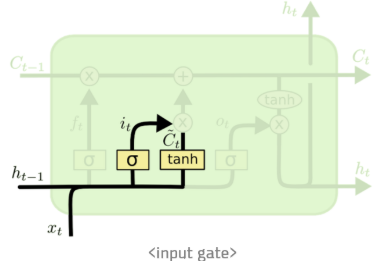
input gate



sigmoid

0~1사이의 값을 곱해서 input 으로 넣어줌 -> input으로 사용하지않을 것 0

output gate

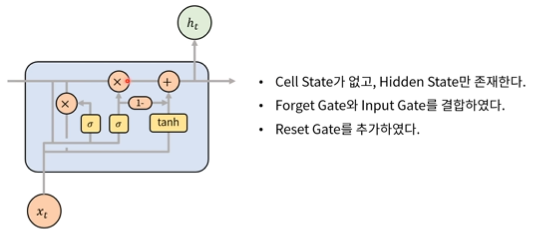


sigmoid, 0~1

tanh

출력값들이 누적되므로 그 값을 일정한 범위 수준으로 유지시켜 over되는 상황을 막기 위함

1. GRU



cell state 가 없고 hidden state 만 존재

forget으로 잊은 index-1해서 그 부분을 새로운 기억으로 채움

reset gate

새로운 feature를 뽑는 tanh으로 가기 전에, 0에 가까운 값이 되면 reset이 된다.

[참고]

<https://excelsior-cjh.tistory.com/89>