

PyTorch로 딥러닝 제대로 배우기

- 중급편 -

Part7. 순환 신경망(RNN) 이론

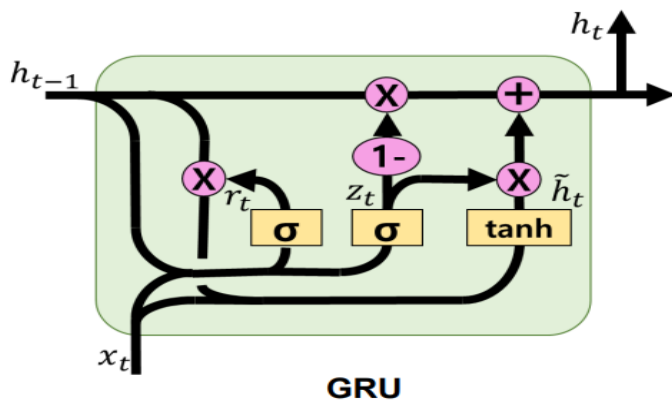
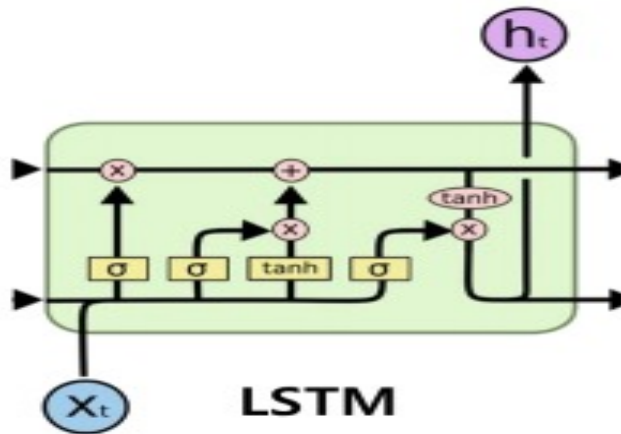
강사: 김 동 희



III. GRU(Gated Recurrent Unit, 2014)

1. GRU

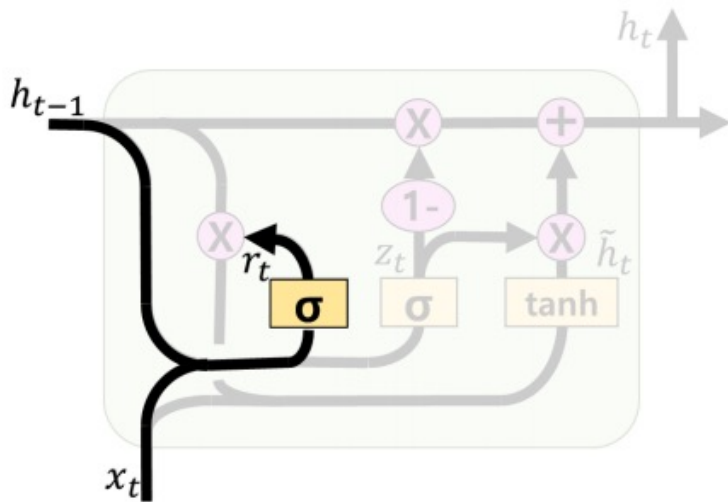
□ LSTM vs. GRU



- 단기기억을 위한 **Hidden Stage**와
장기기억을 위한 **Cell Stage**로 나누어짐
 - Hidden Stage 내에서도 Forget, Input, Output Gate 각각 역할 수행
 - 복잡한 구조로 학습에 시간 ↑
 - Cell State가 없고 **Hidden State**만 존재
 - Forget Gate와 Input Gate가 **결합** (Update Gate)
 - Reset Gate가 추가
- => **구조 단순화**로 인한 계산량 감소

2. GRU 구조

Reset Gate



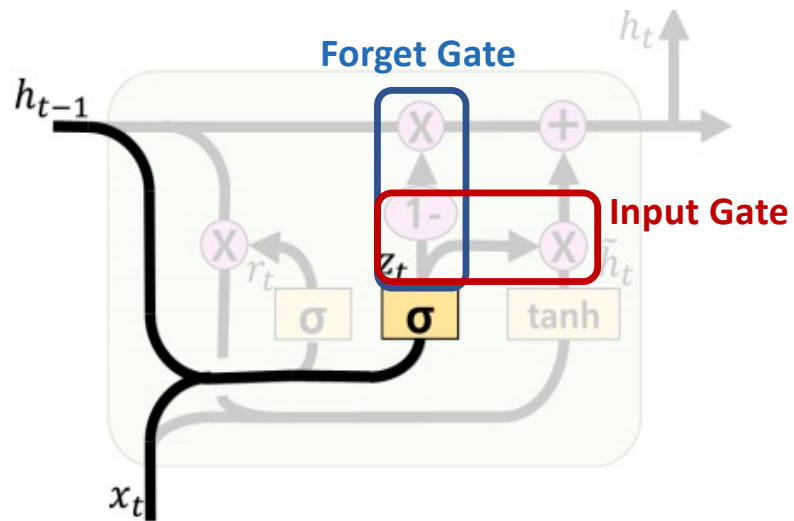
- **이전 Hidden State를 얼마나 사용할지 결정**
- 현재기억 후보를 만들기 위해 과거 기억을 재조정하는 비율
- Sigmoid 활성화함수로 0~1의 값 가진다 (0에 가까운 값이 되면 Reset)

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$

과거 현재
기억 입력

2. GRU 구조

□ Update gate



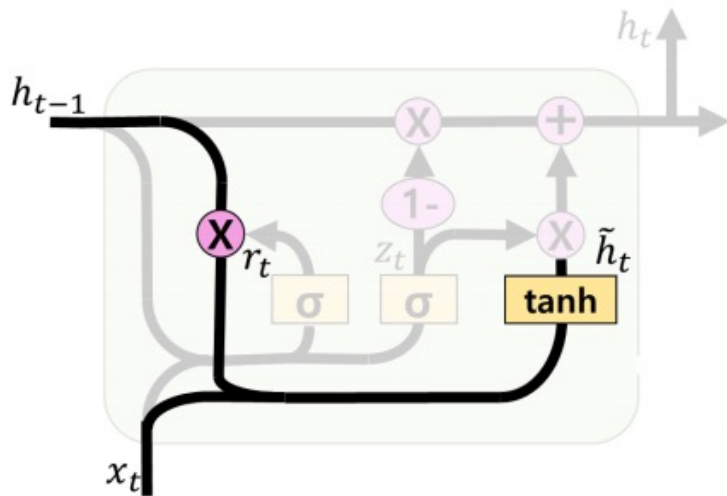
- Forget과 Input Gate 모두 제어
- 0 출력 시
Forget Gate 열림, Input Gate 닫힘
- 1 출력 시
Forget Gate 닫힘, Input Gate 열림
- (t-1)의 기억이 저장될 때, t의 입력 삭제

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$

과거 현재
기억 입력

2. GRU 구조

□ Memory Candidate



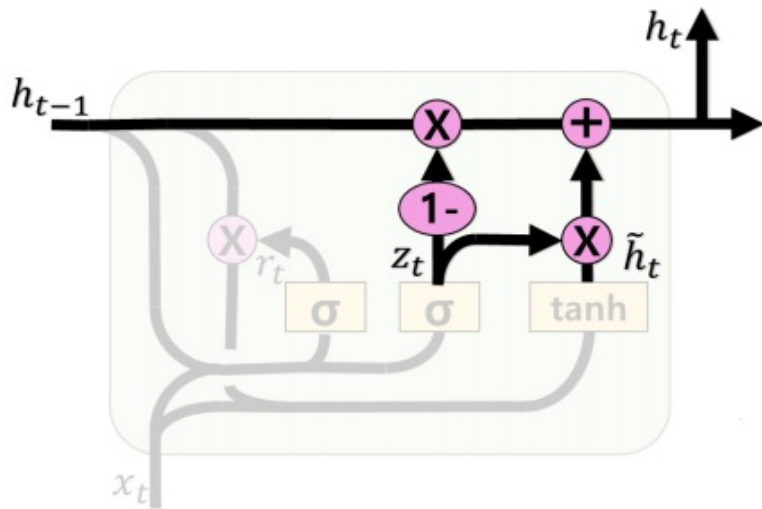
- 현재 기억의 후보를 만들기 위해 **현재 입력과 재조정(Reset)된 과거 기억의 조합**
- r_t 가 0에 가까울 수록 과거 기억을 지우고(reset), 현재 입력 위주로 현재 기억 후보를 만든다

$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$

재조정 비율	과거 기억	현재 입력
-----------	----------	----------

2. GRU 구조

□ Final memory



- 갱신 비율(z_t)에 따른 **과거의 기억**과 **현재 기억 후보의 조합**
- $z_t=0$: 과거 기억 위주의 최종 업데이트
- $z_t=1$: 현재 기억 후보 위주의 최종 업데이트

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

갱신 비율	과거 기억	갱신 비율	현재 기억 후보
----------	----------	----------	----------------

감사합니다.