



6. 게이트 순환 신경망(GRU)

6-1. GRU 구조

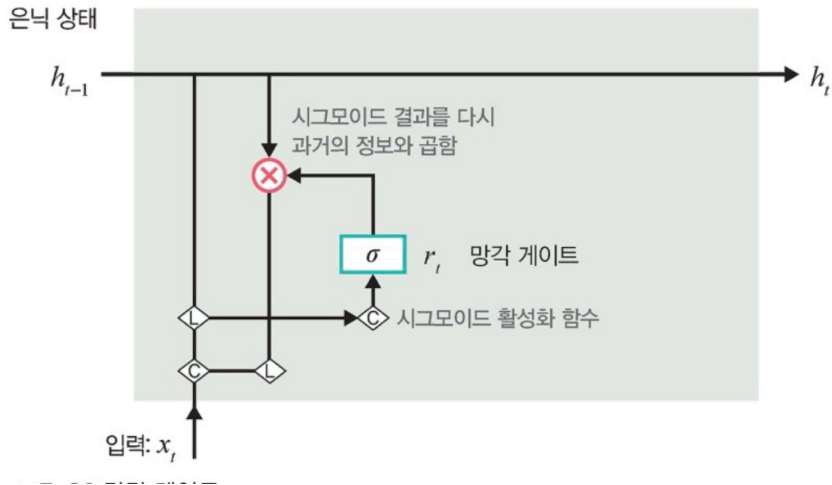
GRU(Gated Recurrent Unit)

- 게이트 메커니즘이 적용된 RNN 프레임워크의 한 종류
 - LSTM에서 사용하는 망각 게이트와 입력 게이트를 하나로 합친 것 → 더 간단한 구조
 - 별도의 업데이트 게이트로 구성되어 있음
 - 하나의 게이트 컨트롤러가 망각 게이트와 입력 게이트를 모두 제어
 - 게이트 컨트롤러가 **1**을 출력하면 망각 게이트는 열리고 입력 게이트는 닫힘
 - 반대로 **0**을 출력하면 망각 게이트는 닫히고 입력 게이트는 열림
- ⇒ 이전 기억이 저장될 때마다 단계별 입력은 삭제됨
- 출력 게이트가 x
 - 전체 상태 벡터가 매 단계마다 출력됨
 - 이전 상태의 어느 부분이 출력될지 제어하는 새로운 게이트 컨트롤러가 별도로 존재

망각 게이트

- 과거 정보를 적당히 초기화 시키려는 목적
- 시그모이드 함수를 출력으로 이용
 - (0,1) 값을 이전 은닉층에 곱함
 - 이전 시점의 은닉층 값에 현시점의 정보에 대한 가중치를 곱한 것

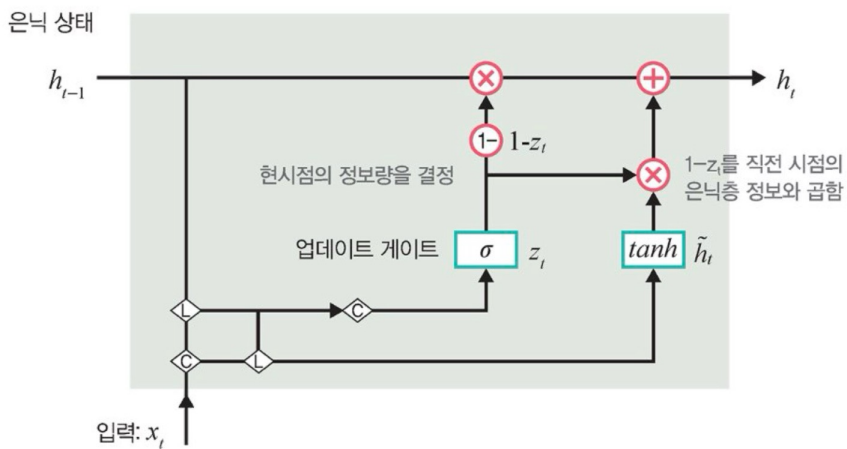
$$r_t = \sigma(W_r * [h_{t-1}, x_t])$$



업데이트 게이트

- 과거와 현재 정보의 최신화 비율을 결정하는 역할
- 시그모이드로 출력된 결과(z_t)는 현시점의 정보량을 결정하고 1에서 뺀 값($1 - z_t$)을 직전 시점의 은닉층 정보와 곱함

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



후보군(candidate)

- 현시점의 정보에 대한 후보군을 계산
- 과거 은닉층의 정보를 그대로 이용하지 x
 - 망각 게이트의 결과를 이용하여 후보군 계산

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

은닉층 계산

- 업데이트 게이트 결과와 후보군 결과를 결합하여 현시점의 은닉층을 계산
 - 시그모이드 함수의 결과는 현시점에서 결과에 대한 정보량을 결정
 - 1 - 시그모이드 함수의 결과는 과거의 정보량을 결정함

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

GRU 전체 내부 구조

