GRU(Gated Recurrent Unit)란?

GRU는 gate 종류가 많은 LSTM의 단점을 보완하기 위한 방법이다. LSTM과 달리, cell state가 없는데도 gradient vanishing 문제를 해결할 수 있다. Cell state가 없는 대신, ht-1이 nonlinear activation을 통과하지 않고 다음 state로 넘어갈 수 있기 때문이다. GRU에는 reset gate, update gate가 있다. Gate 수가 적을 수 있는 이유는 LSTM처럼 4 종류의 게이트를 구하는 것이 아니라, (1-update gate)의 방법으로 다른 게이트를 직접적으로 구하지 않고 모델을 만드는 것이기 때문이다.

GRU의 장점

LSTM에 비해 데이터 및 파라미터 수가 적기 때문에 연산 속도가 빠르다.

GRU의 단점

메모리와 결과값의 컨트롤이 불가능하다. 사실상 성능이 LSTM과 비슷한데, LSTM의 기능이 더 많고, 많이 사용되어왔기 때문에 GRU는 상대적으로 검증이 덜 된 architecture이라 할 수 있다.