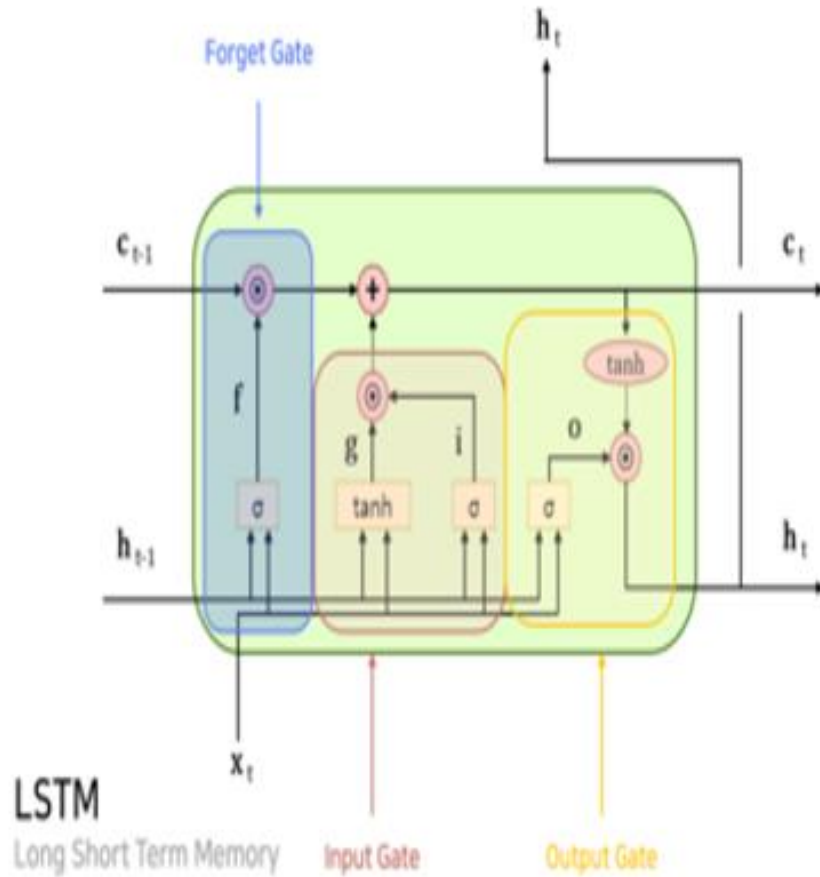


양방향 LSTM

박지훈

기존 LSTM의 구조

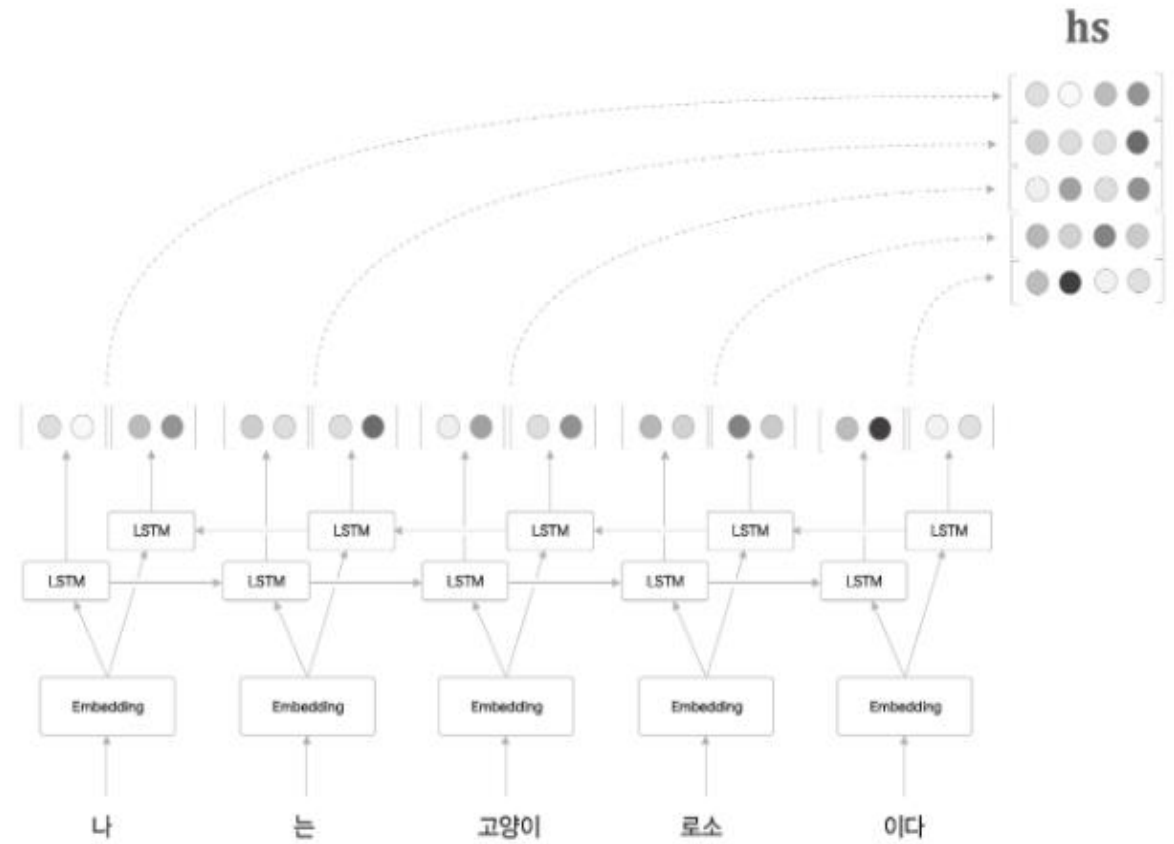
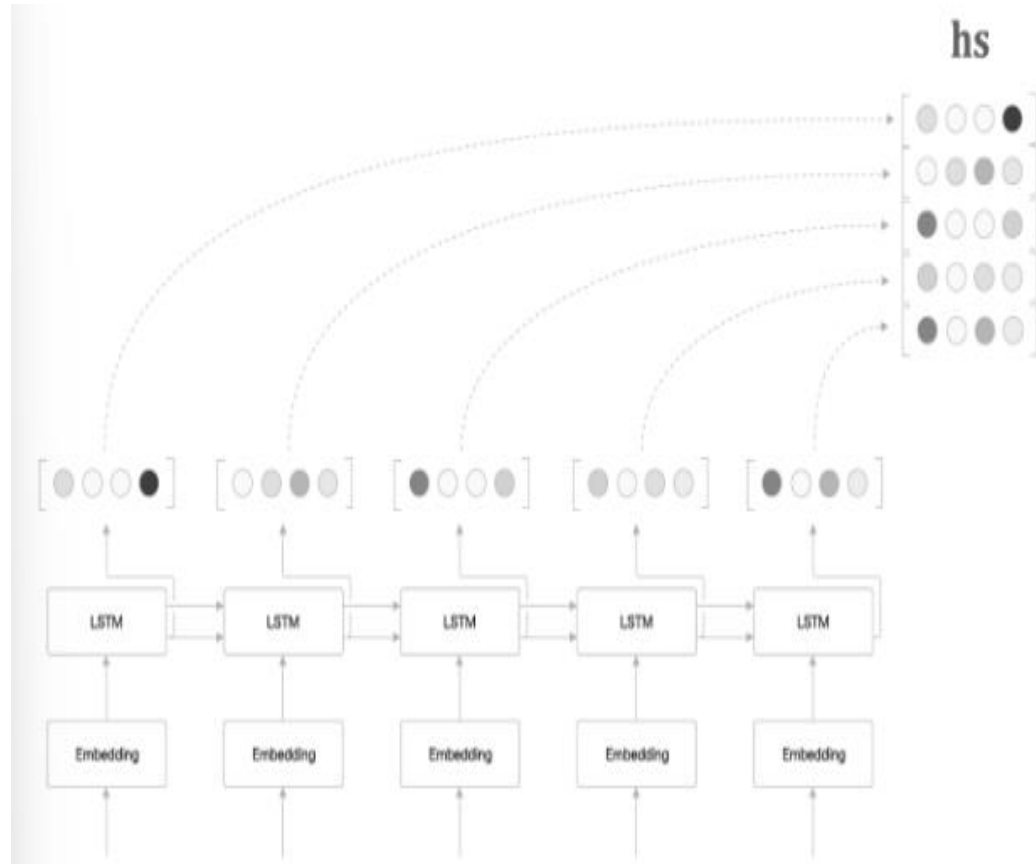


- RNN의 단점을 보완해서 만들어짐
- 입력, 망각, 출력 게이트를 통해 Long short-term dependency를 보완

기존 순환신경망의 문제

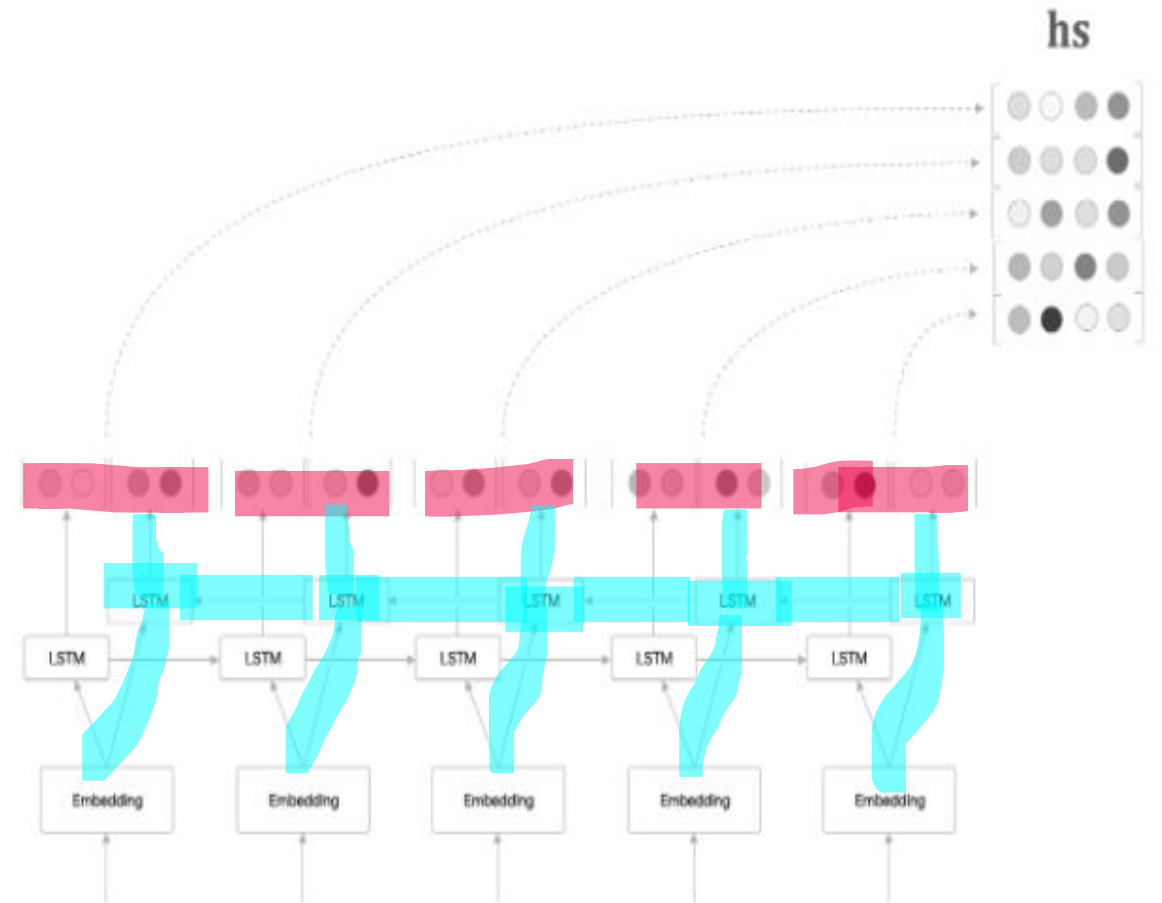
- RNN이나 LSTM과 같은 단방향 순환신경망은 시간순서를 입력하기 때문에 결과물이 **직전 패턴**을 기반으로 수렴하는 경향이 있다.
- 글은 왼쪽에서 오른쪽으로 읽는다.
- “나는 잘생긴 수원대학교 학생이다” 라는 문장이 있을때, “수원대학교”에 대응하는 벡터에는 “나는” “잘생긴”, “수원대학교”까지 총 세 단어의 정보가 인코딩 된다.
 - 문장이 길어질수록 한 벡터에 포함하고 있는 단어의 정보는 점점 많아진다.
 - 따라서 균형있게 담기 위해 “**양방향 LSTM**”을 사용하자.

기존 LSTM과 양방향 LSTM의 아키텍처



양방향 LSTM

- 기존의 LSTM계층에 역방향으로 LSTM계층을 추가한다.
- 최종은닉 상태는 두LSTM계층의 은닉상태를 연결한 벡터를 출력한다.



양방향 LSTM특징

- 최근 머신러닝 분야에서 좋은 성과에 적용된 모델일 정도로 높은 성능의 알고리즘 중 하나.
- 출력값에 대한 손실을 최소화하는 과정에서 모든 파라미터를 동시에 학습하는 **종단간 학습** 가능
- 단어와 구 간 유사성을 입력벡터에 내재화하여 성능 개선
- 데이터 길이가 길어도 성능이 저하되지 않음.

양방향 LSTM 간단하게 구현하는 방법

- 2개의 LSTM계층(A,B)을 사용해서 계층의 단어 순서를 조정
- 첫번째 계층(A)은 기존과 동일하게 입력 문장을 왼쪽에서 오른쪽으로 처리
- 두번째 계층(B)는 입력 문장의 단어순서를 반대로 처리
(나는 박지훈 -> 훈지박 는나)