

자연어 처리 딥러닝

다음 중 _____ 안에 들어갈 단어는 무엇일까요?

I'm becoming an AI _____.

- | | |
|------------|-----------|
| 1. were | 2. expert |
| 3. special | 4. going |

다음 중 _____ 안에 들어갈 단어는 무엇일까요?

I'm becoming an AI _____.

1. were

2. expert

3. special

4. going

다음 중 _____ 안에 들어갈 단어는 무엇일까요?

I'm becoming an AI _____.

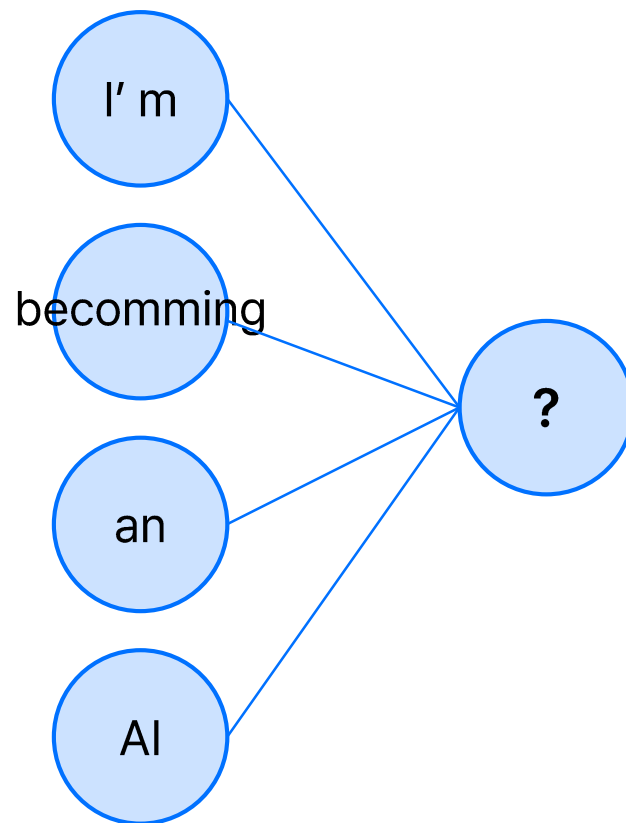
- | | |
|------------|-----------|
| 1. were | 2. expert |
| 3. special | 4. going |

영어문장1	→	빈칸에 들어갈 단어1
영어문장2	→	빈칸에 들어갈 단어2
영어문장3	→	빈칸에 들어갈 단어3
...		...

자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

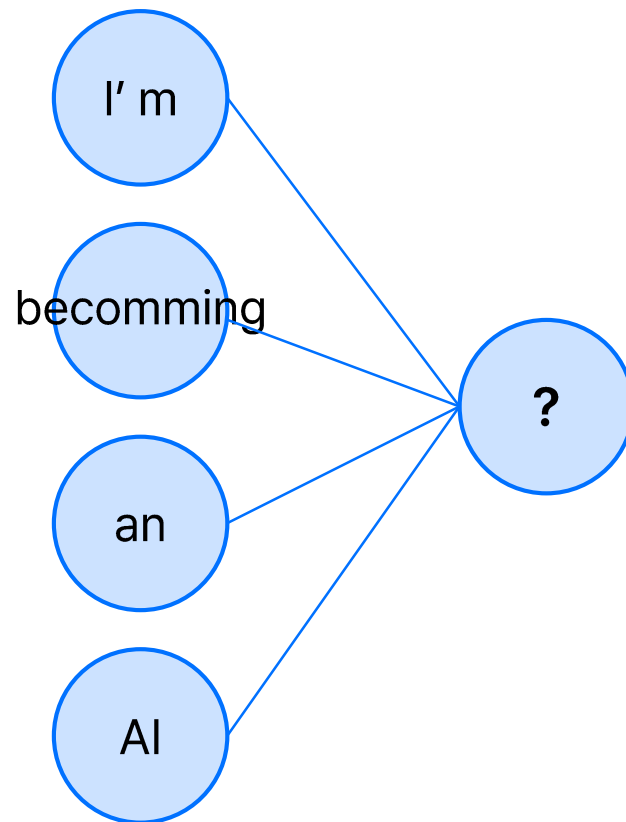
숫자로
넣어줘야함



자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

자연어 -> 숫자
인코딩

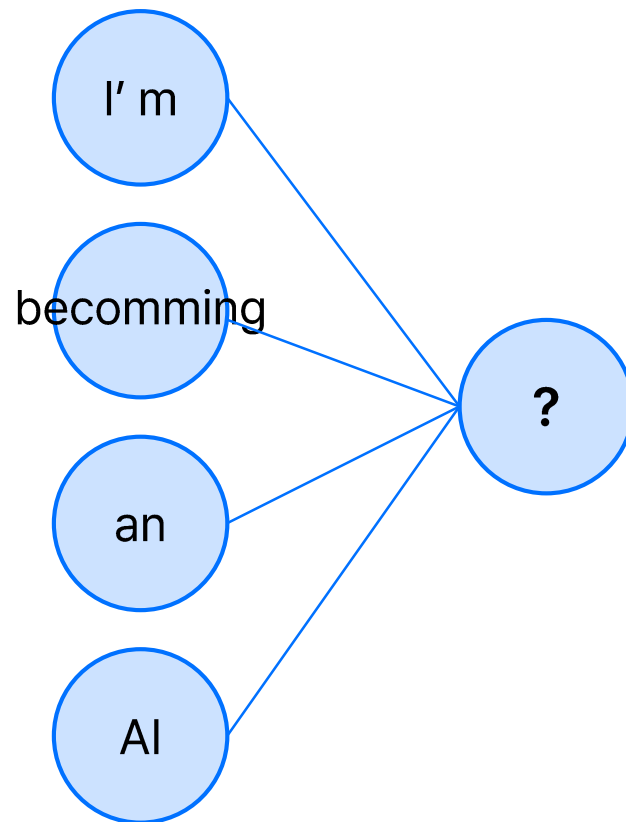


자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

원 핫 인코딩

I'm	0 0 0 0
becoming	0 0 0 0
an	0 0 0 0
AI	0 0 0 0

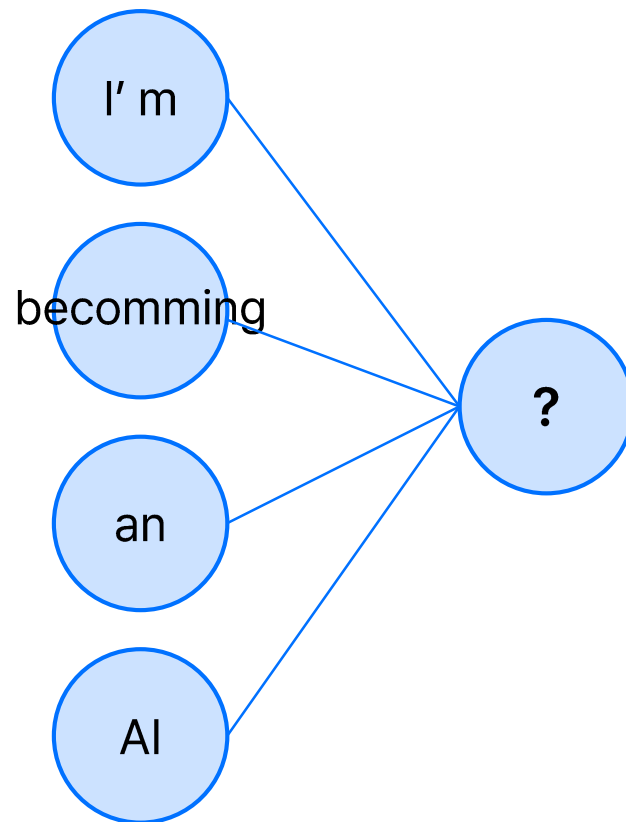


자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

원 핫 인코딩

I'm	1 0 0 0
becoming	0 0 0 0
an	0 0 0 0
AI	0 0 0 0

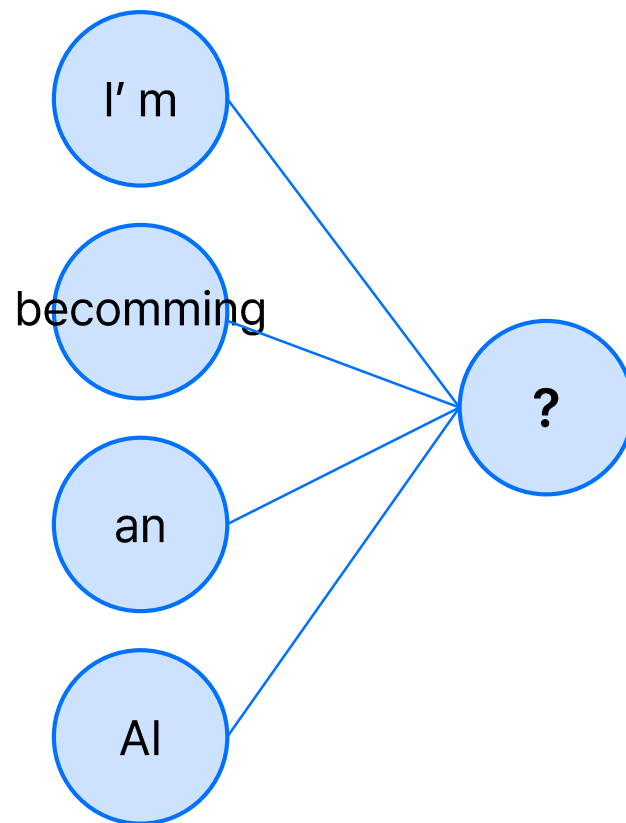


자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

원 핫 인코딩

I'm	1 0 0 0
becoming	0 1 0 0
an	0 0 0 0
AI	0 0 0 0

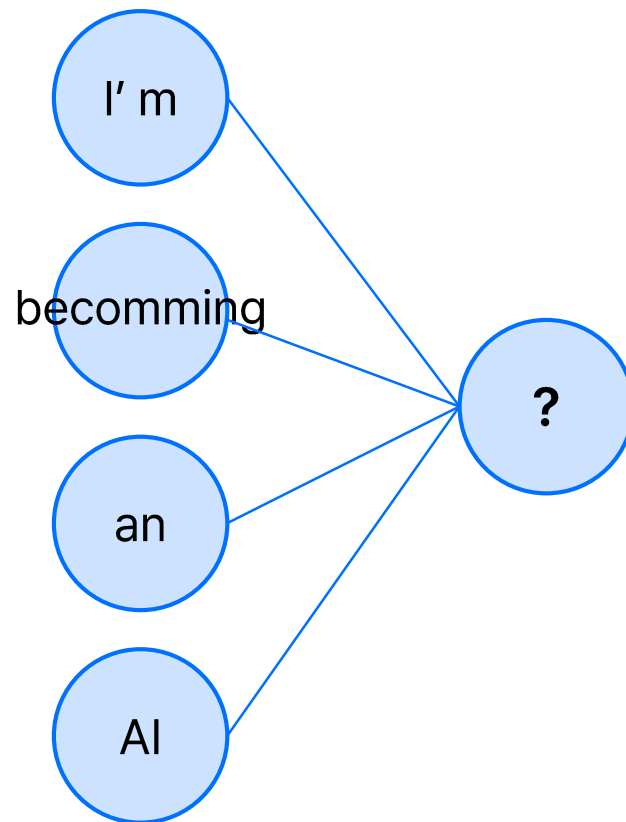


자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

원 핫 인코딩

I'm	1 0 0 0
becoming	0 1 0 0
an	0 0 1 0
AI	0 0 0 0

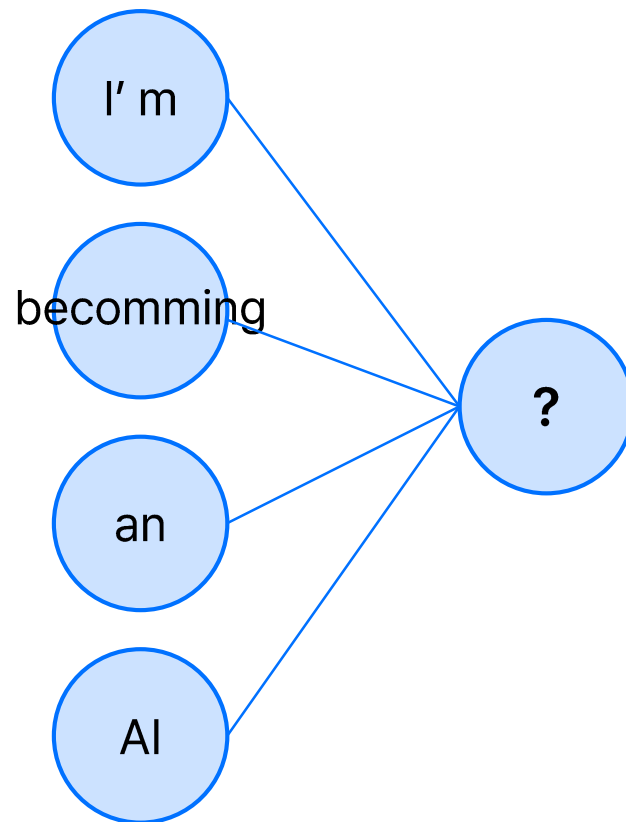


자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

원 핫 인코딩

I'm	1 0 0 0
becoming	0 1 0 0
an	0 0 1 0
AI	0 0 0 1

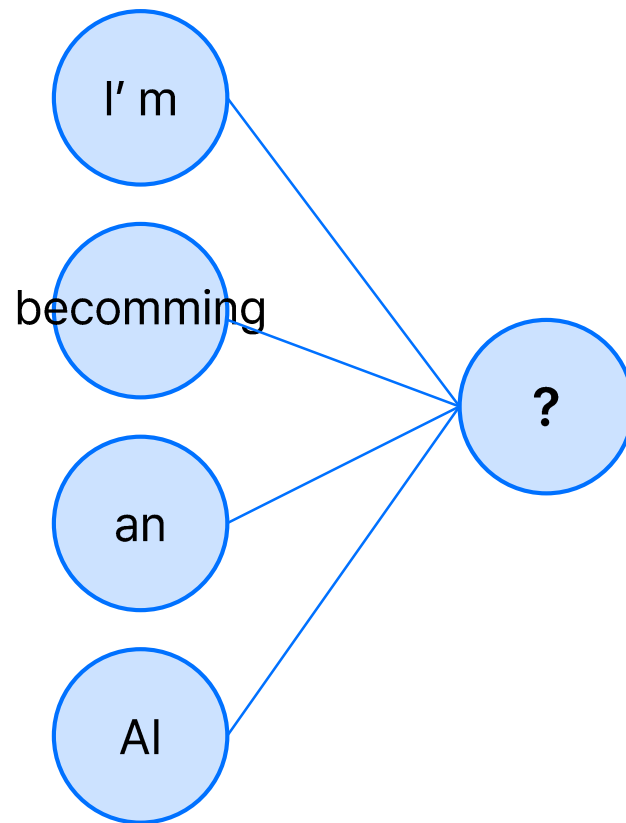


자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

원 핫 인코딩

I'm	1 0 0 0
becoming	0 1 0 0
an	0 0 1 0
AI	0 0 0 1

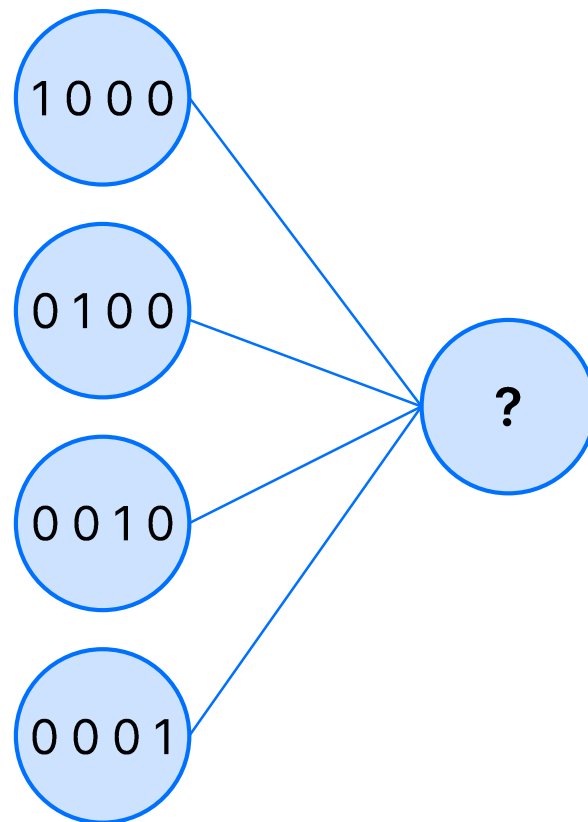


자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

원 핫 인코딩

I'm	1 0 0 0
becoming	0 1 0 0
an	0 0 1 0
AI	0 0 0 1



were 1.5%

expert 55%

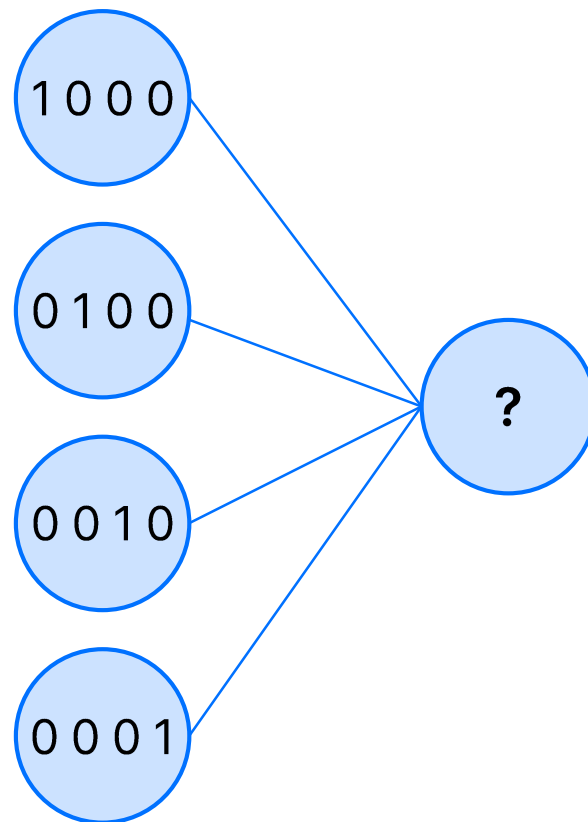
special 42%

going 1.5%

자연어처리 모델의 특징과 대표적 모델

I'm becoming an AI _____.

자연어는 순서가
중요함



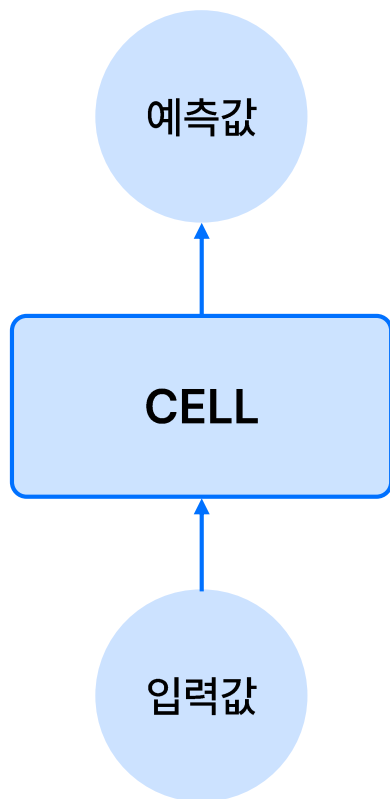
were 1.5%
expert 55%
special 42%
going 1.5%

순서를 기억하는 것에 중점을 둔 모델들

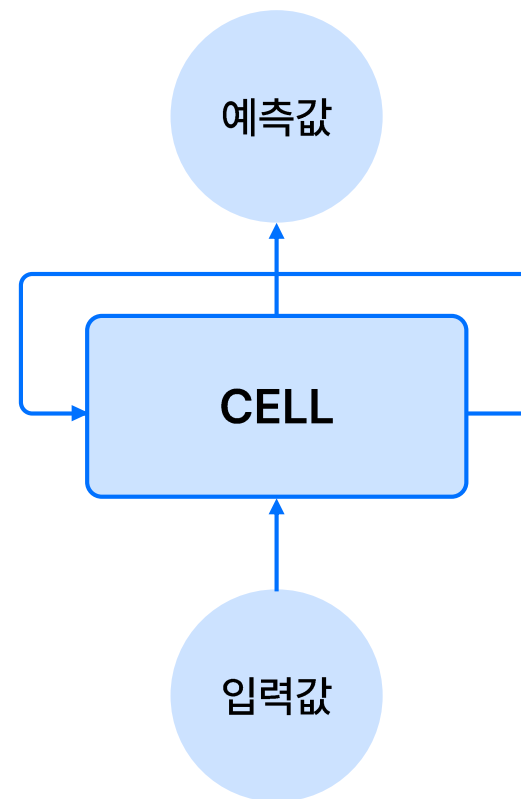
RNN (Recurrent Neural Network)

자연어처리 모델의 특징과 대표적 모델

기존 모델

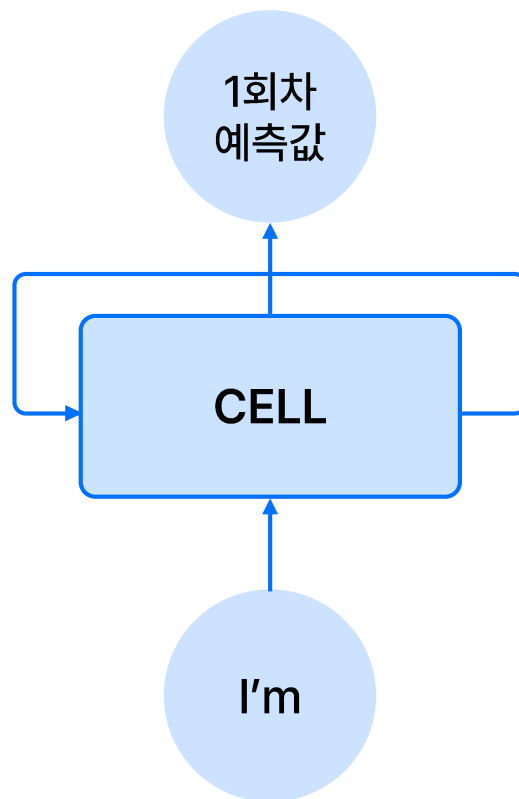


RNN



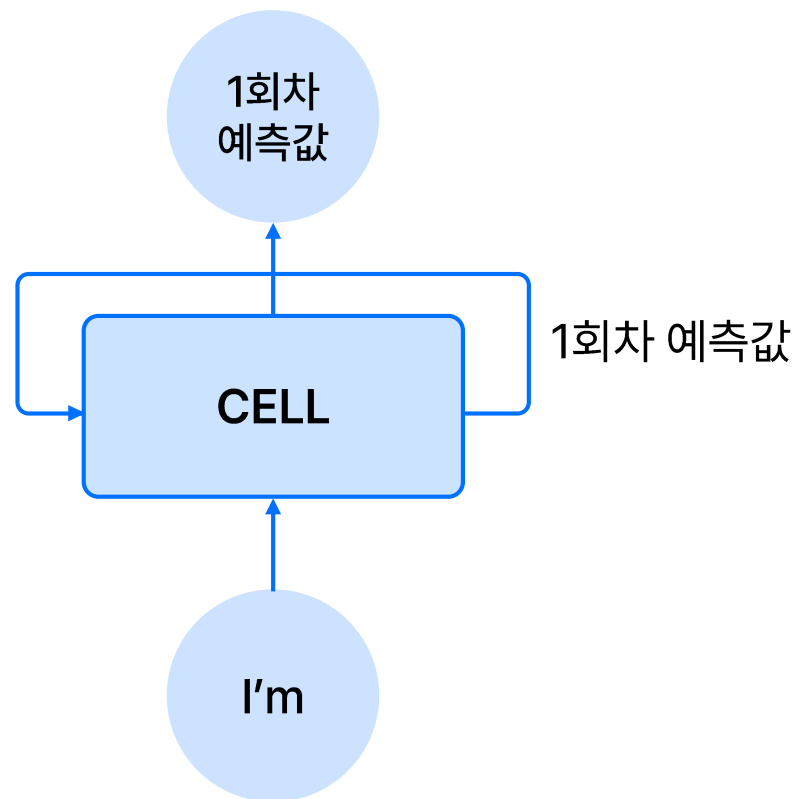
자연어처리 모델의 특징과 대표적 모델

1회차



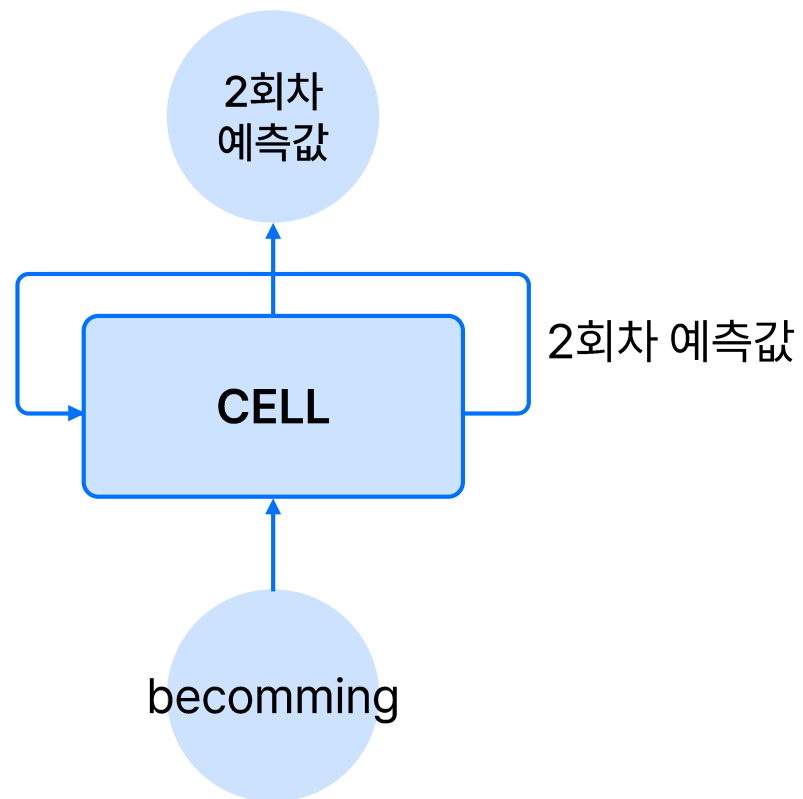
자연어처리 모델의 특징과 대표적 모델

1회차



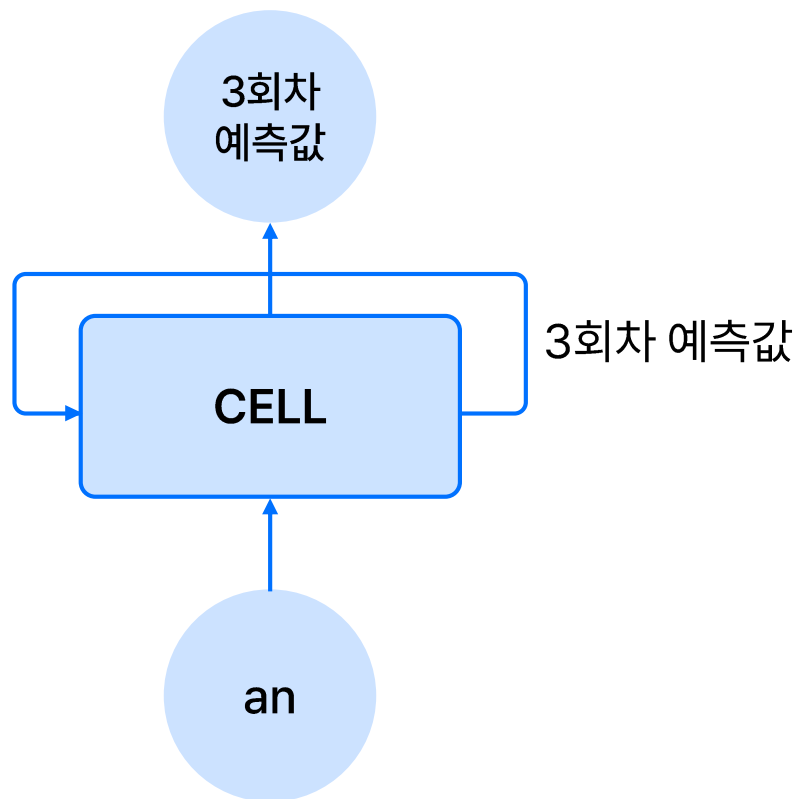
자연어처리 모델의 특징과 대표적 모델

2회차



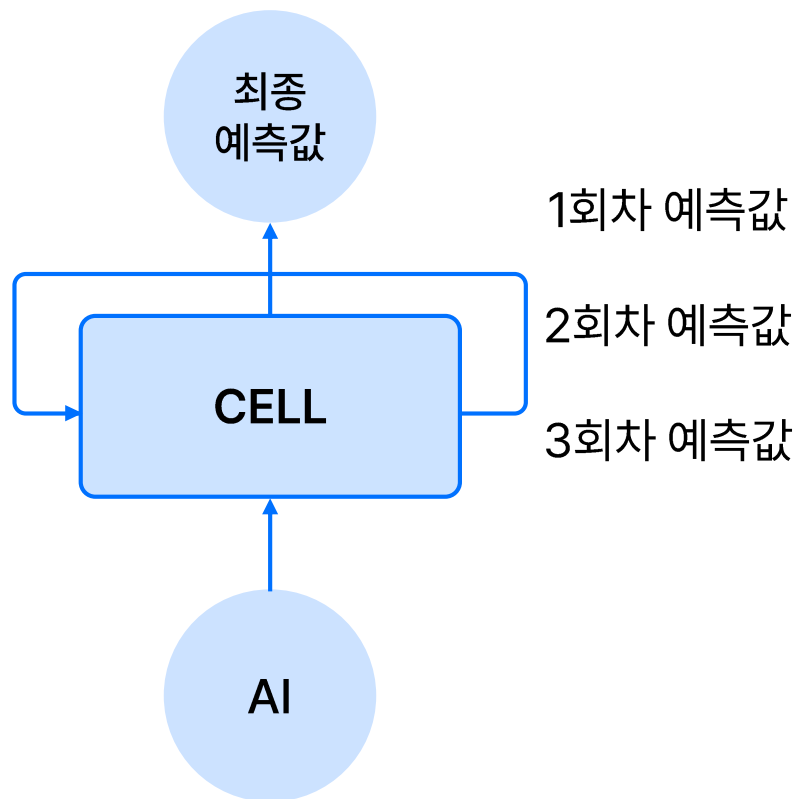
자연어처리 모델의 특징과 대표적 모델

3회차



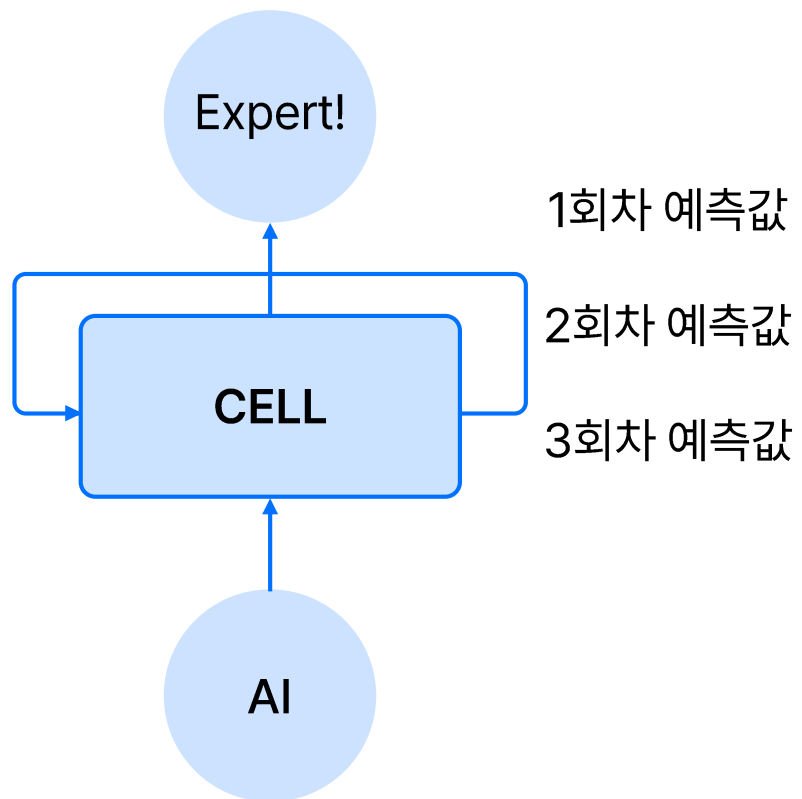
자연어처리 모델의 특징과 대표적 모델

4회차

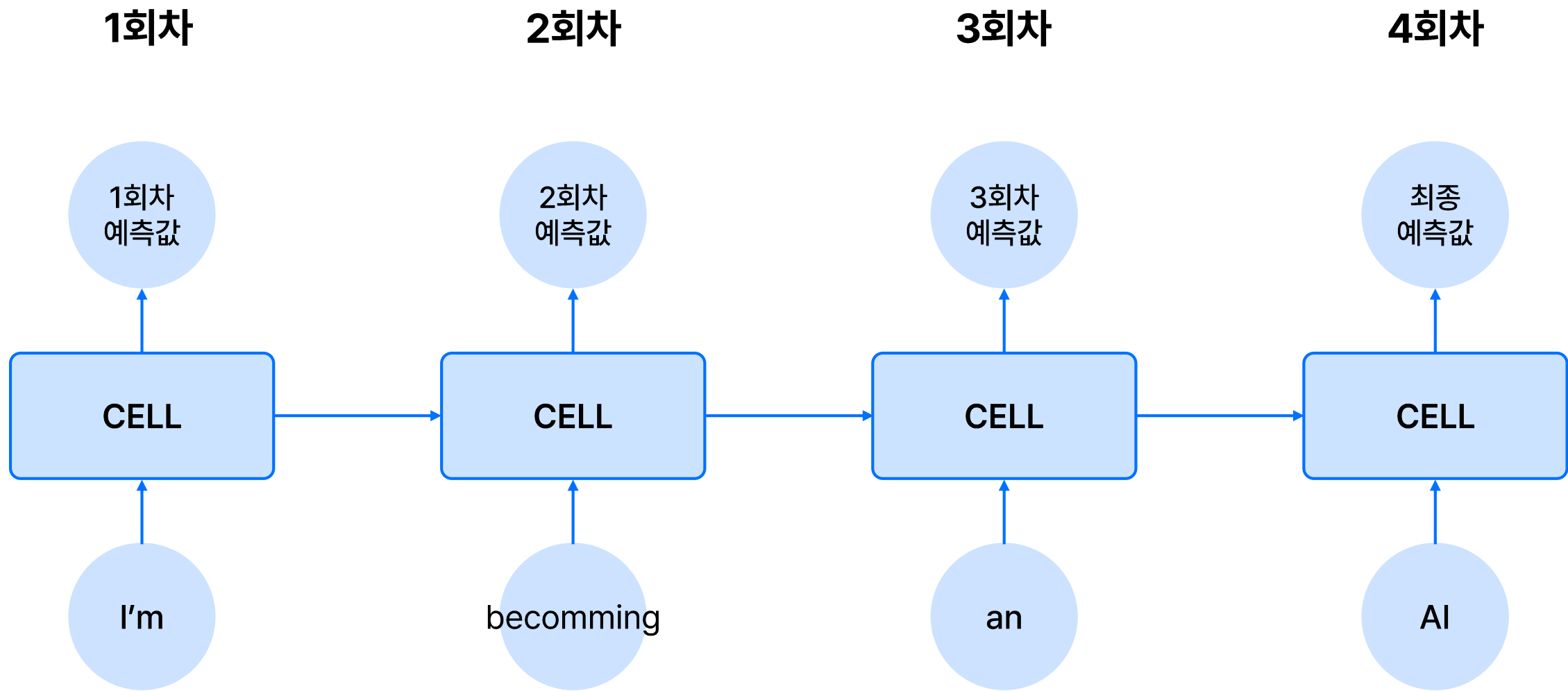


자연어처리 모델의 특징과 대표적 모델

4회차



자연어처리 모델의 특징과 대표적 모델



**RNN 구조는 순서가 있는 데이터에 많이
사용됩니다**

자연어, 시계열 데이터



**RNN 구조는 순서가 있는 데이터에 많이
사용됩니다**

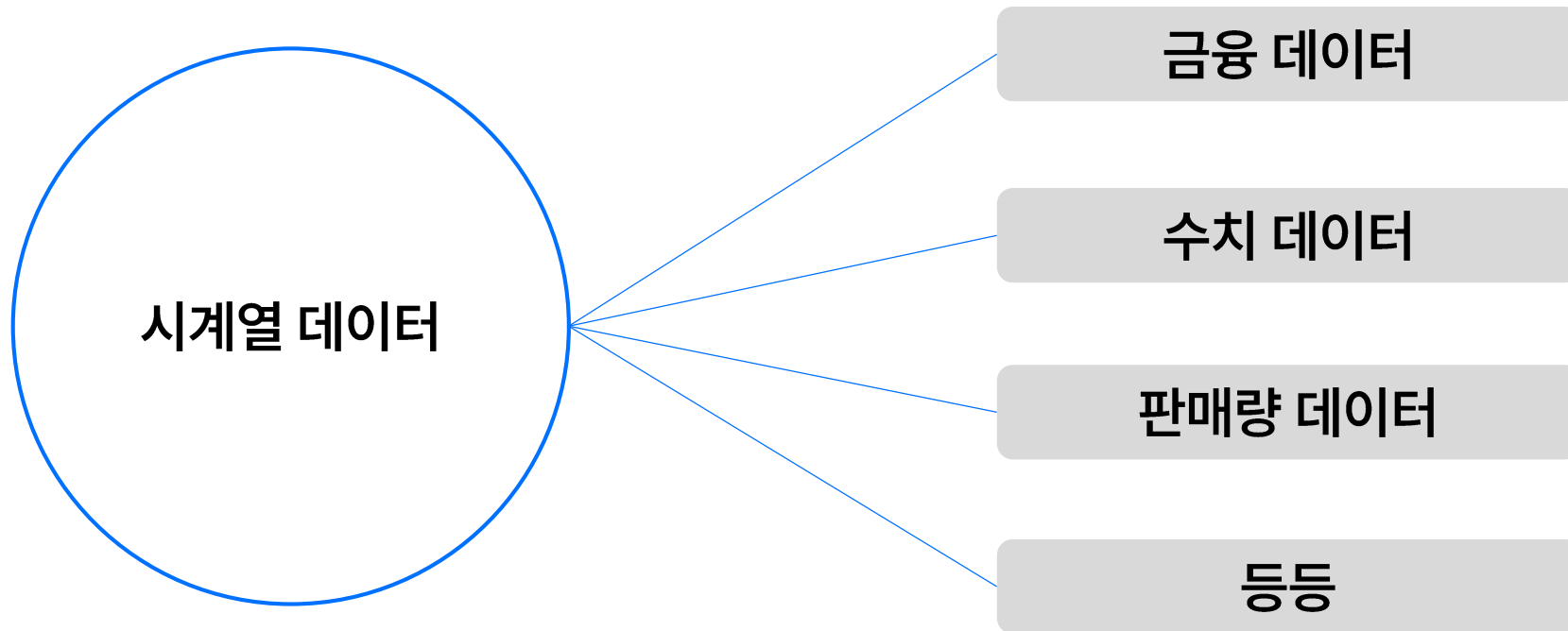
자연어처리 모델의 특징과 대표적 모델



시계열 데이터

순서, 시간 등의 순차적인 정보가
들어가는 데이터

자연어처리 모델의 특징과 대표적 모델



다음 중 _____ 안에 들어갈 단어는 무엇일까요?

탄수화물을 많이 먹으면 _____

- 1. 근육 2. 스트레스
- 3. 지방

다음 중 _____ 안에 들어갈 단어는 무엇일까요?

탄수화물을 많이 먹으면 _____이 늘어난다.

1. 근육

2. 스트레스

3. 지방

다음 중 _____ 안에 들어갈 단어는 무엇일까요?

탄수화물을 많이 먹으면 _____이 늘어난다.

1. 근육

2. 스트레스

3. 지방

BRNN

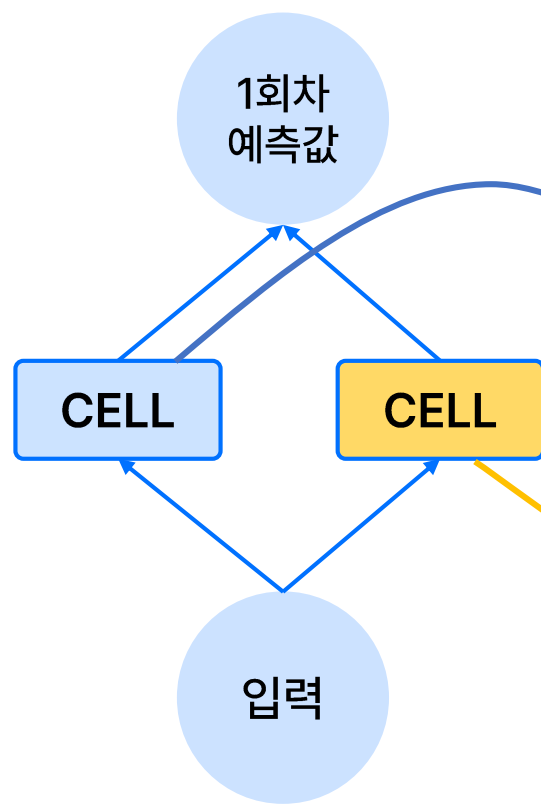
(Bidirectional Recurrent Neural Network)

과거 시점과 미래 시점을 모두 고려

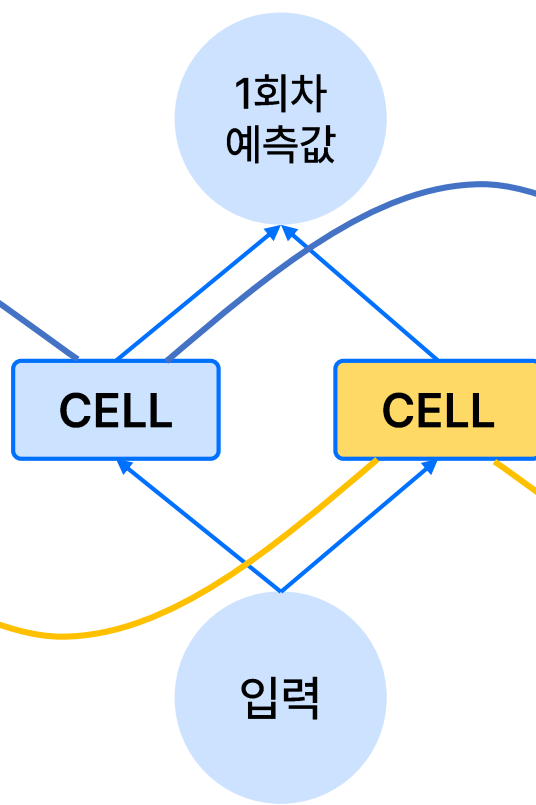
현재 시점의 예측을 더욱 정확하게 만듦

자연어처리 모델의 특징과 대표적 모델

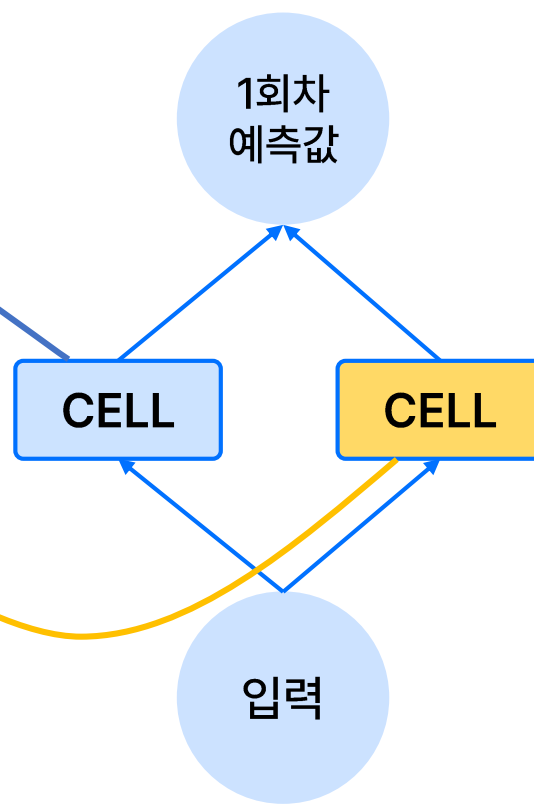
1회차



2회차

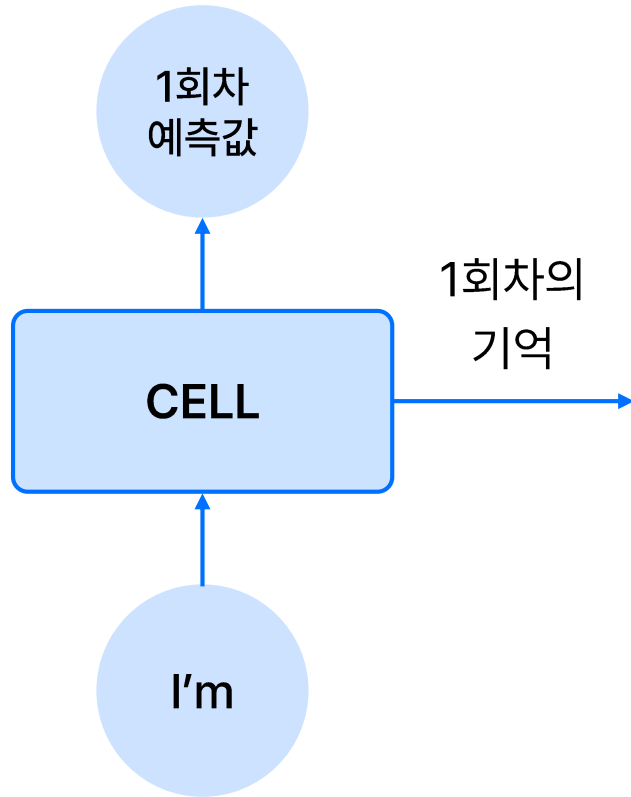


3회차

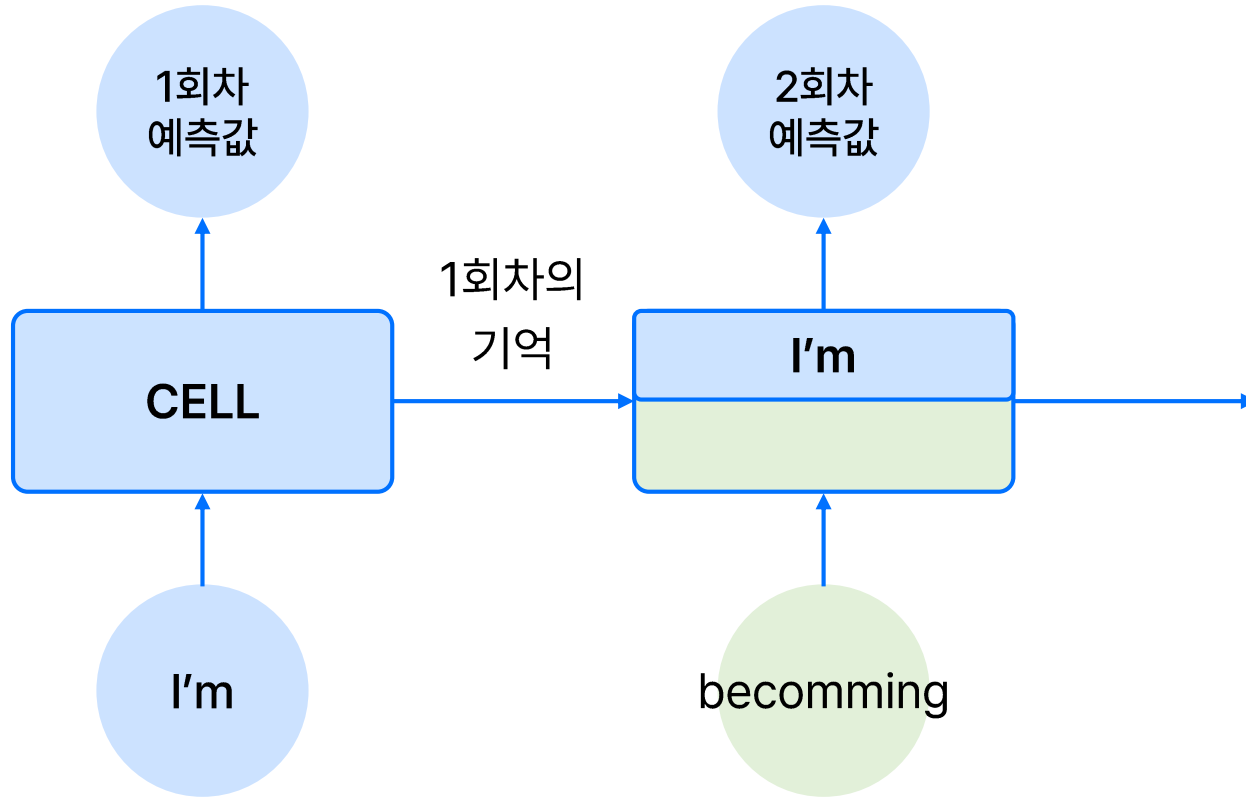


RNN 구조의 문제점

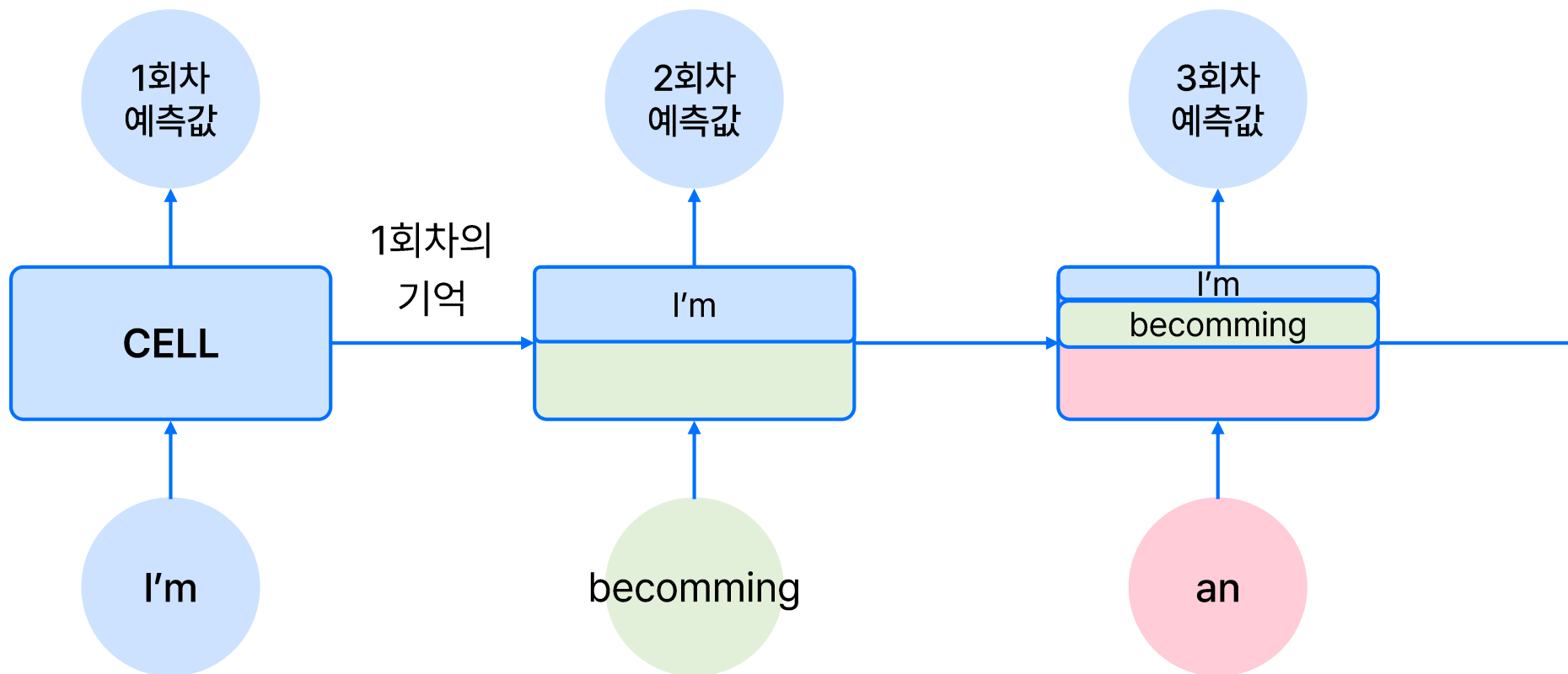
RNN구조의 문제점



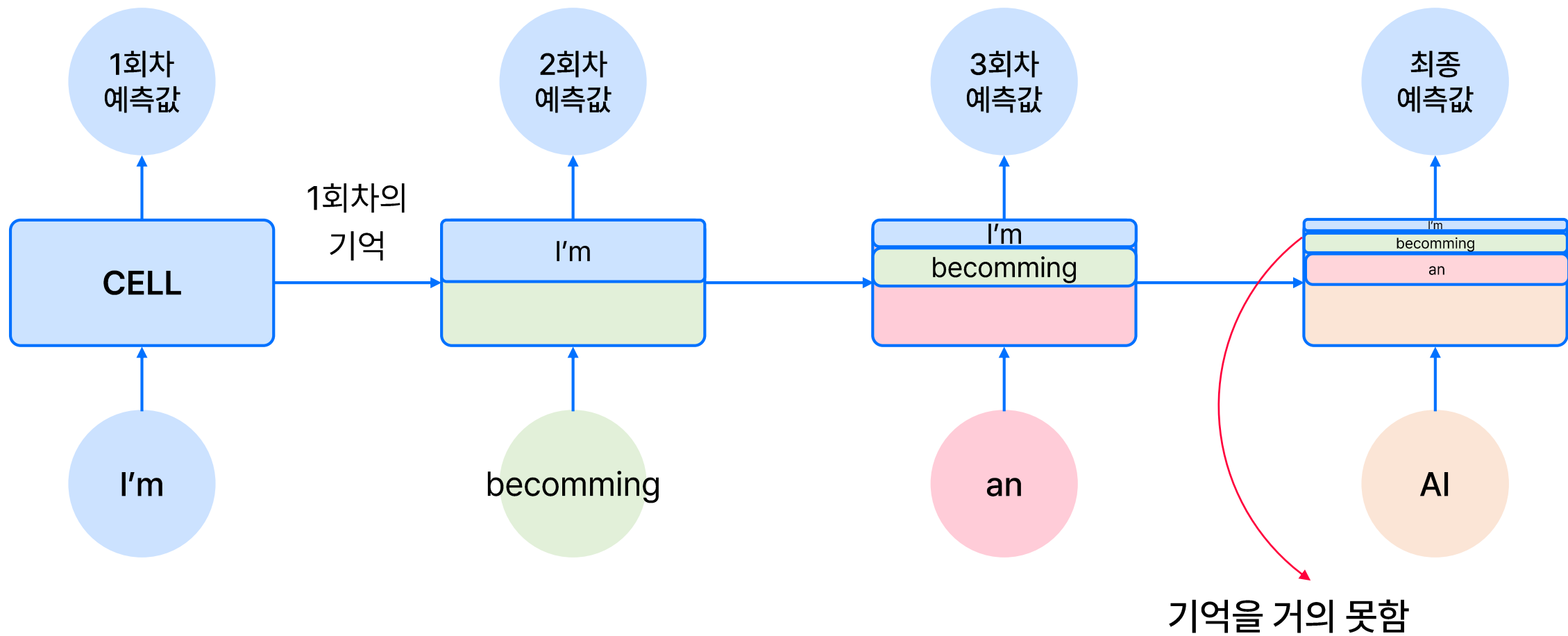
RNN구조의 문제점



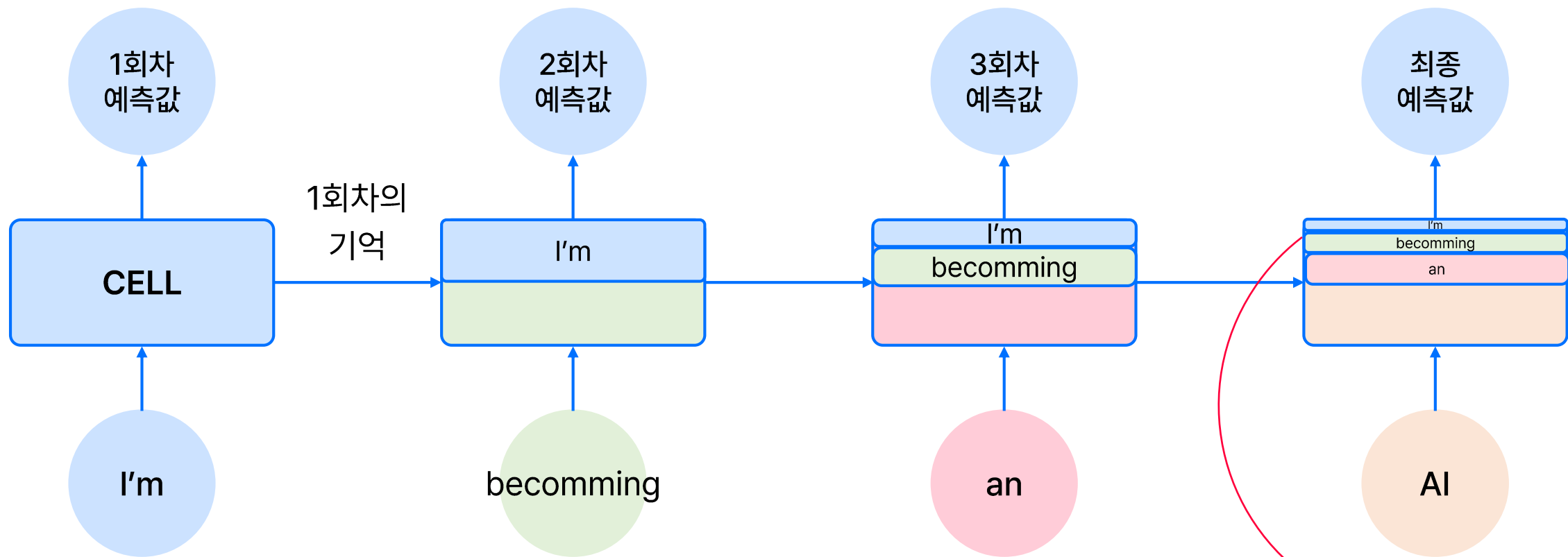
RNN구조의 문제점



RNN구조의 문제점



RNN구조의 문제점



기억을 거의 못함
문장은 5단어보다 훨씬
길 확률이 높음

자연어처리 모델의 특징과 대표적 모델

내가 이번에 일본 교토 여행을 다녀왔는데 건물도 너무 예쁘고
먹을 것도 맛있게 잘 먹었어. 그런데 갑자기 부모님한테 전화가
온거야. 지금 어디냐고 묻더라구 그래서 나는 말했지. 저
여행왔는데요. 여기 ____ 에요.

자연어처리 모델의 특징과 대표적 모델

내가 이번에 **일본 교토** 여행을 다녀왔는데 건물도 너무 예쁘고
먹을 것도 맛있게 잘 먹었어. 그런데 갑자기 부모님한테 전화가
온거야. 지금 어디냐고 묻더라구 그래서 나는 말했지. 저
여행왔는데요. 여기 **일본 교토** 에요.

자연어처리 모델의 특징과 대표적 모델

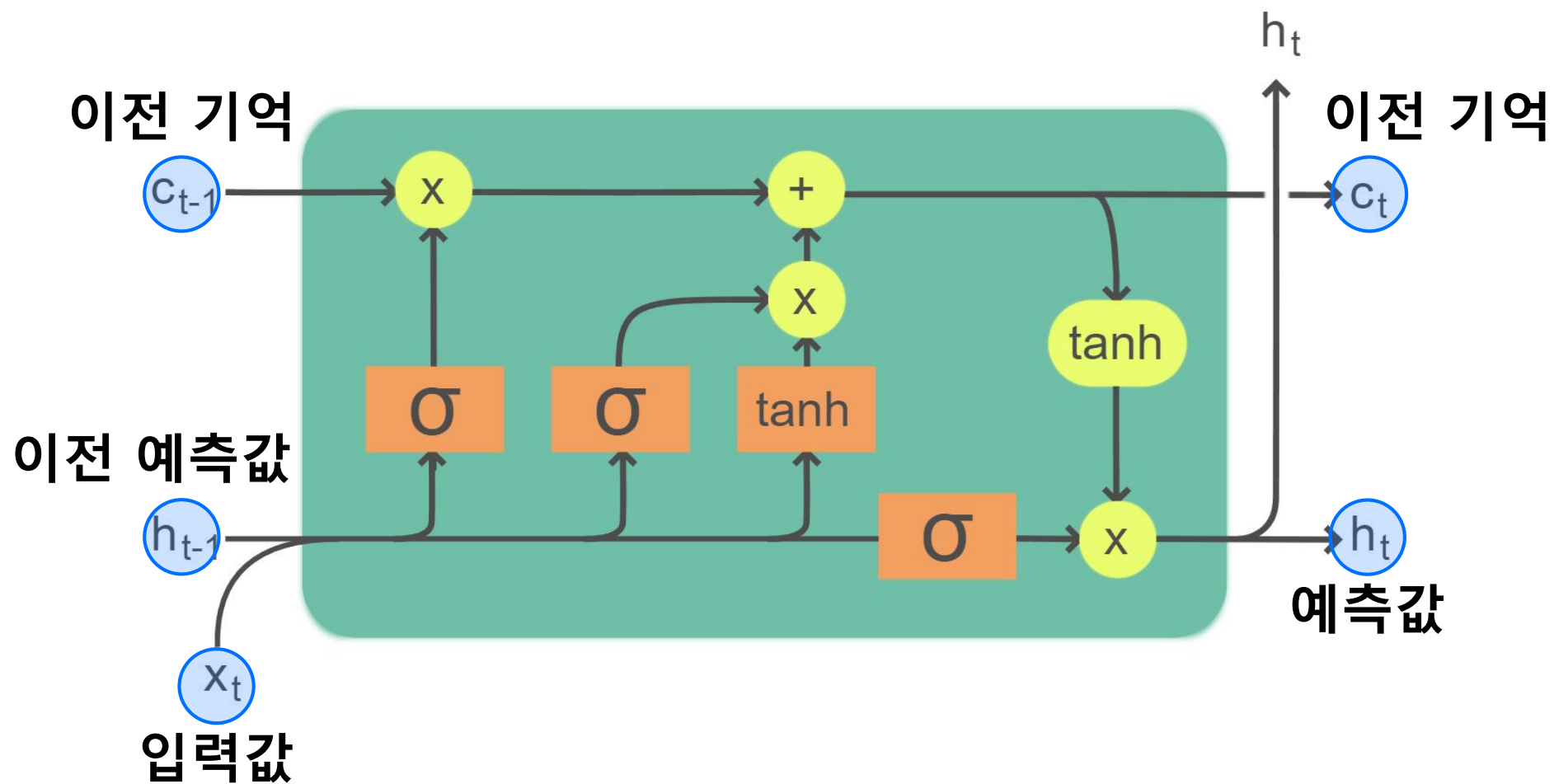
내가 이번에 일본 교토 여행을 다녀왔는데 건물도 너무 예쁘고
먹을 것도 맛있게 잘 먹었어. 그런데 갑자기 부모님한테 전화가
온거야. 지금 어디냐고 묻더라구 그래서 나는 말했지. 저
여행왔는데요. 여기 ????? 에요.

**RNN이 충분한 기억력을 가지고 있지 않다면,
예측이 어려움**

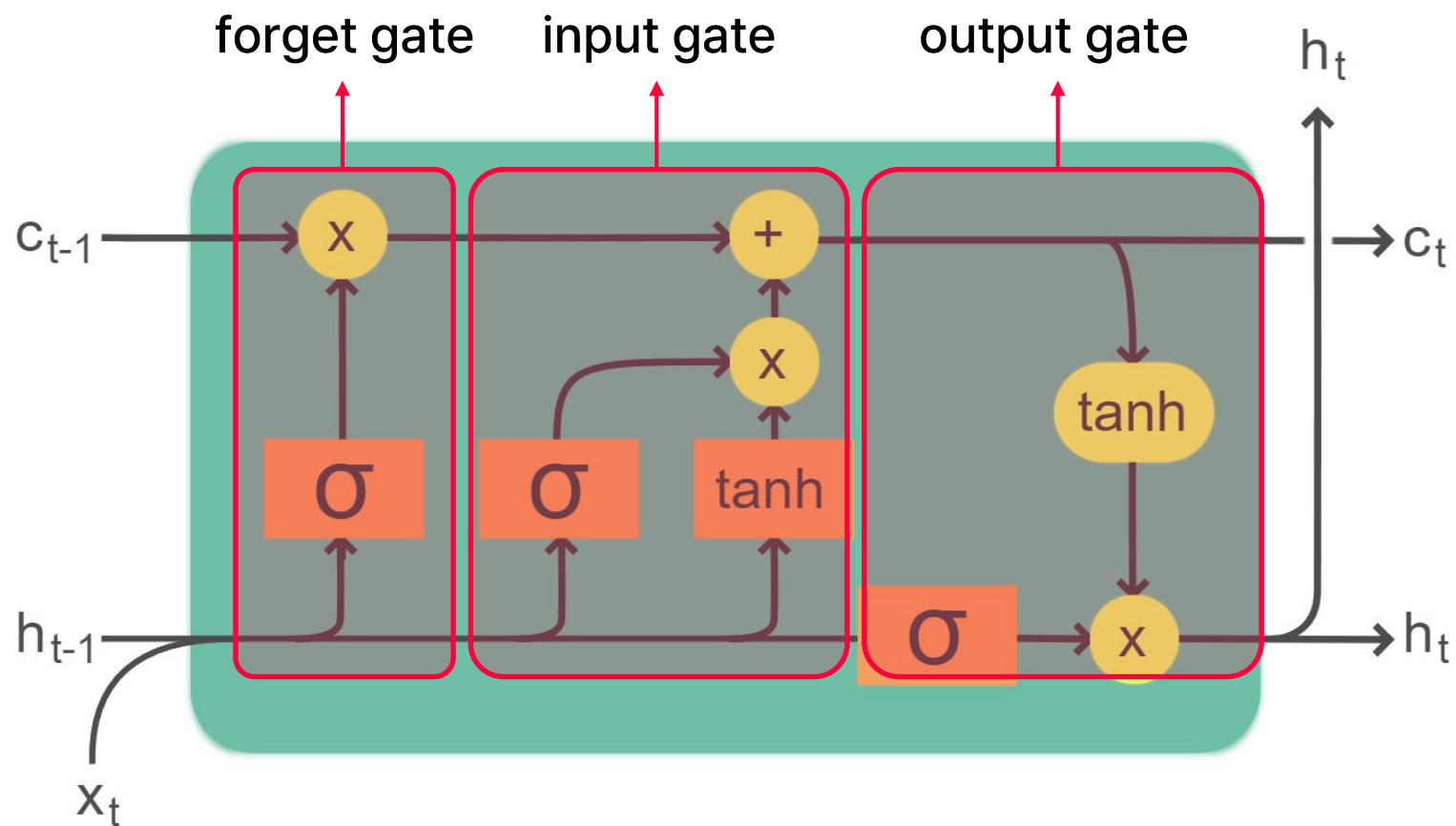
Long Term dependency Problem (장기 의존성 문제)

LSTM(Long Short Term Memory) 모델

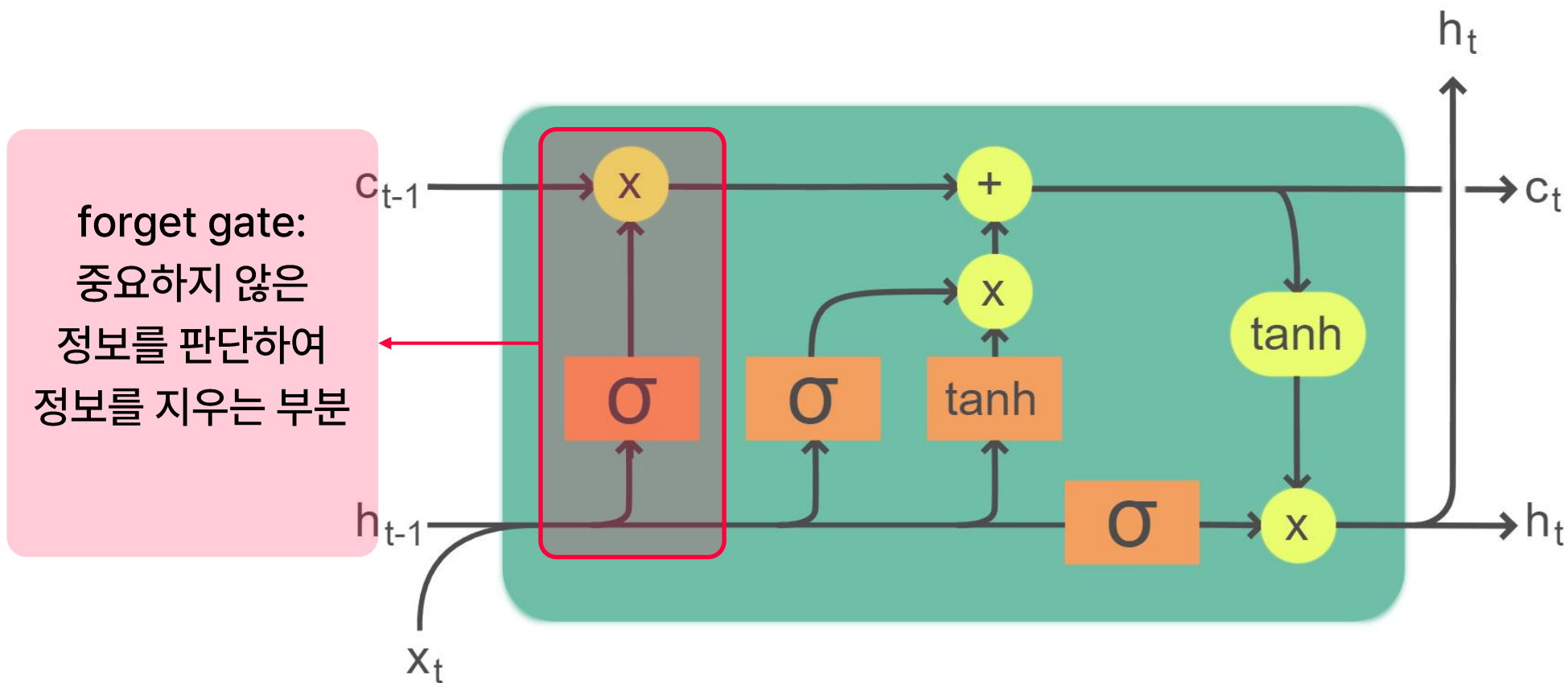
LSTM의 구조



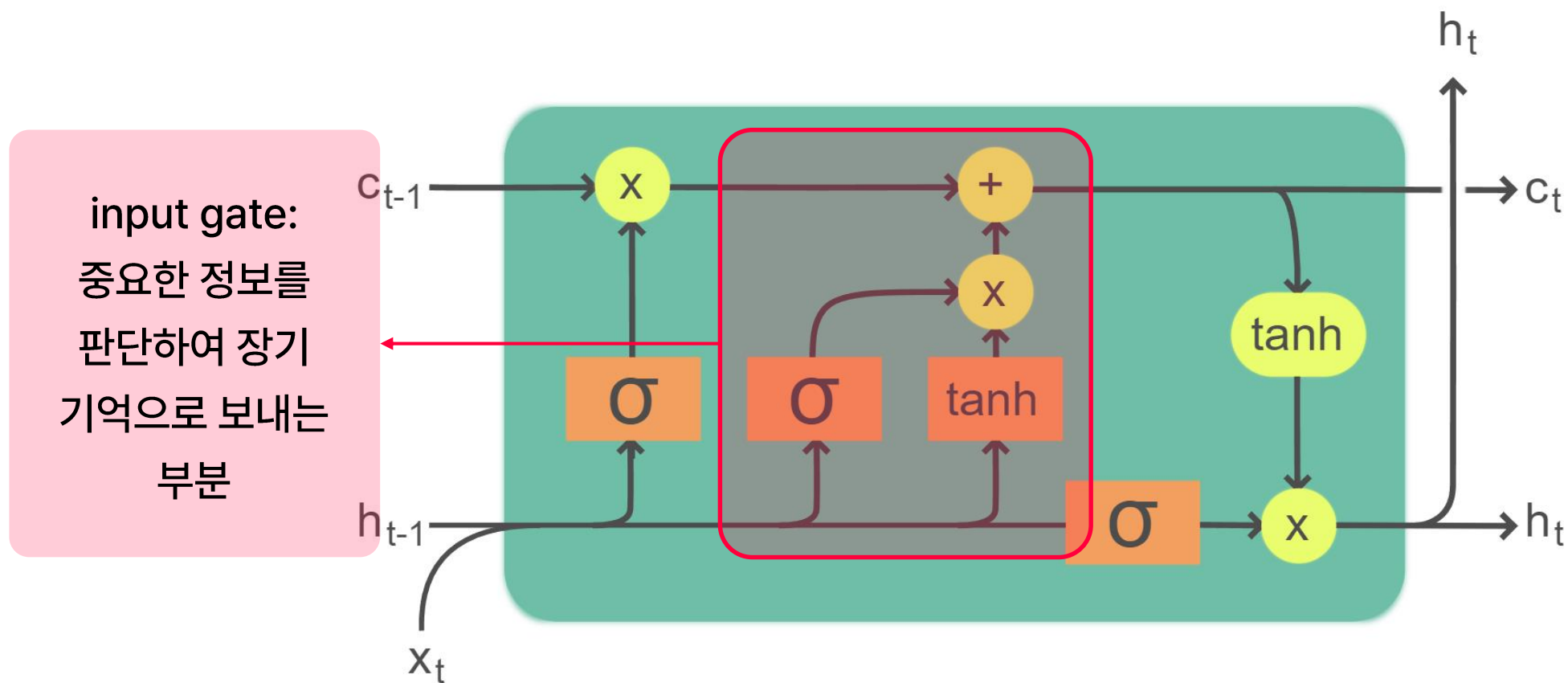
LSTM의 구조



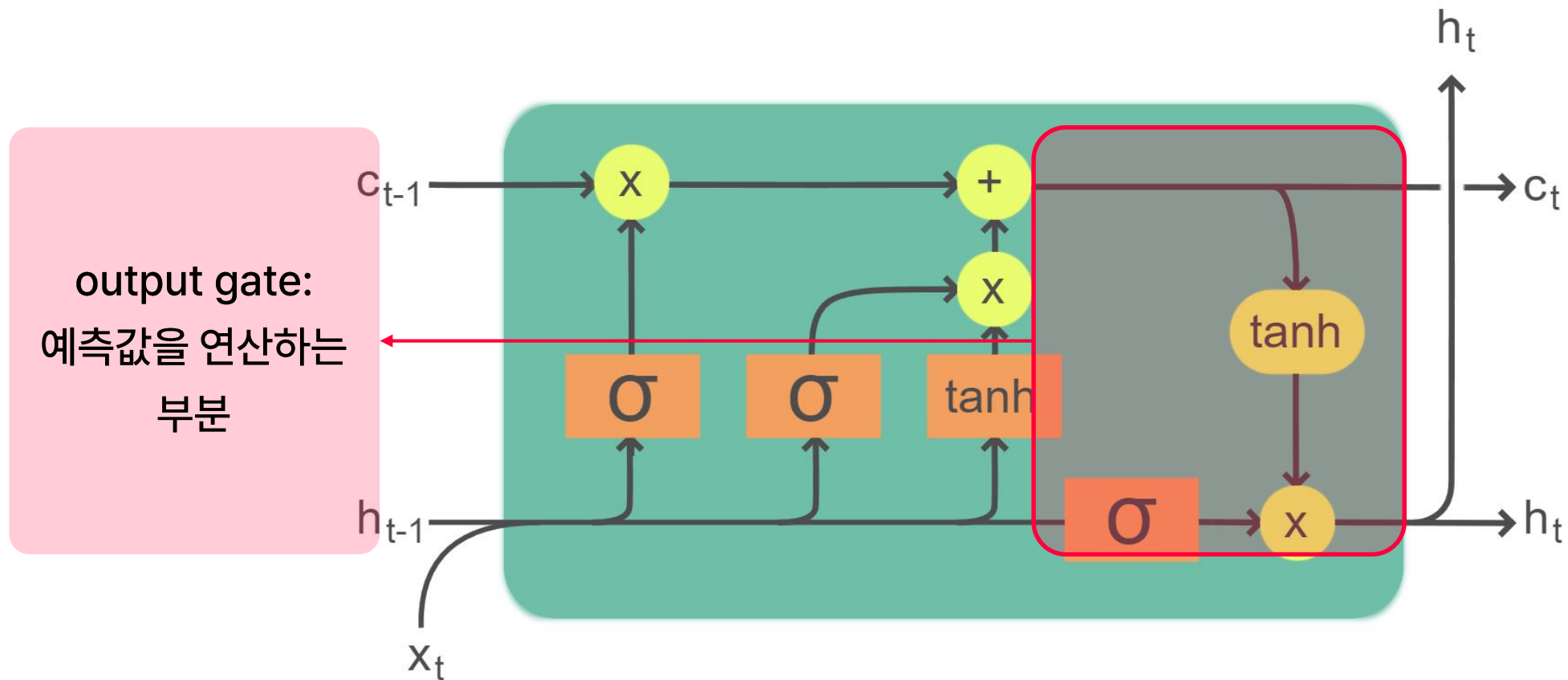
LSTM의 구조



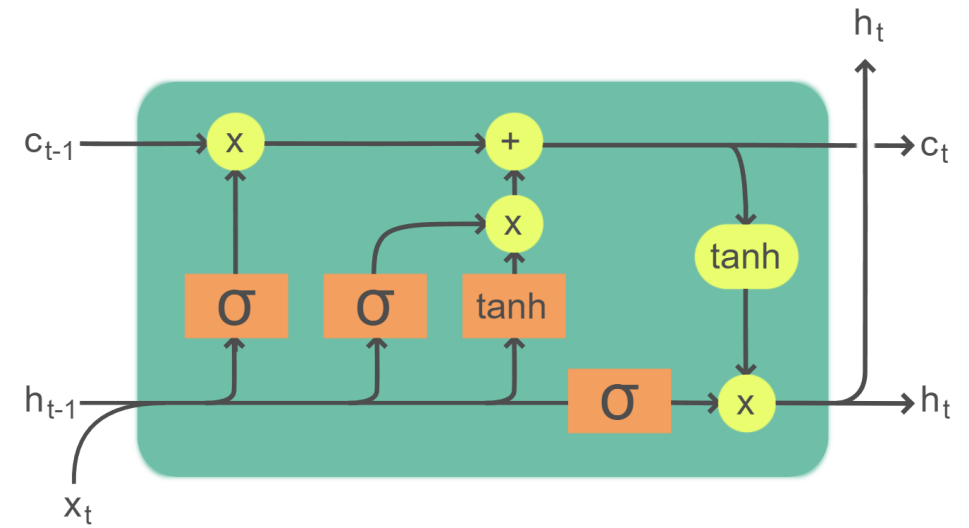
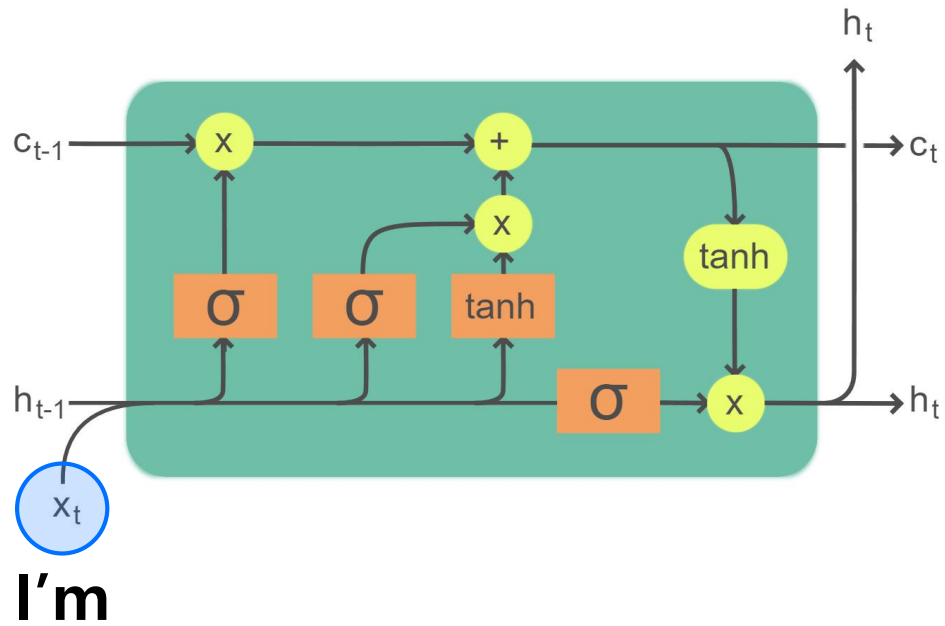
LSTM의 구조



LSTM의 구조

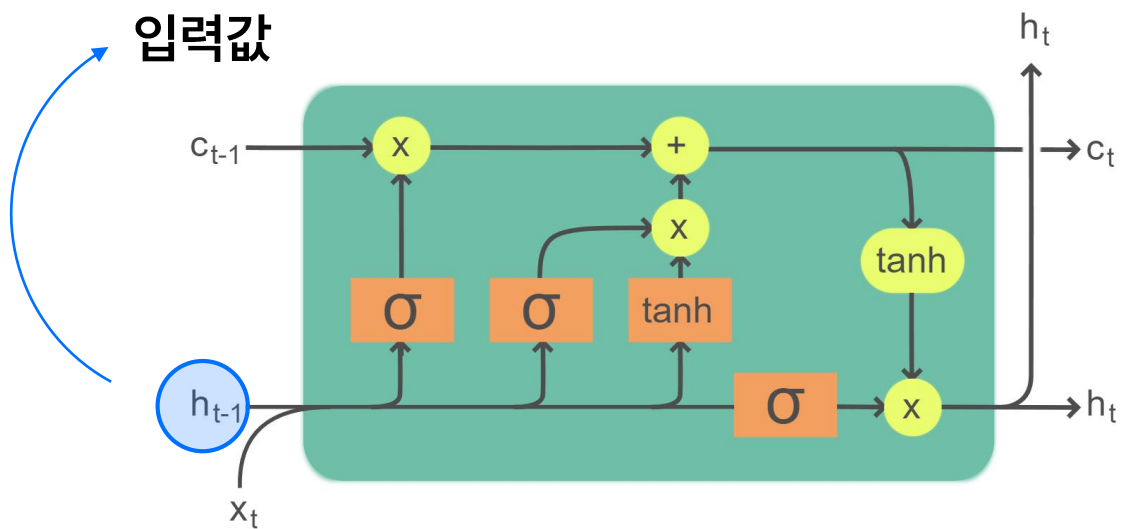


LSTM의 구조

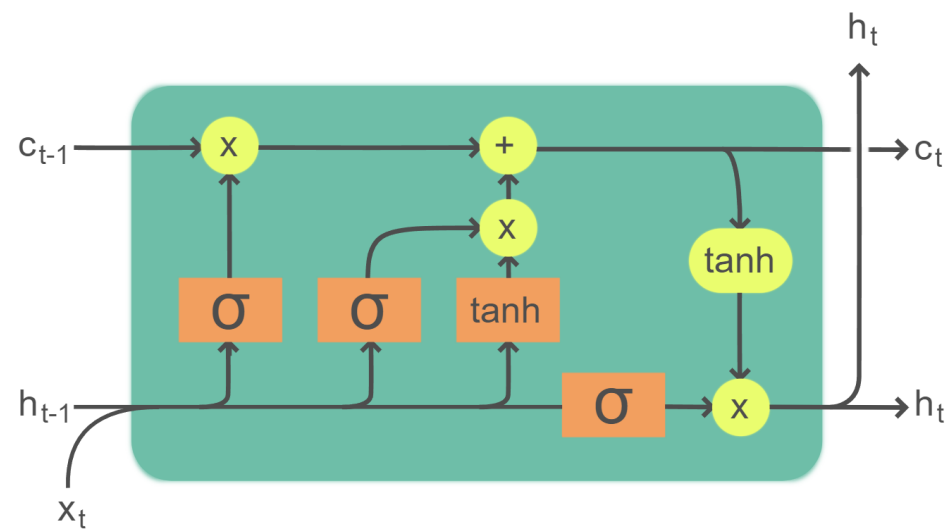


LSTM의 구조

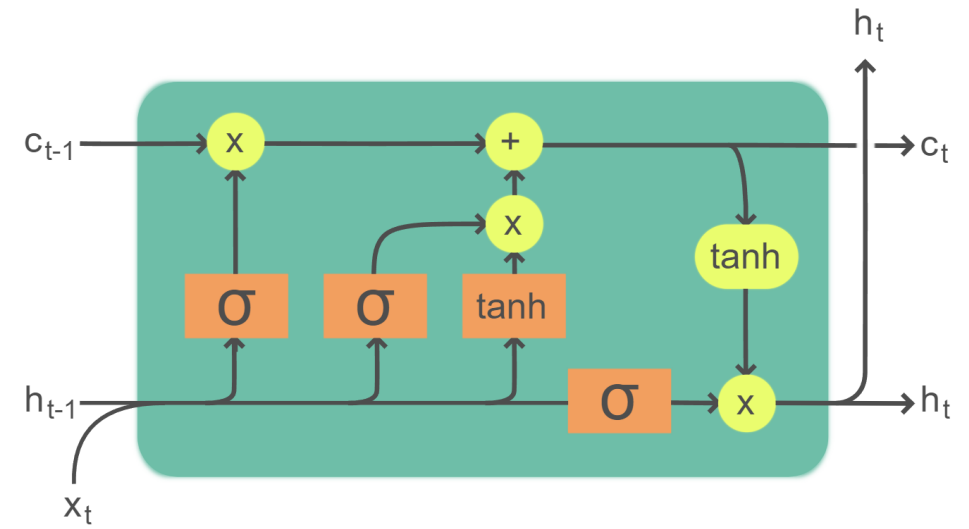
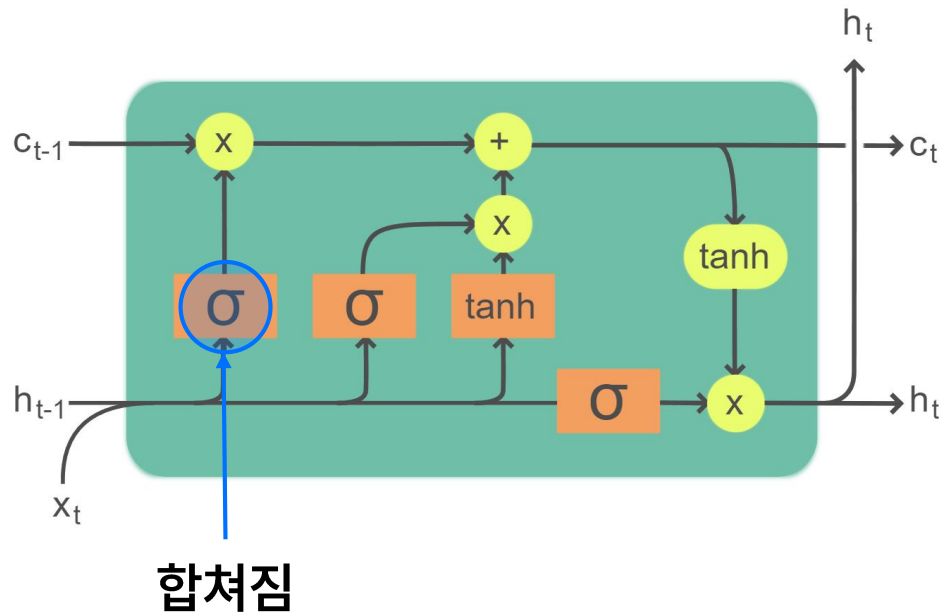
이전에서 연산한
입력값



I'm

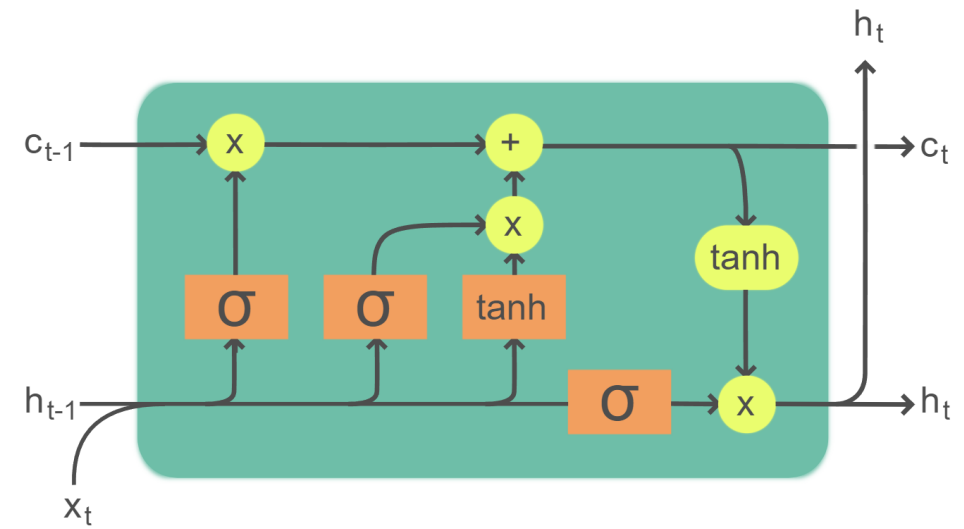
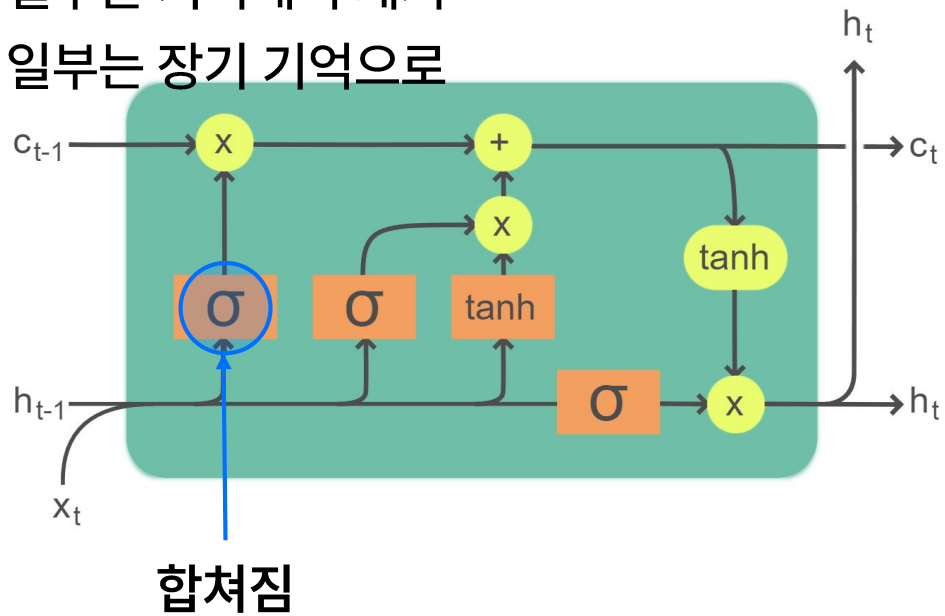


LSTM의 구조

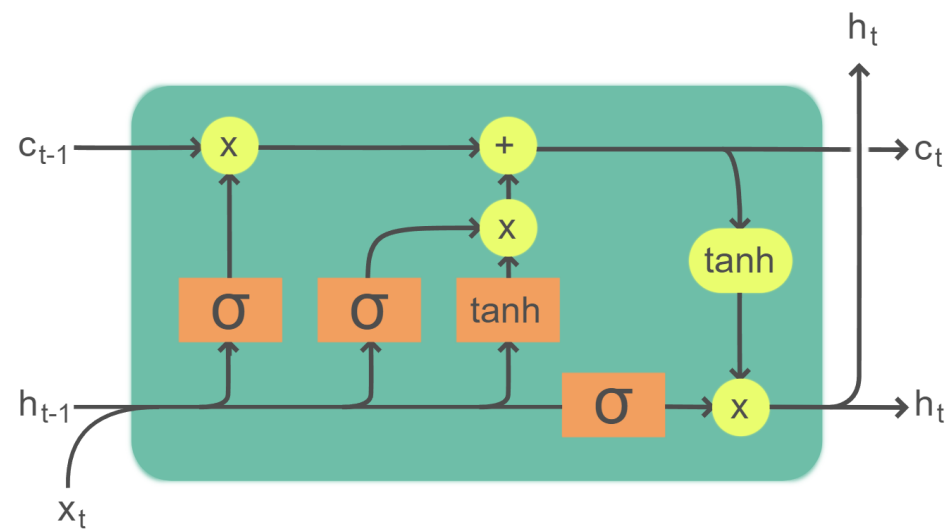
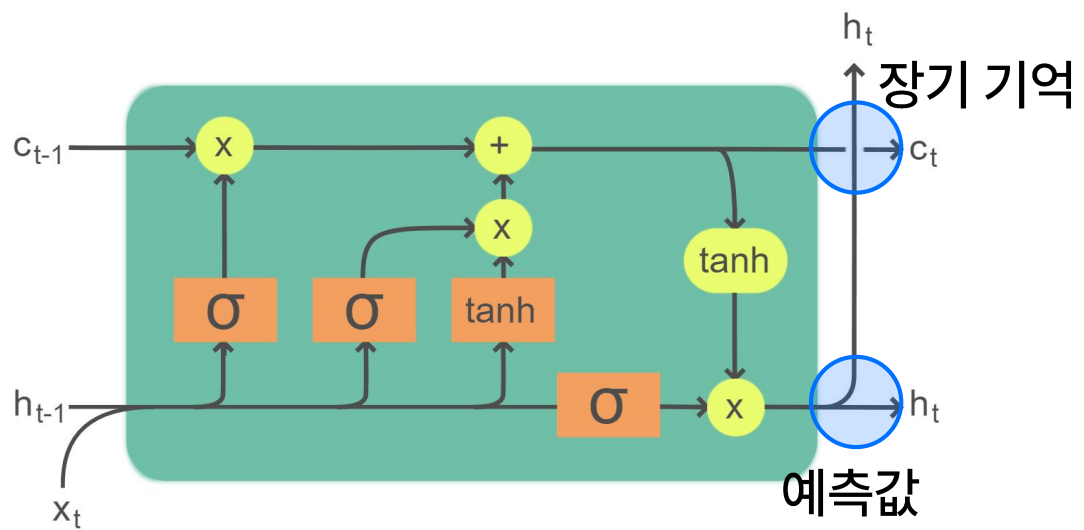


LSTM의 구조

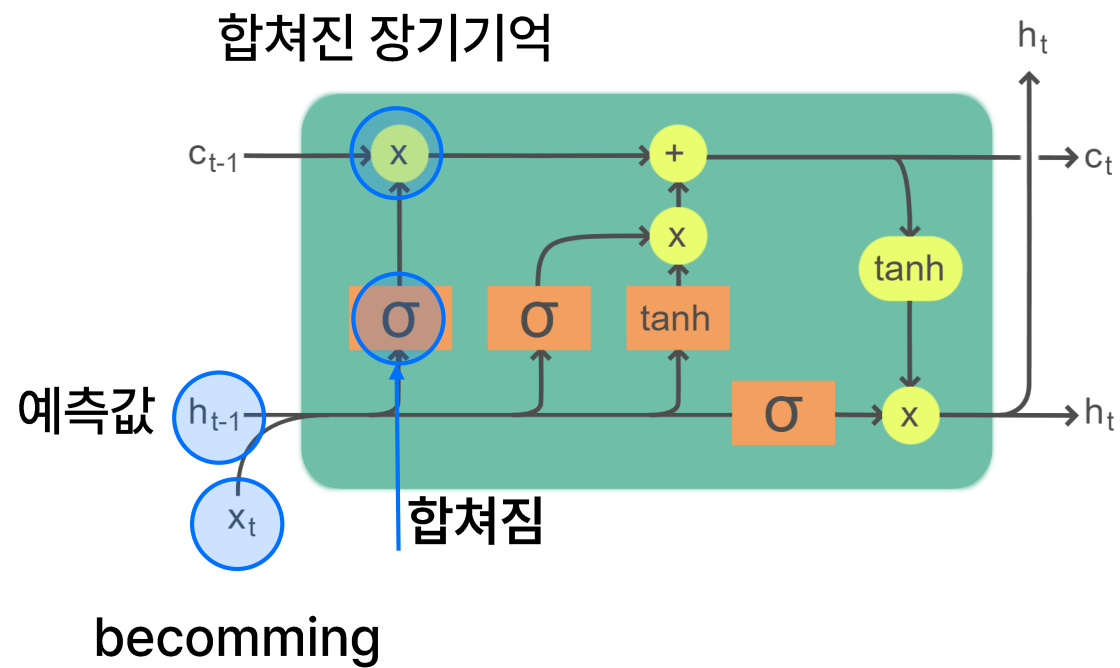
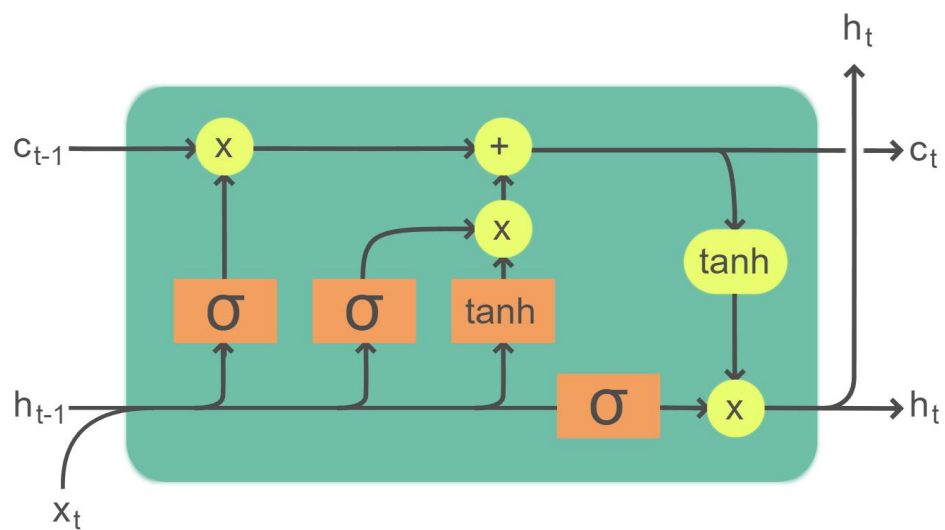
일부는 기억에서 제거
일부는 장기 기억으로



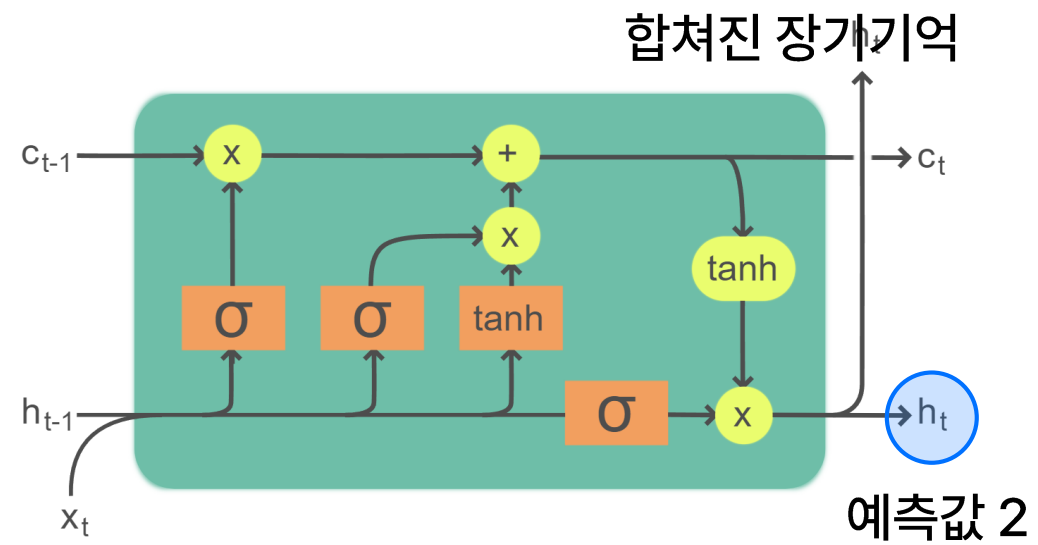
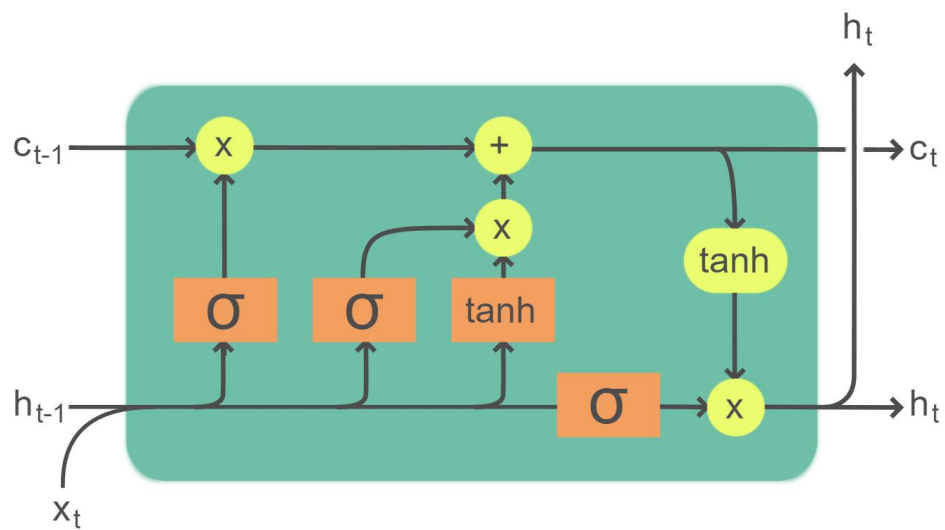
LSTM의 구조



LSTM의 구조

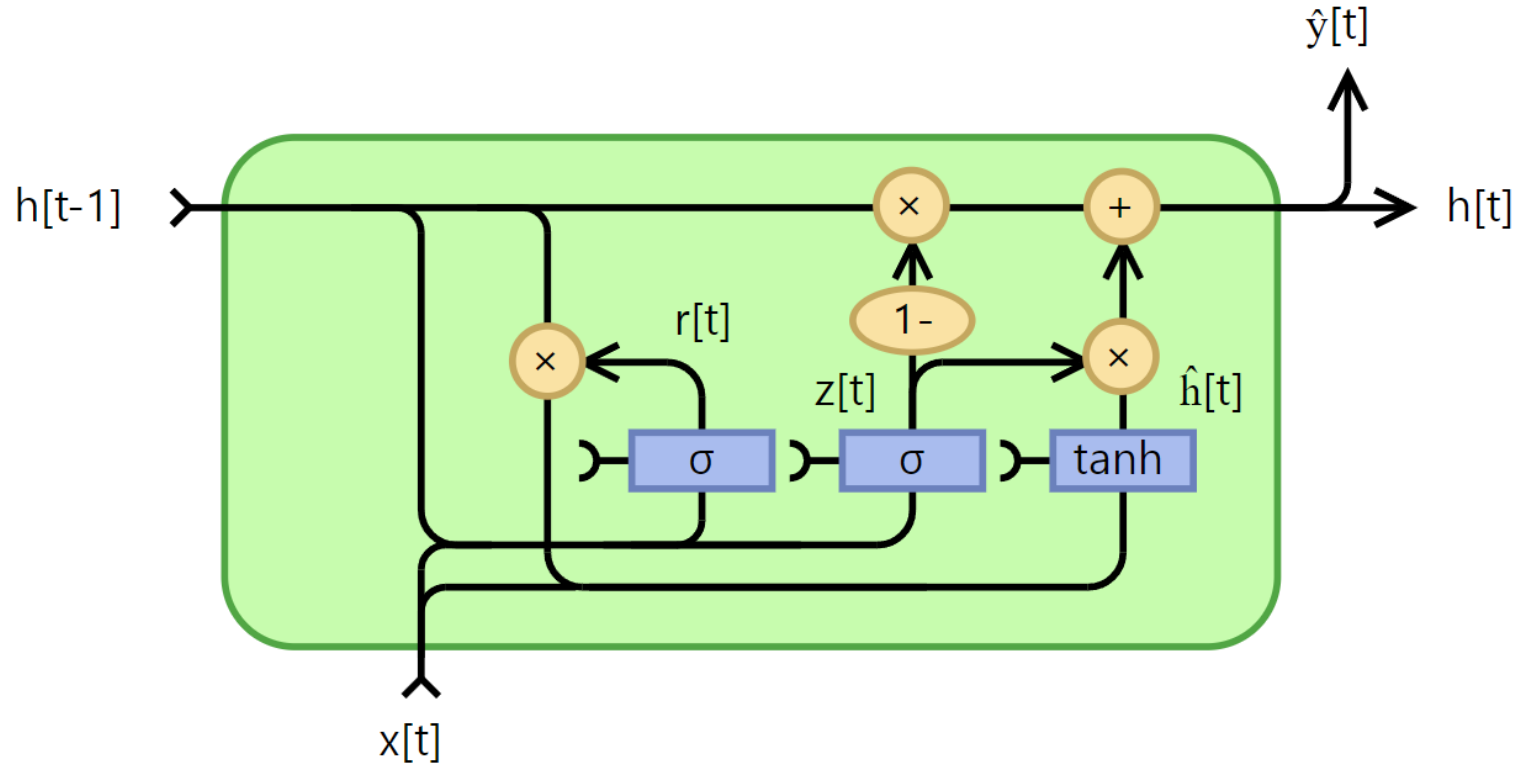


LSTM의 구조

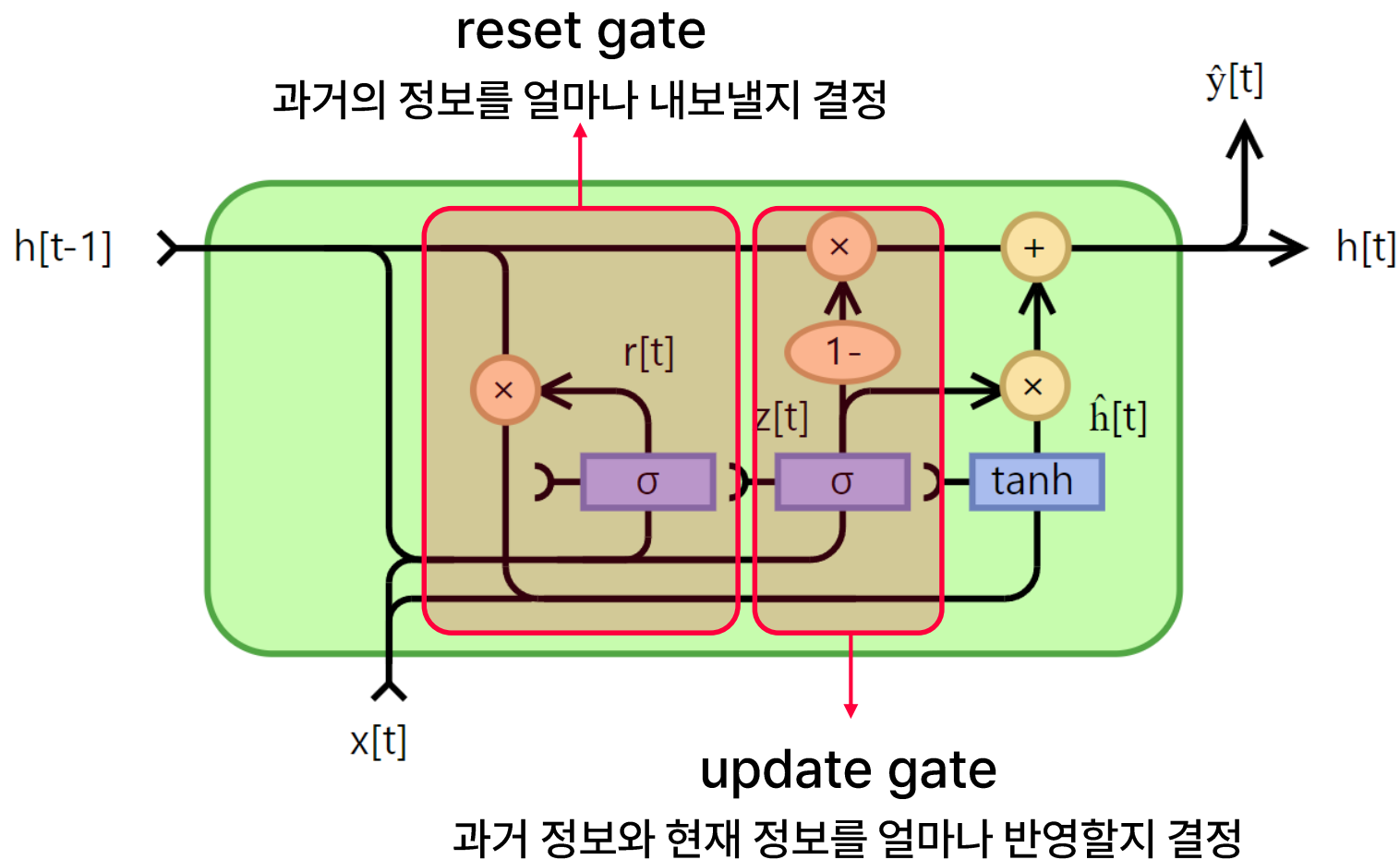


GRU(Gated Recurrent Unit) 모델

LSTM을 단순화한 GRU(Gated recurrent unit)



LSTM을 단순화한 GRU(Gated recurrent unit)



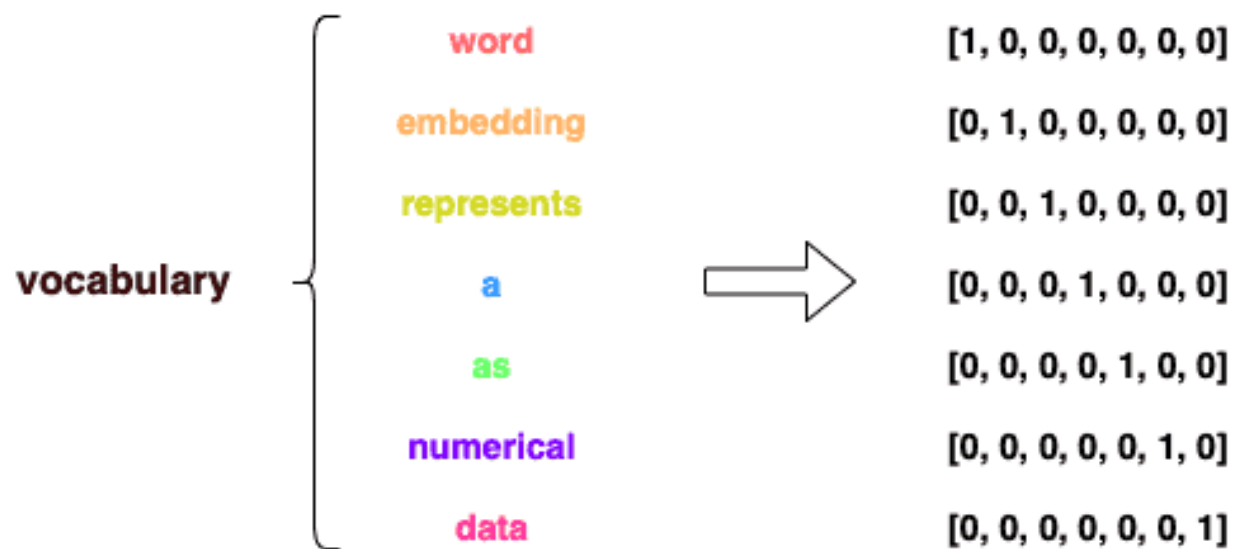
GRU 특징

- 일반적으로 LSTM보다 학습 속도가 더 빠름
- 많은 평가 지표에서 LSTM과 비슷한 성능을 보임
- 그러나, LSTM에서 최적의 하이퍼파라미터를 찾았다면, 굳이 모델을 변경할 필요는 없음

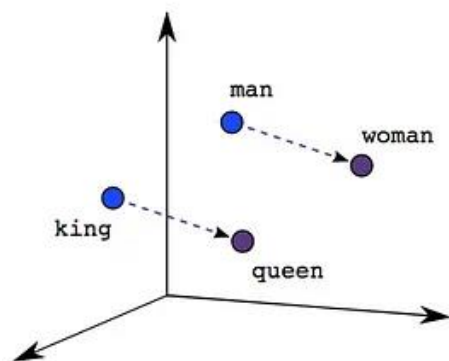
Word Embedding

텍스트를 모델이 이해할 수 있도록 숫자로 바꿔주는 작업

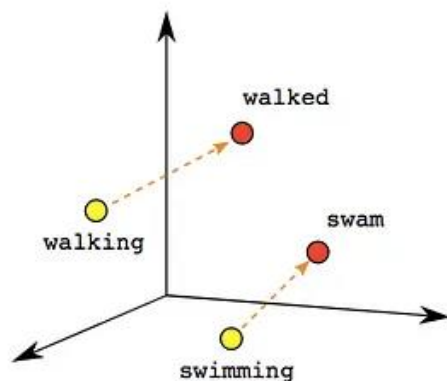
Sentence: Word embedding represents a word as numerical data.



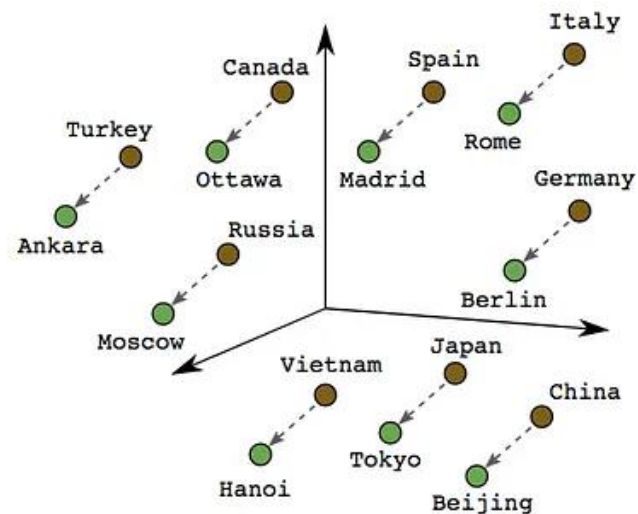
벡터형태이므로 좌표공간에 나타낼 수 있음



Male-Female



Verb Tense



Country-Capital

Sparse embedding vs Dense embedding

• Sparse Embedding

- 대부분의 값이 0, 몇몇 위치만 1인 벡터로 표현
- 문장에 나오는 단어의 빈도를 기준으로 벡터를 만든다.
- tf-idf, BM25 등
- 겹치는 단어가 있다면 유사도가 높게 나오겠지만 단어간 의미적인 관계를 포착하지 못한다.

• Dense Embedding

- 의미를 나타내는 실수 값들로 이루어진 벡터로 표현
- BERT와 같은 Pretrained Language Model이 주로 사용

Sparse embedding vs Dense embedding

“아이들이 공원에서 놀고 있다.”

- Sparse Embedding을 하면
 - 위 2개 문장을 전혀 다르게 봄

“어린이들이 파크에서 뛰어놀고 있어.”

- Dense Embedding을 하면
 - 위 2개 문장의 유사도가 매우 높게 나옴

따라서 요즘에는 Dense Embedding을 사용하는 추세



Q&A