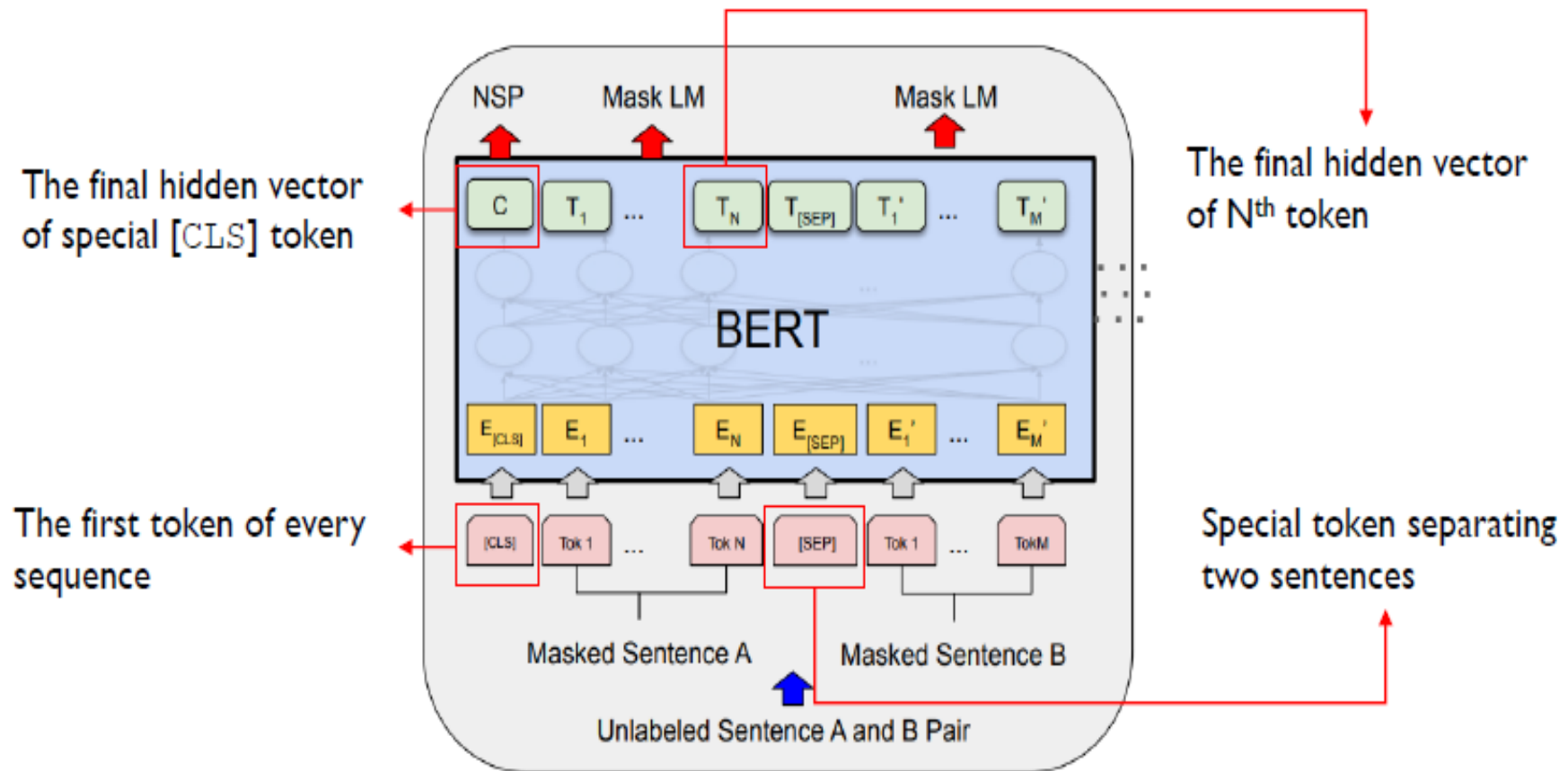


# **BERT(Bidirectional Encoder Representation Transformer)**

# BERT

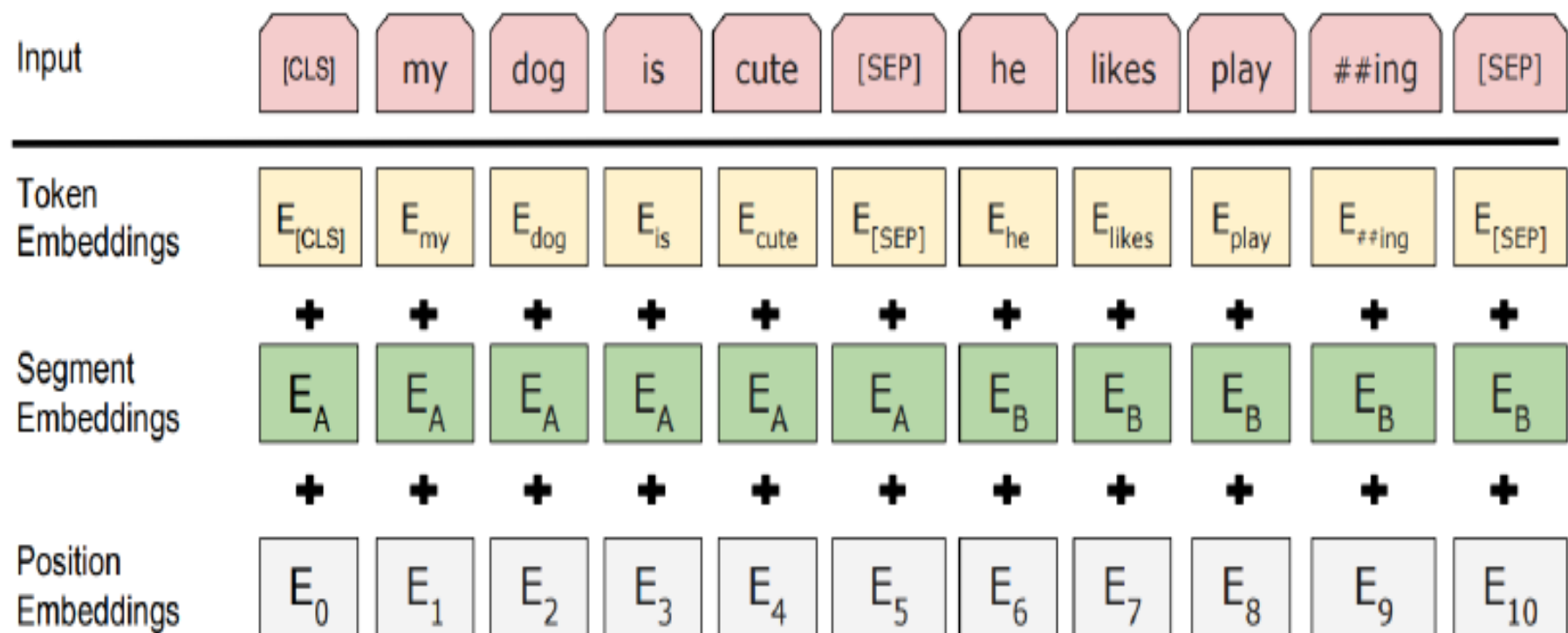
## □ BERT의 기본학습



# BERT

## □ BERT의 기본학습

- (1) 토큰 임베딩
- (2) 세그먼트 임베딩
- (3) 포지션 임베딩



# BERT

## □ BERT의 기본학습

### (1) 세그먼트 임베딩

input 1                      input 2  
" I like cats "      " I like dogs " , 2 inputs

↓ ① concat and tokenize

[CLS] I like cat [SEP] I like dogs , 8 tokens

↓ ② label to distinguish input

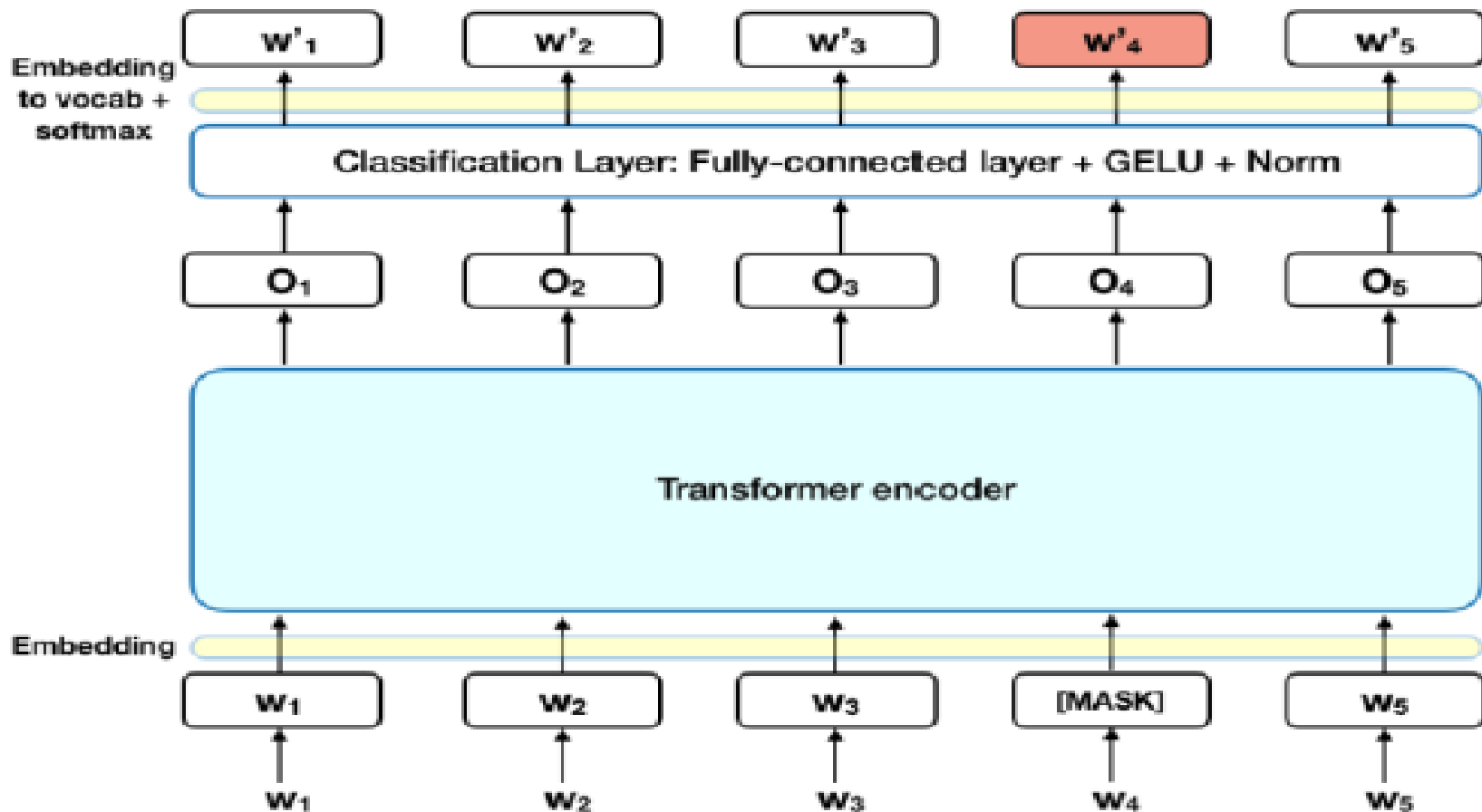
[CLS] I like cat [SEP] I like dogs , 0 : input 1  
0    0    0    0    0    1    1    1                      1 : input 2

↓ ③ Lookup vector representation

2 { Segment Embeddings  
768

# BERT

- BERT의 Task 1 : Masked Language Model (MLM)
  - 각 문자의 15%가 [MASK] 토큰으로 대체된다.
  - Mask가 되기로 결정됐을 때 80% 마스크 10% 랜덤 10% 그대로



# BERT

## □ BERT의 Task 2: 다음 문자 예측 (NSP)

**Monica:** This is harder than I thought it would be.

**Chandler:** Oh, it is gonna be okay.

**Rachel:** Do you guys have to go to the new house right away, or do you have some time?

**Monica:** We got some time.

**Rachel:** Okay, should we get some coffee?

**Chandler:** Sure. Where?

IsNext

NotNext

[C]



[CLS] This is harder than I thought it would be. [SEP] Oh, it is gonna be okay

# BERT

## □ BERT의 Task 2: 다음 문자 예측 (NSP)

**Monica:** This is harder than I thought it would be.

**Chandler:** Oh, it is gonna be okay.

**Rachel:** Do you guys have to go to the new house right away, or do you have some time?

**Monica:** We got some time.

**Rachel:** Okay, should we get some coffee?

**Chandler:** Sure. Where?



# BERT

- BERT의 Task 2: 다음 문자 예측 (NSP)

