

# GPT3 모델 개요

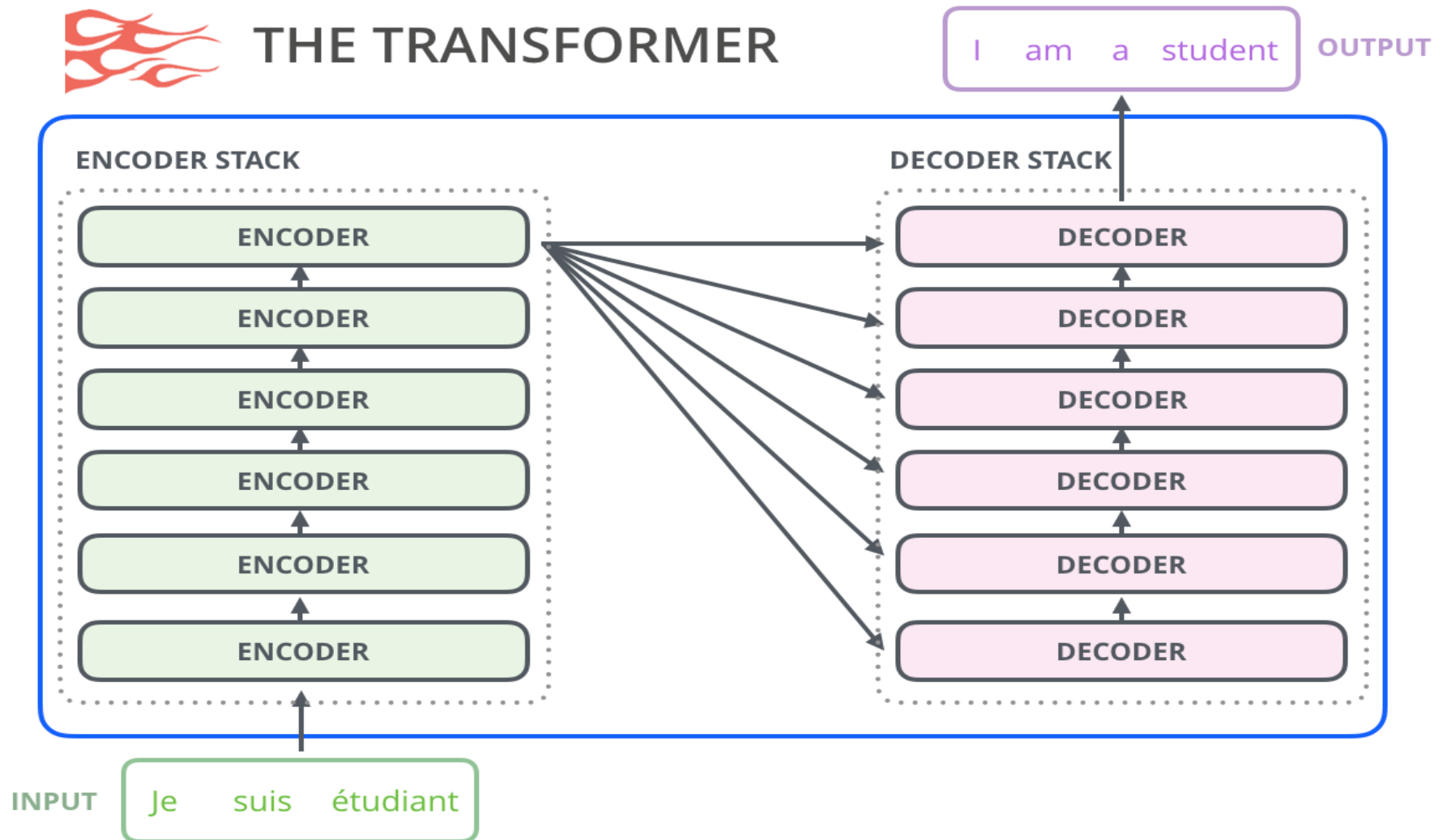
*트랜스포머 디코더*

*Dec 2020*

# GPT3 개요

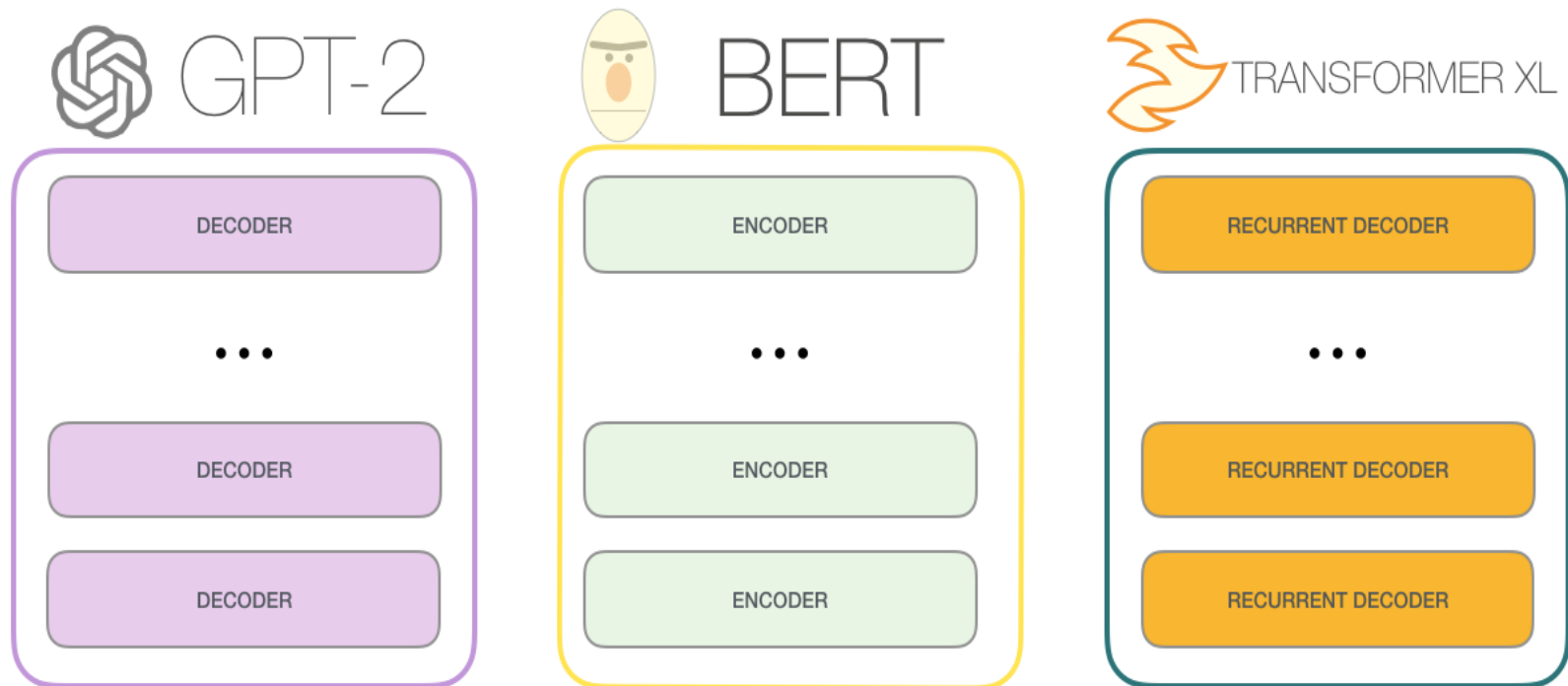
# GPT3 구조

- GPT3 구조는 트랜스포머의 디코더 모델이다.



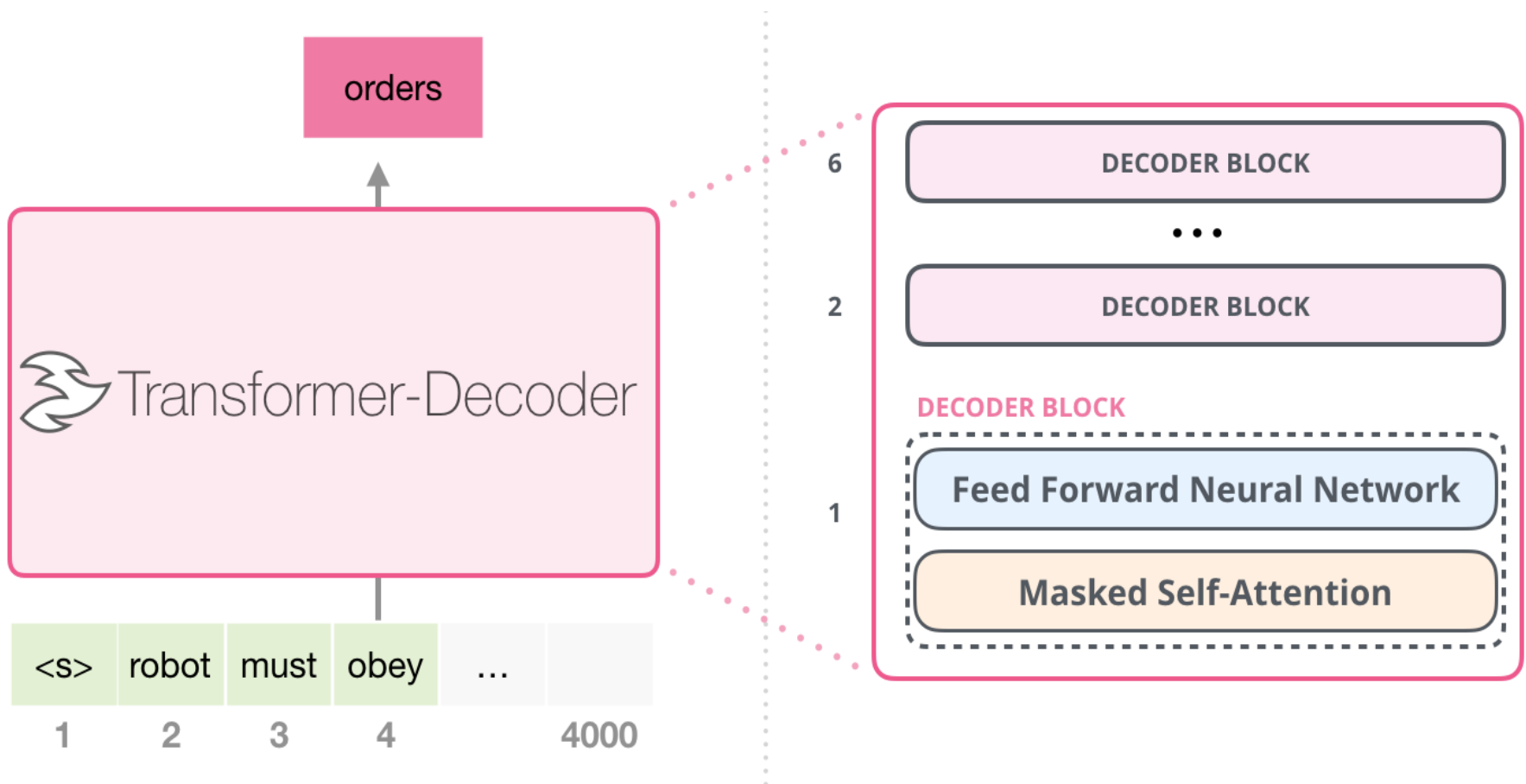
# GPT3 구조

- GPT3 구조는 트랜스포머의 디코더 모델이다.



# GPT3 구조

- GPT3 구조는 트랜스포머의 디코더 모델이다.



# GPT3 언어모델

---

- GPT 언어모델 예제

Input Prompt:

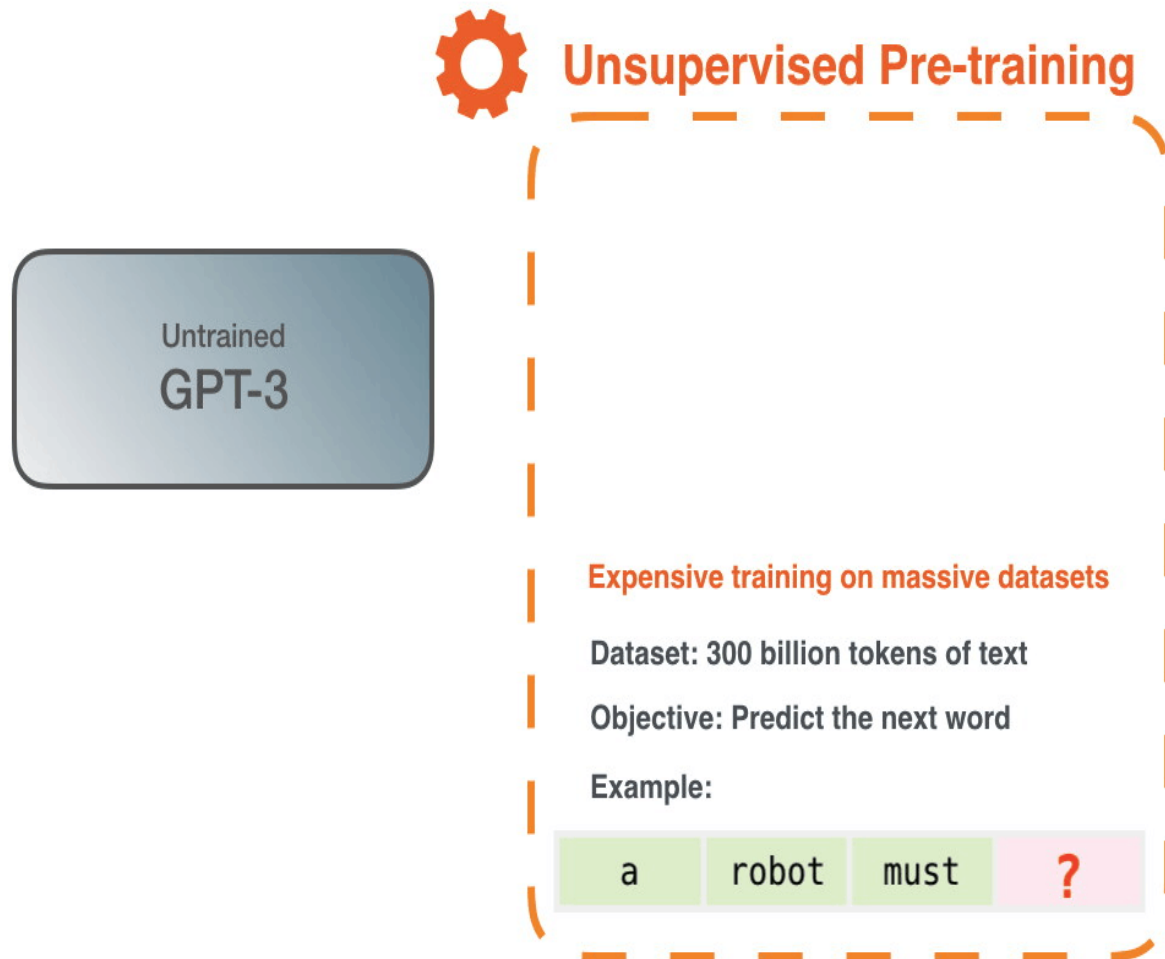
Recite the first law of robotics



Output:

# GPT3 언어모델

- GPT 언어모델 훈련



# GPT3 언어모델

- 훈련예제 – 슬라이딩 윈도우

**Text:** Second Law of Robotics: A robot must obey the orders given it by human beings



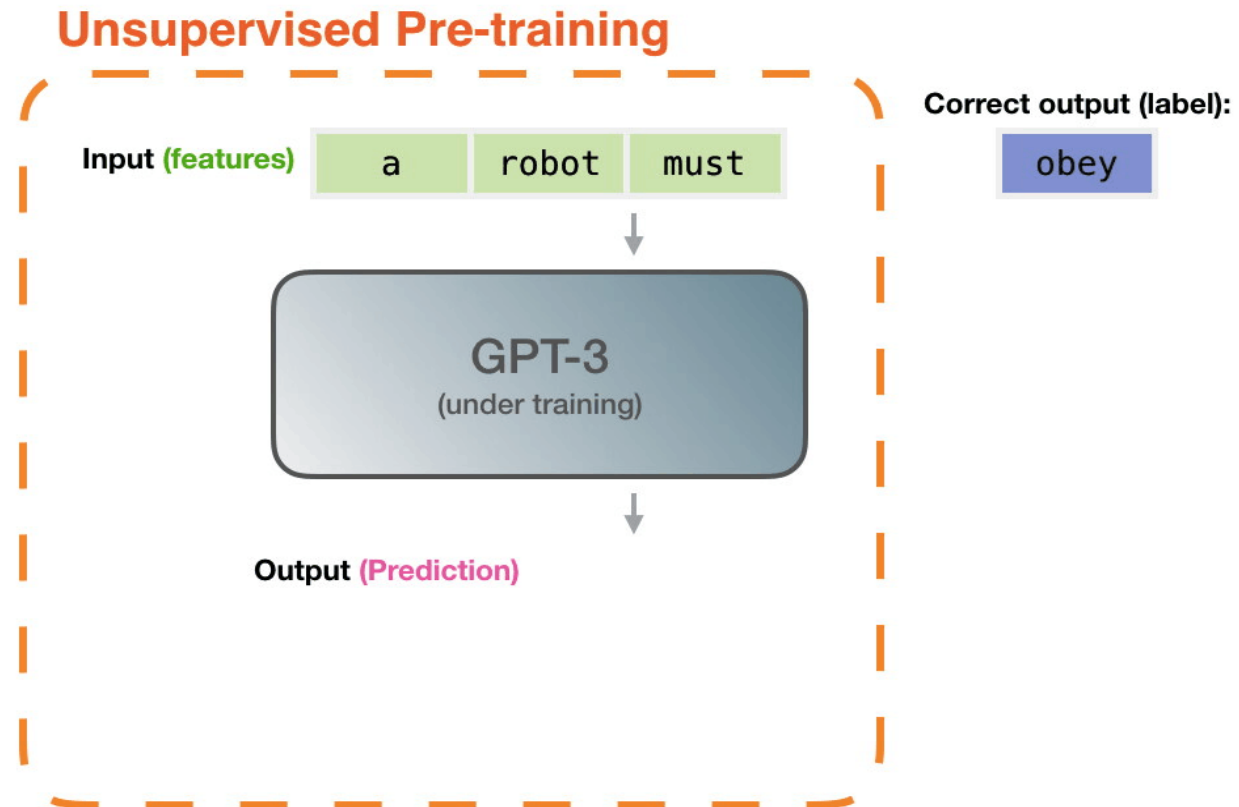
**Generated training examples**

Example #	Input (features)	Correct output (labels)
1	Second law of robotics :	a
2	Second law of robotics : a	robot
3	Second law of robotics : a robot	must
...		



# GPT3 언어모델

- GPT 훈련 스텝 - 역전파



# GPT3 언어모델

---

- GPT3 세부 프로세스 – 한번에 한 토큰(단어)씩 출력

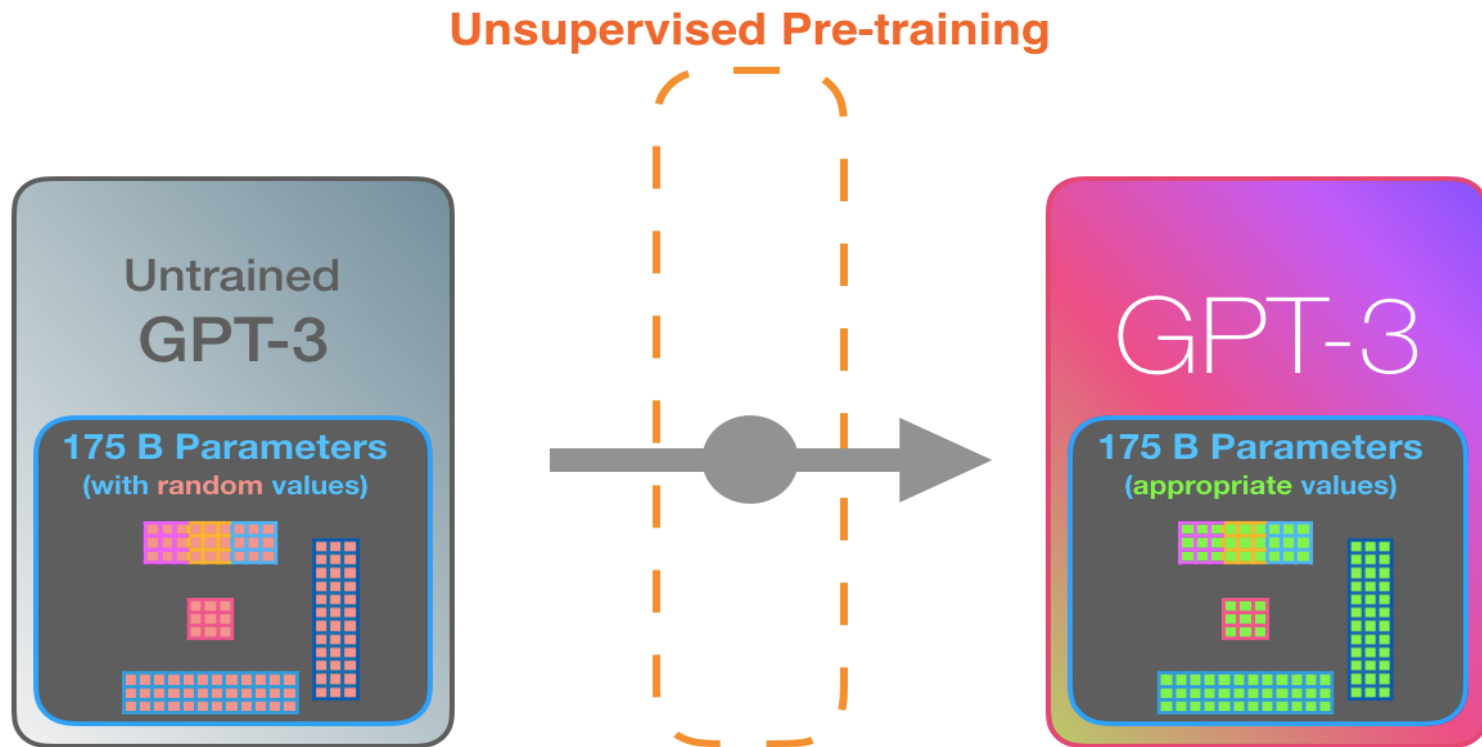
Input Prompt: Recite the first law of robotics



Output:

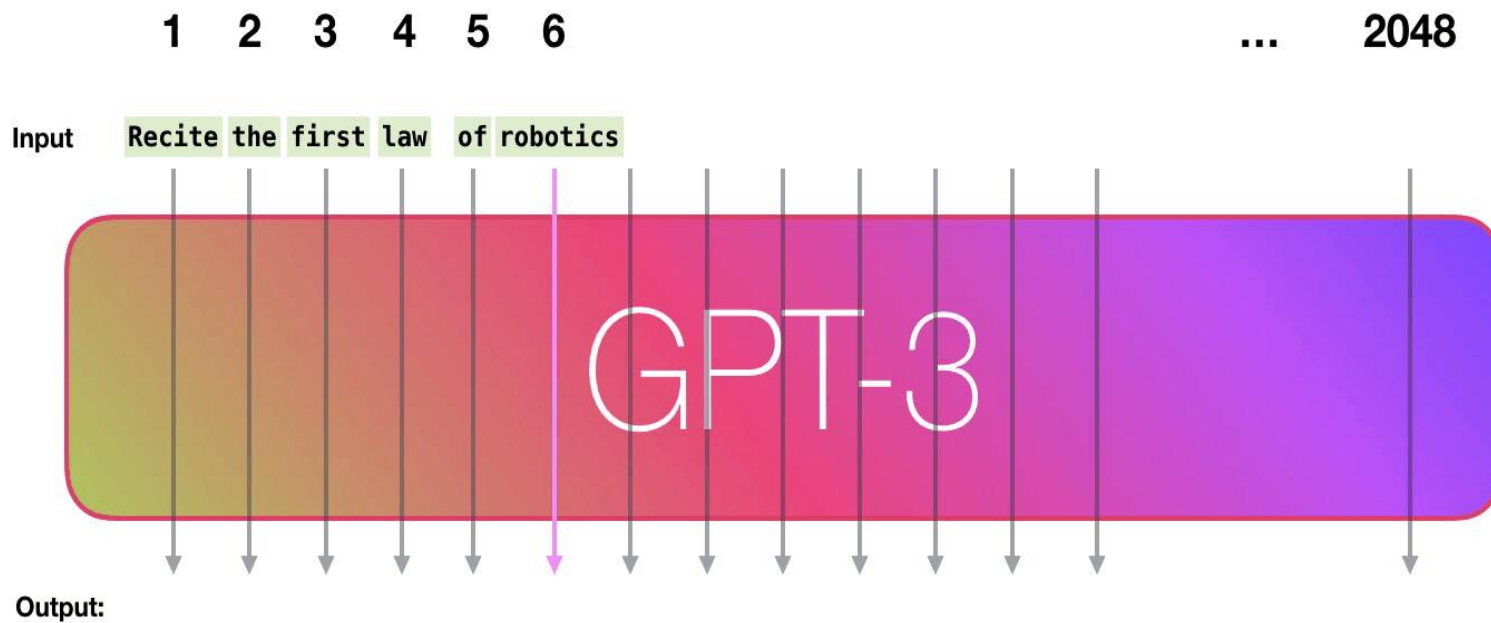
# GPT3 언어모델

- GPT3 파라미터 - 175B개의 파라미터로 다음에 무엇이 나올지를 계산한다.  
(파라미터 가중치는 랜덤 파라미터로 시작한다)



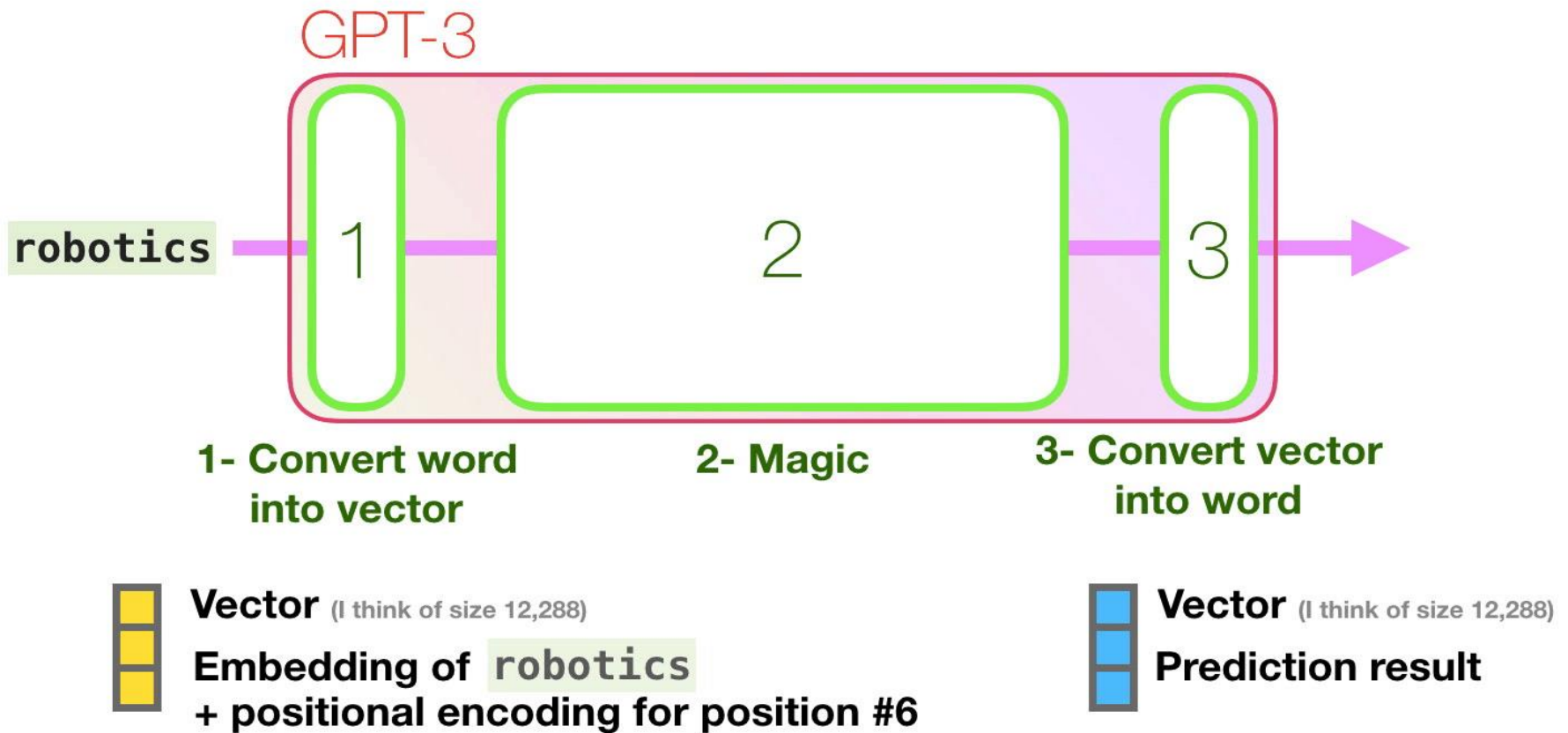
# GPT3 언어모델

- GPT3는 2048 너비의 컨텍스트 윈도우를 갖는다.



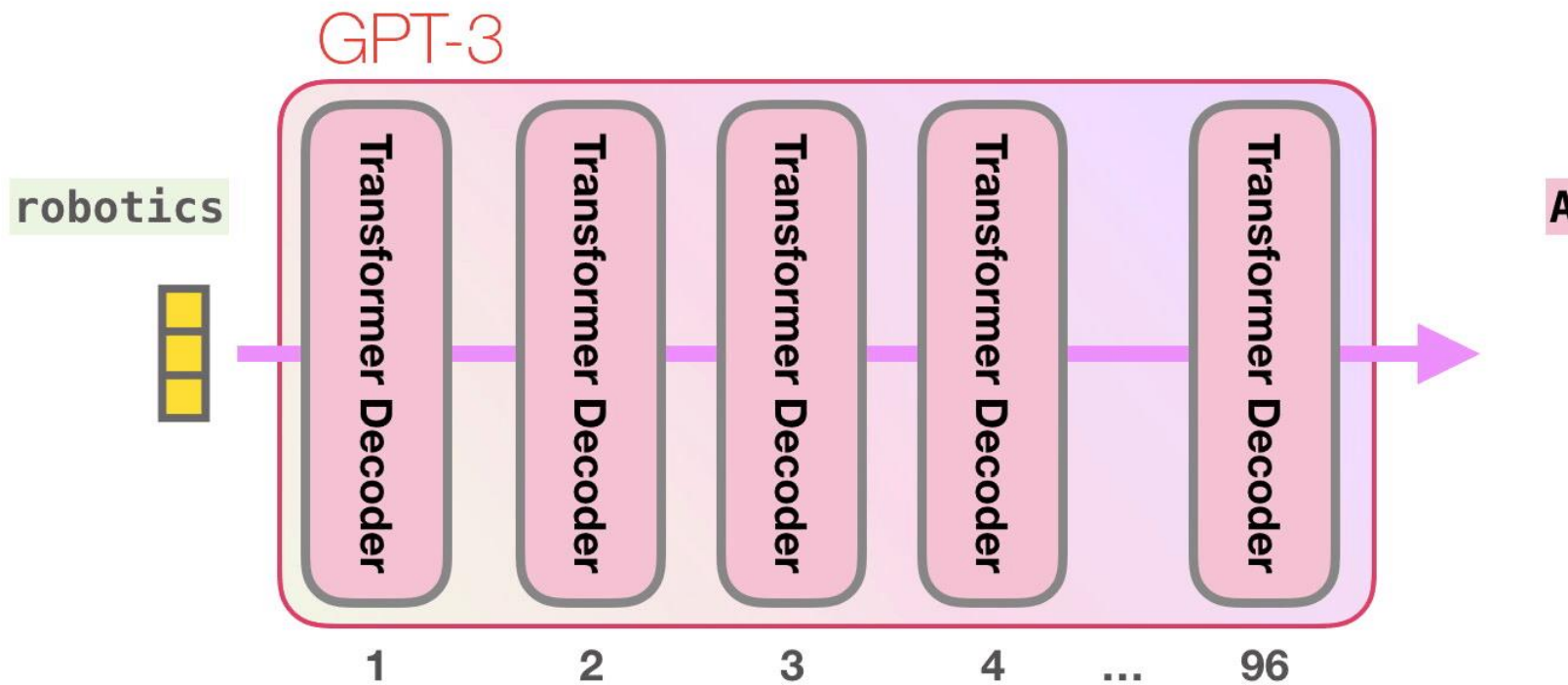
# GPT3 언어모델

- GPT3 임베딩 – 어떻게 단어 (예: robotics)를 처리해서 “A”를 생성하는가?



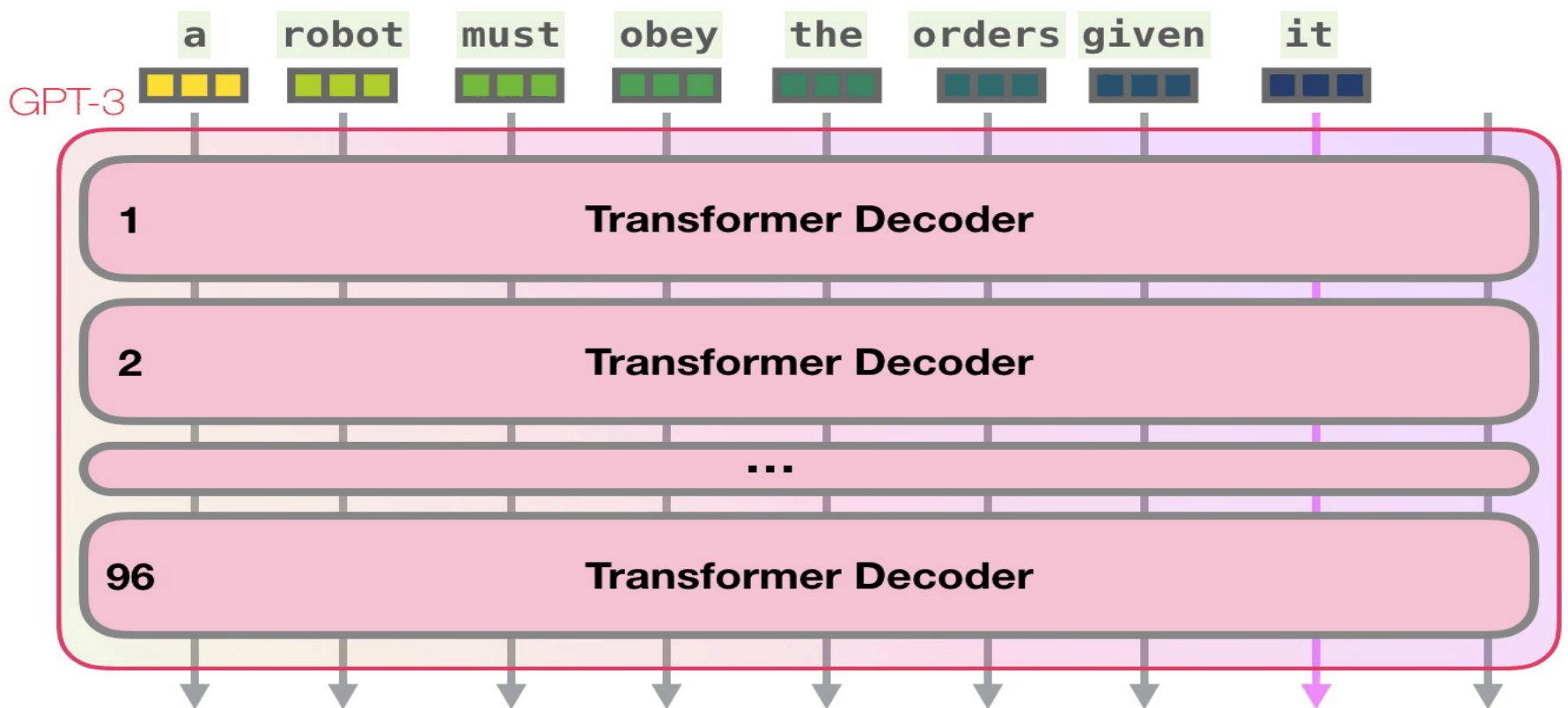
# GPT3 언어모델

- GPT3는 96개의 트랜스포머 블록을 이용 (각층이 1.8B개의 파라미터)



# GPT3 언어모델

- GPT3 토큰의 트랜스포머 블록으로의 입력과 그에 대한 반응("Okay human") 예시



# GPT3 언어모델

- GPT3 리액트 코드 생성 예제

[example] an input that says "search" [toCode] Class App extends React Component... </div> } } }

[example] a button that says "I'm feeling lucky" [toCode] Class App extends React Component...

[example] an input that says "enter a todo" [toCode]



GPT-3





# GPT3 언어모델

- GPT3 파인튜닝 – 특정 작업에 더 좋은 모델을 생성한다.

