

Transformer 모델

학번	201984053	201904059
이름	이상윤	신종훈



목 차

1. 서론

- Transformer 모델 소개
- Transformer 중요성, 역할

2. 이론적 배경

- Self-Attention
- Positional Encoding

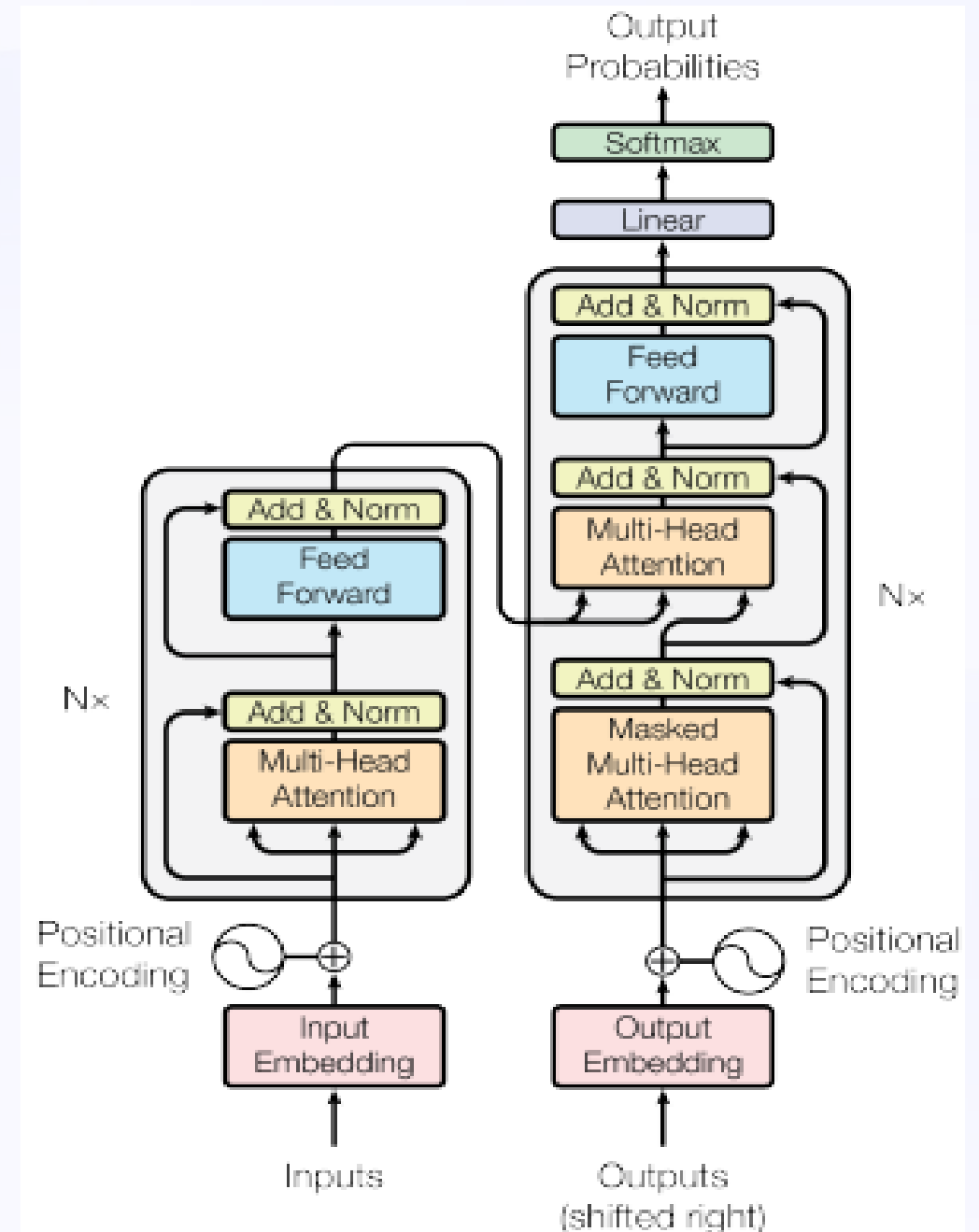
3. Transformer 구조 분석

- Encoder와 Decoder
- Multi-Head Attention
- Feed-Forward Networks
- Residual Connections & Layer Normalization

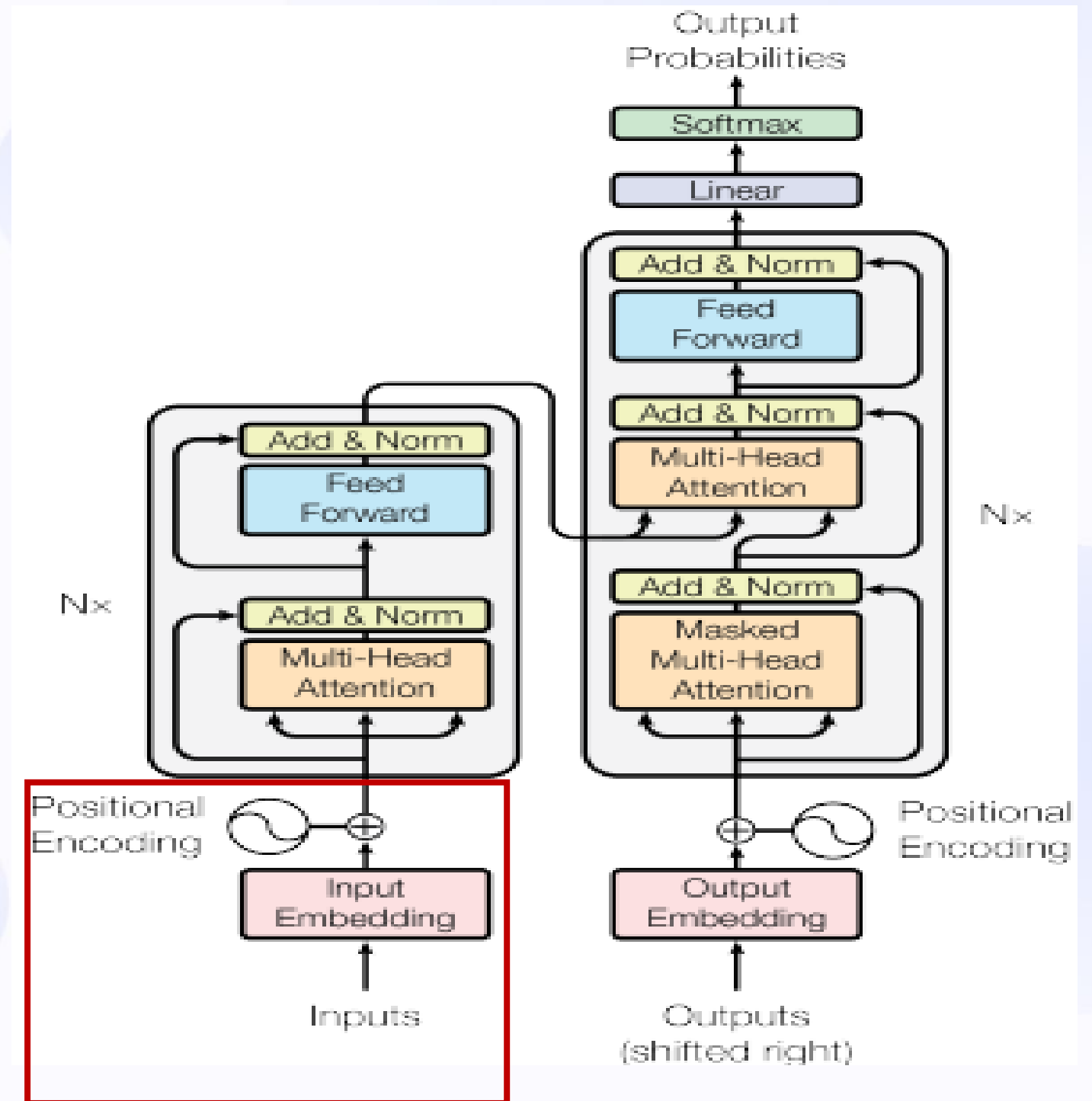
서론



Transformer 모델의 소개



Encoder-Decoder Structure

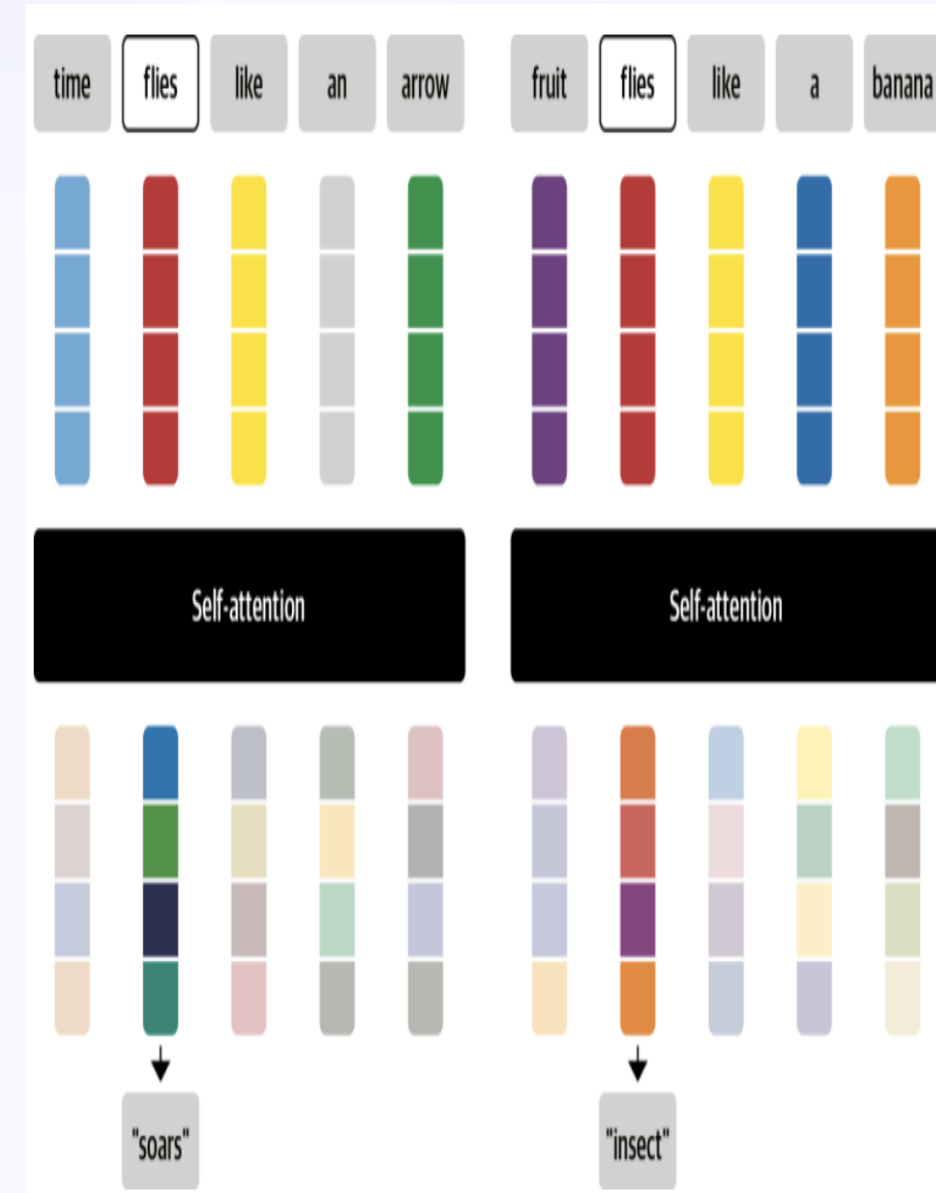
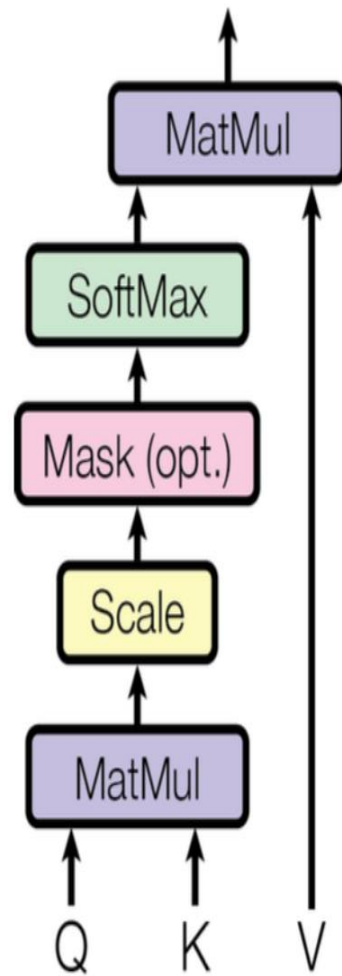


이론적 배경



Self-Attention

Scaled Dot-Product Attention



Positional Encoding

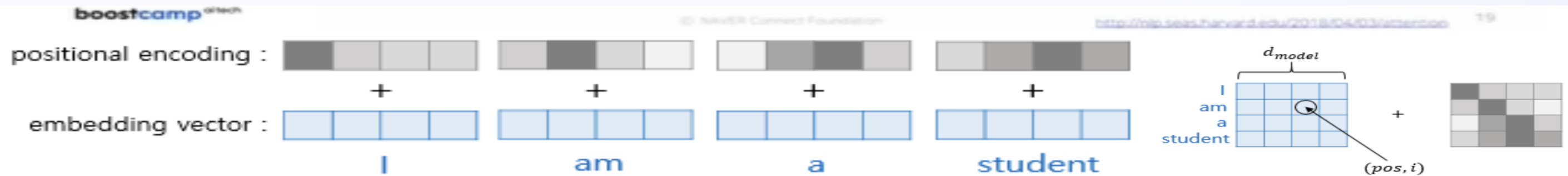
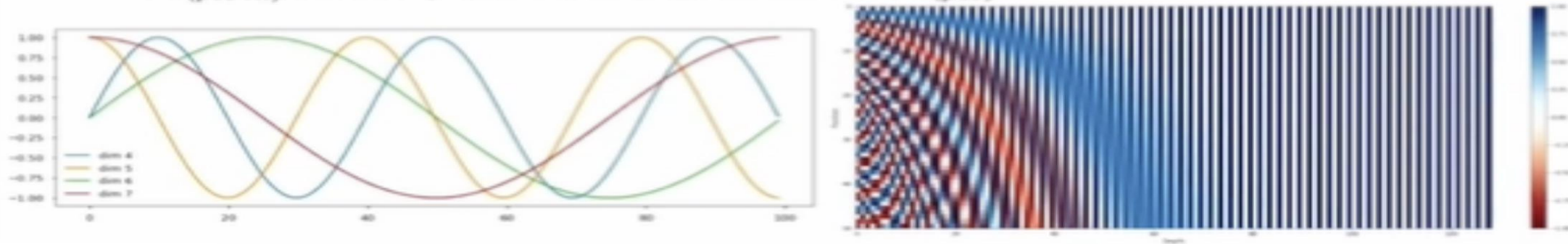
- Use sinusoidal functions of different frequencies

$$PE_{(pos,2i)} = \sin(pos/10000^{2i/d_{model}})$$

$$PE_{(pos,2i+1)} = \cos(pos/10000^{2i/d_{model}})$$

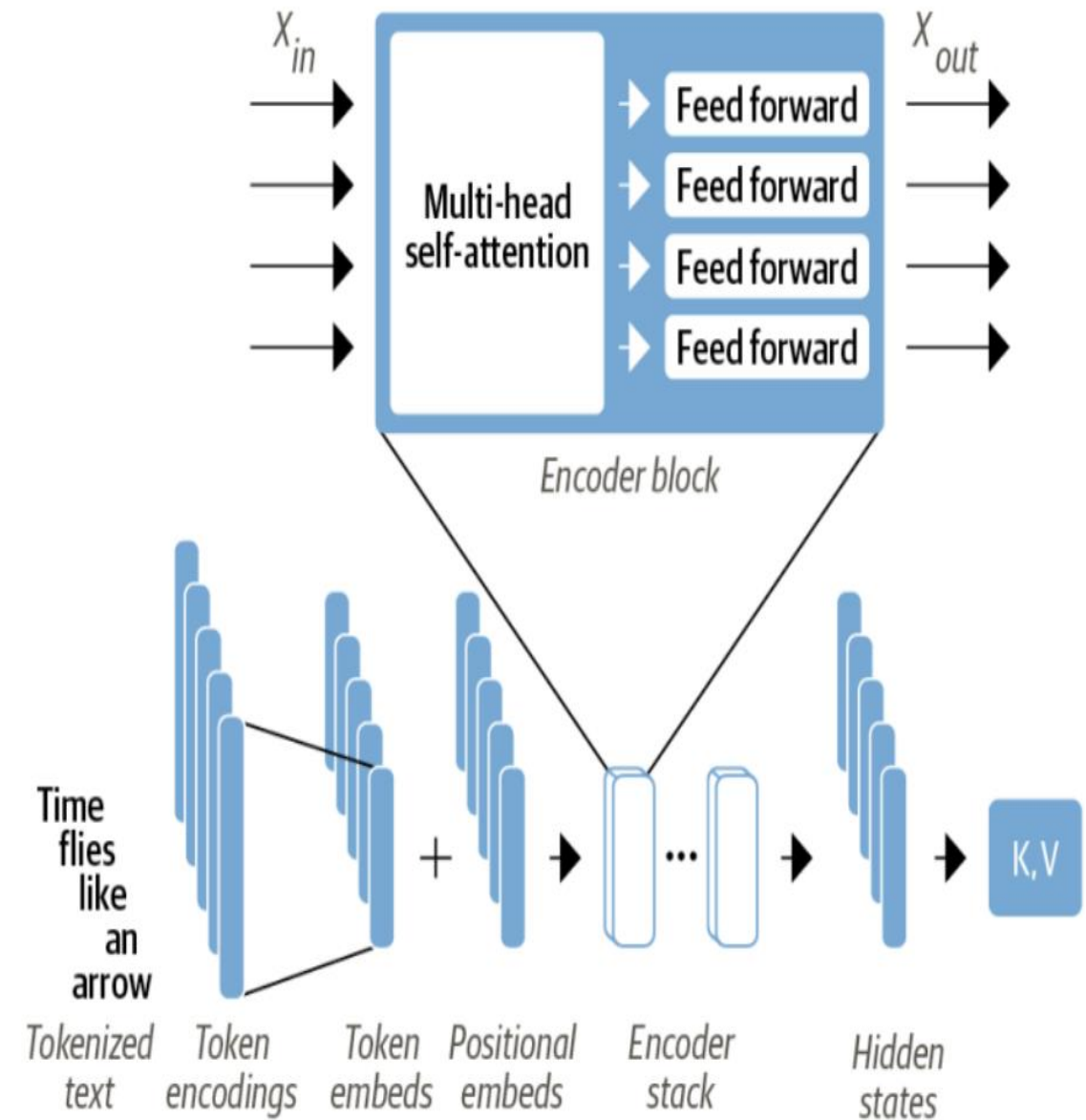
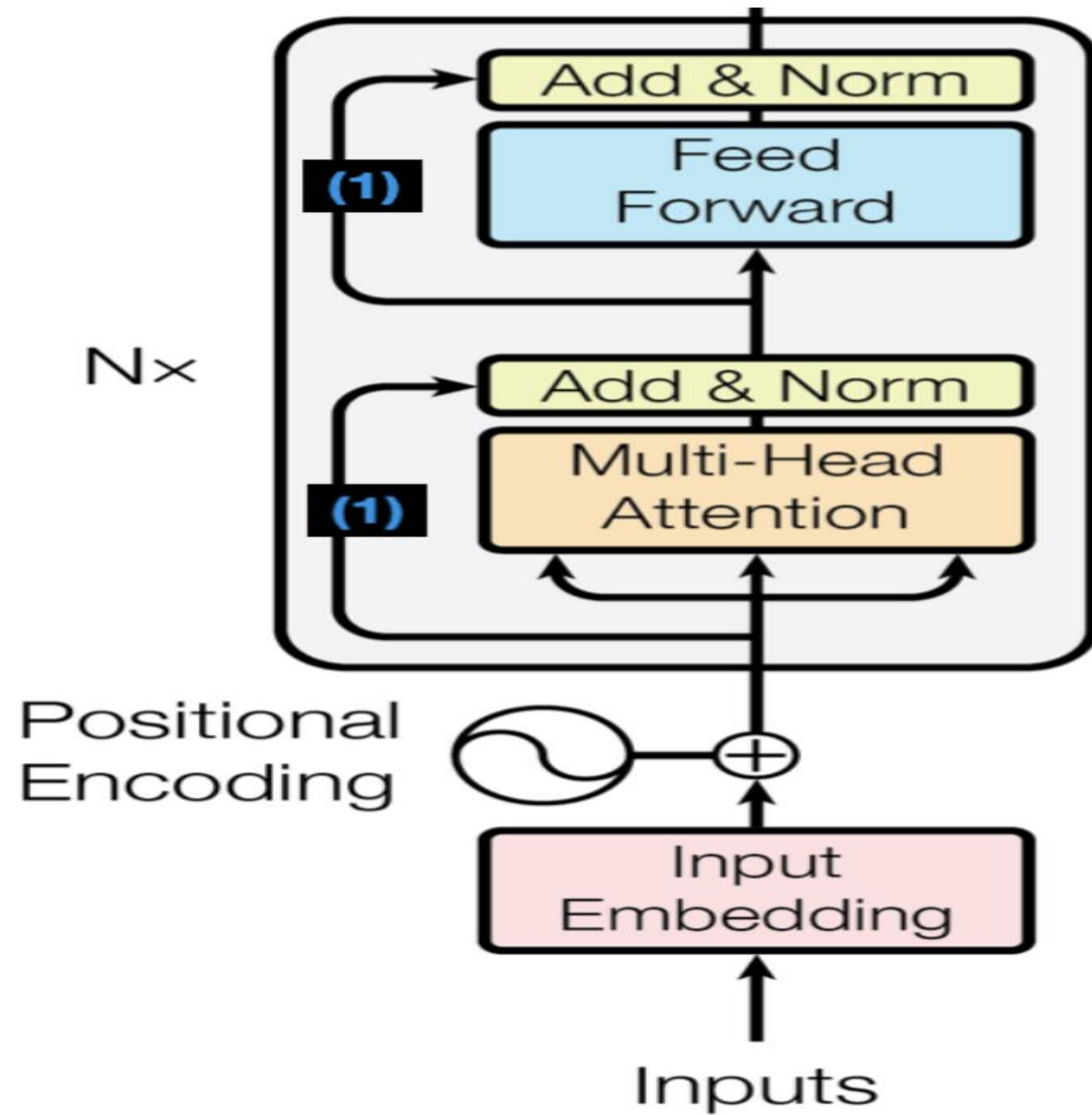
- Easily learn to attend by relative position, since for any fixed offset k ,

$PE_{(pos+k)}$ can be represented as linear function of $PE_{(pos)}$

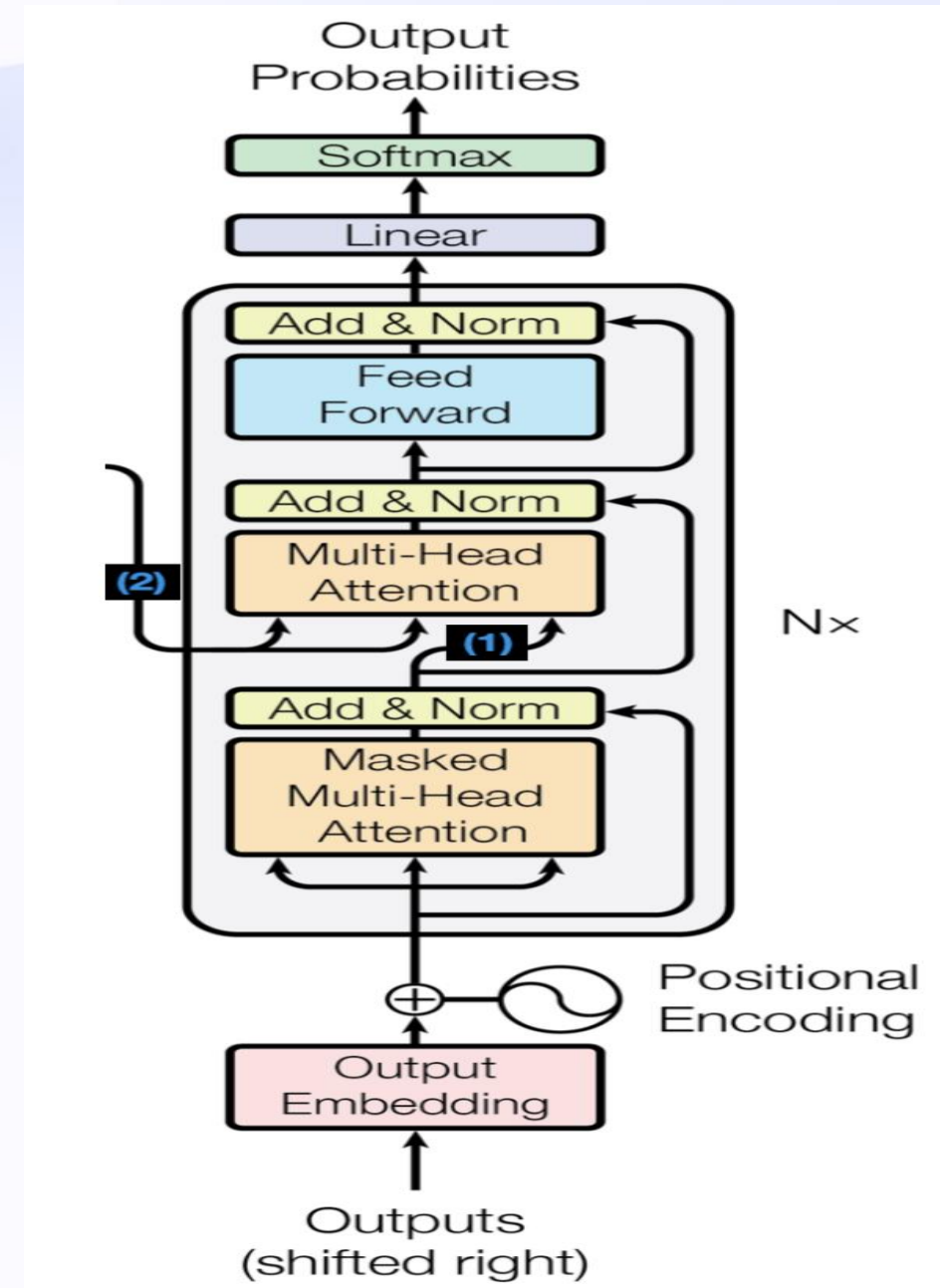


Transformer 구조 분석

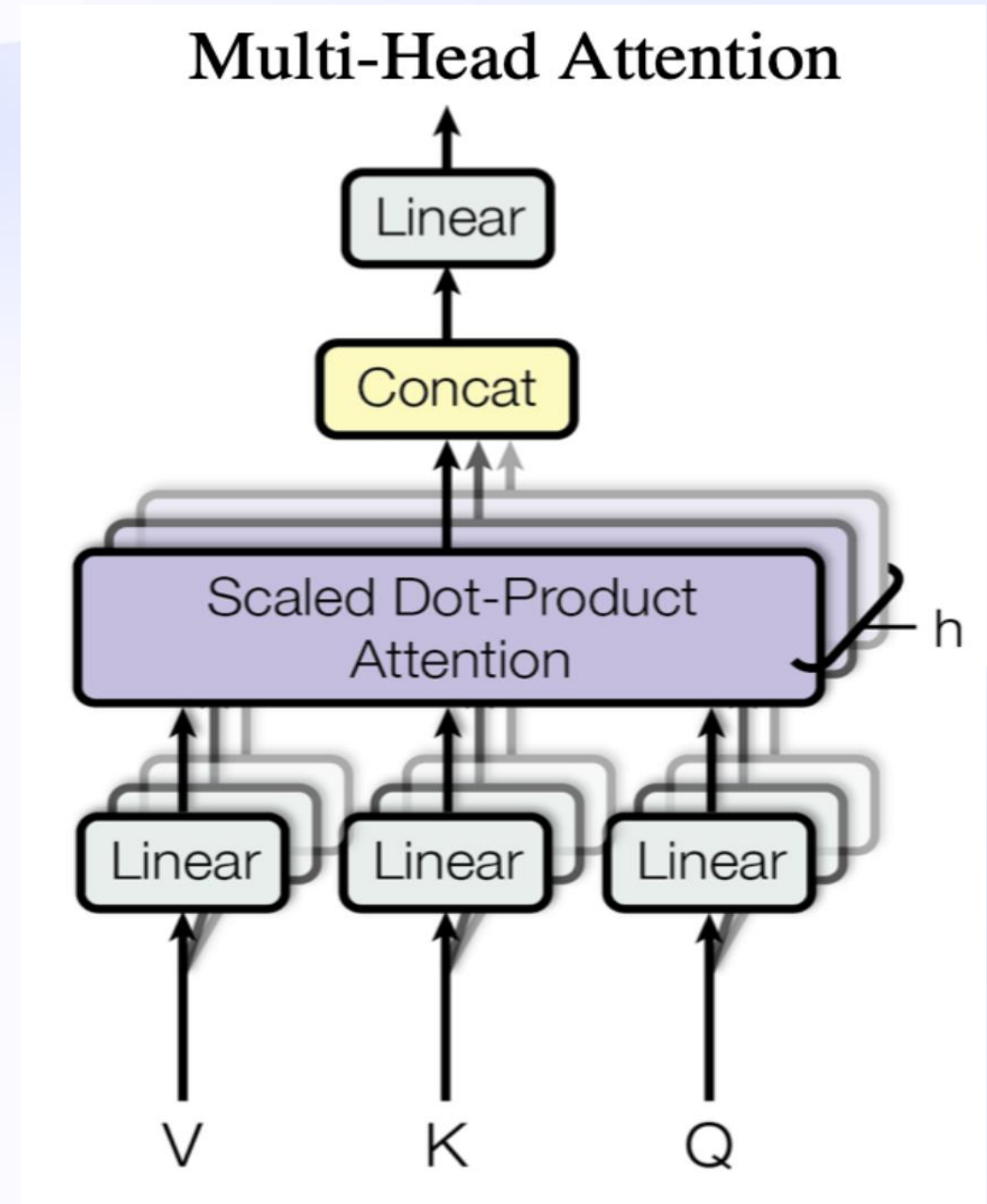
인코더



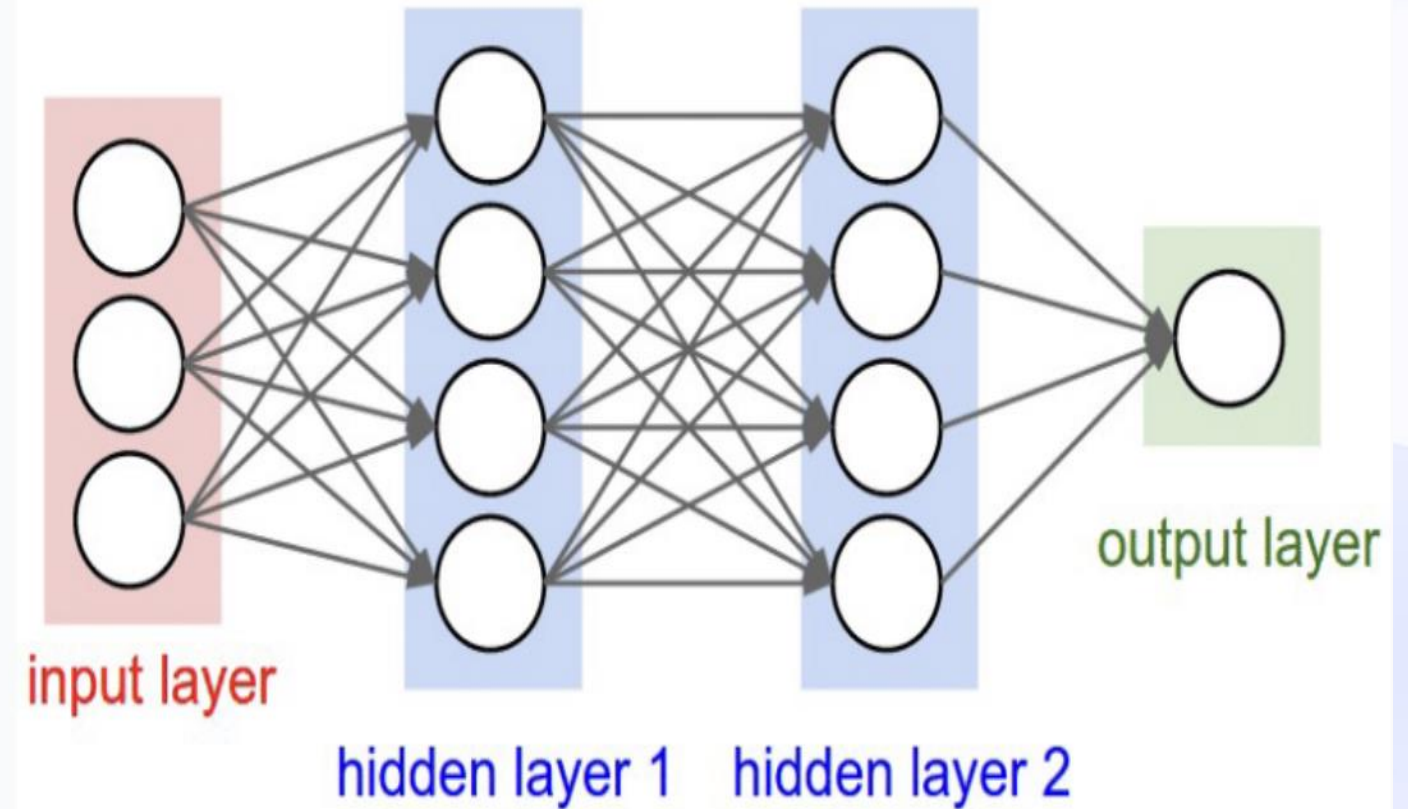
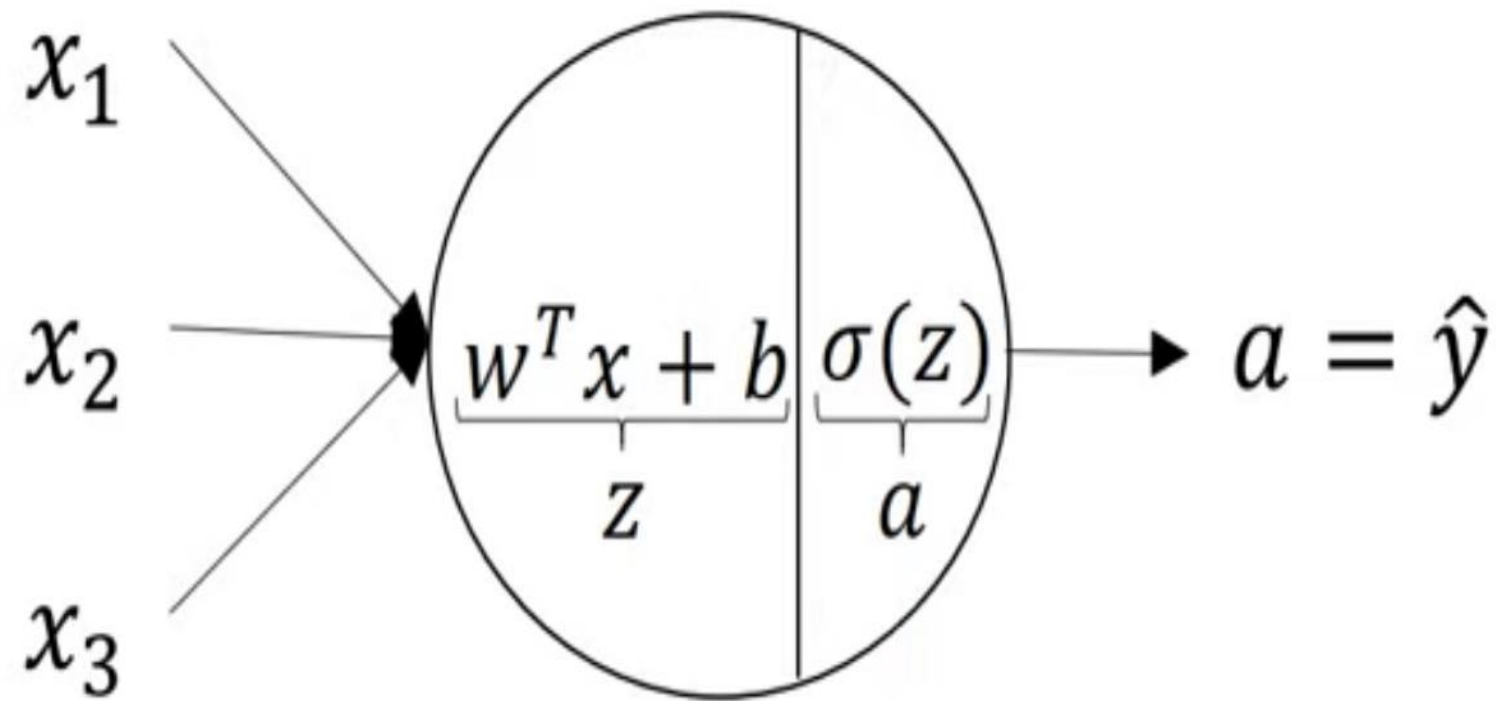
디코더



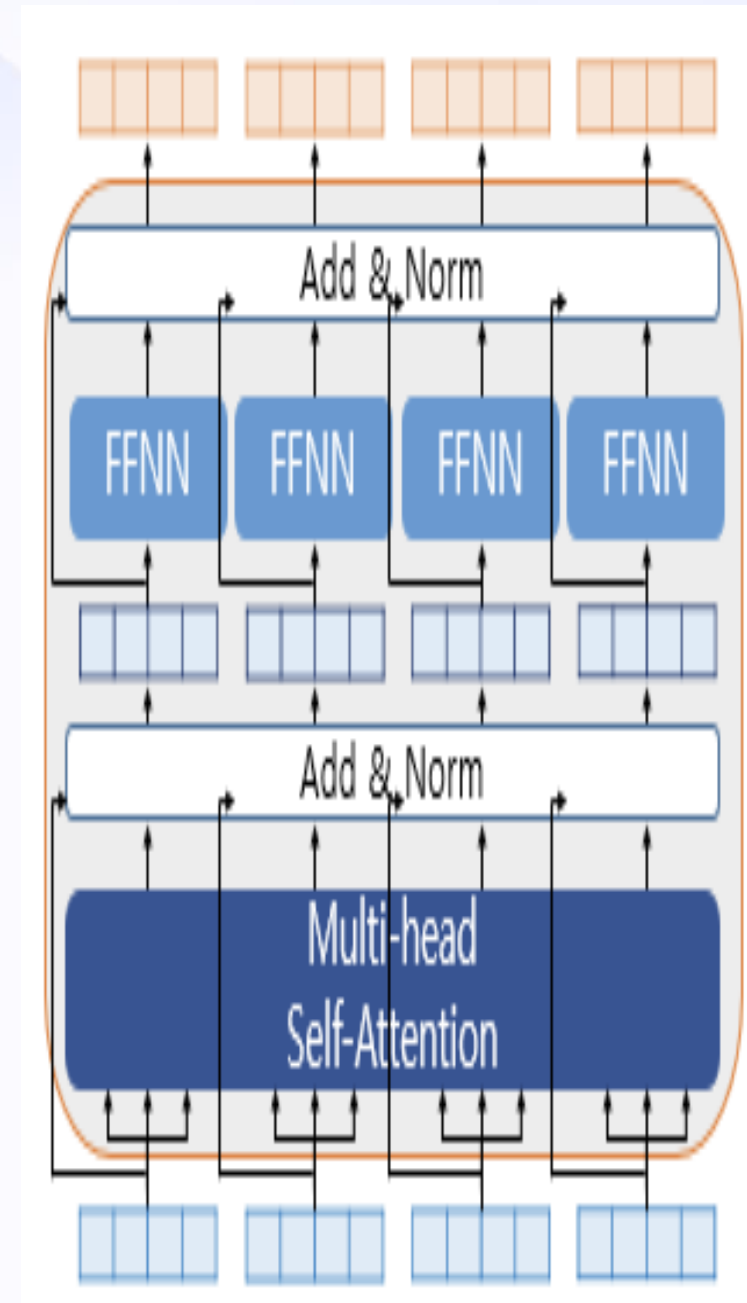
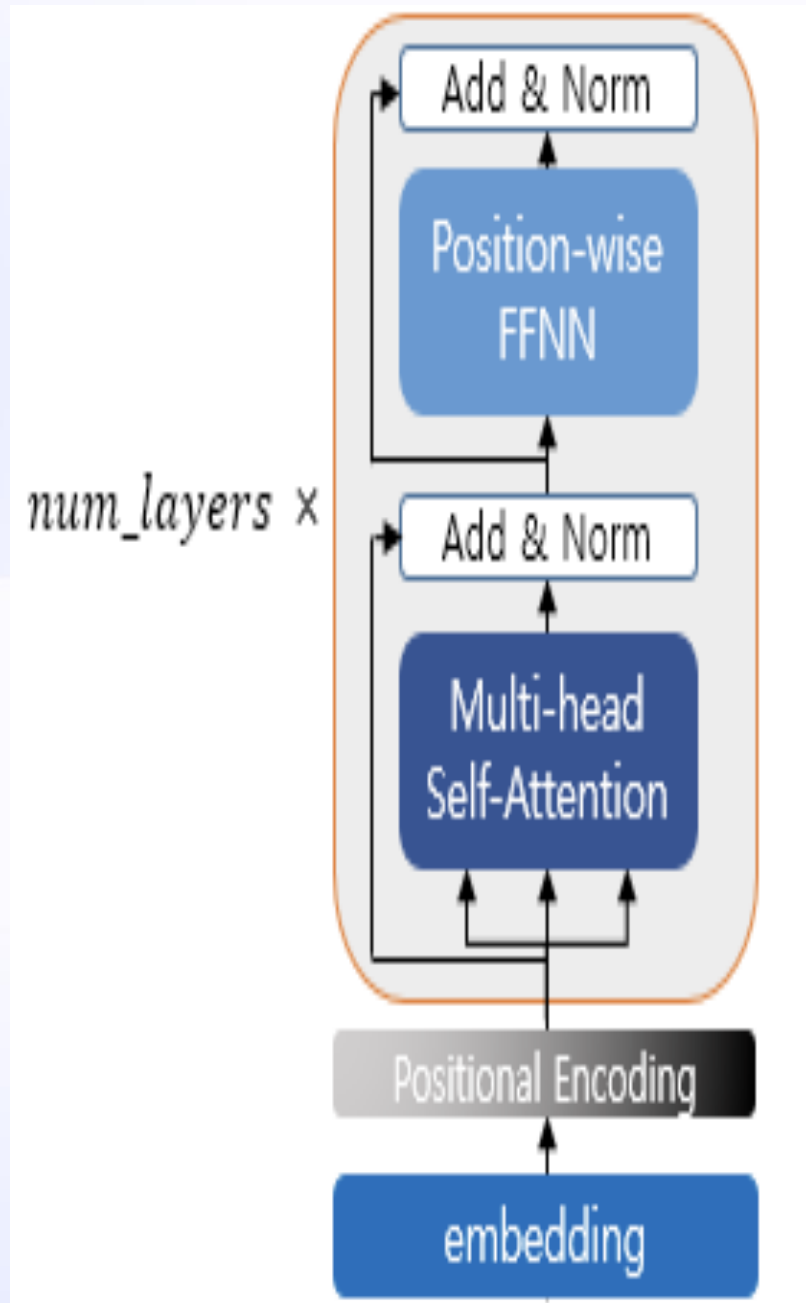
Multi-Head Attention



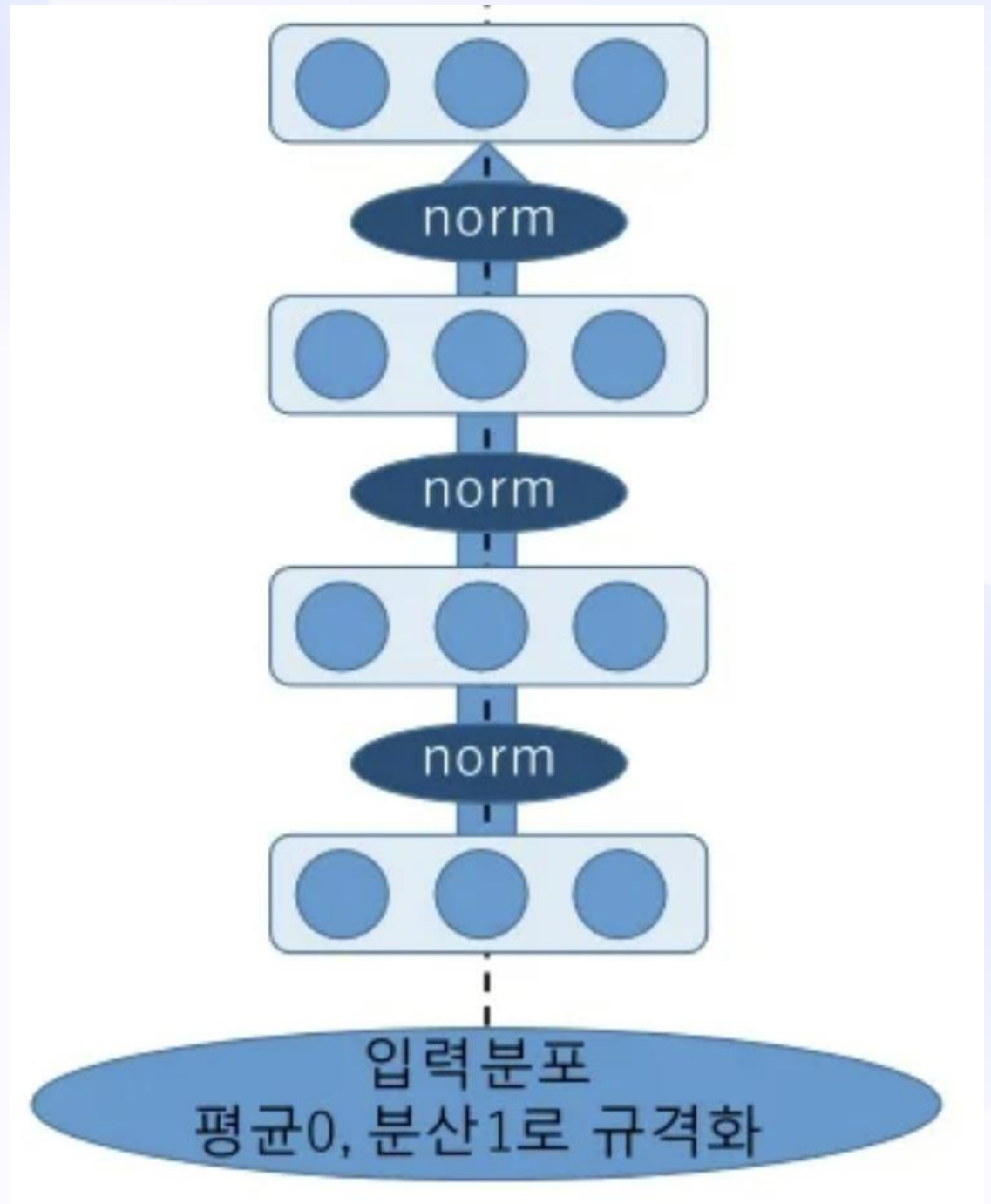
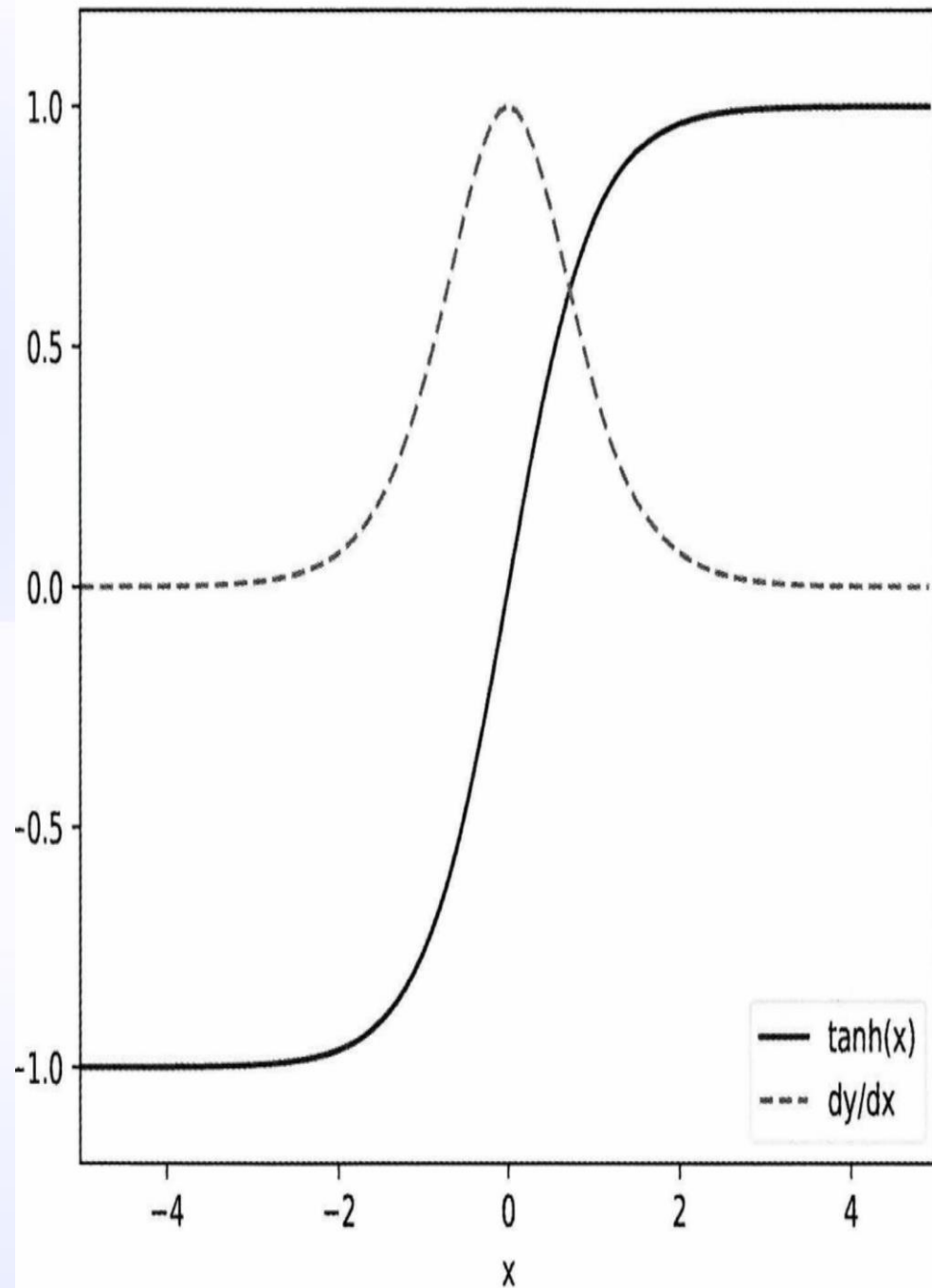
Feed-Forward Networks



Residual Connections (잔차 연결)



Layer Normalization (층 정규화)



결론

Transformer의 핵심 기술

1

성능 향상

트랜스포머는 기계 번역, 문장 생성, 텍스트 요약, 감정 분석 등 다양한 NLP 작업에서 기존 모델을 능가하는 성능을 보인다. 병렬 처리와 긴 의존성 학습의 가능성으로 인해 빠른 학습 속도와 높은 성능이 가능해졌다..

2

새로운 모델의 발전

트랜스포머는 새로운 언어 모델 개발의 기반이 되었으며 트랜스포머 기반을 한 BERT, GPT, T5등의 모델이 개발되었고, 이러한 모델들은 다시 다양한 NLP작업에서 뛰어난 성능을 보인다.

3

사전 훈련된 언어 모델의 사용 증가

트랜스포머를 기반으로 한 사전 훈련된 언어 모델의 사용이 증가했으며, 이 모델들은 다양한 언어 작업에 대해 일반적인 언어 이해능력을 가지고 있으므로, 사전 훈련된 언어 모델을 사용하는 것은 특정 작업에 대해 모델을 처음부터 훈련하는 것보다 효율적이다.

4

인공지능 연구의 방향 변화

병렬 처리와 어텐션 메커니즘의 사용은 다른 인공지능 분야, 이미지 처리나 음성 인식 등에도 적용되어 연구 방향을 변화시켰다

Transformer 모델의 미래

NLP 분야의 혁신

Transformer 모델은 자연어 처리 분야에서 획기적인 성과를 거두며, 기계 번역, 텍스트 요약, 질의 응답 등 다양한 문제에 적용되고 있습니다.

다양한 분야로의 확장

또한 이미지 처리, 음성 인식 등 다른 분야로도 그 적용 범위가 확장되고 있으며, BERT, GPT GPT 등의 사전 학습 모델을 통해 인공지능 전반에 걸쳐 핵심적인 기술로 자리잡고 있습니다.

지속적인 발전

Transformer 모델의 혁신적인 구조와 기술은 앞으로도 지속적으로 발전하여 인공지능 분야에 큰 영향을 미칠 것으로 기대됩니다.