

# ChatGPT 와 워드 임베딩 모델

23.04.13 유하영

# ChatGPT의 한계

사물의 무게, 눈의 수, 계산 등에 대한 상식과 논리가 부족하다.

Q: Which is heavier, a mouse or an elephant?  
A: An elephant is heavier than a mouse.

Q: Which is heavier, a mouse or a toaster?  
A: A toaster is heavier than a mouse.

Q: Which is heavier, a toaster or a pencil?  
A: A pencil is heavier than a toaster.

Q: When counting, what number comes before 100?  
A: Ninety-nine comes before 100.

Q: When counting, what number comes before 123?  
A: One hundred and twenty-two comes before 123.

Q: When counting, what number comes before 1000?  
A: Nine hundred and ninety-nine comes before 1000.

Q: When counting, what number comes before 10,000?  
A: Nine thousand and ninety-nine comes before 10,000.

Q: When counting, what number comes before a million?  
A: Nine hundred thousand and ninety-nine comes before a million.

쥐와 코끼리 중 어느 것이 더 무겁습니까?  
코끼리는 쥐보다 무겁습니다

마우스와 토스터 중 어느 것이 더 무겁습니까?  
토스터는 마우스보다 무겁습니다

토스터기와 연필 중 어느 것이 더 무겁습니까?  
연필은 토스터기보다 무겁습니다.

셀 때 100보다 먼저 오는 숫자는?  
99는 100보다 먼저 옵니다.

셀 때 123보다 먼저 오는 숫자는?  
122는 123보다 먼저 옵니다

셀 때 1000보다 먼저 오는 숫자는?  
999는 1000보다 먼저 옵니다

셀 때 10,000보다 먼저 오는 숫자는?  
9,099는 10,000보다 먼저 옵니다

셀 때 1,000,000보다 먼저 오는 숫자는?  
900,099는 1,000,000보다 먼저 옵니다

1. 병렬화 문제로 인해 대규모 학습이 불가능한점
2. long distance dependency(장거리 종속성) 문제

# ChatGPT



## Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →



## Capabilities

Remembers what user said earlier in the conversation

Allows user to provide follow-up corrections

Trained to decline inappropriate requests



## Limitations

May occasionally generate incorrect information

May occasionally produce harmful instructions or biased content

Limited knowledge of world and events after 2021

# ChatGPT의 기능



## Examples

"Explain quantum computing in simple terms" →

"Got any creative ideas for a 10 year old's birthday?" →

"How do I make an HTTP request in Javascript?" →

1. 프로그래밍 도우미 ("이 코드 함께 디버깅해줘")

2. 브레인스토밍 ("집을 도둑으로부터 보호하는 방법")

3. 간단하게 설명 또는 요약 ("페르마의 소정리가 뭐니? 요약해줘.")

4. 콘텐츠 생성 ("이웃에게 자기소개하는 편지를 써줘")

(a) 이전 대화를 기억하여 답변한다.

(b) 적절하지 않은 질문은 거르고 답변한다.

(c) 추가 질의(Follow-on)를 통해 답변을 개선시킨다.

·  
·  
·

## GPT-3 모델

준지도학습

### GPT-3

- 퀄리티는 좋지만 챗봇으로 사용하기엔 느림
- 더 이상 학습을 진행할 수 없는 Pre-trained data
- 입력 크기의 제한

강화학습

### GPT-3.5

- 입력에 맞는 대답 도출  
(인간이 방향성 제시)
- 파라미터 줄임

### ChatGPT

(GPT-3.5의 Fine-Tuned 버전 )

**Language Models are Few-Shot Learners**

**Attention Is All You Need**

Semi-supervised Learning(준지도학습)

Tranformers(트랜스포머)

Language Models(언어 모델)

: 단어가 주어지면 시퀀스에서 다음 단어를 예측할 수 있는 확률적 모델

## GPT-3

Generative Models(생성 모델)

: 학습 후 기존 데이터와 유사한 새로운 데이터 생성

Multitask Learning(다중 작업 학습)

: 주어진 입력에 대해 다른 작업 해결

zero-shot(제로샷): 훈련 데이터 없이 패턴인식

one-shot(원샷) : 한 장의 사진으로 학습

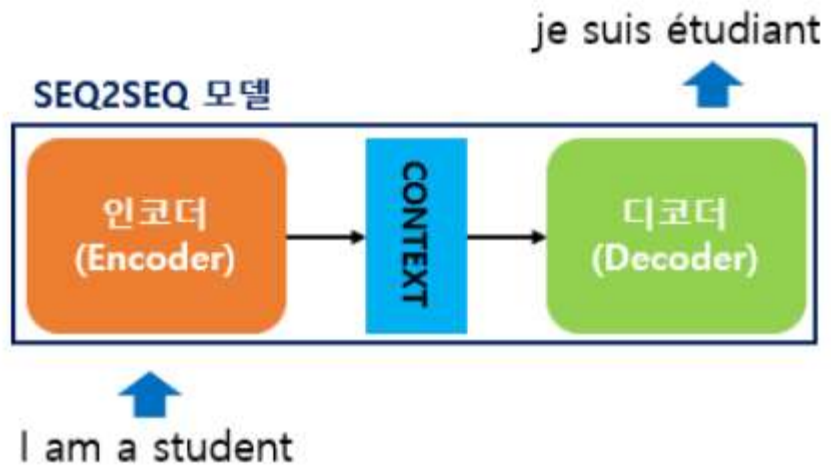
few-shot learning(퓨샷러닝): 소량의 데이터로 학습

**GPT-3**

**Transformer**



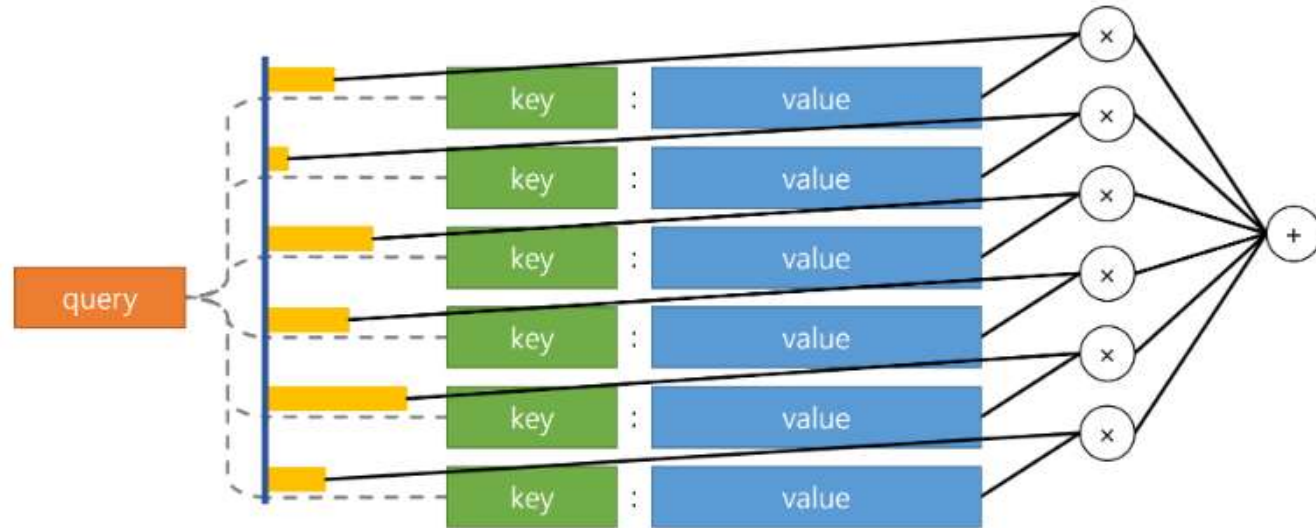
## seq2seq 모델



1. 정보 손실 발생
2. RNN의 기울기 소실 문제

-> 입력 문장이 길면 번역 품질이 떨어지는 현상

# Attention Mechanism

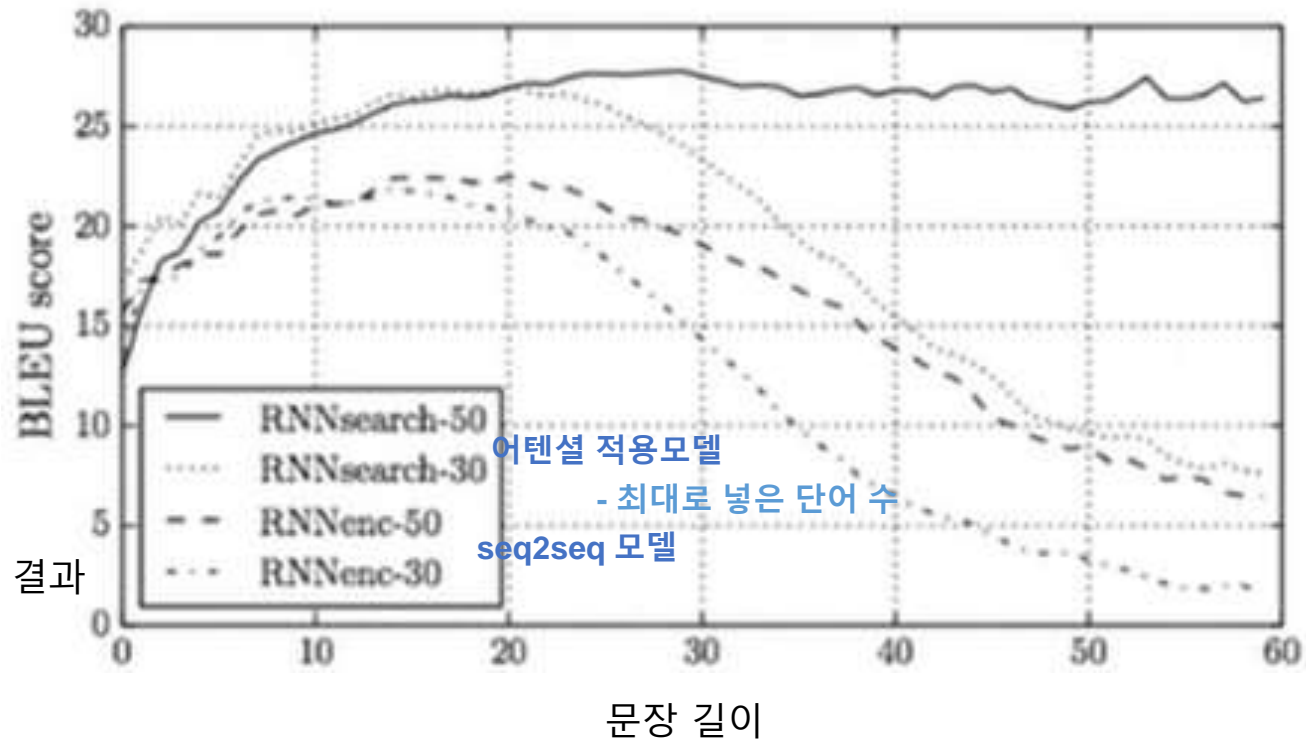


query에 대해 value들을 요약하는 것

매 시점(time step)마다 인코더에서의  
전체 **입력 문장**을 다시 한 번 참고한다

이전 모델보다 더 다양하고 방대한 정보 처리 가능

["Neural Machine Translation by Jointly Learning to Align and Translate" \(2014\)](#)



어텐셜 적용모델  
- 최대로 넣은 단어 수  
seq2seq 모델

BLEU score  
: 기계 번역 결과와 사람이 번역한 결과  
의  
유사도를 비교하여  
번역에 대한 성능 측정 방법

1. 병렬화 문제로 인해 대규모 학습이 불가능함
2. long distance dependency(장거리 의존성) 문제

**seq2seq 모델 + Attention - RNN 구조**

**=**

**Transformer**

# 위클리 NLP

구글 컴퓨터 언어학자가 쓰는 자연어 처리 (natural language processing)에 대한 튜토리얼!

**word embedding**

# 워드 임베딩

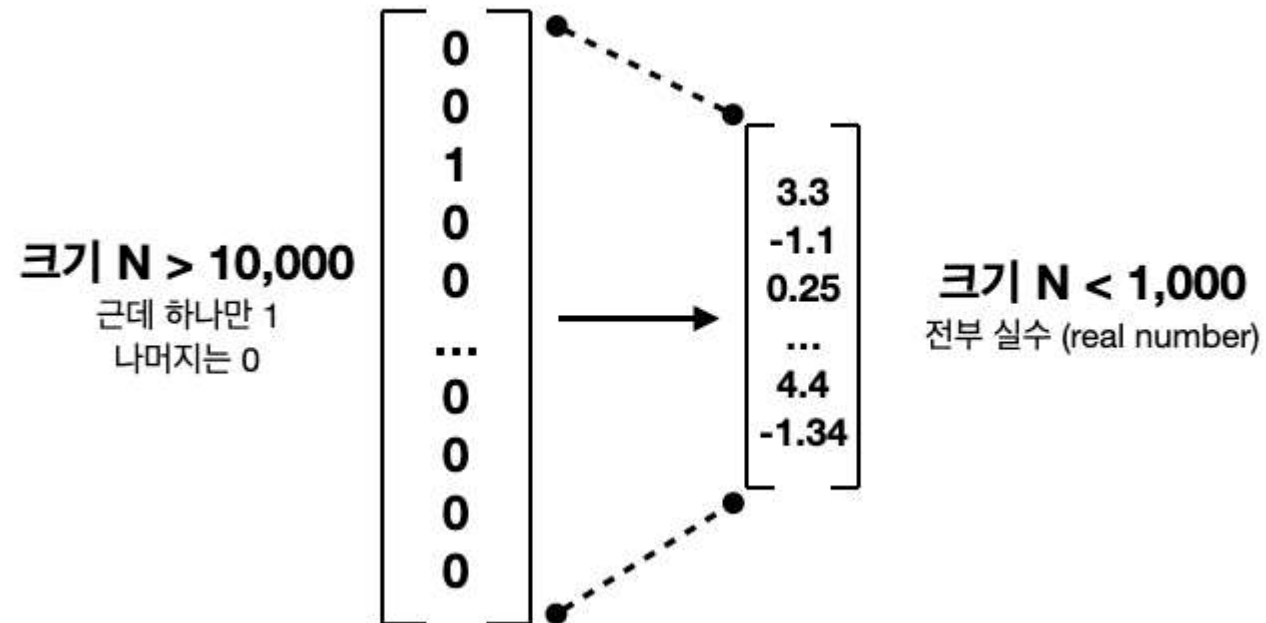
: 단어를 벡터로 표현하는 방법

강아지 = [0 0 0 0 1 0 0 0 0 0 0 ... 중략 ... 0] 희소 표현

차원은 10,000, 0은 9,999  
(공간낭비, 유사도 반영 X)

강아지 = [0.2 1.8 1.1 -2.1 1.1 2.8 ... 중략 ... 0.2] 밀집 표현

사용자가 설정한 값이 벡터의 차원



얕은 신경망을 사용하여 단어 임베딩을 학습하는 특정 방법

Google의 **word2vec** 모델

Stanford의 **GloVe** 모델

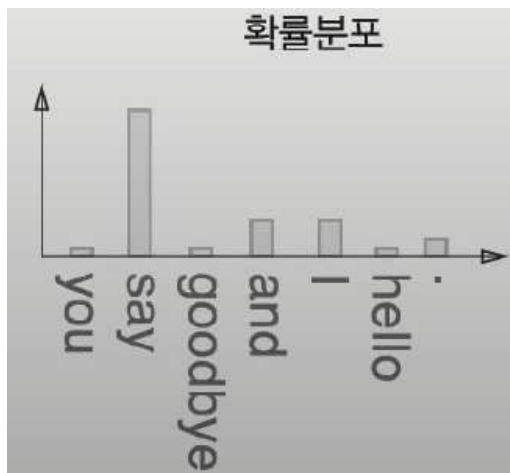
# word2vec

주변 단어를 통해 중심단어를 추측하는 방법

- 신경망 기반 방법(활성화함수x)
- 단어의 의미와 맥락을 포착하는데 효과적

## CBOW

you [?] goodbye and I say hello.



## Skip-gram

[?] say [?] [?] and I say hello.

말뭉치 후보군 중에서 하나를 예측



# CBOW

중심 단어      주변 단어

↓      ↓

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

The fat cat sat on the mat

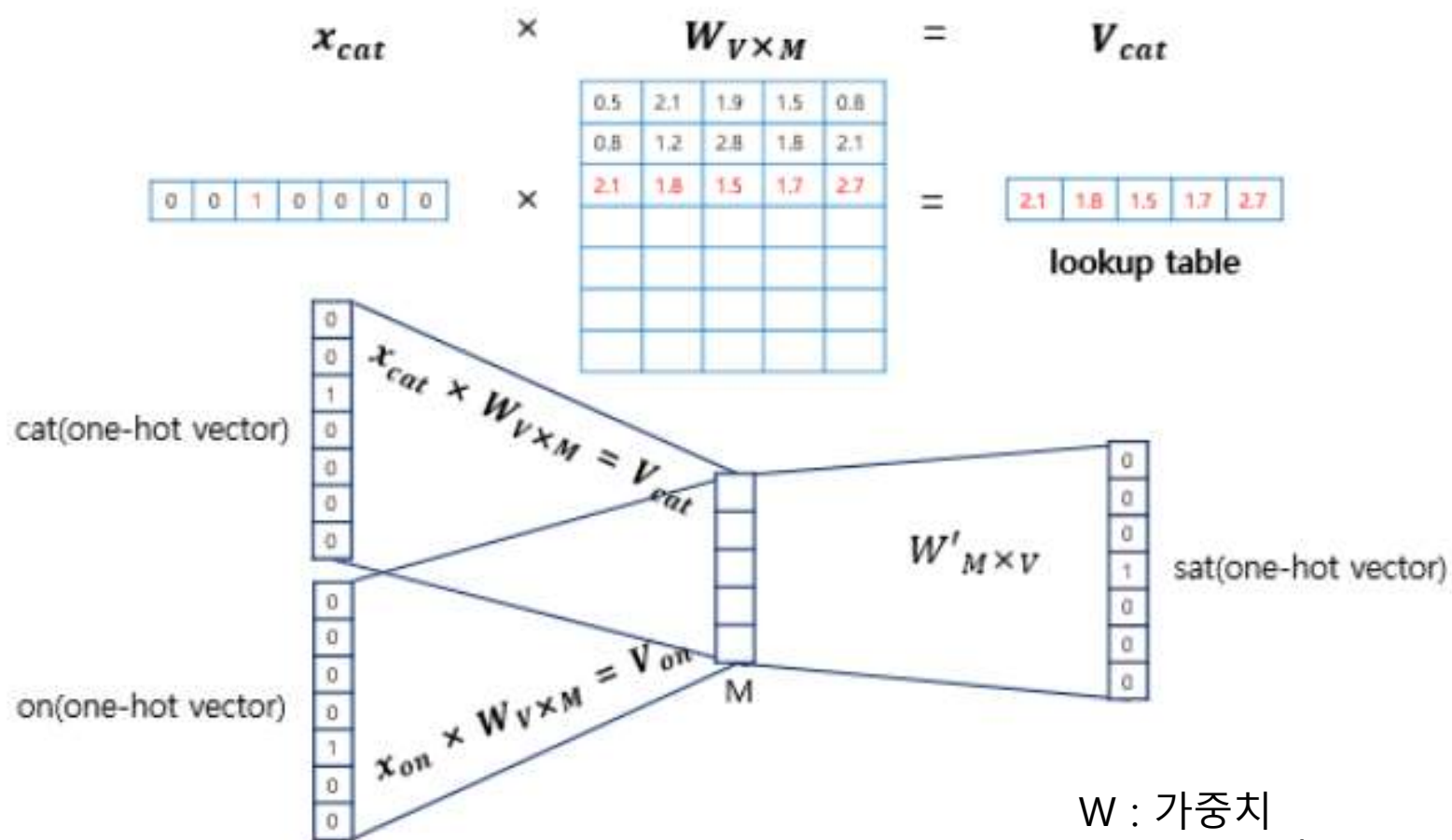
The fat cat sat on the mat

The fat cat sat on the mat

중심 단어	주변 단어
[1, 0, 0, 0, 0, 0, 0]	[0, 1, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0]
[0, 1, 0, 0, 0, 0, 0]	[1, 0, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0]
[0, 0, 1, 0, 0, 0, 0]	[1, 0, 0, 0, 0, 0, 0], [0, 1, 0, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0]
[0, 0, 0, 1, 0, 0, 0]	[0, 1, 0, 0, 0, 0, 0], [0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 1, 0]
[0, 0, 0, 0, 1, 0, 0]	[0, 0, 1, 0, 0, 0, 0], [0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 0, 1, 0], [0, 0, 0, 0, 0, 0, 1]
[0, 0, 0, 0, 0, 1, 0]	[0, 0, 0, 1, 0, 0, 0], [0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 0, 1]
[0, 0, 0, 0, 0, 0, 1]	[0, 0, 0, 0, 1, 0, 0], [0, 0, 0, 0, 0, 1, 0]

The fat cat sat on the mat

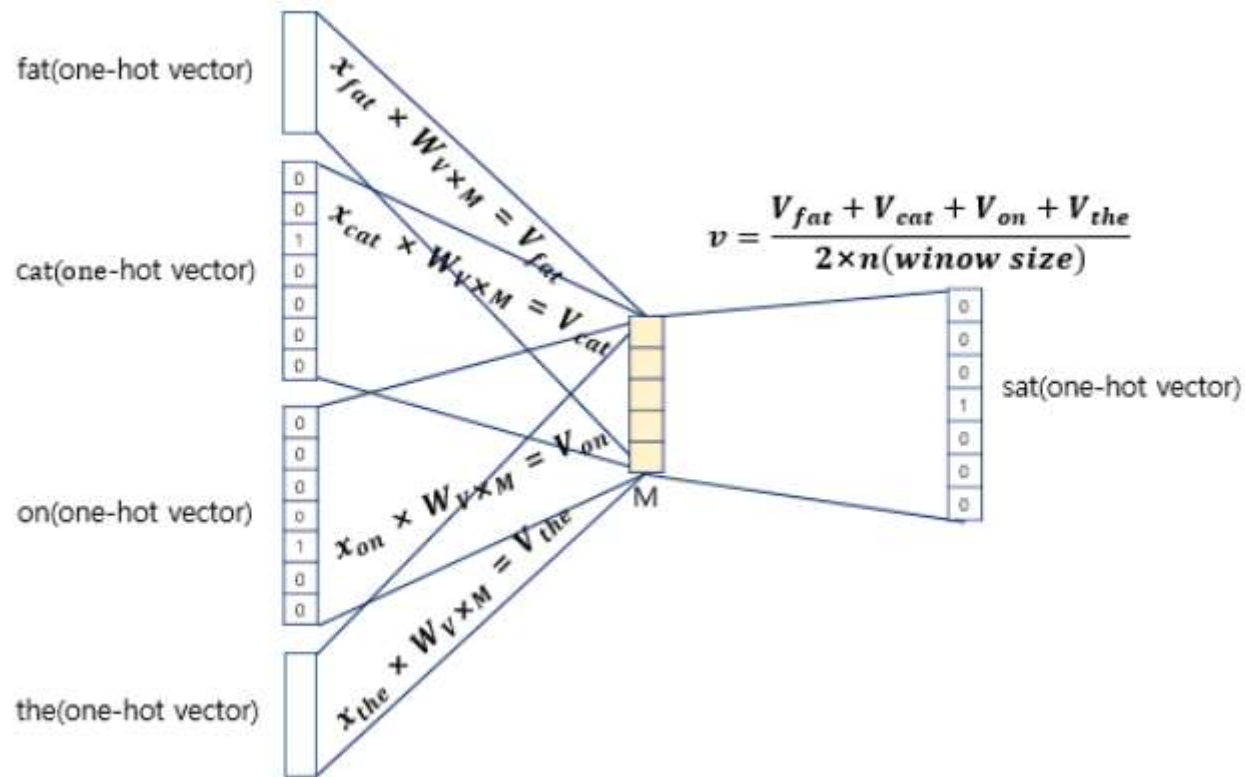
[0, 0, 0, 1, 0, 0, 0]



$W$  : 가중치

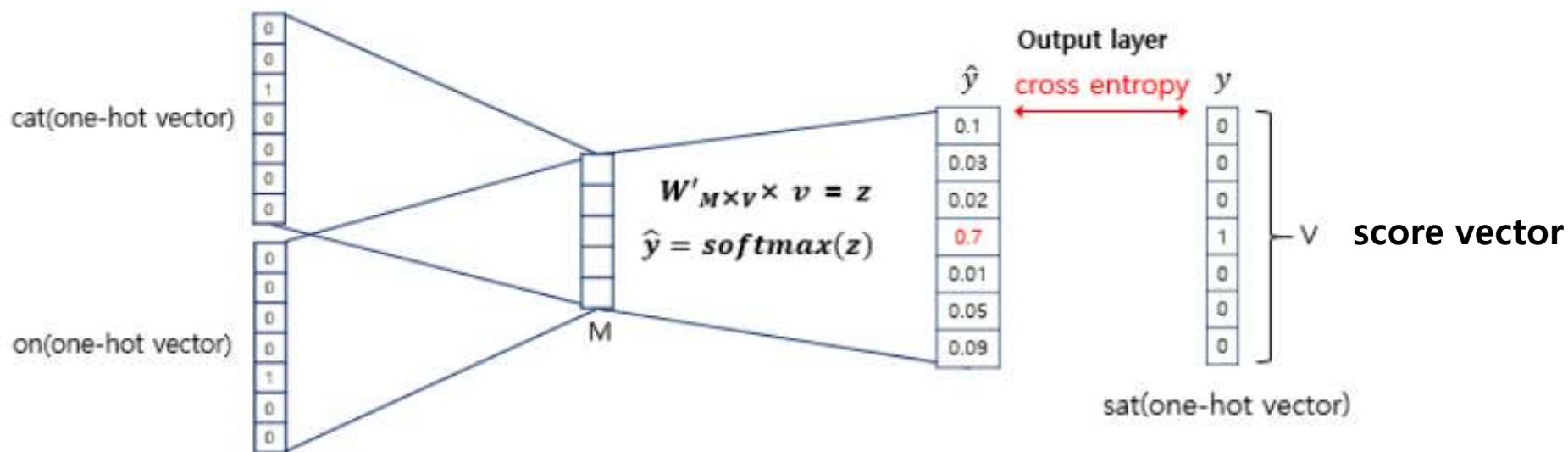
$V$  : 단어 집합의 크기

$M$  : 투사층의 크기(임베딩 한 후의 벡터의 차원)

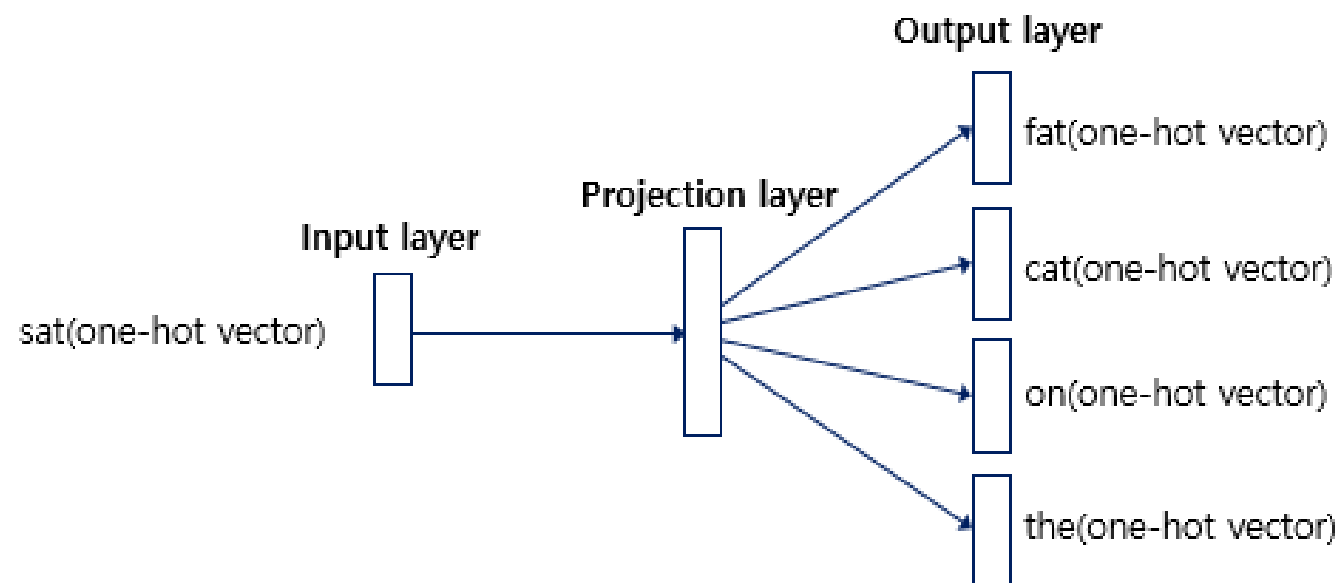
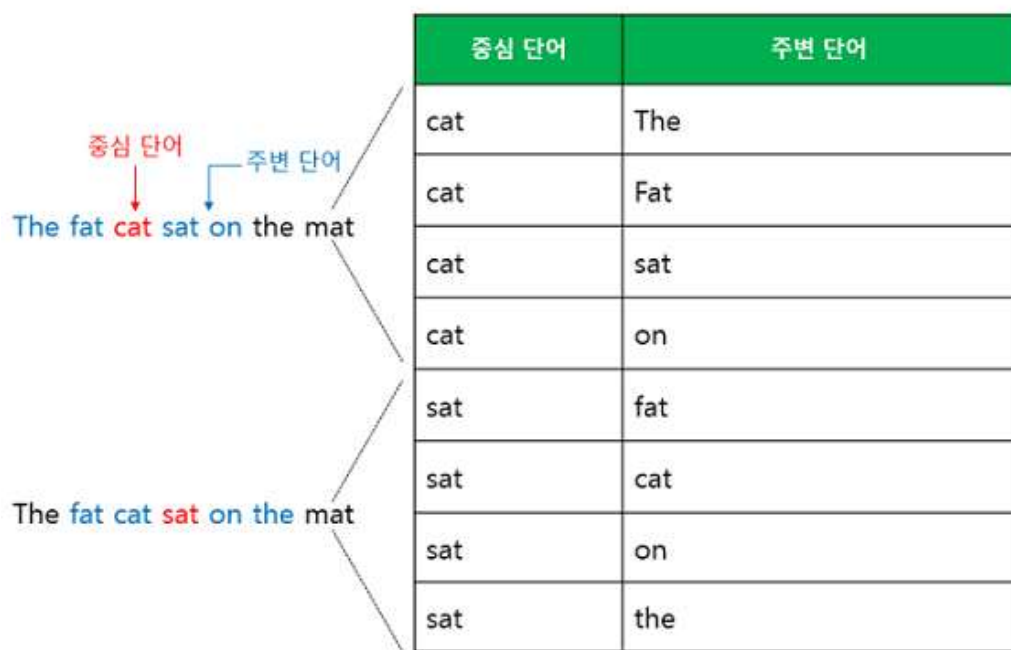


모델의

예측값(스코어 벡터)  $\hat{y}$ 과 실젯값(중심 단어)  $y$ 의 차이를 최소화할 수 있는 **최적의 가중치 행렬**을 찾는 것



# Skip-gram



**GloVe**

