Word2Vec

23.06.01 유하영

Word2Vec

‘ 비슷한 문맥에서 등장하는 단어들은 비슷한 의미를 가진다 ’

유사성 있는 단어를 예측

문장 생성은 X -> 다른 모델 필요

원수는 외나무다리에서 만난다

(연어)

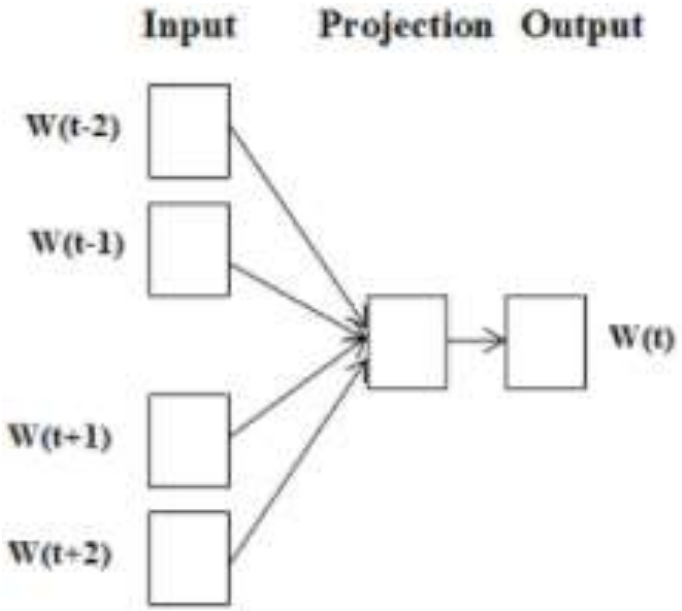
man , woman

word2vec

CBOW

주변단어 -> 중심단어 예측

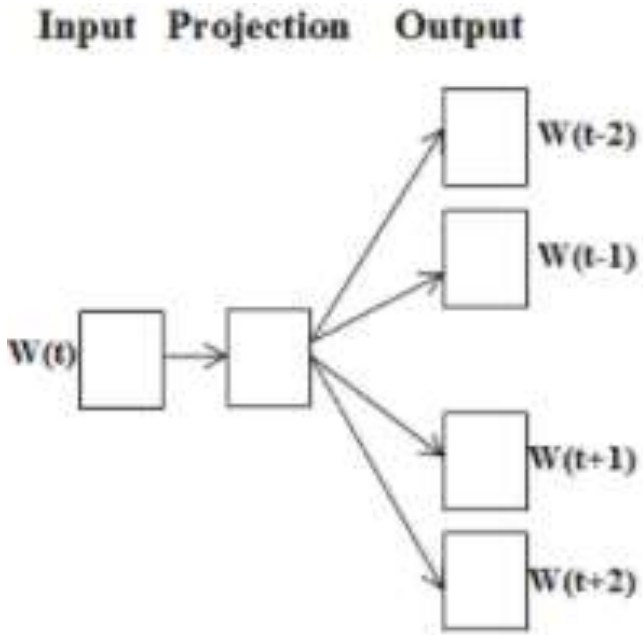
you [?] goodbye and I say hello.



Skip-gram

중심단어 -> 주변단어 예측

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

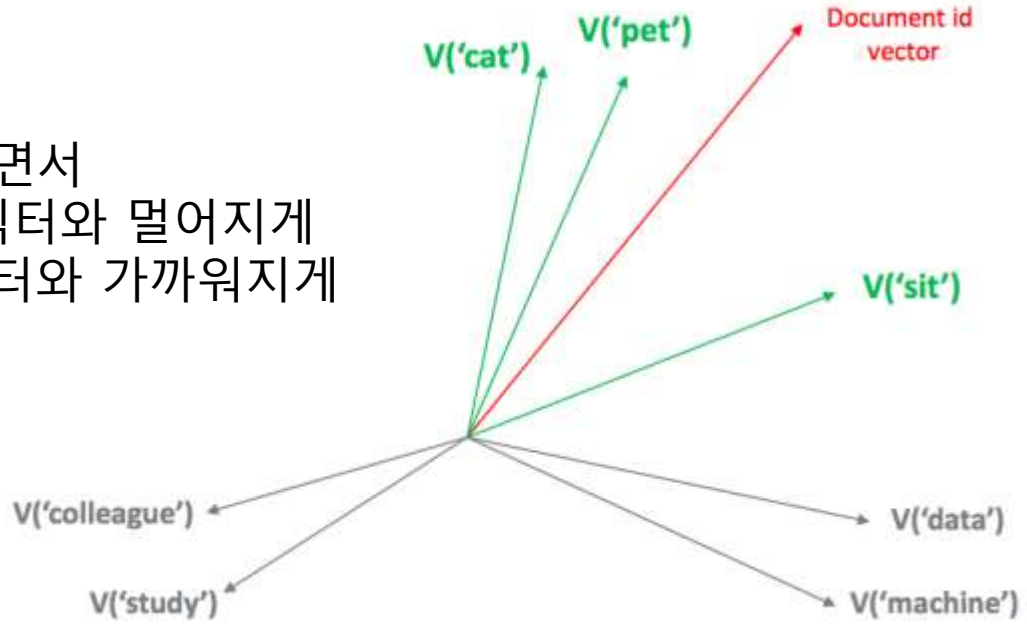


word2vec

CBOW

Skip-gram

코퍼스를 슬라이딩 하면서
등장↓단어의 벡터가 중심단어 벡터와 멀어지게
등장↑단어의 벡터가 중심단어 벡터와 가까워지게



학습데이터 The fat cat sat on the mat

(window size = 2)

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]

학습데이터 The fat cat sat on the mat

```
[('fat', 0.7668752074241638), ('mat', 0.6178626418113708), ('the', 0.10980246961116791), ('on', -0.1959061473608017), ('cat', -0.3469330966472626)]
```

대량의 문장으로 이루어진 문서를
전처리하여 토큰화 한 후
모든 단어에 대한 학습

word2vec

```
model.similarity('this', 'is')
```

```
0.407970363878
```

```
model.similarity('post',
```

```
0.057204389197
7
```

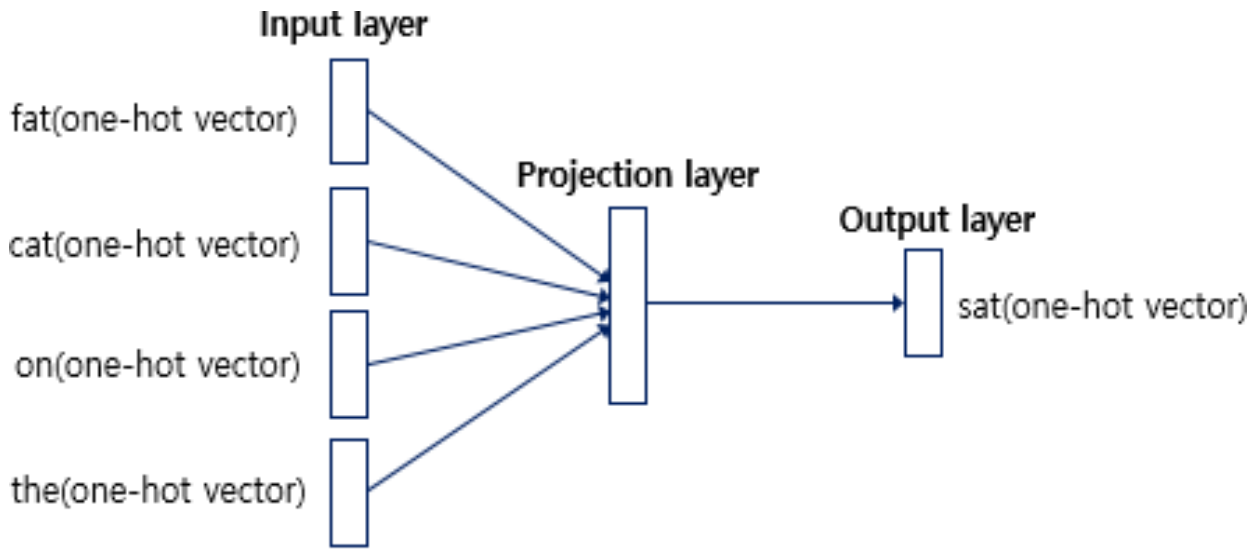
```
'book'
```

```
[ 0.11279297 -0.02612305 -0.04492188
0.06982422 0.140625 0.03039551 -0.04370117
0.24511719 0.08740234 -0.05053711 0.23144531 -
0.07470703 ...
중략 ...
0.03637695 -0.16796875 -0.01483154 0.09667969
-0.05761719 -0.00515747]
```

CBOW

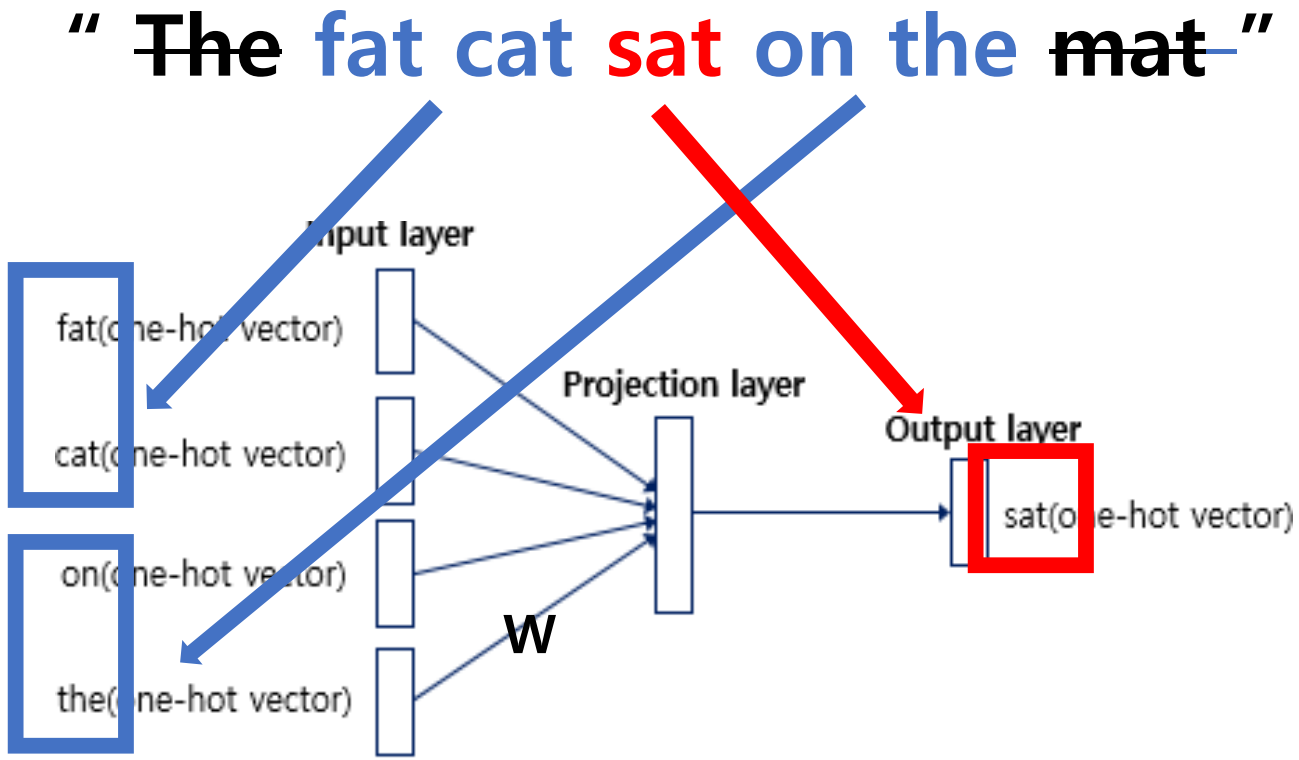
주변단어 -> 중심단어 예측

" ~~The~~ fat cat sat on the ~~mat~~ "

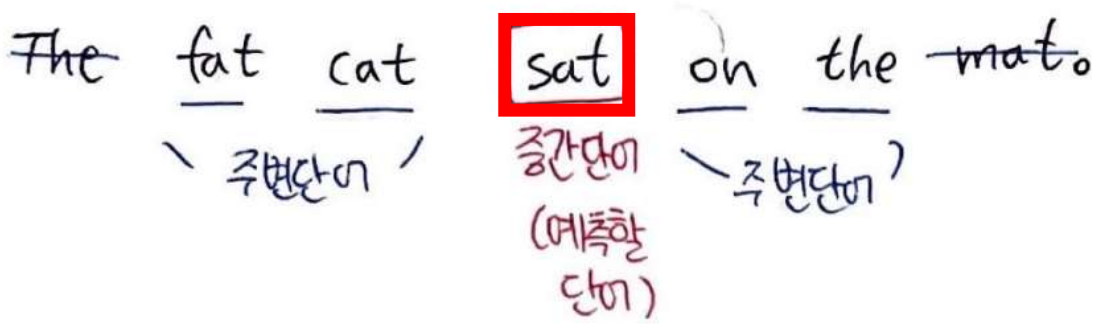


CBOW

주변단어 -> 중심단어 예측



예문)



중심단어	주변단어
Sat	fat
	cat
	on
	the

→ 원-핫 벡터로 변환

중심단어
(예측할 단어)

sat

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

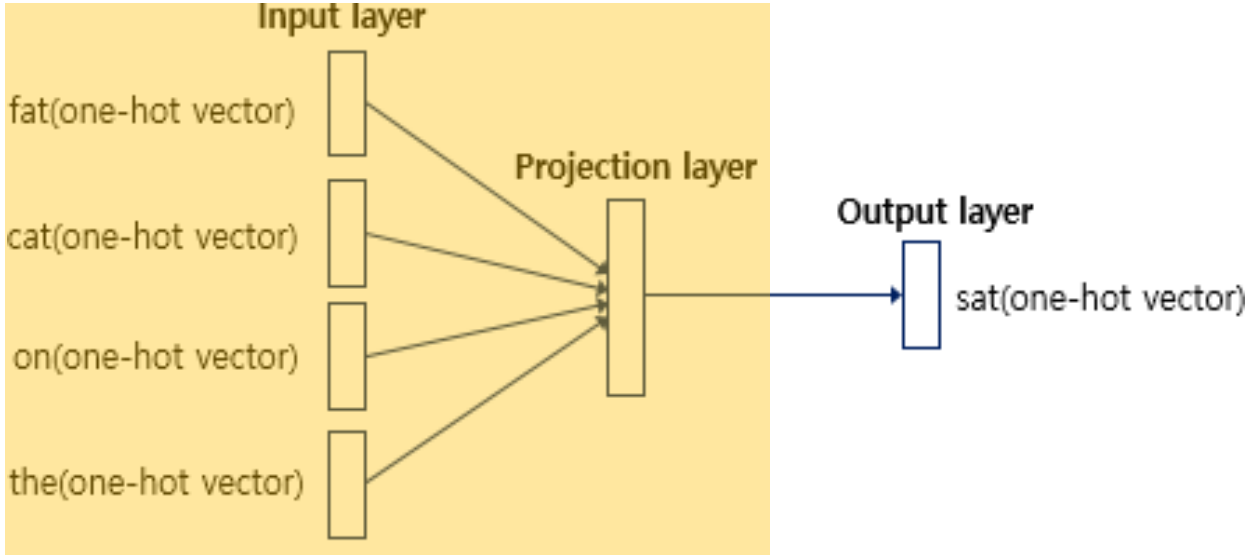
주변단어

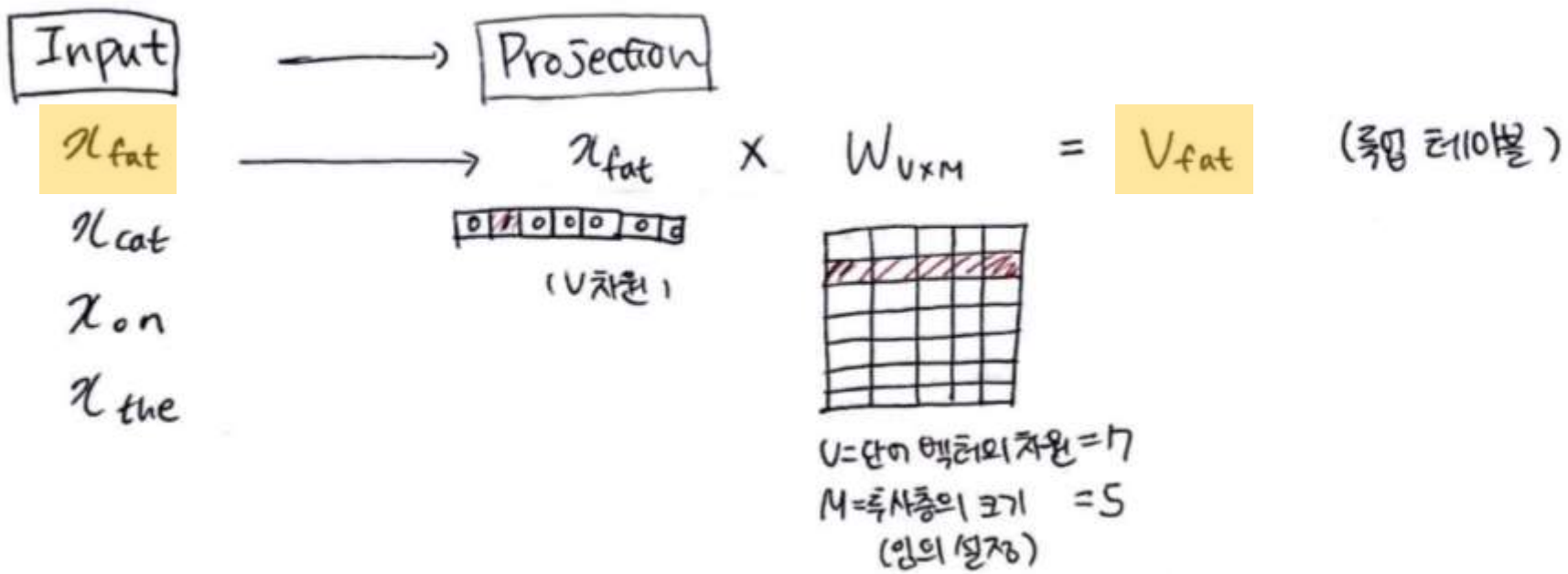
fat [0,1,0,0,0,0,0]

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

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

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


Input layer
fat(one-hot vector)
cat(one-hot vector)
on(one-hot vector)
the(one-hot vector)
Projection layer
Output layer
sat(one-hot vector)


Input
Projection
x_{fat}
x_{cat}
x_{on}
x_{the}
x_{fat} × $W_{V\times M}$ = V_{fat}
(룩업 테이블)
(V차원)
V=단어 벡터의 차원=7
M=투사층의 크기 =5
(임의 설정)

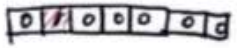
Input → Projection

x_{fat} → x_{fat} × $W_{V\times M}$ = V_{fat} (룩업 테이블)

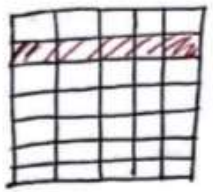
x_{cat}

x_{on}

x_{the}



(V차원)



V=단어 벡터의 차원 =7

M=투사층의 크기 =5

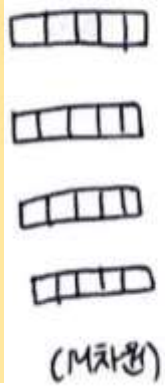
(임의 설정)

$x_{fat} \rightarrow x_{fat} \times W_{V\times M} = V_{fat}$

$x_{cat} \rightarrow x_{cat} \times W_{V\times M} = V_{cat}$

$x_{on} \rightarrow x_{on} \times W_{V\times M} = V_{on}$

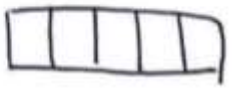
$x_{the} \rightarrow x_{the} \times W_{V\times M} = V_{the}$



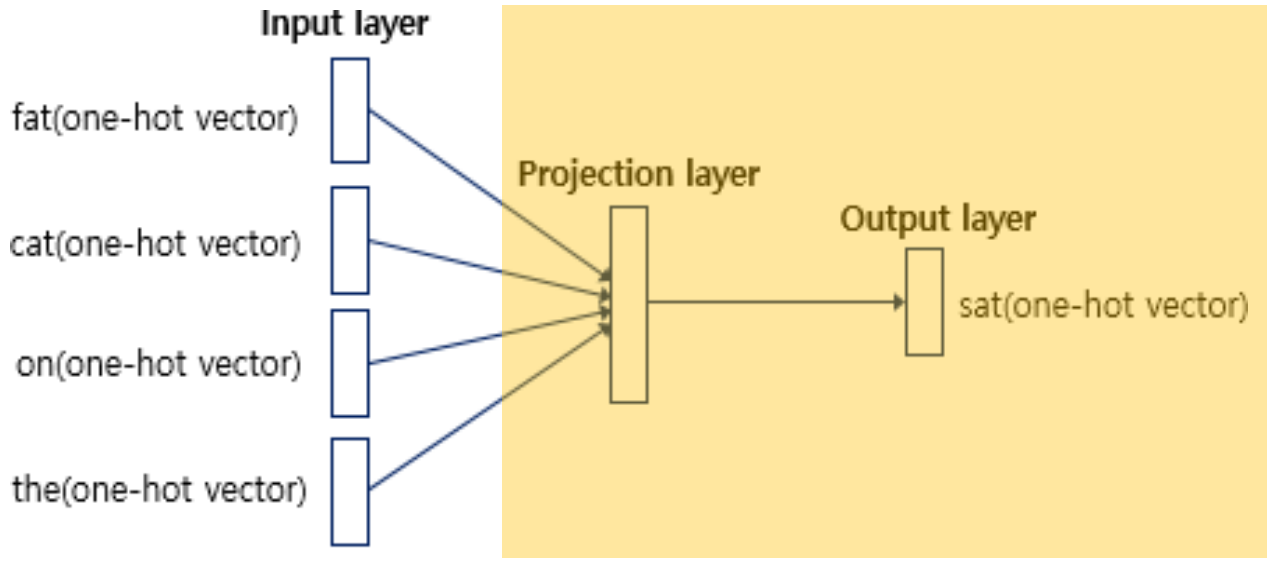
(M차원)

단어벡터들의 평균

$$\hat{V} = \frac{V_{fat} + V_{cat} + V_{on} + V_{the}}{2n}$$



(M차원의 임베딩 벡터)



Projection

단어벡터들의 평균

$\hat{v}$ = -

(M 차원)

x W'_{MxV} = Z

(V차원)
: 원핫벡터와 차원동일

output

$\hat{y}$ = Softmax(z)

V차원

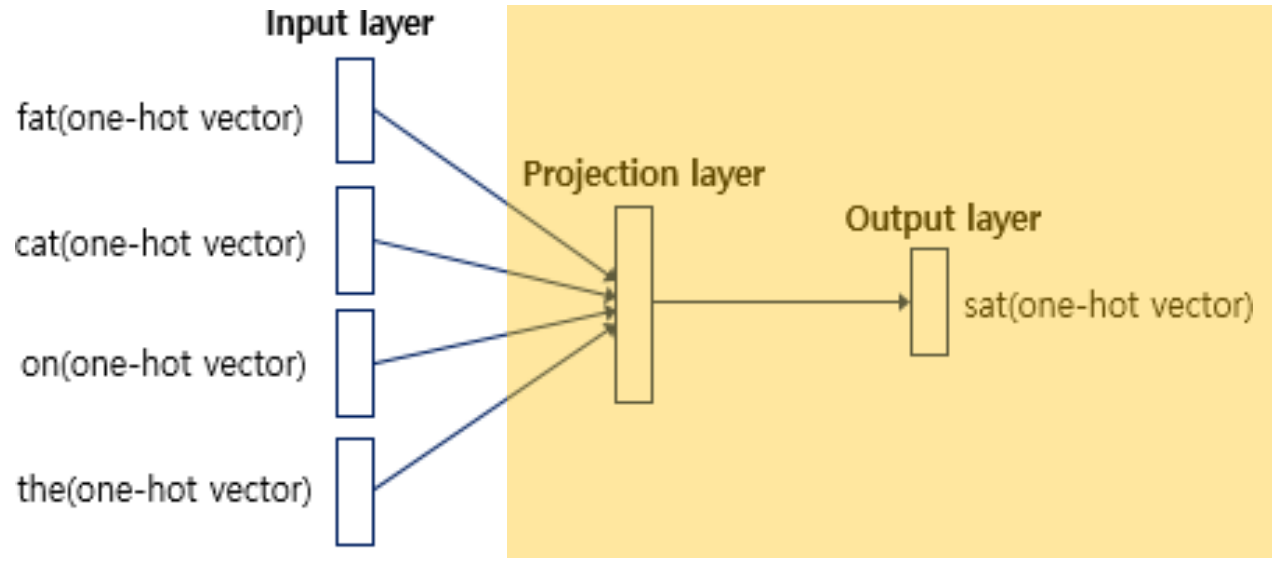
y

(라벨값)
: Sat의 원핫벡터

중간단어

크로스-엔트로피

손실



Projection

$\hat{U}$ (M 차원) $\times$ W'_{MxV} $=$ Z (V차원) : 원핫벡터와 차원동일

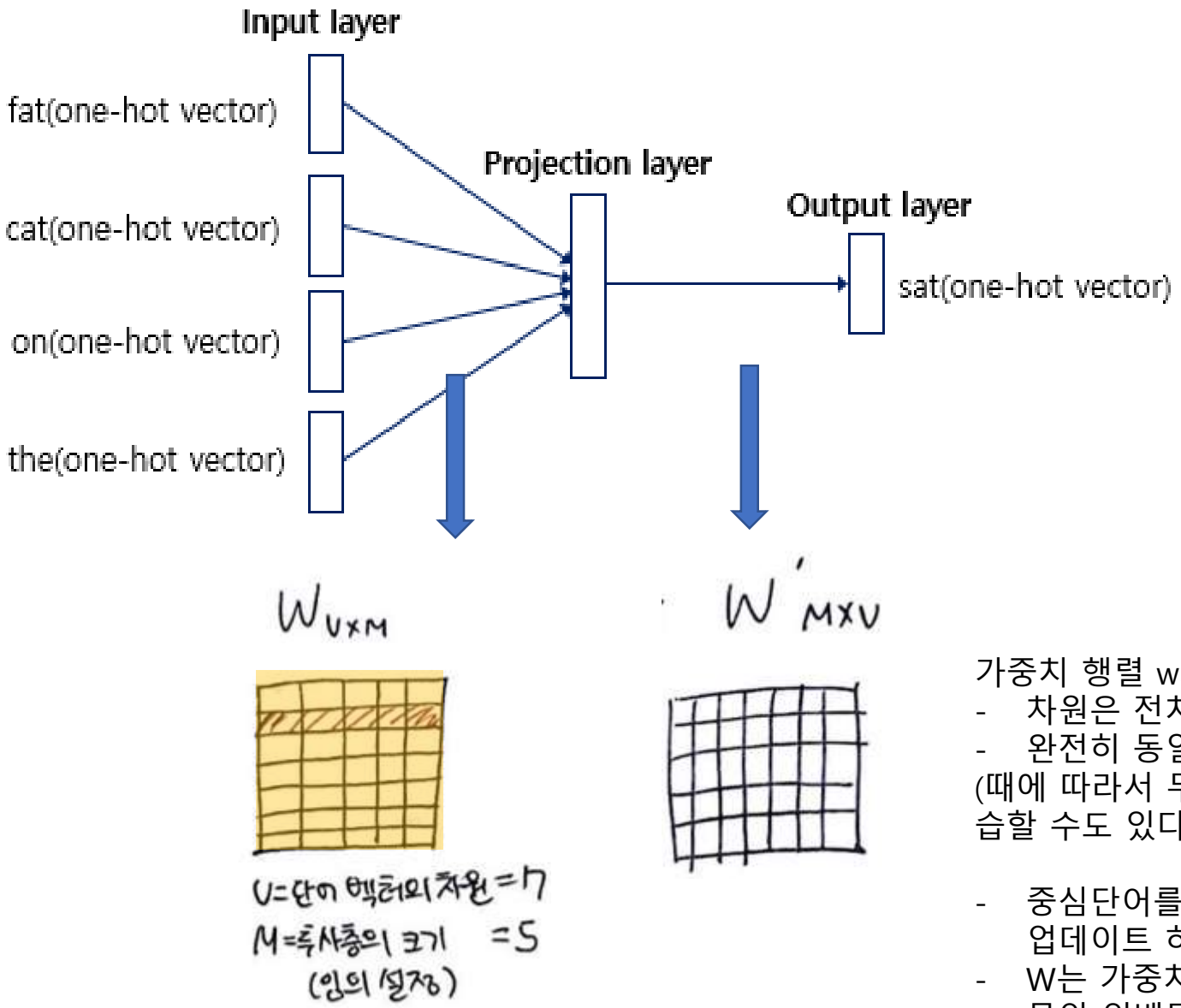
output

$\hat{y} = \text{Softmax}(z)$ V차원

y (라벨값) : Sat의 원핫벡터

중간단어

크로스-엔트로피 손실



가중치 행렬 w와 w′
- 차원은 전치한 것과 같다.
- 완전히 동일한 행렬은 아니다.

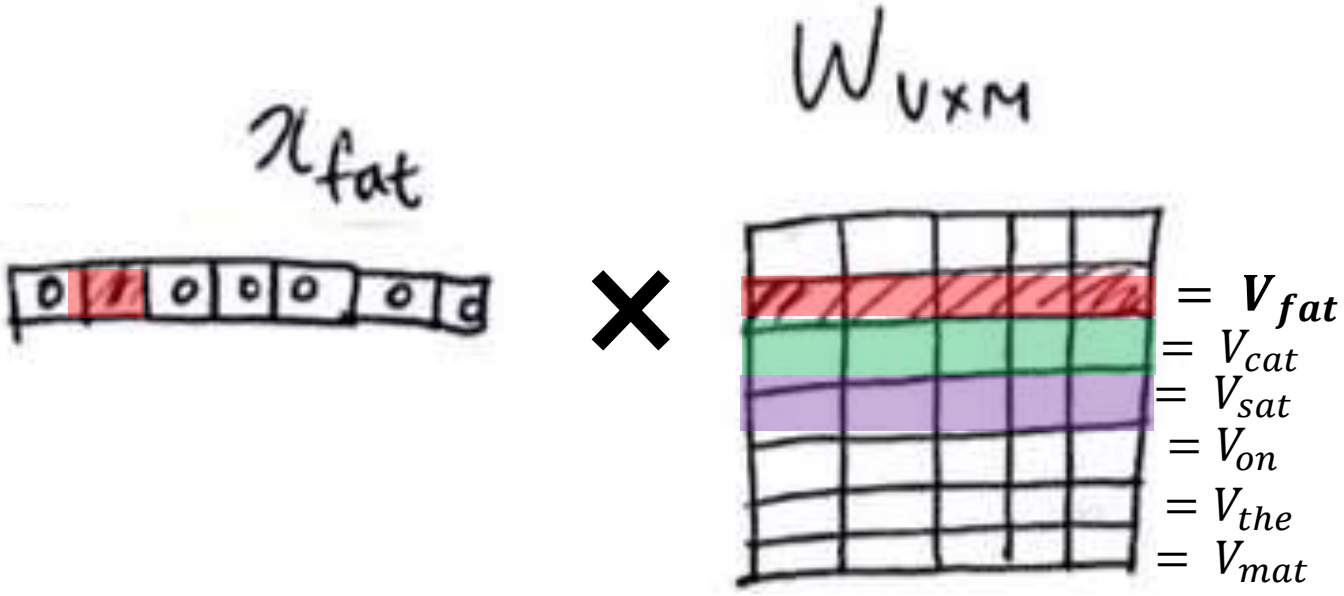
(때에 따라서 두개를 하나의 행렬로 취급해서 학습할 수도 있다.)

- 중심단어를 맞추기 위해 W와 W′를 조금씩 업데이트 하며 학습하는 것
- W는 가중치 행렬이자 word2vec의 최종 결과물인 임베딩 단어벡터의 모음이다.

가중치 선택

lookup table

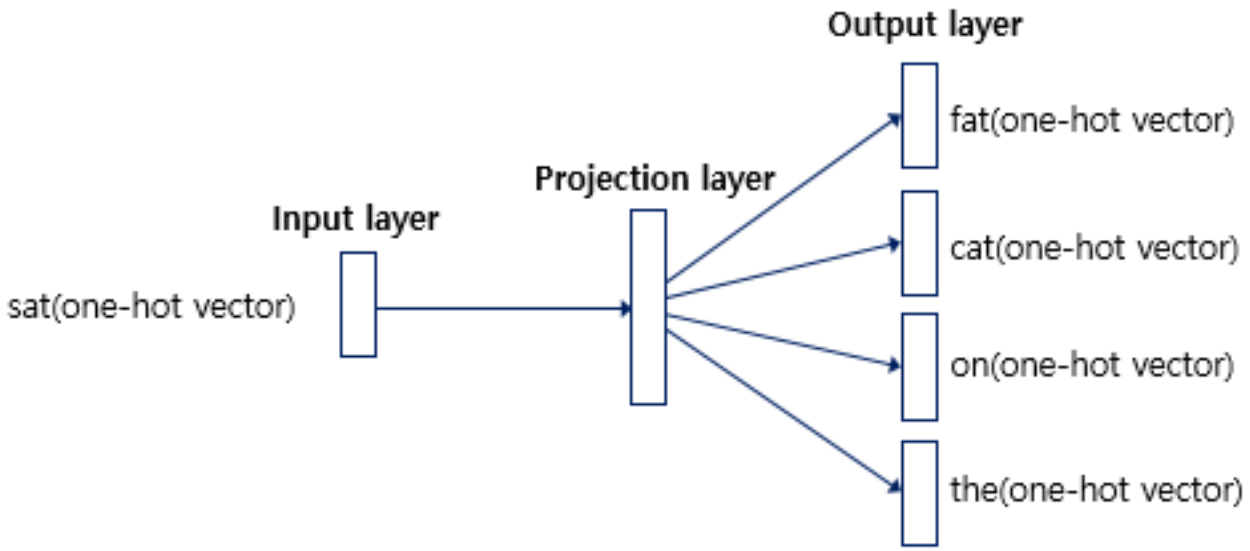
A. 가중치행렬 **W**를 최종단어의 분산 표현으로 이용한다.
B. 가중치행렬 W'를 최종단어의 분산 표현으로 이용한다.
C. 양쪽 가중치 모두 최종단어의 분산표현으로 이용한다.



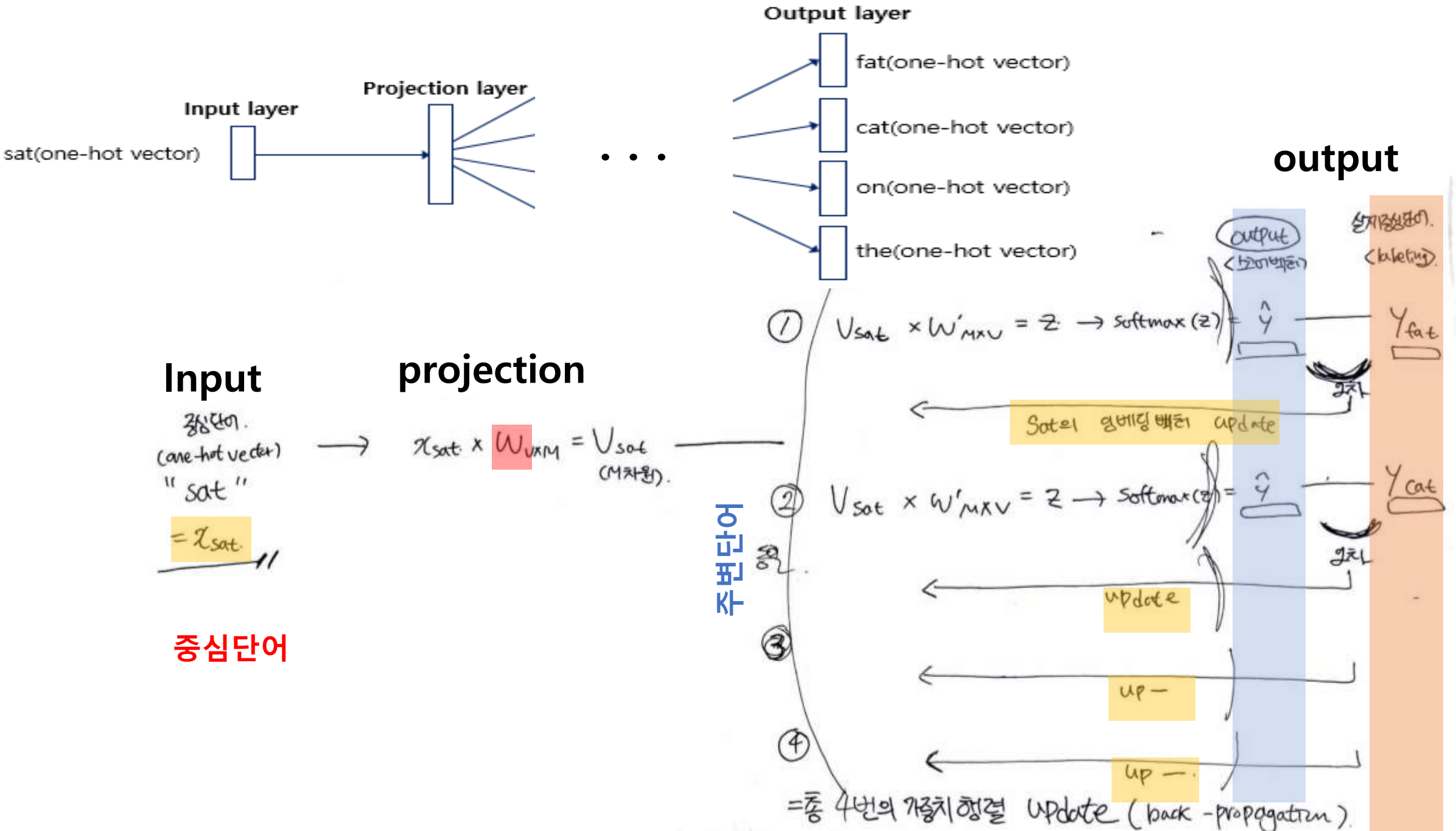
Skip-gram

중심단어 -> 주변단어 예측

~~The~~ fat cat sat on the ~~mat~~



중심단어	주변단어
sat	fat
sat	cat
sat	on
sat	the


Output layer
fat(one-hot vector)
cat(one-hot vector)
on(one-hot vector)
the(one-hot vector)
Projection layer
Input layer
sat(one-hot vector)
output
Input
projection
중심단어
주변단어
update
=총 4번의 가중치 행렬 update (back-propagation)

you got me looking for attention.

SGNS

skip-gram negative sampling

skip-gram	SGNS
전체 단어 집합	일부 단어 집합에 집중
다중 분류류	이중분

negative sampling

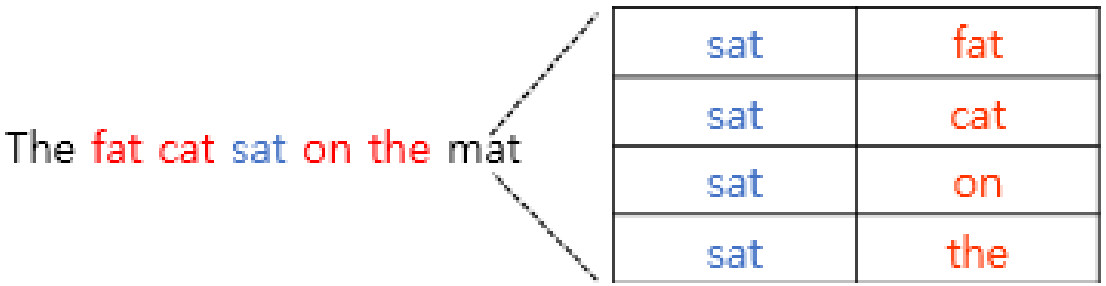
I love you

I, you 주어졌을때 가운데 나올단어가 무엇인가?

I, you란 단어가 주어졌을때 가운데 나올단어가 love인가?

positive 실제로 관측된 단어쌍의 점수 최대화

negative 무작위로 선택된 부정샘플의 집합점수 최소화



skip-gram

입력과 레이블의 변화

입력1	입력2	레이블
cat	The	1
cat	fat	1
cat	sat	1
cat	on	1

입력1	입력2	레이블
cat	The	1
cat	fat	1
cat	**pizza**	**0**
cat	**computer**	**0**
cat	sat	1
cat	on	1

단어 집합에서 랜덤으로
선택된 단어들을
레이블 0의 샘플로 추가.

negative sampling