

$\sqrt{1} = 1$  (used the product)

↳ cleaned, now

$\rightarrow \text{TFIDFvectorizer} \rightarrow \boxed{w} \rightarrow 1$



model.predict( $\sqrt{1}$ )

class cleaner:

—

—

—

—

fit()

transform()

Apple → fruit  
Company

3 million words

↓  
Spacy 3 → 10k .  
↓  
90 contexts

SM .  
lg .  
look .  
↓  
130 contexts

Words → Numbers → Any models

13 81.3 40 73.5

The child is cute

The kid is cute,

- 1.35

BOW  
TFIDF

lion → Leadership → ✓  
→ Warriors → ✓  
→ influential → ✓  
→ animal → ✓  
→ Carnivore → ✓  
→ Wild → ✓  
→ strength → ✓  
→ pit → ✓

30 vectors

child → child is very cute  
→ Don't behave like a child  
↓  
Great → This product is great  
→ Ok battery died, great  
↓  
Sarcastic 23.0

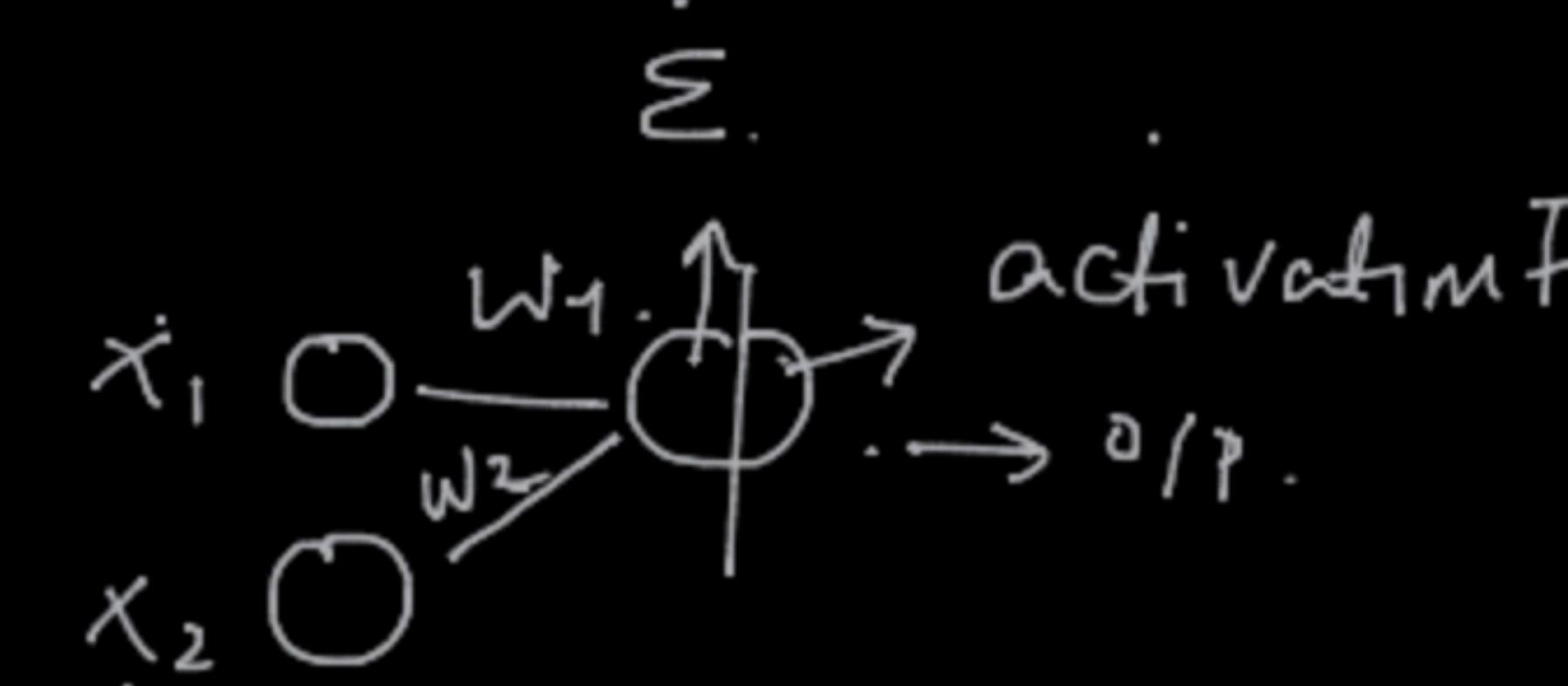
NLU → Natural Language Understanding

child → 81.3, kid → 80.5

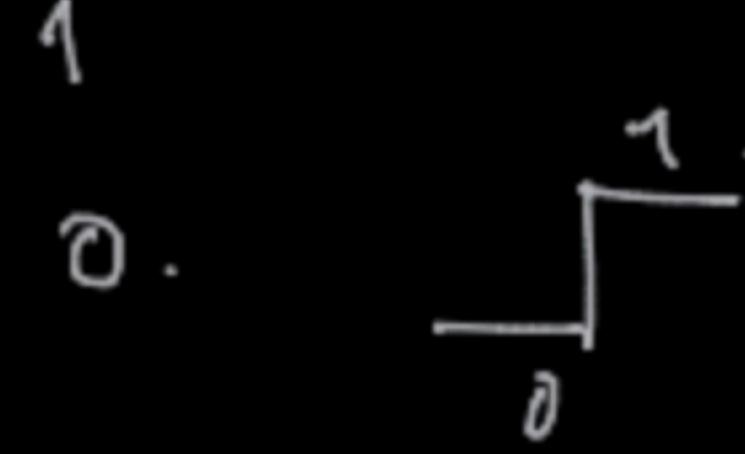
(1) Words which are used in similar contexts should have similar numbers

(2) Every context in which the word is used should have its own number

Neural networks → weights



$$(w_1 x_1 + w_2 x_2) > b$$



ReLU

Sigmoid

Softmax

tanh

Step

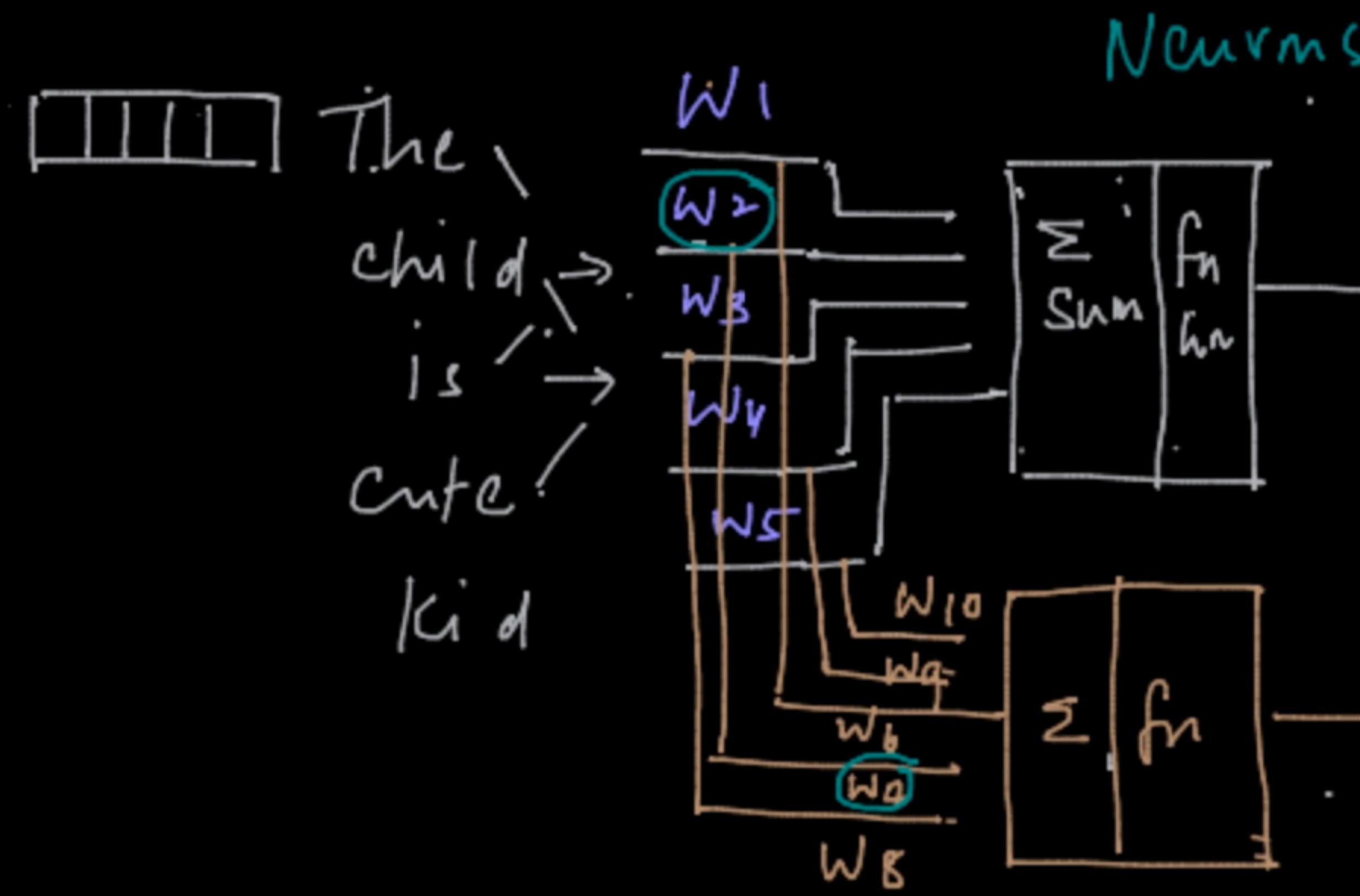
$x \rightarrow$  linear  $\rightarrow x$ .  
(identity fn).

$\begin{cases} \text{The child is cute.} \\ \text{The kid is cute.} \end{cases}$

Training data

Each word is given 2 numbers  $\rightarrow$  2 activation

1 activation  $\rightarrow$  1 number for each word

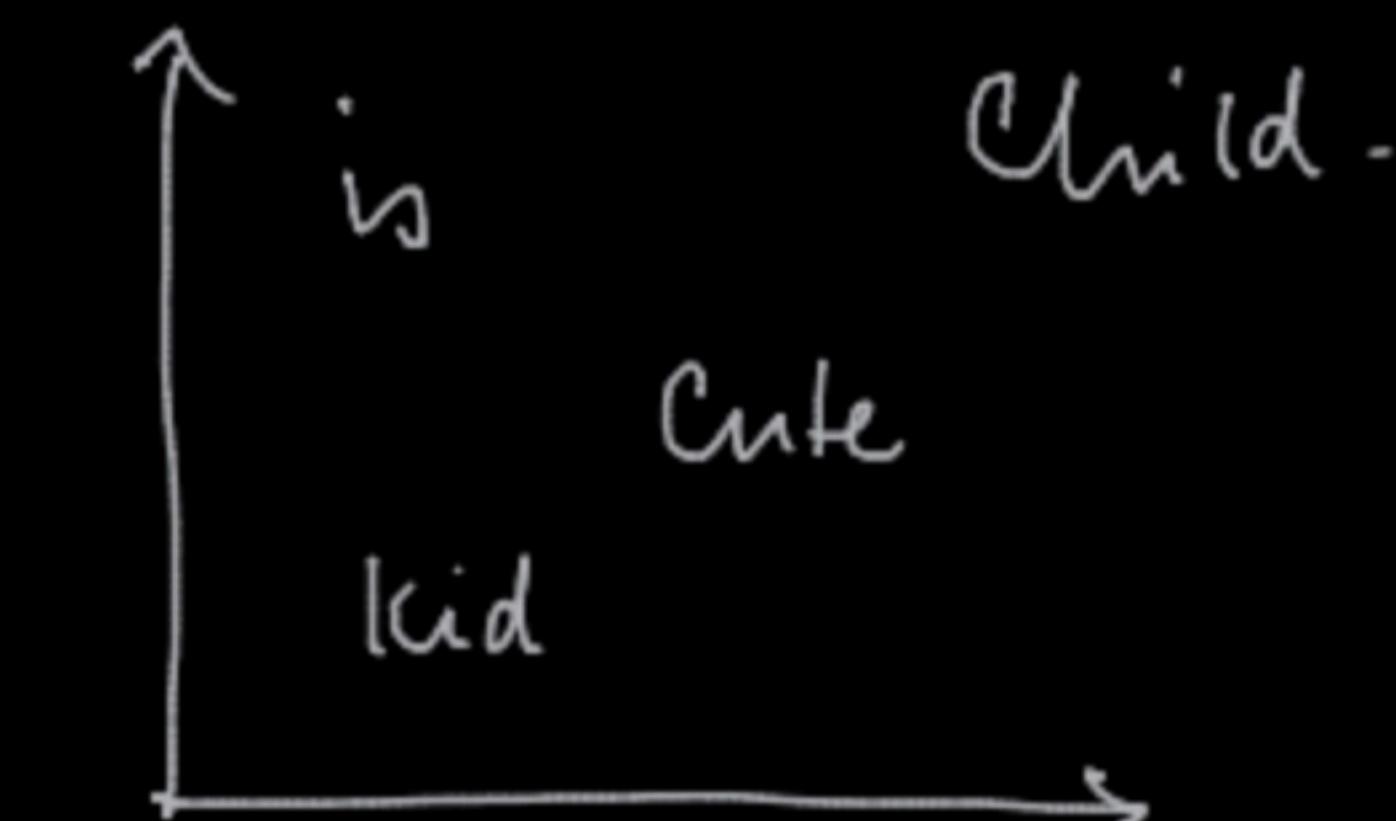


Neurons  
Next word

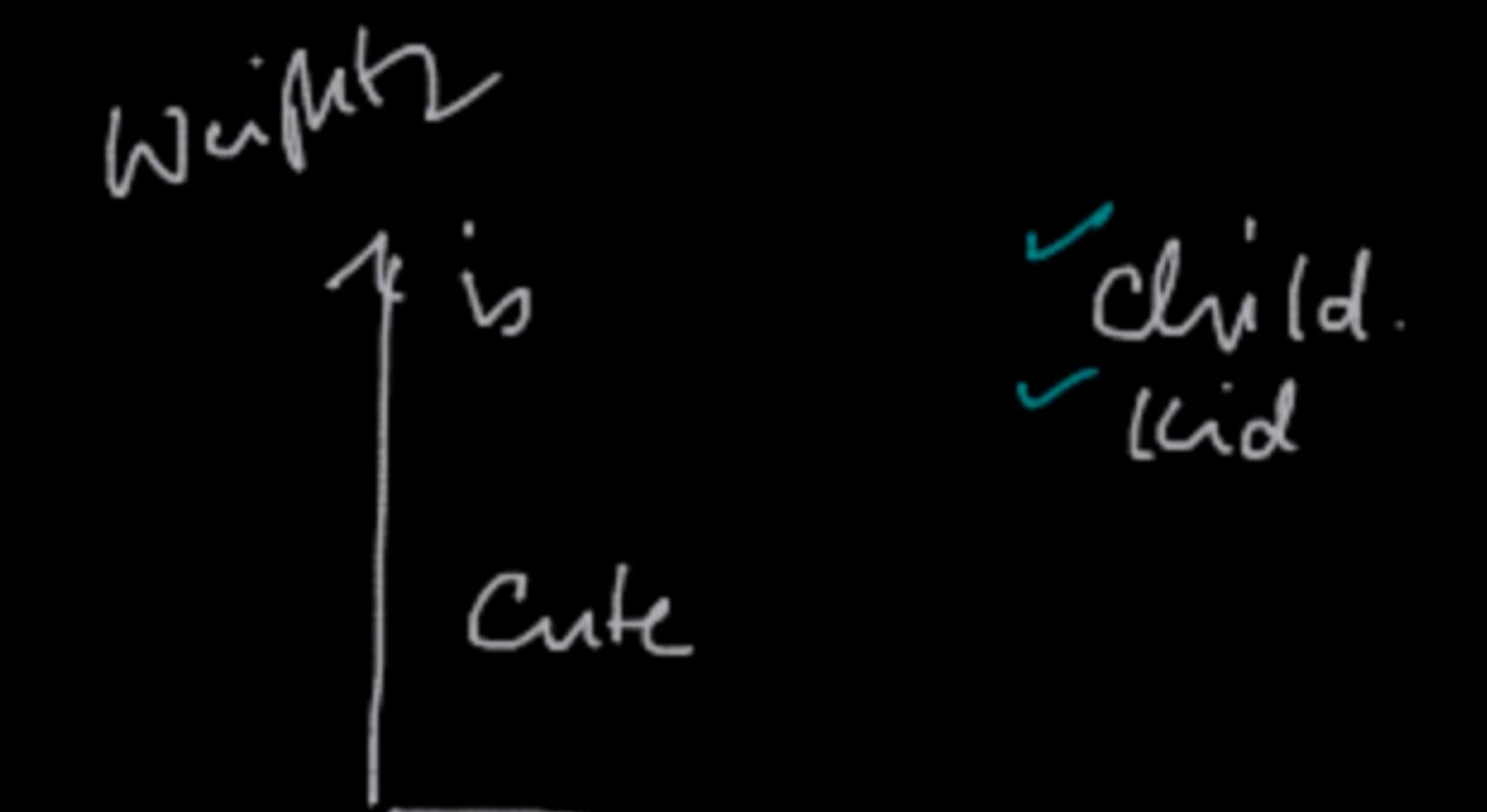
Next word

Embedding  $\leftarrow 2d \rightarrow$

$\begin{aligned} \text{child} &\rightarrow [w_2, w_7] \\ \text{The} &\rightarrow [w_1, w_6] \\ \text{is} &\rightarrow [w_3, w_8] \\ \text{cute} &\rightarrow [w_4, w_9] \\ \text{kid} &\rightarrow [w_5, w_{10}] \end{aligned}$



$\begin{aligned} y \rightarrow \hat{y} &\rightarrow \text{loss function} \\ &\rightarrow \text{gradient} \\ &\rightarrow \text{weights} \end{aligned}$



$\begin{aligned} \downarrow & \text{Center word in window} \\ \rightarrow P_1 &\rightarrow [ \dots ] \\ \rightarrow P_2 &\rightarrow [ \dots ] \end{aligned}$

Center word in window  $\rightarrow$  child. Backprop from here.

$\begin{aligned} \text{Bus} &\rightarrow 0.2 \\ \text{train} &\rightarrow 0.6 \\ \text{cab} &\rightarrow 0.2 \end{aligned}$

← Predict the next word →

Input

0 The ✓  
 1 Child  
 2 Is  
 3 Cute.  
 4 Kid

	Expected -	Predict -	The child is cute . The kid is cute .
0	The	0 0.31	
1	child	0 0.20	
2	is	1 0.43 → ✓	
3	Cute.	0 0.33	1. CBOW - — Continuous BoW .
4	Kid	0 0.25	2. Skipgram -

Window size → 3 .

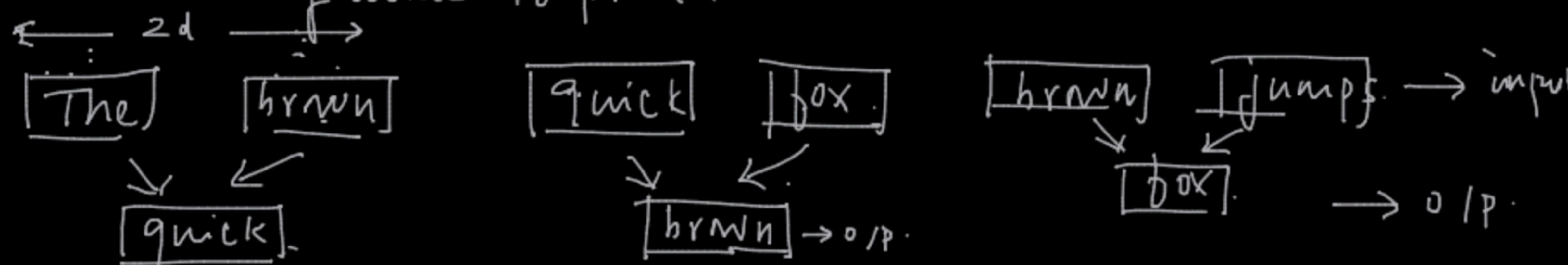
Child [0|1|0|0|b]

is [0|0|1|0|0| ]

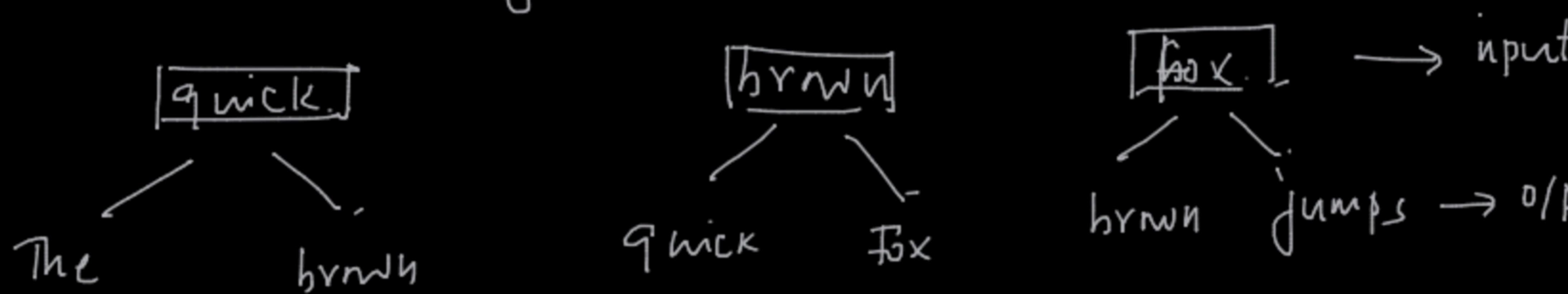
The quick brown fox jumps over the lazy dog → training data

Window size = 3

CBOW → use surrounding words to predict central word



skipgram → use a single word to predict surrounding words



<u>1/p</u>	<u>0/p</u>	<u>1/p</u>	
1	0	0	0 → Truth
0	1	0	0 → will
0	0	1	0 → cut
0	0	0	1 → yes
0	0	0	0 → true

Truth will cut yes → vocab = 5

