

## WORD2VECTOR

### Skip-gram:

Входной слой

c-слово

C1	C2	C3	...
1	0	0	
0	1	0	
0	0	1	

N-слов one-hot

Embedding слой


P-интерпараметр

задаётся случайно (потом с помощью G.D. находят оптимальные параметры)

Выходной слой

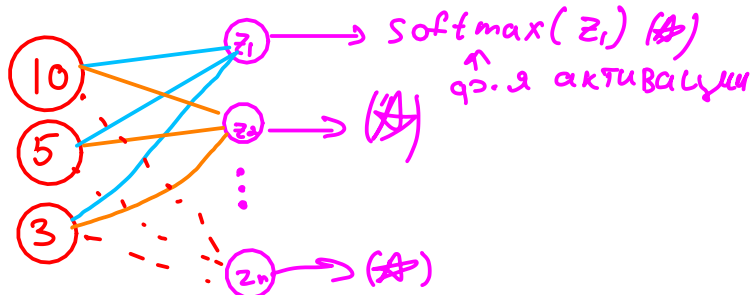
$P_1$
$P_2$
$P_3$
...

$(P_i)$  - вер-ть, что входное слово принадлежит  $i$ -му контексту

Пример:

$[0 | 1 | 0 | 0] \times$

1	4	9
10	5	3
6	8	6
...		



$$z_i = w_{i1} \cdot 10 + w_{i2} \cdot 5 + w_{i3} \cdot 3$$

параметры

$$\sum_{i=1}^n z_i = 1$$

CBOW:  $\leftrightarrow$  тоже самое, что Skip-gram, но входной и выходной слой меняются местами

ОСН. Понятия:

Контекст - это ближайшие слова (близость зависит от window size)