

## Word2Vec

CBOWskip-gram

$v = W \times \text{one\_hot}(v)$

$W \in \mathbb{R}^{V \times D}$

### CBOW

one-hot



$X = (x_{-c}, x_{-c+1}, \dots, x_{-1}, x_{+1}, \dots, x_{+c}) \in \mathbb{R}^{V \times 2C}$

$x_i \in \mathbb{R}^D$   
 $x' = WX + b$   
 $X = (x_{-c}, x_{-c+1}, \dots, x_{-1}, x_{+1}, \dots, x_{+c}) \in \mathbb{R}^{V \times 2C}$   
 $x' = [x_{-c}', x_{-c+1}', \dots, x_{-1}', x_{+1}', \dots, x_{+c}'] \in \mathbb{R}^{D \times 2C}$

$h = \sum_{j=-c, j \neq i}^{+c} \frac{1}{2C} (x_{-c}' + x_{-c+1}' + \dots + x_{-1}' + x_{+1}' + \dots + x_{+c}')$

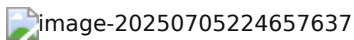
$P = (p_1, p_2, \dots, p_V) \in \mathbb{R}^V$   
 $P = W'h + b'$

### Softmax

$P' = (p'_1, p'_2, \dots, p'_V) \in \mathbb{R}^V$   
 $p'_i = \frac{\exp(p_i)}{\sum_{k=1}^V \exp(p_k)}$   
 $P' = \text{softmax}(P)$

**Cross Entropy**  
 $T = (t_1, t_2, \dots, t_V) \in \mathbb{R}^V$   
 $\text{Loss} = -\sum_{i=1}^V t_i \log(p_i)$

### skip-gram



输入

输入一个向量  $x_i \in \mathbb{R}^{V \times 1}$  的索引

输出

输出一个向量  $x_i \in \mathbb{R}^{V \times 1}$  的索引  $h$   $h = Wx_i + b$

计算

计算  $h$  的索引  $W_j$  的索引  $b_j$  的索引  $S_j \in \mathbb{R}^{V \times 1}$   $S = (S_1, S_2, \dots, S_{2C})^T$   $S_j = W_j' h + b_j'$

**Softmax**

计算  $S_j$  的 softmax 的索引  $P_j$   $P_j = (P_j(0), P_j(1), \dots, P_j(V-1))^T$   $P_j(k) = \text{Softmax}(S_j) = \frac{\exp(S_j(k))}{\sum_{l=0}^{V-1} \exp(S_j(l))}$

计算

计算  $P_j$  的索引  $1$  的索引  $0$  的索引

计算

计算 Cross Entropy 的索引  $P_j(k)$  的索引  $T$   $P_j = (P_1, P_2, \dots, P_{2C})$   $T_j = [t_j(1), t_j(2), \dots, t_j(V)]$   $L_j = -\sum_{k=1}^V t_k \log(P_j(k))$   $\text{Loss} = \sum_{j=1}^{2C} L_j$

输出

输出 word2vec 的索引  $V$  的索引 **Huffman** 的索引  $\log V$

**Softmax**

计算 **Huffman** 的索引

- 计算
- 计算

计算



计算

1. 计算 **softmax** 的索引  $P$  的索引
2. 计算  $P$  的索引  $1-P$
3. 计算
4. 计算

计算

计算

计算

