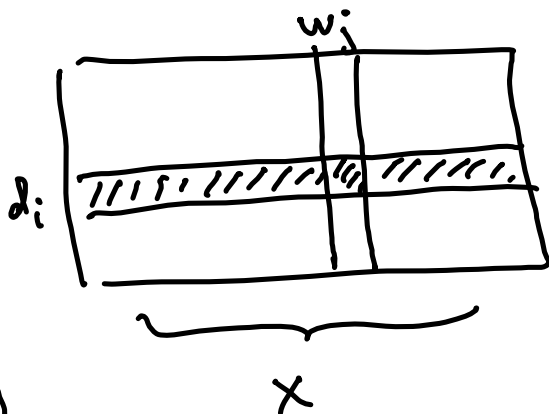
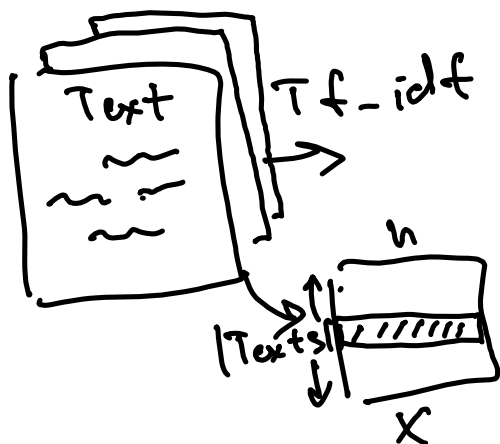


Word2Vec

$$w \rightarrow \boxed{\phantom{\text{vector}}} \in \mathbb{R}^h$$

$$D = \{d_1, \dots, d_{|D|}\}$$

$$d_i = \{w_1, w_2, \dots, w_{|d_i|}\}$$



$$\text{tezguino} \rightarrow \boxed{\phantom{\text{vector}}} v_1$$

$$\text{wine} \rightarrow \boxed{\phantom{\text{vector}}} v_2$$

$$\text{cat} \rightarrow \boxed{\phantom{\text{vector}}} v_3.$$

$$p(v_1, v_2) < p(v_1, v_3)$$

$$\begin{aligned}
 \mathbb{P}(w_1, w_2, \dots, w_T) &= \mathbb{P}(w_1) \cdot \mathbb{P}(w_2 | w_1) \\
 &\cdot \mathbb{P}(w_3 | w_2, w_1) \cdot \dots \cdot \mathbb{P}(w_T | w_1, \dots, w_{T-1}) \\
 &\approx \mathbb{P}(w_1 | w_3) \cdot \cancel{\mathbb{P}(w_3)} \cdot \mathbb{P}(w_2 | w_3) \cdot \\
 &\cdot \mathbb{P}(w_4 | w_3) \cdot \mathbb{P}(w_5 | w_3) \cdot
 \end{aligned}$$

$$L = \prod_{t=1}^T \prod_{\substack{-m \leq j \leq m \\ j \neq 0}} \mathbb{P}(w_{t+j} | w_j)$$

$$\mathbb{P}(\text{out} | \text{center}) = \frac{\exp(u_{\text{out}}^T \cdot v_{\text{center}})}{\sum_{u \in U} \exp(u^T v_{\text{center}})}$$

$$\text{softmax: } \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \rightarrow \begin{bmatrix} \frac{e^{a_1}}{e^{a_1} + e^{a_2} + \dots + e^{a_n}} \\ \frac{e^{a_2}}{e^{a_1} + \dots + e^{a_n}} \\ \vdots \end{bmatrix}$$

$$U = \text{[Diagram: A square representing a matrix of size } |Vocab| \times h \text{]} \\ |Vocab| \times h$$

$$V = \text{[Diagram: A square representing a matrix of size } |Vocab| \times h \text{]} \\ |Vocab| \times h$$

$$-\log P(\text{word}_i | \text{context}) = J$$

$$= -\log \left(\frac{\exp(u_{\text{word}_i}^T v_{\text{context}})}{\sum_{u \in \mathcal{U}} \exp(u^T v_{\text{context}})} \right) =$$

$$= -u_{\text{word}_i}^T v_{\text{context}} + \log \left(\sum_{u \in \mathcal{U}} \exp(u^T v_{\text{context}}) \right)$$

$$\nabla_{v_{\text{context}}} J = -u_{\text{word}_i} + \frac{\sum_u \exp(u^T v_{\text{context}}) \cdot u}{\sum_u \exp(u^T v_{\text{context}})} \\ \underbrace{\hspace{10em}}_{f(u)}$$

$$J = -u_{\text{unseen}}^T v_{\text{cat}} + \log \left(\sum_{u \in \{u_{i_1}, u_{i_2}, \dots, u_{i_k}\}} \exp(u^T v_{\text{cat}}) \right)$$

↑
Skip-Gram NS.

(CBOW)

Text = $w_1 w_2 \dots w_p$

$$\text{Vec}_{\text{text}} = \frac{v_{w_1} \cdot \text{idf}(w_1) + v_{w_2} \cdot \text{idf}(w_2) + \dots}{\text{idf}(w_1) + \text{idf}(w_2) + \dots}$$
