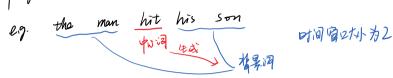
· Skip-gram & CBOW (Continuous Bag Of Words) 的下模型 negative sampling & hierarchical softmax 一份种训练方法

· skip-gram



$$P(W_{0}|W_{c}) = \frac{e^{\vec{u}_{0}^{T}\vec{v}_{c}}}{\sum_{i \in V} e^{\vec{u}_{0}^{T}\vec{v}_{c}}} \cdot i \exists \not \in V_{0}, \dots, \dots |V_{l-1}| i \not \in V_{l}, \dots |V_{l}| i \not \in V_{l}, \dots$$

· 对每个词C, 它在 uand wec 模型中有两个 向量散达:

Vc 017为代日 儿 奶帽

$$\frac{\partial \log P(w_0|w_c)}{\partial \vec{V}_c} = \vec{u}_0 - \sum_{j \in V} \frac{e^{\vec{u}_0 \vec{1} \cdot \vec{V}_c}}{\sum_{i \in V} e^{\vec{u}_i \vec{1} \cdot \vec{V}_c}} \vec{u}_j$$

$$\frac{\partial \log P(w_0|w_c)}{\partial \vec{V}_c} = \log e^{\vec{u}_0 \vec{1} \cdot \vec{V}_c} - \log \sum_{i \in V} e^{\vec{u}_i \vec{1} \cdot \vec{V}_c}$$

$$= \vec{u}_0 \vec{1} \cdot \vec{V}_c - \log \sum_{i \in V} e^{\vec{u}_0 \vec{1} \cdot \vec{V}_c}$$

$$\frac{\partial \log \Gamma(w_0 | w_c)}{\partial V_c} = \overline{U_0} - \frac{\partial \log \Gamma_{ev} e^{\overline{u_i}^T \cdot \overline{v_c}}}{\partial V_c}$$

$$= \overline{U_0} - \frac{1}{\Gamma_{ev} e^{\overline{u_i}^T \cdot \overline{v_c}}} \cdot \sum_{i \in V} \left(e^{\overline{U_i}^T \cdot \overline{v_c}} \cdot \overline{U_j} \right)$$

$$= \overline{U_0} - \frac{1}{\Gamma_{ev} e^{\overline{u_i}^T \cdot \overline{v_c}}} \cdot \sum_{i \in V} \left(e^{\overline{U_i}^T \cdot \overline{v_c}} \cdot \overline{U_j} \right)$$

$$= \overline{\mathcal{U}}_{0} - \frac{1}{\sum_{i \neq i} e^{i \vec{u}_{i}^{T} \cdot \vec{v}_{c}}} \cdot \sum_{i \neq i} e^{i \vec{u}_{i}^{T} \cdot \vec{v}_{c}} \cdot u_{j})$$

$$= \overline{\mathcal{U}}_{0} - \sum_{i \neq i} \frac{e^{i \vec{u}_{i}^{T} \cdot \vec{v}_{c}}}{\sum_{i \neq i} e^{i \vec{u}_{i}^{T} \cdot \vec{v}_{c}}} \cdot u_{j}$$

流感到红框中配成之,所以了简化为:

· CBOW

og. the man hit his son

$$\frac{2)m-\sum_{t=1}^{N}\log\left(w^{(t)}\right)^{N-1},w^{(t)},\dots,w^{(t)}}{\left(\sqrt{v_{0}}\right)^{N-1}+\sqrt{v_{0}}\right)/2m} \approx \frac{e^{iL}(\sqrt{v_{0}})^{N-1}+\sqrt{v_{0}}}{\left(\sqrt{v_{0}}\right)^{N-1}+\sqrt{v_{0}}\right)/2m}$$

$$=\frac{e^{iL}(\sqrt{v_{0}})^{N-1}+\sqrt{v_{0}}}{\left(\sqrt{v_{0}}\right)^{N-1}+\sqrt{v_{0}}\right)/2m}$$

$$=\frac{e^{iL}(\sqrt{v_{0}})^{N-1}+\sqrt{v_{0}}}{\left(\sqrt{v_{0}}\right)^{N-1}+\sqrt{v_{0}}\right)/2m}$$

$$=\frac{e^{iL}(\sqrt{v_{0}})^{N-1}+\sqrt{v_{0}}}{\left(\sqrt{v_{0}}\right)^{N-1}+\sqrt{v_{0}}\right)/2m}$$

$$=\frac{e^{iL}(\sqrt{v_{0}})^{N-1}+\sqrt{v_{0}}}{\left(\sqrt{v_{0}}\right)^{N-1}+\sqrt{v_{0}}\right)/2m}$$

$$=\frac{e^{iL}(\sqrt{v_{0}})^{N-1}+\sqrt{v_{0}}}{\left(\sqrt{v_{0}}\right)^{N-1}+\sqrt{v_{0}}}$$

窗口的旅来Im个词

$$\frac{\partial P(Wc|Wo_1, ---, Wo_{2m})}{\partial \vec{V}_{o_i}} = \frac{1}{2m} \left(\vec{U}_c - \sum_{j \in V} \frac{\vec{U}_j^T \cdot \vec{V}_c}{\sum_{i \neq V} e^{\vec{U}_i^T \cdot \vec{V}_c}} \vec{U}_j \right)$$

$$\frac{\partial P(Wc|Wo_1, ---, Wo_{2m})}{\partial \vec{V}_{o_i}} = \frac{1}{2m} \left(\vec{U}_c - \sum_{j \in V} P(W_j|W_c) \cdot \vec{U}_j \right)$$

- 。近似训练法
 - · 负采梓

- BWORK WO DO NIND / 10 WINZ THAN BURY WIN