CS224n Assignment2 Solution

Shawn

July 2020

1 Written

(a)

$y_w = 0$ for all $w \neq o$, and $y_o = 1$. So $-\sum_{w \in Vocab} y_w log(\hat{y}_o) = -log(\hat{y}_o)$

(b)

$J = -logexp(u_o^T v_c) + log(\sum_{w \in Vocab} exp(u_w^T v_c))$
$\frac{\partial J}{\partial v_c} = -u_o + \sum_{x \in Vocab} \frac{exp(u_x^T v_c)}{\sum_{w \in Vocab} exp(u_w^T v_c)} u_x$
$\frac{\partial J}{\partial v_c} = -Uy + U\hat{y}$

(c)

1) For $w = o$, $\frac{\partial J}{\partial u_w} = \frac{\partial J}{\partial u_o} = -v_c + \frac{exp(u_o^T v_c)}{\sum_{w \in Vocab} exp(u_w^T v_c)} v_c = -v_c + y^T \hat{y} v_c = (y^T \hat{y} - 1) v_c$
2) For $w \neq o$, $\frac{\partial J}{\partial u_w} = \frac{exp(u_w^T v_c)}{\sum_{w \in Vocab} exp(u_w^T v_c)} v_c = \hat{y}_w v_c$

(d)

$\frac{d\sigma(x)}{dx} = \sigma(x) - \sigma(x)^2$

(e)

$\frac{\partial J}{\partial v_c} = -(1 - \sigma(u_o^T v_c)) u_o - \sum_{k=1}^{K} (\sigma(-u_k^T v_c) - 1) u_k$
$\frac{\partial J}{\partial u_o} = -(1 - \sigma(u_o^T v_c)) v_c$
$\frac{\partial J}{\partial u_k} = (1 - \sigma(-u_k^T v_c)) v_c$
Compared with the naive-softmax, the partial derivatives of negative-sampling loss function is much efficient to compute since it dose not involve the matrix and vector multiplication with dimension of total word numbers, which is huge for large dataset.

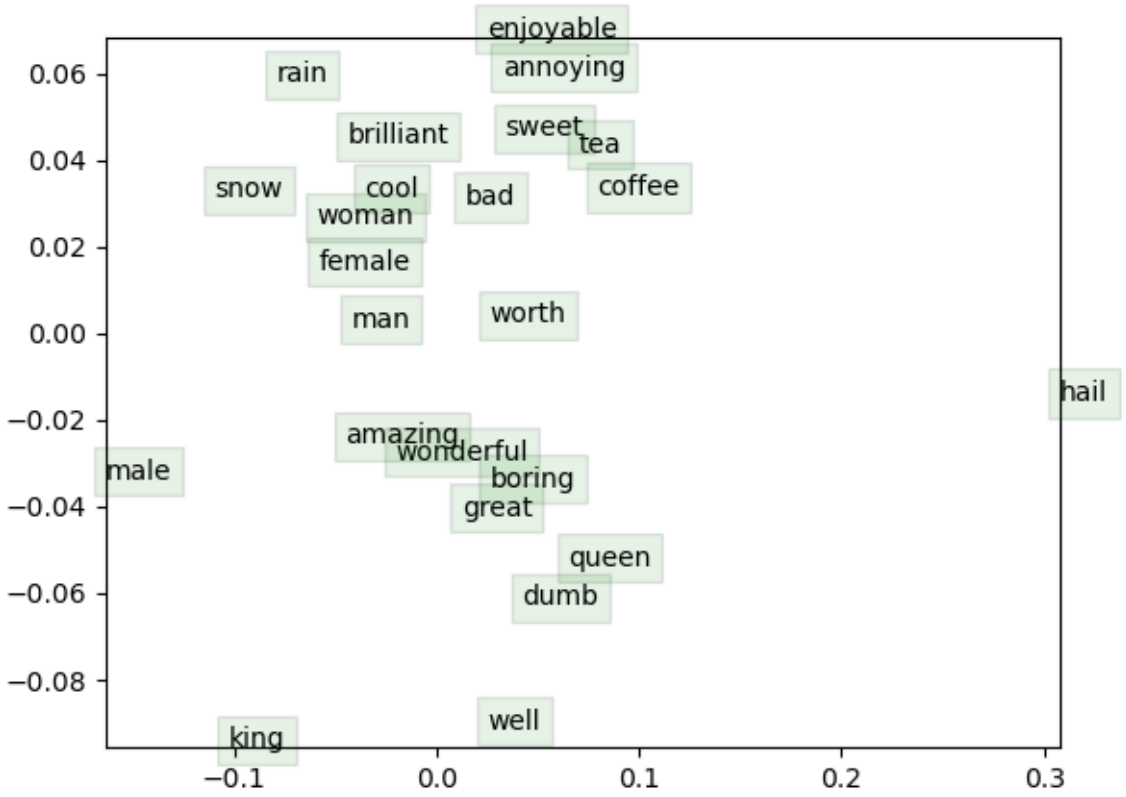

Figure 1: Word Vectors

(f)

$\frac{\partial J_{skip\_gram}(v_c,w_{t-m},...,w_{t+m},U)}{\partial U} = \sum_{-m\leq j\leq m,j\neq 0} \frac{\partial J(v_c,w_{t+j},U)}{\partial U}$

$\frac{\partial J_{skip\_gram}(v_c,w_{t-m},...,w_{t+m},U)}{\partial v_c} = \sum_{-m\leq j\leq m,j\neq 0} \frac{\partial J(v_c,w_{t+j},U)}{\partial v_c}$

$\frac{\partial J_{skip\_gram}(v_c,w_{t-m},...,w_{t+m},U)}{\partial w} = \frac{\partial J(v_c,w,U)}{\partial w}$

2 Word2Vec Implementing Result analysis

The Stanford Sentiment Treebank (SST) dataset word2Vec training result is showed in 1. It is clearly that words with similar semantics or similar class are clustered together, such as "woman" and "female", "tea" and "coffee", "amazing", "wonderful" and "boring".