

パラグラフベクトルの目的関数

$$L = \sum_d L_d$$

$$L_d = \frac{1}{T} \sum_{t=k+1}^T \log p(w_t | w_{t-k}, \dots, w_{t-1}),$$

$$p(w_t | w_{t-k}, \dots, w_{t-1}) = \frac{e^{y_{w_t}}}{\sum_{w'} e^{y_{w'}}},$$

$$y = b + Uh(w_{t-k}, \dots, w_{t-1}, d; W, D)$$

d : 文書または文

w_i : 文書または文中の i 番目の単語

W : 全ての単語のベクトルを含む行列

D : 全ての文書または文のベクトルを含む行列

T : 現在の文書または文に含まれる単語数

k : ウィンドウサイズ

ウィンドウ : ある単語の周辺を表す区間

p : softmax 関数により正規化された、文脈から現在の単語が導かれることの尤度

h : 引数となる単語と文書または文のベクトルを結合したベクトルを返す関数

文ベクトルの重み付け平均の式

$$\mathbf{t}_{i_{part}} = \sum_{i_{sent}} \frac{w(x_{i_{part}}(i_{sent}))}{|\sum_{i'_{sent}} w(x_{i_{part}}(i'_{sent}))|} \mathbf{s}_{i_{sent}},$$

$$x_{i_{part}}(i_{sent}) = \frac{i_{sent}}{\#sent - 1} - \frac{i_{part}}{\#part - 1},$$

$$w(x) = \begin{cases} \frac{1}{2}(\cos(\pi|x|) + 1) & \text{if } |x| \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

$\mathbf{s}_{i_{sent}}$: レビュー内の文ベクトル

$\mathbf{t}_{i_{part}}$: 重み付け平均された文ベクトル

i_{sent} : レビュー内の文ベクトルのインデックス

i_{part} : 重み付け平均された文ベクトルのインデックス

$\#sent$: レビュー内の文ベクトルの数

$\#part$: 重み付け平均された文ベクトルの数

ニューラルネットワークの構成

▶ 活性化関数：シグモイド関数

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

▶ 更新式：Adam

- ▶ 重みについて更新された値を累積していき更新料が少ない重みを大きく更新する手法
- ▶ 確率的勾配降下法（SGD）より収束が早い

▶ ドロップアウト

- ▶ ある層においていくつかのニューロンの出力を確率的に0にする正則化の手法
- ▶ 実験では，入力層で 0.2，中間層で 0.5 の確率

▶ 重み減衰

- ▶ ニューラルネットワークの重みの二乗和を目的関数に加え最小化する正則化の手法
- ▶ 実験では，係数 $5e-5$ で減衰