

$_v anish]GloVe$

$$F(w_i, w_j, w_k^c) = P_{ij}P_{jk}$$

$$\begin{aligned} &word_i X_i \\ &P_{ij} = \\ &P(j|i) = \\ &X_{ij} X_i X_i word_i \\ &Fw_i, w_j word_i, , word_j w_k^c word_k \end{aligned}$$

$$F((w_i-w_j)^T w_k^c) = F(w_i^T w_k^c)F(w_j^T w_k^c)$$

$$\begin{aligned} &w_i w_i^c \\ &y_i = \\ &\sigma(z_i) = \\ &\sigma(w_i x_i + \\ &b_i) \sigma \\ &gradient_v anish.png[rnn_v anish]RNN \end{aligned}$$

$$\begin{aligned} (1) \quad &\frac{\partial C}{\partial b_1} = \frac{\partial C}{\partial y_4} \frac{\partial y_4}{\partial z_4} \frac{\partial z_4}{\partial x_4} \frac{\partial x_4}{\partial z_3} \frac{\partial z_3}{\partial x_3} \frac{\partial x_3}{\partial z_2} \frac{\partial z_2}{\partial x_2} \frac{\partial x_2}{\partial z_1} \frac{\partial z_1}{\partial b_1} \\ (2) \quad &= \frac{\partial C}{\partial y_4} \sigma'(z_4) w_4 \sigma'(z_3) w_3 \sigma'(z_2) w_2 \sigma'(z_1) \\ &\frac{1}{4} \frac{\partial C}{\partial b_1} \\ &_{unit.png[rnn_v anish]LSTM} \end{aligned}$$

$$\begin{aligned} (3) \quad &i_t = \sigma_g(W^i x_t + U_i h_{t-1} + b^i) \\ (4) \quad &f_t = \sigma_g(W^f x_t + U_f h_{t-1} + b^f) \\ (5) \quad &o_t = \sigma_g(W^o x_t + U_o h_{t-1} + b^o) \\ (6) \quad &c_t = f_t \odot c_{t-1} + i_t \odot \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \\ (7) \quad &h_t = o_t \odot \sigma_h(c_t) \\ &\sigma_g \sigma_c, \sigma_h x_t h_t \\ &c_t W, U, b f_t \\ &i_t o_t \\ &_{\tilde{c}_5} ncoding.png[rnn_v anish] \end{aligned}$$

$$(8) \dots [h_1 c_1] = LSTM^{Text}(x_1, h_0, c_0)$$

$$(9) [h_T c_T] = LSTM^{Text}(x_1, h_{T-1}, c_{T-1})$$

$$(10) \dots [h_{T+1} c_{T+1}] = LSTM^{Hypothesis}(x_1, h_0, c_T)$$

$$(11) [h_N c_N] = LSTM^{Hypothesis}(x_1, h_{N-1}, c_{N-1})$$

$$c = \tanh(W h_N)$$

$$(12) (x_1 \dots x_T)(x_{T+1} \dots x_N) h_0, c_0$$

$$_{encode_{lstm}.png[rnn_v anish]}$$

$$(13) \dots [h_1 c_1] = LSTM^{target}(t_1, h_0, c_0)$$

$$(14) [h_M c_M] = LSTM^{target}(t_1, h_{M-1}, c_{M-1})$$

$$(15) \dots [h_1^{forward} c_1^{forward}] = LSTM^{forward}(x_1, h_0, c_T)$$

$$(16) [h_N^{forward} c_N^{forward}] = LSTM^{forward}(x_n, h_{N-1}^{forward}, c_{N-1}^{forward})$$

$$(17) \dots [h_N^{backward} c_N^{backward}] = LSTM^{backward}(x_n, h_0, c_T)$$