

Sentence:

"I like to eat pizza."

Vocabulary: POS Tags

"I" - Pronoun

"like" - Verb

"to" - Preposition

"eat" - Verb

"pizza" - Noun

$$\mathbf{a}^{(t)} = \mathbf{b} + \mathbf{W}\mathbf{h}^{(t-1)} + \mathbf{U}\mathbf{x}^{(t)}$$

$$\mathbf{h}^{(t)} = \tanh(\mathbf{a}^{(t)})$$

$$\mathbf{o}^{(t)} = \mathbf{c} + \mathbf{V}\mathbf{h}^{(t)}$$

$$\hat{\mathbf{y}}^{(t)} = \text{softmax}(\mathbf{o}^{(t)})$$

$$\mathbf{W}_{\text{xh}} = \begin{bmatrix} 0.4 & -0.3 & 0.1 & -0.2 & 0.5 \\ 0.1 & -0.2 & 0.3 & 0.2 & -0.4 \\ 0.2 & -0.1 & 0.5 & 0.4 & -0.3 \end{bmatrix}$$

$$\mathbf{W}_{\text{hy}} = \begin{bmatrix} 0.2 & 0.6 & -0.1 \\ 0.3 & -0.2 & 0.4 \\ -0.4 & 0.1 & 0.5 \\ 0.1 & 0.2 & 0.3 \end{bmatrix}$$

$$\mathbf{W}_{\text{hh}} = \begin{bmatrix} 0.2 & -0.1 & 0.3 \\ -0.1 & 0.4 & -0.2 \\ 0.4 & -0.3 & 0.5 \end{bmatrix}$$

$$\mathbf{h}_0 = [0.1, -0.1, 0.2]$$

Word "I":

$x_t = [1, 0, 0, 0, 0]$ # One-hot encoding for "I" in the vocabulary

$h_t = \tanh(W_{xh} * x_t + W_{hh} * h_{(t-1)})$

$h_t = \tanh([[0.4, -0.3, 0.1, -0.2, 0.5],$
 $[0.1, -0.2, 0.3, 0.2, -0.4],$
 $[0.2, -0.1, 0.5, 0.4, -0.3]] * [1, 0, 0, 0, 0]^T // 3 \times 5 * 5 \times 1$

+ $[[0.2, -0.1, 0.3],$
 $[-0.1, 0.4, -0.2],$
 $[0.4, -0.3, 0.5]] * [0.1, -0.1, 0.2]^T // 3 \times 3 * 3 \times 1$

$h_t = \tanh([0.49, 0.01, 0.37])$

$h_t = [0.4621, 0.0099, 0.3584]$

Tanh

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

$pos_scores = W_{hy} * h_t$

$pos_scores = [[0.2, 0.6, -0.1],$
 $[0.3, -0.2, 0.4],$
 $[-0.4, 0.1, 0.5],$
 $[0.1, 0.2, 0.3]] * [0.4621, 0.0099, 0.3584]$

$pos_scores = [0.06252, 0.28001, 0.19535, 0.15571]$

After softmax $[0.2229, 0.2777, 0.2547, 0.2446]$

Pronoun verb preposition noun

Cross-Entropy Loss (L) = $-\sum(T_i * \log(P_i))$

$L = -(1 * \log(0.2229) + 0 * \log(0.2777) + 0 * \log(0.2547) + 0 * \log(0.2446))$

$L = -(\log(0.2229))$

$L \approx 1.5020$