Sentence:
"I like to eat pizza."

Vocabulary: POS Tags
"I" - Pronoun
"like" - Verb
"to" - Preposition
"eat" - Verb
"pizza" - Noun

$$
\begin{aligned}
\boldsymbol{a}^{(t)} &= \boldsymbol{b} + \boldsymbol{W}\boldsymbol{h}^{(t-1)} + \boldsymbol{U}\boldsymbol{x}^{(t)} \\
\boldsymbol{h}^{(t)} &= \tanh(\boldsymbol{a}^{(t)}) \\
\boldsymbol{o}^{(t)} &= \boldsymbol{c} + \boldsymbol{V}\boldsymbol{h}^{(t)} \\
\hat{\boldsymbol{y}}^{(t)} &= \operatorname{softmax}(\boldsymbol{o}^{(t)})
\end{aligned}
$$

W_xh = [[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]]

W_hy = [[0.2, 0.6, -0.1],
[0.3, -0.2, 0.4],
[-0.4, 0.1, 0.5],
[0.1, 0.2, 0.3]]

W_hh = [[0.2, -0.1, 0.3],
[-0.1, 0.4, -0.2],
[0.4, -0.3, 0.5]]
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 //3x5 *5x1

+ [[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) //3x3 *3*1

h_t =tanh([0.49, 0.01, 0.37])
h_t=[0.4621, 0.0099, 0.3584]

$$Tanh$$

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

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) = -Σ(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 ≈ 1.5020