CS280 Spring 2025 Assignment 3
Part A

RNN and LSTM

April 8, 2025

Name: Chen Xuanxin

Student ID: 2024233125

1. RNN (10 points)

Assume that there is a simple RNN with all the variables as scalars. The hidden state s_t at time step t is computed from the previous hidden state s_t and the current input x_t as follows:

$$s_t = \text{step}(w_1 x_t + w_2 s_{t-1} + b),$$

where $\text{step}(z) = \begin{cases} 1, & z > 0 \\ 0, & \text{otherwise} \end{cases}$. The output is $y_t = s_t$.

(a) Assuming $s_0 = 0$, please provide the values of w_1, w_2 and b that could generate the output sequence

$$[0, 0, 0, 1, 1, 1, 1]$$

given the input sequence

$$[0, 0, 0, 1, 0, 1, 0].$$

(b) Suppose that we want a model to generate the output sequence

$$[1, 1, 1, 0, 0, 0, 1, 1]$$

given the input sequence

$$[0, 0, 0, 1, 0, 0, 1, 0].$$

Do you think it is possible to use the above RNN to implement the mapping from the input sequence to the output sequence? If you answered "no", please explain why. If you answered "yes", please provide the parameters w_1, w_2, and b required to implement the mapping.

Answer 1

(a) We need to find parameters w_1, w_2, and b such that the RNN generates the output sequence $[0, 0, 0, 1, 1, 1, 1]$ given the input sequence $[0, 0, 0, 1, 0, 1, 0]$. The hidden state s_t is computed as: $s_t = \text{step}(w_1 x_t + w_2 s_{t-1} + b)$, where $\text{step}(z) = 1$ if $z > 0$, and 0 otherwise. The output $y_t = s_t$.

To generate the desired output sequence, we analyze the input and output transitions:

1. Time step 1-3: Input is 0, output is 0. This implies:

$$w_1 \cdot 0 + w_2 \cdot 0 + b \leq 0 \implies b \leq 0.$$

2. Time step 4: Input is 1, output transitions to 1. This requires:

$$w_1 \cdot 1 + w_2 \cdot 0 + b > 0 \implies w_1 + b > 0.$$

3. Time step 5-7: Input alternates between 0 and 1, but output remains 1. This requires:

$$w_1 \cdot x_t + w_2 \cdot 1 + b > 0 \quad \text{for all } x_t \in \{0, 1\}.$$

To satisfy these conditions, we can set:

$$w_1 = 1, \quad w_2 = 2, \quad b = -1.$$

Verification:

- For $t = 1 - 3$: $s_t = \text{step}(1 \cdot 0 + 2 \cdot 0 - 1) = 0$.
- For $t = 4$: $s_4 = \text{step}(1 \cdot 1 + 2 \cdot 0 - 1) = \text{step}(0) = 1$.
- For $t \geq 5$: $s_t = \text{step}(1 \cdot x_t + 2 \cdot 1 - 1)$. Since $x_t \geq 0$, $1 \cdot x_t + 2 \cdot 1 - 1 \geq 1 > 0$, so $s_t = 1$.

(b) It is not possible to use the given RNN to generate the output sequence $[1, 1, 1, 0, 0, 0, 1, 1]$ from the input sequence $[0, 0, 0, 1, 0, 0, 1, 0]$. The output sequence starts with 1, which requires $s_0 = 1$. However, the problem states $s_0 = 0$, which is a contradiction. Even if $s_0 = 1$, the output sequence transitions from 1 to 0 at $t = 4$. This would require:

$$w_1 \cdot x_4 + w_2 \cdot s_3 + b \leq 0.$$

But $x_4 = 1$ and $s_3 = 1$, so:

$$w_1 + w_2 + b \leq 0.$$

However, the output transitions back to 1 at $t = 7$, requiring:

$$w_1 \cdot x_7 + w_2 \cdot s_6 + b > 0.$$

Since $x_7 = 1$ and $s_6 = 0$, this implies:

$$w_1 + b > 0.$$

These two conditions are contradictory:

$$w_1 + w_2 + b \leq 0 \quad \text{and} \quad w_1 + b > 0.$$

Therefore, no parameters w_1, w_2, b can satisfy both.

2. LSTM (10 points)

Consider a standard LSTM unit as shown below. Assume that L is the loss function. Given the gradient $\frac{\partial L}{\partial c_t}$ (backpropagated directly from c_{t+1} to c_t) and the gradient $\frac{\partial L}{\partial h_t}$, compute the following gradients: $\frac{\partial L}{\partial x_t}$ and $\frac{\partial L}{\partial h_{t-1}}$. Please provide a complete mathematical derivation and clearly present the formulas for each step.

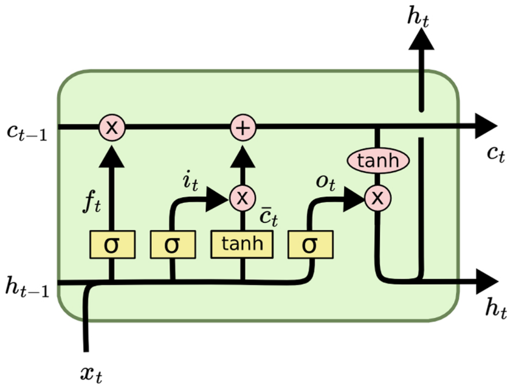

Figure 1: LSTM unit

$$
\begin{aligned}
i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i) \\
f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f) \\
o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o) \\
\bar{c_t} &= \tanh(W_c x_t + U_c h_{t-1} + b_c) \\
c_t &= f_t \odot c_{t-1} + i_t \odot \bar{c_t} \\
h_t &= o_t \odot \tanh(c_t)
\end{aligned}
$$

Answer 2

In order to compute the gradients $\frac{\partial L}{\partial x_t}$ and $\frac{\partial L}{\partial h_{t-1}}$, we start from the given LSTM equations and use backpropagation through time(**BPTT**).

1. LSTM Equations:

$$\begin{aligned}
i_t &= \sigma(W_i x_t + U_i h_{t-1} + b_i), \\
f_t &= \sigma(W_f x_t + U_f h_{t-1} + b_f), \\
o_t &= \sigma(W_o x_t + U_o h_{t-1} + b_o), \\
\bar{c}_t &= \tanh(W_c x_t + U_c h_{t-1} + b_c), \\
c_t &= f_t \odot c_{t-1} + i_t \odot \bar{c}_t, \\
h_t &= o_t \odot \tanh(c_t).
\end{aligned}$$

2. Given Gradients:

We are given:

$$\frac{\partial L}{\partial c_t} \quad \text{(backpropagated from } c_{t+1} \text{ to } c_t) \quad \text{and} \quad \frac{\partial L}{\partial h_t}.$$

3. Step-by-Step Derivation:

- Compute $\frac{\partial L}{\partial c_t}$:

$$\frac{\partial L}{\partial c_t} = \frac{\partial L}{\partial h_t} \odot (o_t \odot (1 - \tanh^2(c_t))) + \frac{\partial L}{\partial c_{t+1}} \odot f_{t+1}.$$

- Compute $\frac{\partial L}{\partial f_t}$:

$$\frac{\partial L}{\partial f_t} = \frac{\partial L}{\partial c_t} \odot c_{t-1} \odot f_t \odot (1 - f_t).$$

- Compute $\frac{\partial L}{\partial i_t}$:

$$\frac{\partial L}{\partial i_t} = \frac{\partial L}{\partial c_t} \odot \bar{c}_t \odot i_t \odot (1 - i_t).$$

- Compute $\frac{\partial L}{\partial \bar{c}_t}$:

$$\frac{\partial L}{\partial \bar{c}_t} = \frac{\partial L}{\partial c_t} \odot i_t \odot (1 - \bar{c}_t{}^2).$$

- Compute $\frac{\partial L}{\partial o_t}$:

$$\frac{\partial L}{\partial o_t} = \frac{\partial L}{\partial h_t} \odot \tanh(c_t) \odot o_t \odot (1 - o_t).$$

- Compute $\frac{\partial L}{\partial x_t}$:

$$\frac{\partial L}{\partial x_t} = \frac{\partial L}{\partial i_t} \cdot W_i^T + \frac{\partial L}{\partial f_t} \cdot W_f^T + \frac{\partial L}{\partial o_t} \cdot W_o^T + \frac{\partial L}{\partial \bar{c}_t} \cdot W_c^T.$$

- Compute $\frac{\partial L}{\partial h_{t-1}}$:

$$\frac{\partial L}{\partial h_{t-1}} = \frac{\partial L}{\partial i_t} \cdot U_i^T + \frac{\partial L}{\partial f_t} \cdot U_f^T + \frac{\partial L}{\partial o_t} \cdot U_o^T + \frac{\partial L}{\partial \bar{c}_t} \cdot U_c^T.$$