

CS280 Fall 2022 Assignment 3

Part A

RNN, LSTM and GRU

December 1, 2022

Name:Dai ZiJia

Student ID:2022233158

1. Parity-check network (16 points)

Note that the initial parity bit is 1, what's the relation between each input and the previous parity bit? Determine the relation between the parity and inputs and complete the parity bits (p_1, p_2, p_3, p_4) and design and draw a RNN to predict parity.

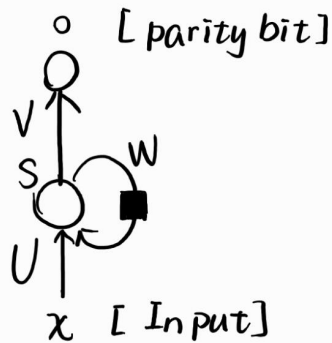
Parity bits :	0	0	0	1	0	1	p_1	p_2	p_3	p_4	\rightarrow
Input :	0	1	1	0	0	0	1	1	0	0	

Solution I

① It's Xor

② $P_t = P_{t-1} \oplus I_t$

$P_1 = 1$ $P_2 = 1$ $P_3 = 0$ $P_4 = 1$



2. GRU (17 points)

1. Draw the diagram of GRU, describe the gates (where? What is the role of each gate?), and point out the differences between GRU and LSTM in the design of gates.
2. In what situations(s) is LSTM/GRU used respectively? Explain your reason.

Solution II
1. Diagram of GRU

• Reset gates help capture short-term dependencies in sequence
• Update gates help capture long-term dependencies in sequence

Difference with LSTM :
LSTM has three gate, GRU only has two gate. LSTM use memory cell to include hidden state, but GRU directly pass the hidden state to the next unit.

2. In general task, I prefer to use GRU, because it's easier to converge with relatively few parameters. However, LSTM performs better when the dataset is large.