COMP 642 Assignment --- MODULE 10

1. These statements describe the advantages and disadvantages of RNN, mark each statement as True or False and give an explanation.

   a. Basic RNN can access information from a long time ago.
      a. False: There are features in the RNN that cause the algorithm to forget information from a long time ago, somewhat simulating the actual brain. This information forgetting is due to the gradient of the loss function as it is backpropagated.

   b. Basic RNN can take into account historical information but also future information.
      a. False: Only historical information, and up to the current time, can be taken into account for a recurrent neural network.

   c. Weights of RNN are shared across time.
      a. True: Sharing of weights is an important feature of RNN as these weights are shared across time and is a from of parameter linking.

   d. The model will be increasing with size of input.
      a. False: The number of parameters for the model does not change over time. RNN is designed to iterate over time period with shared weights across these periods. This means that the parameters will remain constant no matter the size of the input.

   e. Unlike traditional neural network, RNN can process input of any length.
      a. True: Due to the RNN's iteration process, it can ingest inputs of any length. This is used primarily in language models and time-series predictions where the input length can vary widely.

2. In the lecture the forward process of LSTM was outlined. GRU is a "simpler version" of LSTM. In this problem you'll practice stepping through the forward process of GRU. Calculate value of the output state and each gate for t1 and t2 as was done for LSTM in lecture. (calculation should be rounded to four decimals)

   Recall:

$$z = sigmoid(x_t U^z + s_{t-1} W^z)$$
$$r = sigmoid(x_t U^r + s_{t-1} W^r)$$
$$h = tanh(x_t U^h + (s_{t-1} \circ r) W^h)$$
$$s_t = (1 - z) \circ h + z \circ s_{t-1}$$

Time Step	X_0	X_1	y
T1	0.23	0.12	0.8
T2	0.48	0.98	0.3

$$U^z = \begin{vmatrix} 0.35 \\ 0.6 \end{vmatrix} \quad U^r = \begin{vmatrix} 0.45 \\ 0.12 \end{vmatrix} \quad U^h = \begin{vmatrix} 0.56 & 0.1 & 0.3 \\ 0.2 & 0.5 & 0.21 \end{vmatrix}$$

$$W^z = \begin{vmatrix} 0.12 \\ 0.67 \\ 0.34 \end{vmatrix} \quad W^r = \begin{vmatrix} 0.21 \\ 0.22 \\ 0.45 \end{vmatrix} \quad W^h = \begin{vmatrix} 0.54 & 0.1 & 0.2 \\ 0.2 & 0.66 & 0.3 \\ 0.9 & 0.87 & 0.3 \end{vmatrix}$$

Please refer to the Jupyter notebook for the details on the calculations

Output_t1_t2['t1']:

{'z': array([[0.5381]]),

'r': array([[0.5294]]),

'h': array([[0.1516, 0.0828, 0.0939]]),

's_t': array([[0.07 , 0.0383, 0.0434]])}

Output_t1_t2['t2']:

{'z': array([[0.691]]),

'r': array([[0.593]]),

'h': array([[0.4738, 0.5223, 0.3563]]),

's_t': array([[0.1948, 0.1878, 0.1401]])}

3. Execute hw_10.ipynb.

Submit a .doc or .pdf with your written answers. Submit your Python notebook. Submit a PDF of your Python notebook.