Week 8: Assignment 8

The due date for submitting this assignment has passed.

Due on 2024-09-18, 23:59 IST.

Assignment submitted on 2024-09-18, 21:13 IST

1) Match the following: *1 point*

1) One-to-many RNN architecture	i) Sentiment analysis of Netflix reviews
2) Many-to-one RNN architecture	ii) Convert English sentence to Hindi
3) Many-to-many RNN architecture with equal number of inputs and outputs	iii) Named entity recognition
4) Many-to-many RNN architecture with unequal number of inputs and outputs (Encoder and Decoder model)	iv) Music generation

- 1 → iii, 2 → iv, 3 → ii, 4 → i
- 1 → iv, 2 → i, 3 → iii, 4 → ii
- 1 → i, 2 → ii, 3 → iii, 4 → iv
- 1 → i, 2 → iii, 3 → ii, 4 → iv

Yes, the answer is correct.
Score: 1
Accepted Answers:
1 → iv, 2 → i, 3 → iii, 4 → ii

2) Consider an attention model for the image captioning task. At a particular time step t, the context vector z is obtained from the feature map of the final layer of the CNN and attention weights obtained from RNN. If RNN's attention output is $\begin{pmatrix} 0.25 & 0.5 \\ 0.25 & 0.0 \end{pmatrix}$ and feature map from CNN is $\begin{pmatrix} v_1 & v_2 \\ v_3 & v_4 \end{pmatrix}$ where v_1, v_2, v_3, v_4 are d-dimensional vectors, what is the value of the context vector z under soft and hard attention mechanisms respectively? *1 point*

- $\frac{v_1+2v_2+v_3}{4}, v_2$
- $\frac{2v_1+v_2+2v_3}{4}, \frac{2v_1+v_2+2v_3}{4}$
- $\frac{v_1+v_2+v_3+v_4}{4}, \frac{v_1+v_2+v_3+v_4}{4}$
- $\frac{v_1+v_2+v_3+v_4}{4}, v_2$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\frac{v_1+2v_2+v_3}{4}, v_2$

3) Which of the following statements are **true**? (Select all options that apply) *1 point*

- [ ] The number of learnable parameters in an RNN grows exponentially with input sequence length considered.
- [ ] An image classification task with a batch size of 16 is a sequential learning problem.
- [x] In an RNN, the current hidden state h_t not only depends on the previous hidden state h_{t-1} but implicitly depends on earlier hidden states also.
- [x] Generating cricket commentary for a corresponding video snippet is a sequence learning problem.

Yes, the answer is correct.
Score: 1
Accepted Answers:
In an RNN, the current hidden state h_t not only depends on the previous hidden state h_{t-1} but implicitly depends on earlier hidden states also.
Generating cricket commentary for a corresponding video snippet is a sequence learning problem.

4) Which one of the following statements is **true**? *1 point*

- Attention mechanisms cannot be applied to the bidirectional RNN model
- An image captioning network cannot be trained end-to-end even though we are using 2 different modalities to train the network
- One of the key components in the vanilla transformer are the recurrent connections that help them to deal with variable input length.
- All of the above
- None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
None of the above

5) Match the following ways of boosting image captioning techniques with attributes. Here, I =Image; A = Image Attributes; $f(.)$ is the function applied on them. *1 point*

(1) A1	(i) $x^{-2} = f(A)$ and $x^{-1} = f(I)$
(2) A5	(ii) $x^{-1} = f(A)$ and $x^{t} = f(I)$
(3) A2	(iii) $x^{-1} = f(A)$
(4) A4	(iv) $x^{-1} = f(I)$ and $x^{t} = f(A)$
(5) A3	(v) $x^{-2} = f(I)$ and $x^{-1} = f(A)$

- 1→iii, 2→ii, 3→v, 4→iv,5→i
- 1→iii, 2→iv, 3→i, 4→ii,5→v
- 1→iii, 2→iv, 3→v, 4→ii,5→i
- 1→i, 2→iii, 3→iv, 4→ii,5→v

No, the answer is incorrect.
Score: 0
Accepted Answers:
1→iii, 2→iv, 3→v, 4→ii,5→i

6) Which of the following statements are **true**? (Select all possible correct options) *1 point*

- [ ] Autoencoder can be equivalent to Principal Component Analysis (PCA) provided we make use of non-linear activation functions
- [x] When using global attention on temporal data, alignment weights are learnt for encoder hidden representations for all time steps
- [x] Positional encoding is an important component of the transformer architecture as it conveys information about order in a given sequence
- [ ] It is not possible to generate different captions for the same image that have similar meaning but different tone/style
- [ ] Autoencoders can not be used for data compression as its input and output dimensions are different

Yes, the answer is correct.
Score: 1
Accepted Answers:
When using global attention on temporal data, alignment weights are learnt for encoder hidden representations for all time steps
Positional encoding is an important component of the transformer architecture as it conveys information about order in a given sequence

7) Which of the following is true regarding Hard Attention and Soft Attention? *1 point*

- Hard Attention is smooth and differentiable
- Variance reduction techniques are used to train Hard Attention models
- Soft Attention is computationally cheaper than Hard Attention when the source input is large
- All of the above
- None of the above

Yes, the answer is correct.
Score: 1
Accepted Answers:
Variance reduction techniques are used to train Hard Attention models

8) Match the following attention mechanisms to their corresponding alignment score functions: *1 point*

(1) General Attention	(i) $v_a^T \tanh(W_a[s_t; h_i])$
(2) Content-Based Attention	(ii) $\alpha_{t,j} = \text{softmax}(W_a s_t)$
(3) Dot-Product Attention	(iii) $s_t^T h_i$
(4) Additive Attention	(iv) $\cos(s_t, h_i)$
(5) Location-Based Attention	(v) $s_t^T W_a h_i$

- 1→iii, 2→ii, 3→v, 4→iv,5→i
- 1→iii, 2→iv, 3→i, 4→ii,5→v
- 1→i, 2→iii, 3→iv, 4→ii,5→v

1→v, 2→iv, 3→iii, 4→i,5→ii

Yes, the answer is correct.
Score: 1

Accepted Answers:
1→v, 2→iv, 3→iii, 4→i,5→ii

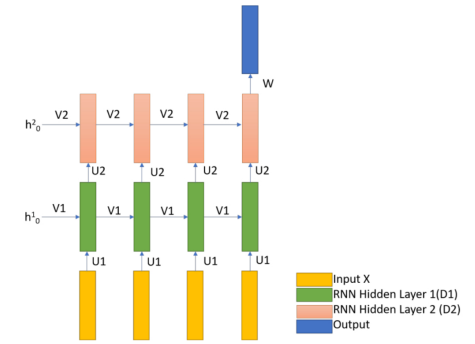
9) Which of the following statements is **true** (select all that apply): ***1 point***

- [ ] The number of learnable parameters in an RNN grows exponentially with input sequence length considered
- [x] Long sentences give rise to the vanishing gradient problem
- [x] Electrocardiogram signal classification is a sequence learning problem
- [x] RNNs can have more than one hidden layer

Yes, the answer is correct.
Score: 1

Accepted Answers:
Long sentences give rise to the vanishing gradient problem
Electrocardiogram signal classification is a sequence learning problem
RNNs can have more than one hidden layer

The RNN given below is used for classification:



The dimensions of the layers of RNN are as follows:

- Input $X \in \mathbb{R}^{132}$
- Hidden Layer 1 D1 $\in \mathbb{R}^{256}$
- Hidden Layer 2 D2 $\in \mathbb{R}^{128}$
- Number of classes:15

Note: Do not consider the bias term

10) Number of weights in Weight Matrix U1 is

33792

Yes, the answer is correct.
Score: 0.2

Accepted Answers:
(Type: Numeric) 33792

0.2 points

11) Number of weights in Weight Matrix V1 is

65536

Yes, the answer is correct.
Score: 0.2

Accepted Answers:
(Type: Numeric) 65536

0.2 points

12) Number of weights in Weight Matrix U2 is

32768

Yes, the answer is correct.
Score: 0.2

Accepted Answers:
(Type: Numeric) 32768

0.2 points

13) Number of weights in Weight Matrix V2 is

16383

No, the answer is incorrect.
Score: 0

Accepted Answers:
(Type: Numeric) 16384

0.2 points

14) Number of weights in Weight Matrix W is

1919

No, the answer is incorrect.
Score: 0

Accepted Answers:
(Type: Numeric) 1920

0.2 points

Consider an LSTM cell, and the data given below:

x_t =3
h_{t-1} =2
W_f =[0.1,0.2]
b_f =0
W_i=[-0.1,-0.3]
b_i =-1
W_o=[-1.2,-0.2]
b_o=1.5
W_C =[3,-1]
b_C =0.5
C_{t-1} =-0.5

Compute the following quantities (round upto 3 decimal places, refer formulas from lecture slides for computation). Note that we use scalars and vectors for ease of calculation here; in a realistic setup, this will be matrices and not scalars.

15) Forget Gate f_t

0.69

Yes, the answer is correct.
Score: 0.16

Accepted Answers:
(Type: Range) 0.679,0.699

0.16 points

16) Input Gate i_t

0.109

Yes, the answer is correct.
Score: 0.16

Accepted Answers:
(Type: Range) 0.099,0.119

0.16 points

17) Output Gate i_t

0.182

Yes, the answer is correct.
Score: 0.16

Accepted Answers:
(Type: Range) 0.172,0.192

0.16 points

18) New cell content C_t

0.998

Yes, the answer is correct.
Score: 0.16
Accepted Answers:
(Type: Range) 0.988,1.008

0.16 points

19) Cell State C_t

-0.236

Yes, the answer is correct.
Score: 0.16
Accepted Answers:
(Type: Range) -0.246,-0.226

0.16 points

20) Hidden state h_t

-0.042

Yes, the answer is correct.
Score: 0.2
Accepted Answers:
(Type: Range) -0.052,-0.032

0.2 points

Consider a GRU cell, and the following data:

- x_t =1.5
- h_{t-1} =-0.5
- W_z =[1.1,1.2]
- W_r =[-1.1,-1.3]
- W =[-1,-0.5]

Compute the following quantities (round upto 3 decimal places, bias values can be considered zero, refer formulas from lecture slides for computation). Note that we use scalars and vectors for ease of calculation here; in a realistic setup, this will be matrices and not scalars.

21) Update Gate z_t

0.777

Yes, the answer is correct.
Score: 0.25
Accepted Answers:
(Type: Range) 0.767,0.787

0.25 points

22) Reset Gate r_t

0.197

Yes, the answer is correct.
Score: 0.25
Accepted Answers:
(Type: Range) 0.187,0.207

0.25 points

23) New hidden state content h_t

-0.572

Yes, the answer is correct.
Score: 0.25
Accepted Answers:
(Type: Range) -0.582,-0.562

0.25 points

24) hidden state h_t

-0.516

No, the answer is incorrect.
Score: 0
Accepted Answers:
(Type: Range) -0.566,-0.546

0.25 points