X

swayam **(https://swayam.gov.in)** **(https://swayam.gov.in/nc_details/NPTEL)**

harshaldharpure9922@gmail.com ⌄

NPTEL (https://swayam.gov.in/explorer?ncCode=NPTEL) » Introduction to Large Language Models (LLMs) (course)

☰

If already registered, click to check your payment status

Course outline

About NPTEL ()

How does an NPTEL online course work? ()

Week 1 ()

Week 2 ()

Week 3 ()

Week 5 : Assignment 5

The due date for submitting this assignment has passed.
Due on 2025-02-26, 23:59 IST.

Assignment submitted on 2025-02-22, 10:22 IST

1) Which of the following is a disadvantage of Recurrent Neural Networks (RNNs)?   *1 point*

○ Can only process fixed-length inputs.

○ Symmetry in how inputs are processed.

◉ Difficulty accessing information from many steps back.

○ Weights are not reused across timesteps.

Yes, the answer is correct.
Score: 1
Accepted Answers:
Difficulty accessing information from many steps back.

2) Why are RNNs preferred over fixed-window neural models?   *1 point*

○ They have a smaller parameter size.

◉ They can process sequences of arbitrary length.

○ They eliminate the need for embedding layers.

○ None of the above.

Yes, the answer is correct.

Score: 1
Accepted Answers:
They can process sequences of arbitrary length.

3) What is the primary purpose of the cell state in an LSTM?   *1 point*

○ Store short-term information.

○ Control the gradient flow across timesteps.

◉ Store long-term information.

○ Perform the activation function.

Yes, the answer is correct.
Score: 1
Accepted Answers:
Store long-term information.

4) In training an RNN, what technique is used to calculate gradients over multiple   *1 point*
timesteps?

◉ Backpropagation through Time (BPTT)

○ Stochastic Gradient Descent (SGD)

○ Dropout Regularization

○ Layer Normalization

Yes, the answer is correct.
Score: 1
Accepted Answers:
Backpropagation through Time (BPTT)

5) Consider a simple RNN: *2 points*

● Input vector size: 3
● Hidden state size: 4
● Output vector size: 2
● Number of timesteps: 5
How many parameters are there in total, including the bias terms?

○ 210

◉ 190

○ 90

○ 42

No, the answer is incorrect.
Score: 0
Accepted Answers:
42

6) What is the time complexity for processing a sequence of length 'N' by an RNN, if the   *1 point*
input embedding dimension, hidden state dimension, and output vector dimension are all 'd'?

○ O(N)

○ O(N²d)

○ O(Nd)

◉ O(Nd²)

Yes, the answer is correct.
Score: 1

Accepted Answers:
O(Nd²)

7) Which of the following is true about Seq2Seq models?   *1 point*

(i) Seq2Seq models are always conditioned on the source sentence.
(ii) The encoder compresses the input sequence into a fixed-size vector representation.
(iii) Seq2Seq models cannot handle variable-length sequences.

◉ (i) and (ii)

○ (ii) only

○ (iii) only

○ (i), (ii), and (iii)

Yes, the answer is correct.
Score: 1

Accepted Answers:
(i) and (ii)

8) Given the following encoder and decoder hidden states, compute the attention   *2 points*
scores. (Use dot product as the scoring function)

Encoder hidden states: h_1 = [1,2], h_2 = [3,4], h_3 = [5,6]
Decoder hidden state: s = [0.5,1]

○ 0.00235,0.04731,0.9503

◉ 0.0737,0.287,0.6393

○ 0.9503,0.0137,0.036

○ 0.6393,0.0737,0.287

No, the answer is incorrect.
Score: 0

Accepted Answers:
0.00235,0.04731,0.9503