

Indian Institute of Technology, Patna
End Semester Examination (2024-25)
CS411: Deep Learning for Natural Language Processing

Time: 3 Hrs.

Max. Marks: 100

Attempt all Questions. Please be precise and concrete in your answers.

Q1. Draw the architecture of GRU cell and define each component mathematically. [10]

Q2. What is the drawback of Vanilla Encoder-Decoder architecture and what solutions have been proposed before the year 2017? [4+6]

Q3. In Luong Attention Mechanism, how weights are calculated for different parts of the input in order to generate the output? Discuss and express it mathematically. [10]

Q4. Discuss the benefits of multi-head attention over single-head attention in transformer architecture. Also discuss the purpose of masked multi-head attention in the Transformer architecture. [5+5]

Q5. What is Autoregressive mode in transformer? Suppose you are working on the task of machine translation and you would like to predict the output for the sequence "Open the door". How the given sequence would be processed in Transformer architecture?

[3+7]

Q6. Find the positional encoding in six dimensional vector of the second position words in the given sequence "Open the door". [10]

Q7. What is Transfer Learning? In ULMFiT transfer learning method, what techniques has been proposed to enhance the performance of the model. [4+6]

Q8. What is Dynamic Memory Network? Draw its architecture and discuss each components. [4+6]

Q9 a. What is information retrieval? How is it different from data retrieval? [2+2]

9b. What do you mean by Vector Space Model? Derive the expression of its ranking function. [6]

Q10 a. A system returns documents having following relevance values in the following order: <2, 0, 1, 0, 0, 2, 1, 0, 2, 1> (2 = highly relevant, 1 = partially relevant, 0 = non-relevant). Find nCG and nDCG. [3+3]

10 b. Consider non-zero (both 1 and 2) scores above as relevant and 0 as non-relevant in binary relevance settings. Find the natural average precision for the query (assume number of relevant docs for the query = 20). [4]

