

CSL 7340 - Natural Language Understanding
Jan - April 2022 End Semester Exams
Time - 2 hours Full marks - 20

(Answers should be to the point and precise. Diagrams should be neat.)

1. When it comes to handling previously unseen or rare words for text analytics tasks, what kind of behaviour do you expect from each of the following pre-trained word embeddings :- Word2vec, GLOVE and BERT. Justify your answer. (3)
2. Explain how attention based mechanisms are employed by transformers to ensure better performance than LSTMs, when used to build machine translation systems. (3)
3. You have to build a feedback analysis system for your company. The aspects that people talk about can change from year to year. The organization needs that your software will provide the top 5 issues that people are talking about along with their relative sentiments.
 - (a). Provide a modular diagram for the system - indicating how information flows within the system from ingestion of feedbacks to generating the analytical report. Give a one-line description of each module's functionality.
 - (b). What is the backend mechanism to identify and measure the sentiment for each issue?
 - (b). How will the end-user be sure that the software is yielding correct results - in other words, what kind of validations will you put in place? (2 + 2 + 1 = 5)
4. What is a knowledge graph? What are its uses? (1 + 1 = 2)
5. You are required to build a chatbot which will cater to travellers and provide the following functionalities :
 - (i). Answer user queries about possible travel destinations based on their preferences
 - (ii). Provide information about accommodation in their chosen destination
 - (iii). Assist in booking

Provide a high-level schema diagram for designing the solution. Clearly indicate the purpose of each module in your diagram along with the basic technology that you will use to power such a chatbot. How can the effectiveness of the system be measured? (3 + 1 = 4)

6. The following boxes show three short essays written by three school children. Explain how you will design an automated scoring system for grading the essays. (3)

Tom loves reading books. He prefers reading books at library. So he always goes to library. (a)	Tom love read book. He prefer to reading books at library. So he always goes to library. (b)	Tom loves reading books. He missed his lunch today. So he always goes to library (c)
--	---	---