| Name:<br>Enrolment No: | |
|---|---|

**UPES**

**End Semester Examination, May 2023**

**Course: Natural Language Processing** **Semester: VI**

**Program: B.Tech (CSE)** **Time : 03 hrs.**

**Course Code: CSEG 3043P** **Max. Marks: 100**

**Instructions: All questions are compulsory. Question no. 9 of Section B and Question no. 10 of Section C have internal choice.**

**SECTION A**
**(5Qx4M=20Marks)**

| S. No. | | Marks | CO |
|---|---|---|---|
| Q 1 | Write 4 examples of words with prefix or suffix. Perform stemming and lemmatization on these words. | **4** | **CO1** |
| Q 2 | Tri-grams are used in language models. Provide 4 examples of tri-gram words. | **4** | **CO2** |
| Q 3 | How is probability used in machine translation? | **4** | **CO3** |
| Q 4 | Differentiate between top-down and bottom-up parser in NLP. | **4** | **CO4** |
| Q 5 | Write 4 recent platforms/applications of Question Answering. | **4** | **CO5** |

**SECTION B**
**(4Qx10M= 40 Marks)**

| | | | |
|---|---|---|---|
| Q 6 | How is corpus imported using NLTK? Write code to import corpus and discuss output of code in brief. | **10** | **CO1** |
| Q 7 | Describe techniques for sentiment analysis. Write python code for sentiment analysis. | **10** | **CO2** |
| Q 8 | Differentiate between rule-based and probability-based machine translations? How probability-based machine translation calculates the best possible target text? | **10** | **CO3** |
| Q 9 | Discuss "Word sense disambiguation" problem in NLP and explain Dictionary and knowledge-based method to resolve the word sense ambiguity.<br><br>or<br>Explain vector space model of Information Retrieval. Suppose there exist two documents:<br>1) $d_f$ with fried chicken recipe, and<br>2) $d_p$ with poached chicken recipe. | **10** | **CO4** |

| | | | |
|---|---|---|---|
| | Discuss why a query for *fried chicken* will match document $d_f$ rather than document $d_p$? | | |
| **SECTION-C (2Qx20M=40 Marks)** | | | |
| Q 10 | Explain the need for probabilistic parsing in CKY algorithm. Using PCKY parser, calculate probability for the sentence "the flight includes a meal" using the rules given below:<br><br>$S \rightarrow NP\ VP$ .80 $\quad Det \rightarrow the$ .40<br>$NP \rightarrow Det\ N$ .30 $\quad Det \rightarrow a$ .40<br>$VP \rightarrow V\ NP$ .20 $\quad N \rightarrow meal$ .01<br>$V \rightarrow includes$ .05 $\quad N \rightarrow flight$ .02<br><br>or<br><br>Using same grammar given in above question, construct the possible parse trees for the sentence "the flight includes a meal" using top-down parser. Calculate probability using statistical parsing for every possible parse tree. | **20** | **CO4** |
| Q 11 | Elaborate applications of document clustering. Discuss with real example where document clustering is used. How k-means clustering is used to create clusters of documents? Explain with example. | **20** | **CO5** |