

Course Name: Computational Linguistics and Natural Language Processing

Course Outcome

- CO1: Apply basic concepts of Text processing
 CO2: Apply the concept of Vectors and text similarity
 CO3: Formulate text based classification
 CO4: Design and develop applications of NLP

Printed Pages: 2

University Roll No.

Mid Term Examination, Odd Semester 2023-24
B.Tech (CSE-AIML), 3rd Year, V Semester

**Subject Code & Subject Name- COMPUTATIONAL LINGUISTICS
 AND NATURAL LANGUAGE PROCESSING (BCSE 0704)**

Maximum Marks: 30

Time: 2 Hours

Instruction for students:
 All questions are compulsory

Section - A

3 X 5 = 15 Marks

Attempt All Questions

No.	Detail of Question	Marks	CO	BL	KL
1	Consider the following sentences: S1= "GLA is University in India" S2= "Flies like a flower" S3= "I prefer a morning flight" Apply stemming, lemmatization, stop-words removal and POS tagging on these sentences and write sentences after performing these text processing.	3	CO1	E	P
2	Identify any scenario where NLP is to be applied such as Chatbot, Voice assistant, Sentiment analysis etc. Describe the steps used for NLU, NLP and NLG for selected scenario.	3	CO1	A	C
3	Consider an example to find the similarity between two vectors – 'x' and 'y', using Cosine Similarity. The 'x' vector has values, $x = \{3, 2, 0, 5\}$ The 'y' vector has values, $y = \{1, 0, 0, 0\}$	3	CO2	An	P
4	Write the name of POS tag for following symbols: JJ, IN, VB, MD, RB	3	CO1	R	F
5	Consider a scenario of next word prediction. 3-gram approach should be used to predict text. Given a word, how the chances of this word can be calculated using 3- gram.	3	CO2	U	C

Section – B

Attempt All Questions

5 X 3 = 15 Marks

No.	Detail of Question	Marks	CO	BL	KL
6	In email spam classification, write regular expressions to remove punctuation symbols, extract phone numbers, email id, valid email id, dates, HTML tags, valid URLs	5	CO1	U	C
7	You are building a recommendation system for a news website. How can you use Word2Vec to find similar words to a user's search query and suggest relevant articles based on those similar words?	5	CO2	C	M
8	D1= “Cricket is one of the most famous sports in India” D2= “NIFTY 50 is stock index in India” D3= “GLA is University in India” Q= “What is the name of stock index in India” Find the similarity between query and document corpus using TF/IDF and rank the similarity in descending order.	5	CO2	E	P

Section - B

5 X 3 = 15 Marks

Attempt All Questions

No.	Detail of Question	Marks	CO	BL	KL
6	In email spam classification, write regular expressions to remove punctuation symbols, extract phone numbers, email id, valid email id, dates, HTML tags, valid URLs	5	CO1	U	C
7	You are building a recommendation system for a news website. How can you use Word2Vec to find similar words to a user's search query and suggest relevant articles based on those similar words?	5	CO2	C	M
8	D1= "Cricket is one of the most famous sports in India" D2= "NIFTY 50 is stock index in India" D3= "GLA is University in India" Q= "What is the name of stock index in India" Find the similarity between query and document corpus using TF/IDF and rank the similarity in descending order.	5	CO2	E	P