347631

Course Name:

Course Outcome

CO1-Unders and and core task of NLP and its application

CO2- Understand the human language, demonstrate similarity computation for documents

CO3-Analyzing indexing and pre-processing of textual documents for NLP

CO4- Understand the algorithmic description the main language levels: morphology, syntax, semantics and pragmatics

Printed Pages : **University Roll No. .....**

Mid Term Examination, Odd Semester 2024-25

B.Tech (AIML), III Year, V Semester

BCSE0704: COMPUTATIONAL LINGUISTICS AND NATURAL LANGUAGE PROCESSING

Time: 2 Hours **Maximum Marks: 30**

Section – A

Attempt All Questions 3 X 5 = 15 Marks

No.	Detail of Question	Marks	CO	BL	KL
1	Differentiate between Jaccard and cosine similarity? Compute the Jaccard coefficient between words W1 and W2 based on char bigram model? W1=night and W2=nicht	3	CO2	R	F
2	Write short notes on the following: a) Tokenization b) Lemmatization c) Stemming	3	CO1	R	F
3	Identify any scenario where NLP is to be applied such as Sentiment analysis. Describe the steps used for NLP for selected scenario.	3	CO1	A	C
4	Define Parser. Why is it used in NLP? Differentiate between top down and bottom up approach in parser.	3	CO3	An	C
5	Write the names of pos for following Symbols: JJ, IN, VB, MD, RB, PDT	3	CO2	An	C

Section – B

Attempt All Questions 5 X 3 = 15 Marks

No.	Detail of Question	Marks	CO	BL	KL
6	Define the language. Differentiate between semantics and pragmatics? Construct CFG using the sentence –"The man read this book" where S→ NP VP S→Aux NP VP S→VP NP→Det Noun VP→V VP→V NP	5	CO4	U	C
7	Consider the following four documents: Doc1: Breakthrough drug for covid Doc2: new covid drug Doc3: new approach for treatment of covid Doc4: new hopes for covid patients a) Draw the term document incidence matrix for the document collection b) Draw the inverted index represented for this collection c) Process the query covid AND NOT new. Differentiate between BOW and Word2vec.	5	CO2	C	C
8	T1=Julie loves me more than Heena loves me T2= Jasmin likes me more than Julie loves me We want to know how similar these texts are, purely in the term of word counts ignoring word order. a) What is the size of vector generated for above text in term of similarity computation b) Compute the number of times each of these word appears in each text. c) Find the cosine similarity between T1 nad T2 d) When to use jaccard similarity and when to use cosine similarity e) What is jaccard distance	5	CO2	E	M