Hall Ticket Number:

Code No. : 18621 A

VASAVI COLLEGE OF ENGINEERING *(AUTONOMOUS)*, HYDERABAD
Accredited by NAAC with A++ Grade
B.E. (I.T.) VIII-Semester Main & Backlog Examinations, June-2022
Natural Language Processing (PE-V)

Time: **3 hours**   Max. Marks: **60**

*Note: Answer all questions from **Part-A** and any **FIVE** from **Part-B***

Part-A (10 × 2 = 20 Marks)

Q. No.	Stem of the question	M	L	CO	PO
1.	What are the tasks applied in Text Normalization?	2	1	1	1
2.	Define extrinsic evaluation and intrinsic evaluation of language models?	2	1	1	1
3.	Discuss the confusion matrix in terms of precision and recall.	2	1	2	1
4.	How sigmoid function does is useful in logistic regression?	2	1	2	1
5.	How to compute cosine similarity metric between two vectors v and w?	2	1	3	1
6.	Explain Point wise Mutual Information (PMI) w.r.t to two independent events x and y.	2	1	3	1
7.	Give the formulas for sigmoid, tanh and ReLU activation functions.	2	2	4	1
8.	What is the role of bi-directional RNN?	2	1	4	1
9.	Illustrate the encoder-decoder networks components.	2	2	5	1
10.	Discuss about task –based dialogue system.	2	1	5	1

Part-B (5 × 8 = 40 Marks)

		M	L	CO	PO
11. a)	Discuss how Hidden Markov Model is used for Part-of-Speech Tagging?	3	3	1	2
b)	Write regular expressions for the following languages. By "word", we mean an alphabetic string separated from other words by whitespace, any relevant punctuation, line breaks, and so forth.	5	4	1	3
	1. The set of all strings with two consecutive repeated words (e.g., "Humbert Humbert" and "the the" but not "the bug" or "the big bug");				
	2. All strings that start at the beginning of the line with an integer and that end at the end of the line with a word.				
12. a)	Analyze how Naive Bayes is used for multi-class text classification.	3	2	2	2
b)	Write down how logistic regression perform in NLP and briefly discuss the algorithm of stochastic gradient descent.	5	4	2	2

			M	L	CO	PO
13.	a)	Demonstrate term-document matrix for four words.	3	2	3	1
	b)	Apply Learning skip-gram embedding with suitable scenario.	5	4	3	3
14.	a)	Explain how LSTMs and GRUs are using language models with neat diagram.	5	2	4	1
	b)	Apply chain rule on backward differentiation on computation graphs for neural networks.	3	4	4	3
15.	a)	How to relate encoder-decoder with RNN?	3	3	5	2
	b)	Explain how RNNs are used for machine translation?	5	4	5	2
16.	a)	Explain the statistical hypothesis testing and briefly analyze the important of minimum edit distance algorithms in NLP.	5	2	1	1
	b)	What is the naive assumption behind Naive Bayes and does this assumption hold for language-based classification tasks?	3	3	2	2
17.		Answer any *two* of the following:				
	a)	Define TF-IDF discuss its applications?	4	2	3	1
	b)	Explain feed forward neural networks with neat diagram. How do we model it for regression task	4	2	4	1
	c)	Illustrate Encoder-Decoder with Transformers.	4	2	5	1

M : Marks;　L: Bloom's Taxonomy Level;　CO; Course Outcome;　PO: Programme Outcome

i)	Blooms Taxonomy Level – 1	20%
ii)	Blooms Taxonomy Level – 2	40%
iii)	Blooms Taxonomy Level – 3 & 4	40%

*****