Hall Ticket Number:

Code No. : 18631 (A) N/O

VASAVI COLLEGE OF ENGINEERING *(AUTONOMOUS)*, HYDERABAD

Accredited by NAAC with A++ Grade

B.E. (I.T.) VIII-Semester Main & Backlog Examinations, May-2023

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.	For each sentence, identify whether the different meanings arise from structural ambiguity, semantic ambiguity or pragmatic ambiguity? a. Time flies like an arrow b. He crushed the key to my heart	2	1	1	1
2.	Explain why CFG is used to represent natural language in parsing	2	2	1	1
3.	Given a trained classifier, how could you set up an experiment to evaluate its performance on some new text?	2	1	2	1
4.	Please comment about squared loss and log loss in logistic regression?	2	1	2	1
5.	Explain about Skip Gram and Common Bag Of Words (CBOW) models?	2	2	3	1
6.	How do you evaluate vector models?	2	1	3	1
7.	Why bias used in activation functions?	2	2	4	1
8.	What are limitations of LSTM and explain what is long term memory and short term memory w.r.t language models?	2	1	4	1
9.	Define the encoder-decoder networks components and its applications?	2	2	5	1
10.	What are the various metrics used to evaluate language models?	2	1	5	1
	Part-B (5 × 8 = 40 Marks)				
11. a)	Apply different smoothing techniques to handle data sparseness problem in n-gram model?	3	3	1	2
b)	Use Bi-Gram and Tri-gram model on following training corpus: 1. Thank you so much for your help. 2. I really appreciate your help. 3. Excuse me, do you know what time it is? 4. I'm really sorry for not inviting you. 5. I really like your watch. Which one is used to optimize any machine learning model for text classification? Justify your answer	5	3	1	2
12. a)	Discuss about logistic regression w.r.t. text classification.	4	2	2	1

Contd... 2

b)	Assume the following likelihoods for each word being part of a positive or negative movie review, and equal prior probabilities for each class. | | pos | neg | | I | 0.09 | 0.16 | | always | 0.07 | 0.06 | | like | 0.29 | 0.06 | | foreign | 0.04 | 0.15 | | movies | 0.08 | 0.11 | What class will Naive bayes assign to the sentence "I always like foreign films."?	4	3	2	2
13. a)	Demonstrate applications of TF-IDF and PPMI ?	3	2	3	1
b)	Apply Learning skip-gram embedding with suitable scenario.	5	3	3	2
14. a)	Explain the significance of different gates in LSTM. Also explain its limitations	4	2	4	1
b)	Describe how recurrent neural networks are modelled for sentiment classification?	4	3	4	1
15. a)	Discuss GRUs with neat diagram?	4	2	5	1
b)	Explain how Frame based dialogue system are used for machine translation?	4	2	5	1
16. a)	Discuss how Hidden Markov Model is used for Part-of-Speech Tagging?	5	2	1	1
b)	Analysis the optimization problem in Logistic Regression and how do you solve it using gradient decent method?	3	3	2	2
17.	Answer any ***two*** of the following:				
a)	Analysis the important of Word Embedding techniques in NLP with example?	4	3	3	1
b)	What X-OR problem how do we solve it?	4	1	4	1
c)	Apply the beam search algorithm in machine translation, discuss with diagram?	4	3	5	2

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%

*****