23AIPC401

MANGALORE INSTITUTE OF TECHNOLOGY & ENGINEERING

(A Unit of Rajalaxmi Education Trust®, Mangalore)
Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE, New Delhi
Accredited by NAAC with A+ Grade & ISO 9001:2015 Certified Institution

Model Question Paper

Seventh Semester BE Degree Examination

Natural Language Processing

Time: 3 Hours(180 Minutes) **Max. Marks: 100**

Note: 1. Answer any FIVE full questions, choosing ONE full question from each module.
2. M: Marks, L: RBT (Revised Bloom's Taxonomy) level, C: Course outcomes.

		Module -1	M	L	C
Q1	a.	Explain Natural Language Processing (NLP) and distinguish between the rationalist and empiricist approaches to NLP	06	L2	CO1
	b.	Describe the need of lexical, syntactic, and semantic levels in language processing using suitable examples.	06	L2	CO1
	c.	Explain the Byte-Pair Encoding (BPE) technique for subword tokenization with a suitable example.	08	L2	CO1
		OR			
Q2	a.	Differentiate between a word and a morpheme. Explain free and bound morphemes with suitable examples.	06	L2	CO1
	b.	Explain rule-based tokenization used in NLP with suitable examples.	06	L2	CO1
	c.	Explain the Minimum Edit Distance algorithm. Illustrate insertion, deletion and substitution operations using a suitable pair of words.	08	L2	CO1
		Module- 2			
Q3	a.	Given the following corpus: • `<s> I like NLP </s>` • `<s> I like AI </s>` determine the bigram counts and calculate P(AI|like) and P(NLP|like).	06	L3	CO2
	b.	A language model assigns probabilities 0.5, 0.25, 0.2 and 0.4 to four successive words in a sentence. Calculate the perplexity of the sentence using the perplexity formula and interpret the obtained value in terms of the model's predictive performance.	06	L3	CO2
	c.	A bigram language model contains the following probabilities: P(I|<s>)=0.6,P(like|I)=0.5,P(NLP|like)=0.4 Calculate the probability of generating the sequence: **<s> I like NLP** and explain the significance of the obtained probability.	08	L3	CO2
		OR			
Q4	a.	A corpus contains the following word counts given in table 1: **Table 1 : Example dataset 1** **Word** / **Count**: NLP 40; AI 30; ML 20; Data 10 Calculate the unigram probability of each word.	06	L3	CO2

	b.	A bigram language model has a vocabulary of 8 words. The bigram "data science" occurs 0 times, while the word "data" occurs 20 times. Using Laplace smoothing, calculate the probability P(science|data).	06	L3	CO2
	c.	A bigram "natural language" occurs 3 times, and the word "natural" occurs 30 times in a corpus. The vocabulary size is 20. **a.** State the general formula for Add-k smoothing. **b.** Calculate P(language|natural) using Add-1 smoothing. **c.** Calculate P(language|natural) using Add-0.5 smoothing. **d.** Compare the two probabilities and justify your answer on the effect of changing the value of k.	08	L3	CO2
Module - 3					
Q5	a.	Given the sentence "the student reads a book carefully" and a context window of two words around the target word "reads", apply the Skip-gram with Negative Sampling approach to identify the positive training pairs and generate negative training pairs when k=2. Illustrate how these training pairs are used to train the Skip-gram embeddings.	06	L3	CO3
	b.	Consider the following word-context count matrix given in table 2 : Table 2 : Example Dataset 2 Word | data | model | learning machine | 20 | 15 | 25 algorithm | 18 | 20 | 22 banana | 2 | 30 | 5 Calculate the cosine similarity between "machine" and "algorithm", and between "machine" and "banana". Determine which word is more semantically similar to "machine" based on the obtained cosine values.	06	L3	CO3
	c.	Consider the sentence "The movie was not excellent". Illustrate with a neat diagram how a feedforward neural-network classifier using hand-built features can represent the sentence and classify its sentiment.	08	L3	CO3
OR					
Q6	a.	Consider the following corpus: 1. machine learning improves prediction 2. machine learning improves accuracy 3. deep learning improves prediction 4. deep learning improves accuracy Using a context window of one word on either side, construct the word-context count vectors for "machine" and "deep" using the context vocabulary [machine, deep, learning, improves, prediction, accuracy] & also Illustrate how these vectors can be used to identify the contextual similarity between the two words.	06	L3	CO3
	b.	Given the sentence "lemon a tablespoon of apricot jam a pinch", apply the Skip-gram approach with a context window of two words around the target word "apricot" and illustrate the resulting target-context pairs.	06	L3	CO3
	c.	Given the context words "the", "student", and "reads", illustrate with a neat diagram how a feedforward neural language model converts the context words into embeddings and uses them through the hidden layer and softmax output layer to predict the next word.	08	L3	CO3
Module - 4					
Q7	a.	Explain with a suitable example the concept of a Large Language Model (LLM) and describe how it differs from a traditional statistical language model.	06	L2	CO4
	b.	Differentiate between encoder, decoder, and encoder-decoder architectures of language models with suitable examples.	06	L2	CO4
	c.	Explain with a suitable example how conditional text generation and prompting	08	L2	CO4

		are used to obtain a desired response from a Large Language Model.			
OR					
Q8	a.	Describe the three architectures of language models and explain the information flow in each architecture.	06	L2	CO4
	b.	Explain with suitable examples the difference between greedy generation and sampling in text generation.	06	L2	CO4
	c.	Describe the major stages involved in training and evaluating a Large Language Model.	08	L2	CO4
Module - 5					
Q9	a.	Explain the concept of self-attention in a Transformer with a suitable example.	06	L2	CO4
	b.	Describe the structure and major components of a Transformer block with a suitable diagram.	06	L2	CO4
	c.	Explain how token embeddings and positional embeddings are used to represent the input to a Transformer.	08	L2	CO4
OR					
Q10	a.	Explain the role of the Language Modeling Head in a Transformer-based language model.	06	L2	CO4
	b.	Describe the role of sampling in Transformer-based text generation.	06	L2	CO4
	c.	Explain how a Transformer can be interpreted using attention information to understand relationships between tokens.	08	L2	CO4

-*****-