CBCS SCHEME

USN | | | | | | | | | | |

BAI601

Sixth Semester B.E./B.Tech. Degree Examination, Dec.2025/Jan.2026
Natural Language Processing

Time: 3 hrs. Max. Marks: 100

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

		Module – 1	M	L	C
Q.1	a.	Illustrate with suitable examples the different levels of NLP.	8	L3	CO1
	b.	List and explain the challenges of NLP.	6	L3	CO1
	c.	Apply the bi-gram model for the test sentence and estimate the probability Training Sentence : — The quick brown for jumps over the lazy dog — A fast and agile for can legs very high — The dog watches the fox from a distance Test Sentence : — The quick and agile fox jumps over the dog.	6	L3	CO1
		OR			
Q.2	a.	Explain the role of transformationed rules in transformational grammar with the help of an example.	8	L2	CO1
	b.	Explain applications of NLP.	6	L2	CO1
	c.	List the problems associated with n-gram model. Explain how these problems are handled.	6	L2	CO1
		Module – 2			
Q.3	a.	Explain the working of two-step morphological parser. Write a simple Finite State Transducer (FST) for mapping English nouns.	8	L2	CO2
	b.	Illustrate parts of speech tagging and explain different categories.	6	L3	CO2
	c.	Explain Minimum Edit Distance. Compute the minimum edit distance between EXECUTION and INTENTION.	6	L2	CO2
		OR			
Q.4	a.	Design CYK Algorithm and tabulate the sequence of states created while parsing "A pilot likes flying planes". Consider the following simplified grammar in CNF : S → NP VP NN → pilot NP → DT NN NNS → planes NP → JJ NNS JJ → flying VP → VBG NNS DT → a VP → VBZ NP VBZ → likes VBG → flying	8	L4	CO2
	b.	Explain top-down parsing and bottom up parsing with examples.	8	L2	CO2
	c.	List out the disadvantages of Probabilistic Context Free Grammar (PCFG)	4	L2	CO2

		Module – 3			
Q.5	a.	Apply Naïve Bayes classifier and classify the test document "love gloomy fun" for the green training data. Use add-1 smoothing. D_1 : Love fun sunshine - Positive D_2 : Rain sad gloomy – Negative D_3 : Love rain fun – Positive D_4 : Gloomy sad sunshine – Negative	8	L4	CO3
	b.	What are the challenges in optimizing Naïve Bayes for sentiment analysis? How can they be addressed?	6	L1	CO3
	c.	Explain how Naïve Bayes can be adopted as a simple language model.	6	L2	CO3
		OR			
Q.6	a.	Explain Naïve Bayes assumption in text classification. Derive the Naïve Bayes formula with a simple example.	8	L2	CO3
	b.	Describe the step involved in training a Naïve Bayes classifier for text data.	6	L2	CO3
	c.	What are the limitations of using Naïve Bayes for text classification and sentiment analysis?	6	L2	CO3
		Module – 4			
Q.7	a.	Explain design features of information retrieval systems, with a neat diagram.	8	L2	CO4
	b.	Define term weighting, consider a document represented by 3 terms {tornado, swirl, wind} with the term frequency 4, 1 and 1 respectively. In a collection of 100 documents, 15 documents contains the term tornado, 20 contain swirl and 10 contain wind. Find the IDF and term weight of these terms.	6	L3	CO4
	c.	Explain the benefits of eliminatory stop words. Give examples where stop word elimination may be harmful.	6	L2	CO4
		OR			
Q.8	a.	List different Information Retrieval Models. Explain classical IR models.	10	L2	CO4
	b.	Explain WordNet and list the applications.	10	L2	CO4
		Module – 5			
Q.9	a.	Describe general architecture of a Encoder Decoder model used in Neural Machine Translation. Explain the roles and responsibilities of encoder and decoder components.	10	L2	CO5
	b.	Explain the concept of bias in Machine Translation Systems. What are the potential sources of bias in the training data and model architecture?	10	L2	CO5
		OR			
Q.10	a.	Explain the concept of language divergences in the context of Machine Translation. Provide at least three distinct types of linguistic divergences with illustrative examples.	10	L2	CO5
	b.	Explain the ethical issues associated with the development and deployment of machine Translation technologies. Provide at least two distinct ethical concerns.	10	L2	CO5

* * * * *

C.M.R. INSTITUTE OF TECHNOLOGY CMR BANGALORE
CMRIT LIBRARY BANGALORE - 560 037