

BE Sem 8, Choice based, 2019-20, C-8 scheme, May 2025  
Date :- 15/5/2025, Branch - E X T C

Time: 3 Hours

Marks:80

- Question Number 1 is Compulsory.
- Attempt ANY THREE out of the remaining five questions.
- Figures to the right indicate Maximum Marks.

- 1 Attempt ANY FOUR 20
- (a) List out five points that explain why natural language processing is hard.
- (b) With respect to Porter Stemmer, compute the number of vowels in the following words: {brachydactyly, employability, psychology, hieroglyphology, crystallography}
- (c) Point out, with justification, the referents of the reflexive pronouns in the following discourse:
  - I myself will answer this question, which is an impossible activity in itself.
  - Writing the report was itself a major task and the teacher handed it over to him.
- (d) Translate the following sentences into vectors using Count-Vectorization, by identifying unique words in the vocabulary, arranging them in alphabetical order, indexing them and finally counting the frequency of each word in the sentence and converting the sentence into vector.
  - It was the best of times.
  - It was the worst of times.
- (e) Describe open and closed class words with suitable examples.
- (f) Identify the referents and the referring expressions in the following two sentences:
  - Riya parked her Toyota in the garage. She had driven it around for hours.
  - Riya parked her Toyota in the garage. It is incredibly messy, with old bike and car parts lying around everywhere.
- 2(a) Explain morphological analysis with suitable examples 10
- 2(b) Consider the following corpus and list all possible bigrams. Compute the conditional probabilities for the word following "I". 10  
**NOTE:** Consider <s> and </s> as the start and end of sentence markers.  
 I like to sleep. I like to play. I like to eat. And I want to say that I like Mangoes. And I want Apples. But I got Peanuts.
- 3(a) Describe the process of parts of speech tagging using Hidden Markov Model. 10
- 3(b) Explain the measure of a word as described in Porter Stemmer. Use it to compute the measure of the following words: 10  
 {Supercalifragilisticexpialidocious, Pseudopseudohypoparathyroidism, Dermatoglyphics, Uncopyrightable}
- 4(a) Explain sense of a word. Illustrate word sense disambiguation task with suitable examples. 10
- 4(b) Explain term-frequency and inverse document frequency (TF-IDF) with an example 10
- 5(a) Explain co-reference resolution. Analyze the following sentences and point out, with justification, the referents of "she": 10  
 Mary had a little lamb. Mary always takes it to school. Angela saw Mary in the school garden playing with her lamb. She was happy.
- 5(b) Describe syntactic and semantic constraints on co-reference. 10
- 6(a) Describe in brief, the Question Answering System 10
- 6(a) Explain the steps involved in a Generic NLP system. 10

\*\*\*\*\*