

B.E. INFORMATION TECHNOLOGY FOURTH YEAR SECOND SEMESTER - 2024

NLP AND TEXT MINING(HONS.)

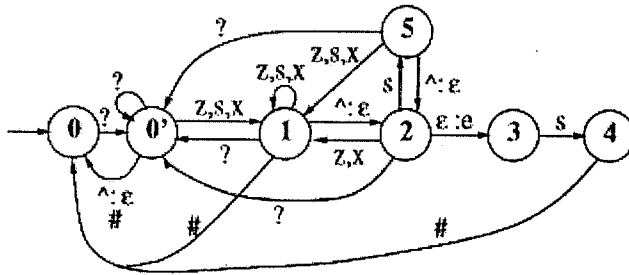
Time : 3hr

Full Marks : 100

Answer all Questions

CO1 (20)

- Describe the pipeline of NLP to be deployed in the Machine Learning model. (8)
- What are the steps of Morphological Parsing? Explain with an example. (6)
-



Apply this backwards to translate from surface to int. form. Starting from state 0, how many sequences of transitions are compatible with the input string 'asses'? explain. (6)

CO2 (20)

- Compute the edit distance between *different* and *difference*. explain and execute the algorithm. (3+7)
- Why do we need POS tags? What are the types of Lexical Ambiguity? (4+6)

CO3 (20)

- What are the importance of Quantification and Scoping in the context of NLP semantic representation. Explain Graph-Based Representation Frameworks with an example. (4+6)
- How First Order Predicate Logic is applied in NL semantics. (6)
- Translate the given statement into First-order logic- "None of my friends are perfect." (4)

[Turn over

CO4 (20)

- a. Discuss on the pipeline of question answering. (8)
- b. What do you mean by argumentation mining? Give an example. (5+2)
- c. What is the importance of tf-idf in text extraction summarization? (5)

CO5 (20)

- a. What do you mean by Precision, Recall, F1 Score, and Accuracy in the context of text mining. (2+2+2+2)
 - b. How can a model have a high accuracy but a low F1 score? Provide an example where this might happen? (6)
 - c. What do you mean by confusion matrix? Explain with an example. (6)
-