Reg. No. : | | | | | | | | | | | | |

Question Paper Code: 2437248

B.E. / B.Tech. DEGREE EXAMINATIONS, NOV/ DEC 2024
Seventh Semester
Artificial Intelligence and Data Science
U20AI703 – NATURAL LANGUAGE PROCESSING
(Regulation 2020)

Time: Three Hours Maximum: 100 Marks

Answer ALL questions

PART – A (10 x 2 = 20 Marks)

1. Explain the origins and challenges of Natural Language Processing (NLP).

2. Discuss the differences between grammar-based language models and statistical language models.

3. Describe the concept of unsmoothed N-grams and their limitations in NLP.

4. Explain the process of Part-of-Speech tagging and its importance in NLP.

5. What are Context-Free Grammars (CFG) and how are they used in syntactic analysis?

6. Discuss the role of treebanks in syntactic parsing.

7. Describe First-Order Logic and its application in semantic representation.

8. Explain the concept of Word Sense Disambiguation (WSD) and its methods.

9. What is discourse segmentation and why is it important for coherence in text?

10. Discuss the role of lexical resources like WordNet in Natural Language Processing.

PART – B (5 x 16 = 80 Marks)

11. (a) Explain in detail the different types of Language Models, comparing Grammar-based and Statistical Language Models. (16)

(OR)

(b) Discuss Finite-State Automata and their application in NLP, particularly in tokenization and English morphology. (16)

12. (a) Analyze the different smoothing techniques for N-grams and their impact on model performance. (16)

(OR)

(b) Compare rule-based, stochastic, and transformation-based approaches to Part-of-Speech tagging. (16)

13. (a) Describe the components of Context-Free Grammar (CFG) and explain how CFG can be applied to construct Treebanks for syntactic parsing. (16)

(OR)

(b) Discuss dynamic programming parsing techniques and their application in resolving ambiguity. (16)

14. (a) Developing a machine translation system for translating English to Spanish, define Word Sense Disambiguation (WSD) and explain how different methods, specifically Supervised Learning, Dictionary & Thesaurus-based approaches, and Bootstrapping—can be applied to improve the accuracy of translations. Provide examples of how each method might handle ambiguous words in sentences. (16)

(OR)

(b) Discuss thematic roles and selectional restrictions in semantic analysis. (16)

15. (a) Development of a chatbot designed to assist users with customer service inquiries, describe the contributions of key lexical resources such as WordNet, Penn Treebank, and Brill's Tagger. How can these resources enhance the chatbot's ability to understand user queries and generate appropriate responses? Provide specific examples of how each resource can be utilized in this context. (16)

(OR)

(b) Explain the processes involved in Anaphora Resolution, highlighting the Hobbs and Centering algorithms. (16)

----------xxxxx----------