UNIVERSITY OF PETROLEUM AND ENERGY STUDIES
Online End Semester Examination, Dec 2020

Course: Computational Linguistics & Natural Language Processing **Semester: V**
Program: B.Tech. CSE MAD **Time 03 hrs.**
Course Code: CSEG 3024P **Max. Marks: 100**

SECTION A
1. **Each Question will carry 5 Marks**
2. **Instruction: Complete the statement / Select the correct answer(s)**

S. No.	Question	CO
Q 1	Different phases of NLP are ------------- , ---------------, --------------, ---------, ----------------, ---------------, ----------------.	CO1
Q2	Which of the following is free morpheme a. books b. redo c. bounded d. drink	CO1
Q3	The differences in rule based and stochastic based POS tagging are --------------, ----------, -------------, --------------, ---------------, ---------------, --------------, -------------, ----------------, ------------------, --------------------.	CO2
Q4	Select the correct options a. Constituency parsing is based on grammar rules. b. Dependency parsing is based on word relationship. c. Cosine similarity is maximum for 90 degree angle between documents.	CO2
Q5	Select the correct statements a. Treebank is annotated corpus. b. Canned text is used in Surface realization c. POS tagging is used to analyze probability of sentences. d. Word sense disambiguation is based on semantic ambiguity.	CO2
Q6	Applications of Natural Language Processing are ----------------------- , ------------------------, --------------------------, -------------------- , ------------------------	CO1

SECTION B
1. **Each question will carry 10 marks**
2. **Instruction: Write short / brief notes**

Q 7	What are differences between archive and corpus? Write various standard corpus and steps to retrieve words and sentences (Use any NLP library such as NLTK) .	CO1

Q 8	Explain Boolean, Algebraic and Probabilistic Information retrieval models? Write applications of NLP information retrieval using real examples. OR Consider 3 documents are used as dataset and query is written in natural language. Implement TF/IDF and cosine similarity between query and documents (Use any NLP library such as NLTK).	**CO2**

Consider the following context-free grammar:

S -> NP VP

NP -> Det N

VP -> V

VP -> V NP

N -> dog

N -> cat

N -> mouse

Det -> the

V -> sees

V -> hates

V -> sneezes

Which of the following sentences are recognized by this grammar and why?

i) The dog sneezes the cat

ii) The mouse hates

iii) The cat the mouse hates

iv) The mouse hates the mouse

Q 9   v) The mouse sees the mouse

CO2

| Q 10 | Phases of Natural Language Processing with examples. | **CO1** |

Q 11	Explain document planning, micro planning and surface realization in detail. Use relevant examples for these phases of Natural Language Generation.	**CO1**

<table>
<tr><td colspan="3" align="center">Section C</td></tr>
</table>

1. **Each Question carries 20 Marks.**
2. **Instruction: Write long answer.**

Q12	What is Information answering? Write various datasets used for information answering. What are IR and knowledge based information answering techniques? Explain in detail with the use of examples. <center>OR</center> Explain use of POS tagging in NLP. What are different symbols used in tagging. Explain in detail using examples.	**CO2**