Final Assessment Test (FAT) - May 2024

Programme	**B.Tech.**	Semester	**WINTER SEMESTER 2023 - 24**
Course Title	**NATURAL LANGUAGE PROCESSING**	Course Code	**BCSE409L**
Faculty Name	**Prof. Radhika Selvamani**	Slot	**G1+TG1**
		Class Nbr	**CH2023240501677**
Time	**3 Hours**	Max. Marks	**100**

General Instructions:

- Write only Register Number in the Question Paper where space is provided (right-side at the top) & do not write any other details.

Answer all questions (10 X 10 Marks = 100 Marks)

01. Illustrate the different levels of Natural Language Processing techniques involved in designing an Automatic Grammar Checker with suitable examples for each level. **[10]**

02. Identify the affixes within the following words and determine whether each morpheme is free or bound. Additionally, specify whether they are inflectional or derivational, and provide justification. (5*2=10 Marks) **[10]**

a) Childishness
b) Employment
c) Running
d) Inexpensive
e) Uncomfortable

03. Your project team member developed an NLP application to match employment opportunities suited to job seekers based on their biodata. As a team member, you are asked to identify the suitable word embedding approach for the given application and illustrate it with proper justifications. **[10]**

04. Draw the phrase structured tree by assuming suitable production rules for the given sentence, "I walked towards the girl with the knife". Draw the dependency parse trees for the same sentence. Note that, If it is ambiguous then it will have more than one acceptable syntactic structure and you need to draw all possible forms for the same. **[10]**

05. Identify and discuss the lexeme-sense relations for the following pairs: (5*2=10) **[10]**

a) Chased – Pursued
b) Tool – Screwdriver
c) Sofa – Furniture
d) Tough – Delicate
e) Document – Introduction

06) Consider the probabilistic context-free grammar for the sentence "Visnu saw a beautiful bird in the garden". Using the given information, draw at least two parse trees and identify the probability score for the most probable tree. [10]

Rules	Probability	Rules	Probability
S→NP VP	1.0	NP→Det Noun	0.6
NP→NP PP	0.4	NP→Det ADJ Noun	0.5
VP→V NP	0.6	ADJ→beautiful	0.7
VP→V PP	0.5	V→likes	0.16
P →in	0.6	NP→Noun	0.6
PP→P NP	1.0	Det→ a | the	0.6|0.4
Verb→saw	0.4	Noun→Visnu | bird| garden	0.4|0.4|0.4

07) Learn a Bi-gram language model using the given corpus (including start and end symbol): [10]

<S>Brazilian Zika Brazilian</S>
<S>Brazilian Zika Virus</S>
<S>Virus Mild Brazilian Victim</S>
<S>Mild Symptoms Brazilian</S>

Perform the bigram probabilities and Laplace smoothing for the given sentence, "Brazilian Zika Virus Mild Symptoms".

08) Tag the sentence **"A pilot likes flying planes"** using Hidden Markov model with the given information: [10]

P(A/DT) =0.7	P(V/N)=0.8
P(Pilot/N)=0.8	P(N/V)=0.6
P(flying/ADJ) =0.9	P(ADJ/N)=0.2
P(flying/V)= 0.5	P(ADJ/V)=0.8
P(flying/N)=0.3	P (N/ADJ)=0.6
P(planes/N)=0.7	P(likes/V)=0.7
P(planes/V)=0.2	P(N/DT)=0.9
P(DT/N)=0.4	---

09) The digital books in the field of civil law are huge and challenging to read. You are informed to develop an app, capable of searching for answers to questions related to civil law in India. As an NLP expert, design a Question and Answering (QA) system for the given application by appropriate steps with suitable algorithms. [10]

10. Identify the thematic roles in the following sentences:(5*2=10) [10]

a) **We** persuaded **Mary** to apply for the job as a bus driver.
b) **Vinoth** has been earning **money** through his poetry.
c) Veena supplied **the reporters** with **information**.
d) The pirates sent **a message** to **the ship**.
e) **Nirmala** visited **California** before her marriage.