Name: **Enrolment No:**	

UPES

End Semester Examination, May 2023

Course: Natural Language Processing **Semester: II**

Program: M.Tech (CSE) **Time : 03 hrs.**

Course Code: CSAI 7006P **Max. Marks: 100**

Instructions: All questions are compulsory. Question no. 9 of Section B and Question no. 10 of Section C have internal choice.

SECTION A
(5Qx4M=20Marks)

S. No.		Marks	CO
Q 1	"I went to bank to withdraw money". Which type of ambiguity exists in this sentence?	**4**	**CO1**
Q 2	"Sentiment Analysis is application of NLP. Machine translation is also an application of NLP". Perform NLG aggregation on these sentences.	**4**	**CO2**
Q 3	How phrases are formed from source text in machine translation?	**4**	**CO3**
Q 4	Write formula of Cosine Similarity and explain terms used in formula.	**4**	**CO4**
Q 5	In top-down parser, when sentence is marked as correct?	**4**	**CO1**

SECTION B
(4Qx10M= 40 Marks)

Q 6	Describe Wordnet corpus and applications of this corpus in finding synsets of words. How similarity between words is calculated using Wordnet? Write code to find similarity.	**10**	**CO1**
Q 7	How text, sentence, document, and micro planning are performed in NLG?	**10**	**CO2**
Q 8	Various possibilities exist to convert source text to target text in machine translation. How probability-based models are used to find the most correct target text?	**10**	**CO3**
Q 9	Cosine similarity is used in Vector Space model. Draw diagram to demonstrate the use of cosine angle to find similarity between documents. Write formula with detailed explanation of terms. or Explain the use of indexing in search engine. How forward and inverted index are created?	**10**	**CO4**

SECTION-C
(2Qx20M=40 Marks)

<table>
<tr>
<td>Q 10</td>
<td>D1= “Cricket match between India and Australia was very close match”
D2= “NIFTY 50 is stock index”
D3= “In India, several reputed Universities are opened now”
Q= “What is the name of stock index in India”
Find the similarity between query and document corpus using TF/IDF and rank the similarity in descending order.
or

Consider the following productions:
NP ⟶ NP PP
NP ⟶ Det N
NP ⟶ Delhi
N ⟶ flight
VP ⟶ VP PP
VP ⟶ Verb NP
Verb ⟶ Book
PP ⟶ Prep NP
Prep ⟶ through
Det ⟶ the
Where;
NP – noun phrase
VP –verb phrase
PP -preposition phrase.
a) Use the CYK parsing algorithm to find if the sentence "Book the flight through Delhi" belongs to the above grammar.
Explain the CYK algorithm.</td>
<td>20</td>
<td>CO4</td>
</tr>
<tr>
<td>Q 11</td>
<td>How probabilistic based grammar is used for resolving ambiguity?

If P(Art|Φ)=0.7, P(N|Art)=1, P(V|N)=0.4, P(N|V)=0.3, P(N|Φ)=0.30, 29P(Art|V)=0.7, P(V|Φ)=0.01, P(P|V)=0.1, P(V|V)=0.1, P(Art|P)=0.2

and occurrences of terms are as follows.

<table>
<tr><td></td><td>N</td><td>V</td><td>ART</td><td>P</td><td>Total</td></tr>
<tr><td>Flies</td><td>37</td><td>30</td><td>0</td><td>0</td><td>67</td></tr>
<tr><td>like</td><td>5</td><td>25</td><td>0</td><td>12</td><td>42</td></tr>
<tr><td>A</td><td>2</td><td>0</td><td>30</td><td>0</td><td>30</td></tr>
<tr><td>flower</td><td>50</td><td>10</td><td>0</td><td>0</td><td>60</td></tr>
<tr><td>Total</td><td>94</td><td>65</td><td>30</td><td>12</td><td>199</td></tr>
</table>

How ambiguity in sentence “Flies like a flower” can be resolved using Probabilistic grammar.</td>
<td>20</td>
<td>CO1</td>
</tr>
</table>