Total No. of Questions: 6 *Total No. of Printed Pages:3*

Enrollment No.....

Faculty of Engineering

End Sem Examination May-2023

CS3EA06 Natural Language Processing

Programme: B.Tech. Branch/Specialisation: CSE All

Duration: 3 Hrs. **Maximum Marks: 60**

Note: All questions are compulsory. Internal choices, if any, are indicated. Answers of Q.1 (MCQs) should be written in full instead of only a, b, c or d. Assume suitable data if necessary. Notations and symbols have their usual meaning.

Q.1 i. Google Translate is one of the _____ applications.? **1**

(a) Machine translation (b) Information Retrieval

(c) Information Extraction (d) Summarization

ii. What are the input and output of an NLP system? **1**

(a) Speech and noise (b) Speech and Written Text

(c) Noise and Written Text (d) Noise and value

iii. Singular: knife, Plural: Knives Which rules are applied? **1**

(a) Morphological Rule (b) Orthographic rule

(c) Mechanical Rule (d) Dynamic Rule

iv. Which step is the process of breaking down documents into smaller units of analysis? **1**

(a) Ngrams (b) Stopwords (c) Corpus (d) Tokenization

v. What kind of signal is used in speech recognition? **1**

(a) Electromagnetic signal (b) Electric signal

(c) Acoustic signal (d) Radar

vi. What is viewed as a problem of probabilistic inference? **1**

(a) Speech recognition (b) Speaking

(c) Hearing (d) Utterance

vii. Which model gives the probability of each word following each other word? **1**

(a) Bigram model (b) Diagram model

(c) Gram model (d) Speech model

P.T.O.

viii. N-grams are defined as the combination of N keywords together. How many bi-grams can be generated from the given sentence: Gandhiji is the father of our nation- **1**

(a) 7 (b) 6 (c) 8 (d) 9

ix. Which is the main Python package we use for NLP? **1**

(a) NLTK (b) NLP-LIB (c) Scikit-Learn (d) PyNLP

x. Which of the following is the major problem in Machine Translation? **1**

(a) Referential Ambiguity
(b) Stop word
(c) Emoticons
(d) Proper Noun

Q.2 i. Design a finite automata for given regular expression - aa*b. **2**

ii. Why regular expression is important for text processing? Justify with an example. **3**

iii. Explain the difference between formal language and natural language with appropriate example. **5**

OR iv. What is natural language processing? Explain two major approaches to natural language processing. **5**

Q.3 i. Explain Part of speech tagging (POS). **2**

ii. Design a finite state transducer with E-insertion orthographic rule that parses from surface level "foxes" to lexical level "fox+N+PL" using FST. **8**

OR iii. What is morphological parsing? Explain the two steps of morphological parser. **8**

Q.4 i. Explain spelling correction algorithm. **3**

ii. What do you mean by Computational Phonology? How speech and phonetics are interrelated with an example? **7**

OR iii. Analyze the naive Bayes classifier approach to Word Sense Disambiguation in NLP. **7**

Q.5 i. Explain any two N-gram models. **4**

ii. Write a short note on the importance of Smoothing and Perplexity. **6**

OR iii. Explain statistical parsing and probabilistic parsing with an example. **6**

Q.6 Attempt any two:

i. Describe transfer model of Machine Translation. List out its three phases. **5**

ii. Explain the term "Ambiguity", also explain different level of ambiguity occurs in natural language processing. **5**

iii. What are the methods used for spelling correction in natural language processing? **5**

******

Marking Scheme

CS3EA06[T]-Natural Language Processing

Q.1	i)	a). Machine translation		**1**
	ii)	b). Speech and Written Text		**1**
	iii)	b).Orthographic rule		**1**
	iv)	d).Tokenization		**1**
	v)	a). Electromagnetic signal		**1**
	vi)	a). Speech recognition		**1**
	vii)	a). Bigram model		**1**
	viii)	b). 6		**1**
	ix)	a).NLTK		**1**
	x)	a).Referential Ambiguity		**1**
Q.2	i.	Design a finite automata for given regular expression - aa*b 2 marks		**2**
	ii.	regular expression is important for text processing example	2 marks 1 mark	**3**
	iii.	difference between formal language and natural language for each difference (1*5)	1 mark	**5**
OR	iv.	definition two major approaches to NLP	2 marks 3 marks	**5**
Q.3	i.	definition	2 marks	**2**
	ii.	Design a finite state transducer with E-insertion orthographic 8 marks		**8**
OR	iii.	What is morphological parsing Explain the two steps of morphological parser	4 marks 4 marks	**8**
Q.4	i.	Explain spelling correction algorithm	3 marks	**3**
	ii.	Computational Phonology,and phonetics example	4 marks 3 marks	**7**

OR	iii.	Analyze the naive Bayes classifier approach to Word Sense Disambiguation in NLP	7 marks	**7**
Q.5	i.	two N-gram models	2 marks for each(2*2)	**4**
	ii.	Smoothing Perplexity	3 marks 3 marks	**6**
OR	iii.	statistical parsing and probabilistic parsing example	4 marks 2 marks	**6**
Q.6				
	i.	transfer model of Machine Translation three phases (1*3)	2 marks 3 marks	**5**
	ii.	definition of Ambiguity different level of ambiguity occurs in NLP	2 marks 3 marks	**5**
	iii.	methods used for spelling correction in NLP	5 marks	**5**

******