



Library
Brainware University
33, Ramkrishnapur Road, Barasat
Kolkata, West Bengal-700125

BRAINWARE UNIVERSITY

Term End Examination 2024-2025

Programme – MCA-2022/MCA-2023

Course Name – Natural Language Processing

Course Code - MCA401A

(Semester IV)

Full Marks : 60

Time : 2:30 Hours

[The figure in the margin indicates full marks. Candidates are required to give their answers in their own words as far as practicable.]

Group-A

(Multiple Choice Type Question)

1 x 15=15

1. Choose the correct alternative from the following :

- (i) Distinguish the field of Natural Language Processing (NLP).
 - a) Computer Science
 - b) Artificial Intelligence
 - c) Linguistics
 - d) All of the mentioned
- (ii) Choose the main challenge of NLP.
 - a) Handling Ambiguity of Sentences
 - b) Handling Tokenization
 - c) Handling POS-Tagging
 - d) All of the mentioned
- (iii) Distinguish from the following areas where NLP can be useful.
 - a) Automatic Text Summarization
 - b) Automatic Question-Answering Systems
 - c) Information Retrieval
 - d) All of the mentioned
- (iv) Distinguish the best definition for Coreference Resolution.
 - a) Anaphora Resolution
 - b) Given a sentence or larger chunk of text, determine which words ("mentions") refer to the same objects ("entities")
 - c) All of the mentioned
 - d) None of the mentioned
- (v) What is Machine Translation?
 - a) Converts one human language to another
 - b) Converts human language to machine language
 - c) Converts any human language to English
 - d) Converts Machine language to human language
- (vi) What is Morphological Segmentation?
 - a) Does Discourse Analysis
 - b) Separate words into individual morphemes and identify the class of the morphemes
 - c) Is an extension of propositional logic
 - d) None of the mentioned
- (vii) Distinguish the best match. _____ is one of the goals of sentiment analysis is to identify sentiment among several posts or even in the same post where emotion is not always explicitly expressed.

- a) Automatic Summarization
c) Text Classification
- b) Sentiment Analysis
d) All of the mentioned
- (viii) "I am tired." Contains which type of ambiguity?
a) Semantic ambiguity
c) Lexical ambiguity
- b) Syntax Level ambiguity
d) None of the mentioned
- (ix) The study of the construction of words from primitive meaningful units, is labelled as _____.
a) Phonology
c) Morpheme
- b) Phonology
d) Morphology
- (x) Determine the meaning of Morphology.
a) The study of word format
c) The study of syntax of sentence
- b) The study of sentence format
d) The study of semantics of sentence.
- (xi) Predict the derivational prefixes do not change the category of the word to which they are attached.
a) Re- & -Un
c) -ize
- b) -er
d) -ing
- (xii) Select the Python library used to implement natural language processing.
a) NLTK
c) Matplotlib
- b) Scrapy
d) Pydot
- (xiii) Choose the correct option for the statement. Porter Stemmer algorithm is used for _____.
a) Lemmatization
c) Stemming
- b) Syntax Analysis
d) Part of speech tagging
- (xiv) Predict the example of free morphemes.
a) Dog
c) Re-(Reschedule)
- b) Un-(unhappy)
d) -y(smiley)
- (xv) Express the difference between Finite State Automata (FSA) and Finite State Transducers (FST).
a) FSA contains single tape and FST also contain single tape.
c) FSA contains single input tape and FST contain input. output pair tapes
- b) FSA contains single input tape and FST contain single output tape.
d) Both FSA and FST contain output tapes only

Group-B

(Short Answer Type Questions)

3 x 5=15

2. Describe the properties of finite automata. (3)
3. Describe the properties of NFA. (3)
4. Explain statistical machine translation. (3)
5. Reframe the steps of Text Normalization. (3)
6. Express the process of Word2Vec. (3)

OR

- Integrate TF and IDF in word analyzing. (3)

Group-C

(Long Answer Type Questions)

5 x 6=30

7. Construct a document vector table for the given corpus: Document 1: We are going to Mumbai Document 2: Mumbai is a famous place. Document 3: We are going to a famous place. Document 4: I am famous in Mumbai. (5)
8. Explain the steps of text Normalization. (5)

9. Calculate TFIDF for the given corpus and mention the word(s) having the highest value. (5)
Document 1: We are going to Mumbai Document 2: Mumbai is a famous place. Document 3: We are going to a famous place. Document 4: I am famous in Mumbai.
10. Assess the Turing Test in NLP. (5)
11. Justify Semantic Analysis with a proper example. (5)
12. Discriminate the steps involved in preprocessing data for NLP. (5)

OR

Summarize Bag of Words (BOW). (5)

Library
Brainware University
398, Ramkrishnapur Road, Barasat
Kolkata, West Bengal-700125