



LOYOLA COLLEGE (AUTONOMOUS), CHENNAI – 600 034

M.Sc. DEGREE EXAMINATION – DATA SCIENCE

THIRD SEMESTER – NOVEMBER 2022

PDS 3507 – NATURAL LANGUAGE PROCESSING

Date: 28-11-2022

Dept. No.

Max. : 100 Marks

Time: 09:00 AM - 12:00 NOON

PART – A

Q. No **Answer ALL questions** **(10 x 2 = 20 Marks)**

- 1 What is language modeling?
- 2 What is text classification?
- 3 What do you mean by word sense disambiguation?
- 4 Define a Named Entity.
- 5 What is POS tagging?
- 6 Define a Hidden Markov Model.
- 7 What is parsing?
- 8 Define unification of two features.
- 9 Differentiate between word sense and lemma.
- 10 What is a discourse?

PART – B

Answer ALL questions **(5 x 8 = 40 Marks)**

- 11 (a) Briefly explain any two core applications of NLP.
or
(b) Explain the origin and challenges of NLP.
- 12 (a) Briefly Explain: (i) Stop Word Removal (ii) Lemmatization
or
(b) Briefly Explain: (i) Phonemes (ii) Morphemes
- 13 (a) Explain N-Grams and their uses in language processing.
or
(b) Explain Weighted Finite State Automaton and Markov Chain.
- 14 (a) Explain Syntactic Parsing and Parse Trees.
or
(b) Explain Context-Free Grammar with an example.
- 15 (a) Explain Lexical Semantics with an example.
or
(b) Explain Dialog Conversational Agents.

PART – C

Answer any TWO questions

(2 x 20 = 40 Marks)

- 16 (a) Explain how Machine Learning algorithms are used in NLP.
(b) Explain TF-IDF Word Embedding method with an example.
- 17 (a) Explain Word Frequency calculation and Sentence Score calculation.
(b) Explain Rule-based POS tagging algorithm.
- 18 (a) Explain Bottom-up parsing with an example.
(b) Explain the features of WordNet with an example.

\$\$\$\$\$\$