PRESIDENCY UNIVERSITY
BENGALURU

Department of Research & Development

Mid - Term Examinations - AUGUST 2024

Odd Semester: Ph.D. Course Work   **Date**: 13-08-2024

Course Code: CSE900   **Time**: 09.30am to 11.00am

Course Name: Natural Language Processing   **Max Marks**: 50

Department: Computer Science Engineering   **Weightage**: 25%

Instructions:
(i) Read the all questions carefully and answer accordingly.
(ii) Do not write any matter on the question paper other than roll number.

Part A

Answer **ALL** the Questions. **Each** question carries **4** marks.   (4Q x 5M = 20M)

1. Define Lemmatization, Tokenization, sentence segmentation.
2. Give five examples of typical regular expression patterns.
3. Describe Word Embeddings with the help of an example.
4. Explain the need for classifiers in natural language processing. List a few typical classifiers used in NLP.

Part B

Answer **ALL** Questions. **Each** question carries **15** marks.   (2Q x 15=30M)

5. Demonstrate the working of minimum edit distance algorithm.
6. Illustrate the key concepts of n-grams and how they are used in NLP.