

Time: 3 Hours

Max. Marks: 80

- **Question No.1 is compulsory.**
- Solve **ANY THREE** questions from the **remaining** five questions.
- Assume suitable data wherever required, but justify the same.

1. Solve **ANY FOUR** questions from the following: (20)
 - a. List out five applications in Natural Language Processing.
 - b. With respect to Porter Stemmer, evaluate the measure of the following words: { queue, cryogenic, onomatopoeia, neuropsychoticdisorder, larynx }
 - c. Compute the cosine distance between the two given documents:
D₁ = For want of a nail, the shoe was lost.
D₂ = For want of a shoe, the horse was lost.
 - d. Compute the conditional probability P(okay | okay) from the following input data:

And it's okay okay to feel this way. Feel like you're lost. Dreaming and hoping. And it's okay okay to feel this way. But just know that it's all gonna be okay. And it's okay okay to feel this way. Feel like you're lost. Dreaming and hoping. And it's okay okay to feel this way. But just know that it's all gonna be okay.

- e. Identify derivational morphemes from the given data:

With every day passing. She felt distant. Don't know if it was her anxiety, her thoughts, insecurities, these feelings suppressed her

2. (a) Describe Derivational Morphology (10)
(b) Describe Finite State Transducer as a tool in morphological analysis. (10)
3. (a) Illustrate the process of stemming and lemmatization with suitable examples. (10)
(b) Describe the process of Parts of Speech Tagging using Hidden Markov Model. (10)
4. (a) Explain the following terms with suitable examples: (10)
Hyponymy, Hypernym, Meronym, Holonym, Zeugma
(b) Illustrate with suitable examples the following types of referring expressions: (10)
Indefinite noun phrases, Definite noun phrases, Pronouns, Demonstratives, One-Anaphora
5. (a) List and explain syntactic and semantic constraints on co-reference (10)
(b) Explain cosine distance between documents with suitable examples. (10)
6. (a) Explain Question Answering System in detail. (10)
(b) Write short notes on ANY TWO of the following: (10)
 - (i) Text Summarization
 - (ii) Named Entity Recognition
 - (iii) Machine Translation
