

GUJARAT TECHNOLOGICAL UNIVERSITY
BACHELOR OF ENGINEERING – SEMESTER 7 – EXAMINATION – SUMMER 2026

Subject Code: 3174602

Date: 03/06/2026

Subject Name: NATURAL LANGUAGE PROCESSING

Time: 02:30 PM TO 05:00 PM

Total Marks: 70

Instructions:

- 1. Attempt all questions.**
- 2. Make suitable assumptions wherever necessary.**
- 3. Figures to the right indicate full marks.**
- 4. Simple and non-programmable scientific calculators are allowed.**

- Q.1**
- (a) Mention any four popular NLP libraries and their uses. **03**
 - (b) Describe the phases of an NLP pipeline in brief. **04**
 - (c) Discuss why NLP is considered a difficult task. Explain Zipf's Law and Heaps Law. **07**
- Q.2**
- (a) What are stop words? Explain the process of stop word removal with example. **03**
 - (b) Differentiate between stemming and lemmatization. **04**
 - (c) Describe the Text Preprocessing steps: Stop word removal, Punctuation removal, and Lowercase conversion. **07**
- OR**
- Q.2**
- (c) Explain the need for smoothing and describe advanced techniques like Laplace, Good-Turing, and Kneser-Ney smoothing. **07**
- Q.3**
- (a) Differentiate between POS tagging and NER. **03**
 - (b) Define Named Entity Recognition (NER). Explain its working and applications. **04**
 - (c) Explain Unigram, Bigram, and Trigram models with suitable examples. **07**
- OR**
- Q.3**
- (a) Define Part-of-Speech (POS) Tagging. Explain its significance. **03**
 - (b) Explain the concept of Morphology in NLP with examples. **04**
 - (c) Explain Natural Language Generation (NLG). Describe its architecture and challenges in generating human-like text. **07**
- Q.4**
- (a) Define Word Sense Disambiguation (WSD). Why is it important? **03**
 - (b) Differentiate between Extractive and Abstractive Text Summarization. **04**
 - (c) Explain the challenges in Multilingual Question Answering systems. **07**
- OR**
- Q.4**
- (a) Explain the working of Sentiment Mining with an example. **03**
 - (b) Differentiate between CBOW and Skip-gram. **04**
 - (c) Explain the Continuous Bag-of-Words (CBOW) architecture. **07**
- Q.5**
- (a) Explain the working of Direct Machine Translation with an example. **03**
 - (b) Explain how the Expectation-Maximization (EM) algorithm is used in SMT parameter learning. **04**

(c) Describe the Encoder-Decoder architecture used in Neural Machine Translation. **07**

OR

Q.5 (a) Explain the role of Dirichlet distribution in LDA. **03**

(b) Explain Gibbs Sampling and its role in LDA. **04**

(c) Explain Rule-Based Machine Translation (RBMT) in detail with advantages and limitations. **07**
